Moments matrices, real algebraic geometry and polynomial optimization

Marta Abril Bucero

UNIVERSITÉ DE NICE-SOPHIA ANTIPOLIS -UFR Sciences

Ecole doctorale de Sciences Fondamentales et appliquées

T H È S E

pour obtenir le titre

Docteur en Sciences

de l'UNIVERSITÉ de Nice - Sophia Antipolis

Discipline: (ou spécialité) Mathematiques Appliquées

présentée par

AUTEUR: Marta Abril Bucero

Matrices de Moments, Géométrie algébrique réelle et Optimisation polynomiale

Moments matrices, Real Algebraic Geometry and polynomial optimization

Thèse dirigée par: Bernard Mourrain

soutenue le 12 Décembre de 2014

Jury :

Rapporteurs :	Didier Henrion	-	CNRS-LAAS, Toulouse
	Markus Schweighofer	-	Universität Konstanz, Allemagne
	Mohab Safey el Din	-	LIP6, Paris
Directeur :	Bernard Mourrain	-	INRIA Sophia Antipolis
President :	André Galligo	-	Université de Nice
Examinateur:	Mariemi Alonso	-	Universidad Complutense, Madrid

Remerciements

Tout d'abord, je voudrais remercier mon directeur de thèse Bernard Mourrain pour me transmettre sa passion pour la recherche, pour son soutien, son encouragement, son humanité et sa patience pour me faire comprendre les choses. Pour m'avoir donneé l'opportunité de participer à beaucoup de conférences où j'ai eu la chance de rencontrer et discuter avec des chercheurs reconnus dans le domaine de l'optimization polynomial.
Ensuite, je remercie Didier Henrion, Markus Schweighofer et Mohab Safey El Din pour avoir accepté de faire un rapport de ce manuscrit malgré leur très lourd emploi du temps.
Je remercie tout aussi chaleureusement Andre Galligo et Mariemi Alonso pour avoir bien voulu être membre de mon jury et spécialement parce que sans eux je n'aurais pas pu avoir la posibilité de faire cette thèse. Merci de croire en moi.
Je voudrais remercier aussi mon directeur de travail du fin de master et professeur pendant mes années d'université Enrique Arrondo, pour m'avoir transmis sa passion pour les maths et m'avoir encouragé à faire le master.
Puis je voudrais remercier tous les galadiens et toutes les galadiennes que j'ai connus tout au long de ma thèse. En commencant pour ceux que m'ont accuelli, merci Alexandra, Nicolas, Mathieu, Meriadeg pour votre gentillese, votre aide et pour toutes ces soirées qu'on a partagé ensemble. Merci Matthieu pour ta joie de vivre, merci Valentin pour ta vision gaussienne et merci Meng pour me faire voir la vie avec d'autres yeux.
Et qu'aurait été cette thèse sans mes amis Emma, Rachid, Clement et Anais, merci pour me faire sourire chaque jour et pour votre soutien inconditionel.
Ensuite, je remercie chaleureusement ma famille et mes amis de Madrid pour me montrer que malgré la distance, ils ont toujours eté là.
Et enfin le plus important, merci Javi, parce que cette aventure n'aurais jamais eté posible sans toi. Merci pour toute la force et tout l'amour que tu m'as donnés tout au long de ces trois années.

Contents

1 Ideals, dual space, hankel matrices and quotient algebra **5**
- 1.1 Ideals and varieties . . . 5
- 1.2 Dual space and Hankel operators . . . 7
- 1.3 Artinian algebra . . . 11
- 1.4 Artinian Gorenstein algebra and positive linears forms . . . 12

2 Minimization problem and varieties of critical points **19**
- 2.1 The gradient variety . . . 20
- 2.2 The Karush-Kuhn-Tucker variety . . . 21
- 2.3 The Fritz John variety . . . 23
- 2.4 The minimizer variety . . . 26

3 Relation between Optimization problem and Moment matrices **31**
- 3.1 Positive Polynomials . . . 32
- 3.2 Moment matrices . . . 37
- 3.3 Lasserre relaxation . . . 41

4 Finite Convergence Certification **45**
- 4.1 Representation of positive polynomials . . . 47
- 4.2 Finite convergence . . . 53
- 4.3 Consequences . . . 58
  - 4.3.1 Global optimization . . . 58
  - 4.3.2 General case . . . 59
  - 4.3.3 Regular case . . . 59
  - 4.3.4 Zero dimensional real variety . . . 60
  - 4.3.5 Smooth real variety . . . 61
  - 4.3.6 Known minimum . . . 62
  - 4.3.7 Radical computation. . . . 62

5 Border basis relaxation for polynomial optimization **65**
- 5.1 Border basis . . . 66
- 5.2 Border basis hierarchy . . . 69
  - 5.2.1 Optimal linear form . . . 71
- 5.3 Convergence certification . . . 73
  - 5.3.1 Flat extension criterion . . . 74
  - 5.3.2 Flat extension algorithm . . . 75

5.3.3 Computing the minimizers 80
5.4 Minimizer border basis algorithm 82
5.4.1 Example in detail 84
5.4.2 Examples 86
5.4.3 Examples of theorical results for ideals non zero-dimensional 88

6 Experimentations, Applications and Implementation 93
6.1 Experimentations 93
6.2 Applications 97
6.2.1 Best low-rank tensor approximation 97
6.2.2 Factors in the growth of the plant roots 106
6.2.3 Marx generators design 111
6.3 Implementation 117
6.3.1 Input arguments, Input data file and Output data file . 119

Bibliography 125

Introduction

L'optimisation, c'est à dire le calcul du minimum d'une fonction f à valeurs réelles, est un problème important des mathématiques "numériques" et qui a des applications dans de nombreux domaines. L'approche classique pour résoudre ce problème est basée sur des techniques de descente du gradient de f. L'inconvénient de ces techniques reside dans le fait qu'elles calculent un optimun local qui n'est pas nécessairement un optimum global. Depuis une quinzaine d'années, de nouvelles techniques algébriques, dites de relaxation, visant un meilleur contrôle des résultats calculés ont été mises au point.

Ces méthodes sont basées sur une reformulation de la question en termes de matrices de moments. Ainsi, la recherche de solutions réelles (sans tenir compte des solutions complexes) d'un problème d'optimisation ou d'un système d'équations polynomiale est remplacée par le calcul de mesures ayant des propriétés spécifiques. Cette approche repose sur des techniques de programmation semidéfinie (SDP) et d'algèbre linéaire numérique. Toute l'information nécessaire est alors contenue dans la matrice des moments, dont les lignes et les colonnes sont indexées par une base de monômes. L'inconvénient majeur de cette approche est que la taille de la matrice des moments, égale au nombre de monômes d'un degré particulier, augmente à chaque boucle de l'algorithme et devient potentiellement importante.

Récemment, certaines améliorations ont été proposées pour réduire la taille de la matrice des moment et donc la taille du problème soumis au solveur SDP. Elles s'inspirent de la méthode des bases de bord pour la résolution de systèmes d'équations polynomiales. L'idée est de sélectionner certains monômes, considérés comme des candidats à une base de l'espace quotient : Au cours de l'algorithme, les dimensions des systèmes linéaires à résoudre sont alors liées au nombre de monômes associés à la base de bord et sont donc mieux controllées. Une fonctionnalité intéressante de cette approche (en contraste avec les approches de type base de Groebner) est sa robustesse par rapport aux perturbations de coéfficients dans le système d'origine.

Le but du présent manuscrit est d'étudier la combinaison des méthodes de base de bord, de l'approche de relaxation et des techniques de programmation semidéfinie, afin de calculer l'optimum d'un polynôme sur un ensemble semi-algébrique. Plus précisément calculer :

$$\begin{aligned} \inf_{\mathbf{x}\in\mathbb{R}^n} \quad & f(\mathbf{x}) \\ s.t. \quad & g_1^0(\mathbf{x}) = \cdots = g_{n_1}^0(\mathbf{x}) = 0 \\ & g_1^+(\mathbf{x}) \geq 0, ..., g_{n_2}^+(\mathbf{x}) \geq 0 \end{aligned} \tag{1}$$

Notre manuscrit comporte une introduction et six chapitres.

Dans le premier chapitre, nous fixons les notations et nous rappelons les concepts et théorèmes sur les idéaux de polynômes, les opérateurs de Hankel, les formes positives et l'algèbre quotient.

Dans le deuxième chapitre nous définissons notre problème d'optimisation et nous précisons les définitions de variétés de points critiques (variétés gradient, de Karush Kuhn Tucker ou de Fritz John) ainsi que les relations qui existent entre elles. Dans tous les travaux précédents, les auteurs supposent que le minimum doit être un point de la variété de Karush Kuhn Tucker. Nous pouvons éliminer cette hypothèse en utilisant une variété de Fritz John. Ceci constitue une partie de notre première contribution à l'optimisation polynomiale.

Dans le troisième Chapitre nous expliquons comment les polynômes positifs, et les matrices de moments interviennent dans l'optimisation polynomiale et (dans la dernière section) nous rappelons la méthode de relaxation de Lasserre.

Dans le quatrième chapitre nous traitons de la représentation des polynômes positifs et nous montrons une propriété de convergence finie (convergence en un nombre fini de pas) dans un cadre plus général que celui habituellement considéré. Puis, nous déduisons les conséquences de cette convergence dans des cas particuliers intéressants.

Dans le cinquième chapitre nous expliquons notre algorithme. Nous utilisons la méthode de relaxation de Lasserre combinée avec les bases de bord pour réduire la taille des matrices de moments et aussi les nombres de paramètres à chercher dans notre SDP. Nous donnons aussi un nouveau critère de terminaison qui vérifie que l'extension est plate et permet ainsi de savoir quand le minimum est atteint. Dans la dernière partie de ce chapitre nous expliquons en détail comment fonctionne notre algorithme, en illustrant avec des exemples.

Dans le sixième et dernier chapitre, nous montrons les expérimentations que nous avons réalisées, puis nous les comparons avec les résultats fournis par un logiciel déjà commercialisé. Nous donnons trois applications de notre algorithme dans trois domaines différents, ce qui éclaire la façon dont notre travail peut servir. La fin de ce chapitre fournit des détails sur la manière dont nous avons implémente notre algorithme et sur la manière de l'utiliser.

Les résultats des deuxième et quatrième chapitre d'une part et du cinquième chapitre d'autre part font l'objet de deux pre-publications que nous avons soumises pour publication [Abril Bucero 2013, Abril Bucero 2014].

CHAPTER 1

Ideals, dual space, hankel matrices and quotient algebra

In this chapter, in the first section we set our notation and we recall definitions and theorems about ideals and varieties. In Section 2 we give the definition of Hankel Matrix and its properties and theorems. In Section 3 and 4 we study the properties of the quotient ring obtained from quotient by the kernel of the an Hankel operator.

1.1 Ideals and varieties

Let $\mathbb{K}[\mathbf{x}]$ be the set of the polynomials in the variables $\mathbf{x} = (x_1, \ldots, x_n)$, with coefficients in the field $\mathbb{K}$. Hereafter, we will choose $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. Let $\overline{\mathbb{K}}$ denotes the algebraic closure of $\mathbb{K}$. For $\alpha \in \mathbb{N}^n$, $\mathbf{x}^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ is the monomial with exponent α and degree $|\alpha| = \sum_i \alpha_i$. The set of all monomials in $\mathbf{x}$ is denoted $\mathcal{M} = \mathcal{M}(\mathbf{x})$. For a polynomial $f = \sum_\alpha f_\alpha \mathbf{x}^\alpha$, its support is $supp(f) := \{\mathbf{x}^\alpha \mid f_\alpha \neq 0\}$, the set of monomials occurring with a nonzero coefficient in f.
For $t \in \mathbb{N}$, $\mathbb{N}^n_t = \{\alpha \in \mathbb{N}^n \mid\mid \alpha \mid := \sum_{i=1}^n \alpha_i \leq t\}$.

For $t \in \mathbb{N}$ and $D \subseteq \mathbb{K}[\mathbf{x}]$, we introduce the following sets:

- D_t is the set of elements of D of degree $\leq t$,
- $\langle D \rangle = \left\{ \sum_{f \in S} \lambda_f f \mid f \in S, \lambda_f \in \mathbb{K} \right\}$ is the linear span of S,
- $\langle S \mid t \rangle = \left\{ \sum_{f \in S_t} p_f f \mid p_f \in \mathbb{K}[\mathbf{x}]_{t-\deg(f)} \right\}$ is the vector space spanned by $\{\mathbf{x}^\alpha f \mid f \in S_t, |\alpha| \leq t - \deg(f)\}$,
- $S_t = S \cap \mathbb{K}[\mathbf{x}]_t$
- $S^{[t]} = \{x^\alpha f \mid f \in S, |\alpha| \leq t\}$,
- $\mathcal{Q}_t^+ = \left\{ \sum_{i=1}^l p_i^2 \mid l \in \mathbb{N}, p_i \in \mathbb{R}[\mathbf{x}]_t \right\}$ is the set of finite sums of squares of polynomials of degree $\leq t$; $\mathcal{Q}^+ = \mathcal{Q}_\infty^+$ (sum of squares SOS).

Remark 1.1.1 $\langle S \mid t \rangle \subseteq (S) \cap \mathbb{K}[\mathbf{x}]_t = (S)_t$, *but the inclusion may be strict.*

Given an ideal $I \subseteq \mathbb{K}[\mathbf{x}]$ and a field $\mathbb{K}$, we denote by

$$V_{\mathbb{K}}(I) := \{x \in \mathbb{K}^n \mid f(x) = 0 \ \forall f \in I\}$$

its associated variety in $\mathbb{L}^n$. By convention $V(I) = V_{\overline{\mathbb{K}}}(I)$. For a set $V \subseteq \mathbb{K}^n$, we define its vanishing ideal

$$I(V) := \{f \in \mathbb{K}[\mathbf{x}] \mid f(v) = 0 \ \forall v \in V\}.$$

Furthermore, we denote by

$$\sqrt{I} := \{f \in \mathbb{K}[\mathbf{x}] \ \mid \ f^m \in I \ \text{ for some } m \in \mathbb{N} \setminus \{0\}\}$$

the radical of I.

For $\mathbb{K} = \mathbb{R}$, we have $V(I) = V_{\mathbb{C}}(I)$, but one may also be interested in the subset of real solutions, namely the real variety $V_{\mathbb{R}}(I) = V(I) \cap \mathbb{R}^n$. The corresponding vanishing ideal is $I(V_{\mathbb{R}}(I))$ and the *real radical ideal* is
$\sqrt[\mathbb{R}]{I} := \{p \in \mathbb{R}[\mathbf{x}] \mid p^{2m} + \sum_j q_j^2 \in I \ \text{ for some } q_j \in \mathbb{R}[\mathbf{x}], m \in \mathbb{N} \setminus \{0\}\}$.
Obviously,

$$I \subseteq \sqrt{I} \subseteq I(V_{\mathbb{C}}(I)), \quad I \subseteq \sqrt[\mathbb{R}]{I} \subseteq I(V_{\mathbb{R}}(I)).$$

An ideal I is said to be *radical* (resp., *real radical*) if $I = \sqrt{I}$ (resp. $I = \sqrt[\mathbb{R}]{I}$). Obviously, $I \subseteq I(V(I)) \subseteq I(V_{\mathbb{R}}(I))$. Hence, if $I \subseteq \mathbb{R}$ is real radical, then I is radical and moreover, $V(I) = V_{\mathbb{R}}(I) \subseteq \mathbb{R}^n$ if $|V_{\mathbb{R}}(I)| < \infty$.

The following two famous theorems relate vanishing and radical ideals:

Theorem 1.1.2

(i) **Hilbert's Nullstellensatz** $\sqrt{I} = I(V_{\mathbb{C}}(I))$ *for an ideal* $I \subseteq \mathbb{C}[\mathbf{x}]$.

(ii) **Real Nullstellensatz** $\sqrt[\mathbb{R}]{I} = I(V_{\mathbb{R}}(I))$ *for an ideal* $I \subseteq \mathbb{R}[\mathbf{x}]$.

By convention, a set of constrains $C = \{c_1^0, \ldots, c_{n_1}^0; c_1^+, \ldots, c_{n_2}^+\} \subset \mathbb{R}[\mathbf{x}]$ is a finite set of polynomials composed of a subset $C^0 = \{c_1^0, \ldots, c_{n_1}^0\}$ corresponding to the equality constraints and a subset $C^+ = \{c_1^+, \ldots, c_{n_2}^+\}$ corresponding to the non-negativity constraints. For two set of constraints $C, C' \subset \mathbb{R}[\mathbf{x}]$, we say that $C \subset C'$ if $C^0 \subset C'^0$ and $C^+ \subset C'^+$.

Definition 1.1.3 *For* $t \in \mathbb{N} \cup \{\infty\}$ *and a set of constraints* $C = \{c_1^0, \ldots, c_{n_1}^0; c_1^+, \ldots, c_{n_2}^+\} \subset \mathbb{R}[\mathbf{x}]$, *we define the (truncated) quadratic module of* C *by*

$$\mathcal{Q}_t(C) = \{\sum_{i=1}^{n_2} c_i^0 \, h_i + s_0 + \sum_{j=1}^{n_2} c_j^+ \, s_j \mid h_i \in \mathbb{R}[\mathbf{x}]_{2t-\deg(c_i^0)}, s_0 \in \mathcal{Q}_t^+, s_i \in \mathcal{Q}_{t-\lceil \deg(c_i^+)/2 \rceil}^+\}.$$

If $\tilde{C}$ is such that $\tilde{C}^0 = C^0$ and $\tilde{C}^+ = \{\prod (c_1^+)^{\varepsilon_1} \cdots (c_{n_2}^+)^{\varepsilon_{n_2}} \mid \varepsilon_i \in \{0,1\}\}$, $\mathcal{Q}_t(\tilde{C})$ is also called the (truncated) preordering of C and denoted $\mathcal{P}_t(C)$. When $t = \infty$, $\mathcal{P}(C) := \mathcal{P}_\infty(C)$ is the preordering of C. The (truncated) preordering generated by the positive constraints is denoted $\mathcal{P}^+(C) = \mathcal{P}(C^+)$.

Definition 1.1.4 *For a set of constrains $C = (C^0; C^+) \subset \mathbb{R}[\mathbf{x}]$,*

$$\begin{aligned} \mathcal{S}(C) &:= \{\mathbf{x} \in \mathbb{R}^n \mid c^0(\mathbf{x}) = 0 \ \forall c^0 \in C^0,\ c^+(\mathbf{x}) \geq 0 \ \forall c^+ \in C^+\}, \\ \mathcal{S}^+(C) &:= \{\mathbf{x} \in \mathbb{R}^n \mid c^+(\mathbf{x}) \geq 0 \ \forall c^+ \in C\}. \end{aligned}$$

To describe the vanishing ideal of these sets, we introduce the following ideals:

Definition 1.1.5 *For a set of constraints $C = (C^0; C^+) \subset \mathbb{R}[\mathbf{x}]$,*

$$\begin{aligned} \sqrt{C^0} &= \{p \in \mathbb{R}[\mathbf{x}] \mid p^m \in (C^0) \ \text{ for some } m \in \mathbb{N} \setminus \{0\}\} \\ \sqrt[\mathbb{R}]{C^0} &= \{p \in \mathbb{R}[\mathbf{x}] \mid p^{2m} + q \in (C^0) \ \text{ for some } m \in \mathbb{N} \setminus \{0\}, q \in \mathcal{Q}^+\} \\ \sqrt[C^+]{C^0} &= \{p \in \mathbb{R}[\mathbf{x}] \mid p^{2m} + q \in (C^0) \ \text{ for some } m \in \mathbb{N} \setminus \{0\}, q \in \mathcal{P}^+(C)\} \end{aligned}$$

These ideals are called respectively the radical of C^0, the real radical of C^0, the C^+-radical of C^0.

Remark 1.1.6 *If $C^+ = \emptyset$, then $\sqrt[C^+]{C^0} = \sqrt[\mathbb{R}]{C^0}$.*

The following three famous theorems relate vanishing and radical ideals:

Theorem 1.1.7 *Let $C = (C^0; C^+)$ be a set of constraints of $\mathbb{R}[\mathbf{x}]$.*

(i) **Hilbert's Nullstellensatz** *(see, e.g., [Cox 2005, §4.1])* $\sqrt{C^0} = \mathcal{I}(\mathcal{V}^{\mathbb{C}}(C^0))$.

(ii) **Real Nullstellensatz** *(see, e.g., [Bochnak 1998, §4.1])* $\sqrt[\mathbb{R}]{C^0} = \mathcal{I}(\mathcal{V}^{\mathbb{R}}(C^0))$.

(iii) **Positivstellensatz** *(see, e.g., [Bochnak 1998, §4.4])* $\sqrt[C^+]{C^0} = \mathcal{I}(\mathcal{S}(C)) = \mathcal{I}(\mathcal{V}^{\mathbb{R}}(C^0) \cap \mathcal{S}^+(C))$.

1.2 Dual space and Hankel operators

This Section is an introduction to the dual space and Hankel operators which are basic elements in our study.

Definition 1.2.1 $\mathbb{K}[\mathbf{x}]^* := Hom_{\mathbb{K}}(\mathbb{K}[\mathbf{x}], \mathbb{K})$ *called dual space of* $\mathbb{K}[\mathbf{x}]$*, is the set of* $\mathbb{K}$*-linear forms from* $\mathbb{K}[\mathbf{x}]$ *to* $\mathbb{K}$*.*

There exists a natural isomorphism between the ring of formal power series and the dual space ring of polynomials $\mathbb{K}[\mathbf{x}]$. It is given the following pairing:

$$\begin{array}{rcl} \mathbb{K}[[\boldsymbol{z}]] \times \mathbb{K}[\mathbf{x}] & \to & \mathbb{K} \\ (\boldsymbol{z}^\alpha, \mathbf{x}^\beta) & \mapsto & \langle \boldsymbol{z}^\alpha | \mathbf{x}^\beta \rangle = \left\{ \begin{array}{ll} \alpha! & \text{if } \alpha = \beta \\ 0 & \text{otherwise}. \end{array} \right. \end{array}$$

If $\Lambda \in \mathrm{Hom}_{\mathbb{K}}(\mathbb{K}[\mathbf{x}], \mathbb{K}) = \mathbb{K}[\mathbf{x}]^*$ is an element of the dual of $\mathbb{K}[\mathbf{x}]$, it can be represented by the series:

$$\Lambda(\boldsymbol{z}) = \sum_{\alpha \in \mathbb{N}^n} \Lambda(\mathbf{x}^\alpha) \frac{\boldsymbol{z}^\alpha}{\alpha!} \in \mathbb{K}[[z_1, \ldots, z_n]], \tag{1.1}$$

so that we have $\langle \Lambda(\boldsymbol{z}) | \mathbf{x}^\alpha \rangle = \Lambda(\mathbf{x}^\alpha)$.
This map $\Lambda \in R^* \mapsto \sum_{\alpha \in \mathbb{N}^n} \Lambda(\mathbf{x}^\alpha) \frac{\boldsymbol{z}^\alpha}{\alpha!} \in \mathbb{K}[[\boldsymbol{z}]]$ is an isomorphism. And therefore any series defined as $\Lambda(\boldsymbol{z}) = \sum_{\alpha \in \mathbb{N}^n} \Lambda_\alpha \frac{\boldsymbol{z}^\alpha}{\alpha!} \in \mathbb{K}[[\boldsymbol{z}]]$ can be interpreted as a linear form in $\mathbb{K}[\mathbf{x}]$

$$p(\mathbf{x}) = \sum_{\alpha \in A \subset \mathbb{N}^n} p_\alpha \mathbf{x}^\alpha \in \mathbb{K}[\mathbf{x}] \mapsto \langle \Lambda \mid p(\mathbf{x}) \rangle = \sum_{\alpha \in A \subset \mathbb{N}^n} p_\alpha \Lambda_\alpha.$$

From now on, we identify the dual $\mathrm{Hom}_{\mathbb{K}}(\mathbb{K}[\mathbf{x}], \mathbb{K})$ with $\mathbb{K}[[\boldsymbol{z}]]$. Using this identification, the dual basis of the monomial basis $(\mathbf{x}^\alpha)_{\alpha \in \mathbb{N}^n}$ is $\left(\frac{\boldsymbol{z}^\alpha}{\alpha!}\right)_{\alpha \in \mathbb{N}^n}$. The coefficients $\sigma_\alpha = \langle \Lambda \mid \mathbf{x}^\alpha \rangle$ are called the *moments* of Λ.

Among interesting elements of $\mathrm{Hom}(\mathbb{K}[\mathbf{x}], \mathbb{K}) \equiv \mathbb{K}[[\boldsymbol{z}]]$, we have the evaluations at points of $\mathbb{C}^n$:

Definition 1.2.2 *The evaluation at a point* $\xi \in \mathbb{K}^n$ *is:*

$$\begin{array}{rcl} \mathbf{1}_\xi : \mathbb{K}[x_1, \ldots x_n] & \to & \mathbb{K} \\ p(\mathbf{x}) & \mapsto & p(\xi) \end{array}$$

which corresponds to the formal series:

$$\mathbf{1}_\xi(\boldsymbol{z}) \;=\; \sum_{\alpha \in \mathbb{N}^n} \xi^\alpha \frac{\boldsymbol{z}^\alpha}{\alpha!} \;=\; e^{\langle \xi \mathbf{z} \rangle}.$$

Using this formalism, the series $\Lambda(\boldsymbol{z}) = \sum_{i=1}^r \omega_i \mathbf{1}_{\xi_i}(\boldsymbol{z})$ can be interpreted as a linear combination of evaluations at the points ξ_i which coefficients are ω_i, for $i = 1, \ldots, r$.

Notice that the product of $\boldsymbol{z}^\alpha \mathbf{1}_\xi(\boldsymbol{z})$ with a monomial $\mathbf{x}^{\alpha+\beta} \in \mathbb{C}[\mathbf{x}]$ is given by

$$\langle \boldsymbol{z}^\alpha \mathbf{1}_\xi(\boldsymbol{z}) | \mathbf{x}^{\alpha+\beta} \rangle = \frac{(\alpha+\beta)!}{\beta!} \xi^\beta = \partial_{x_1}^{\alpha_1} \cdots \partial_{x_n}^{\alpha_n} \mathbf{x}^{\alpha+\beta}(\xi),$$

so that $\Lambda(\boldsymbol{z}) = \sum_{i=1}^r \omega_i(\boldsymbol{z}) \mathbf{1}_{\xi_i}(\boldsymbol{z})$can be seen as a sum of *polynomial differential operators* $\omega_i(\partial)$ "at" the points ξ_i, that we call *infinitesimal operators*: $\forall p \in \mathbb{C}[\mathbf{x}], \langle \Lambda(\boldsymbol{z}) | p(\mathbf{x}) \rangle = \sum_{i=1}^r \omega_i(\partial) p(\xi)$.

Definition 1.2.3 *For any $\Lambda(\boldsymbol{z}) \in \mathbb{K}[[\boldsymbol{z}]]$, the inner product associated to $\Lambda(\boldsymbol{z})$ on $\mathbb{K}[\mathbf{x}]$ is*

$$\begin{array}{rcl} \mathbb{K}[\mathbf{x}] \times \mathbb{K}[\mathbf{x}] & \to & \mathbb{K} \\ (p(\mathbf{x}), q(\mathbf{x})) & \mapsto & \langle p(\mathbf{x}), q(\mathbf{x}) \rangle_\Lambda := \langle \Lambda(\boldsymbol{z}) | p(\mathbf{x}) q(\mathbf{x}) \rangle = \Lambda(pq). \end{array}$$

The dual space $\mathrm{Hom}(\mathbb{K}[\mathbf{x}], \mathbb{K}) \equiv \mathbb{K}[[\boldsymbol{z}]]$ has a natural structure of $\mathbb{K}[\mathbf{x}]$-module, defined as follows: $\forall \sigma(\boldsymbol{z}) \in \mathbb{K}[[\boldsymbol{z}]], \forall p(\mathbf{x}), q(\mathbf{x}) \in \mathbb{K}[\mathbf{x}]$,

$$\langle p(\mathbf{x}) \star \Lambda(\boldsymbol{z}) \mid q(\mathbf{x}) \rangle = \langle \Lambda(\boldsymbol{z}) \mid p(\mathbf{x}) q(\mathbf{x}) \rangle = \langle p(\mathbf{x}), q(\mathbf{x}) \rangle_\Lambda = \Lambda(pq).$$

We easily check that $\forall \Lambda \in \mathbb{K}[[\boldsymbol{z}]], \forall p, q \in \mathbb{K}[\mathbf{x}]$, $(pq) \star \Lambda = p \star (q \star \Lambda)$.

Example 1.2.4 *If $\Lambda(\boldsymbol{z}) = \sum_{i=1}^r \omega_i \mathbf{1}_{\xi_i}(\boldsymbol{z})$, with $\omega_i \in \mathbb{K}$ and $\xi_i \in \mathbb{K}^n$ and $p(\mathbf{x}) \in \mathbb{K}[\mathbf{x}]$, we have*

$$p(\mathbf{x}) \star \Lambda(\boldsymbol{z}) = \sum_{i=1}^r \omega_i p(\xi_i) \mathbf{1}_{\xi_i}(\boldsymbol{z}). \tag{1.2}$$

An interesting property of this external product is that polynomials act as differentials on the series:

Lemma 1.2.5 $\forall p \in \mathbb{K}[\mathbf{x}], \forall \Lambda \in \mathbb{K}[[\boldsymbol{z}]]$, $p(\mathbf{x}) \star \Lambda(\boldsymbol{z}) = p(\partial_{z_1}, \ldots, \partial_{z_n})(\Lambda)$.

Proof. We first prove the relation for $p = x_i$ and $\Lambda = \boldsymbol{z}^\alpha$. Let $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$ be the exponent vector of x_i. $\forall \beta \in \mathbb{N}^n$, we have

$$\begin{array}{rcl} \langle x_i \star \boldsymbol{z}^\alpha | \mathbf{x}^\beta \rangle & = & \langle \boldsymbol{z}^\alpha | x_i \mathbf{x}^\beta \rangle = \quad \alpha! \quad \text{if} \quad \alpha = \beta + e_i \quad \text{and} \quad 0 \quad \text{otherwise} \\ & = & \alpha_i \langle \boldsymbol{z}^{\alpha - e_i} | \mathbf{x}^\beta \rangle. \end{array}$$

with the convention that $\boldsymbol{z}^{\alpha - e_i} = 0$ if $\alpha_i = 0$. This shows that $x_i \star \boldsymbol{z}^\alpha = \alpha_i \boldsymbol{z}^{\alpha - e_i} = \partial_{z_i}(\boldsymbol{z}^\alpha)$ as elements of $R^* \equiv \mathbb{K}[[\boldsymbol{z}]]$.

By transitivity and bilinearity of the product $\star$, we deduce that $\forall p \in \mathbb{K}[\mathbf{x}], \forall \Lambda \in \mathbb{K}[[\boldsymbol{z}]]$, $p(\mathbf{x}) \star \Lambda(\boldsymbol{z}) = p(\partial_{z1}, \ldots, \partial_{z_n})(\Lambda)$. ■

For a subset $D \subset \mathbb{K}[[\boldsymbol{z}]]$, the *inverse system* generated by D is the vector space spanned by the elements $p(\mathbf{x}) \star \delta(\boldsymbol{z})$ for $\delta(\boldsymbol{z}) \in D$ and $p(\mathbf{x}) \in \mathbb{K}[\mathbf{x}]$. By Lemma 1.2.5, the inverse system of D is the space generated by the elements of D and all their derivative in the variables $\boldsymbol{z}$ at any order.

The external product $\star$ allows us to define an Hankel operator as a multiplication operator by a dual element $\in \mathbb{K}[[\boldsymbol{z}]]$:

Definition 1.2.6 *The Hankel operator associated to an element $\Lambda(\boldsymbol{z}) \in \mathbb{K}[[\boldsymbol{z}]]$ is*

$$\begin{array}{rcl} H_\Lambda : \mathbb{K}[\mathbf{x}] & \to & \mathbb{K}[[\boldsymbol{z}]] \\ p(\mathbf{x}) & \mapsto & p(\mathbf{x}) \star \Lambda(\boldsymbol{z}). \end{array}$$

Definition 1.2.7 *Given a subspace $E \subset \mathbb{R}[\mathbf{x}]$, we define truncated Hankel operator defined on the subspace E, associated to an element $\Lambda \in \langle E \cdot E \rangle$ as*

$$\begin{array}{rcl} H_\Lambda^E : E & \to & E^* \\ p(\mathbf{x}) & \mapsto & p(\mathbf{x}) \star \Lambda. \end{array}$$

In particular if $E = \mathbb{R}[\mathbf{x}]_t$ we define H_Λ^t.

Definition 1.2.8 *The kernel of the Hankel operator associated to an element $\Lambda(\boldsymbol{z}) \in \mathbb{K}[[\boldsymbol{z}]]$ is*

$$\ker H_\Lambda = \{p(\mathbf{x}) \in \mathbb{K}[\mathbf{x}] \mid p(\mathbf{x}) \star \Lambda = 0\} \tag{1.3}$$

It is also denoted I_Λ.

Definition 1.2.9 *We say that the series Λ has a finite rank $r \in \mathbb{N}$ if* $\operatorname{rank} H_\Lambda = r < \infty$.

Example 1.2.10 *If $\Lambda = \mathbf{1}_\xi$ is the evaluation at a point $\xi \in \mathbb{K}^n$, then*

$$\begin{array}{rcl} H_{\mathbf{1}_\xi} : \mathbb{K}[\mathbf{x}] & \to & \mathbb{K}[[\boldsymbol{z}]] \\ p(\mathbf{x}) & \mapsto & p(\xi)\mathbf{1}_\xi \end{array}$$

Remark 1.2.11 *The matrix of the operator H_Λ in the bases $(\mathbf{x}^\alpha)_{\alpha \in \mathbb{N}^n}$ and $\left(\frac{\boldsymbol{z}^\alpha}{\alpha!}\right)_{\alpha \in \mathbb{N}^n}$ is*

$$[H_\Lambda] \;=\; (\Lambda_{\alpha+\beta})_{\alpha,\beta \in \mathbb{N}^n} = (\langle \Lambda | \mathbf{x}^{\alpha+\beta} \rangle)_{\alpha,\beta \in \mathbb{N}^n} = \Lambda(\mathbf{x}^{\alpha+\beta})_{\alpha,\beta \in \mathbb{N}^n}.$$

In the case $n = 1$, the coefficients of $[H_\Lambda]$ depends only the sum of the indices indexing the rows and columns, which explains why they are called *Hankel* operators.

1.3 Artinian algebra

In this Section, we consider an ideal $I \subset \mathbb{K}[\mathbf{x}]$, with $\mathbb{K}$ algebraically closed (i,e, $\mathbb{K} = \overline{\mathbb{K}}$) and the associated quotient algebra $\mathcal{A} = \mathbb{K}[\mathbf{x}]/I$

Definition 1.3.1 *The quotient algebra $\mathcal{A}$ is artinian if $dim_{\mathbb{R}}(\mathcal{A}) < \infty$*

A classical result states that the quotient algebra $\mathcal{A} = \mathbb{K}[\mathbf{x}]/I$ is finite dimensional, i.e, Artinian if and only if $\mathcal{V}(I)$ is finite, that is, I defines a finite number of isolated points in $\mathbb{K}^n$.

Theorem 1.3.2 *Let $\mathcal{A}$ be an Artinian algebra of dimension r defined by an ideal I. Then we have a direct sum*

$$\mathcal{A} = \mathcal{A}_{\xi_1} \oplus \cdots \oplus \mathcal{A}_{\xi_{r'}} \tag{1.4}$$

where

- $\mathcal{V}(I) = \{\xi_1, ..., \xi_{r'}\} \subset \mathbb{K}^n$ *with* $r' \leq r$,
- $I = Q_1 \cap \cdots \cap Q_{r'}$ *is a minimal primary decomposition of I with Q_i $m_{\xi_i}-$ primary,*
- $\mathcal{A}_{\xi_{r'}} \equiv \mathbb{K}[\mathbf{x}]/Q_i$ *and* $\mathcal{A}_{\xi_i} \cdot \mathcal{A}_{\xi_j} \equiv 0$ *if $i = j$. The multiplicity of an isolated point ξ_i of $\mathcal{V}(I)$ is the dimension over $\mathbb{R}$ of $\mathcal{A}$ localized at ξ_i, that is, $\mathcal{A}_{\xi_i}$*

Definition 1.3.3 *The dual $\mathcal{A}^* = Hom_{\mathbb{K}}(\mathcal{A}, \mathbb{K})$ of $\mathcal{A}$ is naturally identified with the subspace*

$$I^{\perp} = \{\Lambda \in \mathbb{K}[\mathbf{x}]^* \mid \forall p \in I,\ p \star \Lambda = 0\}. \tag{1.5}$$

called inverse system of I, with I the ideal of $\mathbb{K}[\mathbf{x}]$ such that $\mathcal{A} = \mathbb{K}[\mathbf{x}]/I$.

Remark 1.3.4 *As I is stable by multiplication by the variables x_i, the orthogonal $I^{\perp} = \mathcal{A}^*$ is stable by the derivations $\frac{d}{\mathrm{dz}_i}$.*

Proposition 1.3.5 *Let Q be a primary ideal for the maximal ideal $\mathfrak{m}_{\xi}$ of the point $\xi \in \mathbb{K}^n$ and let $\mathcal{A}_{\xi} = \mathbb{K}[\mathbf{x}]/Q$. Then there exists a vector space $D \subset \mathbb{K}[\boldsymbol{z}]$ stable by the derivations $\frac{d}{\mathrm{dz}_i}$ such that*

$$Q^{\perp} = \mathcal{A}_{\xi}^* = D \cdot \mathbf{1}_{\xi}(\boldsymbol{z}).$$

Theorem 1.3.6 *Let $\mathcal{A}$ be an artinian algebra of dimension r with $\mathcal{V}(I) = \{\xi_1, \ldots, \xi_{r'}\} \subset \mathbb{K}^n$. There exists vector spaces $D_i \subset \mathbb{K}[\mathbf{z}]$ stable by derivation of dimension μ_i with $\sum_{i=1}^{r'} \mu_i = r$, such that the elements of $\mathcal{A}^*$ are the elements $\Lambda \in \mathbb{K}[[\mathbf{z}]]$ of the form*

$$\Lambda(\mathbf{z}) = \sum_{i=1}^{r'} \omega_i(\mathbf{z}) \mathbf{1}_{\xi_i}(\mathbf{z}),$$

with $\omega_i(\mathbf{z}) \in D_i$.

Definition 1.3.7 *Let g be a polynomial in $\mathcal{A}$. The g-multiplication operator M_g is defined by*

$$\begin{array}{rccc} M_g : & \mathcal{A} & \longrightarrow & \mathcal{A} \\ & h & \longmapsto & M_g(h) = gh \end{array} \tag{1.6}$$

The transpose application M_g^T of the g-multiplication operator M_g is defined by

$$\begin{array}{rccc} M_g^T : & \mathcal{A}^* & \longrightarrow & \mathcal{A}^* \\ & \Lambda & \longmapsto & M_g^T(\Lambda) = g \cdot \Lambda \end{array} \tag{1.7}$$

Let B be a monomial basis in $\mathcal{A}$ and B^* its dual basis in $\mathcal{A}^*$. As the matrix M_g^T of the transpose application M_g^T in the dual basis B^* in $\mathcal{A}^*$ is the transpose of the matrix of the application M_g in the basis B in $\mathcal{A}$, the eigenvalues are the same for both matrices.

The main property (see [Elkadi 2007]) that we need is the following

Proposition 1.3.8 *Let I be an ideal of $\mathbb{K}[\mathbf{x}]$ and suppose that $\mathcal{V}(I) = \{\xi_1, ..., \xi_r\}$. Then*

- *for all $g \in \mathcal{A}$, the eigenvalues of M_g and M_g^T are the evaluations of the polynomial g at the roots, namely $g(\xi_1), ..., g(\xi_r)$.*
- *the eigenvectors common to all M_g^T with $g \in \mathcal{A}$ are -up to scalar - the evaluations $\mathbf{1}_{\xi_1}, ..., \mathbf{1}_{\xi_r}$*

1.4 Artinian Gorenstein algebra and positive linears forms

In this Section, we analyze the properties of an artinian algebra obtained as a quotient by the kernel of an Hankel operator H_Λ. It is obvious I_Λ defined in the Section before is an ideal of $\mathbb{K}[\mathbf{x}]$. We construct the quotient algebra

$\mathcal{A}_\Lambda = \mathbb{K}[\mathbf{x}]/I_\Lambda$. By construction, $\mathcal{A}_\Lambda^* = I_\Lambda^\perp$ contains the element $p \star \Lambda$ and $\operatorname{Im} H_\Lambda \subset \mathcal{A}_\Lambda^*$. The Hankel operator H_Λ is a map from $\mathbb{K}[\mathbf{x}]$ into $\mathcal{A}^*$:

$$0 \to I_\Lambda \to \mathbb{K}[\mathbf{x}] \xrightarrow{H_\Lambda} \mathcal{A}_\Lambda^* \tag{1.8}$$

The variety defined by I_Λ in $\mathbb{K}^n$ is denoted hereafter $\mathcal{V}_\mathbb{K}(I_\Lambda)$ or simply $\mathcal{V}(I_\Lambda)$ when $\mathbb{K}$ is algebraically closed.

If $\Lambda(\boldsymbol{z}) = \sum_{i=1}^r \omega_i(\boldsymbol{z})\mathbf{1}_{\xi_i}(\boldsymbol{z})$ then, by Lemma 1.2.5, the kernel I_σ is the set of polynomials $p \in \mathbb{K}[\mathbf{x}]$ such that $\forall q \in \mathbb{K}[\mathbf{x}]$, p is a solution of the following partial differential equation:

$$\sum_{i=1}^r \omega_i(\partial)(pq)(\xi_i) = 0.$$

Since $\forall p(\mathbf{x}), q(\mathbf{x}) \in \mathbb{K}[\mathbf{x}]$, $\langle p(\mathbf{x}) + I_\Lambda, q(\mathbf{x}) + I_\Lambda \rangle_\Lambda = \langle p(\mathbf{x}), q(\mathbf{x}) \rangle_\Lambda$, $\langle ., . \rangle_\Lambda$ induces an inner product on $\mathcal{A}_\Lambda$.

Theorem 1.4.1 *Let $\Lambda \in \mathbb{K}[\mathbf{x}]^* = \mathbb{K}[[z]] \setminus 0$*

- $\operatorname{rank} H_\Lambda = dim_\mathbb{K}(\mathcal{A}_\Lambda) < \infty$*, if and only if,*

$$\Lambda(z) = \sum_{i=1}^{r'} w_i(z)\mathbf{1}_{\xi_1}(z) \tag{1.9}$$

  with $w_i(z) \in \mathbb{K}[z] \setminus 0$ and $\xi_i \in \mathbb{K}^n$ pairwise distinct.

