This is to certify that the

dissertation entitled

Balancing Lifting Values to Improve Numerical Stability of

Polyhedral Homotopy Continuation Methods

presented by

Mengnien Wu

has been accepted towards fulfillment
of the requirements for

Ph.D degree in Mathematics

Major professor

Date 9/27/00

PLACE IN RETURN BOX to remove this checkout from your record.
TO AVOID FINES return on or before date due.
MAY BE RECALLED with earlier due date if requested.

DATE DUE	DATE DUE	DATE DUE
DEC 2 3 2005		

6/01 c:/CIRC/DateDue.p65-p.15

Balancing Lifting Values to Improve Numerical Stability of Polyhedral Homotopy Continuation Methods

By

Mengnien Wu

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Mathematics

2000

ABSTRACT

Balancing Lifting Values to Improve Numerical Stability of Polyhedral Homotopy Continuation Methods

By

Mengnien Wu

Polyhedral homotopy continuation methods exploit the sparsity of polynomial systems so that the number of solution curves to reach all isolated solutions is optimal for generic systems. The numerical stability of tracing solution curves of polyhedral homotopies is mainly determined by the height of the powers of the continuation parameter. To reduce this height we propose a procedure that operates as an intermediate stage between the mixed-volume computation and the tracing of solution curves. This procedure computes new lifting values of the support of a polynomial system. These values preserve the structure of the mixed-cell configuration obtained from the mixed-volume computation and produce better-balanced powers of the continuation parameter in the polyhedral homotopies.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to Professor Tien-Yien Li, my dissertation advisor, for his constant encouragement and kind guidance. His knowledge, insights and enthusiasm were invaluable.

I would also like to thank Professor Chichia Chiu, Professor Ronald Fintushel, Professor Michael Frazier and Professor William Sledd, for their time and advices.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Introduction

During the last two decades, homotopy continuation methods have proven to be reliable and efficient to compute numerical approximations to all isolated zeros of polynomial systems. To exploit sparsity of polynomial systems, polyhedral homotopy continuation methods emerged in [HuSt95, VVC]. For generic polynomial systems, the number of solution curves to be traced is optimal. We refer to [Li97] for a survey on recent advances in homotopy continuation methods for polynomial systems.

Polyhedral homotopies are nonlinear in the continuation parameter. Powers of the continuation parameter too close to zero can be scaled away from zero by a suitable scalar multiplication. After scaling, if very high powers exist in a polyhedral homotopy, small step sizes must be taken in order to successfully trace a solution curve of the polyhedral homotopy. Although the end of the solution curve can be reached as long as the required step size is not smaller than the available machine precision, the efficiency of the curve-tracing is greatly reduced. A more serious problem occurs when the continuation parameter, starting from 0, is not yet close enough to 1, some terms of the polyhedral homotopy with high powers of the continuation parameter have values smaller than the machine precision and some solution curves may come close to "valleys" where the values of the homotopy are numerically zero, but no solution curves exist inside the "valleys". This situation can easily cause the curve-tracings to be trapped in these "valleys" with no chance of reaching the ends of solution curves unless the curves are retraced with smaller step sizes.

Two known geometric approaches to control the numerical stability of polyhedral homotopy continuation methods are recursive liftings (as in Bernshtein's algorithm [Bern, VVC]) and dynamic liftings [VGC]. However, because of using multiple liftings or flattenings, these approaches both require more expensive construction of subdivisions and create more homotopy curves which need to be traced than a random floating-point lifting.

To minimize the height of the powers of the continuation parameter in polyhedral homotopies, we search in the cone of all lifting vectors that induce the same mixed-cell configuration to obtain better-balanced powers of the continuation parameter of polyhedral homotopies. This idea can also be found in the proof of Proposition 1.11 in [St96].

This dissertation is organized in the following manner: In Chapter 1, we start with the polyhedral homotopy method and its numerical stability; in Chapter 2, the idea originated from the construction of the polyhedral homotopy to balance lifting values is given, followed by the setup of the linear programming model for balancing lifting values as well as the reduction of the size of this LP model; and finally, some numerical experiments are presented in Chapter 3.

CHAPTER 1

Polyhedral Homotopy

To solve a polynomial system $P(\mathbf{x}) = (p_1(\mathbf{x}), \cdots, p_n(\mathbf{x}))$ in $\mathbb{C}^n$ by the homotopy continuation method, the homotopy $H(\mathbf{x}, t) : \mathbb{C}^n \times [0,1] \to \mathbb{C}^n$ needs to posess the following properties:

- Property 1 (*Triviality*). The solutions of $H(\mathbf{x}, 0) = \mathbf{0}$ are known.
- Property 2 (*Smoothness*). The solution set of $H(\mathbf{x}, t) = \mathbf{0}$ for $0 \leq t \leq 1$ consists of a finite number of smooth paths, each parametrized by $t \in [0, 1)$.
- Property 3 (*Accessibility*). Every isolated solution of $H(\mathbf{x}, 1) = P(\mathbf{x}) = \mathbf{0}$ can be reached by some path originating at $t = 0$, i.e. this path starts at a solution of $H(\mathbf{x}, 0) = \mathbf{0}$.

When $H(\mathbf{x}, t) = \mathbf{0}$ defines a homotopy that satisfies above properties, the number of isolated zeros of $H(\mathbf{x}, 0)$ must be no fewer than the number of isolated zeros of $P(\mathbf{x})$. Unfortunately, the former is in general much greater than the later, resulting in a considerable waste of computational effort in following extraneous paths in practice.

The Bernshtein theory on root count of polynomial systems is essential for our attempt to reduce the number of homotopy curves need to be traced when the homotopy continuation method is employed to find all isolated zeros of polynomial systems.

In the first section of this chapter, the Bernshtein theory on root count in $(\mathbb{C}^*)^n$, where $\mathbb{C}^* = \mathbb{C}\setminus\{0\}$, as well as its extension to root count in $\mathbb{C}^n$ are presented. In the second section, the polyhedral homotopy, based on the Bershtein theory, for finding all isolated zeros of a polynomial system is introduced. In the last section, we will discuss the method to solve a binomial system, so that initial solutions of a polyhedral homotopy can be identified.

1.1 Bernshtein Theory

Let the given polynomial system be $P(\mathbf{x}) = (p_1(\mathbf{x}), \cdots, p_n(\mathbf{x})) \in \mathbb{C}[\mathbf{x}]$, where $\mathbf{x} = (x_1, \cdots, x_n)$. With $\mathbf{x}^{\mathbf{q}} = x_1^{q_1} \cdots x_n^{q_n}$ where $\mathbf{q} = (q_1, \cdots, q_n)$, write

$$\begin{aligned} p_1(\mathbf{x}) &= \sum_{\mathbf{q}\in S_1} c_{1,\mathbf{q}}\mathbf{x}^{\mathbf{q}}, \\ &\vdots \\ p_n(\mathbf{x}) &= \sum_{\mathbf{q}\in S_n} c_{n,\mathbf{q}}\mathbf{x}^{\mathbf{q}}, \end{aligned} \tag{1.1}$$

where $S_1, \cdots, S_n$ are fixed subsets of $\mathbb{N}^n$ with cardinals $k_j = \#S_j$, and $c_{j,\mathbf{q}} \in \mathbb{C}^*$ for $\mathbf{q} \in S_j$, $j = 1, \cdots, n$. We call S_j the *support* of $p_j(\mathbf{x})$, denoted by $\mathrm{spt}(p_j)$, its convex hull $Q_j = \mathrm{conv}(S_j)$ in $\mathbb{R}^n$ the *Newton polytope* of p_j, and $S = (S_1, \cdots, S_n)$ the *support* of $P(\mathbf{x})$, denoted by $\mathrm{spt}(P)$.

We now embed the system (1.1) in the system $P(\mathbf{c}, \mathbf{x}) = (p_1(\mathbf{c}, \mathbf{x}), \cdots, p_n(\mathbf{c}, \mathbf{x}))$, where

$$\begin{aligned} p_1(\mathbf{c}, \mathbf{x}) &= \sum_{\mathbf{q}\in S_1} c_{1,\mathbf{q}}\mathbf{x}^{\mathbf{q}}, \\ &\vdots \\ p_n(\mathbf{c}, \mathbf{x}) &= \sum_{\mathbf{q}\in S_n} c_{n,\mathbf{q}}\mathbf{x}^{\mathbf{q}}, \end{aligned} \tag{1.2}$$

and the coefficients $c_{j,\mathbf{q}}$ with $\mathbf{q} \in S_j$, for $j = 1, \cdots, n$ in the system are taken to be a set of $M := k_1 + \cdots + k_n$ variables. Namely, the system $P(\mathbf{x})$ in (1.1) is considered as a system in (1.2) corresponding to a set of specified values of coefficients $\mathbf{c} = (c_{j,\mathbf{q}})$ or $P(\mathbf{x}) = P(\mathbf{c}, \mathbf{x})$.

We shall refer to the total number of isolated zeros, counting multiplicities, of a polynomial system as the *root count* of the system.

Lemma 1 [Hu96] *For polynomial systems $P(\mathbf{c}, \mathbf{x})$ in (1.2), there exists a polynomial system $G(\mathbf{d}) = (g_1(\mathbf{d}), \cdots, g_n(\mathbf{d}))$ in the variables $\mathbf{d} = (d_{j,\mathbf{q}})$ for $\mathbf{q} \in S_j$ and $j = 1, \cdots, n$ such that for those coefficients $\mathbf{c} = (c_{j,\mathbf{q}})$ for which $G(\mathbf{d})\Big|_{\mathbf{d} = \mathbf{c}} \neq \mathbf{0}$, the root count in $(\mathbb{C}^*)^n$ of the corresponding polynomial systems in (1.2) is a fixed number. And the root count in $(\mathbb{C}^*)^n$ of any other polynomial systems in (1.2) is bounded above by this number.*

Remark 1 Since the zeros of the polynomial system $G(\mathbf{d})$ in the above lemma form an algebraic set with dimension smaller than M, its complement is open and dense with full measure in $\mathbb{C}^M$. Therefore, with probability one, $G(\mathbf{d}) \neq \mathbf{0}$ for randomly chosen coefficients $\mathbf{d} = (d_{j,\mathbf{q}}) \in \mathbb{C}^M$. Hence, polynomial systems $P(\mathbf{c}, \mathbf{x})$ in (1.2) with $G(\mathbf{c}) \neq \mathbf{0}$ are said to be **in general position**.

Theorem 1 (([Bern], Theorem A)) *The root count in $(\mathbb{C}^*)^n$ of a polynomial system $P(\mathbf{x}) = (p_1(\mathbf{x}), \ldots, p_n(\mathbf{x}))$ in general position equals to the* **mixed volume** *of its support.*

The terminology in this theorem needs explanation. For non-negative variable $\lambda_1, \cdots, \lambda_n$ and the Newton polytopes Q_j of p_j, for $j = 1, \cdots, n$, let $\lambda_1 Q_1 + \cdots + \lambda_n Q_n$ denote the *Minkowski sum* of $\lambda_1 Q_1, \cdots, \lambda_n Q_n$, that is,

$$\lambda_1 Q_1 + \cdots + \lambda_n Q_n = \left\{ \lambda_1 r_1 + \cdots + \lambda_n r_n \,\middle|\, r_j \in Q_j,\, j = 1, \cdots, n \right\}.$$

It can be shown that the n-dimensional volume $\mathrm{vol}_n(\lambda_1 Q_1 + \cdots + \lambda_n Q_n)$ of this polytope is a homogeneous polynomial of degree n in $\lambda_1, \cdots, \lambda_n$. The coefficient of the term $\lambda_1 \times \cdots \times \lambda_n$ in this homogeneous polynomial is called the *mixed volume* of the polytopes $Q_1, \cdots, Q_n$, denoted by $\mathcal{M}(Q_1, \cdots, Q_n)$, or the mixed volume of the support of the system $P(\mathbf{x}) = (p_1(\mathbf{x}), \cdots, p_n(\mathbf{x}))$, denoted by $\mathcal{M}(S_1, \cdots, S_n)$ where $S_j = \mathrm{spt}(p_j)$ for $j = 1, \cdots, n$. Sometimes, when no ambiguities exist, it is called the mixed volume of $P(\mathbf{x})$.

In [CaRo], this root count was nicknamed the BKK bound after its inventors, Bernshtein [Bern], Kushnirenko [Ku76] and Khovanskii [Kh78]. In general, it provides a much tighter bound compared to variant Bézout bounds [MoSo, Shaf]. An apparent limitation of the theorem is that it only counts the isolated zeros of polynomial systems in $(\mathbb{C}^*)^n$ rather than all the isolated zeros in the affine space $\mathbb{C}^n$. For the purpose of finding all the isolated zeros of a polynomial system in $\mathbb{C}^n$, a generalized version of the theorem which counts the roots in $\mathbb{C}^n$ is strongly desirable. This problem was first attempted in [CaRo] where the notion of the *shadowed* sets was introduced and a bound for the root count in $\mathbb{C}^n$ was obtained. Later, a significantly much tighter bound was discovered in the following theorem.

Theorem 2 ([LiWa96]) *The root count in $\mathbb{C}^n$ of a polynomial system $P(\mathbf{x}) = (p_1(\mathbf{x}), \cdots, p_n(\mathbf{x}))$ with supports $S_j = \mathrm{spt}(p_j)$, $j = 1, \cdots, n$, is bounded above by the mixed volume $\mathcal{M}(S_1 \cup \{\mathbf{0}\}, \ldots, S_n \cup \{\mathbf{0}\})$.*

In other words, the root count in $\mathbb{C}^n$ of a polynomial system $P(\mathbf{x}) = (p_1(\mathbf{x}), \cdots, p_n(\mathbf{x}))$ is bounded above by the root count in $(\mathbb{C}^*)^n$ of a polynomial system in general position with support equal to $\mathrm{spt}(p_j) \cup \{\mathbf{0}\}$ for all $j = 1, \cdots, n$. As a corollary, when $\mathbf{0} \in \mathrm{spt}(p_j)$ for all $j = 1, \cdots, n$, namely, all $p_j(\mathbf{x})$ in $P(\mathbf{x})$ have constant terms, then the mixed volume of $P(\mathbf{x})$ also serves as a bound for the root count of $P(\mathbf{x})$ in $\mathbb{C}^n$, rather than in $(\mathbb{C}^*)^n$ as Theorem 1 asserts.

This theorem was further extended in several different ways [HuSt95, RoWa].

1.2 Polyhedral Homotopy

In light of Theorem 2 given in the last section, to find all isolated zeros of a given polynomial system $P(\mathbf{x}) = (p_1(\mathbf{x}), \cdots, p_n(\mathbf{x}))$ in $\mathbb{C}^n$ with support $S = (S_1, \cdots, S_n)$, we first append the monomial $\mathbf{x}^{\mathbf{0}} = 1$ to those p_j's which do not have constant terms. Followed by choosing coefficients of all the monomials in the system generically, a new system with support $\mathcal{A} = (\mathcal{A}_1, \cdots, \mathcal{A}_n)$ is obtained, where, of course, $\mathcal{A}_j = S_j \cup \{\mathbf{0}\}$ for $j = 1, \cdots, n$. In this section, we may assume $\mathbf{0} \in \mathcal{A}_j := \operatorname{spt}(p_j)$ and write

$$P(\mathbf{x}) = \begin{cases} p_1(\mathbf{x}) & = \displaystyle\sum_{\mathbf{q} \in \mathcal{A}_1} c_{1,\mathbf{q}} \mathbf{x}^{\mathbf{q}}, \\ & \vdots \\ p_n(\mathbf{x}) & = \displaystyle\sum_{\mathbf{q} \in \mathcal{A}_n} c_{n,\mathbf{q}} \mathbf{x}^{\mathbf{q}}, \end{cases} \tag{1.3}$$

with all generically chosen coefficients $c_{j,\mathbf{q}}$ for $\mathbf{q} \in \mathcal{A}_j$ and $j = 1, \cdots, n$, this system may be considered as a system in general position. Namely, there exists a polynomial system

$$G(\mathbf{d}) = (g_1(\mathbf{d}), \cdots, g_m(\mathbf{d})) \tag{1.4}$$

in the variables $\mathbf{d} = (d_{j,\mathbf{q}})$, for $\mathbf{q} \in \mathcal{A}_j$, $j = 1, \cdots, n$, such that polynomial systems with coefficient $\mathbf{c} = (c_{j,\mathbf{q}})$ for which $G(\mathbf{d})\Big|_{\mathbf{d} = \mathbf{c}} \neq \mathbf{0}$ reach the maximum root count in $(\mathbb{C}^*)^n$ for the support $\mathcal{A} = (\mathcal{A}_1, \cdots, \mathcal{A}_n)$.

Let t denote a new complex variable and consider the polynomial system $\widehat{P}(\mathbf{x}, t) = (\widehat{p}_1(\mathbf{x}, t), \cdots, \widehat{p}_n(\mathbf{x}, t))$ in the $n+1$ variables $(\mathbf{x}, t)$ given by

$$\widehat{P}(\mathbf{x}, t) = \begin{cases} \widehat{p}_1(\mathbf{x}, t) & = \displaystyle\sum_{\mathbf{q} \in \mathcal{A}_1} c_{1,\mathbf{q}} \mathbf{x}^{\mathbf{q}} t^{\omega_1(\mathbf{q})}, \\ & \vdots \\ \widehat{p}_n(\mathbf{x}, t) & = \displaystyle\sum_{\mathbf{q} \in \mathcal{A}_n} c_{n,\mathbf{q}} \mathbf{x}^{\mathbf{q}} t^{\omega_n(\mathbf{q})}, \end{cases} \tag{1.5}$$

where each $\omega_j : \mathcal{A}_j \to \mathbb{R}$ for $j = 1, \cdots, n$, known as a *lifting* is chosen generically on

$\mathcal{A}_j$. For a fixed τ, we rewrite the system in (1.5) as

$$\widehat{P}(\mathbf{x},\tau) = \begin{cases} \widehat{p}_1(\mathbf{x},\tau) &= \sum\limits_{\mathbf{q}\in\mathcal{A}_1} (c_{1,\mathbf{q}}\tau^{\omega_1(\mathbf{q})})\mathbf{x}^{\mathbf{q}}, \\ & \vdots \\ \widehat{p}_n(\mathbf{x},\tau) &= \sum\limits_{\mathbf{q}\in\mathcal{A}_n} (c_{n,\mathbf{q}}\tau^{\omega_n(\mathbf{q})})\mathbf{x}^{\mathbf{q}}. \end{cases}$$

This system is in general position if for $G(\mathbf{d})$ in (1.4),

$$\Theta(\tau) \stackrel{\text{def}}{=} G(\mathbf{d})\Big|_{\mathbf{d}\,=\,\left(c_{j,\mathbf{q}}\tau^{\omega_j(\mathbf{q})}\right)} \neq \mathbf{0}, \text{ for } \mathbf{q}\in\mathcal{A}_j \text{ and } j = 1,\cdots,n.$$

The system $\Theta(t) = \mathbf{0}$ can have only finitely many solutions since $\Theta(1) = G(\mathbf{c}) \neq \mathbf{0}$ implies $\Theta(t)$ is not identically $\mathbf{0}$. Let

$$t_1 = r_1e^{i\theta_1},\ \cdots,\ t_k = r_ke^{i\theta_k}$$

be all solutions of $\Theta(t) = \mathbf{0}$. Then, for any $\theta \neq \theta_j$, $j = 1,\cdots,k$, the systems $\overline{P}(\mathbf{x},t) = (\overline{p}_1(\mathbf{x},t),\cdots,\overline{p}_n(\mathbf{x},t))$ given by

$$\overline{P}(\mathbf{x},t) = \begin{cases} \overline{p}_1(\mathbf{x},t) &= \sum\limits_{\mathbf{q}\in\mathcal{A}_1} (c_{1,\mathbf{q}}e^{i\omega_1(\mathbf{q})\theta})\mathbf{x}^{\mathbf{q}}t^{\omega_1(\mathbf{q})}, \\ & \vdots \\ \overline{p}_n(\mathbf{x},t) &= \sum\limits_{\mathbf{q}\in\mathcal{A}_n} (c_{n,\mathbf{q}}e^{i\omega_n(\mathbf{q})\theta})\mathbf{x}^{\mathbf{q}}t^{\omega_n(\mathbf{q})}, \end{cases}$$

are in general position for all real $t > 0$ because

$$c_{j,\mathbf{q}}e^{i\omega_j(\mathbf{q})\theta}t^{\omega_j(\mathbf{q})} = c_{j,\mathbf{q}}(te^{i\theta})^{\omega_j(\mathbf{q})}$$

and

$$G(\mathbf{d})\Big|_{\mathbf{d}\,=\,\left(c_{j,\mathbf{q}}(te^{i\theta})^{\omega_j(\mathbf{q})}\right)} = \Theta(te^{i\theta}) \neq \mathbf{0},$$

Therefore, without loss of generality (choose an angle θ at random and change the coefficients $c_{j,\mathbf{q}}$ into $c_{j,\mathbf{q}}e^{i\omega_j(a)\theta}$ if necessary), we may suppose the systems $\widehat{P}(\mathbf{x},t)$ in (1.5) are in general position for all $t > 0$. Together with Lemma 1 given in the last section, it follows that for all $t > 0$ the systems $\widehat{P}(\mathbf{x}, t)$ in (1.5) have the same number of isolated zeros in $(\mathbb{C}^*)^n$. This number, say k, should equal to the mixed volume of the support of $P(\mathbf{x})$ in (1.3) by Theorem 1. We shall skip this fact temporarily and will reach this assertion at the end of this section.

Now consider $\widehat{P}(\mathbf{x}, t) = \mathbf{0}$ as a homotopy, known as the *polyhedral homotopy*, defined on $(\mathbb{C}^*)^n \times [0, 1]$. We have $\widehat{P}(\mathbf{x}, 1) = P(\mathbf{x})$, and the zero set of this homotopy is made up of k homotopy paths, say, $\mathbf{x}^1(t), \cdots, \mathbf{x}^k(t)$, since for each $0 < t \leq 1$, $\widehat{P}(\mathbf{x}, t)$ has exactly k isolated zeros from the argument given above. Since each $\widehat{p}_j(\mathbf{x}, t)$ has nonzero constant term for all $j = 1, \cdots, n$, by a standard application of generalized Sard's Theorem [ChMPYo], all those homotopy paths are smooth with no bifurcations. Therefore, both Property 2 (*Smoothness*) and Property 3 (*Accessibility*) mentioned at the beginning of this chapter hold for this homotopy. However, when $t = 0$, $\widehat{P}(\mathbf{x}, 0) \equiv \mathbf{0}$. Consequently, the starting points $\mathbf{x}^1(0), \cdots, \mathbf{x}^k(0)$ of those homotopy paths can not be identified, causing the breakdown of the standard homotopy continuation algorithm. This major obstacle can be overcome by the devise we describe below.

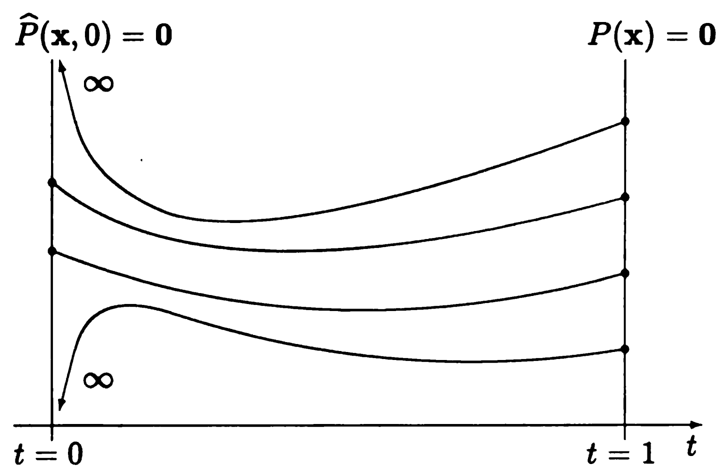

Figure 1.1: Solution curves of $\widehat{P}(\mathbf{x}, t) = \mathbf{0}$

For $\alpha = (\alpha_1, \cdots, \alpha_n) \in \mathbb{R}^n$, consider the transformation $\mathbf{x} = \mathbf{y}t^{\alpha}$ defined by

$$\begin{aligned} x_1 &= y_1 t^{\alpha_1}, \\ &\vdots \\ x_n &= y_n t^{\alpha_n}. \end{aligned} \tag{1.6}$$

For $\mathbf{q} = (q_1, \cdots, q_n) \in \mathbb{N}^n$, we have

$$\begin{aligned} \mathbf{x}^{\mathbf{q}} &:= x_1^{q_1} \cdots x_n^{q_n} \\ &= (y_1 t^{\alpha_1})^{q_1} \cdots (y_n t^{\alpha_n})^{q_n} \\ &= y_1^{q_1} \cdots y_n^{q_n}\, t^{\alpha_1 q_1 + \cdots + \alpha_n q_n} \\ &= \mathbf{y}^{\mathbf{q}}\, t^{\langle \alpha, \mathbf{q} \rangle}. \end{aligned} \tag{1.7}$$

Here, $\langle \cdot, \cdot \rangle$ stands for the usual inner product in $\mathbb{R}^n$. Substituting (1.7) into (1.5) yields, for $j = 1, \cdots, n$,

$$\begin{aligned} h_j^{\alpha}(\mathbf{y}, t) &\stackrel{\text{def}}{=} \widehat{p}_j(\mathbf{y}t^{\alpha}, t) = \sum_{\mathbf{q} \in \mathcal{A}_j} c_{j,\mathbf{q}} \mathbf{y}^{\mathbf{q}} t^{\langle \alpha, \mathbf{q} \rangle} t^{\omega_j(\mathbf{q})} \\ &= \sum_{\mathbf{q} \in \mathcal{A}_j} c_{j,\mathbf{q}} \mathbf{y}^{\mathbf{q}} t^{\langle (\alpha, 1), (\mathbf{q}, \omega_j(\mathbf{q})) \rangle} \\ &= \sum_{\mathbf{q} \in \mathcal{A}_j} c_{j,\mathbf{q}} \mathbf{y}^{\mathbf{q}} t^{\langle \widehat{\alpha}, \widehat{\mathbf{q}} \rangle}, \end{aligned} \tag{1.8}$$

where $\widehat{\mathbf{q}} := (\mathbf{q}, \omega_j(\mathbf{q}))$ for $\mathbf{q} \in \mathcal{A}_j$ and $\widehat{\alpha} := (\alpha, 1) \in \mathbb{R}^{n+1}$. The new homotopy

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

retains most of the properties of the homotopy $\widehat{P}(\mathbf{x}, t) = \mathbf{0}$, in particular, $H^{\alpha}(\mathbf{y}, 1) = \widehat{P}(\mathbf{y}, 1) = P(\mathbf{y})$ and both Properties 2 (*Smoothness*) and 3 (*Accessibility*) stand. Let

$$m_j \stackrel{\text{def}}{=} \min_{\mathbf{q} \in \mathcal{A}_j} \langle \widehat{\alpha}, \widehat{\mathbf{q}} \rangle, \; j = 1, \cdots, n$$

and define the homotopy

$$\overline{H}^{\alpha}(\mathbf{y}, t) = (\overline{h}_1^{\alpha}(\mathbf{y}, t), \cdots, \overline{h}_n^{\alpha}(\mathbf{y}, t)) = \mathbf{0}$$

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

$$\begin{aligned} \overline{h}_j^{\alpha}(\mathbf{y},t) &= t^{-m_j} h_j^{\alpha}(\mathbf{y},t) = \sum_{\mathbf{q}\in\mathcal{A}_j} c_{j,\mathbf{q}} \mathbf{y}^{\mathbf{q}} t^{\langle \hat{\alpha},\hat{\mathbf{q}}\rangle - m_j} \\ &= \sum_{\substack{\mathbf{q}\in\mathcal{A}_j \\ \langle \hat{\alpha},\hat{\mathbf{q}}\rangle = m_j}} c_{j,\mathbf{q}} \mathbf{y}^{\mathbf{q}} + \sum_{\substack{\mathbf{q}\in\mathcal{A}_j \\ \langle \hat{\alpha},\hat{\mathbf{q}}\rangle > m_j}} c_{j,\mathbf{q}} \mathbf{y}^{\mathbf{q}} t^{\langle \hat{\alpha},\hat{\mathbf{q}}\rangle - m_j}. \end{aligned} \tag{1.9}$$

Evidently, for any path $\tilde{\mathbf{y}}(t)$ defined on $[0,1]$, we have, for all $t > 0$,

$$\overline{H}^{\alpha}(\tilde{\mathbf{y}}(t),t) = \mathbf{0} \quad \Longleftrightarrow \quad H^{\alpha}(\tilde{\mathbf{y}}(t),t) = \mathbf{0}.$$

Therefore, the zero set of $\overline{H}^{\alpha}(\mathbf{y},t) = \mathbf{0}$ consists of the same homotopy paths of the homotopy $H^{\alpha}(\mathbf{y},t) = \mathbf{0}$ in (1.8). The difference is, the starting points of the homotopy paths considered in the homotopy $\overline{H}^{\alpha}(\mathbf{y},t) = \mathbf{0}$ are solutions of the system

$$\overline{H}^{\alpha}(\mathbf{y},0) = \begin{cases} \overline{h}_1^{\alpha}(\mathbf{y},0) = \displaystyle\sum_{\substack{\mathbf{q}\in\mathcal{A}_1 \\ \langle \hat{\alpha},\hat{\mathbf{q}}\rangle = m_1}} c_{1,\mathbf{q}} \mathbf{y}^{\mathbf{q}} = \mathbf{0}, \\ \qquad \vdots \\ \overline{h}_n^{\alpha}(\mathbf{y},0) = \displaystyle\sum_{\substack{\mathbf{q}\in\mathcal{A}_n \\ \langle \hat{\alpha},\hat{\mathbf{q}}\rangle = m_n}} c_{n,\mathbf{q}} \mathbf{y}^{\mathbf{q}} = \mathbf{0}. \end{cases} \tag{1.10}$$

As shown below, when this system is in certain desired form, its isolated nonsingular solutions that lie in $(\mathbb{C}^*)^n$ can be constructively identified. In those situations, Property 1 (*Triviality*) becomes partially valid for those homotopy paths of $\overline{H}^{\alpha}(\mathbf{y},t) = \mathbf{0}$ that emanate from those nonsingular solutions of (1.10) in $(\mathbb{C}^*)^n$, and we may follow those paths to reach a partial set of isolated zeros of $P(\mathbf{y})$ at $t = 1$.