- *If $\Lambda(z) = \sum_{i=1}^{r'} w_i(z)\mathbf{1}_{\xi_1}(z)$ with $w_i(z) \in \mathbb{K}[z] \setminus 0$, then*
  - *the map $\mathcal{H}_\Lambda : \mathcal{A} \to \mathcal{A}^*$ induced by H_Λ is an isomorphism.*
  - *the inner product $\langle \cdot, \cdot \rangle_\Lambda$ si non -degenerate on $\mathcal{A} = \mathbb{K}[\mathbf{x}]/I_\Lambda$*
  - *the rank of H_Λ is $\sum_{i=1}^{r'} \mu_i$ where μ_i is the dimension of the vector space spanned by $w_i(z)$ and all its derivatives $\partial_{z_1}^{\alpha_1} \cdots \partial_{z_n}^{\alpha_n} w_i(z)$ for $\alpha = (\alpha_1, ..., \alpha_n) \in \mathbb{N}^n$.*
  - *the variety $\mathcal{V}(I_\Lambda)$ is the set of points $\xi_1, ..., \xi_{r'} \in \mathbb{K}^n$, with multiplicity $\mu_1, ..., \mu_{r'}$*

Proof. By definition of I_Λ and by short exact sequence

$$0 \to I_\Lambda \to \mathbb{K}[\mathbf{x}] \xrightarrow{H_\Lambda} \mathcal{A}_\Lambda^* \tag{1.10}$$

we have $\mathcal{A} = \mathbb{K}[\mathbf{x}]/I_\Lambda \sim Im(H_\Lambda)$. If rank $H_\Lambda = dim\ Im(H_\Lambda) = r < \infty$, then $dim(\mathcal{A}) = dim(\mathbb{K}[\mathbf{x}]/I_\Lambda)$ and $\mathcal{A}$ is artinian algebra (of dimension r over $\mathbb{K}$). By Theorem 1.3.2, it can be decomposed as a direct sum of sub-algebras

$$\mathcal{A}_\Lambda = \mathcal{A}_{\xi_1} \oplus \cdots \oplus \mathcal{A}_{\xi_{r'}}$$

where $\mathcal{V}_\mathbb{K}(I_\Lambda) = \{\xi_1, \ldots, \xi_{r'}\}$ and $\mathcal{A}_{\xi_i}$ is a local algebra for the maximal ideal $\boldsymbol{m}_{\xi_i}$ defining the root $\xi_i \in \mathbb{K}^n$: $\mathcal{A}_{\xi_i} = \mathbb{K}[\mathbf{x}]/Q_i$ with Q_i an $\boldsymbol{m}_{\xi_i}$-primary ideal of $\mathbb{K}[\mathbf{x}]$. Moreover, we have the minimal primary decomposition $I_\Lambda = Q_1 \cap \cdots \cap Q_r$.

The series $\Lambda(\mathbf{z})$ represent an element of the dual $\mathcal{A}_\Lambda^* = I_\Lambda^\perp$, which by Theorem 1.3.6 can be decomposed as

$$\Lambda(\mathbf{z}) = \sum_{i=1}^{r'} \omega_i(\mathbf{z}) \mathbf{1}_{\xi_i}(\mathbf{z}) \tag{1.11}$$

with $\omega_i(\mathbf{z}) \in \mathbb{C}[\mathbf{z}]$. The polynomial $\omega_i(\mathbf{z})$ cannot be zero, otherwise $Q_i \subset \ker H_\Lambda = I_\Lambda$. As $I_\Lambda = Q_1 \cap \cdots \cap Q_r$, we deduce that $I_\Lambda = Q_i$ and that $\Lambda(\mathbf{z}) = \omega_i(\mathbf{z})\mathbf{1}_{\xi_i}(\mathbf{z}) = 0$, which contradicts the hypothesis.

Conversely, if $\Lambda(\mathbf{z}) = \sum_{i=1}^r \omega_i(\mathbf{z})\mathbf{1}_{\xi_i}(\mathbf{z})$ with $\omega_i(\mathbf{z}) \in \mathbb{K}[\mathbf{z}] \setminus \{0\}$ and $\xi_i \in \mathbb{K}^n$ pairwise distinct, we easily check that I_Λ contains $\cap_{i=1}^r \boldsymbol{m}_{\xi_i}^{d_{i+1}}$ where d_i is the degree of $\omega_i(\mathbf{z})$. Thus $\mathcal{V}(I_\Lambda) \subset \{\xi_1, \ldots, \xi_r\}$.

The ideal I_Λ contains in particular univariate polynomials in each variable x_i. Thus $\mathcal{A}_\Lambda = \mathbb{K}[\mathbf{x}]/I_\Lambda$ is of finite dimension over $\mathbb{K}$ and rank $H_\Lambda < \infty$.

Let us assume now that $\Lambda(\mathbf{z}) = \sum_{i=1}^{r'} \omega_i(\mathbf{z})\mathbf{1}_{\xi_i}(\mathbf{z})$ with $\omega_i(\mathbf{z}) \in \mathbb{K}[\mathbf{z}] \setminus \{0\}$ so that $\mathcal{A}_\Lambda = \mathbb{K}[\mathbf{x}]/I_\Lambda$ is of dimension r over $\mathbb{K}$.

As $\mathcal{A}_\Lambda = \mathbb{K}[\mathbf{x}]/I_\Lambda \sim \mathrm{Im}(H_\Lambda)$, H_Λ induces an injection from $\mathcal{A}_\Lambda$ into $\mathcal{A}_\Lambda^*$ which is of dimension r. We deduce that H_Λ induces an isomorphism between $\mathcal{A}_\Lambda$ and $\mathcal{A}_\Lambda^*$, and we have the short exact sequence:

$$0 \to I_\Lambda \to \mathbb{K}[\mathbf{x}] \xrightarrow{H_\Lambda} \mathcal{A}_\Lambda^* \to 0.$$

This shows that $\mathcal{A}_\Lambda^*$ is generated by elements $p \star \Lambda$ for $p \in \mathbb{K}[\mathbf{x}]$, that is, $\mathcal{A}_\Lambda^*$ is the inverse system generated by Λ.

By definition of I_Λ, if $p \in \mathbb{K}[\mathbf{x}]$ is such that $\forall q \in \mathbb{K}[\mathbf{x}]$

$$\langle p(\mathbf{x}), q(\mathbf{x}) \rangle_\Lambda = \langle p \star \Lambda(\mathbf{z}) | q(\mathbf{x}) \rangle = 0,$$

then $p \star \Lambda(\mathbf{z}) = 0$ and $p \in I_\Lambda$. We deduce that the inner product $\langle \cdot, \cdot \rangle_\Lambda$ is non-generate on $\mathcal{A}_\Lambda = \mathbb{K}[\mathbf{x}]/I_\Lambda$.

By Theorem 1.3.6, $\Lambda \in \mathcal{A}_\Lambda^*$ has a decomposition of the form (1.9) which must coincide with the given one: $\Lambda(\mathbf{z}) = \sum_{i=1}^{r'} \omega_i(\mathbf{z})\mathbf{1}_{\xi_i}(\mathbf{z})$. Thus $\mathcal{A}_\Lambda^* =$

$\mathcal{A}^*_{\xi_1} \oplus \cdots \oplus \mathcal{A}^*_{\xi_{r'}}$ where $I_\Lambda = Q_1 \cap \cdots \cap Q_{r'}$ and $\mathcal{A}^*_{\xi_i} = Q_i^\perp$ is the inverse system generated by $\omega_i(\mathbf{z})\mathbf{1}_{\xi_i}(\mathbf{z})$ for $i = 1, \ldots, r'$.

The dimension of $\mu_i = \dim \mathcal{A}^*_{\xi_i} = \dim \mathcal{A}_{\xi_i}$ of the inverse system $\mathcal{A}^*_{\xi_i}$ is the multiplicity of ξ_i; it is also the dimension of the vector space spanned by $\omega_i(\mathbf{z})$ and all its derivatives $\partial^{\alpha_1}_{z_1} \cdots \partial^{\alpha_n}_{z_n} \omega_i(\mathbf{z})$ for $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$. We deduce that $\dim \mathcal{A}_\Lambda = \dim \mathcal{A}^*_\Lambda = r = \sum_{i=1}^{r'} \mu_i$.

As $I_\Lambda = Q_1 \cap \cdots \cap Q_{r'}$, we deduce that $\mathcal{V}(I_\Lambda) = \{\xi_1, \ldots, \xi_{r'}\}$, which concludes the proof of this theorem. ∎

Remark 1.4.2 *An algebra $\mathcal{A}$ is called Gorenstein if $\mathcal{A}$ and its dual $\mathcal{A}^*$ are isomorphic $\mathcal{A}$-modules. Then the quotient space $\mathcal{A}_\Lambda = K[\mathbf{x}]/\ker H_\Lambda$ is a Gorenstein algebra*

A special case of interest is when the roots are simple. We characterize it as follows:

Proposition 1.4.3 *Let $\Lambda \in \mathbb{K}[\mathbf{x}]^*$. The following conditions are equivalent:*

1. $\Lambda = \sum_{i=1}^r \omega_i \mathbf{1}_{\xi_i}$, *with $\omega_i \in \mathbb{K} \setminus \{0\}$ and $\xi_i \in \mathbb{K}^n$ pairwise distinct.*
2. *The rank of H_Λ is r and the multiplicity of the points $\xi_1, \ldots, \xi_r$ in $\mathcal{V}(I_\Lambda)$ is 1.*
3. *A basis of $\mathcal{A}^*_\Lambda$ is $\mathbf{1}_{\xi_1}, \ldots, \mathbf{1}_{\xi_r}$.*

Proof. $1 \Rightarrow 2$. The dimension of the vector space spanned by $\omega_i \in \mathbb{K} \setminus \{0\}$ and its derivatives is 1. By Theorem 1.4.1, the rank $\mathcal{A}_\Lambda$ is $r = \sum_{i=1}^r 1$ and the multiplicity of the roots $\xi_1, \ldots, \xi_r$ in $\mathcal{V}(I_\Lambda)$ is 1.

$2 \Rightarrow 3$. By Theorem 1.4.1, $\mathcal{A}^*_\Lambda$ is the inverse system spanned by Λ. As $\forall p \in \mathbb{K}[\mathbf{x}]$, $p \star \Lambda = \sum_{i=1}^r \omega_i p(\xi_i)\mathbf{1}_{\xi_i}$, $\mathcal{A}^*_\Lambda$ is in the vector space spanned by $\mathbf{1}_{\xi_1}, \ldots, \mathbf{1}_{\xi_r}$. As $\dim(\mathcal{A}^*_\Lambda) = r$, it is a basis.

$3 \Rightarrow 1$. As $\Lambda \in \mathcal{A}^*_\Lambda$, there exists $\omega_i \in \mathbb{K}$ such that $\Lambda = \sum_{i=1}^r \omega_i \mathbf{1}_{\xi_i}$. If one of these coefficients ω_i vanishes that $\dim(\mathcal{A}^*_\Lambda) < r$, which is contradicting point 3. Thus $\omega_i \in \mathbb{K} \setminus \{0\}$. ∎

In the case where all the coefficients of Λ are in $\mathbb{R}$, we can consider the following notion of positivity:

Definition 1.4.4 *An element $\Lambda \in \mathbb{R}[\mathbf{x}]^*$ is positive if $\forall p \in \mathbb{R}[\mathbf{x}]$, $\langle p, p\rangle = \langle \Lambda \mid p^2\rangle = \Lambda(p^2) \geqslant 0$. It is denoted $\Lambda \succcurlyeq 0$.*

The positivity of Λ induces the following property on its decomposition that we can find also in [Lasserre 2012],:

Proposition 1.4.5 *Let $\Lambda \in \mathbb{R}[\mathbf{x}]^*$ such that $rank H_\Lambda = r$, $\Lambda \succcurlyeq 0$ iff*

$$\Lambda = \sum_{i=1}^{r} \omega_i \, \mathbf{1}_{\xi_i}$$

with $\omega_i > 0$, $\xi_i \in \mathbb{R}^n$, for $i = 1, .., r$.

Proof. If $\Lambda = \sum_{i=1}^{r} \omega_i \, \mathbf{1}_{\xi_i}$ with $\omega_i > 0$, $\xi_i \in \mathbb{R}^n$, for $i = 1, .., r$, then clearly $\forall p \in \mathbb{R}[\mathbf{x}]$,

$$\langle \Lambda \mid p^2 \rangle = \Lambda(p^2) = \sum_{i=1}^{r} \omega_i \, p^2(\xi_i) \geqslant 0$$

and $\Lambda \succcurlyeq 0$.

Conversely suppose that $\forall p \in \mathbb{R}[\mathbf{x}]$, $\langle \Lambda \mid p^2 \rangle = \Lambda(p^2) \geqslant 0$. Then $p \in I_\Lambda$ iff $\langle \Lambda \mid p^2 \rangle = \Lambda(p^2) = 0$. We check that I_Λ is real radical: If $p^{2k} + \sum_j q_j^2 \in I_\Lambda$ for some $k \in \mathbb{N}$, $p, q_j \in \mathbb{R}[\mathbf{x}]$ then

$$\langle \Lambda \mid p^{2k} + \sum_j q_j^2 \rangle = \Lambda(p^{2k} + \sum_j q_j^2) = \Lambda(p^{2k}\rangle + \sum_j \Lambda(q_j^2) = 0$$

which implies that $\langle \Lambda \mid p^2 k \rangle = \Lambda(p^{2k}) = 0$, $\langle \Lambda \mid q_j^2 \rangle = \Lambda(q_j^2) = 0$ and $p^k, q_j \in I_\Lambda$.
Let $k' = \lceil \frac{k}{2} \rceil$. We have $\langle \Lambda \mid p^{2k'} \rangle = \Lambda(p^{2k'})\rangle = 0$, which implies that $p^{k'} \in I_\Lambda$. Iterating this reduction, we deduce that $p \in I_\Lambda$. This shows that I_Λ is real radical and $\mathcal{V}(I_\Lambda) \subset \mathbb{R}^n$. By Proposition 1.4.3, we deduce that $\Lambda = \sum_{i=1}^{r} \omega_i \, \mathbf{1}_{\xi_i}$ with $\omega_i \in \mathbb{C} \setminus \{0\}$ and $\xi_i \in \mathbb{R}^n$. Let $p_i \in \mathbb{R}[\mathbf{x}]$ be interpolation polynomials at $\xi_i \in \mathbb{R}^n$: $p_i(\xi_i) = 1$, $p_i(\xi_j) = 0$ for $j \neq i$. Then $\langle \Lambda \mid p_i^2 \rangle = \Lambda(p_i^2) = \omega_i \in \mathbb{R}_+$. This proves that $\Lambda = \sum_{i=1}^{r} \omega_i \, \mathbf{1}_{\xi_i}$ with $\omega_i > 0$, $\xi_i \in \mathbb{R}^n$, for $i = 1, .., n$. ∎

Chapter 2

Minimization problem and varieties of critical points

Let $f, g_1^0, \ldots, g_{n_1}^0, g_1^+, \ldots, g_{n_2}^+ \in \mathbb{R}[\mathbf{x}]$ be polynomials functions. The minimization problem that we consider all along the manuscrit is the following:

$$\begin{array}{ll} \inf\limits_{\mathbf{x}\in\mathbb{R}^n} & f(\mathbf{x}) \\ s.t. & g_1^0(\mathbf{x}) = \cdots = g_{n_1}^0(\mathbf{x}) = 0 \\ & g_1^+(\mathbf{x}) \geq 0, ..., g_{n_2}^+(\mathbf{x}) \geq 0 \end{array} \tag{2.1}$$

More precisely, the objectives of the method we describe are to compute the minimum value when f is bounded by below and the points where this minimum value is reached if they exists.

Hereafter, we fix the set of constraints

$$\mathbf{g} = \{\mathbf{g}^0, \mathbf{g}^+\} = \{g_1^0, \ldots, g_{n_1}^0; g_1^+, \ldots, g_{n_2}^+\} \tag{2.2}$$

and we denote by

$$S := \mathcal{S}(\mathbf{g}) = \{\mathbf{x} \in \mathbb{R}^n \mid g_1^0(\mathbf{x}) = 0, \ldots, g_{n_1}^0(\mathbf{x}) = 0; g_1^+(\mathbf{x}) \geq 0, \ldots, g_{n_2}^+(\mathbf{x}) \geq 0\} \tag{2.3}$$

the basic semi-algebraic set defining the points which satisfy the constraints of our minimization problem (2.1). And

$$S^+(\mathbf{g}) = \{\mathbf{x} \in \mathbb{R}^n \mid g_1^+(\mathbf{x}) \geq 0, \ldots, g_{n_2}^+(\mathbf{x}) \geq 0\} \tag{2.4}$$

is the basic semi-algebraic set defining the points which satisfy the nonnegative constraints of our minimization problem (2.1).

When $n_1 = n_2 = 0$, there is no constraint and $S = \mathbb{R}^n$. In this case, we are considering a global minimization problem.

The points $\mathbf{x}^* \in S$ which satisfy $f(\mathbf{x}^*) = \inf_{\mathbf{x}\in S} f(\mathbf{x})$ are called the *minimizers* of f on S. If the set of minimizers is not empty, we say that the *minimization problem is feasible*.

Before describing how to compute the minimizer points, we analyse the geometry of this minimization problem and the varieties associated to its

critical points. In the following, we denote by $\mathbf{y} = (\mathbf{x}, \mathbf{u}, \mathbf{v})$ and $\mathbf{z} = (\mathbf{x}, \mathbf{u}, \mathbf{v}, \mathbf{s})$, the $n + n_1 + n_2$ and $n + n_1 + 2n_2$ variables of these problems. For any ideal $J \subset \mathbb{R}[\mathbf{z}]$, we denote $J^{\mathbf{x}} = J \cap \mathbb{R}[\mathbf{x}]$. The projection of $\mathbb{C}^n \times \mathbb{C}^{n_1+2\,n_2}$ (resp. $\mathbb{C}^n \times \mathbb{C}^{n_1+\,n_2}$) on $\mathbb{C}^n$ is denoted $\pi^{\mathbf{x}}$.

2.1 The gradient variety

A natural approach to deal with constraints in optimization problems is to introduce Lagrangian multipliers. Replacing the inequalities $g_i^+ \geq 0$ by the equalities $g_i^+ - s_i^2 = 0$ (adding new variables s_i) and introducing new parameters for all the equality constraints yields the following minimization problem:

$$\begin{array}{ll} \inf\limits_{(\mathbf{x},\mathbf{u},\mathbf{v},\mathbf{s})\in\mathbb{R}^n\times\mathbb{R}^{n_1+2\,n_2}} & f(\mathbf{x}) \\ s.t. & \nabla F(\mathbf{x},\mathbf{u},\mathbf{v},\mathbf{s}) = 0 \end{array} \tag{2.5}$$

where $F(\mathbf{x}, \mathbf{u}, \mathbf{v}, \mathbf{s}) = f(\mathbf{x}) - \sum_{i=1}^{n_1} u_i g_i^0(\mathbf{x}) - \sum_{j=1}^{n_2} v_j (g_j^+(\mathbf{x}) - s_j^2)$,
$\mathbf{u} = (u_1, ..., u_{n_1})$,
$\mathbf{v} = (v_1, ..., v_{n_2})$ and $\mathbf{s} = (s_1, ..., s_{n_2})$.

Definition 2.1.1 *The gradient ideal of $F(\mathbf{z})$ is:*

$$I_{grad} = (\nabla F(\mathbf{z})) = (F_1, ..., F_n, g_1^0, ..., g_{n_1}^0, g_1^+ - s_1^2, ..., g_{n_2}^+ - s_{n_2}^2, v_1 s_1, ..., v_{n_2} s_{n_2}) \subset \mathbb{R}[\mathbf{z}]$$

where $F_i = \frac{\partial f}{\partial x_i} - \sum_{j=1}^{n_1} u_j \frac{\partial g_j^0}{\partial x_i} - \sum_{j=1}^{n_2} v_j \frac{\partial g_j^+}{\partial x_i}$.
The gradient variety is $V_{grad} := \mathcal{V}(I_{grad})$ and the real gradient variety is $V_{grad}^{\mathbb{R}} := V_{grad} \cap (\mathbb{R}^n \times \mathbb{R}^{n_1+2n_2})$.
Its projection on $\mathbf{x}$ is $V_{grad}^{\mathbf{x}} := \overline{\pi^{\mathbf{x}}(V_{grad})}$, where $\pi^{\mathbf{x}}$ is the projection of $\mathbb{C}^n \times \mathbb{C}^{n_1+2n_2}$ onto $\mathbb{C}^n$.

Example 2.1.2

$$\begin{array}{ll} \inf\limits_{(x_1,x_2)\in\mathbb{R}^2} & -12x_1 - 7x_2 + x_2^2; \\ s.t. & 2x_1^4 - 2 + x_2 = 0 \\ & -x_1 + 3 \geq 0, -x_2 + 2 \geq 0, x_1 \geq 0, x_2 \geq 0 \end{array}$$

His gradient ideal is $I_{grad} = (\nabla F(\mathbf{z})) = (-12 + 8u_1x_1^3 - v_1 + v_3, -7 + 2x_2 + u_1 - v_2 + v_4, 2x_1^4 - 2 + x_2, -x_1 + 3 - s_1^2, -x_2 + 2 - s_2^2, x_1 - s_3^2, x_2 - s_4^2, v_1s_1, v_2s_2, v_3s_3, v_4s_4)$

Definition 2.1.3 *For any $F \in \mathbb{R}[\mathbf{z}]$, the values of F at the (resp. real) points of $\mathcal{V}(\nabla F) = V_{grad}$ are called the (resp. real) critical values of F.*

We easily check the following property:

Lemma 2.1.4 $F\mid_{V_{grad}} = f\mid_{V_{grad}}$.

Thus minimizing f on V_{grad} is the same as minimizing F on V_{grad}, that is computing the minimal critical value of F.

2.2 The Karush-Kuhn-Tucker variety

A variant of the gradient variety that we can use in constrained problems is the well-known Karush-Kuhn-Tucker (KKT) variety which have been used in several approches about polynomial optimization (see [Demmel 2007, Ha 2010, Nie 2011].

Definition 2.2.1 *A point* $\mathbf{x}^*$ *is called a KKT point if there exists* $u_1, \ldots, u_{n_1}$, $v_1, \ldots, v_{n_2} \in \mathbb{R}$ *s.t.*

$$\nabla f(\mathbf{x}^*) - \sum_{i=1}^{n_1} u_i \nabla g_i^0(\mathbf{x}^*) - \sum_{j=0}^{n_2} v_j \nabla g_j^+(\mathbf{x}^*) = 0,\ g_i^0(\mathbf{x}^*) = 0,\ v_j g_j^+(\mathbf{x}^*) = 0.$$

The corresponding minimization problem is the following:

$$\begin{aligned} \inf_{(\mathbf{x},\mathbf{u},\mathbf{v})\in\mathbb{R}^{n+n_1n_2}} \quad & f(\mathbf{x}) \\ s.t. \quad & F_1 = \cdots = F_n = 0 \\ & g_1^0 = \cdots = g_{n_1}^0 = 0 \\ & v_1\, g_1^+ = \cdots = v_{n_2}\, g_{n_2}^+ = 0 \\ & g_1^+ \geq 0, \ldots, g_{n_2}^+ \geq 0 \end{aligned} \tag{2.6}$$

where $F_i = \frac{\partial f}{\partial x_i} - \sum_{j=1}^{n_1} u_j \frac{\partial g_j^0}{\partial x_i} - \sum_{j=1}^{n_2} v_j \frac{\partial g_j^+}{\partial x_i}$.

This leads to the following definitions:

Definition 2.2.2 *The Karush-Kuhn-Tucker (KKT) ideal associated to Problem (2.1) is*

$$I_{KKT} = (F_1, ..., F_n, g_1^0, ..., g_{n_1}^0, v_1 g_1^+, ..., v_{n_2} g_{n_2}^+) \subset \mathbb{R}[\mathbf{y}]. \tag{2.7}$$

The KKT variety is $V_{KKT} := \mathcal{V}(I_{KKT}) \subset \mathbb{C}^n \times \mathbb{C}^{n_1+n_2}$ *and the real KKT variety is* $V_{KKT}^{\mathbb{R}} := V_{KKT} \cap (\mathbb{R}^n \times \mathbb{R}^{n_1+n_2})$.
Its projection on $\mathbf{x}$ *is* $V_{KKT}^{\mathbf{x}} := \overline{\pi^{\mathbf{x}}(V_{KKT})}$, *where* $\pi^{\mathbf{x}}$ *is the projection of* $\mathbb{C}^n \times \mathbb{C}^{n_1+n_2}$ *onto* $\mathbb{C}^n$.

The set of KKT points of S is denoted S_{KKT} and a KKT-minimizer of f on S is a point $\mathbf{x}^* \in S_{KKT}$ such that $f(\mathbf{x}^*) = \min_{\mathbf{x}\in S_{KKT}} f(\mathbf{x})$.

Notice that $V_{KKT}^{\mathbf{x},\mathbb{R}} = \overline{\pi^{\mathbf{x}}(V_{KKT})}^{\mathbb{R}} = \overline{\pi^{\mathbf{x}}(V_{KKT}^{\mathbb{R}})}$, since any linear dependency relation between real vectors can be realized with real coefficients.

Example 2.2.3 *The KKT ideal asociated to the optimization problem 2.1.2 is*
$I_{KKT} = (-12 + 8u_1x_1^3 - v_1 + v_3, -7 + 2x_2 + u_1 - v_2 + v_4, 2x_1^4 - 2 + x_2, v_1(-x_1 + 3), v_2(-x_2 + 2), v_3x_1, v_4x_2)$

The KKT ideal is related to the gradient ideal as follows:

Proposition 2.2.4 $I_{KKT} = I_{grad} \cap \mathbb{R}[\mathbf{y}]$.

Proof. As $s_i(s_iv_i) + v_i(g_i^+ - s_i^2) = v_ig_i^+ \ \forall i = 1, ..., n_2$, we have $I_{KKT} \subset I_{grad} \cap \mathbb{R}[\mathbf{y}]$.

In order to prove the equality, we use the property that if K is a Groebner basis of I_{grad} for an elimination ordering such that $\mathbf{s} \gg \mathbf{x}, \mathbf{u}, \mathbf{v}$ then $K \cap \mathbb{R}[\mathbf{y}]$ is the Groebner basis of $I_{grad} \cap \mathbb{R}[\mathbf{y}]$ (see [Cox 2005]). Notice that $s_i(s_iv_i) + v_i(g_i^+ - s_i^2) = v_ig_i^+$ $(i = 1, ..., n_2)$ are the only S-polynomials involving the variables $s_1, \ldots, s_{n_2}$ which may have a non-trivial reduction. Thus $K \cap \mathbb{R}[\mathbf{y}]$ is also the Groebner basis of $F_1, ..., F_n, g_1^0, ..., g_{n_1}^0, v_1g_1^+, ..., v_{n_2}g_{n_2}^+$ and we have $(K) \cap \mathbb{R}[\mathbf{y}] = I_{grad} \cap \mathbb{R}[\mathbf{y}] = I_{KKT}$. ■

The KKT points on S are related to the real points of the gradient variety as follows:

Lemma 2.2.5 $S_{KKT} = V_{grad}^{\mathbf{x},\mathbb{R}} = V_{KKT}^{\mathbf{x},\mathbb{R}} \cap \mathcal{S}^+(\mathbf{g})$.

Proof. A real point $\mathbf{y} = (\mathbf{x}, \mathbf{u}, \mathbf{v})$ of $V_{KKT}^{\mathbb{R}}$ lifts to a point $\mathbf{z} = (\mathbf{x}, \mathbf{u}, \mathbf{v}, \mathbf{s})$ in $V_{grad}^{\mathbb{R}}$, if and only if, $g_i^+(\mathbf{x}) \geq 0$ for $i = 1, \ldots, n_2$. This implies that $V_{grad}^{\mathbf{y},\mathbb{R}} = V_{KKT}^{\mathbb{R}} \cap \mathcal{S}^+(\mathbf{g})$, which gives by closed projection the equalities $S_{KKT} = V_{KKT}^{\mathbf{x},\mathbb{R}} \cap \mathcal{S}^+(\mathbf{g}) = V_{grad}^{\mathbf{x},\mathbb{R}}$ since a point $\mathbf{x}^*$ of $V_{grad}^{\mathbf{x},\mathbb{R}}$ satisfies $\mathbf{g}_j^+(\mathbf{x}^*) \geq 0$ for $j \in [1, n_2]$. ■

Remark 2.2.6 *This shows that if a minimizer point of f on S is a KKT point, then it is the projection of a real critical point of F.*

2.3 The Fritz John variety

A minimizer of f on S is not necessarily a KKT point as show in the example 2.4.1

$$\begin{array}{ll} \min & x \\ s.t & x^3 \geq 0 \end{array}$$

The KKT ideal is $I_{KKT} = (1 - 3v_1x^2, v_1x^3) = (1)$. As $\mathcal{V}(1) = \emptyset$, there is not KKT minimizers but the minimizer is 0.

In order to solve this kind of problems, more general conditions that are satisfied by minimizers were given by F. John for polynomial non-negativity constraints and further refined for general polynomial constraints [John 1948, Mangasarian 1967].

To describe these conditions, we introduce a new variable u_0 and denote by $\mathbf{y}'$ the set of variables $\mathbf{y}' = (\mathbf{x}, u_0, \mathbf{u}, \mathbf{v})$.
Let

$$F_i^{u_0} = u_0 \frac{\partial f}{\partial x_i} - \sum_{j=1}^{n_1} u_j \frac{\partial g_j^0}{\partial x_i} - \sum_{j=1}^{n_2} v_j \frac{\partial g_j^+}{\partial x_i}.$$

Definition 2.3.1 *For any $\gamma \subset [1, n_1]$, let*

$$I_{FJ}^{\gamma} = (F_1^{u_0}, ..., F_n^{u_0}, g_1^0, ..., g_{n_1}^0, v_1 g_1^+, ..., v_{n_2} g_{n_2}^+, u_i, i \notin \gamma) \subset \mathbb{R}[\mathbf{y}']. \tag{2.8}$$

For $m \in \mathbb{N}$, the m^{th} Fritz-John (FJ) ideal associated to Problem (2.1) is

$$I_{FJ}^m = \cap_{|\gamma|=m} I_{FJ}^{\gamma}. \tag{2.9}$$

Let $V_{FJ}^{\gamma} := \mathcal{V}(I_{FJ}^{\gamma}) \subset \mathbb{C}^n \times \mathbb{P}^{n_1+n_2}$.
The m^{th} FJ variety is $V_{FJ}^m := \mathcal{V}(I_{FJ}^m) = \cup_{|\gamma|=m} V_{FJ}^{\gamma}$, and the real FJ variety is $V_{FJ}^{m,\mathbb{R}} := V_{FJ}^m \cap \mathbb{R}^n \times \mathbb{RP}^{n_1+n_2}$.
Its projection on $\mathbf{x}$ is $V_{FJ}^{m,\mathbf{x}} := \overline{\pi^{\mathbf{x}}(V_{FJ}^m)}$.
When $m = \max_{\mathbf{x} \in S} \text{ rank}([\nabla g_1^0(\mathbf{x}), \ldots, \nabla g_{n_1}^0(\mathbf{x})])$, the m^{th} FJ variety is denoted V_{FJ}.

Proposition 2.3.2 *Any minimizer $\mathbf{x}^*$ of f on S is the projection of a real point of $V_{FJ}^{\mathbb{R}}$.*

Proof. The proof is similar to Theorem 4.3.2 of [Bazaraa 2006]. For any minimizer point $\mathbf{x}^*$ (if it exists), we consider a maximal set of linearly independent gradients $\nabla g_j^0(\mathbf{x}^*)$ for $j \in \gamma$ (with $|\gamma| \leq m$) and apply the same proof as [Bazaraa 2006][Theorem 4.3.2]. This shows that $\mathbf{x}^* \in V_{FJ}^{\gamma,\mathbb{R}} \subset V_{FJ}^{\mathbb{R}}$. ∎

Remark 2.3.3 *If $n_1 = 0$, the m^{th} Fritz-John (FJ) ideal is*

$$I_{FJ} = (F_1^{u_0}, ..., F_n^{u_0}, v_1 g_1^+, ..., v_{n_2} g_{n_2}^+) \subset \mathbb{R}[\mathbf{y}']. \tag{2.10}$$

where

$$F_i^{u_0} = u_0 \frac{\partial f}{\partial x_i} - \sum_{j=1}^{n_2} v_j \frac{\partial g_j^+}{\partial x_i}$$

Notice that this definition slightly differs from the classical one [John 1948, Lasserre 2009a, Mangasarian 1967], which does not provide any information when the gradient vectors $\nabla g_i^0(\mathbf{x}), i = 1 \dots n_1$ are linearly dependent on S.

Definition 2.3.4 *We denote by $V_{sing} = V_{FJ} \cap \mathcal{V}(u_0)$ the intersection of V_{FJ} with the hyperplane $u_0 = 0$.*

We easily check that the "affine part" of V_{FJ} corresponding to $u_0 \neq 0$ is the variety V_{KKT}. Thus, we have the decomposition

$$\boxed{V_{FJ} = V_{sing} \cup V_{KKT}} \tag{2.11}$$

Its projection on $\mathbb{C}^n$ decomposes as

$$\boxed{V_{FJ}^{\mathbf{x}} = V_{sing}^{\mathbf{x}} \cup V_{KKT}^{\mathbf{x}}} \tag{2.12}$$

Let us describe more precisely the projection $V_{FJ}^{\mathbf{x}}$ onto $\mathbb{C}^n$. For $\nu = \{j_1, \dots, j_k\} \subset [1, n2]$, we define

$A_\nu = [\nabla f(\mathbf{x}), \nabla g_1^0(\mathbf{x}), \dots, \nabla g_{n_1}^0(\mathbf{x}), \nabla g_{j_1}^+(\mathbf{x}), \dots, \nabla g_{j_k}^+(\mathbf{x})]$

$V_\nu = \{\mathbf{x} \in \mathbb{C}^n \mid g_i^0(\mathbf{x}) = 0, i = 1 \dots n_1, g_j^+(\mathbf{x}) = 0, j \in \nu, \text{ rank}(A_\nu) \leq m + |\nu|\}$.

Let $\Delta_1^\nu, \dots, \Delta_{l_\nu}^\nu$ be polynomials defining the variety $\{\mathbf{x} \in \mathbb{C}^n \mid g_j^+(\mathbf{x}) = 0, j \in \nu, \text{ rank}(A_\nu) \leq m + |\nu|\}$. If $n \geq m + |\nu|$, these polynomials can be defined as $g_j^+, j \in \nu$ and as linear combinations of $(m + |\nu| + 1)$-minors of the matrix A_ν, as described in [Bruns 1988, Nie 2011]. If $n < m + \mid \nu \mid$, we take $l_\nu = 0$, $\Delta_i^\nu = 0$.

Let Γ_{FJ} be the union of $\mathbf{g}^0$, and the set of polynomials

$$\phi_{\nu,i} := \Delta_i^\nu \prod_{j \notin \nu} g_j^+, \tag{2.13}$$

for $i = 1, \dots, l_\nu, \nu \subset [0, n_2]$.

Lemma 2.3.5 $V_{FJ}^{\mathbf{x}} = \cup_{\nu\subset[1,n_2]} V_\nu = \mathcal{V}(\Gamma_{FJ})$.

Proof. For any $\mathbf{x} \in \mathbb{C}^n$, let $\nu(\mathbf{x}) = \{j \in [1, n_2] \mid g_j^+(\mathbf{x}) = 0\}$.

Let $\mathbf{y}'$ be a point of V_{FJ}, $\mathbf{x}$ its projection on $\mathbb{C}^n$ and $\nu(\mathbf{x}) = \nu = \{j_1, \ldots, j_k\}$. We have $g_j^+(\mathbf{x}) \neq 0$, $v_j = 0$ for $j \notin \nu$ and $\Delta_i^\nu = 0$ for $i = 1, \ldots, l_\nu$. This implies that $\text{rank}(A_\nu(\mathbf{x})) \leq m + |\nu|$ and there exists $(u_0, u_1, \ldots, u_{n_1}, v_1, \ldots, v_{n_2}) \neq 0$ and $\gamma \subset [1, n_1]$ of size $|\gamma| \leq m$ such that

$$u_0\nabla f + u_1\nabla g_1^0 + \cdots + u_{n_1}\nabla g_{n_1}^0 + v_1\nabla g_{j_1}^+ + \cdots + v_{n_2}\nabla g_{j_k}^+ = 0,$$

with $u_i = 0$, $i \notin \gamma \subset [1, n_1]$. Therefore $\mathbf{x} \in \pi^{\mathbf{x}}(V_{FJ})$, which proves that $\mathcal{V}(\mathbf{g}^0, \phi_{\nu,i}, \nu \subset [0, n_2], i = 1 \ldots l_\nu) \subset \pi^{\mathbf{x}}(V_{FJ})$.

Conversely, if $\mathbf{x} \in \pi^{\mathbf{x}}(V_{FJ})$ then $\mathbf{x} \in V_{\nu(\mathbf{x})} \subset \cup_\nu V_\nu$ which is defined by the polynomials $g_1^0, \ldots, g_{n_1}^0$ and $\phi_{\nu,i} := \Delta_i^\nu \prod_{j\notin\nu} g_j^+$, for $i = 1, \ldots, l_\nu, \nu \subset [0, n_2]$. ■

Remark 2.3.6 *The real variety $\pi^{\mathbf{x}}(V_{FJ}^{\mathbb{R}}) = V_{FJ}^{\mathbf{x}} \cap \mathbb{R}^n$ can also be defined by $\mathbf{g}^0$ and the set Φ_{FJ} of polynomials*

$$\varrho_\nu := \Delta^\nu \prod_{j\notin\nu} g_j^+ \text{ where } \Delta^\nu = \det(A_\nu A_\nu^T), \tag{2.14}$$

for $\nu \subset [1, n_2]$ and $n \geq m + \mid \nu \mid$, as described in [Ha 2010].

Similarly the projection $V_{sing}^{\mathbf{x}}$ onto $\mathbb{C}^n$ can be described as follows. For $\nu = \{j_1, \ldots, j_k\} \subset [1, n_2]$,

$$B_\nu = [\nabla g_1^0(\mathbf{x}), \ldots, \nabla g_{n_1}^0(\mathbf{x}), \nabla g_{j_1}^+(\mathbf{x}), \ldots, \nabla g_{j_k}^+(\mathbf{x})]$$

$W_\nu = \{\mathbf{x} \in \mathbb{C}^n \mid g_i^0(\mathbf{x}) = 0, i = 1 \ldots n_1, g_j^+(\mathbf{x}) = 0, j \in \nu, \text{ rank}(B_\nu) \leq m + |\nu| - 1\}$. Let $\Theta_1^\nu, \ldots, \Theta_{l_\nu}^\nu$ be polynomials defining the variety $\{\mathbf{x} \in \mathbb{C}^n \mid$

$g_j^+(\mathbf{x}) = 0, j \in \nu, \text{rank}(B_\nu) \leq m + |\nu| - 1\}$ and let Γ_{sing} be the union of $\mathbf{g}^0$ and the set of polynomials

$$\sigma_{\nu,i} := \Theta_i^\nu \prod_{j\notin\nu} g_j^+, \tag{2.15}$$

for $\nu \subset [1, n_2], i = 1 \ldots l_\nu$.

With similar arguments, we prove the following

Lemma 2.3.7 $V_{sing}^{\mathbf{x}} = \cup_{\nu\subset[1,n_2]} W_\nu = \mathcal{V}(\Gamma_{sing})$.

If we come back to the example in the begining of the section

Example 2.3.8

$$\begin{array}{ll} \min & x \\ s.t & x^3 \geq 0 \end{array}$$

The ideal de Fritz-John is :

$$I_{FJ} = (u_0 - v_1 3x^2, v_1 x^3)$$

- *If* $u_0 \neq 0$ *then* I_{FJ} *correspond to the ideal* $I_{KKT} = (1 - \frac{v_1}{u_0}3x^2, v_1 x^3)$ *and* $V_{KKT} = \emptyset$.
- *If* $u_0 = 0$ *then* $I_{FJ} \cap (u_0) = (-v_1 3x^2, v_1 x^3, u_0)$, $V_{sing} = \{(0, (0:1))\}$.

Then

$$V_{FJ} = V_{sing} \cup V_{KKT} = V_{sing}$$

Its projection on $\mathbb{C}^n$ *is equal to*

$$V_{FJ}^{\mathbf{x}} = V_{sing}^{\mathbf{x}} \cup V_{KKT}^{\mathbf{x}}$$

To compute $V_{FJ}^{\mathbf{x}}$ *we apply Lemma 2.3.5:*
There is not equalities so $m = 0$.
For $\nu = 0$, $A_0 = [1]$, $V_0 = \{\mathbf{x} \in \mathbb{C} \mid rank(A_0) \leq 0\} = \emptyset$ *then* $\Delta_1^0 = 1$, $\phi_{0,1} = 1 \cdot x^3$. *For* $\nu = 1$, $A_1 = [1, -3x^2]$, $V_1 = \{\mathbf{x} \in \mathbb{C} \mid x^3 = 0, rank(A_1) \leq 1\} = \{0\}$ *then* $\Delta_1^1 = x^3$, $\phi_{1,1} = x^3$, *and* $\Gamma_{FJ} = (x^3)$.
Then $V_{FJ}^{\mathbf{x}} = V_0 \cup V_1 = \mathcal{V}(\Gamma_{FJ}) = \{0\}$.

To compute $V_{sing}^{\mathbf{x}}$ *we apply Lemma 2.3.7:*
For $\nu = 0$, $B_0 = []$, $W_0 = \{\mathbf{x} \in \mathbb{C} \mid rank(B_0) \leq -1\} = \emptyset$ *then* $\Theta^0 = 1$, $\sigma_{0,1} = 1 \cdot x^3$. *For* $\nu = 1$, $B_1 = [-3x^2]$, $W_1 = \{\mathbf{x} \in \mathbb{C} \mid x^3 = 0, rank(B_1) \leq 0\} = \{0\}$ *then* $\Theta_1^1 = x^3$, $\Theta_2^1 = x^2$,$\sigma_{1,1} = x^3$, $\sigma_{1,2} = x^2$ *and* $\Gamma_{sing} = (x^2)$.
Then $V_{sing}^{\mathbf{x}} = W_1 = \mathcal{V}(\Gamma_{sing}) = \{0\}$.