The system (1.10) is known as the *binomial system* if each $\overline{h}_j^{\alpha}(\mathbf{y},0)$ consists of exactly two terms, that is,

$$\begin{aligned} \overline{h}_1^{\alpha}(\mathbf{y},0) &= c_1 \mathbf{y}^{\mathbf{a}_1} + c_1' \mathbf{y}^{\mathbf{a}_1'} = \mathbf{0}, \\ &\vdots \\ \overline{h}_n^{\alpha}(\mathbf{y},0) &= c_n \mathbf{y}^{\mathbf{a}_n} + c_n' \mathbf{y}^{\mathbf{a}_n'} = \mathbf{0}, \end{aligned} \tag{1.11}$$

where $\mathbf{a}_j, \mathbf{a}_j' \in \mathcal{A}_j$, $c_j = c_{j,\mathbf{a}_j}$ and $c_j' = c_{j,\mathbf{a}_j'}$ for $j = 1, \cdots, n$. And in this case, $C = (\{\mathbf{a}_1, \mathbf{a}_1'\}, \cdots, \{\mathbf{a}_n, \mathbf{a}_n'\})$ is called *a* **mixed cell** *of $\mathcal{A}$ with inner normal α*, or

$\widehat{C} = (\{\widehat{\mathbf{a}}_1, \widehat{\mathbf{a}}_1'\}, \cdots, \{\widehat{\mathbf{a}}_n, \widehat{\mathbf{a}}_n'\})$ *is a lifted mixed cell with inner normal* $\widehat{\alpha} = (\alpha, 1)$. Alternatively, we say α *supports cell* C or $\widehat{\alpha}$ *supports lifted cell* $\widehat{C}$. $\overline{H}^{\alpha}(\mathbf{y}, t) = \mathbf{0}$ is called the *polyhedral homotopy on cell* C *induced by* ω and α. The collection of all mixed cells is called the **mixed cell configuration** $\mathcal{M}_\omega$ *induced by* ω.

Proposition 1 *The binomial system in (1.11) has*

$$k_\alpha := \left| \det \begin{pmatrix} \mathbf{a}_1 - \mathbf{a}_1' \\ \vdots \\ \mathbf{a}_n - \mathbf{a}_n' \end{pmatrix} \right| \tag{1.12}$$

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

The number k_α is called the *volume* of the mixed cell $(\{\mathbf{a}_1, \mathbf{a}_1'\}, \cdots, \{\mathbf{a}_n, \mathbf{a}_n'\})$. The proof of this proposition is constructive and therefore provides an algorithm for solving the binomial system (1.11) in $(\mathbb{C}^*)^n$. We will come back to this matter in the next section.

In summary, for given $\alpha = (\alpha_1, \cdots, \alpha_n) \in \mathbb{R}^n$, by changing variables $\mathbf{x} = \mathbf{y}t^\alpha$, as in (1.6), in the homotopy $\widehat{P}(\mathbf{x}, t) = (\widehat{p}_1(\mathbf{x}, t), \cdots, \widehat{p}_n(\mathbf{x}, t)) = \mathbf{0}$ in (1.5), the homotopy $H^\alpha(\mathbf{y}, t) = (h_1^\alpha(\mathbf{y}, t), \cdots, h_n^\alpha(\mathbf{y}, t)) = \mathbf{0}$ in (1.8) is obtained, where $h_j^\alpha(\mathbf{y}, t) = \widehat{p}_j(\mathbf{y}t^\alpha, t)$. Followed by factoring out the lowest power t^{m_j} of t among all monomials in each individual $h_j^\alpha(\mathbf{y}, t) = 0$ for $j = 1, \cdots, n$ we arrive at the homotopy $\overline{H}^{\alpha}(\mathbf{y}, t) = \mathbf{0}$ in (1.9). When the start system $\overline{H}^{\alpha}(\mathbf{y}, 0) = \mathbf{0}$ of this homotopy is binomial, its nonsingular solutions in $(\mathbb{C}^*)^n$, k_α (as given in (1.12)) of them, become available. We may then follow those homotopy paths of $\overline{H}^{\alpha}(\mathbf{y}, t) = \mathbf{0}$ originated from those k_α regular solutions of $\overline{H}^{\alpha}(\mathbf{y}, 0) = \mathbf{0}$ in $(\mathbb{C}^*)^n$, and reach k_α isolated zeros of $P(\mathbf{y})$ at $t = 1$. Worth notifying here is the fact that the system $P(\mathbf{x})$, or $P(\mathbf{y})$, stays invariant at $t = 1$ during the process.

Now, the existence of $\alpha \in \mathbb{R}^n$ for which the start system $\overline{H}^{\alpha}(\mathbf{y}, 0) = \mathbf{0}$ is binomial is warranted by the following

Proposition 2 *For all the real functions $\omega_j : \mathcal{A}_j \to \mathbb{R}$, $j = 1, \cdots, n$ being generically chosen, there must exist $\alpha \in \mathbb{R}^n$, for which the start system $\overline{H}^{\alpha}(\mathbf{y}, 0) = \mathbf{0}$ of the homotopy $\overline{H}^{\alpha}(\mathbf{y}, t) = \mathbf{0}$ in (1.9) is binomial with a nonempty set of nonsingular solutions in $(\mathbb{C}^*)^n$, i.e., $k_\alpha \neq 0$ in (1.12).*

The assertion of this proposition was proved implicitly in [HuSt95] with terminologies and machineries developed in combinatorial geometry, such as, *random liftings*, *fine mixed subdivisions*, *lower facets of convex polytopes*, etc., see also [Li97]. Here, we elect to reinterpret the result without those specialized terms.

Now, different $\alpha \in \mathbb{R}^n$ given in Proposition 2 lead to different homotopy $\overline{H}^{\alpha}(\mathbf{y}, t) = \mathbf{0}$ in (1.9). Henceforth, following homotopy paths of those different homotopies will reach different sets of isolated zeros of $P(\mathbf{y})$. By taking the Puiseux series expansions of those homotopy paths of $\overline{H}^{\alpha}(\mathbf{y}, t) = \mathbf{0}$ originated at $(\mathbb{C}^*)^n$ into consideration, it is not hard to see that those different sets of isolated zeros of $P(\mathbf{y})$ reached by different sets of homotopy paths are actually disjoint from each other. Most importantly, it can be shown that every isolated zero of $P(\mathbf{y})$ can be obtained this way by following certain homotopy curve of the homotopy $\overline{H}^{\alpha}(\mathbf{y}, t) = \mathbf{0}$ associated with certain $\alpha \in \mathbb{R}^n$ given by Proposition 2. Thus the total number of isolated zeros of $P(\mathbf{y})$ must equal to the sum of those k_α's corresponding to all the possible α's provided by Proposition 2, respectively. In [HuSt95], it was shown that this sum actually equal to the mixed volume of $P(\mathbf{y})$. This yields an alternative proof of Theorem 1, it is very different from Bernshtein's original approach [Bern].

Solving Binomial Systems

Another major step in solving polynomial systems by using the polyhedral homotopy method as we described in the previous section is finding the solutions of the corresponding binomial system

$$\begin{aligned} c_1 \mathbf{y}^{\mathbf{a}_1} + c_1' \mathbf{y}^{\mathbf{a}_1'} &= 0, \\ &\vdots \\ c_n \mathbf{y}^{\mathbf{a}_n} + c_n' \mathbf{y}^{\mathbf{a}_n'} &= 0, \end{aligned} \tag{1.13}$$

produced by the mixed cell $(\{\mathbf{a}_1, \mathbf{a}_1'\}, \cdots, \{\mathbf{a}_n, \mathbf{a}_n'\})$ as in (1.11). We now discuss the method for solving (1.13) in $(\mathbb{C}^*)^n$. Let

$$v_j = \mathbf{a}_j - \mathbf{a}_j', \quad j = 1, \cdots, n,$$

and, with $\mathbf{y} \in (\mathbb{C}^*)^n$ in mind, we rewrite the system (1.13) as

$$\begin{aligned} \mathbf{y}^{v_1} &= b_1, \\ &\vdots \\ \mathbf{y}^{v_n} &= b_n, \end{aligned} \tag{1.14}$$

where $b_j = -\dfrac{c_j'}{c_j}$ for $j = 1, \cdots, n$. Let

$$V = \left[\; v_1^T \;\middle|\; \cdots \;\middle|\; v_n^T \;\right] \tag{1.15}$$

and for brevity, write

$$\mathbf{y}^V = (\mathbf{y}^{v_1}, \cdots, \mathbf{y}^{v_n}) \quad \text{and} \quad \mathbf{b} = (b_1, \cdots, b_n).$$

Then, (1.14) becomes,

$$\mathbf{y}^V = \mathbf{b}. \tag{1.16}$$

With this notation, it is easy to verify that for an $n \times n$ integral matrix U, we have,

$$(\mathbf{y}^V)^U = \mathbf{y}^{(VU)}.$$

Now, when the matrix V in (1.15) is an upper triangular matrix, i.e.,

$$V = \begin{bmatrix} v_{11} & \cdots & v_{1n} \\ & \ddots & \vdots \\ 0 & & v_{nn} \end{bmatrix},$$

then the equation in (1.16) becomes

$$\begin{array}{lcl} y_1^{v_{11}} & = & b_1, \\ y_1^{v_{12}} y_2^{v_{22}} & = & b_2, \\ & \vdots & \\ y_1^{v_{1n}} y_2^{v_{2n}} \cdots y_n^{v_{nn}} & = & b_n. \end{array} \tag{1.17}$$

By forward substitutions, all the solutions of the system (1.17) in $(\mathbb{C}^*)^n$ can be found, and the total number of solutions is $|v_{11}| \times \cdots \times |v_{nn}| = |\det V|$.

In general, we may upper triangularize V in (1.15) by the following process. Recall that the greatest common divisor d of two nonzero integers a and b, denoted by $\gcd(a, b)$, can be written as

$$d := \gcd(a, b) = k\,a + \ell\,b,$$

for certain nonzero integers k and ℓ. Let

$$B = \begin{bmatrix} k & \ell \\ -\frac{b}{d} & \frac{a}{d} \end{bmatrix}.$$

We have $\det(B) = 1$, and

$$B \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} k & \ell \\ -\frac{b}{d} & \frac{a}{d} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} d \\ 0 \end{bmatrix}.$$

Similar to using Givens rotation to produce zeros in a matrix for its QR factorization, the matrix B may be used to upper triangularize V as follows. For $\mathbf{v} \in \mathbb{Z}^n$, let a and b be its i^{th} and the j^{th} (nonzero) components where $i < j$, that is,

$$\mathbf{v} = \begin{array}{cl} \begin{bmatrix} \vdots \\ a \\ \vdots \\ b \\ \vdots \end{bmatrix} & \begin{array}{l} \\ \to i^{\text{th}} \\ \\ \to j^{\text{th}}. \\ \\ \end{array} \end{array}$$

Let $d := \gcd(a, b)$, and

$$U_{ij} := \left[\begin{array}{ccc|ccccc|ccc}
1 & & & & & & & & & & \\
& \ddots & & & & & & & & & \\
& & 1 & & & & & & & & \\
\hline
& & & k & & & & \ell & & & \\
& & & & 1 & & & & & & \\
& & & & & \ddots & & & & & \\
& & & & & & 1 & & & & \\
& & & -\frac{b}{d} & & & & \frac{a}{d} & & & \\
\hline
& & & & & & & & 1 & & \\
& & & & & & & & & \ddots & \\
& & & & & & & & & & 1
\end{array}\right]
\begin{array}{l}
\\ \\ \\ i^{\text{th}} \\ \\ \\ \\ j^{\text{th}} \\ \\ \\ .
\end{array}
\tag{1.18}$$

Evidently, U_{ij} is an integeral matrix with $|\det U_{ij}| = 1$ and

$$U_{ij}\,\mathbf{v} = \begin{bmatrix} \vdots \\ d \\ \vdots \\ 0 \\ \vdots \end{bmatrix}
\begin{array}{l} \\ i^{\text{th}} \\ \\ j^{\text{th}} \\ . \end{array}$$

Thus a series of matrices in the form of U_{ij} in (1.18) may be used to successively produce zeros in the lower triangular part of the matrix V in (1.15), resulting in an upper triangular matrix. In simple terms, we may construct an integeral matrix U , as a product of those U_{ij}'s, with $|\det U| = 1$ and UV is an upper triangular integeral matrix.

Now, as mentioned above, the solutions of the system

$$(\mathbf{z}^U)^V = \mathbf{z}^{UV} = \mathbf{b} \tag{1.19}$$

in $(\mathbb{C}^*)^n$ can be found by forward substitutions, since UV is an upper triangular integeral matrix. And the total number of solutions in $(\mathbb{C}^*)^n$ is

$$|\det(UV)| = |\det U| \cdot |\det V| = |\det V|.$$

By letting $\mathbf{y} = \mathbf{z}^U$ for each solution $\mathbf{z}$ of (1.19) in $(\mathbb{C}^*)^n$, we obtain all the solutions of the system (1.19) in $(\mathbb{C}^*)^n$, and hence, solve the system (1.13) in $(\mathbb{C}^*)^n$.

1.3 Numerical Stability of Polyhedral Homotopies

To solve a specific polynomial system $P^*(\mathbf{x})$ of support $\mathrm{spt}(P^*) = \mathcal{A} = (\mathcal{A}_1, \cdots, \mathcal{A}_n)$, where $\mathcal{A}_j \subset \mathbb{Z}^n$, $j = 1, \cdots, n$, we first solve a corresponding generic system $P(\mathbf{x})$ with $\mathrm{spt}(P) = \mathrm{spt}(P^*)$, and then use the *linear homotopy* $\mathcal{H}(\mathbf{x}, t) = (1-t)P(\mathbf{x}) + tP^*(\mathbf{x}) = \mathbf{0}$ to solve $P^*(\mathbf{x}) = \mathbf{0}$. To solve the generic system $P(\mathbf{x}) = \mathbf{0}$, we use the polyhedral homotopy method.

Let's briefly recall the procedure: A random lifting $\omega = (\omega_1, \cdots, \omega_n)$ will induce the *mixed cell configuration* $\mathcal{M}_\omega$ of support $\mathcal{A}$. For a cell $C = (C_1, \ldots, C_n) \in \mathcal{M}_\omega$ with *inner normal* α, the corresponding polyhedral homotopy $\overline{H}^\alpha(\mathbf{y}, t) = \mathbf{0}$ yields that $\overline{H}^\alpha(\mathbf{y}, 0)$ is a binomial system with support C and $\overline{H}^\alpha(\mathbf{y}, 1) = P(\mathbf{y})$. If powers of t in the polyhedral homotopy $\overline{H}^\alpha(\mathbf{y}, t) = \mathbf{0}$ are too high or too close to zero, instability might occur in tracing solution curves of $\overline{H}^\alpha(\mathbf{y}, t) = \mathbf{0}$ unless very small step sizes are taken. For the standard prediction-correction method [AlGe93] in tracing a curve, the number of steps required for homotopy curves with high powers of t is usually much greater than those for other curves. This may severely reduce the efficiency of the algorithm. A more serious problem can occur when curve-tracings are trapped into "valleys" where the values of the polyhedral homotopy are numerically zeros, but no solution curves exist inside the "valleys".

Example 1 Consider the polynomial system $P(\mathbf{x}) = (p_1(\mathbf{x}), p_2(\mathbf{x})) = \mathbf{0}$ with $\mathbf{x} = (x_1, x_2)$, where

$$\begin{cases} p_1(\mathbf{x}) = c_1 x_1^2 + c_2 x_2^2 + c_3 = 0, \\ p_2(\mathbf{x}) = c_4 x_1^3 x_2^3 + c_5 x_1 + c_6 x_2 = 0, \end{cases}$$

with support $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$, where $\mathcal{A}_1 = \{(2,0), (0,2), (0,0)\}$ and $\mathcal{A}_2 = \{(3,3), (1,0), (0,1)\}$. This system has mixed volume $\mathcal{M}(\mathcal{A}) = 12$.

Choose a random lifting $\omega = (\omega_1, \omega_2)$ where $\omega_1 : \mathcal{A}_1 \to \mathbb{R}$ and $\omega_2 : \mathcal{A}_2 \to \mathbb{R}$ with

$$\omega_1((2,0)) = 0.655416, \quad \omega_1((0,2)) = 0.200995, \quad \omega_1((0,0)) = 0.893622,$$
$$\omega_2((3,3)) = 0.281886, \quad \omega_2((1,0)) = 0.525000, \quad \omega_2((0,1)) = 0.314127.$$

Then the mixed cell configuration $\mathcal{M}_\omega$ in the fine mixed subdivision of $\mathcal{A}$ induced by ω consists of two mixed cells:

$$C = (\{(2,0),(0,2)\},\{(3,3),(1,0)\}) \quad \text{with inner normal } \alpha,$$
$$C' = (\{(0,2),(0,0)\},\{(1,0),(0,1)\}) \quad \text{with inner normal } \beta.$$

To construct the system $P(\mathbf{x}) = \mathbf{0}$, we choose the following set of randomly generated coefficients,

$$c_1 = -0.434847 - 0.169859i, \quad c_2 = 0.505911 + 0.405431i,$$
$$c_3 = 0.0738596 + 0.177798i, \quad c_4 = -0.0906755 + 0.208825i,$$
$$c_5 = 0.175313 - 0.163549i, \quad c_6 = 0.527922 - 0.364841i.$$

The polyhedral homotopy induced by ω, α, C is $\overline{H}^\alpha(\mathbf{y},t) = (\overline{h}_1^\alpha(\mathbf{y},t), \overline{h}_2^\alpha(\mathbf{y},t)) = \mathbf{0}$, where

$$\begin{cases} \overline{h}_1^\alpha(\mathbf{y},t) &= c_1 y_1^2 + c_2 y_2^2 + c_3 t^{50.63523}, \\ \overline{h}_2^\alpha(\mathbf{y},t) &= c_4 y_1^3 y_2^3 + c_5 y_1 + c_6 y_2 t^2. \end{cases}$$

There are ten solution curves of $\overline{H}^\alpha(\mathbf{y},t) = \mathbf{0}$ emanating from ten solutions of $\overline{H}^\alpha(\mathbf{y},0) = \mathbf{0}$. At $t = 0.65$, five of those ten curves have phase space tangent vectors $(\frac{dy_1}{dt}, \frac{dy_2}{dt})$ all roughly pointing to $\mathbf{y} = (0,0)$ at the points on the curves. For the standard prediction-correction method, starting at these points, the prediction step with step size 0.025 will give the predicted points close to $(\mathbf{y},t) = (0,0,0.675)$ for all those five curves. Since $t^{50.63523}$ is about 10^{-9} for $t = 0.675$, the function values of $H(\mathbf{y},t)$ in a very small neighborhood of $(0,0,0.675)$ (the "valley") are almost zero. Starting from these predicted points at $t = 0.675$, Newton iterations for the correction step will converge to the "valley" rather than the points on the curves at $t = 0.675$. But there are no solution curves of $H^\alpha(\mathbf{y},t) = \mathbf{0}$ passing through the "valley". $\square$

CHAPTER 2

Balancing the Powers by the Sandwich Model

Given the mixed cell configuration $\mathcal{M}_\omega$ induced by a generic lifting $\omega = (\omega_1, \cdots, \omega_n)$, assume $C = (C_1, \cdots, C_n) \in \mathcal{M}_\omega$. We shall use the short hand notations:

$$\widehat{\mathbf{q}} := (\mathbf{q}, \omega_i(\mathbf{q})),\ \mathbf{q} \in \mathcal{A}_i \text{ for a lifted point,}$$
$$\widehat{\mathcal{A}}_i := \{\widehat{\mathbf{q}} \,|\, \mathbf{q} \in \mathcal{A}_i\} \text{ for a lifted support,}$$
$$\widehat{C}_i := \{\widehat{\mathbf{q}} \,|\, \mathbf{q} \in C_i\}, \text{ and } \widehat{C} := (\widehat{C}_1, \cdots, \widehat{C}_n) \text{ a lifted cell.}$$

For a vector $\widehat{\alpha} := (\alpha, 1) \in \mathbb{R}^{n+1}$, we say "$\widehat{\alpha}$ supports $\widehat{C}$", "α supports C", or α is the inner normal of C when the hyperplane with inner normal $\widehat{\alpha}$ supports $\mathrm{conv}(\widehat{\mathcal{A}}_1), \cdots, \mathrm{conv}(\widehat{\mathcal{A}}_n)$ at $\widehat{C}_1, \cdots, \widehat{C}_n$, respectively.

2.1 A Fundamental Observation

When constructing the polyhedral homotopy for cell C with inner normal α, we actually execute the following transformations:

(here vectors are regarded as in columns)

$$\begin{array}{llllll}
\text{exponents of} & (\mathbf{x},t) \text{ in } \widehat{p}_i & & (\mathbf{y},t) \text{ in } h_i^\alpha & & (\mathbf{y},t) \text{ in } \overline{h}_i^\alpha \\
& \widehat{\mathbf{q}} \in \widehat{\mathcal{A}}_i & \stackrel{\mathbf{U}}{\longmapsto} & (\mathbf{q}, \langle \widehat{\alpha}, \widehat{\mathbf{q}} \rangle) & \stackrel{\mathbf{S}_i}{\longmapsto} & (\mathbf{q}, \langle \widehat{\alpha}, \widehat{\mathbf{q}} \rangle - m_i), \\
(\text{ see} & (1.5) & & (1.8) & & (1.9) \quad)
\end{array} \tag{2.1}$$

where

$$\mathbf{U} := \left[\begin{array}{ccc|c} 1 & & 0 & 0 \\ & \ddots & & \vdots \\ 0 & & 1 & 0 \\ \hline \alpha_1 & \cdots & \alpha_n & 1 \end{array}\right], \quad \mathbf{S}_i : \begin{array}{l} \text{shift of the last coordinate of elements} \\ \text{in } \mathbf{U}\widehat{\mathcal{A}}_i \text{ by } m_i := \min_{\mathbf{q}\in\mathcal{A}_i} \langle \widehat{\alpha}, \widehat{\mathbf{q}} \rangle. \end{array}$$

Clearly, $\det \mathbf{U} = 1$, $\mathbf{U}^{-1} = \left[\begin{array}{ccc|c} 1 & & 0 & 0 \\ & \ddots & & \vdots \\ 0 & & 1 & 0 \\ \hline -\alpha_1 & \cdots & -\alpha_n & 1 \end{array}\right]$, and $\mathbf{U}^{-T}\widehat{\alpha} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$.

When a vector $\widehat{\beta} := (\beta, 1) \in \mathbb{R}^{n+1}$ supports another cell $\widehat{C}'$, then

$$\mathbf{U}^{-T}\widehat{\beta} = \begin{bmatrix} \beta_1 - \alpha_1 \\ \vdots \\ \beta_n - \alpha_n \\ 1 \end{bmatrix}$$

with the last coordinate > 0. For

$$\mathbf{s}_i := \begin{bmatrix} 0 \\ \vdots \\ 0 \\ m_i \end{bmatrix},$$

we have $\mathbf{U}^{-1}\mathbf{s}_i = \mathbf{s}_i$. Now, for any $\widehat{\gamma}$, $\widehat{\mathbf{q}}$, we have $\langle \widehat{\gamma}, \widehat{\mathbf{q}} \rangle = \langle \mathbf{U}^{-T}\widehat{\gamma}, \mathbf{U}\widehat{\mathbf{q}} \rangle$. It follows that

$$\begin{aligned}
\langle \mathbf{U}^{-T}\widehat{\gamma}, \mathbf{S}_i\mathbf{U}\widehat{\mathbf{q}} \rangle &= \langle \mathbf{U}^{-T}\widehat{\gamma}, \mathbf{U}\widehat{\mathbf{q}} - \mathbf{s}_i \rangle \\
&= \langle \mathbf{U}^{-T}\widehat{\gamma}, \mathbf{U}\widehat{\mathbf{q}} \rangle - \langle \mathbf{U}^{-T}\widehat{\gamma}, \mathbf{s}_i \rangle \\
&= \langle \widehat{\gamma}, \widehat{\mathbf{q}} \rangle - \langle \widehat{\gamma}, \mathbf{U}^{-1}\mathbf{s}_i \rangle \\
&= \langle \widehat{\gamma}, \widehat{\mathbf{q}} \rangle - \langle \widehat{\gamma}, \mathbf{s}_i \rangle,
\end{aligned}$$

So,

$$\min_{\mathbf{q}\in\mathcal{A}_i}\langle \mathbf{U}^{-T}\widehat{\gamma}, \mathbf{S}_i\mathbf{U}\widehat{\mathbf{q}}\rangle = \min_{\mathbf{q}\in\mathcal{A}_i}\langle \widehat{\gamma}, \widehat{\mathbf{q}}\rangle - \langle \widehat{\gamma}, \mathbf{s}_i\rangle.$$

Therefore, $\mathbf{U}^{-T}\widehat{\alpha}$ supports $\mathbf{U}\widehat{C}$ (thus, $(\mathbf{S}_1\mathbf{U}\widehat{C}_1, \ldots, \mathbf{S}_n\mathbf{U}\widehat{C}_n)$), and at the same time, $\mathbf{U}^{-T}\widehat{\beta}$ also supports $\mathbf{U}\widehat{C}'$ (thus, $(\mathbf{S}_1\mathbf{U}\widehat{C}'_1, \ldots, \mathbf{S}_n\mathbf{U}\widehat{C}'_n)$). □

It is not clear how powers of t will be distributed in the homotopy $\overline{H}^{\alpha}$ until $\mathbf{S}_i\mathbf{U}$ transformations applied. The consequence of applying $\mathbf{U}$ is to "turn cell C horizontal" without changing existing mixed cells. When the powers of t in homotopy $\overline{H}^{\alpha}$ are considered as another sets of liftings, obviously it gives the same mixed cells. On the other hand, if we slightly alter lifting values, mixed cell configuration can still be the same but powers of t may become more desirable. Accordingly, there may exist better liftings that preserves the mixed cell configuration previously found but induces polyhedral homotopies with more balanced powers of t.

Example 2 Given a polynomial system $P = (p_1, p_2)$ with

$$\begin{array}{lcl} \mathrm{spt}(p_1) = \mathcal{A}_1 = \{\mathbf{a}, \mathbf{b}, \mathbf{c}\}, & \text{and} & \mathbf{a} = (0,1), \quad \mathbf{b} = (0,0), \quad \mathbf{c} = (1,0), \\ \mathrm{spt}(p_2) = \mathcal{A}_2 = \{\mathbf{d}, \mathbf{e}, \mathbf{f}\}, & & \mathbf{d} = (0,1), \quad \mathbf{e} = (1,0), \quad \mathbf{f} = (1,1). \end{array}$$

A lifting $\omega = (\omega_1, \omega_2)$, $\omega_1 : \mathcal{A}_1 \longrightarrow \mathbb{R}$, $\omega_2 : \mathcal{A}_2 \longrightarrow \mathbb{R}$,

$$\begin{cases} \omega_1(\mathbf{a}) = 1.1, & \omega_1(\mathbf{b}) = 1, & \omega_1(\mathbf{c}) = 11, \\ \omega_2(\mathbf{d}) = 12, & \omega_2(\mathbf{e}) = 2.1, & \omega_2(\mathbf{f}) = 2, \end{cases}$$

will result in two mixed cells:

$$C = (\{\mathbf{a}, \mathbf{b}\}, \{\mathbf{d}, \mathbf{f}\}) \text{ with inner normal } \alpha = (10, -0.1),$$
$$C' = (\{\mathbf{b}, \mathbf{c}\}, \{\mathbf{e}, \mathbf{f}\}) \text{ with inner normal } \beta = (-10, 0.1).$$

Since $\widehat{\alpha} = (10, -0.1, 1)$ supports $\widehat{C}$, in the homotopy $H^{\alpha} = (h_1^{\alpha}, h_2^{\alpha})$ induced by α, the distribution of powers of t are as follows:

$$\left\{\begin{array}{l} \mathbf{a}:\langle(10,-0.1,1),\ (0,1,1.1)\rangle=1 \\ \mathbf{b}:\langle(10,-0.1,1),\ (0,0,1)\rangle=1 \quad \text{in } h_1^\alpha, \\ \mathbf{c}:\langle(10,-0.1,1),\ (1,0,11)\rangle=21 \end{array}\right. \quad \left\{\begin{array}{l} 0 \\ 0 \quad \text{in } \overline{h}_1^\alpha \\ 20 \end{array}\right.$$

$$\left\{\begin{array}{l} \mathbf{d}:\langle(10,-0.1,1),\ (0,1,12)\rangle=11.9 \\ \mathbf{e}:\langle(10,-0.1,1),\ (1,0,2.1)\rangle=12.1 \quad \text{in } h_2^\alpha, \\ \mathbf{f}:\langle(10,-0.1,1),\ (1,1,2)\rangle=11.9 \end{array}\right. \quad \left\{\begin{array}{l} 0 \\ 0.2 \quad \text{in } \overline{h}_2^\alpha \\ 0 \end{array}\right.$$

Namely, we have

$$\left\{\begin{array}{l} \overline{h}_1^\alpha(\mathbf{y},t)=c_{\mathbf{a}}\mathbf{y}^{\mathbf{a}}+c_{\mathbf{b}}\mathbf{y}^{\mathbf{b}}+c_{\mathbf{c}}\mathbf{y}^{\mathbf{c}}t^{20}, \\ \overline{h}_2^\alpha(\mathbf{y},t)=c_{\mathbf{d}}\mathbf{y}^{\mathbf{d}}+c_{\mathbf{e}}\mathbf{y}^{\mathbf{e}}t^{0.2}+c_{\mathbf{f}}\mathbf{y}^{\mathbf{f}}. \end{array}\right.$$

Since $\widehat{\beta}=(-10,0.1,1)$ supports $\widehat{C}'$, in the homotopy $H^\beta=(h_1^\beta,h_2^\beta)$ induced by β, the distribution of powers of t are

$$\left\{\begin{array}{l} \mathbf{a}:\langle(-10,0.1,1),\ (0,1,1.1)\rangle=1.2 \\ \mathbf{b}:\langle(-10,0.1,1),\ (0,0,1)\rangle=1 \quad \text{in } h_1^\beta, \\ \mathbf{c}:\langle(-10,0.1,1),\ (1,0,11)\rangle=1 \end{array}\right. \quad \left\{\begin{array}{l} 0.2 \\ 0 \quad \text{in } \overline{h}_1^\beta \\ 0 \end{array}\right.$$

$$\left\{\begin{array}{l} \mathbf{d}:\langle(-10,0.1,1),\ (0,1,12)\rangle=12.1 \\ \mathbf{e}:\langle(-10,0.1,1),\ (1,0,2.1)\rangle=-7.9 \quad \text{in } h_2^\beta, \\ \mathbf{f}:\langle(-10,0.1,1),\ (1,1,2)\rangle=-7.9 \end{array}\right. \quad \left\{\begin{array}{l} 20 \\ 0 \quad \text{in } \overline{h}_2^\beta \\ 0 \end{array}\right.$$

That is,

$$\left\{\begin{array}{l} \overline{h}_1^\beta(\mathbf{y},t)=c_{\mathbf{a}}\mathbf{y}^{\mathbf{a}}t^{0.2}+c_{\mathbf{b}}\mathbf{y}^{\mathbf{b}}+c_{\mathbf{c}}\mathbf{y}^{\mathbf{c}}, \\ \overline{h}_2^\beta(\mathbf{y},t)=c_{\mathbf{d}}\mathbf{y}^{\mathbf{d}}t^{20}+c_{\mathbf{e}}\mathbf{y}^{\mathbf{e}}+c_{\mathbf{f}}\mathbf{y}^{\mathbf{f}}. \end{array}\right.$$