2.4 The minimizer variety

By the decomposition (2.12) and Proposition 2.3.2, we know that the minimizer points of f on S are in

$$\boxed{S_{FJ} = S_{KKT} \cup S_{sing}} \tag{2.16}$$

where

$$S_{FJ} = V_{FJ}^{\mathbf{x},\mathbb{R}} \cap S = V_{FJ}^{\mathbf{x},\mathbb{R}} \cap \mathcal{S}^+(\mathbf{g}), \tag{2.17}$$

$$S_{KKT} = V_{KKT}^{\mathbf{x},\mathbb{R}} \cap S = V_{KKT}^{\mathbf{x},\mathbb{R}} \cap \mathcal{S}^+(\mathbf{g}) \tag{2.18}$$

$$S_{sing} = V_{sing}^{\mathbf{x},\mathbb{R}} \cap S = V_{sing}^{\mathbf{x},\mathbb{R}} \cap \mathcal{S}^+(\mathbf{g}) \tag{2.19}$$

Therefore, we can decompose the initial optimization problem (2.1) into two subproblems:

1. find the infimum of f on S_{KKT};
2. find the infimum of f on S_{sing};

and take the least of these two infima. Since the second problem is of the same type as (2.1) but with the additional constraints $\sigma_{\nu,i} = 0$ described in (2.15), we analyse only the first subproblem. The approach developed for this first sub-problem is applied recursively to the second subproblem, in order to obtain the solution of Problem (2.1).

Example 2.4.1 *We consider the "ill-posed" problem*

$$\min\ x\ s.t\ x^3 \geq 0.$$

The ideal I_{KKT} is $I_{KKT} = (1-3v_1x^2, v_1x^3) = (1)$. Thus $V_{KKT} = \emptyset$. According to the decomposition (2.16), $S_{FJ} = S_{sing}$ and we compute the minimum of x on S_{sing}, which is defined by $x^2 = 0$:

$$\min\ x\ s.t\ x^2 = 0.$$

As we will see in the Section 4.3.4, the relaxation associated to this problem is exact because $V^{\mathbb{R}}(x^2) = 0$ is finite and yields the solution $x = 0$.

Definition 2.4.2 *We define the KKT-minimizer set and ideal of f on S as:*

$$\begin{aligned} S_{min} &= \{\mathbf{x}^* \in S_{KKT} \text{ s.t. } \forall \mathbf{x} \in S_{KKT}, f(\mathbf{x}^*) \leq f(\mathbf{x})\} \\ I_{min} &= \mathcal{I}(S_{min}) \subset \mathbb{R}[\mathbf{x}]. \end{aligned}$$

A point $\mathbf{x}^*$ in S_{min} is called a KKT-minimizer. Notice that $I_{KKT} \subset I_{min}$ and that I_{min} is a real radical ideal.

We have $I_{min} \neq (1)$, if and only if, the KKT-minimum f^* is reached in S_{KKT}.

If $n_1 = n_2 = 0$, I_{min} is the vanishing ideal of the *critical points* $\mathbf{x}^*$ of f (satisfying $\nabla f(\mathbf{x}^*) = 0$) where $f(\mathbf{x}^*)$ reaches its minimal critical value.

Remark 2.4.3 *If we take $f = 0$ in the minimization problem (2.1), then all the points of S are KKT-minimizers and $I_{min} = \mathcal{I}(S) = \sqrt[\mathbf{g}^+]{\mathbf{g}^0}$. Moreover, $I_{KKT} \cap \mathbb{R}[\mathbf{x}] = (g_1^0, \ldots, g_{n_1}^0) = (\mathbf{g}^0)$ since $F_1, \ldots, F_n, v_1 g_1^+, \ldots, v_{n_2} g_{n_2}^+$ are homogeneous of degree 1 in the variables $\mathbf{u}, \mathbf{v}$.*

Chapter 3

Relation between Optimization problem and Moment matrices

In this chapter we connect the minimization problem (2.1) with the theory of Moment matrices and the set of positive polynomials.

For the case in one variable, that is, $n = 1$, Shor [Shor 1987] showed that the unconstrained minimization problem

$$f^* = min_{x \in \mathbb{R}^n} f(x) \tag{3.1}$$

reduces to a convex problem. Afterwards Nesterov [Nesterov 2000], through a representation of univariate non negative polynomials as a sum of squares, provided a self-concordant barrier for the cone of nonnegative univariate polynomials so that efficient interior point algorithms become available to compute the global minimum.

However, the multivariate case is very different from the one-dimensional case because not every nonnegative polynomial can be written as a sum of squares of polynomials. For instance the Robinson polynomial and Motzkin polynomial are such famous polynomials in 2 variables which are non-negatives but not sum of squares. For the last one we will give the proof in the first Section of this Chapter. As mentioned by Nesterov in [Nesterov 2000], the global unconstrained minimization problem of a 4-degree polynomial is a NP-hard problem. For constrained minimization problem as (2.1), Shor [Shor 1998] transforms this problem, via succesive changes of variables, into a quadratic constrained optimization problem that he solves through a standard convex linear matrix inequality (LMI) relaxation to obtain a good lower bound. If we add redundant quadratic constraints we can improve the lower bound and sometimes obtain the optimal value.

About a decade ago, a relaxation approach was proposed by Lasserre in [Lasserre. 2001] to solve this difficult problem. Instead of searching points where the polynomial f reaches its minimum that we will call f^*, a probability measure which minimizes the function f is searched. This problem is relaxed into a hierarchy of finite dimensional convex minimization problems, which can be solved by Semi-Definite Programming (SDP) techniques. The sequence

of SDP minima converges to the minimum f^* [Lasserre. 2001] under some hypothesis that we will see at the end of the chapter. This hierarchy of SDP problems can be formulated in terms of linear matrix inequalities on moment matrices associated to the set of monomials of degree t or less, for increasing values of t. The dual hierarchy can be described as a sequence of maximization problems over the cone of polynomials that are Sums of Squares (SoS). A feasibility condition is needed to prove that this dual hierarchy of maximization problems also converges to the minimum f^*, i.e. that there is no duality gap.

This chapter is organized as follows, in the first Section we give the definition of positive and sum of squares polynomials, the definition of preordering and quadratic module and the theorems which rely on these concepts due to Putinar and Schmudgen. We also talk about the sequence of maximization problems over the cone of SoS polynomials. In the second Section we recall the concepts of Moment matrices and the theorems related and the last Section we explain the relaxation method of Lasserre [Lasserre. 2001].

3.1 Positive Polynomials

First of all we recall definitions as sum of squares, quadratic module and preordering that we have seen in Chapter 1.
We say that a polynomial $p \in \mathbb{R}[\mathbf{x}]$ is a *sum of squares of polynomials* (**SOS**) if p can be written as $p = \sum_{i=1}^{m} u_j^2$ for some $u_1, ..., u_m \in \mathbb{R}[\mathbf{x}]$.

Let $S \subset \mathbb{R}^n$ be a basic semialgebraic set, defined as in Chapter 2, $S := S(\mathbf{g}) = S(\mathbf{g}^0; \mathbf{g}^+)$,

$$S = \{\mathbf{x} \in \mathbb{R}^n \mid g_1^0(\mathbf{x}) = 0, \ldots, g_{n_1}^0(\mathbf{x}) = 0; g_1^+(\mathbf{x}) \geq 0, \ldots, g_{n_2}^+(\mathbf{x}) \geq 0\}$$

Definition 3.1.1

- *A polynomial $f(\mathbf{x})$ is called nonnegative on S if*

$$f(u) \geq 0 \ \forall u \in S.$$

- *A polynomial $f(\mathbf{x})$ is called positive on S if*

$$f(u) > 0 \ \forall u \in S.$$

Definition 3.1.2

- *The preordering $\mathcal{P}(\mathbf{g})$ is the set*

$$\mathcal{P}(\mathbf{g}) = \{\sum_{i=1}^{n_1} \phi_i(\mathbf{x}) g_i^0 + \sum_{\nu\in\{0,1\}^m} \sigma_\nu(\mathbf{x}) g_1^+(\mathbf{x})^{\nu_1} \cdots g_{n_2}^+(\mathbf{x})^{\nu_{n_2}} \mid \sigma_\nu(\mathbf{x}) \text{ is sos}, \ \phi_i(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]\} \tag{3.2}$$

  The preordering generated by the positive constraints is denoted $\mathcal{P}^+(\mathbf{g}) = \mathcal{P}(\mathbf{g}^+)$.

- *For a finite dimensional subspace $E \subset \mathbb{R}[\mathbf{x}]$ we can define the truncated preordering*

$$\mathcal{P}_E(\mathbf{g}) = \{\sum_{i=1}^{n_1} \phi_i(\mathbf{x}) g_i^0 + \sum_{\nu\in\{0,1\}^m} \sigma_\nu(\mathbf{x}) g_1^+(\mathbf{x})^{\nu_1} \cdots g_{n_2}^+(\mathbf{x})^{\nu_{n_2}} \mid \sigma_\nu(\mathbf{x}) \text{ is sos}(E), \sigma_\nu(\mathbf{x}) g_1^+(\mathbf{x})^{\nu_1} \cdots g_{n_2}^+(\mathbf{x})^{\nu_{n_2}} \in \langle E, E\rangle, \phi_i(\mathbf{x}) \in E, \phi_i(\mathbf{x}) g_i^0 \in \langle E, E\rangle\} \tag{3.3}$$

Definition 3.1.3

- *The quadratic module $\mathcal{Q}(S) := \mathcal{Q}(\mathbf{g})$ is the set*

$$\mathcal{Q}(\mathbf{g}) = \{\sum_{i=1}^{n_1} \phi_i(\mathbf{x}) g_i^0 + \sigma_0 + \sum_{j=1}^{n_2} \sigma_j(\mathbf{x}) g_j^+(\mathbf{x}) \mid \sigma_j \text{ is sos}, \ \phi_i(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]\} \tag{3.4}$$

  The quadratic module generated by the positive constraints is denoted $\mathcal{Q}^+(\mathbf{g}) = \mathcal{Q}(\mathbf{g}^+)$.

- *For a finite dimensional subspace $E \subset \mathbb{R}[\mathbf{x}]$ we can define the truncated quadratic module*

$$\mathcal{Q}_E(\mathbf{g}) = \{\sum_{i=1}^{n_1} \phi_i(\mathbf{x}) g_i^0 + \sum_{i=1}^{n_2} \sigma_\nu(\mathbf{x}) g_i^+(\mathbf{x})^{\nu_1} \mid \sigma_\nu(\mathbf{x}) \text{ is sos}(E), \sigma_\nu(\mathbf{x}) g_i^+(\mathbf{x})^{\nu_1} \in \langle E, E\rangle, \phi_i(\mathbf{x}) \in E, \phi_i(\mathbf{x}) g_i^0 \in \langle E, E\rangle\} \tag{3.5}$$

Proposition 3.1.4

- *The sets $\mathcal{P}(\mathbf{g})$ and $\mathcal{Q}(\mathbf{g})$ are convex.*
- *If $f(\mathbf{x}) \in \mathcal{P}(\mathbf{g})$, then $f(\mathbf{x})$ is nonnegative on S.*
- *If $f(\mathbf{x}) \in \mathcal{Q}(\mathbf{g})$, then $f(\mathbf{x})$ is nonnegative on S.*

Theorem 3.1.5 *[Schmüdgen 1991] Let $S = \{\mathbf{x} \in \mathbb{R}^n \mid g_1^+(\mathbf{x}) \geq 0, \ldots, g_{n_2}^+(\mathbf{x}) \geq 0\}$ be a compact set. If $f(\mathbf{x})$ is positive on S, then $f(\mathbf{x}) \in \mathcal{P}^+(\mathbf{g})$.*

Example 3.1.6 *The quadratic polynomial* $f(x_1, x_2) = x_1x_2 + 1$ *is positive on the unit sphere* $S = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1^2 + x_2^2 \leq 1\}$ *which is compact. Then we have that* $f(\mathbf{x}) \in \mathcal{P}^+(\mathbf{g})$

$$x_1x_2 + 1 = \frac{1}{2}(x_1 + x_2)^2 + \frac{1}{2} + \frac{1}{2}(1 - x_1^2 - x_2^2)$$

Remark 3.1.7 *If* $f(\mathbf{x})$ *is only nonnegative on* S*, then Schmudgen's theorem may not be true. We can see it in the following counter-example.*

Example 3.1.8

$$f(x) = 1 - x^2 \text{ and } K = \{(1 - x^2)^3 \geq 0\}.$$

$f(x)$ *is nonnegative on* K*. Suppose there are SOS polynomials* $s_1(x), s_2(x)$ *such that*

$$1 - x^2 = s_1(x) + s_2(x)(1 - x^2)^3 \tag{3.6}$$

then -1 *must be a root of* $s_1(x)$ *therefore of multiplicity 2, but* -1 *has multiplicity 1 on the left.*

Definition 3.1.9 *The module quadratic* $\mathcal{Q}^+(\mathbf{g})$ *sastifies the Archimedean condition (AC) if there exits* $N > 0$ *such that*

$$N - || \mathbf{x} ||_2^2 \in \mathcal{Q}^+(\mathbf{g}). \tag{3.7}$$

Remark 3.1.10 *If the AC holds, the set* S *must be compact, but the reverse might not be true. We can see it in the following counter-example.*

Example 3.1.11

$$S = \{\mathbf{x} \in \mathbb{R}^2 \mid x_1 - \frac{1}{2} \geq 0, x_2 - \frac{1}{2} \geq 0, 1 - x_1x_2 \geq 0\}$$

This set is clearly compact.

The Archimedean Condition is not verified because there does not exist N *such that*

$$N - (x_1^2 + x_2^2) = \sigma_0 + (x_1 - \frac{1}{2})\sigma_1 + (x_2 - \frac{1}{2})\sigma_2 + (1 - x_1x_2)\sigma_3, \text{ with } \sigma_i \text{ sos.}$$

If they exist, then $D = max(deg(\sigma_0), deg(\sigma_3)+2) \geq 1+max(deg(\sigma_1), deg(\sigma_2))$ *When* $D = 2$*, it does not work. When* $D > 2$*, the highest even term of* $\sigma_0 + (1 - x_1x_2)\sigma_3$ *must vanish, which is not possible.*

Theorem 3.1.12 *[Putinar 1993] Let* $S := \{\mathbf{x} \in \mathbb{R}^n \mid g_1^+(\mathbf{x}) \geq 0, \ldots, g_{n_2}^+(\mathbf{x}) \geq 0\}$ *be a compact set. Suppose the quadratic module* $\mathcal{Q}^+(\mathbf{g})$ *verify the Archimedean Condition. If* $f(\mathbf{x})$ *is positive on* S*, then* $f(\mathbf{x}) \in \mathcal{Q}^+(\mathbf{g})$

Remark 3.1.13 *If AC fails, the conclusion of Putinar's theorem migth not be true.*

Example 3.1.14 *We take the example 3.1.11,* S *is compact and the* $\mathcal{Q}^+(\mathbf{g})$ *does not verify the Archimedean Condition. The polynomial* $N - (x_1^2 + x_2^2)$ *with* $N > 2$ *is positive on* S *but* $N - (x_1^2 + x_2^2) \notin \mathcal{Q}^+(\mathbf{g})$

The relation between positive polynomials and our minimization problem is the following. We can reformulate our problem (2.1) as:

$$\boxed{f^* = \sup \rho \ s.t \ f(\mathbf{x}) - \rho \geq 0 \ on \ S} \tag{3.8}$$

In order to manage this hard problem, we can relaxe it into the following simpler problem for $S = \mathbb{R}^n$

$$\boxed{f^{sos} = \sup \rho \ s.t \ f(\mathbf{x}) - \rho \ is \ SOS} \tag{3.9}$$

The following lemma shows that we can tackl it with Semidefinite Programming.

Lemma 3.1.15 *Let* $f \in \mathbb{R}[\mathbf{x}]$, $f = \sum_{\alpha \in \mathbb{N}^n_{2d}} f_\alpha \mathbf{x}^\alpha$, *be a polynomial of degree* $\leq 2d$*. The following assertions are equivalent:*

1. f *is a sum of squares.*
2. *The following system in the matrix variable* $X = (X_{\alpha,\beta})_{\alpha,\beta \in \mathbb{N}^n_d}$ *is feasible:*

$$\begin{cases} X \succcurlyeq 0 \\ \sum_{\alpha,\gamma \in \mathbb{N}^n_d | \beta+\gamma=\alpha} X_{\beta,\gamma} = f_\alpha (\mid \alpha \mid \leq 2d). \end{cases} \tag{3.10}$$

Proof. Let $z_d := (x^\alpha \mid\mid \alpha \mid \leq d)$ is the vector containing all monomials of degree at most d. Then for polynomials $u_j \in \mathbb{R}[\mathbf{x}]_d$, we have $u_j = coeff(u_j)^T z_d$ and thus $\sum_j u_j^2 = z_d^T (\sum_j coeff(u_j) coeff(u_j)^T) z_d$. Therefore, f is a sum of squares of polynomials if and only if $f = z_d^T X z_d$ for some positive semidefinite matrix X. Comparing the coefficients of f and $z_d^T X z_d$ we find the system (3.10). ∎

Example 3.1.16 *We want to verify if the polynomial* $f = x^4+2x^3y+3x^2y^2+2xy^2+2y^4$ *which is nonnegative is SOS. We can decompose* f *as:*

$$f = \begin{pmatrix} x^2 & xy & y^2 \end{pmatrix} \cdot \underbrace{\begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix}}_{X} \cdot \begin{pmatrix} x^2 \\ xy \\ y^2 \end{pmatrix}$$

We look for $X \succcurlyeq 0$ *such that:*

$$\begin{aligned} x^4 = x^2 \cdot x^2 \quad & 1 = a \\ x^3y = x^2 \cdot xy \quad & 2 = 2b \\ x^2y^2 = xy \cdot xy = x^2 \cdot y^2 \quad & 3 = d + 2c \\ xy^3 = xy \cdot y^2 \quad & 2 = 2e \\ y^4 = y^2 \cdot y^2 \quad & 2 = f \end{aligned}$$

Then

$$X = \begin{pmatrix} 1 & 1 & c \\ 1 & 3-2c & 1 \\ c & 1 & 2 \end{pmatrix} \succcurlyeq 0 \Longleftrightarrow -1 \leq c \leq 1$$

For $c = -1$ *we have the Gram decomposition and we get*

$$X = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 5 & 1 \\ -1 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 2 \\ -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & -1 \\ 0 & 2 & 1 \end{pmatrix}$$

and one decomposition as sum of squares of p *is:* $f = (x^2+xy-y^2)^2+(2xy+y^2)^2$

Remark 3.1.17 *It is obvious that* $f^{sos} \leq f^*$. *But not all the non negative polynomials are sum of squares as we can see in the next example.*

Example 3.1.18 *The Motzkin polynomial* $f(x,y) = x^2y^2(x^2+y^2-3)+1$, *is non negative but not a sum of squares of polynomials (SOS).*

Indeed, $f(x,y) \geq 0$ *if* $x^2+y^2 \geq 3$. *Otherwise, there exists* $z \in \mathbb{R}$ *such that* $z^2 = -x^2-y^2+3$ *and by arithmetic geometric mean inequality, we have* $\frac{x^2+y^2+z^2}{3} \geq \sqrt[3]{x^2y^2z^2}$, *giving again* $f(x,y) \geq 0$.

Now we verify that we can not decompose $f(x,y)$ *as a sum of squares of polynomials. We suppose* $f = \sum_k p_k^2$, *where* $p_k = a_kx_1^3 + b_kx_1^2x2 + c_kx_1x_2^2 + d_kx_2^3 + e_kx_1^2 + f_kx_1x_2 + g_kx_2^2 + h_kx_1 + i_kx_2 + j_k$, *with* $a_k, ..., j_k \in \mathbb{R}$. *Looking at the coefficients of* x_1^6, x_2^6, x_1^4, x_2^4, x_1^2, x_2^2 *in* f, *we deduce that* $a_k = d_k = e_k = g_k = h_k = i_k = 0 \ \forall\ k$. *And if we look at the coefficient of* $x_1^2x_2^2$ *in* f, *we obtain* $-3 = \sum f_k^2$ *which implies a contradiction.*

In the case where $S = S(\mathbf{g})$ is a semialgebraic set

$$\boxed{f_{\mathbf{g}}^{sos} = \sup \rho \ s.t. \ f - \rho \in \mathcal{P}(\mathbf{g})} \tag{3.11}$$

Remark 3.1.19 *It this case with $S = S(\mathbf{g})$, it is obvious also that in his case $f_{\mathbf{g}}^{sos} \leq f^*$*

This formulation does not lead directly to a semidefinite program of finite size. We need to bound the degree of the polynomials. We can consider for any integer t with
$2t \geq max(deg(f), deg(g_1^0), \ldots, deg(g_{n_1}^0), deg(g_1^+), \ldots, deg(g_{n_2}^+))$ the following semidefinite program

$$\boxed{f_{t,\mathbf{g}}^{sos} = \sup \rho \ s.t. \ f - \rho \in \mathcal{P}_t(\mathbf{g})} \tag{3.12}$$

where the *truncated preordering* $\mathcal{P}_t(\mathbf{g})$, which is the particular case of $\mathcal{P}_E(\mathbf{g})$ where $E = \mathbb{R}[\mathbf{x}]_t$, is the set

$$\mathcal{P}_t(\mathbf{g}) = \{\sum_{i=1}^{n_1} \phi_i(\mathbf{x}) g_i^0 + \sum_{\nu \in \{0,1\}^m} \sigma_\nu(\mathbf{x}) g_1^+(\mathbf{x})^{\nu_1} \cdots g_{n_2}^+(\mathbf{x})^{\nu_{n_2}} \mid \sigma_\nu(\mathbf{x}) \ is \ sos_{t - \lceil \frac{\prod g_i^+}{2} \rceil}, \ \phi_i(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]_{2t - deg(g_i^0)}\} \tag{3.13}$$

We can also define the *truncated quadratic module* $\mathcal{Q}_t(\mathbf{g})$, which is the particular case of $\mathcal{Q}_E(\mathbf{g})$ where $E = \mathbb{R}[\mathbf{x}]_t$, is the set

$$\mathcal{Q}_t(\mathbf{g}) = \{\sum_{i=1}^{n_1} \phi_i(\mathbf{x}) g_i^0 + \sum_{i=1}^{n_2} \sigma_\nu(\mathbf{x}) g_i^+(\mathbf{x}) \mid \sigma_\nu(\mathbf{x}) \ is \ sos_{t - \lceil \frac{deg(g_i^+)}{2} \rceil}, \ \phi_i(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]_{2t - deg(g_i^0)}\} \tag{3.14}$$

Proposition 3.1.20 $f_{t,\mathbf{g}}^{sos} \leq f_{t+1,\mathbf{g}}^{sos} \leq f_{\mathbf{g}}^{sos} \leq f^*$ *and* $\lim_{t \to \infty} f_{t,\mathbf{g}}^{sos} = f_{\mathbf{g}}^{sos}$

3.2 Moment matrices

Before defining the relation between our minimization problem and the Moment Matrices we introduce some definitions. We consider nonnegative Borel measures on $\mathbb{R}^n$, thus, all our measures will be nonnegative. A probability measure μ is a measure with total mass $\mu(\mathbb{R}^n) = 1$.

Definition 3.2.1 *We call the* ***support*** *of a measure μ on $\mathbb{R}^n$, denoted $supp(\mu)$, the smallest closed set $C \subset \mathbb{R}^n$ for which $\mu(\mathbb{R}^n \setminus C) = 0$.*

Definition 3.2.2 *We say that μ is a measure **on** K or a measure **supported by** $K \subset \mathbb{R}^n$ if $supp(\mu) \subset K$.*

Definition 3.2.3 *Given $\mathbf{x} \in \mathbb{R}^n$, $\delta_{\mathbf{x}}$ denotes the **Dirac measure** at $\mathbf{x}$, with support, $supp(\delta_{\mathbf{x}}) = \{\mathbf{x}\}$ and:*

$$\delta_{\mathbf{x}}(\mathbf{u}) = \begin{cases} 1 & si\ \mathbf{u} = \mathbf{x} \\ 0 & sinon. \end{cases}$$

Definition 3.2.4 *If the support of a measure μ is finite, $supp(\mu) = \{x_1, ..., x_r\}$ then μ is of the form:*

$$\mu = \sum_{i=1}^{r} \lambda_i \delta_{x_i} \ for\ some\ \lambda_1, ..., \lambda_r > 0 \tag{3.15}$$

*where x_i are called the **atoms** of μ.*

*If the measure μ has a representation as 3.15 we say that μ is **r-atomic** measure.*

Definition 3.2.5 *We define the **moment of order** α of a measure μ on $\mathbb{R}^n$ is:*

$$y_\alpha = \int \mathbf{x}^\alpha (d\mu). \tag{3.16}$$

- *The sequence $(y_\alpha)_{\alpha \in \mathbb{N}^n}$ is called the **sequence of moments of the measure** μ.*
- *For $t \in \mathbb{N}$, the sequence $(y_\alpha)_{\alpha \in \mathbb{N}^n_t}$ is called the **truncated sequence of moments of the measure** μ **up to order** t.*
- *If the sequence y is the sequence of moments of a measure, we also say that μ is a **representing measure** for y.*

Definition 3.2.6

- *Given a sequence $y = (y_\alpha)_{\alpha \in \mathbb{N}^n} \in \mathbb{R}^{\mathbb{N}^n}$ its **moment matrix** is the infinite matrix $M(y)$ indexed by $\mathbb{N}^n$, with $(\alpha, \beta)th$ entry $y_{\alpha+\beta}$, for $\alpha, \beta \in \mathbb{N}^n$.*
- *For a $t \in \mathbb{Z}$, given a truncated sequence $y = (y_\alpha)_{\alpha \in \mathbb{N}^n_{2t}} \in \mathbb{R}^{\mathbb{N}^n_{2t}}$, its **moment matrix of order** t is the matrix $M_t(y)$ indexed by $\mathbb{N}^n_t$, with $(\alpha, \beta)th$ entry $y_{\alpha,\beta}$ for $\alpha, \beta \in \mathbb{N}^n_t$.*

Example 3.2.7 *The Moment matrix of order 2 of a measure is of the form:*

$$H^2(y) = \left[\begin{array}{c|cc|ccc} 1 & y_{1,0} & y_{0,1} & y_{2,0} & y_{1,1} & y_{0,2} \\ \hline y_{0,1} & y_{2,0} & y_{1,1} & y_{3,0} & y_{2,1} & y_{1,2} \\ y_{1,0} & y_{1,1} & y_{0,2} & y_{2,1} & y_{1,2} & y_{0,3} \\ \hline y_{2,0} & y_{3,0} & y_{2,1} & y_{4,0} & y_{3,1} & y_{2,2} \\ y_{1,1} & y_{2,1} & y_{1,2} & y_{3,1} & y_{2,2} & y_{1,3} \\ y_{0,2} & y_{1,2} & y_{0,3} & y_{2,2} & y_{1,3} & y_{0,4} \end{array}\right]$$

where $y_{i,j}$ represents the $(i+j)$-order moment $\int x^i y^j \mu(d(x,y)$

Definition 3.2.8 *For a certain degree t we called the **truncated localizing matrix** to the matrix $H^t(g \star y)$ with entries:*

$$H^t(g \star y)(i,j) = \sum_{\alpha} g_\alpha y_{(i,j)+\alpha}$$

Example 3.2.9 *The Moment matrix of order 1 of a measure is of the form:*

$$H^1(y) = \begin{bmatrix} 1 & y_{1,0} & y_{0,1} \\ y_{0,1} & y_{2,0} & y_{1,1} \\ y_{1,0} & y_{1,1} & y_{0,2} \end{bmatrix}$$

IF we take $g = 2 - x_1^2 - x_1 x_2$, then

$$H^1(g \star y) = \begin{bmatrix} 2 - y_{2,0} - y_{1,1} & 2y_{1,0} - y_{3,0} - y_{2,1} & 2y_{0,1} - y_{2,1} - y_{1,2} \\ 2y_{1,0} - y_{3,0} - y_{2,1} & 2y_{2,0} - y_{4,0} - y_{3,1} & 2y_{1,1} - y_{3,1} - y_{2,2} \\ 2y_{0,1} - y_{2,1} - y_{1,2} & 2y_{1,1} - y_{3,1} - y_{2,2} & 2y_{0,2} - y_{2,2} - y_{1,3} \end{bmatrix}$$

Remark 3.2.10 *We can easily verifier that Moment matrices of a degree t corresponding a truncated hankel matrix where $E = \mathbb{R}[\mathbf{x}]_t$ and its equivalent in notation is $H^t_{g \cdot \Lambda}$.*

Now we relate the Moments matrices with the Hankel Matrix

Definition 3.2.11 *Given $y \in \mathbb{R}^{\mathbb{N}^n}$, the linear form $\Lambda_y \in \mathbb{R}[\mathbf{x}]^*$ is defined by:*

$$\Lambda_y(f) = y^T vec(f) = \sum_{\alpha} y_\alpha f_\alpha = \sum f_\alpha \Lambda_y(x^\alpha) = vect(1)^T H(y) vec(f) \quad (3.17)$$

for $f = \sum_\alpha f_\alpha x^\alpha$.

Lemma 3.2.12 *[Laurent 2009a] Let $g \in \mathbb{R}[\mathbf{x}]$ and $d_g = \lceil deg(g)/2 \rceil$*

1. *If* $y \in \mathbb{R}^{\mathbb{N}^n_{2t}}$ *is the sequences of moments (up to order 2t) of a measure* μ, *then* $H^t(y) \succcurlyeq 0$ *and* $\text{rank}H^t(y) \leq |\ supp(\mu)\ |$. *Moreover, for* $f \in \mathbb{R}[\mathbf{x}]$, $H^t(y)f = 0$ *implies* $supp(\mu) \subset V_{\mathbb{R}}(f) = \{x \in \mathbb{R}^n \mid f(x) = 0\}$. *Therefore,* $supp(\mu) \subset V_{\mathbb{R}}(\ker H^t(y))$.

2. *If* $y \in \mathbb{R}^{\mathbb{N}^n_{2t}} (t \geq d_g)$ *is the sequence of moments of a measure* μ *supported in the set* $K = \{x \in R^n \mid g(x) \geq 0\}$ *then* $H^{t-d_g}(g \star y) \succcurlyeq 0$.

3. *If* $y \in \mathbb{R}^{\mathbb{N}^n}$ *is the sequence of moments of measure* μ, *then* $H(y) \succcurlyeq 0$. *Moreover, if* $supp(\mu) \subset \{x \in \mathbb{R}^n \mid g(x) \geq 0\}$, *then* $H(g \star y) \succcurlyeq 0$ *and, if* μ *is r-atomic, then* $\text{rank}H(y) = r$.

Proof.

1. For $f \in \mathbb{R}[\mathbf{x}]_t$,

$$f^T H^t(y) f = \sum_{\alpha,\beta \in \mathbb{N}^n_t} f_\alpha f_\beta y_{\alpha+\beta} = \sum_{\alpha,\beta \in \mathbb{N}^n_t} \int x^{\alpha+\beta} \mu(dx) = \int f(x)^2 \mu(dx) \geq 0$$

which shows that $H^t(y) \succcurlyeq 0$.
If $H^t(y)f = 0$, then

$$0 = f^t H^t f = \int f^2(x) \mu(dx).$$

As $V^{\mathbb{R}}(f)$ is a closed set, $supp(\mu) \subset V^{\mathbb{R}}(f)$ holds if we can show that $\mu(\mathbb{R}^n \setminus V^{\mathbb{R}}(f)) = 0$. Indeed, $\mathbb{R}^n \setminus V^{\mathbb{R}}(f) = \bigcup_{k \geq 0} U_k$, setting $U_k = \{x \in R^n \mid f(x)^2 \geq \frac{1}{k}\}$ for positive $k \in \mathbb{N}$. As

$$0 = \int f(x)^2 \mu(dx) = \int_{\mathbb{R}^n \setminus V^{\mathbb{R}}(f)} f(x)^2 \mu(dx) \geq \int_{U_k} f(x)^2 \mu(dx) \geq \frac{1}{k} \mu(U_k)$$

this implies $\mu(U_k) = 0$ for all k and thus $\mu(\mathbb{R}^n \setminus V^{\mathbb{R}}(f)) = 0$.
The inequality $\text{rank}H^t(y) \leq |\ supp(\mu \mid$ is trivial if μ has a infinite support. So assume that μ is r-atomic , say, $\mu = \sum_{i=1}^r \lambda_i \delta_{x_i}$ where $\lambda_i > 0$ and $x_i \in \mathbb{R}^n$ for $i = 1, ..n$. Then $H^t(y) = \sum_{i=1}^r \lambda_i \zeta_{t,x_i} \zeta_{t,x_i}^T$, where $\zeta_{x_i} = (x_i^\alpha)_{\alpha \in \mathbb{N}^n}$ is the sequence of moments of the Dirac measure δ_{x_i}, called the Zeta vector of x_i and $\zeta_{t,x_i} = (x_i^\alpha)_{\alpha \in \mathbb{N}^n_t}$ denotes the truncated zeta vector. It shows that $\text{rank}H^t(y) \leq r$.

2. For $p \in \mathbb{R}[\mathbf{x}]_t$,

$$p^T H^{t-d_g}(g \star y) p = \sum_{\alpha,\beta \in \mathbb{N}^n_{t-d_g}} \sum_{\gamma \in \mathbb{N}^n} p_\alpha p_\beta g_\gamma y_{\alpha+\beta+\gamma} = \int_K g(x) p(x)^2 \mu(dx) \geq 0$$

which shows that $H^{t-d_g}(g \star y) \succcurlyeq 0$

3. The first two claims follow directly from 1,2. Assume now $\mu = \sum_{i=1}^{r} \lambda_i \delta_{x_i}$ where $\lambda_i > 0$ and $x_i \in \mathbb{R}^n$ for $i = 1, ..n$. Then, as $H(y) = \sum_{i=1}^{r} \lambda_i \zeta_{x_i} \zeta_{x_i}^T$, rank$H(y) = r$.

■

Corollary 3.2.13 *[Laurent 2009a] If $y \in \mathbb{R}^{\mathbb{N}^n_{2t}}$ is the sequences of moments (up to order 2t) of a measure supported by the set S then, for any $t \geq d_S = maxdeg(g_i)$,*

$$H^t(y) \succcurlyeq 0, \; H^{t-d_{g_j}}(g_j \star y) \succcurlyeq 0 \; (j = 1, \ldots, m) \tag{3.18}$$

We will study in Chapter 5, several results of Curto and Fialkow showing that, under certain restrictions on the rank of the matrix $H_t(y)$, the condition 3.18 is sufficient for ensuring that y is the sequence of moments of a measure supported by S

3.3 Lasserre relaxation

Based in the results on Moments Matrices of the previous Section, Lasserre in [Lasserre. 2001] proposed to solve the optimization problem (2.1) as a sequence of truncated convex optimization problems which converges to the minimum.

Proposition 3.3.1 *The problem (2.1)*

$$\boxed{f^* = \inf_{x \in S} f(x)} \tag{3.19}$$

is equivalent to the following problem:

$$\boxed{f^* = \inf_{\mu} \int_S f(x)\mu(dx)} \tag{3.20}$$

where the infimun is taken over all probability measures μ on $\mathbb{R}^n$ supported by the set S.

Proof. For any $x_0 \in S$, $f(x_0) = \int f(x)\mu(dx)$ for the dirac measure $\mu = \delta_{x_0}$ then $f^* \geq \inf_{\mu} \int_S f(x)\mu(dx)$. Conversely $f(x) \geq f^*$ for all $x \in S$, $\int_S f(x)\mu(dx) \geq \int_S f^*\mu(dx) = f^*$, since μ is a probability measure. ■

As $\int f(x)\mu(dx) = \sum_\alpha f_\alpha \int x^\alpha \mu(dx) = f^T y$, where $y = \int x^\alpha \mu(dx)$ denotes the sequence of moments of μ, we can reformulated 6.3.1 as:

$$\boxed{f^* = \inf f^T y \; s.t. \; y_0 = 1, \; y \; has \; a \; representing \; measure \; on \; S}. \tag{3.21}$$

According to Lemma 3.2.12 we can imposse the conditions over the Moment matrices that lead to the following lower bound:

$$\boxed{f_{\mathbf{g}}^{\mu} = \inf_{y\in\mathbb{R}^{\mathbb{N}^n}} f^T y \ s.t.\ y_0 = 1,\ H(y) \succcurlyeq 0,\ H(g_j \star y) \succcurlyeq 0} \tag{3.22}$$

This problem is equivalent to:

$$\boxed{f_{\mathbf{g}}^{\mu} = \inf_{\Lambda\in\mathbb{R}[\mathbf{x}]^*} \Lambda(f)\ s.t\ \Lambda(1) = 1,\ \Lambda(p) \geq 0\ \forall p \in \mathcal{P}(\mathbf{g})} \tag{3.23}$$

We cannot solve this problem because it involves infinite moment matrices so in order to obtain finite dimensional semidefinite problem we consider truncated moment matrices and we obtain the following hierarchy of truncated convex optimization problem that we can solve with semidefinite programming methods.

$$\boxed{f_{t,\mathbf{g}}^{\mu} = \inf_{y\in\mathbb{R}^{\mathbb{N}^n_{2t}}} f^T y \ s.t.\ y_0 = 1,\ H_t(y) \succcurlyeq 0,\ H_{t-d_{g_j}}(g_j \star y) \succcurlyeq 0 (j = 1, ..., n_2)} \tag{3.24}$$

or

$$\boxed{f_{t,\mathbf{g}}^{\mu} = \inf_{\Lambda\in\mathbb{R}[\mathbf{x}]^*_{2t}} \Lambda(f)\ s.t\ \Lambda(1) = 1,\ \Lambda(p) \geq 0\ \forall p \in \mathcal{P}_t(\mathbf{g})} \tag{3.25}$$

where $t \geq max(d_f, d_S)$, and

$$\mathcal{L}_t(\mathbf{g}) := \{\Lambda \in \mathbb{R}[\mathbf{x}]^*_{2t} \mid \Lambda(p) \geq 0,\ \forall p \in \mathcal{P}_t(\mathbf{g}), \Lambda(1) = 1\}. \tag{3.26}$$

By this definition, for any element $\Lambda \in \mathcal{L}_t(\mathbf{g})$ and any $g \in \langle \mathbf{g}^0 \rangle \cap \mathbb{R}[\mathbf{x}]_t$, we have $\Lambda(g) = 0$.

Thus (3.24) is equivalent to:

$$\boxed{f_{t,\mathbf{g}}^{\mu} = \inf_{\Lambda\in\mathbb{R}[\mathbf{x}]^*_{2t}} \Lambda(f)\ s.t\ \Lambda \in \mathcal{L}_t(\mathbf{g})} \tag{3.27}$$

This hierarchy of truncated convex optimization problem converges to f^{μ}:

$$f_{t,\mathbf{g}}^{\mu} \leq f_{t+1,\mathbf{g}}^{\mu} \leq \cdots \leq f^{\mu} \leq f^* \tag{3.28}$$

Moreover, we have

Proposition 3.3.2 $f_{t,\mathbf{g}}^{sos} \leq f_{t,\mathbf{g}}^{\mu} \leq f^*$.

Proof. We easily check that $f_{t,\mathbf{g}}^{sos} \leq f_{t,\mathbf{g}}^{\mu}$, since if there exists $\rho \in \mathbb{R}$ such that $f - \rho = q \in \mathcal{P}_t(\mathbf{g})$ then $\forall \Lambda \in \mathcal{L}_t(\mathbf{g})$, $\Lambda(f - \gamma) = \Lambda(f) - \gamma = \Lambda(q) \geq 0$.

We also have $f_t^{\mu}(\mathbf{g}) \leq f^*$ since for any $\mathbf{s} \in S$, the evaluation $\mathbf{1}_{\mathbf{s}} : p \in \mathbb{R}[\mathbf{x}] \mapsto p(\mathbf{s})$ verify that $\mathbf{1}_{\mathbf{s}}(1) = 1$ and $\mathbf{1}_{\mathbf{s}}(q) = q(s) \geq 0\ \forall q \in \mathcal{P}_t(\mathbf{g})$. ■

The relaxation hierarchies introduced in [Lasserre. 2001] correspond to the case where we take $\langle \mathbf{g}^0 | 2t \rangle$ instead of $(\mathbf{g}^0)$.

Hereafter, we will also call Lasserre hierarchy, the full moment matrix relaxation hierarchy. It corresponds to the sequences

$$\cdots \subset \mathcal{L}_{t+1}(\mathbf{g}) \subset \mathcal{L}_t(\mathbf{g}) \subset \cdots \text{ and } \cdots \subset \mathcal{Q}_t(\mathbf{g}) \subset \mathcal{Q}_{t+1}(\mathbf{g}) \subset \cdots$$

which yield the following increasing sequences for $t \in \mathbb{N}$:

$$\cdots f^{\mu}_{t,\mathbf{g}} \leq f^{\mu}_{t+1,\mathbf{g}} \leq \cdots \leq f^* \text{ and } \cdots f^{sos}_{t,\mathbf{g}} \leq f^{sos}_{t+1,\mathbf{g}} \leq \cdots \leq f^*.$$

The foundation of Lasserre's method is to show that these sequences converge to f^*. This is proved under some conditions in [Lasserre. 2001].

Remark 3.3.3 *The same results hold if we replace* $\mathbf{g}$ *by any other set of constraints such that* $S(C) = S(\mathbf{g})$.

Chapter 4

Finite Convergence Certification

We have seen in the previous chapter that the approach proposed by Lasserre in [Lasserre. 2001], yields a sequence of SDP minima that converges to the minimimun of f under some hypothesis. We will say that the relaxation problem is exact if converge to the minimimun in a finite number of steps. One may wonder if using this approach, the relaxation problem is exact and how to compute the minimizer points when the mimization problem is feasible.

In order to answers to this question, the following strategy has been considered: add polynomial inequalities or equalities satisfied by the points where the function f is minimum.

A first family of methods are used when the set S is compact or when the minimizer set can be bounded easily. By adding an inequality constraint, one can then transform S into a compact subset of $\mathbb{R}^n$, for which exact hierarchies can be used [Lasserre. 2001], [Marshall 2003]. It is shown in [Laurent 2007] that if the complex variety defined by the equalities $\mathbf{g}^0 = 0$ is finite (and thus S is compact), then the hierarchy of relaxation problems introduced by Lasserre in [Lasserre. 2001] is exact. It is also proved that there is no duality gap if the generators of this ideal satisfy some regularity conditions. In [Laurent 2009a], it is proved that if the real variety defined by the equalities $\mathbf{g}^0 = 0$ is finite, then the hierarchy of relaxation problems introduced by Lasserre is exact. It is also proved in [Nie 2013a] using differents techniques.

In a second family of methods, equality constraints which are naturally satisfied by the minimizer points are added. These constraints are for instance the gradient of f when $S = \mathbb{R}^n$ or the Karush-Kuhn-Tucker (KKT) constraints, obtained by introducing Lagrange multipliers. In [Nie 2006], it is proved that a relaxation hierarchy using the gradient constraints is exact when the gradient ideal is radical. In [Marshall 2009], it is shown that this gradient hierarchy is exact, when the global minimizers satisfy the Boundary Hessian condition. In [Demmel 2007], it is proved that a relaxation hierarchy which involves the KKT constraints is exact when the KKT ideal is radical. In [Ha 2010], a relaxation hierarchy obtained by projection of the KKT constraints is proved to be exact under a regularity condition on the *real* minimizer points[1]. In [Nie 2011], a similar relaxation hierarchy is shown to be

[1] The results of this paper are true but a problem appears in the proof which we fix in

exact under a stronger regularity condition for the *complex* points of associated KKT varieties. These regularity conditions require that the gradient of the active constraints evaluated at the points of S or of some complex varieties are linearly independent. Thus they cannot be used for general semi algebraic sets S, for instance when S is a real non-complete intersection variety.

Moreover, the assumption that the minimum is reached at a KKT point is required. Unfortunately, in some cases the set of KKT points of S can be empty. As we have seen in Chapter 2, this obstacle can be removed using Fritz John variety (see [John 1948, Mangasarian 1967]). There is not much work dedicated to this issue (see [Lasserre 2009a]).

The case where the infimum value is not reached has also been studied. In [Schweighofer 2006], relaxation techniques are studied for functions for which the minimum is not reached and which satisfy some special properties "at infinity". In [Ha 2008], tangency constraints are used in a relaxation hierarchy which converges to the global minimum of a polynomial, when the polynomial is bounded by below over $\mathbb{R}^n$. In [Guo 2010], generic changes of coordinates and a partial gradient ideal are used in a relaxation hierarchy which also converges to the global minimum of f on $\mathbb{R}^n$.

Notice that Problem (2.1) can be attacked from a purely algebraic point of view. It reduces to the computation of a (minimal) critical value and polynomial system solvers can be used to tackle it (see e.g. [Parrilo 2003], [Greuet 2011]). But in this case, the complex solutions of the underlying algebraic system come into play and additional computation efforts should be spent to remove these extraneous solutions. Semi-algebraic techniques such as Cylindrical Algebraic Decomposition or extensions [Safey El Din 2008] may also be considered here, providing algorithms to solve Problem (2.1), but suffering from similar issues.

In the cases studied so far, the exactness of the relaxation is proved under a genericity condition or a compactness property. From an algorithmic point of view, the flat extension condition of Curto-Fialkow [Curto 1996] is used in most of the works [Henrion 2005, Laurent 2007, Lasserre 2009b, Laurent 2009a] to detect the exactness of the hierarchy, when the number of minimizers is finite. In [Lasserre 2012], a sparse extension [Laurent 2009b] of this flat extension condition is used to compute zero-dimensional real radical ideals.

Our aim is to show that for the general polynomial optimization problem (2.1), exact SDP relaxations can be constructed, which either detect that the problem is infeasible (which means there not exists points on S which minimize

this Chapter.

f) or compute the minimal value and the ideal associated to the minimizer points. In this Chapter we talk about KKT minimizer ideal S_{KKT} but how we have said in Chapter 2 we can apply the same theorems for the singuliers minimizer ideal S_{sing} to obtain the result on S. The main contributions of this Chapter are:

- We prove that exact relaxation hierarchies depending on the variables $\mathbf{x}$ can be constructed for solving the optimization problem (2.1) (see Theorem 4.3.3 and Theorem 4.2.10).
- We prove that if the set of KKT minimizers is empty, the SDP relaxation will eventually be empty (Theorem 4.3.3).
- We prove that the KKT minimizer ideal can be constructed from the moment matrix of an optimal linear form, when the corresponding relaxation is exact, even if the ideal is not zero-dimensional (Theorem 4.2.10).
- We prove that the exactness of the relaxation depends only on the real points which satisfy these constraints (Theorem 4.2.10).
- We provide a general approach which allows us to treat in a uniform way and to extend results on the representation of polynomials which are positive (resp. non-negative) on the critical points (see [Demmel 2007] and Theorem 4.1.9) and on the exactness of relaxation hierarchies (see [Nie 2006], [Ha 2008], [Lasserre 2009b], [Nie 2011], [Lasserre 2012], [Nie 2013a] and Theorem 4.3.2, Theorem 4.3.5, Theorem 4.3.6, Theorem 4.3.7).

4.1 Representation of positive polynomials

In this Section, we analyse the decomposition of polynomials as sum of squares modulo the gradient ideal. Hereafter, J_{grad} is an ideal of $\mathbb{R}[\mathbf{z}]$ such that $\mathcal{V}(J_{grad}) = V_{grad}$ and C is a set of constraints in $\mathbb{R}[\mathbf{x}]$ such that $\mathcal{S}^+(C) = \mathcal{S}^+(\mathbf{g})$.

The first steps consists in decomposing V_{grad} in components on which f has a constant value. We recall here a result, which also appears (with slightly different hypotheses) in [Nie 2006, Lemma 3.3][2].

[2]In its proof, the Mean Value Theorem is applied for a complex valued function, which is not valid. We correct the problem in the proof of Lemma 4.1.1.

Lemma 4.1.1 *Let $f \in \mathbb{R}[\mathbf{x}]$ and let V be an irreducible subvariety contained in $\mathcal{V}^{\mathbb{C}}(\nabla f)$. Then $f(x)$ is constant on V.*

Proof. If V is irreducible in the Zariski topology induced from $\mathbb{C}[\mathbf{x}]$, then it is connected in the strong topology on $\mathbb{C}^n$ and even piecewise smoothly path-connected [Shafarevich 1974]. Let x, y be two arbitrary points of V. There exists a piecewise smooth path $\varphi(t)$ $(0 \leq t \leq 1)$ lying inside V such that $x = \varphi(0)$ and $y = \varphi(1)$. Without loss of generality, we can assume that φ is smooth between x and y in order to prove that $f(x) = f(y)$. By the Mean Value Theorem, it holds that for some $t_1 \in (0, 1)$

$$\mathrm{Re}(f(y) - f(x)) = \mathrm{Re}(f(\varphi(t)))'(t_1) = \mathrm{Re}((\nabla f(\varphi(t_1)) * \varphi'(t_1))) = 0$$

since ∇f vanishes on V. Then $\mathrm{Re}(f(y)) = \mathrm{Re}(f(x))$. We have the same result for the imaginary part: for some $t_2 \in (0, 1)$

$$\mathrm{Im}(f(y)) - \mathrm{Im}(f(x)) = \mathrm{Im}(f(\varphi(t)))'(t_2) = \mathrm{Im}((\nabla f(\varphi(t_2)) * \varphi'(t_2))) = 0$$

since ∇f vanishes on V. Then $\mathrm{Im}(f(y)) = \mathrm{Im}(f(x))$. We conclude that $f(y) = f(x)$ and hence f is constant on V. ∎

Lemma 4.1.2 *The ideal J_{grad} can be decomposed as $J_{grad} = J_0 \cap J_1 \cap \cdots \cap J_s$ with $V_i = \mathcal{V}(J_i)$ and $W_i = \overline{\pi^{\mathbf{x}}(V_i)}$ where $\pi^{\mathbf{x}}(V_i)$ is the projection of V_i on $\mathbb{C}^n$ such that*

- *$f(V_j) = f_j \in \mathbb{C}$, $f_i \neq f_j$ if $i \neq j$,*
- *$W_i^{\mathbb{R}} \cap \mathcal{S}^+(C) \neq \emptyset$ for $i = 0, \ldots, r$,*
- *$W_i^{\mathbb{R}} \cap \mathcal{S}^+(C) = \emptyset$ for $i = r+1, \ldots, s$,*
- *$f_0 < \cdots < f_r$.*

Proof. Consider a minimal primary decomposition of J_{grad}:

$$J_{grad} = Q_0 \cap \cdots \cap Q_{s'},$$

where Q_i is a primary component, and $\mathcal{V}(Q_i)$ is an irreducible variety in $\mathbb{C}^{n+n_1+2n_2}$ included in V_{grad}. By Lemma 4.1.1, F is constant on $\mathcal{V}(Q_i)$. By Lemma 2.1.4, it coincides with f on each variety $\mathcal{V}(Q_i)$. We group the primary components Q_i according to the values $f_0, \ldots, f_s$ of f on these components, into $J_0, \ldots, J_s$ so that $f(\mathcal{V}(J_j)) = f_i$ with $f_i \neq f_j$ if $i \neq j$.

We can number them so that $\overline{\pi^{\mathbf{x}}(V_i)}^{\mathbb{R}} \cap \mathcal{S}^+(C)$ is empty for $i = r+1, \ldots, s$ and contains a real point $\mathbf{x}_i$ for $i = 0, \ldots, r$. Notice that such a point $\mathbf{x}_i$ is in $\mathcal{S}$, since it satisfies $g^0(\mathbf{x}_i) = 0 \ \forall g^0 \in C^0$ and $g^+(\mathbf{x}_i) \geq 0 \ \forall g^+ \in C^+$. As it is the limit of the projection of points in $\mathcal{V}(J_i)$ on which f is constant, we have $f_i = f(\mathbf{x}_i) \in \mathbb{R}$ for $i = 0, \ldots, r$. We can then order $J_0, \ldots, J_r$ so that $f_0 < \cdots < f_r$. ■

Remark 4.1.3 *If the minimum of f on S is reached at a KKT-point, then we have $f_0 = \min_{\mathbf{x} \in S} f(\mathbf{x})$.*

Remark 4.1.4 *If $V^{\mathbb{R}}_{grad} = \emptyset$, then for all $i = 0, \ldots, s$, $W_i^{\mathbb{R}} \cap \mathcal{S}^+(C) = \emptyset$ and by convention, we take $r = -1$.*

Lemma 4.1.5 *There exist $p_0, \ldots, p_s \in \mathbb{C}[\mathbf{x}]$ such that*

- $\sum_{i=0}^{s} p_i = 1 \mod J_{grad}$,
- $p_i \in \bigcap_{j \neq i} J_j$,
- $p_i \in \mathbb{R}[\mathbf{x}]$ *for* $i = 0, \ldots, r$.

Proof. Let $(L_i)_{i=0,\ldots,s}$ be the univariate Lagrange interpolation polynomials at the values $f_0, \ldots, f_s \in \mathbb{C}$ and let $q_i(\mathbf{x}) = L_i(f(\mathbf{x}))$.

The polynomials q_i are constructed so that

- $q_i(V_j) = 0$ if $j \neq i$,
- $q_i(V_i) = 1$,

where $V_i = \mathcal{V}(J_i)$. As the set $\{f_{r+1}, \ldots, f_s\}$ is stable by conjugation and $f_0, \ldots, f_r \in \mathbb{R}$, by construction of the Lagrange interpolation polynomials we deduce that $q_0, \ldots, q_r \in \mathbb{R}[\mathbf{x}]$.

By Hilbert's Nullstellensatz, there exists $N \in \mathbb{N}$ such that $q_i^N \in \bigcap_{j \neq i} J_j$. As $\sum_{j=0}^{s} q_j^N = 1$ on V_{grad} and $q_i^N q_j^N = 0 \mod \bigcap_i J_i = J_{grad}$ for $i \neq j$, we deduce that there exists $N' \in \mathbb{N}$ such that

$$\begin{aligned} 0 &= (1 - \sum_{j=0}^{s} q_j^N)^{N'} \mod J_{grad} \\ &= 1 - \sum_{j=0}^{s} (1 - (1 - q_j^N)^{N'}) \mod J_{grad}. \end{aligned}$$

As the polynomial $p_j = 1-(1-q_j^N)^{N'} \in \mathbb{C}[\mathbf{x}]$ is divisible by q_j^N, it belongs to $\bigcap_{j\neq i} J_j$. Since $q_j \in \mathbb{R}[\mathbf{x}]$ for $j = 0,\ldots,r$, we have $p_j \in \mathbb{R}[\mathbf{x}]$ for $j = 0,\ldots,r$, which ends the proof of this lemma. ■

Lemma 4.1.6 $-1 \in \mathcal{P}^+(C) + (\bigcap_{i>r} J_i^{\mathbf{x}})$.

Proof. As $\bigcup_{i>r} \overline{\pi^{\mathbf{x}}(V_i)}^{\mathbb{R}} \cap \mathcal{S}^+(C) = \mathcal{V}^{\mathbb{R}}(\bigcap_{i>r} J_i \cap \mathbb{R}[\mathbf{x}]) \cap \mathcal{S}^+(C) = \mathcal{V}^{\mathbb{R}}(\bigcap_{i>r} J_i^{\mathbf{x}}) \cap \mathcal{S}^+(C) = \emptyset$, we have $\mathcal{I}(\mathcal{V}^{\mathbb{R}}(\bigcap_{i>r} J_i^{\mathbf{x}}) \cap \mathcal{S}^+(C)) = \mathbb{R}[\mathbf{x}] \ni 1$ and by the Positivstellensatz (Theorem 1.1.7 (iii)),

$$-1 \in \mathcal{P}^+(C) + (\bigcap_{i>r} J_i^{\mathbf{x}}).$$

■

Corollary 4.1.7 *If $S_{min} = \emptyset$, then $-1 \in \mathcal{P}^+(C) + J_{grad}^{\mathbf{x}}$.*

Proof. If $S_{min} = \emptyset$, then f has no real KKT critical value on $S(C)$ and $r = -1$. Lemma 4.1.6 implies that $-1 \in \mathcal{P}^+(C) + (\bigcap_{i=0}^{s} J_i^{\mathbf{x}}) = \mathcal{P}^+(C) + J_{grad}^{\mathbf{x}}$. ■

In this case, $\forall p \in \mathbb{R}[\mathbf{x}]$, $p = \frac{1}{4}((p+1)^2 - (p-1)^2) \in \mathcal{P}^+(C) + J_{grad}^{\mathbf{x}}$. If C^0 is chosen such that $V(C^0) \subset V_{grad}^{\mathbf{x}}$ then $S_{min} = \emptyset$ if and only if $-1 \in \mathcal{P}(C)$.

We recall another useful result on the representation of positive polynomials (see for instance [Demmel 2007]):

Lemma 4.1.8 *Let $J \subset \mathbb{R}[\mathbf{z}]$ and $V = \mathcal{V}(J)$ such that $f(V) = f^*$ with $f^* \in \mathbb{R}^+$. There exists $t \in \mathbb{N}$, s.t. $\forall \varepsilon > 0$, $\exists\, q \in \mathbb{R}[\mathbf{x}]$ with $\deg(q) \leq t$ and $f + \varepsilon = q^2$ mod J.*

Proof. We know that $\frac{f+\varepsilon}{f^*+\varepsilon} - 1$ vanishes on V. By Hilbert's Nullstellensatz $(\frac{f+\varepsilon}{f^*+\varepsilon} - 1)^l \in J$ for some $l \in \mathbb{N}$. From the binomial theorem, it follows that

$$(1 + (\frac{f+\varepsilon}{f^*+\varepsilon} - 1))^{1/2} \equiv \sum_{i}^{l-1} \binom{1/2}{k} (\frac{f+\varepsilon}{f^*+\varepsilon} - 1)^k \stackrel{def}{=} \frac{q}{\sqrt{f^*+\varepsilon}} \; mod\; J$$

Then $f + \varepsilon = q^2 \; mod \; J$. ■

In particular, if $f^* > 0$ this lemma implies that $f = (f - \frac{1}{2}f^*) + \frac{1}{2}f^* = q^2$ mod J for some $q \in \mathbb{R}[\mathbf{x}]$.

Theorem 4.1.9 *Let $C \subset \mathbb{R}[\mathbf{x}]$ be a set of constraints such that $\mathcal{S}^+(C) = \mathcal{S}^+(\mathbf{g})$, let $f \in \mathbb{R}[\mathbf{x}]$, let $f_0 < \cdots < f_r$ be the real KKT critical values of f on S and let $p_0, \ldots, p_r$ be the associated polynomials defined in Lemma 4.1.5.*

1. $f - \sum_{i=0}^{r} f_i\, p_i^2 \in \mathcal{P}^+(C) + \sqrt{J^{\mathbf{x}}_{grad}}$.
2. *If $f \geq 0$ on S_{KKT}, then $f \in \mathcal{P}^+(C) + \sqrt{J^{\mathbf{x}}_{grad}}$.*
3. *If $f > 0$ on S_{KKT}, then $f \in \mathcal{P}^+(C) + J^{\mathbf{x}}_{grad}$.*

Proof. By Lemma 4.1.5, we have

$$1 = (\sum_{i=0}^{s} p_i)^2 = \sum_{i=0}^{s} p_i^2 \mod J_{grad}.$$

Thus $f = \sum_{i=0}^{s} f\, p_i^2 \mod J_{grad}$.

By Lemma 4.1.6, $-1 \in \mathcal{P}^+(C) + (\bigcap_{j>r} J_j^{\mathbf{x}})$ so that $f = \frac{1}{4}((f+1)^2 - (f-1)^2) \in \mathcal{P}^+(C) + \cap_{j>r} J_j^{\mathbf{x}}$ and

$$\sum_{i>r} f\, p_i^2 \in \mathcal{P}^+(C) + \bigcap_{j=0}^{s} J_j^{\mathbf{x}} = \mathcal{P}^+(C) + J^{\mathbf{x}}_{grad}. \tag{4.1}$$

As the polynomial $(f - f_i)\, p_i^2$ vanishes on V_{grad}, we deduce that

$$f = \sum_{i=0}^{r} f_i\, p_i^2 + \sum_{i=r+1}^{s} f\, p_i^2 + \sqrt{J^{\mathbf{x}}_{grad}} = \sum_{i=0}^{r} f_i\, p_i^2 + \mathcal{P}^+(C) + \sqrt{J^{\mathbf{x}}_{grad}},$$

which proves the first point.

If $f \geq 0$ on S_{KKT}, then $f_i \geq 0$ for $i = 0, \ldots, r$ and $\sum_{i=0}^{r} f_i\, p_i^2 \in \mathcal{P}^+(C)$ so that

$$f \in \mathcal{P}^+(C) + \sqrt{J^{\mathbf{x}}_{grad}},$$

which proves the second point.

If $f > 0$ on S_{KKT} by Lemma 4.1.8, we have $f p_i^2 = q_i^2 \mod J^{\mathbf{x}}_{grad}$ with $q_i \in \mathbb{R}[\mathbf{x}]$, which shows that

$$\sum_{i=0}^{r} f\, p_i^2 = \sum_{i=0}^{r} q_i^2 \mod J^{\mathbf{x}}_{grad}$$

Therefore, $\sum_{i=0}^{r} f\, p_i^2 \in \mathcal{P}^+(C) + J^{\mathbf{x}}_{grad}$ and $f \in \mathcal{P}^+(C) + J^{\mathbf{x}}_{grad}$ by (4.1), which proves the third point. ∎

This theorem involves only polynomials in $\mathbb{R}[\mathbf{x}]$ and the points (2) and (3) generalize results of [Demmel 2007] on the representation of positive polynomials.

Let us give now a refinement of Theorem 4.1.9 with a control of the degrees of the polynomials involved in the representation of f as an element of $\mathcal{P}^+(C) + J^{\mathbf{x}}_{grad}$.

Theorem 4.1.10 *Let $C \subset \mathbb{R}[\mathbf{x}]$ be a set of constraints such that $\mathcal{V}(C^0) \subset V^{\mathbf{x}}_{grad}$ and $\mathcal{S}^+(C) = \mathcal{S}^+(\mathbf{g})$. If $f \geq 0$ on S_{KKT}, then there exists t_0 such that $\forall \varepsilon > 0$,*

$$f + \varepsilon \in \mathcal{P}_{t_0}(C).$$

Proof. Let $J_{grad} = (C^0) \cap I_{grad} \subset \mathbb{R}[\mathbf{z}]$, so that $\mathcal{V}(J_{grad}) = V_{grad}$ since $\mathcal{V}(C^0) \subset V^{\mathbf{x}}_{grad}$. Using the decomposition (4.1) obtained in the proof of Theorem 4.1.9, we can choose $t_0' \in \mathbb{N}$ and $t_0 \geq t_0' \in \mathbb{N}$ big enough such that $\deg(p_i) \leq t_0/2$ and

$$\sum_{i>r} f\, p_i^2 \in \mathcal{P}^+_{t_0'}(C) + \ J_{grad} \cap \mathbb{R}[\mathbf{x}]_{t_0'} \subset \mathcal{P}_{t_0}(C),$$

since $J^{\mathbf{x}}_{grad} = (C^0) \cap I^{\mathbf{x}}_{grad} \subset (C^0)$. Then $\forall \varepsilon > 0$,

$$\sum_{i>r} (f+\varepsilon)\, p_i^2 = \sum_{i>r} f\, p_i^2 + \sum_{i>r} \varepsilon\, p_i^2 \in \mathcal{P}_{t_0}(C). \tag{4.2}$$

As $\forall\ \varepsilon > 0,\ f + \varepsilon > 0$ on S_{KKT}, i.e, $f_i + \varepsilon > 0$ for $i = 0, \ldots, r$, we deduce from Lemma 4.1.8 that if t_0 is big enough, we have

$$(f+\varepsilon)\, p_i^2 = q_i^2 \mod \langle C^0 \,|\, t_0 \rangle \cap \mathbb{R}[\mathbf{x}] \tag{4.3}$$

with $deg(q_i) \leq t_0/2$ for $i = 0, \ldots, r$.

Since $1 - \sum_{i=0}^{s} p_i^2 = 0 \mod (C^0)$, we can choose t_0 big enough so that

$$(f+\varepsilon) - \sum_{i=0}^{s} (f+\varepsilon)\, p_i^2 \in \langle C^0 \,|\, t_0 \rangle \cap \mathbb{R}[\mathbf{x}]. \tag{4.4}$$

From Equations (4.2), (4.3), (4.4), we deduce that if $t_0 \in \mathbb{N}$ is big enough, $\forall \varepsilon > 0$

$$f + \varepsilon \in \mathcal{P}_{t_0}(C),$$

which concludes the proof of the theorem. ∎

4.2 Finite convergence

In this Section, we show that the sequence of relaxation problems attains its limit in a finite number of steps and that the minimizer ideal can be recovered from an optimal solution of the corresponding relaxation problem. We recall the following notation:

- $f^* = \inf_{\mathbf{x}\in S_{KKT}} f(\mathbf{x})$
- $S_{min} = \{\mathbf{x}^* \in S_{KKT} \mid f(\mathbf{x}^*) = f^*\}$
- $\mathcal{L}_t(C) := \{\Lambda \in \mathbb{R}[\mathbf{x}]^*_{2t} \mid \Lambda(p) \geq 0,\ \forall p \in \mathcal{P}_t(C), \Lambda(1) = 1\}$.

We first show that $S_{min} = \emptyset$ can be detected from an adapted relaxation sequence:

Proposition 4.2.1 *Let $C = (C^0; C^+)$ be a set of constraints of $\mathbb{R}[\mathbf{x}]$, such that $S_{min} \subset \mathcal{S}(C)$ and $\mathcal{V}(C^0) \subset V^{\mathbf{x}}_{grad}$ and $C^+ = \mathbf{g}^+$. Then $S_{min} = \emptyset$, if and only if, there exits $t_0 \in \mathbb{N}$ such that $\forall t \geq t_0$, $\mathcal{L}_t(C) = \emptyset$.*

Proof. Let $J_{grad} = (C^0) \cap I_{grad}$ and let C' be a set of constraints such that $(C'^0) = J_{grad} \cap \mathbb{R}[\mathbf{x}] = J^{\mathbf{x}}_{grad}$ and $C'^+ = \mathbf{g}^+$ be a finite set. By hypothesis, $\mathcal{V}(J_{grad}) = V_{grad}$. We deduce from Corollary 4.1.7 that if $S_{min} = \emptyset$, then

$$-1 \in \mathcal{P}^+(C') + (C'^0) \subset \mathcal{P}(C) = \cup_{t\in\mathbb{N}} \mathcal{P}_t(C).$$

Thus there exists t_0 such that $-1 \in \mathcal{P}_t(C)$ for $t \geq t_0$, which implies that $\mathcal{L}_t(C) = \emptyset$, since if there exists $\Lambda \in \mathcal{L}_t(C)$, then $\Lambda(1) = 1$ and $\Lambda(-1) \geq 0$.

Conversely, suppose that $S_{min} \neq \emptyset$ contains a point $\mathbf{x}^*$. As $S_{min} \subset \mathcal{S}(C)$, for all $t \in \mathbb{N}$ the evaluation $\underline{\mathbf{1}}_{\mathbf{x}^*}$ at $\mathbf{x}^*$ restricted to $\mathbb{R}[\mathbf{x}]_{2t}$ is an element of $\mathcal{L}_t(C) \neq \emptyset$. ∎

This proposition gives a way to check whether $S_{min} = \emptyset$, using the relaxation sequence $\mathcal{L}_t(C)$. We are now going to analyse the case where f has KKT minimizers on S.

From now on, we assume that $S_{min} \neq \emptyset$.

First, we recall a property similar to [Lasserre 2008, Claim 4.7]:

Proposition 4.2.2 *Let $C = (C^0; C^+)$ be a set of constraints of $\mathbb{R}[\mathbf{x}]$. There exists $t_0 \in \mathbb{N}$ such that $\forall t \geq t_0$, $\forall \Lambda \in \mathcal{L}_t(C)$, $\sqrt[C^+]{C^0} \subset (\ker M^t_\Lambda)$.*

Proof. Let $C^0 = \{g_1, \ldots, g_l\}$ and let $q_1, \ldots, q_k$ be generators of $J := \sqrt[C^+]{C^0}$. By the Positivestllensatz, for $j \in 1, \ldots, k$, there exist $m_j \in \mathbb{N}^*$ and polynomials $u_r^{(j)} \in \mathbb{R}[\mathbf{x}]$ and $\sigma_j \in \mathcal{P}^+(C)$ such that

$$q_j^{2m_j} + \sigma_j = \sum_{r=1}^{l} u_r^{(j)} g_r.$$

Let us take $t_0 \in \mathbb{N}$ big enough such that $u_r^{(j)} g_r \in \langle C \,|\, t_0 \rangle$ and $\sigma_j \in \mathcal{P}_{t_0}^+(C)$. Then for all $t \geq t_0$ and all $\Lambda \in \mathcal{L}_t(C)$, we have $\Lambda(u_r^{(j)} g_r) = 0$, $\Lambda(q_j^{2m_j}) \geq 0$, $\Lambda(\sigma_j) \geq 0$ and $\Lambda(q_j^{2m_j}) + \Lambda(\sigma_j) = 0$, which implies that $\Lambda(q_j^{2m_j}) = 0$ and $q_j \in \ker H_\Lambda^t$. This proves that $(q_1, \ldots, q_k) = J \subset (\ker H_\Lambda^t)$. ∎

Remark 4.2.3 *With the same arguments, we can show that for any $t' \in \mathbb{N}$, there exists $t'_0 \geq t'$ such that $\forall t \geq t'_0$, $\forall \Lambda \in \mathcal{L}_t(C)$,*

$$\langle Q \,|\, t' \rangle \subset \ker H_\Lambda^t,$$

where $Q = \{q_1, \ldots, q_k\}$ generates $J = \sqrt[C^+]{C^0}$.

The next result shows that in the sequence of optimization problems that we consider, the minimum of f on S_{KKT} is reached from some degree.

Theorem 4.2.4 *Let C be a set of constraints of $\mathbb{R}[\mathbf{x}]$ such that $S_{min} \subset \mathcal{S}(C) \subset V_{KKT}^{\mathbf{x},\mathbb{R}}$. There exists $t_1 \geq 0$ such that $\forall t \geq t_1$,*

1. *$f_{t,C}^\mu = f^*$ is reached for some $\Lambda^* \in \mathcal{L}_t(C)$,*
2. *$\forall \Lambda^* \in \mathcal{L}_t(C)$ with $\Lambda^*(f) = f_{t,C}^\mu = f^*$, we have $p_i \in \ker H_{\Lambda^*}^t$, $\forall i = 1, \ldots, r$,*
3. *if $\mathcal{V}(C^0) \subset V_{grad}^{\mathbf{x}}$ then $f_{t,C}^{sos} = f_{t,C}^\mu = f^*$.*

Proof. By Theorem 4.1.9(1) applied to $f - f^*$, we can write

$$f - f^* = \sum_{i=1}^{r} (f_i - f^*)\, p_i^2 + h + g.$$

with $h \in \mathcal{P}^+(C)$ and $g \in \sqrt{I_{grad}} \cap \mathbb{R}[\mathbf{x}] = \sqrt{I_{KKT}} \cap \mathbb{R}[\mathbf{x}] \subset \sqrt[\mathbb{R}]{I_{KKT}} \cap \mathbb{R}[\mathbf{x}]$ (by Proposition 2.2.4). Since $\mathcal{S}(C) \subset V_{KKT}^{\mathbf{x},\mathbb{R}} = \overline{\pi^{\mathbf{x}}(V_{KKT}^{\mathbb{R}})}$, we have $\sqrt[\mathbb{R}]{I_{KKT}} \cap \mathbb{R}[\mathbf{x}] \subset \mathcal{I}(\mathcal{S}(C)) = \sqrt[C^+]{(C^0)}$ by the Positivstellensatz. We deduce that $g \in$

$\sqrt[C^+]{(C^0)}$. By proposition 4.2.2, there exists $t_1 \geq t_0$ such that for all $t \geq t_1$, for all $\Lambda \in \mathcal{L}_t(C)$, $\Lambda(g) = 0$, $\Lambda(h) \geq 0$.

Let us fix $t \geq t_1$ and $\Lambda^* \in \mathcal{L}_t(C)$ such that $\Lambda^*(f) = f^\mu_{t,C}$. Then

$$\Lambda^*(f - f^*) = \sum_{i=1}^{r} (f_i - f^*)\Lambda^*(p_i^2) + \Lambda^*(h).$$

As $f_i - f^* = f_i - f_0 > 0$ for $i = 1, .., r$, $\Lambda^*(p_i^2) \geq 0$ and $\Lambda^*(h) \geq 0$ ($h \in \mathcal{P}_t^+(C)$), we deduce that $\Lambda^*(f - f^*) = \Lambda^*(f) - f^* \geq 0$.

As $\emptyset \neq S_{min} \subset \mathcal{S}(C)$, we have $\Lambda^*(f) = f^\mu_{t,C} =\leq f^*$ (by Proposition 3.3.2), so that $\Lambda^*(f) = f^\mu_{t,C} = f^*$, which proves the first point. Hence for $i = 1, \ldots, r$, $\Lambda^*(p_i^2) = 0$ and $p_i \in \ker H^t_{\Lambda^*}$, which proves the second point.

To prove that $f^{sos}_{t,C} = f^*$ when $\mathcal{V}(C^0) \subset V^{\mathbf{x}}_{grad}$, we apply Theorem 4.1.10 to $f - f^*$ which is positive on S_{KKT}. Let us take $J_{grad} = (C^0) \cap I_{grad} \subset \mathbb{R}[\mathbf{z}]$. We denote by $\tilde{C}$ the set of constraints such that $\tilde{C}^0$ is a finite family of generators of $J_{grad} \cap \mathbb{R}[\mathbf{x}]$ and $\tilde{C}^+ = C^+$.

By Theorem 4.1.10, there exists t_0 such that $\forall \varepsilon > 0$,

$$f - f^* + \varepsilon \in \mathcal{P}_{t_0}(\tilde{C}).$$

As $(\tilde{C}^0) = (C^0) \cap I_{grad} \subset (C^0)$, we can choose $t_1 \geq t_0$ such that $\langle \tilde{C} \,|\, t_0 \rangle \subset \langle C \,|\, t_1 \rangle$ and $\mathcal{P}_{t_0}(\tilde{C}) \subset \mathcal{P}_{t_1}(C)$.

Then $\forall t \geq t_1$, $f - f^* + \varepsilon \in \mathcal{P}_t(C)$. Hence by maximality, $\forall \varepsilon > 0, f^* - \varepsilon \leq f^{sos}_{t,C}$. We deduce that $f^* \leq f^{sos}_{t,C}$, which implies that $f^{sos}_{t,C} = f^\mu_{t,C} = f^*$ and proves the third point. ∎

As for the construction of generators of $\sqrt[C^+]{I_{KKT}}$ (Proposition 4.2.2), we can construct generators of I_{min} from the kernel of a truncated Hankel operator associated to any linear form which minimizes f, using the following propositions:

Proposition 4.2.5 $I_{min} = (p_1, ..., p_r) + \sqrt[C^+]{I^{\mathbf{x}}_{KKT}}$.

Proof. First of all, we prove that $I^{\mathbf{z}}_{min} = (p_1, ..., p_r) + \sqrt[C^+]{I_{grad}} = (p_1, ..., p_r) + \sqrt[\mathbb{R}]{I_{grad}}$.

Using the decomposition of Lemma 4.1.2 and the polynomials p_i of Lemma 4.1.5, we have

$$V^{\mathbb{R}}_{grad} = (V_0 \cup V_1 \cup \cdots \cup V_s) \cap \mathbb{R}^{n+n_1+2\,n_2} = V_0^{\mathbb{R}} \cup \cdots \cup V_r^{\mathbb{R}},$$

By construction, $\mathcal{I}(V_0^{\mathbb{R}}) = I^{\mathbf{z}}_{min}$, $p_i(V_0^{\mathbb{R}}) = 0$ for $i = 1, \ldots, s$ and $p_i \in \mathbb{R}[\mathbf{x}]$ for $i = 0, \ldots, r$. This implies that $p_i \in I^{\mathbf{z}}_{min}$ for $i = 1, \ldots, r$.

As $V_0^{\mathbb{R}} \subset V_{grad}^{\mathbb{R}}$, we also have ${}^{c+}\sqrt{I_{grad}} \subset I_{min}^{\mathbf{z}}$.

We have proved so far that $(p_1, \ldots, p_r) + {}^{c+}\sqrt{I_{grad}} \subset I_{min}^{\mathbf{z}}$. In order to prove the reverse inclusion, we denote by $q_1, \ldots, q_m$ a family of generators of the ideal $I_{min}^{\mathbf{z}}$. Take one of these generators q_j $(1 \leq j \leq m)$. By construction, $q_j\, p_0(V_0^{\mathbb{R}}) = 0$ and $q_j p_0(V_i^{\mathbb{R}}) = 0$ for $i = 1, \ldots, r$, which implies that $q_j p_0 \in {}^{c+}\sqrt{I_{grad}}$.

By Lemma 4.1.5, we have the decomposition

$$q_j \equiv q_j(p_0 + p_1 + \cdots + p_s) \mod I_{grad} \subset {}^{c+}\sqrt{I_{grad}}.$$

Moreover $(p_{r+1} + \cdots + p_s) \in \mathbb{R}[\mathbf{z}]$ and vanishes on $V_k^{\mathbb{R}}$ for $k = 0, \ldots, r$. Thus $(p_{r+1} + \cdots + p_s) \in {}^{c+}\sqrt{I_{grad}}$ and we deduce that $q_j \in (p_1, \ldots, p_r) + {}^{c+}\sqrt{I_{grad}}$. This proves the other inclusion and the first equality.

As $V_{grad}^{\mathbb{R}} = V_{grad}^{\mathbb{R}} \cap \mathcal{S}^+(C)$ (Lemma 2.2.5), by the Positivstellensatz, ${}^{c+}\sqrt{I_{grad}} = {}^{\mathbb{R}}\sqrt{I_{grad}}$, which proves the second equality.

By the Positivstellensatz and Lemma 2.2.5, we have

$${}^{c+}\sqrt{I_{grad}} \cap \mathbb{R}[\mathbf{x}] = {}^{\mathbb{R}}\sqrt{I_{grad}} \cap \mathbb{R}[\mathbf{x}] = \mathcal{I}(\pi^{\mathbf{x}}(V_{grad}^{\mathbb{R}})) = \mathcal{I}(\pi^{\mathbf{x}}(V_{KKT})^{\mathbb{R}} \cap \mathcal{S}^+(C)) = {}^{c+}\sqrt{I_{KKT}^{\mathbf{x}}}.$$

and

$$I_{min} = I_{min}^{\mathbf{z}} \cap \mathbb{R}[\mathbf{x}] = (p_1, ..., p_r) \cap \mathbb{R}[\mathbf{x}] + {}^{c+}\sqrt{I_{grad}} \cap \mathbb{R}[\mathbf{x}] = (p_1, ..., p_r) + {}^{c+}\sqrt{I_{KKT}^{\mathbf{x}}}.$$

which proves the equality. ∎

Theorem 4.2.6 *For C set of constraints of $\mathbb{R}[\mathbf{x}]$ with $S_{min} \subset \mathcal{S}(C) \subset V_{KKT}^{\mathbf{x},\mathbb{R}}$, there exists $t_2 \in \mathbb{N}$ such that $\forall t \geq t_2$, for $\Lambda^* \in \mathcal{L}_t(C)$ with $\Lambda^*(f) = f_{t,C}^{\mu}$, we have $I_{min} \subset (\ker H_{\Lambda^*}^t)$.*

Proof. To prove the inclusion we take $t_2 = \max\{t_0, t_1\}$ and we combine Proposition 4.2.5 with Proposition 4.2.2 for $C \subset \mathbb{R}[\mathbf{x}]$ and Theorem 4.2.4. ∎

We introduce now the notion of *optimal linear form for f*. Such a linear form allows us to compute I_{min}.

Proposition 4.2.7 *For $\Lambda^* \in \mathcal{L}_t(C)$ and $p \in \mathbb{R}[\mathbf{x}]$, the following assertions are equivalent:*

(i) $\operatorname{rank} H_{\Lambda^*}^t = \max_{\Lambda \in \mathcal{L}_t(C), \Lambda(p) = p_{t,C}^{\mu}} \operatorname{rank} H_{\Lambda}^t$.

(ii) $\forall \Lambda \in \mathcal{L}_t(C)$ *such that* $\Lambda(p) = p_{t,C}^{\mu}$, $\ker H_{\Lambda^*}^t \subset \ker H_{\Lambda}^t$.

We say that $\Lambda^* \in \mathcal{L}_t(C)$ is optimal for p if it satisfies one of the equivalent conditions (i)-(ii).

A proof of this proposition can be found in [Lasserre 2012](Proposition 4.7).

Remark 4.2.8 *A linear form $\Lambda^* \in \mathcal{L}_t(C)$ optimal for p can be computed by solving a Semi-Definite Programming problem by an interior point method [Lasserre 2009a]. In this case, the solution Λ^* obtained by convex optimization is in the interior of the face of linear forms that minimize f.*

The next result, which refines Theorem 4.2.6, shows that only elements in I_{min} are involved in the kernel of a truncated Hankel operator associated to an optimal linear form for f.

Theorem 4.2.9 *Let $t \in \mathbb{N}$ such that $f \in \mathbb{R}[\mathbf{x}]_{2t}$ and let $C \subset \mathbb{R}[\mathbf{x}]_{2t}$ with $S_{min} \subset \mathcal{S}(C)$. If $\Lambda^* \in \mathcal{L}_t(C)$ is optimal for f and such that $\Lambda^*(f) = f^*$, then* $\ker H^t_{\Lambda^*} \subset I_{min}$.

Proof. It is similar to proof of Theorem 4.9 in [Lasserre 2012]. Let $p \in \ker H^t_{\Lambda^*}$ and $\mathbf{x}^* \in S_{min}$: $f(\mathbf{x}^*) = f^*$. Let $\underline{\mathbf{1}}_{\mathbf{x}^*}$ denotes the evaluation at $\mathbf{x}^*$ restricted to $\mathbb{R}[\mathbf{x}]_{2t}$. Our objective is to show that $p(\mathbf{x}^*) = 0$. Suppose for contradiction that $p(\mathbf{x}^*) \neq 0$. We know that $\underline{\mathbf{1}}_{\mathbf{x}^*} \in \mathcal{L}_t(C)$ since $S_{min} \subset \mathcal{S}(C)$ and $\underline{\mathbf{1}}_{\mathbf{x}^*}(f) = f(\mathbf{x}^*) = f^*$. We define $\tilde{\Lambda} = \frac{1}{2}(\Lambda^* + \underline{\mathbf{1}}_{\mathbf{x}^*})$. By construction, $\tilde{\Lambda} \in \mathcal{L}_t(C)$ and $\tilde{\Lambda}(f) = \frac{1}{2}(\Lambda^*(f) + \underline{\mathbf{1}}_{\mathbf{x}^*}(f)) = \frac{1}{2}(\Lambda^*(f) + f(\mathbf{x}^*)) = f^*$. As $p \in \ker H^t_{\Lambda^*}$,

$$\tilde{\Lambda}(p^2) = \frac{1}{2}(\Lambda^*(p^2) + \underline{\mathbf{1}}_{\mathbf{x}^*}(p^2)) = \frac{1}{2}p^2(\mathbf{x}^*) \neq 0$$

thus $p \in \ker H^t_{\Lambda^*} \setminus \ker H^t_{\tilde{\Lambda}}$ and by the maximality of the rank of $H^t_{\Lambda^*}$, $\ker H^t_{\tilde{\Lambda}} \not\subset \ker H^t_{\Lambda^*}$. Hence there exits $\tilde{p} \in \ker H^t_{\tilde{\Lambda}} \setminus \ker H^t_{\Lambda^*}$. Then $0 = H^t_{\tilde{\Lambda}}(\tilde{p}) = \frac{1}{2}(H^t_{\Lambda^*}(\tilde{p}) + H^t_{\underline{\mathbf{1}}_{\mathbf{x}^*}}(\tilde{p})) = \frac{1}{2}(H^t_{\Lambda^*}(\tilde{p}) + \tilde{p}(\mathbf{x}^*) \cdot \underline{\mathbf{1}}_{\mathbf{x}^*})$. As $H^t_{\Lambda^*}(\tilde{p}) \neq 0$ implies $\tilde{p}(\mathbf{x}^*) \neq 0$. On the other hand,

$$0 = H^t_{\tilde{\Lambda}}(\tilde{p})(p) = \tilde{\Lambda}(p\tilde{p}) = \frac{1}{2}(\Lambda^*(p\tilde{p}) + p(\mathbf{x}^*)\tilde{p}(\mathbf{x}^*)) =$$

$$= \frac{1}{2}(H^t_{\Lambda^*}(p)(\tilde{p}) + p(\mathbf{x}^*)\tilde{p}(\mathbf{x}^*)).$$

As $p \in \ker H^t_{\Lambda^*}$, we have

$$0 = H^t_{\tilde{\Lambda}}(\tilde{p})(p) = \frac{1}{2}(p(\mathbf{x}^*)\tilde{p}(\mathbf{x}^*)).$$

As $\tilde{p}(\mathbf{x}^*) \neq 0$ and we have supposed that $p(\mathbf{x}^*) \neq 0$, it yields a contradiction. ■

The last result of this Section shows that an optimal linear form for f yields the generators of the minimizer ideal I_{min} in high enough degree.