We wish to have certain lifting $\begin{array}{rcl} \omega_i: \mathcal{A}_i & \longrightarrow & \mathbb{R} \\ \mathbf{q} & \mapsto & \omega_i(\mathbf{q})=\omega_{\mathbf{q}} \end{array}$ that induces polyhedral homotopies in which every power of t lies between 1 and some μ :

$$1\leq \quad \langle(\alpha,1),(\mathbf{c},\omega_{\mathbf{c}})\rangle-\langle(\alpha,1),(\mathbf{a},\omega_{\mathbf{a}})\rangle \quad \leq\mu,$$

$$1\leq \quad \langle(\alpha,1),(\mathbf{e},\omega_{\mathbf{e}})\rangle-\langle(\alpha,1),(\mathbf{d},\omega_{\mathbf{d}})\rangle \quad \leq\mu,$$

where α is the solution to

$$\begin{cases} \langle (\alpha, 1), (\mathbf{a}, \omega_{\mathbf{a}}) \rangle = \langle (\alpha, 1), (\mathbf{b}, \omega_{\mathbf{b}}) \rangle, \\ \langle (\alpha, 1), (\mathbf{d}, \omega_{\mathbf{d}}) \rangle = \langle (\alpha, 1), (\mathbf{f}, \omega_{\mathbf{f}}) \rangle, \end{cases} \tag{2.2}$$

i.e. α supports cell $C = \{\{\mathbf{a}, \mathbf{b}\}, \{\mathbf{d}, \mathbf{f}\}\}$, and

$$\begin{aligned} 1 \leq \quad & \langle (\beta, 1), (\mathbf{a}, \omega_{\mathbf{a}}) \rangle - \langle (\beta, 1), (\mathbf{b}, \omega_{\mathbf{b}}) \rangle \quad \leq \mu, \\ 1 \leq \quad & \langle (\beta, 1), (\mathbf{d}, \omega_{\mathbf{d}}) \rangle - \langle (\beta, 1), (\mathbf{e}, \omega_{\mathbf{e}}) \rangle \quad \leq \mu, \end{aligned}$$

where β is the solution to

$$\begin{cases} \langle (\beta, 1), (\mathbf{b}, \omega_{\mathbf{b}}) \rangle = \langle (\beta, 1), (\mathbf{c}, \omega_{\mathbf{c}}) \rangle, \\ \langle (\beta, 1), (\mathbf{e}, \omega_{\mathbf{e}}) \rangle = \langle (\beta, 1), (\mathbf{f}, \omega_{\mathbf{f}}) \rangle, \end{cases} \tag{2.3}$$

that is, β supports $C' = \{\{\mathbf{b}, \mathbf{c}\}, \{\mathbf{e}, \mathbf{f}\}\}$. Hence, C and C' remain mixed cells.

Consider a linear programming problem with the variables μ and $\omega_{\mathbf{q}}$'s:

$$\begin{array}{lll} \text{LP0}: & \min \mu & \\ & & \qquad\qquad \text{(exponents of } t\text{)} \\ & \text{s.t.} & 1 \leq \ \langle (\alpha, 1), (\mathbf{c}, \omega_{\mathbf{c}}) \rangle - \langle (\alpha, 1), (\mathbf{a}, \omega_{\mathbf{a}}) \rangle \ \leq \mu, \\ & & 1 \leq \ \langle (\alpha, 1), (\mathbf{e}, \omega_{\mathbf{e}}) \rangle - \langle (\alpha, 1), (\mathbf{d}, \omega_{\mathbf{d}}) \rangle \ \leq \mu, \\ & & 1 \leq \ \langle (\beta, 1), (\mathbf{a}, \omega_{\mathbf{a}}) \rangle - \langle (\beta, 1), (\mathbf{b}, \omega_{\mathbf{b}}) \rangle \ \leq \mu, \\ & & 1 \leq \ \langle (\beta, 1), (\mathbf{d}, \omega_{\mathbf{d}}) \rangle - \langle (\beta, 1), (\mathbf{e}, \omega_{\mathbf{e}}) \rangle \ \leq \mu, \\ \text{where} & \multicolumn{2}{l}{\alpha \text{ is the solution of } \begin{cases} \langle \alpha, \mathbf{a} - \mathbf{b} \rangle = & \omega_{\mathbf{b}} - \omega_{\mathbf{a}} \\ \langle \alpha, \mathbf{d} - \mathbf{f} \rangle = & \omega_{\mathbf{f}} - \omega_{\mathbf{d}} \end{cases} \text{by (2.2)},} \\ & \multicolumn{2}{l}{\beta \text{ is the solution of } \begin{cases} \langle \beta, \mathbf{b} - \mathbf{c} \rangle = & \omega_{\mathbf{c}} - \omega_{\mathbf{b}} \\ \langle \beta, \mathbf{e} - \mathbf{f} \rangle = & \omega_{\mathbf{f}} - \omega_{\mathbf{e}} \end{cases} \text{by (2.3)},} \\ \text{variables:} & \multicolumn{2}{l}{\mu, \omega_{\mathbf{a}}, \omega_{\mathbf{b}}, \omega_{\mathbf{c}}, \omega_{\mathbf{d}}, \omega_{\mathbf{e}}, \omega_{\mathbf{f}}.} \end{array}$$

The powers of t in $\overline{H}^{\alpha}$ can be regarded as another lifting which induces the same mixed cells C and C' with $\widehat{C}$ staying "on the ground". By restricting $\omega_{\mathbf{a}}, \omega_{\mathbf{b}}, \omega_{\mathbf{d}}, \omega_{\mathbf{f}} = 0$, we have $\alpha = (0, 0)^T$ in LP0 and obtain a new linear programming model LP1:

$$
\begin{array}{llccc}
\text{LP1}: & \min\ \mu & & & \\
& & \text{(exponents of } t) & & \\
& \text{s.t.}\ \ 1 \leq & \omega_{\mathbf{c}} & \leq \mu, & \\
& \phantom{\text{s.t.}}\ \ 1 \leq & \omega_{\mathbf{e}} & \leq \mu, & \\
& \phantom{\text{s.t.}}\ \ 1 \leq & \langle \beta, \mathbf{a}\rangle - \langle \beta, \mathbf{b}\rangle & \leq \mu, & \\
& \phantom{\text{s.t.}}\ \ 1 \leq & \langle \beta, \mathbf{d}\rangle - \langle (\beta, 1), (\mathbf{e}, \omega_{\mathbf{e}})\rangle & \leq \mu, & \\
\text{where} & \beta \text{ is the solution of } & \left\{ \begin{array}{rl} \langle \beta, \mathbf{b}-\mathbf{c}\rangle = & \omega_{\mathbf{c}} \\ \langle \beta, \mathbf{e}-\mathbf{f}\rangle = & -\omega_{\mathbf{e}} \end{array} \right. & \text{by (2.3)}, & \\
\text{variables:} & \mu, \omega_{\mathbf{c}}, \omega_{\mathbf{e}}. & & &
\end{array}
$$

LP1 is of course preferable if its optimal solution μ and corresponding exponents of t are better than the ones in LP0. $\square$

2.2 Setup of The Sandwich Model

As indicated in the last section, generic random liftings can induce highly nonlinear polyhedral homotopies which may produce numerical instabilities. To overcome this problem, our strategy is to find a new lifting function $\nu = (\nu_1, \ldots, \nu_n)$ where $\nu_i : \mathcal{A}_i \to \mathbb{R}$ for $i = 1, \ldots, n$, *based on the already computed mixed-cell configuration* $\mathcal{M}_\omega$. The mixed-cell configuration $\mathcal{M}_\nu$ induced by the new lifting ν will be the same as $\mathcal{M}_\omega$, but the highest power of the continuation parameter t in the polyhedral homotopies induced by ν is the smallest among all those polyhedral homotopies induced by liftings which keep the mixed-cell configuration $\mathcal{M}_\omega$ invariant. In this way, reidentifying the mixed-cell configuration $\mathcal{M}_\nu$, which is very time consuming, becomes unnecessary.

Let $C = (C_1, \ldots, C_n) \in \mathcal{M}_\omega$ where $C_i = \{\mathbf{a}_i, \mathbf{b}_i\} \subset \mathcal{A}_i$, $i = 1, \ldots, n$. To keep $\mathcal{M}_\omega$ invariant, we impose on any new lifting function $\nu = (\nu_1, \ldots, \nu_n)$ the conditions:

$$
\left\{ \begin{array}{ll} \langle (\gamma, 1), (\mathbf{a}_i, \nu_i(\mathbf{a}_i))\rangle = \langle (\gamma, 1), (\mathbf{b}_i, \nu_i(\mathbf{b}_i))\rangle, & \\ \langle (\gamma, 1), (\mathbf{a}_i, \nu_i(\mathbf{a}_i))\rangle < \langle (\gamma, 1), (\mathbf{q}, \nu_i(\mathbf{q}))\rangle, \quad \forall\ \mathbf{q} \in \mathcal{A}_i \backslash \{\mathbf{a}_i, \mathbf{b}_i\}, & \end{array} \right. \quad i = 1, \ldots, n,
$$

or,

$$\begin{aligned} \langle \mathbf{a}_i, \gamma\rangle + \nu_i(\mathbf{a}_i) &= \langle \mathbf{b}_i, \gamma\rangle + \nu_i(\mathbf{b}_i), && (2.4)\\ \langle \mathbf{a}_i, \gamma\rangle + \nu_i(\mathbf{a}_i) &< \langle \mathbf{q}, \gamma\rangle + \nu_i(\mathbf{q}), \quad \forall\, \mathbf{q} \in \mathcal{A}_i\backslash\{\mathbf{a}_i, \mathbf{b}_i\}, && (2.5) \end{aligned}$$

for $i = 1, \ldots, n$, where $(\gamma, 1)$ is the inner normal of C in $\mathcal{M}_\nu$.

From (2.4), we have $\langle \mathbf{a}_i - \mathbf{b}_i, \gamma\rangle = \nu_i(\mathbf{b}_i) - \nu_i(\mathbf{a}_i)$ for $i = 1, \ldots, n$. Since $C = (C_1, \ldots, C_n)$ is a mixed cell in the fine mixed subdivision S_ω, the edges spanned by $\{\mathbf{a}_i, \mathbf{b}_i\}$, $i = 1, \ldots, n$, determine a full-dimensional parallelepiped in $\mathbb{R}^n$. Equivalently, the matrix

$$A = \begin{bmatrix} \mathbf{a}_1 - \mathbf{b}_1 \\ \vdots \\ \mathbf{a}_n - \mathbf{b}_n \end{bmatrix}$$

is nonsingular, and so, γ can be expressed uniquely as a linear combination of the lifting values $\nu_i(\mathbf{a}_i)$ and $\nu_i(\mathbf{b}_i)$, $i = 1, \ldots, n$. Namely,

$$\gamma^T = \begin{bmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{bmatrix} = A^{-1} \begin{bmatrix} \nu_1(\mathbf{b}_1) - \nu_1(\mathbf{a}_1) \\ \vdots \\ \nu_n(\mathbf{b}_n) - \nu_n(\mathbf{a}_n) \end{bmatrix}. \tag{2.6}$$

As in (1.9), the polyhedral homotopy induced by lifting $\nu = (\nu_1, \ldots, \nu_n)$ and mixed cell $C = (\{\mathbf{a}_1, \mathbf{b}_1\}, \ldots, \{\mathbf{a}_n, \mathbf{b}_n\})$ with inner normal $(\gamma, 1)$ is

$$\bar{h}_i^\gamma(\mathbf{y}, t) = c_{i,\mathbf{a}_i}\mathbf{y}^{\mathbf{a}_i} + c_{i,\mathbf{b}_i}\mathbf{y}^{\mathbf{b}_i} + \sum_{\mathbf{q}\in\mathcal{A}_i\backslash\{\mathbf{a}_i,\mathbf{b}_i\}} c_{i,\mathbf{q}}\mathbf{y}^{\mathbf{q}}\, t^{e_{\mathbf{q}}^\gamma}, \quad i = 1, \ldots, n,$$

where $e_{\mathbf{q}}^\gamma$, the powers of t, are given by

$$e_{\mathbf{q}}^\gamma := \langle \mathbf{q}, \gamma\rangle - \langle \mathbf{a}_i, \gamma\rangle + \nu_i(\mathbf{q}) - \nu_i(\mathbf{a}_i), \quad \forall\, \mathbf{q} \in \mathcal{A}_i\backslash\{\mathbf{a}_i, \mathbf{b}_i\},$$

and they are always positive according to (2.5). By (2.6), γ may be removed from the symbol $e^{\gamma}_{\mathbf{q}}$. We denote the resulting expressions by $e_{\mathbf{q}}$. More explicitly,

$$e_{\mathbf{q}} := (\mathbf{q} - \mathbf{a}_i)A^{-1} \begin{bmatrix} \nu_1(\mathbf{b}_1) - \nu_1(\mathbf{a}_1) \\ \vdots \\ \nu_n(\mathbf{b}_n) - \nu_n(\mathbf{a}_n) \end{bmatrix} + \nu_i(\mathbf{q}) - \nu_i(\mathbf{a}_i). \tag{2.7}$$

To avoid the powers $e_{\mathbf{q}}$ being too large or too small, it is natural to consider the following minimization problem **LP0**:

$$\boxed{\begin{array}{lll} \mathbf{LP0}: & \min & \mu \\ & \text{s.t.} & 1 \le e_{\mathbf{q}} \le \mu \quad \forall\, \mathbf{q} \in \mathcal{A}_i \backslash \{\mathbf{a}_i, \mathbf{b}_i\}, \quad i = 1, \ldots, n, \\ & & \forall\, C = (\{\mathbf{a}_1, \mathbf{b}_1\}, \ldots, \{\mathbf{a}_n, \mathbf{b}_n\}) \in \mathcal{M}_\omega \\ & & \text{with } \ e_{\mathbf{q}} \ \text{ defined in (2.7).} \end{array}} \tag{2.8}$$

Apparently, the lifting function $\nu = (\nu_1, \ldots, \nu_n)$ with values obtained by the solution of this minimization problem satisfies (2.4) and (2.5) with γ defined in (2.6). Therefore, the mixed-cell configuration $\mathcal{M}_\nu$ induced by ν coincides with $\mathcal{M}_\omega$. Moreover, the powers of the continuation parameter t in the polyhedral homotopies induced by ν will be much better balanced.

2.3 Reducing the Size of the Sandwich Model

LP0 has $1 + \sum_{i=1}^{n} \#\mathcal{A}_i$ unknowns, namely, μ as well as $\nu_i(\mathbf{q})$ for $\mathbf{q} \in \mathcal{A}_i$, $i = 1, \ldots, n$, and $2(\#\mathcal{M}_\omega) \sum_{i=1}^{n} (\#\mathcal{A}_i - 2)$ inequalities. For practical considerations, we wish to reduce both the number of unknowns and the number of inequalities in **LP0**. In the following, we will show that for a fixed mixed cell $\overline{C} = (\{\overline{\mathbf{a}}_1, \overline{\mathbf{b}}_1\}, \ldots, \{\overline{\mathbf{a}}_n, \overline{\mathbf{b}}_n\})$, where $\{\overline{\mathbf{a}}_i, \overline{\mathbf{b}}_i\} \subset \mathcal{A}_i$, $i = 1, \ldots, n$, in the mixed-cell configuration $\mathcal{M}_\omega$ with inner normal $(\beta, 1)$, we may set $\nu_i(\overline{\mathbf{a}}_i)$ and $\nu_i(\overline{\mathbf{b}}_i)$ to be zero for $i = 1, \ldots, n$ in **LP0**, so the number of unknowns is reduced by $2n$. But the solution of the new minimization problem defines a lifting function $\nu' = (\nu'_1, \ldots, \nu'_n)$ which induces the same mixed-cell configuration as $\mathcal{M}_\omega$.

For the fixed mixed cell $\overline{C} = (\{\bar{\mathbf{a}}_1, \bar{\mathbf{b}}_1\}, \ldots, \{\bar{\mathbf{a}}_n, \bar{\mathbf{b}}_n\}) \in \mathcal{M}_\omega$ with inner normal $(\beta, 1)$, we define a lifting function $\omega' = (\omega'_1, \ldots, \omega'_n)$, where $\omega'_i : \mathcal{A}_i \to \mathbb{R}$, $i = 1, \ldots, n$, as follows: for $i = 1, \ldots, n$ and $\mathbf{q} \in \mathcal{A}_i$,

$$\omega'_i(\mathbf{q}) := \langle \beta, \mathbf{q} \rangle - \langle \beta, \bar{\mathbf{a}}_i \rangle + \omega_i(\mathbf{q}) - \omega_i(\bar{\mathbf{a}}_i). \tag{2.9}$$

Then ω'_i vanishes at both $\bar{\mathbf{a}}_i$ and $\bar{\mathbf{b}}_i$, $i = 1, \ldots, n$. Let $\mathcal{M}_{\omega'}$ be the mixed-cell configuration in the subdivision $S_{\omega'}$ of $\mathcal{A} = (\mathcal{A}_1, \ldots, \mathcal{A}_n)$ induced by ω'.

Lemma 2 *$\mathcal{M}_{\omega'} = \mathcal{M}_\omega$. More precisely, $C = (C_1, \ldots, C_n) \in \mathcal{M}_\omega$ with inner normal $(\alpha, 1)$ with respect to ω if and only if $C = (C_1, \ldots, C_n) \in \mathcal{M}_{\omega'}$ with inner normal $(\alpha - \beta, 1)$ with respect to ω'.*

PROOF: Let $C_i = \{\mathbf{a}_i, \mathbf{b}_i\} \subset \mathcal{A}_i$ for $i = 1, \ldots, n$. Then, $C \in \mathcal{M}_\omega$ with inner normal $(\alpha, 1) \iff$

$$\begin{cases} \langle (\alpha, 1), (\mathbf{a}_i, \omega_i(\mathbf{a}_i)) \rangle = \langle (\alpha, 1), (\mathbf{b}_i, \omega_i(\mathbf{b}_i)) \rangle, \\ \langle (\alpha, 1), (\mathbf{a}_i, \omega_i(\mathbf{a}_i)) \rangle < \langle (\alpha, 1), (\mathbf{q}, \omega_i(\mathbf{q})) \rangle, \quad \forall\, \mathbf{q} \in \mathcal{A}_i \backslash \{\mathbf{a}_i, \mathbf{b}_i\}, \end{cases} \quad i = 1, \ldots, n.$$

Or,

$$\begin{cases} \langle \alpha, \mathbf{a}_i - \mathbf{b}_i \rangle = \omega_i(\mathbf{b}_i) - \omega_i(\mathbf{a}_i), \\ \langle \alpha, \mathbf{q} - \mathbf{a}_i \rangle > \omega_i(\mathbf{a}_i) - \omega_i(\mathbf{q}), \quad \forall\, \mathbf{q} \in \mathcal{A}_i \backslash \{\mathbf{a}_i, \mathbf{b}_i\}, \end{cases} \quad i = 1, \ldots, n.$$

On the other hand, $C \in \mathcal{M}_{\omega'}$ with inner normal $(\alpha - \beta, 1) \iff$

$$\begin{cases} \langle (\alpha - \beta, 1), (\mathbf{a}_i, \omega'_i(\mathbf{a}_i)) \rangle = \langle (\alpha - \beta, 1), (\mathbf{b}_i, \omega'_i(\mathbf{b}_i)) \rangle, \\ \langle (\alpha - \beta, 1), (\mathbf{a}_i, \omega'_i(\mathbf{a}_i)) \rangle < \langle (\alpha - \beta, 1), (\mathbf{q}, \omega'_i(\mathbf{q})) \rangle, \quad \forall\, \mathbf{q} \in \mathcal{A}_i \backslash \{\mathbf{a}_i, \mathbf{b}_i\}, \end{cases} \quad i = 1, \ldots, n.$$

Or, by (2.9),

$$\begin{aligned} \langle \alpha - \beta, \mathbf{a}_i - \mathbf{b}_i \rangle &= \omega'_i(\mathbf{b}_i) - \omega'_i(\mathbf{a}_i) \\ &= \langle \beta, \mathbf{b}_i \rangle - \langle \beta, \bar{\mathbf{a}}_i \rangle + \omega_i(\mathbf{b}_i) - \omega_i(\bar{\mathbf{a}}_i) \\ &\quad -(\langle \beta, \mathbf{a}_i \rangle - \langle \beta, \bar{\mathbf{a}}_i \rangle + \omega_i(\mathbf{a}_i) - \omega_i(\bar{\mathbf{a}}_i)) \\ &= \langle -\beta, \mathbf{a}_i - \mathbf{b}_i \rangle + \omega_i(\mathbf{b}_i) - \omega_i(\mathbf{a}_i), \end{aligned}$$

i.e.,

$$\langle \alpha, \mathbf{a}_i - \mathbf{b}_i \rangle = \omega_i(\mathbf{b}_i) - \omega_i(\mathbf{a}_i), \quad i = 1, \dots, n$$

and, for $\mathbf{q} \in \mathcal{A}_i \backslash \{\mathbf{a}_i, \mathbf{b}_i\}$,

$$\begin{aligned}
\langle \alpha - \beta, \mathbf{q} - \mathbf{a}_i \rangle &> \omega_i'(\mathbf{a}_i) - \omega_i'(\mathbf{q}) \\
&= \langle \beta, \mathbf{a}_i \rangle - \langle \beta, \bar{\mathbf{a}}_i \rangle + \omega_i(\mathbf{a}_i) - \omega_i(\bar{\mathbf{a}}_i) \\
&\quad - \Big(\langle \beta, \mathbf{q} \rangle - \langle \beta, \bar{\mathbf{a}}_i \rangle + \omega_i(\mathbf{q}) - \omega_i(\bar{\mathbf{a}}_i) \Big) \\
&= \langle -\beta, \mathbf{q} - \mathbf{a}_i \rangle + \omega_i(\mathbf{a}_i) - \omega_i(\mathbf{q}),
\end{aligned}$$

i.e.,

$$\langle \alpha, \mathbf{q} - \mathbf{a}_i \rangle > \omega_i(\mathbf{a}_i) - \omega_i(\mathbf{q}), \quad i = 1, \dots, n.$$

So, the proof of the assertion is achieved. □

Most importantly, a straightforward calculation shows that the polyhedral homotopy, as in (1.9), induced by the cell $C = (C_1, \dots, C_n)$ in $\mathcal{M}_\omega$ with inner normal $(\alpha, 1)$ is exactly the same as the one induced by cell $C = (C_1, \dots, C_n)$ in $\mathcal{M}_{\omega'}$ with inner normal $(\alpha - \beta, 1)$. So, we may solve $P(\mathbf{x}) = \mathbf{0}$ by using the polyhedral homotopies produced by mixed cells in $\mathcal{M}_{\omega'}$ together with corresponding inner normals.

Now, with the lifting ω', we consider the minimization problem **LP1** :

$$\boxed{\begin{aligned}
\mathbf{LP1}: \ \min \ & \mu \\
\text{s.t.} \ & 1 \le \langle \gamma, \mathbf{q} - \mathbf{a}_i \rangle + \nu_i(\mathbf{q}) - \nu_i(\mathbf{a}_i) \le \mu \quad \forall\, \mathbf{q} \in \mathcal{A}_i \backslash \{\mathbf{a}_i, \mathbf{b}_i\}, \\
& \forall\, C = (\{\mathbf{a}_1, \mathbf{b}_1\}, \dots, \{\mathbf{a}_n, \mathbf{b}_n\}) \in \mathcal{M}_{\omega'}, \\
& \nu_i(\bar{\mathbf{a}}_i) = \nu_i(\bar{\mathbf{b}}_i) = 0, \quad i = 1, \dots, n.
\end{aligned}} \tag{2.10}$$

Here, γ can be expressed, as in (2.6), as a linear combination of the values of ν_i's:

$$\gamma^T = \begin{bmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 - \mathbf{b}_1 \\ \vdots \\ \mathbf{a}_n - \mathbf{b}_n \end{bmatrix}^{-1} \begin{bmatrix} \nu_1(\mathbf{b}_1) - \nu_1(\mathbf{a}_1) \\ \vdots \\ \nu_n(\mathbf{b}_n) - \nu_n(\mathbf{a}_n) \end{bmatrix}.$$

This problem has $2n$ fewer unknowns than **LP0**, and its solution yields a lifting function $\nu' = (\nu'_1, \ldots, \nu'_n)$. As before, the mixed-cell configuration $\mathcal{M}_{\nu'}$ induced by the lifting ν' is the same as $\mathcal{M}_{\omega'}$.

The following proposition shows the feasibility of this problem.

Proposition 3 **LP1** *has an optimal solution.*

PROOF: Apparently, the values of the lifting function $\omega' = (\omega'_1, \ldots, \omega'_n)$ satisfy

$$0 < \langle \alpha, \mathbf{q} - \mathbf{a}_i \rangle + \omega'_i(\mathbf{q}) - \omega'_i(\mathbf{a}_i), \quad i = 1, \ldots, n, \tag{2.11}$$

for all $C = (\{\mathbf{a}_1, \mathbf{b}_1\}, \ldots, \{\mathbf{a}_n, \mathbf{b}_n\}) \in \mathcal{M}_{\omega'}$ with inner normal $(\alpha, 1)$ and $\mathbf{q} \in \mathcal{A}_i \backslash \{\mathbf{a}_i, \mathbf{b}_i\}$, where

$$\alpha^T = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 - \mathbf{b}_1 \\ \vdots \\ \mathbf{a}_n - \mathbf{b}_n \end{bmatrix}^{-1} \begin{bmatrix} \omega'_1(\mathbf{b}_1) - \omega'_1(\mathbf{a}_1) \\ \vdots \\ \omega'_n(\mathbf{b}_n) - \omega'_n(\mathbf{a}_n) \end{bmatrix}.$$

It follows that the function values of $\nu^{(\ell)} := \ell\omega'$ for $\ell > 0$ also satisfy (2.11). There exists $\ell_0 > 0$ such that

$$1 \leq \langle \ell_0\alpha, \mathbf{q} - \mathbf{a}_i \rangle + \nu_i^{(\ell_0)}(\mathbf{q}) - \nu_i^{(\ell_0)}(\mathbf{a}_i), \quad i = 1, \ldots, n, \tag{2.12}$$

for all $C = (\{\mathbf{a}_1, \mathbf{b}_1\}, \ldots, \{\mathbf{a}_n, \mathbf{b}_n\}) \in \mathcal{M}_{\omega'}$ and $\mathbf{q} \in \mathcal{A}_i \backslash \{\mathbf{a}_i, \mathbf{b}_i\}$. Let μ_0 be the maximum of the right-hand side of (2.12). Then $(\nu^{(\ell_0)}, \mu_0)$ is a feasible solution of **LP1**. So **LP1** has an optimal solution with the value of μ between 1 and μ_0. $\square$

Remark 2 The pair $(\nu^{(\ell_0)}, \mu_0)$ constructed in the proof can serve as a starting point of standard simplex algorithms for solving **LP1**.

The number of variables in each double inequality of **LP1** is no greater than $2n+2$ which is usually much smaller than the number of variables in **LP1**. This sparsity of the linear programming problem **LP1** is exploited in our algorithm and results in a remarkable speed-up. Some of the inequalities in the constraints of **LP1** are exactly the same, and they can easily be detected by comparisons and deleted when the constraints are being generated.

In the rest of this section, we will show that the constraints in **LP1** can also be derived from *circuits* [GKZ, MiVe] of the support $\mathcal{A} = (\mathcal{A}_1, \ldots, \mathcal{A}_n)$. For a mixed cell $C = (\{\mathbf{a}_1, \mathbf{b}_1\}, \ldots, \{\mathbf{a}_n, \mathbf{b}_n\}) \in \mathcal{M}_\omega$ with inner normal $(\alpha, 1)$ and $\mathbf{q} \in \mathcal{A}_j \backslash \{\mathbf{a}_j, \mathbf{b}_j\}$, we have

$$\langle \alpha, \mathbf{a}_j \rangle + \omega_j(\mathbf{a}_j) < \langle \alpha, \mathbf{q} \rangle + \omega_j(\mathbf{q}),$$

or,

$$\langle \alpha, \mathbf{q} - \mathbf{a}_j \rangle + \omega_j(\mathbf{q}) - \omega_j(\mathbf{a}_j) > 0. \tag{2.13}$$

We will call this inequality the **normal inequality** with respect to C and $\mathbf{q}$. This inequality serves as the building block in our sandwich model **LP1** in (2.10).

Note that $(\alpha, 1)$ is also the normal of the n-dimensional subspace H of $\mathbb{R}^{n+1}$ spanned by the vectors

$$\Big(\mathbf{b}_1 - \mathbf{a}_1, \omega_1(\mathbf{b}_1) - \omega_1(\mathbf{a}_1)\Big), \cdots, \Big(\mathbf{b}_n - \mathbf{a}_n, \omega_n(\mathbf{b}_n) - \omega_n(\mathbf{a}_n)\Big).$$

Let $(\mathbf{q} - \mathbf{a}_j, s)$ be the projection of the vector $(\mathbf{q} - \mathbf{a}_j, \omega_j(\mathbf{q}) - \omega_j(\mathbf{a}_j))$ on the subspace H along the direction of its last coordinate. Then $\langle (\alpha, 1), (\mathbf{q} - \mathbf{a}_j, s) \rangle = 0$, or,

$$\langle \alpha, \mathbf{q} - \mathbf{a}_j \rangle + s = 0$$

and (2.13) becomes

$$\omega_j(\mathbf{q}) - \omega_j(\mathbf{a}) - s > 0.$$

So, the left-hand side of (2.13) measures the distance between the points $(\mathbf{q} - \mathbf{a}_j, \omega_j(\mathbf{q}) - \omega_j(\mathbf{a}_j))$ and $(\mathbf{q} - \mathbf{a}_j, s)$ as shown in Figure 2.1.