Theorem 4.2.10 *Let $\mathbf{g} \subset \mathbb{R}[\mathbf{x}]$ be a set of constraints with $S_{min} \neq \emptyset$. For a set of constraints $C \subset \mathbb{R}[\mathbf{x}]$ with $S_{min} \subset \mathcal{S}(C) \subset V_{KKT}^{\mathbf{x},\mathbb{R}}$ (asociated to $\mathbf{g}$), there exists $t_2 \in \mathbb{N}$ (defined in Theorem 4.2.6) such that $\forall t \geq t_2$,*

- *$f_{t,C}^{\mu} = \min_{\mathbf{x} \in S_{KKT}} f(\mathbf{x})$ is reached for some $\Lambda^* \in \mathcal{L}_t(C)$,*
- *$\forall \Lambda^* \in \mathcal{L}_t(C)$ optimal for f, we have $\Lambda^*(f) = f^*$ and $(\ker H_{\Lambda^*}^t) = I_{min}$,*
- *if $\mathcal{V}(C^0) \subset V_{grad}^{\mathbf{x}}$ then $f_{t,C}^{sos} = f_{t,C}^{\mu} = f^*$.*

Proof. We obtain the result as a consequence of Theorem 4.2.4, Theorem 4.2.6 and Theorem 4.2.9. ■

The same results hold if we replace C by any other finite set defining a *real* variety such that $S_{min} \subset \mathcal{S}(C) \subset V_{KKT}^{\mathbf{x},\mathbb{R}}$.

Remark 4.2.11 *We can also replace the initial set of constraints $\mathbf{g}$ by any other set $\tilde{\mathbf{g}}$ defining the same semi-algebraic set $S = \mathcal{S}(\mathbf{g}) = \mathcal{S}(\tilde{\mathbf{g}})$ and consider the KKT variety associated to $\tilde{\mathbf{g}}$.*

4.3 Consequences

Let us describe now some consequences of these results in specific cases, which have been previously studied.

4.3.1 Global optimization

We consider here the case $n_1 = n_2 = 0$. Theorem 4.1.9 implies the following result (compare with [Nie 2006]):

Theorem 4.3.1 *Let $f \in \mathbb{R}[\mathbf{x}]$.*

1. *If f is positive at its real critical points, then $f \in sos + \sqrt{(\frac{\partial f}{\partial_{x_1}}, \ldots, \frac{\partial f}{\partial_{x_n}})}$.*
2. *If f is strictly positive at its real critical points, then $f \in sos + (\frac{\partial f}{\partial_{x_1}}, \ldots, \frac{\partial f}{\partial_{x_n}})$.*

In particular, if there is no real critical value, then $f \in sos + (\frac{\partial f}{\partial_{x_1}}, \ldots, \frac{\partial f}{\partial_{x_n}})$.

A consequence of Proposition 4.2.1 and Theorem 4.2.10 is the following:

Theorem 4.3.2 *Let $f \in \mathbb{R}[\mathbf{x}]$ and $C = \{\frac{\partial f}{\partial_{x_1}}, \ldots, \frac{\partial f}{\partial_{x_n}}\}$. Then, there exists $t_0 \in \mathbb{N}$, such that $\forall t \geq t_0$ either $\mathcal{L}_t(C) = \emptyset$ and $S_{min} = \emptyset$ or*

1. *$f_{t,C}^{sos} = f_{t,C}^{\mu} = f^* = \min_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x})$ is reached for some $\Lambda^* \in \mathcal{L}_t(C)$,*
2. *$\forall \Lambda^* \in \mathcal{L}_t(C)$ optimal for f, $\ker H_{\Lambda^*}^t$ generates I_{min}.*

The first point of this theorem can also be found in [Nie 2006].

4.3.2 General case

A direct consequence of Proposition 4.2.1 and Theorem 4.2.10 is the following:

Theorem 4.3.3 *Let $C \subset \mathbb{R}[\mathbf{x}]$ be a set of constraints such that*

- *$(C^0) = I_{KKT} \cap \mathbb{R}[\mathbf{x}]$,*
- *$C^+ = \mathbf{g}^+$.*

Then there exists $t_0 \in \mathbb{N}$ such that $\forall t \geq t_0$, either $\mathcal{L}_t(C) = \emptyset$ and $S_{min} = \emptyset$ or

- *$f_{t,C}^{sos} = f_{t,C}^{\mu} = min_{\mathbf{x}\in S_{KKT}} f(\mathbf{x})$ is reached for some $\Lambda^* \in \mathcal{L}_t(C)$,*
- *$\forall \Lambda^* \in \mathcal{L}_t(C)$ optimal for f, we have $\Lambda^*(f) = f^*$ and $(\ker H_{\Lambda^*}^t) = I_{min}$.*

The set C^0 is constructed so that $\mathcal{V}(C^0) = V_{KKT}^{\mathbf{x}} = V_{grad}^{\mathbf{x}}$. As we have seen, the weaker condition $S_{min} \subset \mathcal{S}(C) \subset V_{KKT}^{\mathbf{x}}$ is sufficient to have an exact relaxation sequence.

The generators C^0 of $I_{KKT} \cap \mathbb{R}[\mathbf{x}]$ can be computed by elimination techniques (for instance by Groebner basis computation with a product order on monomials [Cox 2005]).

4.3.3 Regular case

We consider here a semi-algebraic set S such that its defining constraints intersect properly. For any $\mathbf{x} \in \mathbb{C}^n$, let $\nu(\mathbf{x}) = \{j \in [1, n_2] \mid g_j^+(\mathbf{x}) = 0\}$.

Definition 4.3.4 *We say that a set of constraints $\mathbf{g} = (g_1^0, \ldots g_{n_1}^0 ; g_1^+, \ldots, g_{n_2}^+)$ is regular if for all points $\mathbf{x} \in \mathcal{S}(\mathbf{g})$ with $\nu(\mathbf{x}) = \{j_1, \ldots, j_k\}$, the vectors $\nabla g_1^0(\mathbf{x}), \ldots, \nabla g_{n_1}^0(\mathbf{x})$, $\nabla g_{j_1}^+(\mathbf{x}), \ldots$, $\nabla g_{j_k}^+(\mathbf{x})$ are linearly independent.*

This condition is used for instance in [Ha 2010]. It implies that $\forall \mathbf{x} \in S$, $|\nu(\mathbf{x})| \leq n - n_1$ and that $B_{\nu(\mathbf{x})}(\mathbf{x})$ is of rank $n_1 + |\nu(\mathbf{x})|$. A stronger condition, called the $\mathbb{C}$-regularity, corresponds to sets of constraints such that $\forall \mathbf{x} \in \mathbb{C}^n$, $B_{\nu(\mathbf{x})}(\mathbf{x})$ is of rank $n_1 + |\nu(\mathbf{x})|$. This condition is used for instance in [Nie 2011]. It is satisfied for semi-algebraic sets defined by "generic" constraints when $n_1 \leq n$ as shown in [Nie 2011].

If $\mathbf{g}$ is regular, then for all points $\mathbf{x}$ in S the rank of $B_{\nu(\mathbf{x})}(\mathbf{x})$ is $n_1 + |\nu(\mathbf{x})|$ and $S_{sing} = \emptyset$. The decomposition (2.16) implies that $S_{FJ} = S_{KKT}$ and that all minimizer points of f on S are KKT points. If moreover $\mathbf{g}$ is $\mathbb{C}$-regular, then $V^{\mathbf{x}}_{FJ} = \mathcal{V}(\Gamma_{FJ}) = V^{\mathbf{x}}_{KKT} = V^{\mathbf{x}}_{grad}$.

We deduce from Theorem 4.2.10 the following result:

Theorem 4.3.5 *Let $\mathbf{g} \subset \mathbb{R}[\mathbf{x}]$ be a regular set of constraints and let $C \subset \mathbb{R}[\mathbf{x}]$ be the set of constraints such that*

- *$C^0 = \Gamma_{FJ}$ defined in (2.13) (resp. $C^0 = \Phi_{FJ}$ defined in (2.14)),*
- *$C^+ = \mathbf{g}^+$.*

Suppose that $\min_{\mathbf{x} \in \mathcal{S}(\mathbf{g})} f(\mathbf{x})$ is reached at some point of $\mathcal{S}(\mathbf{g})$. Then, there exists $t_0 \in \mathbb{N}$ such that $\forall t \geq t_0$,

1. *$f^{\mu}_{t,C} = f^* = \min_{\mathbf{x} \in \mathcal{S}(\mathbf{g})} f(\mathbf{x})$ is reached for some $\Lambda^* \in \mathcal{L}_t(C)$,*
2. *$\forall \Lambda^* \in \mathcal{L}_t(C)$ optimal for f, $\ker H^t_{\Lambda^*}$ generates $I^{\mathbf{x}}_{min}$,*
3. *If $\mathbf{g}$ is $\mathbb{C}$-regular and $C^0 = \Gamma_{FJ}$, then $f^{sos}_{t,C} = f^{\mu}_{t,C} = f^*$.*

By Lemma 2.3.5 and Remark 2.3.6, C is constructed so that $S_{min} \subset \mathcal{S}(C) = S_{KKT} \subset V^{\mathbf{x},\mathbb{R}}_{KKT}$.

Points (1) and (3) are proved for $C^0 = \Gamma_{FJ}$ in [Nie 2011] under the condition that $\mathbf{g}$ is $\mathbb{C}$-regular. These points can also be found in [Ha 2010] for $C^0 = \mathbf{g}^0 \cup \Phi_{FJ}$ under the condition that $\mathbf{g}$ is regular (but a problem appears in the proof: the vanishing of the polynomials Φ_{FJ} at a point $\mathbf{x} \in \mathbb{C}^n$ does not imply that $\operatorname{rank} A_{\nu(\mathbf{x})}(\mathbf{x}) < n_1 + |\nu(\mathbf{x})|$).

In this case, the relaxation constructed with Γ_{FJ} (or Φ_{FJ}) is exact and can be used to compute the minimizer ideals of f on the semi-algebraic set S.

4.3.4 Zero dimensional real variety

Let $\mathbf{g} \subset \mathbb{R}[\mathbf{x}]$ be a set of constraints such that $\mathcal{V}^{\mathbb{R}}(\mathbf{g}^0)$ is finite and let $S := \mathcal{S}(\mathbf{g})$. By Remark 4.2.11, we can assume that S is defined by a set of constraints $\tilde{\mathbf{g}}$ such that $(\tilde{\mathbf{g}}^0)$ is radical. Then $\forall \mathbf{x} \in \mathcal{V}(\mathbf{g}^0) = \mathcal{V}(\tilde{\mathbf{g}}^0)$, the

Jacobian matrix $\tilde{B}_{\nu(\mathbf{x})}(\mathbf{x})$ associated to $\tilde{\mathbf{g}}^0$ is of rank n. Therefore we have $\mathcal{V}(\mathbf{g}^0) = \mathcal{V}(\tilde{\mathbf{g}}^0) = V^{\mathbf{x}}_{KKT}(\tilde{\mathbf{g}}) = V^{\mathbf{x}}_{grad}(\tilde{\mathbf{g}})$ and any point of S is a KKT-point: $S = S_{FJ} = S_{KKT}$. Consequently, if we take $C = \mathbf{g}$, $S_{min} \subset \mathcal{S}(C) \subset V^{\mathbf{x},\mathbb{R}}_{KKT}(\tilde{\mathbf{g}})$, we deduce from Theorem 4.2.10 the following result:

Theorem 4.3.6 *Let $\mathbf{g} = (\mathbf{g}^0, \mathbf{g}^+) \subset \mathbb{R}[\mathbf{x}]$ be a set of constraints such that $\mathcal{V}^{\mathbb{R}}(\mathbf{g}^0)$ is finite. Then there exists $t_0 \in \mathbb{N}$ such that $\forall t \geq t_0$,*

1. *$f^{sos}_{t,\mathbf{g}} = f^{\mu}_{t,\mathbf{g}} = f^* = \min_{\mathbf{x}\in\mathcal{S}} f(\mathbf{x})$ is reached for some $\Lambda^* \in \mathcal{L}_t(\mathbf{g})$,*

2. *$\forall \Lambda^* \in \mathcal{L}_t(\mathbf{g})$ optimal for f, $\ker H^t_{\Lambda^*}$ generates I_{min}.*

This answers an open question in [Laurent 2009a]. The first point was also solved in [Nie 2013a] using dedicated techniques.

4.3.5 Smooth real variety

We consider a set of constraints $\mathbf{g} = \{g^0_1, \ldots, g^0_{n_1}\} \subset \mathbb{R}[\mathbf{x}]$ such that $\mathcal{V}^{\mathbb{R}}(\mathbf{g}^0)$ is equidimensional smooth and $\mathbf{g}^+ = \emptyset$. This means that $S = \mathcal{S}(\mathbf{g}) = \mathcal{V}^{\mathbb{R}}(\mathbf{g}^0)$ is the union of irreducible components of the same dimension d and that for any point $\mathbf{x} \in S$, $B_\emptyset(\mathbf{x}) = [\nabla g^0_1(\mathbf{x}), \ldots, \nabla g^0_{n_1}(\mathbf{x})]$ is of rank $m = \dim S = n - d$. Therefore, $S_{sing} = \emptyset$. In this case, $\nabla f(\mathbf{x})$ is a linear combination of $\nabla g^0_1(\mathbf{x}), \ldots, \nabla g^0_{n_1}(\mathbf{x})$, if and only if, $\text{rank} A_\emptyset(\mathbf{x}) \leq r$.

The set Γ_{FJ} defined in (2.13) (or $C^0 = \mathbf{g}^0 \cup \Phi_{FJ}$ defined in (2.14)), or the union Δ^{n-d} of $\mathbf{g}^0$ and the set of $(n-d+1)\times(n-d+1)$ minors of the Jacobian matrix of $\{f, g^0_1, \ldots, g^0_{n_1}\}$, which contain the first column ∇f define the variety S_{KKT}.

We deduce from Theorem 4.2.10, the following result:

Theorem 4.3.7 *Let $\mathbf{g} = \{g^0_1, \ldots, g^0_{n_1}\} \subset \mathbb{R}[\mathbf{x}]$ such that $S = \mathcal{V}^{\mathbb{R}}(\mathbf{g})$ is an equidimensional and smooth variety of dimension d.*

Let $C \subset \mathbb{R}[\mathbf{x}]$ be the set of constraints such that $C^0 = \Gamma_{FJ}$ defined in (2.13) (or $C^0 = \Phi_{FJ}$ defined in (2.14), $C^0 = \Delta^{n-d}$) Then there exists $t_0 \in \mathbb{N}$ such that $\forall t \geq t_0$, either $\mathcal{L}_t(C) = \emptyset$ and $S_{min} = \emptyset$ or

1. *$f^{\mu}_{t,C} = f^* = \min_{\mathbf{x}\in S} f(\mathbf{x})$ is reached for some $\Lambda^* \in \mathcal{L}_t(C)$,*

2. *$\forall \Lambda^* \in \mathcal{L}_t(C)$ optimal for f, $\ker H^t_{\Lambda^*}$ generates I_{min}.*

4.3.6 Known minimum

In the case where we know the minimum f^* of f on the basic closed semi-algebraic set S, we take $\mathbf{g}'$ with $\mathbf{g}'^0 = \{\mathbf{g}^0, f - f^*\}$ and $\mathbf{g}'^+ = \mathbf{g}^+$. Let $S = \mathcal{S}(\mathbf{g})$, $S' = \mathcal{S}(\mathbf{g}')$. By construction $S_{min} \subset S'$ and $S' = S'_{KKT}$ and $\mathcal{V}(\mathbf{g}^0) \subset V^{\mathbf{x}}_{KKT}(\mathbf{g}'^0)$. Theorem 4.2.10 applied to $\mathbf{g}'$ implies the following result:

Theorem 4.3.8 *Let $\mathbf{g} = \{g_1^0, \ldots, g_{n_1}^0 g_1^+, \ldots, g_{n_2}^+\} \subset \mathbb{R}[\mathbf{x}]$. Let f^* be the minimum of f and $C \subset \mathbb{R}[\mathbf{x}]$ the set of constraints such that $C^0 = \{\mathbf{g}^0, f - f^*\}$ and $C^+ = \mathbf{g}^+$. Then there exists $t_0 \in \mathbb{N}$ such that $\forall t \geq t_0$,*

1. *$f^{sos}_{t,C} = f^{\mu}_{t,C} = f^* = \min_{\mathbf{x} \in \mathcal{S}(C)} f(\mathbf{x})$ is reached for some $\Lambda^* \in \mathcal{L}_{t,C}$,*
2. *$\forall \Lambda^* \in \mathcal{L}_t(C)$ optimal for f, $\ker H^t_{\Lambda^*}$ generates I_{min}.*

4.3.7 Radical computation.

In the case where $f = 0$, by Remark 2.4.3 all the points of S are KKT points and minimizers of f so that $S_{min} = S = S_{KKT}$. Moreover, $I^{\mathbf{x}}_{KKT} = (g_1^0, \ldots, g_{n_1}^0)$ since $F_1, \ldots, F_n, v_1 g_1^+, \ldots, v_{n_2} g_{n_2}^+$ are homogeneous of degree 1 in the variables $u_1, \ldots, u_{n_1}, v_1, \ldots, v_{n_2}$. We deduce the following result:

Theorem 4.3.9 *Let $\mathbf{g} = \{g_1^0, \ldots, g_{n_1}^0; g_1^+, \ldots, g_{n_2}^+\} \subset \mathbb{R}[\mathbf{x}]$. There exists $t_2 \in \mathbb{N}$ such that $\forall t \geq t_2$, $\forall \Lambda^* \in \mathcal{L}_t(\mathbf{g})$ optimal for 0, we have $(\ker H^t_{\Lambda^*}) = \mathcal{I}(\mathcal{S}) = \sqrt[\mathbf{g}^+]{(\mathbf{g}^0)}$.*

This gives a way to compute $\sqrt[C^+]{(C^0)}$ (see also [Ma 2013]), which generalizes the approach of [Lasserre 2009b], [Lasserre 2012] or [Rostalki 2009] to compute the real radical of an ideal.

Chapter 5

Border basis relaxation for polynomial optimization

In Chapter 3 we have seen the relaxation approach proposed by Lasserre in [Lasserre. 2001] which approximates the problem (2.1) by a sequence of finite dimensional convex optimization problems. These optimization problems can be formulated in terms of linear matrix inequalities on moment matrices associated to the set of monomials of degree $\leq t \in \mathbb{N}$ for increasing values of t. They can be solved by Semi-Definite Programming (SDP) techniques. The sequence of minima converges to the actual minimum f^* of the function under some hypotheses [Lasserre. 2001]. In some cases, the sequence even reaches the minimum f^* in a finite number of steps as we saw in Chapter 4. This approach proved to be particularly fruitful in many problems [Lasserre 2009a]. In contrast with numerical methods such as gradient descent methods, which converge to a local extrema but with no guaranty for the global solution.

From an algorithmic and computational perspective, some issues need however to be considered. The size of the SDP problems to be solved is a bottleneck of the method. This size is related to the number of monomials of degree $\leq t$ and is increasing exponentially with the number of variables and the degree t. Many SDP solvers are based on interior point methods which provide an approximation of the optimal moment sequence within a given precision in a polynomial time: namely $\mathcal{O}((p\,s^{3.5} + c\,p^2 s^{2.5} + c\,p^3 s^{0.5}) \log(\varepsilon^{-1}))$ arithmetic operations where $\varepsilon > 0$ is the precision of the approximation, s is bounding the size of the moment matrices, p is the number of parameters (usually of the order s^2) and c is the number of constraints [Nesterov 1994]. Thus reducing the size s or the number of parameters p can improve significantly the performance of these relaxation methods. Some recent works address this issue, using symmetries (see e.g. [Riener 2013]) or polynomial reduction (see e.g. [Lasserre 2012]). We use Border basis in order to obtain a polynomial reduction. As we know by [Mourrain 2005], border basis extend Groebner basis and they have more numerical stability.

While determining the minimum value of a polynomial function on a semi-algebraic set is important, computing the points where this minimum is reached if they exist, is also critical in many applications. Determining when

and how these minimizer points can be computed from the relaxation sequence is a problem that has been addressed for instance in [Henrion 2005, Nie 2012] using full moment matrices. This approach has been used for solving polynomial equations [Laurent 2007, Lasserre 2008, Lasserre 2009b, Lasserre 2009a].

This chapter is organized as follows. In the first section we give the definitions and main propositions and theorems about border basis. In the second section we propose a new method which combines Lasserre's SDP relaxation approach with polynomial algebra, in order to increase the efficiency of the optimization algorithm. Border basis computations are considered for their numerical stability [Mourrain 2005, Mourrain 2008]. In principle, any graded normal form techniques could be used here.

In the third section a new stopping criterion is given to detect when the relaxation sequence reaches the minimum, using a flat extension criterion from [Laurent 2009b]. We also provide a new algorithm to reconstruct a finite sum of weighted Dirac measures from a truncated sequence of moments. This reconstruction method can be used in other problems such as tensor decomposition [Brachat 2010] and multivariate sparse interpolation [Giesbrecht 2009].

In the last section we obtain a new algorithm for polynomial optimization be able to compute zero-dimensional minimizer ideals and the minimizer points, or zero-dimensional G-radical. We show also how the algoritm work on a detailed and on examples which are particular cases of Section 4.3 of the Chapter 4. As we will see in Chapter 6, the impact on the performance of the relaxation approach is significant.

5.1 Border basis

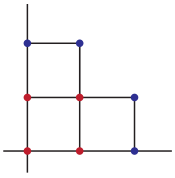
The eigenvalue method for solving polynomial equations from the Section 1.4 requires the knowledge of a basis of $\mathcal{A} = \mathbb{K}[\mathbf{x}]/I$ and an algorithm to compute the normal form of a polynomial with respect to this basis. In this Section we will recall a general method for computing such a basis and a method to reduce polynomials to their normal form.

Throughout $\mathcal{B} \subseteq \mathcal{M}$ is a finite set of monomials in n variables.

Definition 5.1.1 *Let $\mathcal{B}$ be a finite set of monomials in n variables.*

- *$\mathcal{B}$ is said to be* connected to 1 *if $1 \in \mathcal{B}$ and for every monomial $m \neq 1$ in $\mathcal{B}$, $m = x_{i_0} m'$ for some $i_0 \in [1, n]$ and $m' \in \mathcal{B}$.*
- *$\mathcal{B}^+ = \mathcal{B} \cup x_1\mathcal{B} \cup x_2\mathcal{B} \cdots \cup x_n\mathcal{B}$ is the prologantion de $\mathcal{B}$.*
- *$\partial\mathcal{B} = \mathcal{B}^+ \setminus \mathcal{B}$ is the border of $\mathcal{B}$.*

Example 5.1.2 *The monomial set* $\mathcal{B} = \{1, x, y, xy\}$ *is connected to 1 as show the figure (red points). Its prolongation is* $\mathcal{B}^+ = \mathcal{B} \cup x\mathcal{B} \cup y\mathcal{B} = \{1, x, y, xy, x^2, x^2y, xy^2, y^2\}$ *and its border is* $\partial\mathcal{B} = \mathcal{B}^+ \setminus \mathcal{B} = \{x^2, x^2y, xy^2, y^2\}$ *(blue points).*



Definition 5.1.3 *A rewriting family* F *for a (monomial) set* $\mathcal{B}$ *is a set of polynomials* $F = \{f_i\}_{i\in\mathcal{I}}$ *such that*

- $supp(f_i) \subseteq \mathcal{B}^+$,
- f_i *has exactly* **one** *monomial in* $\partial\mathcal{B}$*, denoted as* $\gamma(f_i)$ *and called the leading monomial of* f_i*. (The polynomial* f_i *is normalized so that the coefficient of* $\gamma(f_i)$ *is* 1*.)*
- *if* $\gamma(f_i) = \gamma(f_j)$ *then* $i = j$.

Definition 5.1.4 *We say that the rewriting family* F *is* graded *if* $\deg(\gamma(f)) = \deg(f)$ *for all* $f \in F$.

Definition 5.1.5 *A rewriting family* F *for* $\mathcal{B}$ *is said to be* complete *in degree* t *if it is graded and satisfies* $(\partial\mathcal{B})_t \subseteq \gamma(F)$*; that is, each monomial* $m \in \partial\mathcal{B}$ *of degree at most* t *is the leading monomial of some (necessarily unique)* $f \in F$.

Example 5.1.6 *For the example above for the set* $B = \{1, x, y, xy\}$*, the set of polynomials* $F = \{x^2 - 1, y^2 - 1, x^2y - y, y^2x - x\}$ *is a complete rewriting family of degree 3 for this set.*

Notice that a complete family F for $\mathcal{B}$ in degree t allow us to rewrite the monomials of $\mathcal{B}_t^+$ modulo F as elements of $\mathcal{B}_t$. This lead in fact, to the definition of the projection $\pi_{F,\mathcal{B}}$, asociated to a complete family for a set $\mathcal{B}$ connected to 1.

Definition 5.1.7 *Let* F *be a rewriting family for* $\mathcal{B}$*, complete in degree* t*. Let* $\pi_{F,\mathcal{B}}$ *be the projection on* $\langle\mathcal{B}\rangle$ *along* F *defined recursively on the monomials* $m \in \mathcal{M}_t$ *in the following way:*

- *if* $m \in \mathcal{B}_t$, *then* $\pi_{F,\mathcal{B}}(m) = m$,
- *if* $m \in (\partial\mathcal{B})_t$ $(= (\mathcal{B}^{[1]} \setminus \mathcal{B}^{[0]})_t)$, *then* $\pi_{F,\mathcal{B}}(m) = m - f$, *where* f *is the (unique) polynomial in* F *for which* $\gamma(f) = m$,
- *if* $m \in (\mathcal{B}^{[k]} \setminus \mathcal{B}^{[k-1]})_t$ *for some integer* $k \geq 2$, *write* $m = x_{i_0} m'$, *where* $m' \in \mathcal{B}^{[k-1]}$ *and* $i_0 \in [1, n]$ *is the smallest possible variable index for which such a decomposition exists, then* $\pi_{F,\mathcal{B}}(m) = \pi_{F,\mathcal{B}}(x_{i_0}\, \pi_{F,\mathcal{B}}(m'))$.

If F is a graded rewriting family, one can easily verify that $\deg(\pi_{F,\mathcal{B}}(m)) \leq \deg(m)$ for $m \in \mathcal{M}_t$. The map $\pi_{F,\mathcal{B}}$ extends by linearity to a linear map from $\mathbb{K}[\mathbf{x}]_t$ onto $\langle\mathcal{B}\rangle_t$. By construction, $f = \gamma(f) - \pi_{F,\mathcal{B}}(\gamma(f))$ and $\pi_{F,\mathcal{B}}(f) = 0$ for all $f \in F_t$. The next theorems show that, under some natural commutativity condition, the map $\pi_{F,\mathcal{B}}$ coincides with the linear projection from $\mathbb{K}[\mathbf{x}]_t$ onto $\langle\mathcal{B}\rangle_t$ along the vector space $\langle F \,|\, t\rangle$. It leads to the notion of border bases.

Definition 5.1.8 *Let* $\mathcal{B} \subset \mathcal{M}$ *be connected to* 1. *A family* $F \subset \mathbb{K}[\mathbf{x}]$ *is a border basis for* $\mathcal{B}$ *if it is a rewriting family for* $\mathcal{B}$, *complete in all degrees, and such that* $\mathbb{K}[\mathbf{x}] = \langle\mathcal{B}\rangle \oplus (F)$.

An algorithmic way to check that we have a border basis is based on the following result, that we recall from [Mourrain 2005]:

Theorem 5.1.9 *Assume that* $\mathcal{B}$ *is connected to* 1 *and let* F *be a rewriting family for* $\mathcal{B}$, *complete in degree* $t \in \mathbb{N}$. *Suppose that, for all* $m \in \mathcal{M}_{t-2}$,

$$\pi_{F,\mathcal{B}}(x_i\, \pi_{F,\mathcal{B}}(x_j\, m)) = \pi_{F,\mathcal{B}}(x_j\, \pi_{F,\mathcal{B}}(x_i\, m)) \quad \text{for all } i, j \in [1, n]. \tag{5.1}$$

Then $\pi_{F,\mathcal{B}}$ *coincides with the linear projection of* $\mathbb{K}[\mathbf{x}]_t$ *on* $\langle\mathcal{B}\rangle_t$ *along the vector space* $\langle F \,|\, t\rangle$ *that is,* $\mathbb{K}[\mathbf{x}]_t = \langle\mathcal{B}\rangle_t \oplus \langle F \,|\, t\rangle$.

In order to have a simple test and effective way to test the commutation relations (5.1), we introduce now the commutation polynomials.

Definition 5.1.10 *Let* F *be a rewriting family and* $f, f' \in F$. *Let* m, m' *be the smallest degree monomials for which* $m\,\gamma(f) = m'\,\gamma(f')$. *Then the polynomial* $CP(f, f') := mf - m'f' = m'\pi_{F,\mathcal{B}}(f') - m\pi_{F,\mathcal{B}}(f)$ *is called the* commutation polynomial *of* f, f'.

Definition 5.1.11 *For a rewriting family* F *with respecet to* $\mathcal{B}$, *we denote by* $CP^+(F)$ *the set of polynomials of the form* $m\,f - m'\,f'$, *where* $f, f' \in F$ *and*

- *either* $m\,\gamma(f) = m'\,\gamma(f')$,

- *or* $m\,\gamma(f) \in \mathcal{B}$ *and* $m' = 0$.

Therefore, $CP^+(F) \subset \langle \mathcal{B}^+ \rangle$ and $CP^+(F)$ contains all commutation polynomials $CP(f, f')$ for $f, f' \in F$ whose monomial multipliers m, m' are of degree ≤ 1. The next result can be deduced using Theorem 5.1.9.

Theorem 5.1.12 *Let* $\mathcal{B} \subset \mathcal{M}$ *be connected to* 1 *and let* F *be a rewriting family for* $\mathcal{B}$*, complete in degree* t*. If for all* $c \in CP^+(F)$ *of degree* $\leq t$*,* $\pi_{F,\mathcal{B}}(c) = 0$*, then* $\pi_{F,\mathcal{B}}$ *is the projection of* $\mathbb{K}_t$ *on* $\langle \mathcal{B} \rangle_t$ *along* $\langle F \,|\, t \rangle$*, ie.* $\mathbb{K}_t = \langle \mathcal{B} \rangle_t \oplus \langle F \,|\, t \rangle$.

If such a property is satisfied we say that F is a border basis for B in degree $\leq t$.

Theorem 5.1.13 *[Mourrain 2005] Let* $\mathcal{B} \subset \mathcal{M}$ *be connected to 1 and let* F *be a rewriting family for* $\mathcal{B}$*, complete in any degree. Assume that* $\pi_{F,\mathcal{B}}(c) = 0$ *for all* $c \in CP^+(F)$*.Then* $\mathcal{B}$ *is a basis of* $\mathbb{K}[\mathbf{x}]/(F)$*,* $\mathbb{K} = \langle \mathcal{B} \rangle \oplus (F)$*, and* $(F)_t = \langle F \,|\, t \rangle$ *for all* $t \in \mathbb{N}$ *the set* F *is a* border basis *of the ideal* $I = (F)$ *with respect to* $\mathcal{B}$.

This implies the following characterization of border bases using the commutation property.

Corollary 5.1.14 *[Mourrain 1999] Let* $\mathcal{B} \subset \mathcal{M}$ *be connected to 1 and let* F *be a rewriting family for* $\mathcal{B}$*, complete in any degree. If for all* $m \in \mathcal{B}$ *and all indices* $i, j \in [1, n]$*, we have:*

$$\pi_{F,\mathcal{B}}(x_i\,\pi_{F,\mathcal{B}}(x_j\,m)) = \pi_{F,\mathcal{B}}(x_j\,\pi_{F,\mathcal{B}}(x_i\,m)),$$

then $\mathcal{B}$ *is a basis of* $\mathbb{K}/(F)$*,* $\mathbb{K} = \langle \mathcal{B} \rangle \oplus (F)$*, and* $(F)_t = \langle F \,|\, t \rangle$ *for all* $t \in \mathbb{N}$.

5.2 Border basis hierarchy

The sequence of relaxation problems that we will use hereafter is defined as follows. For each $t \in \mathbb{N}$, we construct the graded border basis F_{2t} of $\mathbf{g}^0$ in degree $2\,t$. Let B be the set of monomials (connected to 1) for which F is a border basis in degree $2t$. We define

- $E_t := \langle B_t \rangle$,
- C_t is the set of constraints such that
  - $C_t^0 = \{m - \pi_{B_t,F_{2t}}(m),\, m \in B_t \cdot B_t\}$

– $C_t^+ = \pi_{B_t,F_{2t}}(\mathbf{g}^+)$

and consider the relaxation sequence

$$\mathcal{Q}_{E_t}(C_t) \subset \langle B_t \cdot B_t \rangle \text{ and } \mathcal{L}_{E_t}(C_t) \subset \langle B_t \cdot B_t \rangle^* \tag{5.2}$$

for $t \in \mathbb{N}$. Since the subsets B_t are not necessarily nested, these convex sets are not necessarily included one in the other one. However, by construction of the graded border basis of $\mathbf{g}$, we have the following inclusions

$$\cdots \subset \langle F_{2t} | 2t \rangle \subset \langle F_{2t+2} | 2t+2 \rangle \subset \cdots (\mathbf{g}^0),$$

and we can relate the border basis relaxation sequences with the corresponding full moment matrix relaxation hierarchy, using the following proposition:

Proposition 5.2.1 *Let $t \in \mathbb{N}$, $B \subset \mathbb{R}[\mathbf{x}]_{2t}$ be a monomial set connected to 1, $F \subset \mathbb{R}[\mathbf{x}]$ be a border basis for B in degree $2t$, $E := \langle B_t \rangle$, $E' := \mathbb{R}[\mathbf{x}]_t$, C, C' be sets of constraints such that $C^0 = \{m - \pi_{B,F}(m),\ m \in B_t \cdot B_t\}$, $C'^0 = \langle F | 2\,t \rangle$, $C^+ = C'^+$. Then for all $\Lambda \in \mathcal{L}_E(C)$, there exists a unique $\Lambda' \in \mathcal{L}_{E'}(C')$ which extends Λ. Moreover, Λ' satisfies $\operatorname{rank} H_{\Lambda'}^{E'} = \operatorname{rank} H_\Lambda^E$ and $\ker H_{\Lambda'}^{E'} = \ker H_\Lambda^E + \langle F \,|\, t \rangle$.*

Proof. As $F \subset \mathbb{R}[\mathbf{x}]$ is a border basis for B in degree $2t$, we have $\mathbb{R}[\mathbf{x}]_{2\,t} = \langle B \rangle_{2\,t} \oplus \langle F \,|\, 2\,t \rangle$. As $\langle B_t \cdot B_t \rangle \subset \langle B \rangle_{2\,t} \oplus \langle C^0 \rangle$, $\langle C^0 \rangle \subset \langle C'^0 \rangle = \langle F \,|\, 2t \rangle$ and $\mathbb{R}[\mathbf{x}]_{2\,t} = \langle B \rangle_{2\,t} \oplus \langle F \,|\, 2t \rangle$, we deduce that for all $\Lambda \in \mathcal{L}_E(C)$, there exists a unique $\Lambda' \in \mathbb{R}[\mathbf{x}]_{2\,t}^*$ s.t. $\Lambda'_{|\langle B \rangle_{2t}} = \Lambda$ and $\Lambda'(\langle F \,|\, 2t \rangle) = 0$.

Let us first prove that $\Lambda' \in \mathcal{L}_{E'}(C') = \mathcal{L}_t(C')$. As any element q' of $\mathcal{Q}_{E'}(C')$ can be decomposed as a sum of an element q of $\mathcal{Q}_E(C)$ and an element $p \in \langle F | 2t \rangle$, we have $\Lambda'(q') = \Lambda'(q) + \Lambda'(p) = \Lambda(q) \geq 0$. This shows that $\Lambda' \in \mathcal{L}_{E'}(C')$.

Let us prove now that $\ker H_{\Lambda'}^{E'} = \ker H_\Lambda^E + \langle F \,|\, t \rangle$ where $E := \langle B_t \rangle$, $E' := \mathbb{R}[\mathbf{x}]_t$. As $E \cdot \langle F \,|\, t \rangle \subset \langle F \,|\, 2t \rangle = C'^0$, we have $\Lambda'(E \cdot \langle F \,|\, t \rangle) = 0$ so that

$$\langle F \,|\, t \rangle \subset \ker H_{\Lambda'}^{E'}. \tag{5.3}$$

For any element $b \in \ker H_\Lambda^E$ we have $\forall b' \in E$, $\Lambda(b\,b') = \Lambda'(b\,b') = 0$. As $\Lambda'(E \cdot \langle F \,|\, t \rangle) = 0$ and $E' = E \oplus \langle F \,|\, t \rangle$, for any element $e \in E$, $\Lambda'(b\,e) = 0$. This proves that

$$\ker H_\Lambda^E \subset \ker H_{\Lambda'}^{E'}. \tag{5.4}$$

Conversely as $E' = E \oplus \langle F \,|\, t \rangle$, any element of E' can be reduced modulo $\langle F \,|\, t \rangle$ to an element of E, which shows that

$$\ker H_{\Lambda'}^{E'} \subset \ker H_\Lambda^E + \langle F \,|\, t \rangle. \tag{5.5}$$

From the inclusions (5.3), (5.4) and (5.5), we deduce that $\ker H^{E'}_{\Lambda'} = \ker H^E_\Lambda + \langle F \,|\, t\rangle$ and that $\operatorname{rank} H^{E'}_{\Lambda'} = \operatorname{rank} H^E_\Lambda$. ∎

We deduce from this proposition that $f^\mu_{E_t,C_t} = f^\mu_{t,\langle F_{2t}|2t\rangle}$. The sequence of convex sets $\mathcal{L}_{E_t}(C_t)$ can be seen as the projections of nested convex sets

$$\cdots \supset \mathcal{L}_t(\mathbf{g}) \supset \mathcal{L}_{t+1}(\mathbf{g}) \supset \cdots$$

so that we have $\cdots \leq f^\mu_{E_t,C_t} \leq f^\mu_{E_{t+1},C_{t+1}} \leq \cdots \leq f^*$. We check that similar properties hold for $\mathcal{Q}_{E_t}(C_t)$, $\mathcal{Q}_t(\mathbf{g})$ and $f^{sos}_{E_t,C_t} = f^{sos}_{t,\mathbf{g}}$, taking the quotient modulo $\langle F_{2t}|2t\rangle$.

5.2.1 Optimal linear form

We introduce now the notion of *optimal linear form for* f, involved in the computation of I_{min} (also called generic linear form when $f = 0$ in [Lasserre 2009b, Lasserre 2012]):

Definition 5.2.2 *$\Lambda^* \in \mathcal{L}_E(C)$ is optimal for f if*

$$\text{rank } H^E_{\Lambda^*} = max_{\Lambda\in\mathcal{L}_{E,C},\Lambda(f)=f^\mu_{E,C}} \text{rank } H^E_\Lambda.$$

The next result shows that only elements in I_{min} are involved in the kernel of a truncated Hankel operator associated to an optimal linear form for f.

Theorem 5.2.3 *Let $E \subset \mathbb{R}[\mathbf{x}]$ such that $1 \in E$ and $f \in \langle E \cdot E\rangle$ and let $C \subset \mathbb{R}[\mathbf{x}]$ be a set of constraints with $S_{min} \subset \mathcal{S}(C)$. If $\Lambda^* \in \mathcal{L}_E(C)$ is optimal for f and such that $\Lambda^*(f) = f^*$, then $\ker H^E_{\Lambda^*} \subset I_{min}$.*

Proof. Let $p \in \ker H^E_{\Lambda^*}$ and $\mathbf{x}^* \in S_{min}$, which means $f(\mathbf{x}^*) = f^*$. Let $\underline{\mathbf{1}}_{\mathbf{x}^*}$ denotes the evaluation at $\mathbf{x}^*$ restricted to $\langle E \cdot E\rangle$. Our objective is to show that $p(\mathbf{x}^*) = 0$. Suppose for contradiction that $p(\mathbf{x}^*) \neq 0$. We know that $\underline{\mathbf{1}}_{\mathbf{x}^*} \in \mathcal{L}_E(C)$ since $S_{min} \subset \mathcal{S}(C)$ and $\underline{\mathbf{1}}_{\mathbf{x}^*}(f) = f(\mathbf{x}^*) = f^*$. We define $\tilde{\Lambda} = \frac{1}{2}(\Lambda^* + \underline{\mathbf{1}}_{\mathbf{x}^*})$. By construction, $\tilde{\Lambda} \in \mathcal{L}_E(C)$ and $\tilde{\Lambda}(f) = \frac{1}{2}(\Lambda^*(f) + \underline{\mathbf{1}}_{\mathbf{x}^*}(f)) = \frac{1}{2}(\Lambda^*(f) + f(\mathbf{x}^*)) = f^*$. As $p \in \ker H^E_{\Lambda^*}$,

$$\tilde{\Lambda}(p^2) = \frac{1}{2}(\Lambda^*(p^2) + \underline{\mathbf{1}}_{\mathbf{x}^*}(p^2)) = \frac{1}{2}p^2(\mathbf{x}^*) \neq 0$$

thus $p \in \ker H^E_{\Lambda^*} \setminus \ker H^E_{\tilde{\Lambda}}$ and by the maximality of the rank of $H^E_{\Lambda^*}$, $\ker H^E_{\tilde{\Lambda}} \not\subset \ker H^E_{\Lambda^*}$. Hence there exits $\tilde{p} \in \ker H^E_{\tilde{\Lambda}} \setminus \ker H^E_{\Lambda^*}$. Then $0 =$

$H^E_{\tilde{\Lambda}}(\tilde{p}) = \frac{1}{2}(H^E_{\Lambda^*}(\tilde{p}) + H^E_{\underline{1}_{\mathbf{x}^*}}(\tilde{p})) = \frac{1}{2}(H^E_{\Lambda^*}(\tilde{p}) + \tilde{p}(\mathbf{x}^*) \cdot \underline{1}_{\mathbf{x}^*})$. As $H^E_{\Lambda^*}(\tilde{p}) \neq 0$ implies $\tilde{p}(\mathbf{x}^*) \neq 0$. On the other hand,

$$0 = H^E_{\tilde{\Lambda}}(\tilde{h})(p) = \tilde{\Lambda}(p\tilde{p}) = \frac{1}{2}(\Lambda^*(p\tilde{p}) + p(\mathbf{x}^*)\tilde{p}(\mathbf{x}^*)) =$$

$$= \frac{1}{2}(H^E_{\Lambda^*}(p)(\tilde{p}) + p(\mathbf{x}^*)\tilde{p}(\mathbf{x}^*)).$$

As $p \in \ker H^E_{\Lambda^*}$, we have

$$0 = H^E_{\tilde{\Lambda}}(\tilde{p})(p) = \frac{1}{2}(p(\mathbf{x}^*)\tilde{p}(\mathbf{x}^*)).$$

As $\tilde{p}(\mathbf{x}^*) \neq 0$ and we have supposed that $p(\mathbf{x}^*) \neq 0$, it yields a contradiction. ■

The proof is similar e.g. to [Lasserre 2012][Theorem 4.9].