On the other hand, the Cayley embedding of $\mathcal{A}_i$ into $\mathbb{R}^{2n-1}$ is

$$\widetilde{\mathcal{A}}_i = \{(\mathbf{a}, \mathbf{e}_{i-1}) \mid \mathbf{a} \in \mathcal{A}_i\}, \quad i = 1, \ldots, n,$$

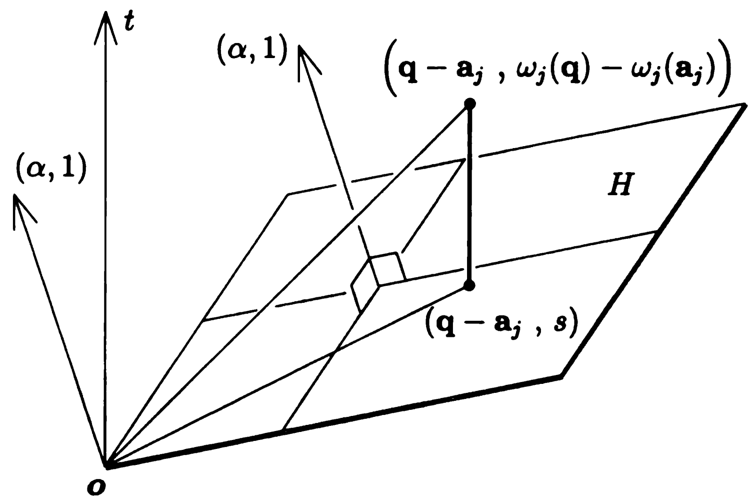

Figure 2.1: An illustration of normal inequality (2.13)

here $\mathbf{e}_0 = \mathbf{0} \in \mathbb{R}^{n-1}$ and $\mathbf{e}_i = (0, \ldots, 1, 0, \ldots, 0)$ is the i^{th} unit vector in $\mathbb{R}^{n-1}$, $i = 1, \ldots, n-1$. Let $\widetilde{\mathcal{A}} = \bigcup_{i=1}^{n} \widetilde{\mathcal{A}}_i$. Define $\widetilde{\omega} : \widetilde{\mathcal{A}} \to \mathbb{R}$ by

$$\widetilde{\omega}(\mathbf{a}, \mathbf{e}_{i-1}) = \omega_i(\mathbf{a}), \quad \forall\, \mathbf{a} \in \mathcal{A}_i, \quad i = 1, \ldots, n.$$

Then the lifting $\widetilde{\omega}$ induces a regular subdivision of $\widetilde{\mathcal{A}}$, denoted by $\widetilde{S}_{\widetilde{\omega}}$.

Proposition 4 (The Cayley Trick [GKZ, St94, VGC]) *S_ω is a fine mixed subdivision of $\mathcal{A} = (\mathcal{A}_1, \ldots, \mathcal{A}_n)$ if and only if $\widetilde{S}_{\widetilde{\omega}}$ is a triangulation of* $\operatorname{conv}(\widetilde{\mathcal{A}})$. *Furthermore, $C = (C_1, \ldots, C_n) \in S_\omega$ if and only if $\bigcup_{i=1}^{n} \widetilde{C}_i \in \widetilde{S}_{\widetilde{\omega}}$.*

By the Cayley embedding, $\mathbf{a}_1, \mathbf{b}_1, \ldots, \mathbf{a}_n, \mathbf{b}_n$ and $\mathbf{q}$ are embedded in $\mathbb{R}^{2n-1}$ as follows:

$$(\mathbf{a}_1, \mathbf{e}_0),\ (\mathbf{b}_1, \mathbf{e}_0),\ \ldots,\ (\mathbf{a}_n, \mathbf{e}_{n-1}),\ (\mathbf{b}_n, \mathbf{e}_{n-1}),\ (\mathbf{q}, \mathbf{e}_{j-1})$$

Since these points are affinely dependent (not necessarily a circuit), there exist $\lambda_1, \lambda_1', \ldots, \lambda_n, \lambda_n' \in \mathbb{R}$ such that

$$\begin{cases} \displaystyle\sum_{i=1}^{n} (\lambda_i + \lambda_i') = 1, \\ \displaystyle(\mathbf{q}, \mathbf{e}_{j-1}) = \sum_{i=1}^{n} \Big(\lambda_i(\mathbf{a}_i, \mathbf{e}_{i-1}) + \lambda_i'(\mathbf{b}_i, \mathbf{e}_{i-1})\Big). \end{cases} \tag{2.14}$$

From (2.14), as in Theorem 4.1 of [MiVe], we have

$$\begin{cases} \lambda_j + \lambda'_j = 1, \\ \lambda_i + \lambda'_i = 0 \quad \text{for } i \neq j. \end{cases} \tag{2.15}$$

It follows that

$$\begin{aligned} (\mathbf{q}, \mathbf{e}_{j-1}) &= \sum_{i=1}^{n} \Big(\lambda_i(\mathbf{a}_i, \mathbf{e}_{i-1}) + \lambda'_i(\mathbf{b}_i, \mathbf{e}_{i-1}) \Big) \\ &= \sum_{i=1}^{n} \Big(\lambda_i(\mathbf{a}_i, \mathbf{e}_{i-1}) + \lambda'_i(\mathbf{a}_i, \mathbf{e}_{i-1}) + \lambda'_i(\mathbf{b}_i, \mathbf{e}_{i-1}) - \lambda'_i(\mathbf{a}_i, \mathbf{e}_{i-1}) \Big) \\ &= (\mathbf{a}_j, \mathbf{e}_{j-1}) + \sum_{i=1}^{n} \lambda'_i(\mathbf{b}_i - \mathbf{a}_i, \mathbf{0}) \end{aligned}$$

and so,

$$\mathbf{q} - \mathbf{a}_j = \sum_{i=1}^{n} \lambda'_i(\mathbf{b}_i - \mathbf{a}_i). \tag{2.16}$$

Now, $\bigcup_{i=1}^{n} \widetilde{C}_i \in \widetilde{S}_{\widetilde{\omega}}$ implies

$$\widetilde{\omega}(\mathbf{q}, \mathbf{e}_{j-1}) > \sum_{i=1}^{n} \Big(\lambda_i \widetilde{\omega}(\mathbf{a}_i, \mathbf{e}_{i-1}) + \lambda'_i \widetilde{\omega}(\mathbf{b}_i, \mathbf{e}_{i-1}) \Big). \tag{2.17}$$

By (2.15) and the definition of $\widetilde{\omega}$, (2.17) can be written as

$$\omega_j(\mathbf{q}) - \omega_j(\mathbf{a}_j) > \sum_{i=1}^{n} \lambda'_i(\omega_i(\mathbf{b}_i) - \omega_i(\mathbf{a}_i)),$$

or,

$$\omega_j(\mathbf{q}) - \omega_j(\mathbf{a}_j) - \sum_{i=1}^{n} \lambda'_i(\omega_i(\mathbf{b}_i) - \omega_i(\mathbf{a}_i)) > 0. \tag{2.18}$$

We will call this inequality the **circuit inequality** with respect to C and $\mathbf{q}$.

Since $(\mathbf{q} - \mathbf{a}_j, s)$ is in the subspace H and thus can be written as a unique linear combination of $(\mathbf{b}_i - \mathbf{a}_i, \omega_i(\mathbf{b}_i) - \omega_i(\mathbf{a}_i))$, $i = 1, \dots, n$, (2.16) leads to

$$s = \sum_{i=1}^{n} \lambda'_i(\omega_i(\mathbf{b}_i) - \omega_i(\mathbf{a}_i)).$$

So, the left-hand side of (2.18) also represents the distance between the point $(\mathbf{q} - \mathbf{a}_j, \omega_j(\mathbf{q}) - \omega_j(\mathbf{a}_j))$ and $(\mathbf{q} - \mathbf{a}_j, s)$. Therefore, the normal inequality (2.13) is the same as the circuit inequality (2.18).

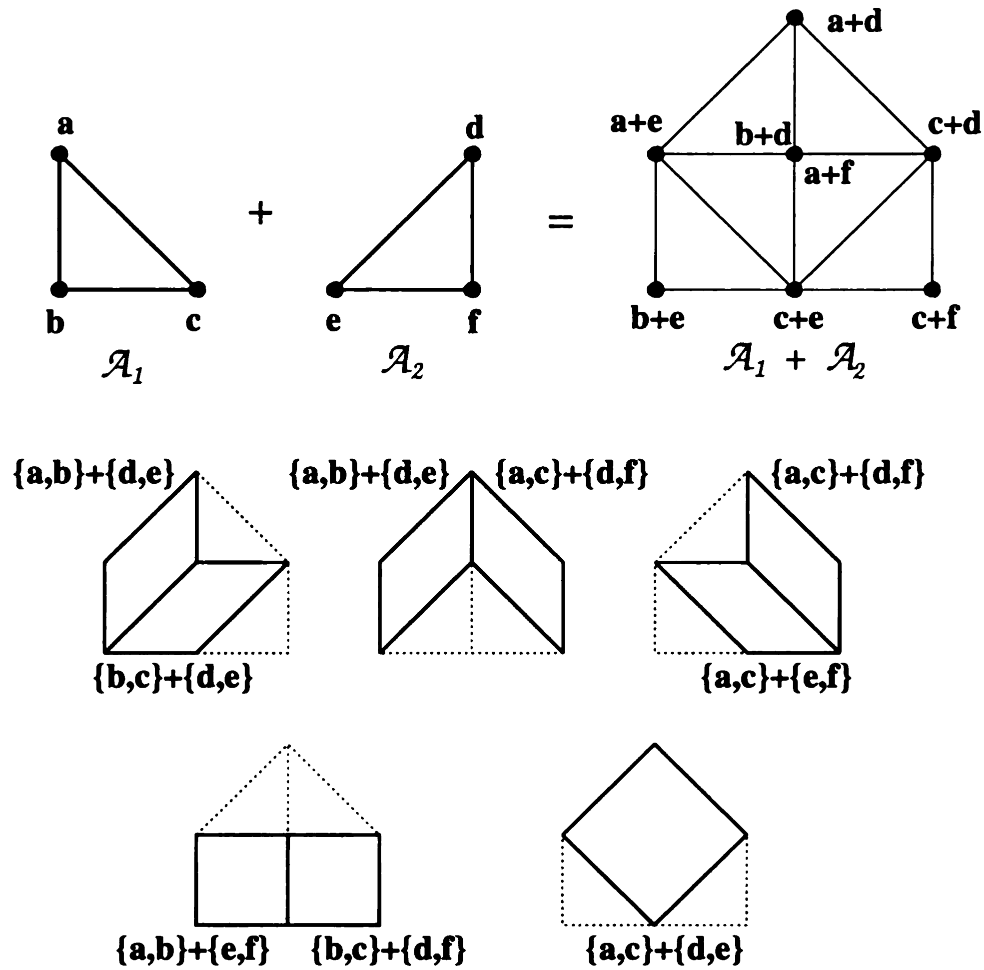

Figure 2.2: Mixed cell configurations of $(\mathcal{A}_1, \mathcal{A}_2)$

Example 3 Let $\mathcal{A}_1 = \{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$, $\mathcal{A}_2 = \{\mathbf{d}, \mathbf{e}, \mathbf{f}\}$ be two supports, where

$$\mathbf{a} = (0,1), \quad \mathbf{b} = (0,0), \quad \mathbf{c} = (1,0),$$

$$\mathbf{d} = (0,1), \quad \mathbf{e} = (1,0), \quad \mathbf{f} = (1,1),$$

Suppose $C = \Big(\{\mathbf{b}, \mathbf{c}\}, \{\mathbf{d}, \mathbf{e}\}\Big)$ and $C' = \Big(\{\mathbf{a}, \mathbf{b}\}, \{\mathbf{d}, \mathbf{e}\}\Big)$ are mixed cells induced by a lifting $\omega = (\omega_1, \omega_2)$. Let $\widetilde{\mathcal{A}} = \widetilde{\mathcal{A}}_1 \cup \widetilde{\mathcal{A}}_2$ be the Cayley embedding of $\mathcal{A}_1$ and $\mathcal{A}_2$, as shown in Figure 2.3. The four points in $\widetilde{C} = \{\widetilde{\mathbf{b}}, \widetilde{\mathbf{c}}, \widetilde{\mathbf{d}}, \widetilde{\mathbf{e}}\}$ form a simplex and it follows that $\{\widetilde{\mathbf{b}}, \widetilde{\mathbf{c}}, \widetilde{\mathbf{d}}, \widetilde{\mathbf{e}}, \widetilde{\mathbf{f}}\}$ is an affinely dependent set (but not a circuit), i.e.

$$\widetilde{\mathbf{f}} = \lambda_1 \widetilde{\mathbf{b}} + \lambda_1' \widetilde{\mathbf{c}} + \lambda_2 \widetilde{\mathbf{d}} + \lambda_2' \widetilde{\mathbf{e}}$$

for some $\lambda_1 + \lambda_1' + \lambda_2 + \lambda_2' = 1$. Since the third coordinate of both $\widetilde{\mathbf{b}}$ and $\widetilde{\mathbf{c}}$ equals

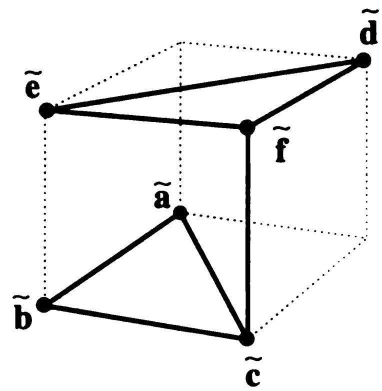

Figure 2.3: Cayley embedding of $\mathcal{A}_1$ and $\mathcal{A}_2$

zero, clearly, $\lambda_2 + \lambda_2' = 1$ and $\lambda_1 + \lambda_1' = 0$. So the equation becomes

$$\widetilde{\mathbf{f}} - \widetilde{\mathbf{d}} = \lambda_1'(\widetilde{\mathbf{c}} - \widetilde{\mathbf{b}}) + \lambda_2'(\widetilde{\mathbf{e}} - \widetilde{\mathbf{d}}),$$

i.e.

$$\mathbf{f} - \mathbf{d} = \lambda_1'(\mathbf{c} - \mathbf{b}) + \lambda_2'(\mathbf{e} - \mathbf{d}). \tag{2.19}$$

By Proposition 4, the projection of the lower facets of

$$\operatorname{conv}\Big((\widetilde{\mathbf{a}}, \omega_1(\mathbf{a})), (\widetilde{\mathbf{b}}, \omega_1(\mathbf{b})), (\widetilde{\mathbf{c}}, \omega_1(\mathbf{c})), (\widetilde{\mathbf{d}}, \omega_2(\mathbf{d})), (\widetilde{\mathbf{e}}, \omega_2(\mathbf{e})), (\widetilde{\mathbf{f}}, \omega_2(\mathbf{f}))\Big)$$

in $\mathbb{R}^4$ gives a regular triangulation of $\operatorname{conv}(\widetilde{\mathbf{a}}, \widetilde{\mathbf{b}}, \widetilde{\mathbf{c}}, \widetilde{\mathbf{d}}, \widetilde{\mathbf{e}}, \widetilde{\mathbf{f}})$, and

$$\Big\{(\widetilde{\mathbf{b}}, \omega_1(\mathbf{b})), (\widetilde{\mathbf{c}}, \omega_1(\mathbf{c})), (\widetilde{\mathbf{d}}, \omega_2(\mathbf{d})), (\widetilde{\mathbf{e}}, \omega_2(\mathbf{e}))\Big\}$$

is one of the lower facets. Let $(\beta, 1) \in \mathbb{R}^4$ be its inner normal, then

$$\langle \beta, \widetilde{\mathbf{b}} \rangle + \omega_1(\mathbf{b}) < \langle \beta, \widetilde{\mathbf{f}} \rangle + \omega_2(\mathbf{f})$$

with

$$\begin{aligned} \langle \beta, \widetilde{\mathbf{b}} \rangle + \omega_1(\mathbf{b}) &= \langle \beta, \widetilde{\mathbf{c}} \rangle + \omega_1(\mathbf{c}) \\ &= \langle \beta, \widetilde{\mathbf{d}} \rangle + \omega_2(\mathbf{d}) \\ &= \langle \beta, \widetilde{\mathbf{e}} \rangle + \omega_2(\mathbf{e}). \end{aligned}$$