Let us describe how optimal linear forms are computed by solving convex optimization problems:

Algorithm 5.2.1: Optimal Linear Form

Input: $f \in \mathbb{R}[\mathbf{x}]$, $B_t = (\mathbf{x}^\alpha)_{\alpha \in A}$ a monomial set containing 1 of degree $\leq t$ with $f = \sum_{\alpha \in A+A} f_\alpha \mathbf{x}^\alpha \in \langle B_t \cdot B_t \rangle$, $C \subset \mathbb{R}[\mathbf{x}]$.

Output: the minimum $f^\mu_{t,C}$ of $\sum_{\alpha \in A+A} \lambda_\alpha f_\alpha$ subject to:

- $H^{B_t}_{\Lambda^*} = (h_{\alpha,\beta})_{\alpha,\beta \in A} \succcurlyeq 0$,
- $H^{B_t}_{\Lambda^*}$ satisfies the Hankel constraints $h_{0,0} = 1$, and $h_{\alpha,\beta} = h_{\alpha',\beta'}$ if $\alpha + \beta = \alpha' + \beta'$,
- $\Lambda^*(g^0) = \sum_{\alpha \in A+A} g^0_\alpha \lambda_\alpha = 0$ $\forall\ g^0 = \sum_{\alpha \in A+A} g^0_\alpha \mathbf{x}^\alpha \in C^0 \cap \langle B_t \cdot B_t \rangle$.
- $H^{B_{t-w}}_{g^+ \cdot \Lambda^*} \succcurlyeq 0$ for all $g^+ \in C^+$ where $w = \lceil \frac{deg(g^+)}{2} \rceil$.

and $\Lambda^* \in \langle B_t \cdot B_t \rangle^*$ represented by the vector $[\lambda_\alpha]_{\alpha \in A+A}$.

This optimization problem is a Semi-Definite Programming problem, corresponding to the optimization of a linear functional on the intersection of a

linear subspace with the convex set of Positive Semi-Definite matrices. It is a convex optimization problem, which can be solved efficiently by SDP solvers. If an Interior Point Method is used, the solution Λ^* is in the interior of a face on which the minimum $\Lambda^*(f)$ is reached so that Λ^* is optimal for f. This is the case for tools such as `csdp`, `sdpa`, `sdpa-gmp` or `mosek`, that we will use in the experimentations.

Example 5.2.4 *We consider the following problem*

$$\begin{aligned} &\min\ x^2+3\\ &s.t.\ f = x^4-x^3-x+1 = (x-1)^2(x^2+x+1) = 0 \end{aligned}$$

In order to solve this problem we solve the equivalent SDP problem for $t=3$

$$\begin{aligned} &\min\ \Lambda(x^2+3) = \Lambda(x^2)+3 \text{ with } \Lambda \text{ s.t.}\\ &\Lambda(1) = 1,\\ &\Lambda(x^4) = \Lambda(x^3)+\Lambda(x)-\Lambda(1),\\ &\Lambda(x^5) = \Lambda(x^3)+\Lambda(x^2)-\Lambda(1),\\ &\Lambda(x^6) = 2\,\Lambda(x^3)-\Lambda(1) \end{aligned}$$

and

$$H_\Lambda := \begin{pmatrix} 1 & a & b & c\\ a & b & c & c+a-1\\ b & c & c+a-1 & c+b-1\\ c & c+a-1 & c+b-1 & 2\,c-1 \end{pmatrix} \succcurlyeq 0$$

where $a=\Lambda(x)$, $b=\Lambda(x^2)$, $c=\Lambda(x^3)$.

Solution: $\Lambda^*(1)=1$, $\Lambda^*(x)=1$, $\Lambda^*(x^2)=1$, $\Lambda^*(x^3)=1$, ...
The minimum is $\Lambda(x^2+3)=\Lambda(x^2)+3=4$.

5.3 Convergence certification

To be able to compute the minimizer points from an optimal linear form, we need to detect when the minimum is reached. In this Section, we describe a new criterion to check when the kernel of a truncated Hankel operator associated to an optimal linear form for f yields the generators of the minimizer

ideal. It involves the flat extension theorem of [Laurent 2009b] and applies to polynomial optimization problems where the minimizer ideal I_{min} is zero-dimensional.

5.3.1 Flat extension criterion

Definition 5.3.1

- *A vector space $E \subseteq \mathbb{K}[\mathbf{x}]$ is said to be connected to 1 if $1 \in E$ and any non-constant polynomial $p \in E$ can be written as $p = p_0 + \sum_{i=1}^{n} x_i p_i$ for some polynomials $p_0, p_i \in E$ with $\deg(p_i) \leq \deg(p) - 1$ for $i \in [0, n]$.*
- *Its prolongation $E^+ := E + x_1 E + \ldots + x_n E$ is again a vector space.*

Remark 5.3.2 *Obviously, E is connected to 1 when $E = \langle \mathcal{C} \rangle$ for some monomial set $\mathcal{C} \subseteq \mathcal{M}$ which is connected to 1. Moreover, $E^+ = \langle \mathcal{C}^+ \rangle$ if $E = \langle \mathcal{C} \rangle$.*

Definition 5.3.3 *Given vector subspaces $E_0 \subset E \subset \mathbb{K}[\mathbf{x}]$ and $\Lambda \in \langle E \cdot E \rangle^*$, H_Λ^E is said to be a flat extension of its restriction $H_\Lambda^{E_0}$ if $\operatorname{rank} H_\Lambda^E = \operatorname{rank} H_\Lambda^{E_0}$.*

We recall here a result from [Laurent 2009b], which extends the result given by Curto and Fialkow in [Curto 1996]. It gives a rank condition for the existence of a flat extension of a truncated Hankel operator [1].

Theorem 5.3.4 *Let $V \subset E \subset \mathbb{R}[\mathbf{x}]$ be vector spaces connected to 1 with $V^+ \subset E$ and let $\Lambda \in \langle E \cdot E \rangle^*$. Assume that $\operatorname{rank} H_\Lambda^E = \operatorname{rank} H_\Lambda^V = \dim V$. Then there exists a (unique) linear form $\tilde{\Lambda} \in \mathbb{R}[\mathbf{x}]^*$ which extends Λ, i.e., $\tilde{\Lambda}(p) = \Lambda(p)$ for all $p \in \langle E \cdot E \rangle$, satisfying $\operatorname{rank} H_{\tilde{\Lambda}} = \operatorname{rank} H_\Lambda^E$. Moreover, we have $\ker H_{\tilde{\Lambda}} = (\ker H_\Lambda^E)$.*

In other words, the condition $\operatorname{rank} H_\Lambda^E = \operatorname{rank} H_\Lambda^V = \dim V$ implies that the truncated Hankel operator H_Λ^E has a (unique) flat extension to a (full) Hankel operator $H_{\tilde{\Lambda}}$ defined on $\mathbb{R}[\mathbf{x}]$.

Theorem 5.3.5 *Let $V \subset E \subset \mathbb{R}[\mathbf{x}]$ be finite dimensional vector spaces connected to 1 with $V^+ \subset E$, $C^0 \cdot V \subset \langle E \cdot E \rangle$, $C^+ \cdot V \cdot V \subset \langle E \cdot E \rangle$.*

Let $\Lambda \in \mathcal{L}_E(C)$ such that $\operatorname{rank} H_\Lambda^E = \operatorname{rank} H_\Lambda^V = \dim V$. Then there exists a linear form $\tilde{\Lambda} \in \mathbb{R}[\mathbf{x}]^*$ which is extending Λ and supported on points of $\mathcal{S}(C)$ with positive weights:

$$\tilde{\Lambda} = \sum_{i=1}^{r} \omega_i \mathbf{1}_{\xi_i} \text{ with } \omega_i > 0, \xi_i \in \mathcal{S}(C).$$

Moreover, $(\ker H_\Lambda^E) = \mathcal{I}(\xi_1, \ldots, \xi_r)$.

[1] In [Laurent 2009b], it is stated with a vector space spanned by a monomial set connected to 1, but its extension to vector spaces connected to 1 is straightforward.

Proof. As $\operatorname{rank} H_\Lambda^E = \operatorname{rank} H_\Lambda^V = \dim V$, Theorem 5.3.4 implies that there exists a (unique) linear function $\tilde{\Lambda} \in \mathbb{R}[\mathbf{x}]^*$ which extends Λ. As $\operatorname{rank} H_{\tilde{\Lambda}} = \operatorname{rank} H_\Lambda^V = |V|$ and $\ker H_{\tilde{\Lambda}} = (\ker H_\Lambda^E)$, any polynomial $p \in \mathbb{R}[\mathbf{x}]$ can be reduced modulo $\ker H_{\tilde{\Lambda}}$ to a polynomial $b \in V$ so that $p - b \in \ker H_{\tilde{\Lambda}}$. Then $\tilde{\Lambda}(p^2) = \tilde{\Lambda}(b^2) = \Lambda(b^2) \geq 0$ since $\Lambda \in \mathcal{L}_E(C)$. By Proposition 1.4.5 $\tilde{\Lambda}$ has a decomposition of the form $\tilde{\Lambda} = \sum_{i=1}^r \omega_i \mathbf{1}_{\xi_i}$ with $\omega_i > 0$ and $\xi_i \in \mathbb{R}^n$.

V is isomorphic to $\mathbb{R}[\mathbf{x}]/\mathcal{I}(\xi_1, \ldots, \xi_r)$ and there exist (interpolation) polynomials $b_1, \ldots, b_r \in V$ satisfying $b_i(\xi_j) = 1$ if $i = j$ and $b_i(\xi_j) = 0$ otherwise. We deduce that for $i = 1, \ldots, r$ and for all elements $g \in C^0$,

$$\Lambda(b_i g) = 0 = \tilde{\Lambda}(b_i g) = \omega_i g(\xi_i).$$

As $\omega_i > 0$ then $g(\xi_i) = 0$. Similarly, for all $h \in C^+$,

$$\Lambda(b_i^2 h) = \tilde{\Lambda}(b_i^2 h) = \omega_i h(\xi_i) \geq 0$$

and $h(\xi_i) \geq 0$, hence $\xi_i \in \mathcal{S}(C)$.

By 1.4.3 and Theorem 5.3.4, we have moreover $\ker H_{\tilde{\Lambda}} = \mathcal{I}(\xi_1, \ldots, \xi_r) = (\ker H_\Lambda^E)$. ∎

This theorem applied to an optimal linear form Λ^* for f gives a convergence certificate to check when the minimum f^* is reached and when a generating family of the minimizer ideal is obtained. It generalizes the flat truncation certificate given in [Nie 2012]. As we will see in the experimentation part, it allows to detect more efficiently when the minimum is reached. Notice that if the test is satisfied, necessarily I_{min} is zero-dimensional.

5.3.2 Flat extension algorithm

In this Section, we describe a new algorithm to check the flat extension property for a linear form for which some moments are known.

Let E be a finite dimensional subspace of $\mathbb{R}[\mathbf{x}]$ connected to 1 and let Λ^* be a linear form defined on $\langle E \cdot E \rangle$ given by its "moments" $\Lambda^*(e_i) := \Lambda_i^*$, where $e_1, \ldots, e_s$ is a basis of $\langle E \cdot E \rangle$ (for instance a monomial basis). In the context of global polynomial optimization over an semialgebraic set that we consider here, this linear form is an optimal linear form for f (see Section 5.2.1) computed by SDP.

We define the linear form Λ^* from its moments as $\Lambda^* : p = \sum_{i=1}^s p_i e_i \in \langle E \cdot E \rangle \mapsto \sum_{i=1}^s p_i \Lambda_i$ and the corresponding inner product defined as in Definition 1.2.3:

$$\begin{array}{rcl} E \times E & \to & \mathbb{R} \\ (p, q) & \mapsto & \langle p, q \rangle_{\Lambda *} := \Lambda^*(p\, q) \end{array} \tag{5.6}$$

To check the flat extension property, we are going to define inductively vector spaces V_i as follows. Start with $V_0 = \langle 1 \rangle$. Suppose V_i is known and compute a vector space L_i of maximal dimension in V_i^+ such that L_i is orthogonal to V_i: $\langle L_i, V_i \rangle_{\Lambda *} = 0$ and $L_i \cap \ker H_{\Lambda *}^{V_i^+} = \{0\}$. Then we define $V_{i+1} = V_i + L_i$.

Suppose that $b_1, \ldots, b_{r_i}$ is an orthogonal basis of V_i: $\langle b_i, b_j \rangle_{\Lambda *} = 0$ if $i \neq j$ and $\langle b_i, b_i \rangle_{\Lambda *} \neq 0$. Then L_i can be constructed as follows: Compute the vectors

$$b_{i,j} = x_j b_i - \sum_{k=1}^{r_i} \frac{\langle x_j b_i, b_k \rangle_{\Lambda *}}{\langle b_k, b_k \rangle_{\Lambda *}} b_k,$$

generating $V_i^\perp$ in V_i^+ and extract a maximal orthogonal family $b_{r_i+1}, \ldots, b_{r_{i+1}}$ for the inner product $\langle ., . \rangle_{\Lambda *}$, that form a basis of L_i. This can be done for instance by computing a QR decomposition of the matrix $[\langle b_{i,j}, b_{i',j'} \rangle_{\Lambda *}]_{1 \leqslant i,i' \leqslant r_i, 1 \leqslant j,j' \leqslant n}$. The process can be repeated until either

- $V_i^+ \not\subset E$ and the algorithm will stop and return `failed`,
- or $L_i = \{0\}$ and $V_i^+ = V_i \oplus \ker H_{\Lambda *}^{V_i^+}$. In this case, the algorithm stops with `success`.

The complete description of the algorithm is as follow:

Algorithm 5.3.1: DECOMPOSITION

Input: a vector space E connected to 1 and a linear form $\Lambda^* \in \langle E \cdot E \rangle^*$.

- Take $B := \{1\}$ $s := 1$; $r := 1$
- While $s > 0$ and $B^+ \subset E$ do
  - compute $b_{j,k} := x_k b_j - \sum_{i=1}^{r} \frac{\langle x_k b_j, b_i \rangle_{\Lambda^*}}{\langle b_i, b_i \rangle_{\Lambda^*}} b_i$ for $j = 1, \ldots, r$, $k = 1, \ldots, n$;
  - compute a maximal subset $B' = \{b'_1, \ldots, b'_s\}$ of $\langle b_{j,k} \rangle$ of orthogonal vectors for the inner product $\langle ., . \rangle_{\Lambda^*}$ and let $B := B \cup B'$, $s =\mid B' \mid$ and $r \mathrel{+}= s$;
- If $B^+ \not\subset E$ then return `failed` else return `success`.

Output: `failed` or `success` with

- a basis $B = \{b_1, \ldots, b_r\} \subset \mathbb{R}[\mathbf{x}]$;
- the relations $x_k b_j - \sum_{i=1}^{r} \frac{\langle x_k b_j, b_i \rangle_{\Lambda^*}}{\langle b_i, b_i \rangle_{\Lambda^*}} b_i$, $j = 1 \ldots r$ $k = 1 \ldots n$.

Let us describe the computation performed on the moment matrix, during the main loop of the algorithm. At each step, the moment matrix of Λ^* on V_i^+ is of the form

$$H_{\Lambda^*}^{V_i^+} = \left[\begin{array}{c|c} H_{\Lambda^*}^{B_i,B_i} & H_{\Lambda^*}^{B_i,\partial B_i} \\ \hline H_{\Lambda^*}^{\partial B_i,B_i} & H_{\Lambda^*}^{\partial B_i,\partial B_i} \end{array}\right]$$

where ∂B_i is a subset of $\{b_{i,j}\}$ such that $B_i \cup \partial B_i$ is a basis of $\langle B_i^+ \rangle$. By construction, the matrix $H_{\Lambda^*}^{B_i,B_i}$ is diagonal since B_i is orthogonal for $\langle \cdot, \cdot \rangle_*$. As the polynomials $b_{i,j}$ are orthogonal to B_i, we have $H_{\Lambda^*}^{B_i,\partial B_i} = H_{\Lambda^*}^{\partial B_i,B_i} = 0$. If $H_{\Lambda^*}^{\partial B_i,\partial B_i} = 0$ then the algorithm stops with `success` and all the elements $b_{i,j}$ are in the kernel of $H_{\Lambda^*}^{B_i,B_i}$. Otherwise an orthogonal basis $b'_1, \ldots, b'_s$ is extracted. It can then be completed in a basis of $\langle b_{i,j} \rangle$ so that the matrix $H_{\Lambda^*}^{\partial B_i,\partial B_i}$ in this basis is diagonal with zero entries after the $(s+1)^{th}$ index. In the next loop of the algorithm, the basis B_{i+1} contains the maximal orthogonal family $b'_1, \ldots, b'_s$ so that the matrix $H_{\Lambda^*}^{B_{i+1},B_{i+1}}$ remains diagonal and invertible.

Proposition 5.3.6 *Let $\Lambda^* \in \mathcal{L}_E(C)$ be optimal for f. If Algorithm 5.3.1 applied to Λ^* and E stops with `success`, then*

1. *there exists a linear form* $\tilde{\Lambda} \in \mathbb{R}[\mathbf{x}]^*$ *which is extending* Λ^* *and supported on points in* $\mathcal{S}(C)$ *with positive weights:*

$$\tilde{\Lambda} = \sum_{i=1}^{r} \omega_i \mathbf{1}_{\xi_i} \text{ with } \omega_i > 0, \xi_i \in \mathbb{R}^n.$$

2. $B = \{b_1, \ldots, b_r\}$ *is a basis of* $\mathcal{A}_{\tilde{\Lambda}} = \mathbb{R}[\mathbf{x}]/I_{\tilde{\Lambda}}$ *where* $I_{\tilde{\Lambda}} = \ker H_{\tilde{\Lambda}}$,

3. $x_k b_j - \sum_{i=1}^{r} \frac{\langle x_k b_j, b_i \rangle_{\Lambda^*}}{\langle b_i, b_i \rangle_{\Lambda^*}} b_i$, $j = 1, \ldots, r$, $k = 1, \ldots, n$ *are generating* $I_{\tilde{\Lambda}} = \mathcal{I}(\xi_1, \ldots, \xi_r)$,

4. $f^{\mu}_{E,C} = f^*$,

5. $V_{min} = \{\xi_1, \ldots, \xi_r\}$.

Proof. When the algorithm terminates with `success`, the set B is such that $\operatorname{rank} H^{B^+}_{\Lambda^*} = \operatorname{rank} H^{B}_{\Lambda^*} = |B|$. By Theorem 5.3.5, there exists a linear form $\tilde{\Lambda} \in \mathbb{R}[\mathbf{x}]^*$ extending Λ^* and supported on points in $\mathcal{S}(C)$ with positive weights:

$$\tilde{\Lambda} = \sum_{i=1}^{r} \omega_i \mathbf{1}_{\xi_i} \text{ with } \omega_i > 0, \xi_i \in \mathcal{S}(C).$$

This implies that $\mathcal{A}_{\tilde{\Lambda}}$ is of dimension r and that $I_{\tilde{\Lambda}} = \mathcal{I}(\xi_1, \ldots, \xi_r)$. As $H^{B}_{\Lambda^*}$ is invertible, B is a basis of $\mathcal{A}_{\tilde{\Lambda}}$ which proves the second point.

Let K be the set of polynomials $x_j b_i - \sum_{k=1}^{r} \frac{\langle x_j b_i, b_k \rangle_{\Lambda^*}}{\langle b_k, b_k \rangle_{\Lambda^*}} b_k$. If the algorithm terminates with `success`, we have $\ker H^{B^+}_{\Lambda^*} = \langle K \rangle$ and by Theorem 5.3.5, we deduce that $(K) = (\ker H^{B^+}_{\Lambda^*}) = I_{\tilde{\Lambda}}$, which proves the third point.

As $\tilde{\Lambda}(1) = 1$, we have $\sum_{i=1}^{r} w_i = 1$ and

$$\tilde{\Lambda}(f) = \sum_{i=1}^{r} \omega_i f(\xi_i) \geq f^*$$

since $\xi_i \in \mathcal{S}(C)$ and $f(\xi_i) \geq f^*$. The relation $f^{\mu}_{E,C} \leq f^*$ implies that $f(\xi_i) = f^*$ for $i = 1, \ldots, r$ and the fourth point is true: $f^{\mu}_{E,C} = f^*$.

As $f(\xi_i) = f^*$ for $i = 1, \ldots, r$, we have $\{\xi_1, \ldots, \xi_r\} \subset V_{min}$. By Theorem 5.2.3, the polynomials of K are in I_{min} so that $V_{min} \subset \mathcal{V}(K) = \{\xi_1, \ldots, \xi_r\}$. This shows that $V_{min} = \{\xi_1, \ldots, \xi_r\}$ and concludes the proof of this proposition ■

Example 5.3.7 *We consider the following global optimization problem*

$$\min\ f(x) = (x-1)^2 \cdot (x-2)^2 \cdot (x^2+1) + (y-1)^2 \cdot (y^2+1)$$

We consider the same problem over his gradient ideal and we compute the optimal linear form solving the SDP problem with t=3.

We obtain the following solution:
$\Lambda^*(1) = 1, \Lambda^*(x) = 1.5, \Lambda^*(y) = 1, \Lambda^*(x^2) = 2.5, \Lambda^*(xy) = 1.5, \Lambda^*(y^2) = 1, \Lambda^*(y^3) = 1, \Lambda^*(xy^2) = 1.5, \Lambda^*(x^2y) = 2.5, \Lambda^*(x^3) = 4.5, \Lambda^*(y^4) = 1, \Lambda^*(xy^3) = 1.5, \Lambda^*(x^2y^2) = 2.5, \Lambda^*(x^3y) = 4.5, \Lambda^*(x^4) = 8.5, ...$

We apply Algorithm 5.3.1 for $E = \mathbb{R}[\mathbf{x}]_3$

- $B_0 = \{1\},\ \partial B_0 = \{x, y\},\ B_0^+ = \{1, x, y\}$

$$H_\Lambda^{B_0^+} = \begin{pmatrix} 1 & 1.5 & 1 \\ 1.5 & 2.5 & 1.5 \\ 1 & 1.5 & 1 \end{pmatrix} \longrightarrow H_\Lambda^{\{1, x-1.5, y-1\}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.25 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$rank\ H_\Lambda^{B_0^+} = 2,$$

$$\{y-1\} \in \ker H_\Lambda^{B_0^+},\ \{x-1.5\} \perp B_0 \text{ and } \{x-1.5\} \notin \ker H_\Lambda^{B_0^+}$$

$$L_0 = \{x-1.5\}$$

- $B_1 = B_0 + L_0 = \{1, x-1.5\},\ \partial B_1 = \{y, x^2-1.5x, xy-1.5y\},$
  $B_1^+ = \{1, x-1.5, y, x^2-1.5x, xy-1.5y\}$

$$H_\Lambda^{B_1^+} = \begin{pmatrix} 1 & 0 & 1 & 0.25 & 0 \\ 0 & 0.25 & 0 & 0.375 & 0.25 \\ 1 & 0 & 1 & 0.25 & 0 \\ 0.25 & 0.375 & 0.25 & 0.625 & 0.375 \\ 0 & 0.25 & 0 & 0.375 & 0.25 \end{pmatrix} \longrightarrow H_\Lambda^{\tilde{B}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0.25 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

where $\tilde{B} = \{1, x - 1.5, y - 1, x^2 - 3x + 2, xy - 1.5y - x + 1.5\}$

rank $H_\Lambda^{B_1^+} = 2$, *and* $L_1 = \{0\}$,

The algorithm stops with sucess, the flat extension is satisfied

$$Ker\ H_\Lambda^{B_1^+} = \{y - 1, x^2 - 3x + 2, xy - 1.5y - x + 1.5\}$$

and

$$I_{min} = (y - 1, x^2 - 3x + 2) = (y - 1, (x - 1) \cdot (x - 2))$$

5.3.3 Computing the minimizers

The remaining step is the computation of the minimizer points, once Algorithm 5.3.1 stops with `success` for $\Lambda^* \in \mathcal{L}_E(C)$ optimal for f. The minimizers can be computed from the eigenvalues of the multiplication operators $M_k : a \in \mathcal{A}_{min} \mapsto x_k a \in \mathcal{A}_{min}$ for $k = 1, \ldots, n$ where $\mathcal{A}_{min} = \mathbb{R}[\mathbf{x}]/I_{min}$ and $I_{min} = I_{\tilde{\Lambda}} = \mathcal{I}(\xi_1, \ldots, \xi_r)$.

Proposition 5.3.8 *The matrix of* M_k *in the basis* B *of* $\mathcal{A}_{min}$ *is* $[M_k] = (\frac{\Lambda^*(x_k\, b_i\, b_j)}{\Lambda^*(b_i\, b_i)})_{1 \le i,j \le r}$. *The operators* M_k, $k = 1 \ldots n$ *have* r *common eigenvectors* $\mathbf{u}_1, \ldots, \mathbf{u}_r$ *which satisfy* $M_k \mathbf{u}_i = \xi_{i,k} \mathbf{u}_i$, *with* $\xi_{i,k}$ *the* k^{th} *coordinate of the minimizer point* $\xi_i = (\xi_{i,1}, \ldots, \xi_{i,n}) \in S$.

Proof. By Proposition 5.3.6 and by definition of the inner-product in Definition 1.2.3 and recall also in (5.6), $B = \{b_1, \ldots, b_r\}$ is a basis of $\mathcal{A}_{\tilde{\Lambda}}$ and

$$x_k b_j \equiv \sum_{i=1}^{r} \frac{\Lambda^*(x_k\, b_i\, b_j)}{\Lambda^*(b_i\, b_i)} b_i \mod I_{min},$$

for $j = 1 \ldots r$, $k = 1 \ldots n$.
This yields the matrix of the operator M_k in the basis B: $[M_k] = (\frac{\Lambda^*(x_k\, b_i\, b_j)}{\Lambda^*(b_i\, b_i)})_{1 \le i,j \le r}$.

As the roots of I_{min} are simple, by [Elkadi 2007][Theorem 4.23] the eigenvectors of all M_k, $k = 1 \ldots n$ are the so-called idempotents $\mathbf{u}_1, \ldots, \mathbf{u}_r$ of $\mathcal{A}_{min}$

and the corresponding eigenvalues are $\xi_{1,k}, \ldots, \xi_{r,k}$. ■

Algorithm 5.3.2: MINIMIZER POINTS

Input: B and the relations as output in Algorithm 5.3.1.

- Compute the matrices $[M_k] = (\frac{\Lambda^*(x_k\, b_i\, b_j)}{\Lambda^*(b_i\, b_i)})_{1\leq i,j\leq r}$.
- For a generic choice of $l_1, \ldots, l_n \in \mathbb{R}$, compute the eigenvectors $\mathbf{u}_1, \ldots, \mathbf{u}_r$ of $l_1[M_1] + \cdots + l_n[M_n]$.
- Compute $\xi_{i,k} \in \mathbb{R}$ such that $M_k\mathbf{u}_i = \xi_{i,k}\mathbf{u}_i$.

Output: the minimizers $\xi_i = (\xi_{i,1}, \ldots, \xi_{i,n})$, $i = 1 \ldots r$.

Example 5.3.9 *With the basis B and the relations in the kernel which are solution of the problem 5.3.7 we can compute the multiplication matrices.*

$$M_x^{B=\{1,x-1.5\}} = \begin{pmatrix} 1.5 & 0.25 \\ 1 & 1.5 \end{pmatrix} \longrightarrow \begin{cases} x = 1.5 \cdot 1 + 1 \cdot (x - 1.5) \\ x(x-1.5) = 0.25 \cdot 1 + 1.5 \cdot (x - 1.5) \end{cases}$$

$$M_y^{B=\{1,x-1.5\}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \longrightarrow \begin{cases} y = 1 \cdot 1 + 0 \cdot (x - 1.5) \\ y(x-1.5) = 0 \cdot 1 + 1 \cdot (x - 1.5) \end{cases}$$

We take a linear combination of these matrices and compute its eigenvalues and eigenvectors

$$M = M_x^B + M_y^B = \begin{pmatrix} 2.5 & 0.25 \\ 1 & 1.5 \end{pmatrix} \longrightarrow \lambda_1 = 2, \lambda_2 = 3$$

$$M \cdot u_1 = \lambda_1 \cdot u_1 \rightarrow u_1^T = (-0.5, 1);\ M \cdot u_2 = \lambda_2 \cdot u_2 \rightarrow u_2^T = (0.5, 1)$$

From these eigenvectors, we compute the eigenvalues asociated to each multiplication matrix. Each eigenvalue computed corresponds to the i-coordinate of each minimizer point as we have seen in Proposition 5.3.8

$$M_x^B \cdot u_1^T = x_1 \cdot u_1^T \rightarrow x_1 = 1;\ M_x^B \cdot u_2^T = x_2 \cdot u_2^T \rightarrow x_2 = 2$$

$$M_y^B \cdot u_1^T = y_1 \cdot u_1^T \rightarrow y_1 = 1;\ M_y^B \cdot u_2^T = y_2 \cdot u_2^T \rightarrow y_2 = 1$$

The minimizers points are $(1, 1)$ *and* $(2, 1)$.

5.4 Minimizer border basis algorithm

In this Section we describe the algorithm to compute the minimum of a polynomial on S, i.e, solve our problem (2.1). It can be seen as a type of border basis algorithm, in which in the main loop we compute the border basis for a fix degree, we contruct the truncated moment problem asociated to this basis, we compute the optimal linear form, solving this truncated moment problem by Semidefinite programming (SDP) methods as we have seen in the Section 5.2.1. In the Chapter 6 we will see how this affects the perfomance of our algorithm. After computing the optimal linear form, we check when the minimun is reached through a new algorithm (explain in Section which verifies the flat extension, which was explained in Section 5.3 and which is an extension of [Laurent 2009a], which come from the flat extension condition of Curto and Fialkow [Curto 1996]. Eventually we can compute the minimizers points by multiplication matrices as we have seen in before subsection. The method is closely connected to the real radical border basis algorithm presented in [Lasserre 2012] but we include this new criterion to verify that the minimum is attained. In the some examples we compare the perfomance of our algorithm with Gloptipoly. Gloptipoly is a Matlab package developped by Jean Bernard Lasserre et Didier Henrion, which implements Lasserre relaxation [Lasserre. 2001] (see Section 3.3 of the Chapter 3). In order to verify the minimum is reached, they use two criterions (see [Henrion 2002]). The first consists in when the solution gives by the SDP problem satisfied all the original problem constraints and reach the objective fonction then the algorithm stops and give the solution. The second one is to verify the flat extension [Curto 1996, Laurent 2009a] by one criterion that consists in taking the complete (with all the monomials) Hankel matrix in a degree t, to compute the submatrices of this matrix and compare its ranks. This criterion can be found in the book [Lasserre 2009a]. There exits two main differences between this criterion and our new algorithm, the first is in many cases our algorithm is more quicker, because our algorithm only require the hankel matrix in a degree t, which is solution of the SDP and we verify the flat extension using the orthogal basis how we explain in the section before. The second difference is that this criterion produces numerical problem in some examples. We will see it in next section and in Chapter 6.

Algorithm 5.4.1: MINIMIZATION OF f ON S

Input: A real polynomial function f and a set of constraints $\mathbf{g} \subset \mathbb{R}[\mathbf{x}]$ with V_{min} non-empty finite.

1. Take $t = max(\lceil \frac{deg(f)}{2} \rceil, d^0, d^+)$ where $d^0 = max_{g^0 \in \mathbf{g}^0}(\lceil \frac{deg(g^0)}{2} \rceil), d^+ = max_{g^+ \in \mathbf{g}^+}(\lceil \frac{deg(g^+)}{2} \rceil)$
2. Compute the graded border basis F_{2t} of $\mathbf{g}^0$ for B in degree $2t$.
3. Let B_t be the set of monomials in B of degree $\leq t$.
4. Let C_t be the set of constraints such that $C_t^0 = \{m - \pi_{B_t,F_{2t}}(m),\ m \in B_t \cdot B_t\}$ and $C^+ = \pi_{B_t,F_{2t}}(g^+)$
5. $[f^*_{C_t,B_t}, \Lambda^*]$:= OPTIMAL LINEAR FORM(f, B_t, C_t).
6. $[c, B', K]$:= DECOMPOSITION(Λ^*, B_t) where $c =$`failed`, $B' = \emptyset, K = \emptyset$ or $c =$`success`, B' is the basis and K is the set of the relations.
7. if $c =$`success` then V=MINIMIZER POINTS(B', K)

   else go to step 2 with $t := t + 1$.

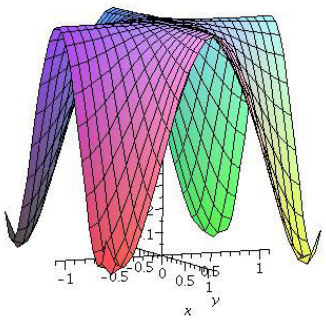
Output: the minimum $f^* = f^*_{C_t,B_t}$, the minimizers $V_{min} = V$, $I_{min} = (K)$ and B' such that K is a border basis for B'.

5.4.1 Example in detail

We consider the Motzkin polynomial,

$$\min \ f(x,y) = 1 + x^4y^2 + x^2y^4 - 3x^2y^2$$

which is non negative on $\mathbb{R}^2$ but not a sum of squares in $\mathbb{R}[x,y]$ as we proved in Example 3.1.18



We minimize with respect its gradient ideal, which is not zero-dimensional.

$$I_{grad}(f) = (-6xy^2 + 2xy^4 + 4x^3y^2, -6yx^2 + 2yx^4 + 4y^3x^2)$$

We also know that the minimizer ideal is zero-dimensional.

FIRST ITERATION

- $g^0 = \{-6xy^2 + 2xy^4 + 4x^3y^2, -6yx^2 + 2yx^4 + 4y^3x^2\}$, then $d_0 = 3$ and $t = 3$
- The border basis is $F_3 = \{x^4y - 3x^2y + 2x^2y^3, y^4x - 3y^2x + 2y^2x^3\}$
- The monomial basis associated to $t = 3$ is:

$$B_3 = \{1, x, y, x^2, xy, y^2, x^3, x^2y, xy^2, y^3\}$$

- The set of constraints

$$C_3^0 = \{\underline{x^5y} - 3x^3y + 2x^3y^3, \underline{y^5x} - 3xy^3 + 2x^3y^3, \underline{x^4y^2} - x^2y^2, \\ \underline{y^4x^2} - x^2y^2, \underline{x^4y} - 3x^2y + 2x^2y^3, \underline{y^4x} - 3y^2x + 2y^2x^3\}$$

hence we have 6 monomials that we can reduce.

- We compute $[f^*_{B_3,C_3}, \Lambda^*]$ =OPTIMAL LINEAR FORM (f, B_3, C_3),

  1. The size of the hankel matrix H^3_Λ is 10x10.
  2. The number of parameters that we look for are:

$$\begin{array}{rc} 27 & (\textit{total number of parameters}) \\ \underline{-6} & (\textit{number of reduced parameters}) \\ 21 & \textit{parameters.} \end{array}$$

  3. We find Λ^* such that $\Lambda^*(f) = -217$.

- When we verify the flat extension through the DECOMPOSITION algorithm:

$$[c, B', K] := \textsc{Decomposition}(\Lambda^*, B_3)$$

  we obtain $c =$ "$failed$", so we go to the second iteration with $t = 4$.

SECOND ITERATION

- We take $t = 4$
- The border basis is $F_4 = \{x^4y - 3x^2y + 2x^2y^3, y^4x - 3y^2x + 2y^2x^3\}$
- The monomial basis associated to $t = 4$ is:

$$B_4 = \{1, x, y, x^2, xy, y^2, x^3, x^2y, xy^2, y^3, x^4, x^3y, x^2y^2, xy^3, y^4\}$$

- The set of constraints C_4 is :

$$\begin{aligned} C^0_4 = \{&\underline{x^7y} - 9x^3y - 8x^3y^3, \underline{x^6y^2} - x^2y^2, \underline{x^5y^3} - x^3y^3, \underline{x^4y^4} - x^2y^2, \\ &\underline{x^3y^5} - x^3y^3, \underline{x^2y^6} - x^2y^2, \underline{xy^7} - 9xy^3 + 8x^3y^3, \underline{x^6y} - 9x^2y + 8x^2y^3, \\ &\underline{x^5y^2} - x^3y^2, \underline{x^4y^3} - x^2y^3, \underline{x^3y^4} - x^3y^2, \underline{x^2y^5} - x^2y^3, \underline{xy^6} - 9xy^2 + 8x^3y^2, \\ &\underline{x^5y} - 3x^3y + 2x^3y^3, \underline{y^5x} - 3xy^3 + 2x^3y^3, \underline{x^4y^2} - x^2y^2, \underline{y^4x^2} - x^2y^2, \\ &\underline{x^4y} - 3x^2y + 2x^2y^3, \underline{y^4x} - 3y^2x + 2y^2x^3\} \end{aligned}$$

  hence we have 19 monomials that we can reduce.

- We compute $[f^*_{B_4,C_4}, \Lambda^*]$ =OPTIMAL LINEAR FORM (f, B_4, C_4),

  1. The size of the hankel matrix H^4_Λ is 15x15

2. The number of parameters that we look for are:

$$\begin{array}{rc} 44 & (total\ number\ of\ parameters) \\ \underline{-19} & (number\ of\ reduced\ parameters) \\ 26 & parameters. \end{array}$$

3. We find Λ^* such that $\Lambda^*(f) = 0$, there is not duality gap, i.e, $f_4^\mu = f_4^{sos}$.

- When we verify the flat extension through the DECOMPOSITION algorithm:

$$[c, B', K] := \text{DECOMPOSITION}(\Lambda^*, B_4)$$

  - $c =$ "$sucess$" which imply $\Lambda^*(f) = f_4^\mu = f_4^{sos} = f^* = 0$
  - $B' = \{1, x, y, xy\}$
  - $K = \{x^2 - 1, y^2 - 1\}$

- We compute the minimizers points through the MINIMIZERS POINTS algorithm:MINIMIZER POINTS$(B', K)$and we obtain

$$\{(x = 1, y = 1), (x = 1, y = -1), (x = -1, y = 1), (x = -1, y = -1)\}$$

Remark 5.4.1 *For this example, Gloptipoly must go to the order $t = 9$ in order to detect that the optimum is reached and to compute the minimizers.*

5.4.2 Examples

Example 5.4.2 *We consider the following problem that corresponds to one example of type 4.3.5*

$$\begin{array}{rl} \min & f(x, y, z) = x^2 + y^2 + z^2; \\ s.t & \text{rank} \begin{pmatrix} x+z+1 & x+y & y+z \\ x+y & y+z & x+z+1 \end{pmatrix} \leq 1 \end{array}$$

or equivalently

$$\begin{array}{rl} \min & f(x, y, z) = x^2 + y^2 + z^2; \\ s.t & (x+z+1)(y+z) - (x+y)^2 = 0; \\ & (x+z+1)^2 - (y+z)(x+y) = 0; \\ & (x+z+1)(x+y) - (y+z)^2 = 0; \end{array}$$

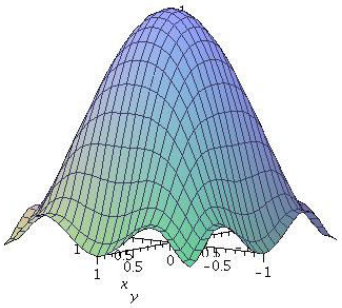
This corresponds to computing the closest point on a twisted cubic defined by 2×2 minors. The set of constraints $\mathbf{g}$ is not regular but $\mathcal{S}(\mathbf{g}) = \mathcal{V}^{\mathbb{R}}(\mathbf{g}^0)$ is a smooth real variety.

In the first iteration of the algorithm, the order is 1, the size of the Hankel matrix M_Λ^1 is 3, $\min \Lambda(f) = 1$ and there is no duality gap, i.e, strong duality holds. The flat extension condition is satisfied for M_Λ^1 and thus we have found the minimum. The algorithm stops and we obtain $I_{min} = (x, y-1, z)$. The minimizer point of f is $\{(x = 0, y = 1, z = 0)\}$.

Example 5.4.3 *We consider the Robinson polynomial*

$$\min \ f(x,y) = 1 + x^6 - x^4 - x^2 + y^6 - y^4 - y^2 - x^4y^2 - x^2y^4 + 3x^2y^2;$$

which is non negative on $\mathbb{R}^2$ but not a sum of squares in $\mathbb{R}[x, y]$.



We minimize f with respect its gradient ideal,
$I_{grad}(f) = (6x^5 - 4x^3 - 2x - 4x^3y^2 - 2xy^4 + 6xy^2, 6y^5 - 4y^3 - 2y - 4y^3x^2 - 2yx^4 + 6yx^2)$ which is not zero-dimensional.

In the first iteration, the order is 3, the size of the Hankel matrix M_Λ^3 is 10, $\min \Lambda(f) = -0.93$. The flat extension condition is not satisfied hence we try with degree 4.

In the second iteration the degree is 4, the size of the Hankel matrix M_Λ^4 is 15, $\min \Lambda(f) = 0$. There is no duality gap, i.e, $f_4^\mu = f_4^{sos} = 0$. The flat extension condition is satisfied.

The algorithm stops and we obtain $f^* = f_4^\mu = f_4^{sos} = 0$, $I_{min} = (x^3 - x, y^3 - y, x^2y^2 - x^2 - y^2 + 1)$. *The points that minimize* f *are* $\{(x = 1, y = 1), (x = 1, y = -1), (x = -1, y = 1), (x = -1, y = -1), (x = 1, y = 0), (x = -1, y = 0), (x = 0, y = 1), (x = 0, y = -1)\}$.