Hence,

$$\begin{aligned}\omega_2(\mathbf{f}) &> \langle\beta,\widetilde{\mathbf{b}}\rangle - \langle\beta,\widetilde{\mathbf{f}}\rangle + \omega_1(\mathbf{b})\\ &= \langle\beta,\widetilde{\mathbf{b}}\rangle - \langle\beta,\widetilde{\mathbf{d}}\rangle - \lambda_1'\langle\beta,\widetilde{\mathbf{c}}-\widetilde{\mathbf{b}}\rangle - \lambda_2'\langle\beta,\widetilde{\mathbf{e}}-\widetilde{\mathbf{d}}\rangle + \omega_1(\mathbf{b})\\ &= \omega_2(\mathbf{d}) - \omega_1(\mathbf{b}) + \lambda_1'(\omega_1(\mathbf{c})-\omega_1(\mathbf{b})) + \lambda_2'(\omega_2(\mathbf{e})-\omega_2(\mathbf{d})) + \omega_1(\mathbf{b})\end{aligned}$$

and it follows the circuit inequality

$$\omega_2(\mathbf{f}) - \omega_2(\mathbf{d}) > \lambda_1'(\omega_1(\mathbf{c})-\omega_1(\mathbf{b})) + \lambda_2'(\omega_2(\mathbf{e})-\omega_2(\mathbf{d})). \tag{2.20}$$

Equation (2.19) along with (2.20) implies that $(\mathbf{f}-\mathbf{d},\omega_2(\mathbf{f})-\omega_2(\mathbf{d}))$ is "above" the subspace H spanned by $(\mathbf{c}-\mathbf{b},\omega_1(\mathbf{c})-\omega_1(\mathbf{b}))$ and $(\mathbf{e}-\mathbf{d},\omega_2(\mathbf{e})-\omega_2(\mathbf{d}))$. Obviously, the right-hand side of (2.20) is the projection of $(\mathbf{f}-\mathbf{d},\omega_2(\mathbf{f})-\omega_2(\mathbf{d}))$ onto H, which equals s in (2.13). Ultimately, the circuit inequality is equivalent to the normal inequality. □

Example 4 In the previous example, we have two mixed cells with mixed volume 2. As depicted in Figure 2.2, there is a mixed cell whose volume equals the mixed volume. Let $\pi : \widetilde{\mathcal{A}} \longrightarrow \{\mathcal{A}_1,\mathcal{A}_2\}$ be the projection onto original supports. Based on a mixed cell $C = (\{\mathbf{b},\mathbf{c}\},\{\mathbf{d},\mathbf{e}\})$ previously found and by Cayley embedding, there exists a circuit $Z := \widetilde{C}\cup\{\mathbf{a}\} = \{\widetilde{\mathbf{a}},\widetilde{\mathbf{b}},\widetilde{\mathbf{c}},\widetilde{\mathbf{d}},\widetilde{\mathbf{e}}\}$ (see Figure 2.4) with two associated triangulations

$$\begin{aligned}T^+ &:= \left\{Z\backslash\{\widetilde{\mathbf{a}}\},\ Z\backslash\{\widetilde{\mathbf{c}}\},\ Z\backslash\{\widetilde{\mathbf{e}}\}\right\},\\ T^- &:= \left\{Z\backslash\{\widetilde{\mathbf{b}}\},\ Z\backslash\{\widetilde{\mathbf{d}}\}\right\}\end{aligned}$$

in $\mathbb{R}^3$ such that

$$\pi(T^+) = \Big\{ \underset{(1)\,:\,(1,1)\text{ type}}{(\{\mathbf{b},\mathbf{c}\},\{\mathbf{d},\mathbf{e}\})},\ \underset{(2)\,:\,(1,1)\text{ type}}{(\{\mathbf{a},\mathbf{b}\},\{\mathbf{d},\mathbf{e}\})},\ \underset{(3)}{(\{\mathbf{a},\mathbf{b},\mathbf{c}\},\{\mathbf{d}\})} \Big\}$$

and

$$\pi(T^-) = \Big\{ \underset{(4)\,:\,(1,1)\text{ type}}{(\{\mathbf{a},\mathbf{c}\},\{\mathbf{d},\mathbf{e}\})},\ \underset{(5)}{(\{\mathbf{a},\mathbf{b},\mathbf{c}\},\{\mathbf{e}\})} \Big\}$$

Apparently, T^- induces fewer number of mixed cell. □

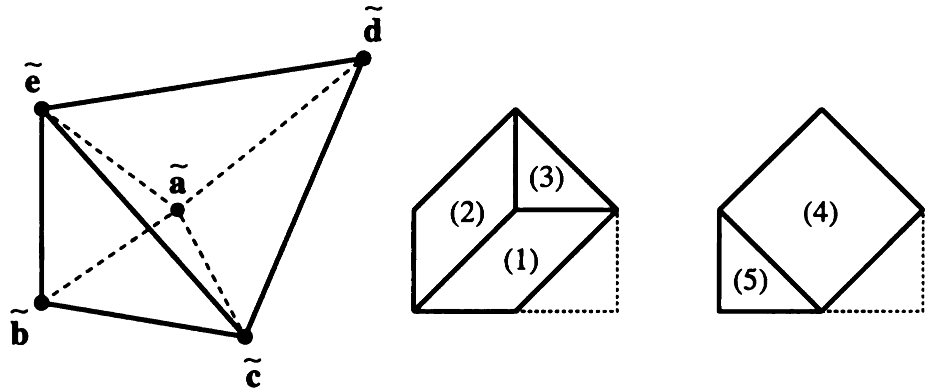

Figure 2.4: Circuit $Z = \{\widetilde{\mathbf{a}}, \widetilde{\mathbf{b}}, \widetilde{\mathbf{c}}, \widetilde{\mathbf{d}}, \widetilde{\mathbf{e}}\}$ and projections of T^{+} and T^{-}

The above observation stimulates the potential that the number of inequalities in **LP1** can be further reduced by verifying the flip operations that involve mixed cells as defined in [MiVe]. Namely, one may check whether the mixed-cell configuration is such that the flips determined by the circuit can be performed. If this is not the case, then the corresponding circuit is not supported and its inequality can be discarded provided that the optimal solution of **LP1** is not interfered. The problem of implementing this idea efficiently remains unclear at this stage. Another possible improvement is to combine the balancing strategy and the flattening method [VGC] when the optimal value μ obtained by the balancing method is still too large for practical computations.

CHAPTER 3

Numerical Experiments

Our algorithm in balancing the powers of t by solving the linear programming problem (2.10) has been successfully implemented. In our numerical experiments, the non-vertex points in the supports of the polynomial systems are deleted before calculating the mixed-cell configurations, and the mixed-cell configurations $\mathcal{M}_\omega$ are generated by our implementation of the Lift-Prune algorithm [EmCa] with random real liftings. The numerical experiments are done on a PC with a Pentium 166MHz processor, 16Mb RAM, 256Kb cache and Linux operating system. The numerical results of applying our algorithm to several well-known polynomial systems as listed below in Tables 3 and 3.

Polynmial Systems

- **Cyclic-n problem** [EmCa] The general formulation goes as follows:

$$P(\mathbf{x}) = \begin{cases} \displaystyle\sum_{i=1}^{n}\prod_{j=1}^{k} x_{(i+j)\bmod n}, & k = 1, \cdots, n-1 \\ \displaystyle\prod_{j=1}^{n} x_j \; - 1. & \end{cases}$$

with variables $\mathbf{x} = (x_1, \cdots, x_n)$.

- **The Cohn-2 system** from PoSSo test suite [PoSSo]:

$$\begin{aligned}
p_1(\mathbf{x}) &= x^3y^2 + 4x^2y^2z - x^2yz^2 + 288x^2y^2 + 207x^2yz + 1152xy^2z \\
&\quad +156xyz^2 + xz^3 - 3456x^2y + 20736xy^2 + 19008xyz + 82944y^2z \\
&\quad +432xz^2 - 497664xy + 62208xz + 2985984x, \\
p_2(\mathbf{x}) &= y^3t^3 + 4y^3t^2 - y^2zt^2 + 4y^2t^3 - 48y^2t^2 - 5yzt^2 \\
&\quad +108yzt + z^2t + 144zt - 1728z, \\
p_3(\mathbf{x}) &= -x^2z^2t + 4xz^2t^2 + z^3t^2 + x^3z + 156x^2zt + 207xz^2t + 1152xzt^2 \\
&\quad +288z^2t^2 + 432x^2z + 19008xzt - 3456z^2t + 82944xt^2 \\
&\quad +20736zt^2 + 62208xz - 497664zt + 2985984z, \\
p_4(\mathbf{x}) &= y^3t^3 - xy^2t^2 + 4y^3t^2 + 4y^2t^3 - 5xy^2t - 48y^2t^2 \\
&\quad +x^2y + 108xyt + 144xy - 1728x,
\end{aligned}$$

with variables $\mathbf{x} = (x, y, z, t)$.

- **Cassou-Nogues system** [Li97]

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

with variables b, c, d, e.

- **Planar 4-bar mechanism system** [MoWa]:

$$\begin{aligned} p_i(\mathbf{x}) &= a_{i,1}x_1^2x_3^2 + a_{i,2}x_1^2x_3x_4 + a_{i,3}x_1^2x_3 + a_{i,4}x_1^2x_4^2 + a_{i,5}x_1^2x_4 \\ &\quad + a_{i,6}x_1^2 + a_{i,7}x_1x_2x_3^2 + a_{i,8}x_1x_2x_3x_4 + a_{i,9}x_1x_2x_3 + a_{i,10}x_1x_2x_4^2 \\ &\quad + a_{i,11}x_1x_2x_4 + a_{i,12}x_1x_3^2 + a_{i,13}x_1x_3x_4 + a_{i,14}x_1x_3 + a_{i,15}x_1x_4^2 \\ &\quad + a_{i,16}x_1x_4 + a_{i,17}x_2^2x_3^2 + a_{i,18}x_2^2x_3x_4 + a_{i,19}x_2^2x_3 + a_{i,20}x_2^2x_4^2 \\ &\quad + a_{i,21}x_2^2x_4 + a_{i,22}x_2^2 + a_{i,23}x_2x_3^2 + a_{i,24}x_2x_3x_4 + a_{i,25}x_2x_3 \\ &\quad + a_{i,26}x_2x_4^2 + a_{i,27}x_2x_4 + a_{i,28}x_3^2 + a_{i,29}x_4^2, \qquad i = 1, \ldots, 4 \end{aligned}$$

with variables $\mathbf{x} = (x_1, x_2, x_3, x_4)$ and generic choice of coefficents $a_{i,j}$'s of the system.

Polynomial System	n	$\#\mathcal{M}_\omega$	Size of **LP0**		Size of **LP1**	
			#var	#ineq	#var	#ineq
Cohn-2 [PoSSo]	4	17	31	748	23	690
Cassou-Noguès [Li97]	4	3	28	114	20	106
Planar 4-bar [MoWa]	4	4	33	192	25	168
Cyclic-6 [EmCa]	6	25	33	1000	21	692
Cyclic-7 [EmCa]	7	126	45	7560	31	4982
Cyclic-8 [EmCa]	8	297	59	24948	43	16118

Table 3.1: Sizes of the Linear Programming problems. Here, n is the number of variables of the polynomial system and $\#\mathcal{M}_\omega$ is the number of mixed cells in the mixed-cell configuration $\mathcal{M}_\omega$.

The data in Table 3.1 are generated by the program with one random lifting function ω for each polynomial system. The fourth and fifth columns give the size of the linear programming problem in (2.8). The last two columns are the size of the linear programming problem in (2.10) after all repeated constraints are deleted. For cyclic-n polynomial systems, about 1/3 of the constraints are deleted, which results in a considerable speed-up.

Polynomial System	Avg highest power of t		Avg CPU time	
	Before balancing	After balancing	Finding mixed cells	Balancing method
Cohn-2 [PoSSo]	1391	85	0.21	0.19
Cassou-Noguès [Li97]	251	11	0.05	0.03
Planar 4-bar [MoWa]	429	8	0.17	0.08
Cyclic-6 [EmCa]	425	31	0.46	0.17
Cyclic-7 [EmCa]	3152	139	7.1	1.9
Cyclic-8 [EmCa]	10281	398	81	16.6

Table 3.2: Height of Powers and CPU Time in Seconds. The averages are obtained from ten different random liftings.

For the data in Table 3.2, we run the algorithm with ten different real random liftings for each polynomial system. We first scale the powers of t in the polyhedral homotopies before balancing such that the lowest power of t in the homotopies is one, and the average of the highest powers of t in the polyhedral homotopies for the ten random liftings are listed in the second column. The third column lists the average of the highest powers of t in the polyhedral homotopies for the ten liftings obtained from the optimal solutions of the corresponding linear programming problems (2.10). The fourth column gives the average time elapsed for finding all mixed cells. The last column is the average time elapsed for finding the optimal lifting functions ν', including the constructing and solving of the linear programming problems (2.10). From these results, we see that the highest powers of t in the polyhedral homotopies are considerablly reduced. The overall reduced powers of t in the polyhedral homotopies greatly limit the chance of running into a "valley" which may cause the failure of curve-tracing.

BIBLIOGRAPHY

BIBLIOGRAPHY

[AlGe93] E.L. Allgower and K. Georg, *Continuation and path following*, Acta Numerica, 1993, 1-64.

[Bern] D.N. Bernshtein, *The number of roots of a system of equations*, Functional Analysis and Appl. **9**(1975), 183-185.

[CaRo] J. Canny and J. M. Rojas (1991), "An optimal condition for determining the exact number of roots of a polynomial system", In *Proceedings of the 1991 International Symposium on Symbolic and Algebraic Computation*, ACM, 96-101.

[ChMPYo] S. N. Chow, J. Mallet-Paret, and J. A. Yorke (1978), "Finding zeros of maps: homotopy methods that are constructive with probability one", *Math. Comput.*, **32**, 887-899.

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

[GKZ] I.M. Gel'fand, M.M. Kapranov, and A.V. Zelevinsky, *Discriminants, Resultants and Multidimensional Determinants*, Birkhäuser, Boston, 1994.

[Hu96] B. Huber (1996), *Solving sparse polynomial systems*, Ph.D. dissertation, Cornell University.

[HuSt95] B. Huber and B. Sturmfels, *A polyhedral method for solving sparse polynomial systems*, Math. Comp. **64**(1995), 1541-1555.

[Kh78] A. G. Khovanskiǐ (1978), "Newton polyhedra and the genus of complete intersections", *Functional Anal. Appl.*, **12**(1), 38-46. Translated from *Funktsional. Anal. i Prilozhen.*, **12**(1), 51-61.

[Ku76] A. G. Kushnirenko (1976), "Newton Polytopes and the Bèzout Theorem", *Functional Anal. Appl.*, **10**(3), 233-235. Translated from *Funktsional. Anal. i Prilozhen.*, **10**(3), 82-83.

[LiWa96] T. Y. Li and X. Wang (1996), "The BKK root count in $\mathbb{C}^n$", *Math. Comp.*, **65**, no. 216, 1477-1484.

[Li97] T.Y. Li, *Numerical solutions of multivariate polynomial systems by homotopy continuation methods*, Acta Numerica **6**(1997), 399-436.

[MiVe] T. Michiels and J. Verschelde, *Enumerating regular mixed-cell configurations*, to appear in Discrete Comput. Geom.

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

[MoWa] A.P. Morgan and C.W. Wampler, *Solving a planar four-bar design problem using continuation*, ASME J. of Mechanical Design **112**(1990), 544-550.

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

[Ro94] J. M. Rojas (1994), "A convex geometric approach to counting the roots of a polynomial system", *Theoret. Comput. Sci.*, **133**, 105-140.

[RoWa] J. M. Rojas and X. Wang (1996), "Counting affine roots of polynomial systems via pointed Newton polytopes", *J. of Complexity*, **12**, no. 2, 116-133.

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

[St94] B. Sturmfels, *On the Newton polytope of the resultant.* J. of Algebraic Combinatorics, 3(1994), 207-236.

[St96] B. Sturmfels, *Gröbner Bases and Convex Polytopes*, volume 8 of University Lecture Series, AMS, Providence, R.I., 1996.

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

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