We remark that for this example, Gloptipoly must go to the order $t = 7$ *in order to detect that the optimum is reached and to compute the minimizers.*

Example 5.4.4 *We consider the homogeneous Motzkin polynomial with a perturbation* $\varepsilon = 0.005$,

$$\begin{array}{rl} \min & f(x,y,z) = x^4y^2 + x^2y^4 - 3x^2y^2z^2 + z^6 + \varepsilon(x^2 + y^2 + z^2); \\ s.t & h(x,y,z) = 1 - x^2 - y^2 - z^2 \geq 0 \end{array}$$

This example coming from [Laurent 2009a, Example 6.25] is a case where the constraints $\mathbf{g}$ *define a compact semi-algebraic set, but the direct relaxation using the associated quadratic module or preordering is not exact.*

We add the projection of the KKT ideal and we solce the following problem

$$\begin{array}{rl} \min & x^4y^2 + x^2y^4 - 3x^2y^2z^2 + z^6 + 0.005(x^2 + y^2 + z^2); \\ s.t & -4zx^4y - 20zx^2y^3 + 12x^2yz^3 - 0.06zy^5 + 12.06yz^5 = 0; \\ & -20zx^3y^2 - 4zxy^4 + 12xy^2z^3 - 0.06zx^5 + 12.06xz^5 = 0; \\ & (4x^3y^2 + 2xy^4 - 6xy^2z^2 + 0.03x^5)(-x^2 - y^2 - z^2 + 1) = 0; \\ & (2x^4y + 4x^2y^3 - 6x^2yz^2 + 0.03y^5)(-x^2 - y^2 - z^2 + 1) = 0; \\ & (-6x^2y^2z + 6.03z^5)(-x^2 - y^2 - z^2 + 1) = 0; \end{array}$$

where the first three equations are the 2×2 *minors of the Jacobian matrix of* f *and* h *and the last three equations are the gradient ideal of* f *multiplied by* h.

In the first iteration the order is 5, the size of the Hankel matrix M_Λ^5 *is 167,* $\min \Lambda(f) = 0$, *there is no duality gap. The flat extension condition is satisfied so the algorithm stops and we obtain* $I_{min} = (x, y, z)$.

The minimizer point of f *is* $(0, 0, 0)$.

For this example, the criterion to verify the flat extension in comparing rank of submatrices of the complete (all the monomials in degree t) Hankel matrix does not hold with Gloptipoly if $\varepsilon \leq 0.01$.

5.4.3 Examples of theorical results for ideals non zero-dimensional

Finally with these two last examples we show that even the minimizer ideal I_{min} is not zero-dimensional we can recover it from a solution of the relaxation problem.

Example 5.4.5 *We consider Motzkin polynomial over the unit ball:*

$$\begin{array}{rl} \min & f(x,y,z) = x^4y^2 + x^2y^4 - 3x^2y^2z^2 + z^6; \\ s.t & h(x,y,z) = 1 - x^2 - y^2 - z^2 \geq 0 \end{array}$$

The polynomial f is homogeneous and non negative on $\mathbb{R}^3$ but not a sum of squares in $\mathbb{R}[x,y,z]$.

We add the projections of KKT ideal and we have the similar problem

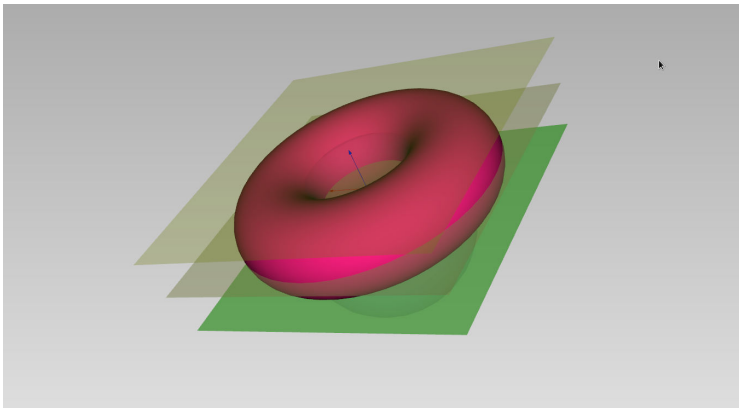
$$\begin{array}{rl} \min & x^4y^2 + x^2y^4 - 3x^2y^2z^2 + z^6; \\ s.t & -4xy^5 + 12xy^3z^2 + 4yx^5 - 12x^3yz^2 = 0; \\ & -4zx^4y - 20zx^2y^3 + 12x^2yz^3 + 12yz^5 = 0; \\ & -20zx^3y^2 - 4zxy^4 + 12xy^2z^3 + 12xz^5 = 0; \\ & (4x^3y^2 + 2xy^4 - 6xy^2z^2)(-x^2 - y^2 - z^2 + 1) = 0; \\ & (2x^4y + 4x^2y^3 - 6x^2yz^2)(-x^2 - y^2 - z^2 + 1) = 0; \\ & (-6x^2y^2z + 6z^5)(-x^2 - y^2 - z^2 + 1) = 0; \end{array}$$

where the first three equations are the 2×2 minors of the Jacobian matrix of f and h and the last three equations are the gradient ideal of f multiplied by h.

In the first iteration the order is 5, the size of the Hankel matrix M_Λ^5 is 156, $\min \Lambda(f) = 0$, there is no duality gap. We compute the kernel of this matrix: $\ker M_\Lambda^5 = \langle z(y^2 - z^2), x(y^2 - z^2), z(x^2 - z^2), y(x^2 - z^2) \rangle$. It generates the minimizer ideal $I_{min} = (z(y^2 - z^2), x(y^2 - z^2), z(x^2 - z^2), y(x^2 - z^2))$ defining 6 lines: $(y \pm z, x \pm z), (x, z), (y, z)$. Here $\mathcal{V}(I_{min})$ is not included in S.

Example 5.4.6 *We consider the minimization of a linear function on a torus:*

$$\begin{array}{rl} \min & f(x,y,z) = z \\ s.t & 9 - 10x^2 - 10y^2 + 6z^2 + x^4 + 2x^2y^2 + 2x^2z^2 + 2y^2z^2 + y^4 + z^4 = 0 \end{array}$$



In the first iteration, the order is 2, the size of the Hankel matrix M_Λ^2 *is* 10, $\min \Lambda(f) = -1$, *there is no duality gap. We compute the kernel of this matrix:* $\ker M_\Lambda^2 = \langle x^2 + y^2 - 4, x(z+1), y(z+1), z(z+1), (z+1)\rangle$ *which generates the minimizer ideal* $I_{min} = (x^2 + y^2 - 4, z + 1)$, *defining a circle which is the intersection of the torus with a tangent plane. Notice that the multiplicity of this intersection has been removed in* I_{min}.

Chapter 6

Experimentations, Applications and Implementation

In the first Section we analyse the practical behavior of this algorithm. In all the examples the minimizer ideal is zero-dimensional hence the algorithm stops in a finite number of steps and yields the minimizer points and generators of the minimizer ideal. In second Section shows some applications of the method in three different areas of research as Signal processing and Telecommunications (Best low-rank tensor approximation), Biology (Factors in the growth of the plant roots) and Electronic (Marx generators' design). In the last Section of this Chapter, we explain in more details the implementation of Algorithm 5.4.1.

6.1 Experimentations

In this Section we show the results of some experiments carried out on an Intel Core i5 2.40GHz. In these experiments, we compare the results between our algorithm 5.4.1 (*bbr*) and the full moment matrix relaxation algorithm (*fmr*) (inside the borderbasix package), that, as we said in the last Section, was described by Lasserre in [Lasserre 2009a], and it was also implemented in the package Gloptipoly of Matlab developed by D. Henrion and J.B. Lasserre.

In Table 6.1 and Table 6.2, we record the problem name or the source of the problem, the number of decision variables (v), the number of inequality and equality constraints (c), the maximum degree between the constraints and the polynomial to minimize (d), the number of minimizer points (sol). For the two algorithms *bbr* and *fmr* we report the total CPU time in seconds using SDPA (t) and using MOSEK ($t_{bbr+msk}$), the order of the relaxation (o), the number of parameters of the SDP problem (p) and the size of the moment matrices (s).

The first part of the table contains examples of positive polynomials, which are not sum of squares. New equality constraints are added following 4.2 to compute the minimizer points in the examples marked with $\diamond$. The fourth part of the table contains examples where the real radical $\sqrt[g^+]{\mathbf{g}^0}$ is computed.

When there are equality constraints, the border basis computation reduces the size of the moment matrices, as well as the number of parameters and the size of the localization matrices associated to the inequalities. This speeds up the SDP computation. In the case where there are only inequalities, the size of the moment matrices is the same but once the optimal linear form is computed using one of the SDP solvers SDPA, SDPA-GMP, CSDP or MOSEK, the DECOMPOSITION algorithm which verify the flat extension and computes the minimizers is more efficient and quicker than the reconstruction algorithm used in the full moment matrix relaxation approach.

The performance is not the only issue: numerical problems can also occur due to the bigger size of the moment matrices in the flat extension test and the reconstruction of minimizers. Such examples where the *fmr* algorithm fails because of the numerical rank problems are marked with *. The examples that Gloptipoly cannot treat due to the high number of variables [Lasserre 2009a] are marked with ** .

These experiments show that when the size of the SDP problems becomes significant, most of the time spent by our algorithm occurs during `sdp` computation and the border basis time and reconstruction time are negligible. We also show that the use of Mosek software reduces the time between 50 % and 80 %. In all the examples, the new border basis relaxation algorithm outperforms the full moment matrix relaxation method. In Table 6.3, we can see in more detail the difference in number of parameters, size of matrices and time between the algorithm *bbr* and *fmr*.

problem	v	c	d	sol	o_{fmr}	p_{fmr}	s_{fmr}	t_{fmr}
◇ Robinson	2	0	6	8	7	119	36	*
◇ Motzkin	2	0	6	4	9	189	55	*
◇ Motzkin perturbed	3	1	6	1	5	286	56	9.57
◇ [Lasserre. 2001], Ex. 1	2	0	4	1	2	14	6	0.050
◇ [Lasserre. 2001], Ex. 2	2	0	4	1	2	14	6	0.050
◇ [Lasserre. 2001], Ex. 3	2	0	6	4	8	152	45	*
[Lasserre. 2001], Ex. 5	2	3	2	3	2	14	6	0.053
[Floudas 1999], Ex. 4.1.4	1	2	4	2	2	4	3	0.040
[Floudas 1999], Ex. 4.1.6	1	2	6	2	3	6	4	0.044
[Floudas 1999], Ex. 4.1.7	1	2	4	1	2	4	3	0.042
[Floudas 1999], Ex. 4.1.8	2	5	4	1	2	14	6	0.077
[Floudas 1999], Ex. 4.1.9	2	6	4	1	4	44	15	0.29
[Floudas 1999], Ex. 2.1.1	5	11	2	1	3	461	56	12.23
[Floudas 1999], Ex. 2.1.2	6	13	2	1	2	209	26	1.29
[Floudas 1999], Ex. 2.1.3	13	35	2	1	2	2379	78	417.96
[Floudas 1999], Ex. 2.1.4	6	15	2	1	2	209	26	1.48
[Floudas 1999], Ex. 2.1.5	10	31	2	1	2	1000	66	44.29
[Floudas 1999], Ex. 2.1.6	10	25	2	1	2	1000	66	43.68
**[Floudas 1999], Ex. 2.1.7(1)	20	30	2	1	2	10625	231	35310.7
** [Floudas 1999], Ex. 2.1.7(5)	20	30	2	1	2	10625	231	36021.3
** [Floudas 1999], Ex. 2.1.8	24	58	2	1	2	20475	325	>14h
[Floudas 1999], Ex. 2.1.9	10	11	2	1	2	1000	55	16.76
[Floudas 1999], Ex. 3.1.3	6	16	2	1	2	209	26	1.42
[Lasserre 2009a] cbms1	3	3	3	5	3	83	20	0.20
[Lasserre 2009a] rediff3	3	3	2	2	2	35	10	0.09
[Lasserre 2009a] quadfor2	4	12	4	2	3	210	35	0.75
** simplex	15	16	2	1	2	3875	136	780.371

Table 6.1: Experimentations with fmr-algorithm

problem	v	c	d	sol	o_{bbr}	p_{bbr}	s_{bbr}	t_{bbr}	$t_{bbr+msk}$
◇ Robinson	2	0	6	8	4	21	15	0.15	0.10
◇ Motzkin	2	0	6	4	4	26	15	0.17	0.080
◇ Motzkin perturbed	3	1	6	1	5	167	56	3.78	0.90
◇ [Lasserre. 2001], Ex. 1	2	0	4	1	2	8	6	0.030	0.022
◇ [Lasserre. 2001], Ex. 2	2	0	4	1	2	8	6	0.030	0.022
◇ [Lasserre. 2001], Ex. 3	2	0	6	4	4	25	15	0.432	0.075
[Lasserre. 2001], Ex. 5	2	3	2	3	2	14	6	0.045	0.037
[Floudas 1999], Ex. 4.1.4	1	2	4	2	2	4	3	0.024	0.023
[Floudas 1999], Ex. 4.1.6	1	2	6	2	3	6	4	0.027	0.023
[Floudas 1999], Ex. 4.1.7	1	2	4	1	2	4	3	0.023	0.022
[Floudas 1999], Ex. 4.1.8	2	5	4	1	2	13	6	0.060	0.031
[Floudas 1999], Ex. 4.1.9	2	6	4	1	4	44	15	0.20	0.11
[Floudas 1999], Ex. 2.1.1	5	11	2	1	3	461	56	7.60	4.61
[Floudas 1999], Ex. 2.1.2	6	13	2	1	2	209	26	1.00	0.46
[Floudas 1999], Ex. 2.1.3	13	35	2	1	2	2379	78	383.97	34.55
[Floudas 1999], Ex. 2.1.4	6	15	2	1	2	209	26	1.01	0.43
[Floudas 1999], Ex. 2.1.5	10	31	2	1	2	1000	66	29.70	12.31
[Floudas 1999], Ex. 2.1.6	10	25	2	1	2	1000	66	28.60	6.05
**[Floudas 1999], Ex. 2.1.7(1)	20	30	2	1	2	10625	231	33219.9	1083.60
** [Floudas 1999], Ex. 2.1.7(5)	20	30	2	1	2	10625	231	33475.2	1117.33
** [Floudas 1999], Ex. 2.1.8	24	58	2	1	2	3875	136	3929.23	311.54
[Floudas 1999], Ex. 2.1.9	10	11	2	1	2	714	44	12.3	1.98
[Floudas 1999], Ex. 3.1.3	6	16	2	1	2	209	26	0.96	0.61
[Lasserre 2009a] cbms1	3	3	3	5	3	26	17	0.16	0.14
[Lasserre 2009a] rediff3	3	3	2	2	2	7	7	0.07	0.06
[Lasserre 2009a] quadfor2	4	12	4	2	3	48	19	0.6	0.45
** simplex	15	16	2	1	2	3059	120	674.534	65.73

Table 6.2: Experimentations with bbr-algorithm

problem	$p_{fmr} - p_{bbr}$	$s_{fmr} - s_{bbr}$	$t_{fmr} - t_{bbr}$
◇ Motzkin perturbed	286-167	56-56	9.57-3.78
◇ [Lasserre. 2001], Ex. 1	14-8	6-6	0.050-0.030
◇ [Lasserre. 2001], Ex. 2	14-8	6-6	0.050-0.030
[Lasserre. 2001], Ex. 5	14-14	6-6	0.053-0.045
[Floudas 1999], Ex. 4.1.4	4-4	3-3	0.040-0.024
[Floudas 1999], Ex. 4.1.6	6-6	4-4	0.044-0.027
[Floudas 1999], Ex. 4.1.7	4-4	3-3	0.042-0.023
[Floudas 1999], Ex. 4.1.8	14-13	6-6	0.077-0.060
[Floudas 1999], Ex. 4.1.9	44-44	15-15	0.29-0.20
[Floudas 1999], Ex. 2.1.1	461-461	56-56	12.23-7.60
[Floudas 1999], Ex. 2.1.2	209-209	26-26	1.29-1.00
[Floudas 1999], Ex. 2.1.3	2379-2379	78-78	417.96-383.97
[Floudas 1999], Ex. 2.1.4	209-209	26-26	1.48-1.01
[Floudas 1999], Ex. 2.1.5	1000-1000	66-66	44.29-29.70
[Floudas 1999], Ex. 2.1.6	1000-1000	66-66	43.68-28.60
**[Floudas 1999], Ex. 2.1.7(1)	10625-10625	231-231	35310.7-33219.9
** [Floudas 1999], Ex. 2.1.7(5)	10625-10625	231-231	36021.3-33475.2
** [Floudas 1999], Ex. 2.1.8	20475-3875	325-136	>14h-3929.23
[Floudas 1999], Ex. 2.1.9	1000-714	55-44	16.76-12.3
[Floudas 1999], Ex. 3.1.3	209-209	26-26	1.42-0.96
[Lasserre 2009a] cbms1	83-26	20-17	0.20-0.16
[Lasserre 2009a] rediff3	35-7	10-7	0.09-0.07
[Lasserre 2009a] quadfor2	210-48	35-19	0.75-0.6
** simplex	3875-3059	136-120	780.371-674.534

Table 6.3: Differents between fmr-bbr algorithms

6.2 Applications

In this Section we present some applications of our algorithm in some differents domains as Signal processing and Telecommunications (Best low-rank tensor approximation), Biology (Factors in the growth of the plant roots) and Electronic (Marx generators' design). First of all we explain in more detail the differents applications and after that we show how to apply our algorithm.

6.2.1 Best low-rank tensor approximation

Tensor problems appears in many context and applications. For instance matrices are an example of tensors of order 2. However in many problems, higher

order tensors are naturally used to collect informations which depend on more than two variables. These data could be observations of some experimentation or of a physical phenomena that depends on several parameters. These parameters are stored in a structure called tensor, according to the dimensional parameters of the problem. For a survey on tensors and applications (see [Comon 2000]). The tensor decomposition problem consits in decomposing the tensor into a minimal sum of indecomposable tensor, i.e, tensors of rank 1. We refer to [Bernardi 2013], [Brachat 2010] for more material in tensor decomposition. This decomposition allows to extract invariants properties and geometry of the tensor. For this reason this tensor decompostion problem appears in many domains. For example Singular Value Decomposition for the case of tensor of rank 2 and for higher rank this approach appears in several domains such as Electrical Engineering, Signal processing, Telecommunications or Data Analysis.
In many problems, tensor coefficients are given with some error. Instead of searching an exact decomposition, one can consider the following approximate decomposition problem: Is there in the neighborhood of a given tensor, a tensor with a small rank? This problem can be seen as an optimization problem. If we fix the rank r, the problem reduces to compute the tensors of rank r which are th closest to the input data. Hereafter we study the best rank-1 and rank-2 tensor approximation. As we will see, in the case of rank-1 mnimizing the value of the distance is equivalent for symmetric tensors to maximize homogeneous polynomial over unit spheres et for nonsymmetric tensors, this is equivalent to maximize multihomogeneous forms over multispheres. In particular we study the examples in the papers [Nie 2013b, Ottaviani 2013] and we compare our results with the results given in those papers. In the first part of this Section we treat the best rank-1 tensor approximation problem and in a second part we treat the best rank-2 tensor approximation problem. Before to explain these problems we introduce some notations.

Definition 6.2.1 *A tensor de order* $m \in \mathbb{Z}^+$ *and dimension* $(n_1, ..., n_m) \in \mathbb{Z}_+^m$ *is an array* $\mathcal{F}$ *that is indexed by integer tuples* $(i_1, ..., i_m)$ *with* $1 \leq i_j \leq n_j (j = 1, ..m)$, *i.e,*

$$\mathcal{F} = (\mathcal{F}_{i_1,...,i_m})_{1 \leq i_1 \leq n_1, ..., 1 \leq i_m \leq n_m}$$

Remark 6.2.2 *The spaces of all such tensors with real entries is denoted as* $\mathbb{R}^{n_1 \times \cdots \times n_m}$

Definition 6.2.3 *A tensor* $\mathcal{F} \in \mathbb{R}^{n_1 \times \cdots \times n_m}$ *is symmetric if* $n_1 = \cdots = n_m$ *and*

$$\mathcal{F}_{i_1,...,i_m} = \mathcal{F}_{j_1,...,j_m} \quad \forall \ (i_1, ..., i_m) \sim (j_1, ..., j_m)$$

where $\sim$ *means that* $(i_1, ..., i_m)$ *is a permutation of* $(j_1, ..., j_m)$.

Remark 6.2.4 *We denote* $S^m(\mathbb{R}^n)$ *the space of real symmetric tensors of order* m *in dimension* n.

Definition 6.2.5 *We define the scalar product of two tensor* $\mathcal{F}, \mathcal{F}' \in \mathbb{R}^{n_1 \times \cdots \times n_m}$ *as:*

$$\langle \mathcal{F}, \mathcal{F}' \rangle = \sum_{i_1=1}^{n_1} \cdots \sum_{i_m=1}^{n_m} \mathcal{F}_{i_1,...,i_m} \mathcal{F}'_{i_1,...,i_m}$$

Definition 6.2.6 *We define the norm of a tensor* $\mathcal{F} \in \mathbb{R}^{n_1 \times \cdots \times n_m}$ *as:*

$$|| \mathcal{F} || = (\sum_{i_1=1}^{n_1} \cdots \sum_{i_m=1}^{n_m} | \mathcal{F}_{i_1,...,i_m} |^2)^{1/2} = \langle \mathcal{F}, \mathcal{F} \rangle$$

Definition 6.2.7 *Every tensor can be expressed as a linear combination of outer products of vectors, i.e,*

$$\mathcal{F} = \sum_{i=1}^{r} u^{i,1} \otimes \cdots \otimes u^{i,m} \quad with\ u^{i,j} \in \mathbb{R}^{n_j}$$

And the smallest positive integer r *is called rank of* $\mathcal{F}$.

6.2.1.1 Best rank-1 tensor approximation

The problem of find the best rank-1 tensor approximation of a tensor $\mathcal{F} \in \mathbb{R}^{n_1 \times \cdots \times n_m}$ can be described as follows.

Problem :*Given a tensor* $\mathcal{F} \in \mathbb{R}^{n_1 \times \cdots \times n_m}$, *find a scalar* λ *and unit vectors* $u^1, \ldots, u^m$ *such that the rank-1 tensor* $\tilde{\mathcal{F}} = \lambda u^1 \otimes \cdots \otimes u^m$ *minimizes the least-squares cost function*

$$f(\tilde{\mathcal{F}}) = || \mathcal{F} - \tilde{\mathcal{F}} ||^2 \tag{6.1}$$

This constrained optimization problem can be analyzed using techninque of Lagrange multipliers. Therefore, we consider the following combination of f with the constraint terms

$$\tilde{f} = \sum_{i_1=1}^{n_1} \cdots \sum_{i_m=1}^{n_m} (\mathcal{F}_{i_1,...,i_m} - \lambda u^1_{i_1} \cdots u^m_{i_m})^2 + \sum_j \lambda_j (\sum_{i_j} (u^j_{i_j})^2 - 1) \tag{6.2}$$

in which $\lambda_j (1 \leq j \leq m)$ are the lagrange multipliers. Setting the derivative with respect to $u_{i_k}^j$ equal to zero yields

$$\lambda \sum_{i_1=1}^{n_1} \cdots \sum_{i_{k-1}=1}^{n_{k-1}} \sum_{i_{k+1}=1}^{n_{k+1}} \cdots \sum_{i_m=1}^{n_m} \mathcal{F}_{i_1,\ldots,i_m} u_{i_1}^1 \cdots u_{i_{k-1}}^{k-1} u_{i_{k+1}}^{k+1} \cdots u_{i_m}^m =$$

$$= \lambda_j u_{i_j}^j + \lambda^2 u_{i_j}^j \sum_{i_1=1}^{n_1} \cdots \sum_{i_{k-1}=1}^{n_{k-1}} \sum_{i_{k+1}=1}^{n_{k+1}} \cdots \sum_{i_m=1}^{n_m} (u_{i_1}^1)^2 \cdots (u_{i_{k-1}}^{k-1})^2 (u_{i_{k+1}}^{k+1})^2 \cdots (u_{i_m}^m)^2. \tag{6.3}$$

If we derive with respect λ_j and λ, we obtain, respectively

$$\sum_{i_j} (u_{i_j}^j)^2 = 1 \tag{6.4}$$

$$\sum_{i_1=1}^{n_1} \cdots \sum_{i_m=1}^{n_m} \mathcal{F}_{i_1,\ldots,i_m} u_{i_1}^1 \cdots u_{i_m}^m = \lambda \sum_{i_1,\ldots,i_m} (u_{i_1}^1)^2 \cdots (u_{i_m}^m)^2. \tag{6.5}$$

We combine (6.4) with (6.5) and with the right-side of (6.3) and we obtain,

$$\sum_{i_1=1}^{n_1} \cdots \sum_{i_m=1}^{n_m} \mathcal{F}_{i_1,\ldots,i_m} u_{i_1}^1 \cdots u_{i_m}^m = \lambda \tag{6.6}$$

$$\lambda \sum_{i_1=1}^{n_1} \cdots \sum_{i_{k-1}=1}^{n_{k-1}} \sum_{i_{k+1}=1}^{n_{k+1}} \cdots \sum_{i_m=1}^{n_m} \mathcal{F}_{i_1,\ldots,i_m} u_{i_1}^1 \cdots u_{i_{k-1}}^{k-1} u_{i_{k+1}}^{k+1} \cdots u_{i_m}^m = (\lambda^2 + \lambda^j) u_{i_j}^j \tag{6.7}$$

Combining (6.4) with the right-side of (6.7) and compraing this to (6.6), yields

$$\sum_{i_1=1}^{n_1} \cdots \sum_{i_{k-1}=1}^{n_{k-1}} \sum_{i_{k+1}=1}^{n_{k+1}} \cdots \sum_{i_m=1}^{n_m} \mathcal{F}_{i_1,\ldots,i_m} u_{i_1}^1 \cdots u_{i_{k-1}}^{k-1} u_{i_{k+1}}^{k+1} \cdots u_{i_m}^m = \lambda u_{i_j}^j \tag{6.8}$$

Thus, the Lagrange equations correspond to $(1 \leq j \leq m)$:

$$\sum_{i_1=1}^{n_1} \cdots \sum_{i_{k-1}=1}^{n_{k-1}} \sum_{i_{k+1}=1}^{n_{k+1}} \cdots \sum_{i_m=1}^{n_m} \mathcal{F}_{i_1,\ldots,i_m} u_{i_1}^1 \cdots u_{i_{k-1}}^{k-1} u_{i_{k+1}}^{k+1} \cdots u_{i_m}^m = \lambda u_{i_j}^j \tag{6.9}$$

$$\sum_{i_1=1}^{n_1} \cdots \sum_{i_m=1}^{n_m} \mathcal{F}_{i_1,\ldots,i_m} u_{i_1}^1 \cdots u_{i_m}^m = \lambda \tag{6.10}$$

$$|| u^j ||= 1, \tag{6.11}$$

De Lathauwer, De Moor and Vandenwalle proved in [De Lathauwer 2000] that the problem (6.1) is is equivalent to maximizer,

$$\begin{aligned} &\max_{u^1 \in \mathbb{R}^{n_1}, \ldots, u^m \in \mathbb{R}^{n_m}} \mid g(u^1, ..., u^m) \mid \\ &s.t \mid\mid u^1 \mid\mid = \cdots = \mid\mid u^m \mid\mid = 1 \end{aligned} \tag{6.12}$$

where $g(u^1, ..., u^m) = \sum_{1 \leq i_1 \leq n_1, \ldots, 1 \leq i_m \leq n_m} \mathcal{F}_{i_1, \ldots, i_m} \cdot (u^1)_{i_1} \cdots (u^m)_{i_m}$.

The exact theorem is the following:

Theorem 6.2.8 *For a tensor $\mathcal{F} \in \mathbb{R}^{n_1 \times \cdots \times n_m}$, the minimization of the cost function (6.1) is equivalent to the maximization problem*

$$\begin{aligned} &\max_{u^1 \in \mathbb{R}^{n_1}, \ldots, u^m \in \mathbb{R}^{n_m}} \mid g(u^1, ..., u^m) \mid \\ &s.t \mid\mid u^1 \mid\mid = \cdots = \mid\mid u^m \mid\mid = 1 \end{aligned}$$

If the scalar λ is chosen in accordance with (6.10), then (6.1) and (6.12) are related by

$$f = \mid\mid \mathcal{F} \mid\mid^2 - g^2 \tag{6.13}$$

Proof. We have the following:

$$f(\tilde{\mathcal{F}}) = \mid\mid \mathcal{F} - \tilde{\mathcal{F}} \mid\mid^2 = \mid\mid \mathcal{F} \mid\mid^2 - 2\langle \mathcal{F}, \tilde{\mathcal{F}} \rangle + \mid\mid \tilde{\mathcal{F}} \mid\mid^2$$

According to the definition of λ, the value taken by $\langle \mathcal{F}, \tilde{\mathcal{F}} \rangle$ equals λ^2. Since $u^1, ..., u^m$ have unit-norm, $\mid\mid \tilde{\mathcal{F}} \mid\mid^2 = 1$ as well. Combining it with the definition of g proves the theorem. ∎

The equivalent result for symetric tensors $S^m(\mathbb{R}^n)$ is the following

Theorem 6.2.9 *For a tensor $\mathcal{F} \in S^m(\mathbb{R}^n)$, find a scalar λ and unit vectors $u^1, \ldots, u^m$ such that the rank-1 tensor $\tilde{\mathcal{F}} = \lambda u^1 \otimes \cdots \otimes u^m$ minimizes the least-squares cost function*

$$f(\tilde{\mathcal{F}}) = \mid\mid \mathcal{F} - \tilde{\mathcal{F}} \mid\mid^2 \tag{6.14}$$

is equivalent to the maximization problem

$$\begin{aligned} &\max_{u \in \mathbb{R}^n} \mid g(u, ..., u) \mid \\ &s.t \mid\mid u \mid\mid = 1 \end{aligned} \tag{6.15}$$

If the scalar λ is chosen in accordance with (6.10), then (6.1) and (6.12) are related by

$$f = \mid\mid \mathcal{F} \mid\mid^2 - g^2 \tag{6.16}$$

Remark 6.2.10 *We notice that the problem* (6.15) *is equivalent to*

$$\max_{u \in \mathbb{S}^{n-1}} \mid F(u, ..., u) \mid \tag{6.17}$$

where $\mathbb{S}^{n-1}$ *is the n-1 dimensional unit sphere.*

In Table 6.4, we apply our algorithm *bbr+msk* to find the best rank-1 approximation for symmetric and non symmetric tensors on examples from [Nie 2013b]. For the problems with several minimizers (which is the case when there are symmetries), the method proposed in [Nie 2013b] cannot certify the result and uses a local method to converge to a local extrema. We apply the border basis relaxation algorithm to find all the global minimizers for the best rank 1 approximation problem.

problem	v	c	d	sol	o_{bbr}	p_{bbr}	s_{bbr}	$t_{bbr+msk}$
[Nie 2013b] Ex. 3.1	2	1	3	1	2	8	5	0.028
[Nie 2013b] Ex. 3.2	3	1	3	1	2	24	9	0.025
[Nie 2013b] Ex. 3.3	3	1	3	1	2	24	9	0.035
[Nie 2013b] Ex. 3.4	4	1	4	2	2	24	9	0.097
[Nie 2013b] Ex. 3.5	5	1	3	1	2	104	20	0.078
[Nie 2013b] Ex. 3.6	5	1	4	2	4	824	105	15.39
[Nie 2013b] Ex. 3.8	3	1	6	4	3	48	16	1.14
[Nie 2013b] Ex. 3.11	8	4	4	8	3	84	25	0.17
[Nie 2013b] Ex. 3.12	9	3	3	4	2	552	52	1.55
[Nie 2013b] Ex. 3.13	9	3	3	12	3	3023	190	223.27
[Ottaviani 2013] Ex. 4.2	6	0	8	4	8	2340	210	59.38

Table 6.4: Best rank-1 and rank-2 approximation tensors

We explain in more details the examples in Table 6.4 which have several minimizers and in particular with all the details the example 3.4:

Example 6.2.11 (Example 3.4) *Consider the tensor* $\mathcal{F} \in S^4(\mathbb{R}^3)$ *with entries:*
$\mathcal{F}_{1111} = 0.2883, \mathcal{F}_{1112} = -0.0031, \mathcal{F}_{1113} = 0.1973, \mathcal{F}_{1122} = -0.2458,$
$\mathcal{F}_{1123} = -0.2939, \mathcal{F}_{1133} = 0.3847, \mathcal{F}_{1222} = 0.2972, \mathcal{F}_{1223} = 0.1862,$
$\mathcal{F}_{1233} = 0.0919, \mathcal{F}_{1333} = -0.3619, \mathcal{F}_{2222} = 0.1241, \mathcal{F}_{2223} = -0.3420,$
$\mathcal{F}_{2233} = 0.2127, \mathcal{F}_{2333} = 0.2727, \mathcal{F}_{3333} = -0.3054$

$$\min \mid\mid \mathcal{F} - \lambda \cdot u_i^{\otimes 4} \mid\mid$$

$$\Updownarrow$$

$$\max\ 0.2883x_0^4-4\cdot0.0031x_0^3x_1+4\cdot0.1973x_0^3x_2-6\cdot0.2485x_0^2x_1^2-12\cdot0.2939x_0^2x_1x_2+$$
$$+6\cdot0.3847x_0^2x_2^2+4\cdot0.2979x_0x_1^3+12\cdot0.1862x_0x_1^2x_2+12\cdot0.0919x_0x_1x_2^2+$$
$$-4\cdot0.3619x_0x_2^3+0.1241x_1^4-4\cdot0.3420x_1^3x_2+6\cdot0.2127x_1^2x_2^2+$$
$$+4\cdot0.2727x_1x_2^3-0.3054x_2^4;$$
$$s.t\ x_0^2+x_1^2+x_2^2=1;$$

We get the rank-1 tensor $\lambda\cdot u_i^{\otimes 3}$ *with:*
$\lambda=-1.0960,\ u_1=(-0.59148,0.7467,0.3042); u_2=$
$(0.59148,-0.7467,-0.3042)$ *and* $||\ \mathcal{F}-\lambda\cdot u_i^{\otimes 4}\ ||=1.9683.$

Example 6.2.12 (Example 3.6) *Consider the tensor* $\mathcal{F}\in S^4(\mathbb{R}^5)$ *with entries:*
$\mathcal{F}_{i_1,i_2,i_3,i_4}=arctan((-1)^{i_1}\frac{i_1}{5})+arctan((-1)^{i_2}\frac{i_2}{5})+arctan((-1)^{i_3}\frac{i_3}{5})+arctan((-1)^{i_4}\frac{i_4}{5})$

We get the rank-1 tensor $\lambda\cdot u^{\otimes 4}$ *with:*
$\lambda=-23.56525,\ u_1=(0.4398,0.2383,0.5604,0.1354,0.6459);$
$u_2=(-0.4398,-0.2383,-0.5604,-0.1354,-0.6459)$ *and*
$||\ \mathcal{F}-\lambda\cdot u_i^{\otimes 4}\ ||=16.8501.$

Example 6.2.13 (Example 3.8) *Consider the tensor* $\mathcal{F}\in S^6(\mathbb{R}^3)$ *with entries:*
$\mathcal{F}_{111111}=2,\mathcal{F}_{111122}=1/3,\mathcal{F}_{111133}=2/5,\mathcal{F}_{112222}=1/3,\mathcal{F}_{112233}=1/6,$
$\mathcal{F}_{113333}=2/5,\mathcal{F}_{222222}=2,\mathcal{F}_{222233}=2/5,\mathcal{F}_{223333}=2/5,\mathcal{F}_{333333}=1$

We get the rank-1 tensor $\lambda\cdot u_i^{\otimes 6}$ *with:*
$\lambda=2,\ u_1=(1,0,0);\ u_2=(-1,0,0);\ u_3=(0,1,0);\ u_4=(0,-1,0)$ *and*
$||\ \mathcal{F}-\lambda\cdot u_i^{\otimes 6}\ ||=20.59.$

Example 6.2.14 (Example 3.11) *Consider the tensor* $\mathcal{F}\in\mathbb{R}^{2\times2\times2\times2}$ *with entries:*
$\mathcal{F}_{1111}=25.1,\mathcal{F}_{1212}=25.6,\mathcal{F}_{2121}=24.8,\mathcal{F}_{2222}=23$

We get the rank-1 tensor $\lambda\cdot u_i^1\otimes u_i^2\otimes u_i^3\otimes u_i^4$ *with:*
$\lambda=25.6,\ u_1^1=(1,0),u_1^2=(0,1),u_1^3=(1,0),u_1^4=(0,1);$
$u_2^1=(-1,0),u_2^2=(0,-1),u_2^3=(-1,0),u_2^4=(0,-1);$
$u_3^1=(-1,0),u_3^2=(0,-1),u_3^3=(1,0),u_3^4=(0,1);$
$u_4^1=(1,0),u_4^2=(0,1),u_4^3=(-1,0),u_4^4=(0,-1);$
$u_5^1=(-1,0),u_5^2=(0,1),u_5^3=(-1,0),u_5^4=(0,1);$
$u_6^1=(1,0),u_6^2=(0,-1),u_6^3=(1,0),u_6^4=(0,-1);$

$u_7^1 = (1, 0), u_7^2 = (0, -1), u_7^3 = (-1, 0), u_7^4 = (0, 1)$;
$u_8^1 = (-1, 0), u_8^2 = (0, 1), u_8^3 = (1, 0), u_8^4 = (0, -1)$.
The distance between $\mathcal{F}$ *and one of these solutions is*
$|| \mathcal{F} - \lambda \cdot u_i^1 \otimes u_i^2 \otimes u_i^3 \otimes u_i^4 || = 42.1195$.

Example 6.2.15 (Example 3.12) *Consider the tensor* $\mathcal{F} \in \mathbb{R}^{3\times3\times3}$ *with entries:*
$\mathcal{F}_{111} = 0.4333, \mathcal{F}_{121} = 0.4278, \mathcal{F}_{131} = 0.4140, \mathcal{F}_{211} = 0.8154, \mathcal{F}_{221} = 0.0199,$
$\mathcal{F}_{231} = 0.5598, \mathcal{F}_{311} = 0.0643, \mathcal{F}_{321} = 0.3815, \mathcal{F}_{331} = 0.8834, \mathcal{F}_{112} = 0.4866,$
$\mathcal{F}_{122} = 0.8087, \mathcal{F}_{132} = 0.2073, \mathcal{F}_{212} = 0.7641, \mathcal{F}_{222} = 0.9924, \mathcal{F}_{232} = 0.8752,$
$\mathcal{F}_{312} = 0.6708, \mathcal{F}_{322} = 0.8296, \mathcal{F}_{332} = 0.1325, \mathcal{F}_{113} = 0.3871, \mathcal{F}_{123} = 0.0769,$
$\mathcal{F}_{133} = 0.3151, \mathcal{F}_{213} = 0.1355, \mathcal{F}_{223} = 0.7727, \mathcal{F}_{233} = 0.4089, \mathcal{F}_{313} = 0.9715,$
$\mathcal{F}_{323} = 0.7726, \mathcal{F}_{333} = 0.5526$

We get the rank-1 tensor $\lambda \cdot u_i^1 \otimes u_i^2 \otimes u_i^3$ *with:*
$\lambda = 2.8166,\ u_1^1 = (0.4279, 0.6556, 0.62209), u_1^2 = (0.5705, 0.6466, 0.5063),$
$u_1^3 = (0.4500, 0.7093, 0.5425);$
$u_2^1 = (0.4279, 0.6556, 0.62209), u_2^2 = (-0.5705, -0.6466, -0.5063)$
$u_2^3 = (-0.4500, -0.7093, -0.5425);$
$u_3^1 = (-0.4279, -0.6556, -0.62209), u_3^2 = (0.5705, 0.6466, 0.5063),$
$u_3^3 = (-0.4500, -0.7093, -0.5425);$
$u_4^1 = (-0.4279, -0.6556, -0.62209), u_4^2 = (-0.5705, -0.6466, -0.5063),$
$u_4^3 = (0.4500, 0.7093, 0.5425),$
The distance between $\mathcal{F}$ *and one of these solutions is* $|| \mathcal{F} - \lambda \cdot u_i^1 \otimes u_i^2 \otimes u_i^3 || =$ 1.3510.

Example 6.2.16 (Example 3.13) *Consider the tensor* $\mathcal{F} \in \mathbb{R}^{3\times3\times3}$ *with entries:*
$\mathcal{F}_{111} = 0.0072, \mathcal{F}_{121} = -0.4413, \mathcal{F}_{131} = 0.1941, \mathcal{F}_{211} = -04413, \mathcal{F}_{221} = 0.0940,$
$\mathcal{F}_{231} = 0.5901, \mathcal{F}_{311} = 0.1941, \mathcal{F}_{321} = -0.4099, \mathcal{F}_{331} = -0.1012, \mathcal{F}_{112} = -0.4413,$
$\mathcal{F}_{122} = 0.0940, \mathcal{F}_{132} = -0.4099, \mathcal{F}_{212} = 0.0940, \mathcal{F}_{222} = 0.2183, \mathcal{F}_{232} = 0.2950,$
$\mathcal{F}_{312} = 0.5901, \mathcal{F}_{322} = 0.2950, \mathcal{F}_{332} = 0.2229, \mathcal{F}_{113} = 0.1941, \mathcal{F}_{123} = 0.5901,$
$\mathcal{F}_{133} = -01012, \mathcal{F}_{213} = -0.4099, \mathcal{F}_{223} = 0.2950, \mathcal{F}_{233} = 0.2229, \mathcal{F}_{313} = -0.1012,$
$\mathcal{F}_{323} = 0.2229, \mathcal{F}_{333} = -0.4891$

We get the rank-1 tensor $\lambda \cdot u_i^1 \otimes u_i^2 \otimes u_i^3$ *with* $\lambda = 1.000$ *and the 12 solutions*
$u_1^1 = (0.7955, 0.2491, 0.5524), u_1^2 = (-0.0050, 0.9142, -0.4051),$

$u_1^3 = (-0.6060, 0.3195, 0.7285)$;
$u_2^1 = (-0.0050, 0.9142, -0.4051), u_2^2 = (-0.6060, 0.3195, 0.7285)$,
$u_2^3 = (0.7955, 0.2491, 0.5524)$;
$u_3^1 = (-0.6060, 0.3195, 0.7285), u_3^2 = (0.7955, 0.2491, 0.5524)$,
$u_3^3 = (-0.0050, 0.9142, -0.4051)$;
$u_4^1 = (0.7955, 0.2491, 0.5524), u_4^2 = (0.0050, -0.9142, 0.4051)$,
$u_4^3 = (0.6060, -0.3195, -0.7285)$;
$u_5^1 = (0.6060, -0.3195, -0.7285), u_5^2 = (0.7955, 0.2491, 0.5524)$,
$u_5^3 = (0.0050, -0.9142, 0.4051)$;
$u_6^1 = (-0.6060, 0.3195, 0.7285), u_6^2 = (-0.7955, -0.2491, -0.5524)$,
$u_6^3 = (0.0050, -0.9142, 0.4051)$;
$u_7^1 = (0.6060, -0.3195, -0.7285), u_7^2 = (-0.7955, -0.2491, -0.5524)$,
$u_7^3 = (-0.0050, 0.9142, -0.4051)$;
$u_8^1 = (-0.7955, -0.2491, -0.5524), u_8^2 = (-0.0050, 0.9142, -0.4051)$,
$u_8^3 = (0.6060, -0.3195, -0.7285)$;
$u_9^1 = (-0.7955, -0.2491, -0.5524), u_9^2 = (0.0050, -0.9142, 0.4051)$,
$u_9^3 = (-0.6060, 0.3195, 0.7285)$;
$u_{10}^1 = (-0.0050, 0.9142, -0.4051), u_{10}^2 = (0.6060, -0.3195, -0.7285)$,
$u_{10}^3 = (-0.7955, -0.2491, -0.5524)$;
$u_{11}^1 = (0.0050, -0.9142, 0.4051), u_{11}^2 = (0.6060, -0.3195, -0.7285)$,
$u_{11}^3 = (0.7955, 0.2491, 0.5524)$;
$u_{12}^1 = (0.0050, -0.9142, 0.4051), u_{12}^2 = (-0.6060, 0.3195, 0.7285)$,
$u_{12}^3 = (-0.7955, -0.2491, -0.5524)$.
The distance between $\mathcal{F}$ *and one of these solutions is* $|| \mathcal{F} - \lambda \cdot u_i^1 \otimes u_i^2 \otimes u_i^3 ||=$ 1.4143.

6.2.1.2 Best rank-2 tensor approximation

Given a tensor $\mathcal{F} \in \mathbb{R}^{n_1 \times \cdots \times n_m}$ that can be symmetric or nonsymmetric, we say that a tensor $\mathcal{T}$ is a best rank-2 approximation of $\mathcal{F}$ if it is a minimizer of the least squares problem

$$\min_{\mathcal{X} \in \mathbb{R}^{n_1 \times \cdots \times n_m}, rank \mathcal{X} = 2} || \mathcal{F} - \mathcal{X} ||^2 \tag{6.18}$$

The last example in Table 6.4 is a best rank-2 tensor approximation example from the paper [Ottaviani 2013]. We explain in more detail this problem

Example 6.2.17 (Example 4.2) *Consider the tensor* $\mathcal{F} \in S^4(\mathbb{R}^3)$ *with entries*

$\mathcal{F}_{1111} = 0.1023, \mathcal{F}_{1112} = -0.002, \mathcal{F}_{1113} = 0.0581, \mathcal{F}_{1122} = 0.0039, \mathcal{F}_{1123} = -0.00032569,$
$\mathcal{F}_{1133} = 0.0407, \mathcal{F}_{1222} = 0.0107, \mathcal{F}_{1223} = -0.0012, \mathcal{F}_{1233} = -0.0011, \mathcal{F}_{1333} = 0.0196,$
$\mathcal{F}_{2222} = 0.0197, \mathcal{F}_{2223} = -0.0029, \mathcal{F}_{2233} = -0.00017418, \mathcal{F}_{2333} = -0.0021,$
$\mathcal{F}_{3333} = 0.1869$

We get the rank-2 tensor $\tilde{\mathcal{F}}(s,t,u) = (as+bt+cu)^4 + (ds+et+fu)^4$ *with the 8 solutions:*
$s_1 = (a,b,c,d,e,f) = (0.01877, 0.006239, -0.6434, -0.5592, 0.008797, -0.3522);$
$s_2 = (-0.01877, -0.006239, 0.6434, 0.5592, -0.008797, 0.3522);$
$s_3 = (0.01877, 0.006239, -0.6434, 0.5592, -0.008797, 0.3522);$
$s_4 = (-0.01877, -0.006239, 0.6434, -0.5592, 0.008797, -0.3522);$
$s_5 = (-0.5592, 0.008797, -0.3522, 0.01877, 0.006239, -0.6434);$
$s_6 = (0.5592, -0.008797, 0.3522, -0.01877, -0.006239, 0.6434);$
$s_7 = (-0.5592, 0.008797, -0.3522, -0.01877, -0.006239, 0.6434);$
$s_8 = (0.5592, -0.008797, 0.3522, 0.01877, 0.006239, -0.6434).$
The distance between $\mathcal{F}$ *and one of these solutions is* $|| \mathcal{F} - \tilde{\mathcal{F}} || = 0.00108483$. *The other possible real rank-2 approximations* $\tilde{\mathcal{F}}(s,t,u) = \pm(as+bt+cu)^4 \pm (ds+et+fu)^4$ *yield solutions which are not as close to* $\mathcal{F}$ *as these solutions. The eight solutions come from the symmetries due to the invariance of the solution set by permutation and negation of the factors.*

6.2.2 Factors in the growth of the plant roots

There exists many factors wich affects in the growth of the plant roots. The root system model is complex due to that the functionnning is linked to dynamics of the architecture. Water and the nutriment uptake depend on the root surface. The root system model is studied by a stochastic model composed of 14 parameters, among which only a subset of 4 must be estimated on images and a total of 15 statistics that give information on the size and shape of the root system, density of black and white pixels in differents area of the image,... The output of this stochastic model is the image of root system model, as shown in the following image. This image has been found in the website http://www.biologie.uni-hamburg.de/b-online/virtualplants/ipimovies.html

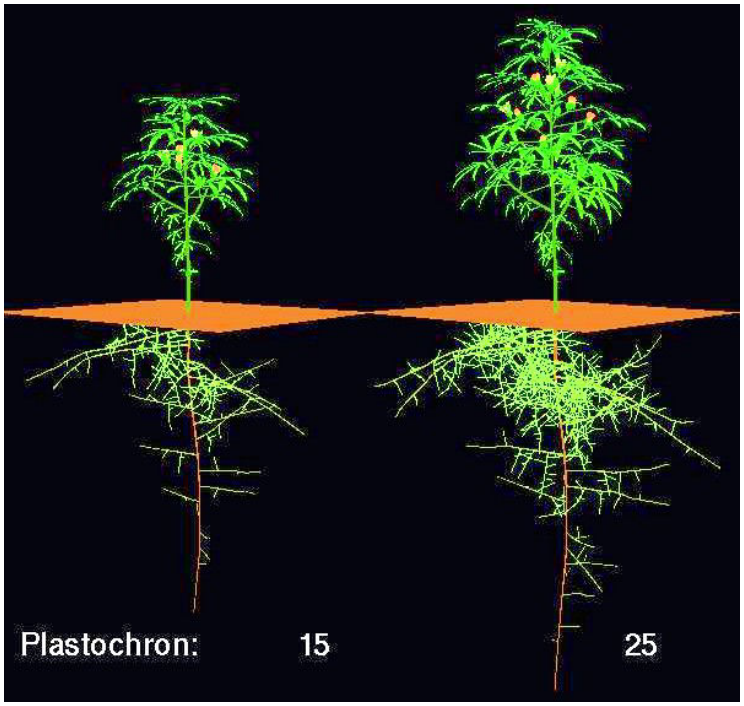


Figure 6.1: virtual images

One of the main objectives of this stochastic method consists in studying which statitics are important and wich are not representative. For this purpose we need to minimize a quadratic problem in 15 variables as we will see in more detail. We can identify each variable with each statitic (even if it is not exactly the same because in our minimization problem the 15 variables correspond the weights of statistics). The variables which are zero correspond to statitics which are not important. This work is the beginning of a collaboration with the team of Claude Bruchou of INRA (Avignon). For more information about this problem see [Cornell 2002, Beaumont 2002, Joyce 2008, Pagès 2011].

In order to determinate the parameters and the statistics and which of them are more important than the other we use the Approximation Bayesian Computing (ABC) method.

Approximation Bayesian Computing is a free likehood method to estimate model parameters. The elements of this method are the followings:

- D = observed data
- D^* = simulated data

- θ vector of parameters with Prior $\pi(\cdot)$
- $S(\cdot)$ function that computes a set of statistics(descriptor)
- $S = S(D)$ =vector of statistics for data D.
- $S^* = S(D^*)$ =vector de statistics for data D^*

In order to estimate if the parameter must be accepted or rejected, we explain in what consists the ABC algorithm:

Algorithm 6.2.1: ALGORITHM A

1. Supose that we have observed data D and $S = S(D)$.
2. Generate θ^* from $\pi(\theta)$.
3. Generate D^* from $f(\cdot \mid \theta^*)$ from $f(D \mid \theta^*)$ where $f(\cdot \mid \theta^*)$ is the probability.
4. Accept θ^* if $D = D^*$ and return to (2).

The problem of this algorithm is that the condition $D = D^*$ is not realistic. Two aproximations are proposed to overcome this issue:

- **First approximation:** Replace the condition $D = D^*$ by a more flexible condition: $d(D, D^*) < \varepsilon$ where $d(\cdot, \cdot)$ defines a distance between the two datasets and uses an threshold ε to accept the simulated θ^*.
- **Second approximation:** If it is impossible to compute $d(D, D^*)$, the statistic $S(D) = (S^1(D), \ldots, S^p(D))$ is defined and the condition $d(D, D^*) < \varepsilon$ is replaced by $d_W(S(D), S(D^*)) < \varepsilon$ where $d_W(\cdot, \cdot) < \varepsilon$ is a weigthed distance between two sets of statistics.

Using the second approximation we have the following algorithm:

Algorithm 6.2.2: ALGORITHM B

1. Suposse we have observed data D and $S = S(D)$.
2. Generate θ^* from $\pi(\theta)$.
3. Generate D^* from $f(\cdot \mid \theta^*)$ from $f(D \mid \theta^*)$ where $f(\cdot \mid \theta^*)$ is the probability.
4. Compute statistics S^* from D^*.
5. Accept θ^* if $d_W(S(D), S(D^*)) < \varepsilon$ and return to (2).

The objectives of this last algorithm are the followings

1. Study the effect of weights W
2. Optimize the choice of statistics weights W.
3. Reduce the number of statistics to improve estimate.

To reach these objectives we study the sensitivity analysis of Mean Squared erreur (MSE). For that we must:

- Find the best weights W of d_W to minimize MSE criterion.
- Point estimate will be: $\hat{\theta} = Mean\{\theta^* : d_W(S, S^*) < \varepsilon\}$ with

$$d_W^2(S(D), S(D^*)) = \sum_{i=1}^{N_S=15} w_i(S_i - S_i^*)^2 \; and \; \sum_{i=1}^{N_S=15} w_i = 1 \; with \;\; w_i \geq 0$$

- Criterion to evaluate point estimate $\hat{(\theta)}$ is:

$$MSE_\theta(W) = \sum_{k=1}^{N_\theta=4} \frac{(\hat{\theta}^{(k)} - \theta^{(k)})^2}{\sigma^2_{\theta^{(k)}}}$$

We use our Algorithm 5.4.1 in the computation of these best weights. The next example have been provided by the team of Claude Bruchou of INRA Avignon:

Example 6.2.18 *Problem*

$$\min\ 918.93287w_1^2 + 1151.12909w_2^2 + 908.27977w_3^2 +$$
$$712.77461w_4^2 + 774.14579w_5^2 + 858.26345w_6^2 + 975.01862w_7^2 +$$
$$+997.92049w_8^2 + 977.09608w_9^2 + 1071.99133w_{10}^2 + 1068.41654w_{11}^2 +$$
$$+951.22177w_{12}^2 + 809.24105w_{13}^2 + 1094.24424w_{14}^2 + 942.00255w_{15}^2 +$$
$$+880.48767w_1w_2 + 1658.93140w_1w_3 + 849.10540w_2w_3 + 390.07161w_1w_4 +$$
$$+470.11656w_2w_4 + 338.43198w_3w_4 + 165.15394w_1w_5 + 604.05698w_2w_5 +$$
$$+72.09737w_3w_5 + 658.16083w_4w_5 + 91.55077w_1w_6 + 975.90816w_2w_6 +$$
$$+91.99926w_3w_6 + 659.09087w_4w_6 + 558.63077w_5w_6 + 456.73235w_1w_7 +$$
$$+1850.68912w_2w_7 + 587.35178w_3w_7 + 282.26909w_4w_7 + 410.73584w_5w_7 +$$
$$+569.83072w_6w_7 + 883.95253w_1w_8 + 2052.51531w_2w_8 + 871.83967w_3w_8 +$$
$$+300.15157w_4w_8 + 147.20287w_5w_8 + 218.10548w_6w_8 + 1509.28027w_7w_8 +$$
$$+1141.08218w_1w_9 + 1707.49981w_2w_9 + 1076.50782w_3w_9 +$$
$$+443.54919w_4w_9 + 438.58698w_5w_9 + 291.41095w_6w_9 + 759.49057w_7w_9 +$$
$$+1161.57519w_8w_9 - 456.10237w_1w_{10} + 1916.74118w_2w_{10} + 86.98672w_3w_{10}$$
$$-74.23206w_4w_{10} + 53.05954w_5w_{10} + 856.63882w_6w_{10} + 1522.76133w_7w_{10} +$$
$$+1648.47669w_8w_{10} - 78.82731w_9w_{10} - 424.93661w_1w_{11} +$$
$$+1989.58896w_2w_{11} + 21.24866w_3w_{11} - 181.48907w_4w_{11} - 96.23656w_5w_{11} +$$
$$+797.67696w_6w_{11} + 1562.63957w_7w_{11} + 1761.24696w_8w_{11} +$$
$$+102.51727w_9w_{11} + 2446.92197w_{10}w_{11} - 79.83487w_1w_{12} +$$
$$+1621.99042w_2w_{12} + 249.81671w_3w_{12} + 80.79488w_4w_{12} - 125.70720w_5w_{12}$$
$$+379.94343w_6w_{12} + 1433.79647w_7w_{12} + 1387.49924w_8w_{12} +$$
$$+431.82652w_9w_{12} + 2043.91638w_{10}w_{12} + 2675.95910w_{11}w_{12} +$$
$$+1276.69957w_1w_{13} + 803.48094w_2w_{13} + 1541.13514w_3w_{13} +$$
$$+485.23186w_4w_{13} + 211.50140w_5w_{13} + 241.29953w_6w_{13} +$$
$$+712.80388w_7w_{13} + 872.96338w_8w_{13} + 965.48511w_9w_{13} +$$
$$413.93988w_{10}w_{13} + 287.60965w_{11}w_{13} + 634.30374w_{12}w_{13} +$$
$$612.56728w_1w_{14} + 1759.42674w_2w_{14} + 423.82649w_3w_{14} + 352.38115w_4w_{14}$$
$$-90.36056w_5w_{14} + 56.76638w_6w_{14} + 929.01717w_7w_{14} + 1224.39898w_8w_{14} +$$
$$818.41864w_9w_{14} + 1095.74584w_{10}w_{14} + 1136.56386w_20w_{14} +$$
$$1397.56386w_{12}w_{14} + 794.17985w_{13}w_{14} + 145.77029w_1w_{15} +$$
$$1788.55675w_2w_{15} + 23.66891w_3w_{15} + 202.22393w_4w_{15} + 867.25256w_5w_{15} +$$
$$+789.49510w_6w_{15} + 926.91090w_7w_{15} + 817.26054w_8w_{15} +$$
$$1074.44452w_9w_{15} + 752.80274w_{10}w_{15} + 688.90470w_20w_{15} +$$
$$776.27853w_{12}w_{15} + 521.71013w_{13}w_{15} - 147.93694w_{14}w_{15};$$
$$s.t. \sum_{i=1}^{15} w_i = 1 \text{ and } w_i \geq 0$$

The solution is:
$w_1 = 0.27119, w_2 = 0, w_3 = 0, w_4 = 0.20533, w_5 = 0.21343,$

$w_6 = 0, w_7 = 0, w_8 = 0, w_9 = 0, w_{10} = 0, w_{11} = 0.30636, w_{12} = 0, w_{13} = 0, w_{14} = 0, w_{15} = 0.00367.$

This solution is very interesting because it implies that in this considered position the statistics $S_2, S_3, S_6, S_7, S_8, S_9, S_{10}, S_{12}, S_{13}, S_{14}$ are not significant. Therefore only the statistics $S_1, S_4, S_5, S_{11}, S_{15}$ are meaningful.

6.2.3 Marx generators design

The optimization problem appears also in physical problem, such as the design of the Marx generators. This design has an specification wich consists a resonance condition that must be satisfied by the circuit so that all the intially energy contained in differents capacitors is carried in a finite time to one only capacitor. We will see that the differents components of this circuit can be represented in a structuered real eigenvalue matrix that we can solve by polynomial optimization. This section is based in the paper [Galeani 2014] We consider the Marx generator network described in the following figure consists in n stages (and $n + 1$ loops) where, disregarding the rigthmost components of the picture, each one of n stages consists in an upper branch with a capacitor and a inductor and a vertical branch with a capacitor only.

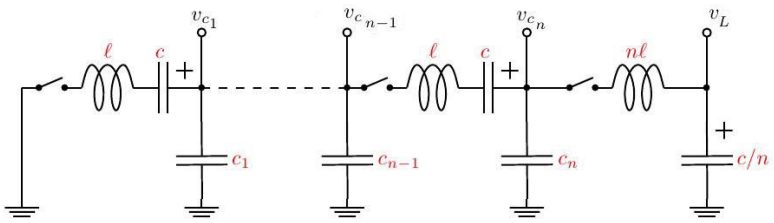


Figure 6.2: network model of a N-stages Marx generator network driving a load capacitor

Following [Antoun 2006, Buchenauer 2010, Zaccarian 2009], we assume that all the capacitors and inductors appearing in the upper branch are the same (corresponding to some fixed positive reals c and l). We call these capacitors "storage capacitors". The design problem addressed here is the selection of vertical capacitors, which are exactly n, where n is the

number of stages of the Marx circuit. These capacitors are called "parasitic capacitors". Following [Antoun 2006, Buchenauer 2010, Zaccarian 2009], the inductor and capacitor apparing in the rigthmost loop take the values nl and c/n, respectively. This capacitor is called "load capacitor". This selection preserves the resonance property (so that the product of any adjacent capacitor/inductor pairs is always lc) in addition to ensuring that the load capacitor is n times larger than each one of the storages capacitors. The problem that we will solve is the following:

Problem 1: *Consider the circuit in Figure 6.2 for a given n and certain values of c and l. Select positive values* $c_i > 0$, $i = 1, ..., n$ *of the parasitic capacitors and a time* $T > 0$ *such that, initializing at t=0 all the storage capacitors with the same voltage* $v(0) = v_0$ *and starting from zero current in all the inductors and zero voltage across the parasitic capacitors and the load capacitor, the circuit response is such that at* $t = T$ *all voltages and currents are zero except for the voltage across the load capacitor that will be* $v_L(T) = nv_0$ *since the circuit is lossless.*

Solution: A solution to this problem can be determined from the solution of a *suitable structured eigenvalue assignment problem* (see [Galeani 2014]). We recall the main theorem of this work:

Theorem 6.2.19 *Consider any set of n distinct positive even integer* $\alpha = (\alpha_1, \cdots, \alpha_n)$, *a matrix* $B \in \mathbb{R}^{n \times n}$ *defined as*

$$B := \begin{bmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 2 & -1 \\ 0 & \cdots & 0 & -1 & \frac{n}{n+1} \end{bmatrix}$$

and any positive definite real diagonal solution $P = diag(p_1, \cdots, p_n)$ *to the structured eigenvalue assignement problem*

$$\sigma(BP) = \{\alpha_1^2 - 1, \cdots, \alpha_n^2 - 1\} \tag{6.19}$$

where $\sigma(BP)$ *denotes the spectrum of BP, i.e, the set of its complex eigenvalues. Then for any value of c, the selection* $c_i = c/p_i$, $i = 1, ..., n$ *solves the Problem 1 for all values of l with* $T = \frac{\pi}{\sqrt{lc}}$.

There exists two methods to solve this problem, using symbolic techniques such that Gröbner basis and using convex polynomial optimization techniques which is of our interest.

In order to find the solution of this problem we have two differents formulations:

- First formulation: It consists in computing the diagonal entries $p = [p_1, \cdots, p_n]$ solution of 6.19. This is equivalent to solve a finite set of n equations with unknown p, each of them correponding to one coefficient of the following polynomial identity in the variable s:

$$det(sI - BP) = \prod_{i=1}^{n}(s - (\alpha_i^2 - 1)),\ \forall \mathbf{s} \in \mathbb{C}. \tag{6.20}$$

For a fixed value n and fixed values in α (we can select $\alpha_i = 2i$, so that the circuit resonates at the lowest possible frequency), we can write a system of n polynomial equations in the variable p with rational coefficients, namely

$$h_i(p) = 0,\ i = 1, ..., n \tag{6.21}$$

Example 6.2.20 *For the case $n = 4$,*

$$B = \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & \frac{5}{4} \end{bmatrix} \text{ and } P = \begin{bmatrix} p_1 & 0 & 0 & 0 \\ 0 & p_2 & 0 & 0 \\ 0 & 0 & p_3 & 0 \\ 0 & 0 & 0 & p_4 \end{bmatrix}$$

So,

$$BP = \begin{bmatrix} 2p_1 & -p_2 & 0 & 0 \\ -p_1 & 2p_2 & -p_3 & 0 \\ 0 & -p_2 & 2p_3 & -p_4 \\ 0 & 0 & -p_3 & \frac{5}{4}p_4 \end{bmatrix}$$

Then,

$$det(sI - BP) = \begin{vmatrix} 2p_1 & -p_2 & 0 & 0 \\ -p_1 & 2p_2 & -p_3 & 0 \\ 0 & -p_2 & 2p_3 & -p_4 \\ 0 & 0 & -p_3 & \frac{5}{4}p_4 \end{vmatrix} =$$

$$= s^4 + s^3(-2p_4 - \frac{15}{8}p_3 - \frac{3}{2}p_2 - \frac{7}{8}p_1) +$$

$$+ s^2(\frac{3}{2}p_4p_3 + 2p_2p_4 + \frac{5}{4}p_2p_3 + \frac{3}{2}p_1p_4 + \frac{5}{4}p_1p_3 + \frac{3}{4}p_1p_2) +$$

$$+ s(-p_2p_3p_4 - p_1p_3p_4 - p_1p_2p_4 - \frac{5}{8}p_1p_2p_3) + \frac{1}{2}p_1p_2p_3p_4$$

and if we take $\alpha = (2, 4, 6, 8)$

$$\prod_{i=1}^{4}(s - (\alpha_i^2 - 1)) = (s-3)(s-15)(s-35)(s-63) =$$

$$= s^4 - 116s^3 + 4014s^2 - 44100s + 99225$$

So, the system of polynomial equations that verifies 6.20 is the following

$$\begin{aligned} 2p_4 + \frac{15}{8}p_3 + \frac{3}{2}p_2 + \frac{7}{8}p_1 &= 116 \\ \frac{3}{2}p_4p_3 + 2p_2p_4 + \frac{5}{4}p_2p_3 + \frac{3}{2}p_1p_4 + \frac{5}{4}p_1p_3 + \frac{3}{4}p_1p_2 &= 4014 \\ p_2p_3p_4 + p_1p_3p_4 + p_1p_2p_4 + \frac{5}{8}p_1p_2p_3 &= 44100 \\ \frac{1}{2}p_1p_2p_3p_4 &= 99225 \end{aligned}$$

- Second formulation: It corresponds to inverting the eigenvalue assigment problem 6.19, thereby obtaining an alternative set of polynomial equations in the unknown $k = [k_1, ..., k_n]^T = [p_1^{-1}, \cdots, p_n^{-1}]$ with rational coefficient, which have the advantage of being linear in the capacitor values, indeed, $k_i = c_i/c$, $i = 1, ..., n$. In particular, for the inverse problem, equation 6.20 becomes

$$det(sI - KB^{-1}) = \prod_{i=1}^{n}(s - (\alpha_i^2 - 1)^{-1}), \ \forall \ s \in \mathbb{C}. \tag{6.22}$$

where $K = diag(k) = P^{-1}$. For a fixed value of n and fixed values in α, one can write a system of n polynomial equations in the variable k with rational coefficients, namely

$$q_i(k) = 0, \ i = 1, ..., n. \tag{6.23}$$

Remark 6.2.21 *All the entries of each solution to this polynomial system are in the interval* $(0, 1)$.

Example 6.2.22 *For the case* $n = 4$,

$$B^{-1} = \begin{bmatrix} \frac{7}{8} & \frac{3}{4} & \frac{5}{8} & \frac{1}{2} \\ \frac{3}{4} & \frac{3}{2} & \frac{5}{4} & 1 \\ \frac{5}{8} & \frac{5}{4} & \frac{15}{8} & \frac{3}{2} \\ \frac{1}{2} & 1 & \frac{3}{2} & 2 \end{bmatrix} \text{ and } K = \begin{bmatrix} k_1 & 0 & 0 & 0 \\ 0 & k_2 & 0 & 0 \\ 0 & 0 & k_3 & 0 \\ 0 & 0 & 0 & k_4 \end{bmatrix}$$

So,

$$KB^{-1} = \begin{bmatrix} \frac{7}{8}k_1 & \frac{3}{4}k_2 & \frac{5}{8}k_3 & \frac{1}{2}k_4 \\ \frac{3}{4}k_1 & \frac{3}{2}k_2 & \frac{5}{4}k_3 & k_4 \\ \frac{5}{8}k_1 & \frac{5}{4}k_2 & \frac{15}{8}k_3 & \frac{3}{2}k_4 \\ \frac{1}{2}k_1 & k_2 & \frac{3}{2}k_3 & 2k_4 \end{bmatrix}$$

Then,

$$det(sI - KB^{-1}) = \begin{vmatrix} s - \frac{7}{8}k_1 & \frac{3}{4}k_2 & \frac{5}{8}k_3 & \frac{1}{2}k_4 \\ \frac{3}{4}k_1 & s - \frac{3}{2}k_2 & \frac{5}{4}k_3 & k_4 \\ \frac{5}{8}k_1 & \frac{5}{4}k_2 & s - \frac{15}{8}k_3 & \frac{3}{2}k_4 \\ \frac{1}{2}k_1 & k_2 & \frac{3}{2}k_3 & s - 2k_4 \end{vmatrix} =$$

$$= s^4 + s^3(-\frac{7}{8}k_1 - \frac{3}{2}k_2 - 2k_4 - \frac{15}{8}k_3) +$$
$$+ s^2(\frac{3}{4}k_1k_2 + \frac{5}{4}k_1k_3 + \frac{3}{2}k_3k_4 + \frac{3}{2}k_1k_4 + 2k_2k_4 + \frac{5}{4}k_2k_3) +$$
$$+ s(-\frac{5}{8}k_1k_2k_3 - k_2k_3k_4 - k_1k_3k_4 - k_1k_2k_4) + \frac{1}{2}k_1k_2k_3k_4$$

and if we take $\alpha = (2, 4, 6, 8)$

$$\prod_{i=1}^{4}(s - (\alpha_i^2 - 1)^{-1}) = (s - \frac{1}{3})(s - \frac{1}{15})(s - \frac{1}{35})(s - \frac{1}{63}) =$$

$$= s^4 - \frac{4}{9}s^3 + \frac{446}{11025}s^2 - \frac{116}{99225}s + \frac{1}{99225}$$

So, the system of polynomial equations that verifies 6.22 is the following

$$\begin{aligned} \frac{7}{8}k_1 + \frac{3}{2}k_2 + 2k_4 + \frac{15}{8}k_3 &= \frac{4}{9} \\ \frac{3}{4}k_1k_2 + \frac{5}{4}k_1k_3 + \frac{3}{2}k_3k_4 + \frac{3}{2}k_1k_4 + 2k_2k_4 + \frac{5}{4}k_2k_3 &= \frac{446}{11025} \\ \frac{5}{8}k_1k_2k_3 + k_2k_3k_4 + k_1k_3k_4 + k_1k_2k_4 &= \frac{116}{99225} \\ \frac{1}{2}k_1k_2k_3k_4 &= \frac{1}{99225} \end{aligned}$$

As we said before we are interested in using convex polynomial optimization thecniques in order to solve this eigenvalue assigment problem. Indeed, we center our attention in the inverse eigenvalue problem because all the real solutions of 6.22 sastify $\mid k_i \mid \leq 1$. A numerical approach solution to this problem is to formulate it as a nonconvex polynomial optimization:

$$\begin{aligned} q^* &= \min_k q_0(k) \\ &s.t\ k \in \mathcal{K} \end{aligned} \tag{6.24}$$

where $q_0 \in \mathbb{Q}[k]$ is a polynomial in $k \in \mathbb{R}^n$ and the feasible set $\mathcal{K}$ is defined as follow:

$$\mathcal{K} = \{k \in \mathbb{R}^n \mid q_i(k) = 0, i = 1, ..., n,\ h_j(k) = k_j - 2k_{j+1} + k_{j+2} \geq 0, j = 1, ..., n-2\}$$

where the inequalities h_j define the fact that we look for regular solutions.

We can choose different objective function $q_0(k)$. But if we want to do the capacitors $c_i = k_i/c$ as identical as possible we choose:

$$q_0(k) = \sum_{i,j=1}^{n} (k_i - k_j)^2 = k^T \begin{bmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 2 & -1 \\ 0 & \cdots & 0 & -1 & 2 \end{bmatrix} k$$

We can solve this non convex polynomial optimization problem thanks to our minimizer border basis algorithm. For each n the regular solution which was given in [Galeani 2014, Buchenauer 2010] is:

n	$n^2\frac{c_1}{c}$	$n^2\frac{c_2}{c}$	$n^2\frac{c_3}{c}$	$n^2\frac{c_4}{c}$	$n^2\frac{c_5}{c}$	$n^2\frac{c_6}{c}$	$n^2\frac{c_7}{c}$	$n^2\frac{c_8}{c}$	cond
1	1.5								1
2	0.63120	1.12660							1.0266
3	0.84408	0.77662	1.41217						1.0387
4	1.13210	0.78731	0.92450	1.60306					1.0440
5	1.47480	0.86342	0.84481	1.07344	1.72179				1.0448
6	1.87892	0.96056	0.85587	0.91518	1.23619	1.77345			1.0592
7	2.07061	1.05669	1.04940	1.05715	1.06861	1.08449	1.85298		1.0502
8	2.39407	1.17326	1.12475	1.11221	1.10440	1.0996	1.0998	1.87252	1.0617

Table 6.5: Regular solutions to the Marx design problem

Indeed, by Theorem 6.2.19, $n^2\frac{c_i}{c} = n^2 k_i$ then the regular solution in k are:

n	k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8	cond
1	1.5								1
2	0.1578	0.28165							1.0266
3	0.09378	0.08629	0.1569						1.0387
4	0.07075	0.0492	0.05778	0.10019					1.0440
5	0.05899	0.03453	0.03379	0.04293	0.06887				1.0448
6	0.0522	0.026682	0.02377	0.02542	0.03433	0.04926			1.0592
7	0.04225	0.02156	0.02141	0.02157	0.02180	0.02213	0.0378		1.0502
8	0.0374	0.01833	0.01757	0.01737	0.01725	0.01718	0.01785	0.02925	1.0617

Table 6.6: Regular solutions to the Marx design problem (valeus k_i)

Example 6.2.23 *In the case $n = 4$ we formulate the following problem*

$$\begin{aligned}
\min \quad & 6k_1^2 - 4k_1k_2 - 4k_1k_3 - 4k_1k_4 + 6k_2^2 - 4k_2k_3 - 4k_2k_4 + 6k_3^2 - 4k_3k_4 + 6k_4^2; \\
s.t \quad & \tfrac{7}{8}k_1 + \tfrac{3}{2}k_2 + 2k_4 + \tfrac{15}{8}k_3 - \tfrac{4}{9} = 0; \\
& \tfrac{3}{4}k_1k_2 + \tfrac{5}{4}k_1k_3 + \tfrac{3}{2}k_3k_4 + \tfrac{3}{2}k_1k_4 + 2k_2k_4 + \tfrac{5}{4}k_2k_3 - \tfrac{446}{11025} = 0; \\
& \tfrac{5}{8}k_1k_2k_3 + k_2k_3k_4 + k_1k_3k_4 + k_1k_2k_4 - \tfrac{116}{99225} = 0; \\
& \tfrac{1}{2}k_1k_2k_3k_4 - \tfrac{1}{99225} = 0; \\
& k_1 - 2k_2 + k_3 \geq 0; \\
& k_2 - 2k_3 + k_4 \geq 0;
\end{aligned}$$

Applying our minimization border basis algorithm 5.4.1 we obtain the following minimizer:

$$sol := (0.0707, 0.0492, 0.0577, 0.1001)$$

which coincides with the regular solution for $n = 4$ of table 6.6 with precision $10e - 4$ therefore we can deduce that the solution is very accurate.

Our experimentation have been done until $n = 8$, the solution obtained is exactly the same that the table 6.6 which improves the solutions not very accurate for $n = 2, 3, 4, 5$ given by Gloptipoly in [Galeani 2014].

6.3 Implementation

The C++ implementation of the previous algorithm has been performed in the BORDERBASIX package of the MATHEMAGIX[1] software, which provides a

[1] www.mathemagix.org

C++ implementation of the border basis algorithm of [Mourrain 2012]. This algorithm was implemented by Philippe Trebuchet and Bernard Mourrain.

For the computation of border basis, we use a choice function which is tolerant to numerical instability i.e. a choice function that chooses as leading monomial a monomial whose coefficient is maximal among the choosable monomials as described in [Mourrain 2008].

The Semi-Definite Programming problems are solved using four different solvers

- SDPA[2] software which work with double precision. It is implemented in C++ language and utilizes the BLAS and LAPACK libraries for matrix computations. It is designed to solve small and medium size SDP problems: usually the number of variables $m < 2000$ and matrix sizes $n < 2000$ but also depend on the available hardware.

- CSDP[3] software which work with double precision. It is implemented in C language and utilizes the BLAS and LAPACK libraries for matrix computations. It is designed to solve small and medium size SDP problems of the same size of SDPA.

- SDPA-GMP[4] software which works with multiple precision to obtain highly accurate solutions. It is implemented in C++ language and utilizes GMP library. It is designed to solve small and medium size SDP problems of the same size as SDPA. On the other side SDPA-GMP is in general ten or hundred times slower than the SDPA.

- MOSEK[5] software. It is implemented in C/C++ language. It is designed to solve small, medium and big size SDP problems. It is in general between two and four times faster than the others three.

For the link with SDPA, SDPA-GMP, CSDP we use a file interface since, input data is the same in the three and the output data change in CSDP. In the case of MOSEK, we use the distributed binary library.

Once we have computed the moment matrix, we verify the flat extension using the Decomposer Algorithm with some differences. The Descomposer Algorithm is in the BORDERBASIX package and it was implemented in C++ by Matthieu Dien. This algorithm decomposes a symmetric tensor in sum of tensor of rank 1 or equivalently, it decomposes a multivariate polynomial in a

[2] http://sdpa.sourceforge.net
[3] https://projects.coin-or.org/Csdp/
[4] http://sdpa.sourceforge.net
[5] http://www.mosek.com

sum of powers of linear forms. We introduce some changes in order to use it for verify the flat extension.

The minimizer points are computed from the eigenvalues of the multiplication matrices. This is performed using Lapack routines inside the Decomposer Algorithm that we can get when we have verify the flat extension.

In order to compare inside the same environment our algorithm with the full moment matrix relaxation algorithm described in [Lasserre 2009a], which is implemented in the package Gloptipoly of Matlab developed by D. Henrion and J.B. Lasserre, we have also implemented in C++ this latter algorithm in the BORDERBASIX package.

6.3.1 Input arguments, Input data file and Output data file

In this subsection we describe the different arguments that we need to give in the input when we want to call to the algorithm 5.4.1. We explain them in the same order than they need to be introduced.

1. *Input data file name.* This file contains all the information of the optimization problem: the number of equalities and inequalities, the polynomial to minimize and the differents constraints. We will describe it in more detail below.

2. *SDP Solver* : -s n where n=(1-SDPA, 2-SDPA-GMP, 3-CSDP, 4-MOSEK). If we do not fill this option the default solver is 1-SDPA.

3. *Method to validate the flat extension*: -m n where n=(0-decomposer algorithm, 1-rank of submatrices). If we do not fill this option the default solver is 0-decomposer algorithm.

4. *Parameter file name for the SDP Solver*: -p f where f is the Parameter file name. Apply only in the cases -s 1, 2 or 3.

5. *Threshold in the Decomposer algorithm*: -t threshold. Apply only in the case -m 0. If we do not fill this option the default threshold is $10e-3$.

6. *Complete border basis or not*: -b n=(0-if we use complet border basis, 1-border basis of degree minimum). If we do not fill this option we computed by default the complet border basis.

The input data file of our algorithm has the following format:

- $n1$; number of equalities

- $n2$; number of inequalities
- f; polynomial function to minimize
- g_i; equality i
- h_j; inequality $(\geq)$ j

Example 6.3.1 *We want to solve this problem:*

$$\begin{array}{ll} \min & -10 + 2x + 6y - 2x^2 + 2xy - 2y^2 \\ s.t. & -x^2 + 2x \geq 0 \\ & -8 - y^2 + 6y \geq 0 \\ & 1 - x^2 + 2xy - y^2 \geq 0 \end{array}$$

The input data file will be as follow:

0;
3;
$-10 + 2 * x0 + 6 * x1 - 2 * x0^2 + 2 * x0 * x1 - 2 * x1^2$;
$-x0^2 + 2 * x0$;
$-8 - x1^2 + 6 * x1$;
$1 - x0^2 + 2 * x0 * x1 - x1^2$;

Table 6.7: **Input data file:"example10"**

The output data file of our algorithm will be called the name of input data file + Solution. For example if the input in the above example is called "example10" the output data file asociated will be called "example10Solution". It has the following format:

- sol:=$[s_1, s_2, ..., s_n]$ where s_i are the different solutions.
- fmin:=f^* where f^* is the minimum of the optimization problem.

Example 6.3.2 *The output data file asociated to the above example will be as follow:*

sol:=[[1, 2], [2, 2], [2, 3]];
fmin:=2

Table 6.8: **Output data file:"example10Solution"**

Conclusion

Pour conclure, notre travail met en avant deux points importants:

- Permettre de traiter les problèmes où la variété KKT est vide bien que le minimum soit atteint dans un point de S. Pour cela, nous utilisons la variété de Fritz-John que nous définissons comme la réunion de la variété KKT et de la variété des points singuliers. Notre problème d'optimisation sur S_{sing} est du même type que sur S_{KKT}, nous avons étudié s'il y a ou pas des points qui minimisent notre fonction f sur S_{KKT}. Nous appliquons récursivement la même méthode sur S_{sing}. Néanmoins, il reste du travail à faire sur ce point, nous avons besoin d'une preuve rigoureuse pour montrer que cette séquence de sous-problèmes est finie. Il faudra regarder si la dimension de la région faisable et/ou la multiplicité des solutions de la région faisable se réduit de plus en plus.

- Nous avons développé et codé en c++ un nouvel algorithme pour résoudre des problèmes d'optimisation de type (2.1) quand le nombre des points qui minimisent notre fonction est fini. Cet algorithme utilise l'algorithme de bases de bord, développé par Mourrain et Trebuchet, qui permet de réduire la taille et le nombre de paramétrés de notre problème SDP, et donc de réduire la complexité de calcul de la SDP. Notre implantation donne la possibilité de résoudre la SDP avec quatre logiciels: SDPA, CSDP, SDPA-GMP et MOSEK. Ce dernier permet une réduction du temps d'exécution comprise entre 50% et 80% par rapport aux trois autres logiciels. Pour vérifier si le minimum est atteint (par une généralisation du critère d'extension plate de Curto-Fialkow), nous avons développé un nouvel algorithme qui utilise des polynômes orthogonaux et que nous peut aussi servir pour la décomposition de tenseurs. Nous avons aussi codé la méthode de Lasserre (Gloptipoly) en c++, afin de bien comparer les temps de calcul. Notre algorithme est librement disponible et sera mis en place sur un site web dans les prochaines mois avec toute l'information nécessaire pour son utilisation. Il reste à réaliser un travail d'optimisation de code ainsi qu'à concevoir un critère d'arrêt de l'algorithme quand le nombre de points qui minimisent notre fonction n'est pas fini.

Bibliography

[Abril Bucero 2013] M. Abril Bucero and B. Mourrain. Exact relaxation for polynomial optimization on semi-algebraic sets. http://arxiv.org/abs/1307.6426, 2013. (Cited on page 2.)

[Abril Bucero 2014] M. Abril Bucero and B. Mourrain. Border basis relaxation for polynomial optimization. http://arxiv.org/abs/1404.5489, 2014. (Cited on page 2.)

[Antoun 2006] N. N. Antoun. *State Space Analysis and Optimization of Marx Generator*. PhD thesis, University of New Mexico, Albuquerque, 2006. (Cited on pages 111 and 112.)

[Bazaraa 2006] Mokhtar S. Bazaraa, Hanif D. Sherali and Shetty C. M. Nonlinear programming:theory and algorithms. J. Wiley and sons, 2006. (Cited on page 24.)

[Beaumont 2002] M. Beaumont, W. Zhang and D.J Balding. *Approximate bayesian computation in population genetics*. Genetics, vol. 162, no. 4, pages 2025–2035, 2002. (Cited on page 107.)

[Bernardi 2013] A. Bernardi, J. Brachat, P. Comon and B. Mourrain. *General tensor decomposition, moment matrices and applications*. J. of Symbolic Computation, vol. 52, pages 51–71, 2013. (Cited on page 98.)

[Bochnak 1998] J. Bochnak, M. Coste and M.-F. Roy. Real algebraic geometry. Springer, 1998. (Cited on page 7.)

[Brachat 2010] J. Brachat, P. Comon, B. Mourrain and E. Tsigaridas. *Symmetric Tensor Decomposition*. Linear Algebra and Applications, vol. 433, pages 1851–1872, 2010. (Cited on pages 66 and 98.)

[Bruns 1988] W. Bruns and U. Vetter. Determinantal rings, volume 1327 of *Lecture Notes in Math.* Springer, Berlin, 1988. (Cited on page 24.)

[Buchenauer 2010] C. J. Buchenauer. *Optimizing compact Marx generators networks*. IEEE Transactions on Plasma Science, vol. 38, no. 10, pages 2771–2784, 2010. (Cited on pages 111, 112 and 116.)

[Comon 2000] P. Comon. *Tensor decomposition - state of art and applications*. Keynote ad. in IMA Conf. in Signal processing, 2000. (Cited on page 98.)

[Cornell 2002] J. Cornell. Experiments and mixtures. J. Wiley and sons, 2002. (Cited on page 107.)

[Cox 2005] D.A. Cox, J.B. Little and D.B. O'Shea. Ideals, varieties, and algorithms : An introduction to computational algebraic geometry and commutative algebra (undergraduate texts in mathematics). Springer, 2005. (Cited on pages 7, 22 and 59.)

[Curto 1996] R.E. Curto and L. Fialkow. *Solution of the truncated complex moment problem for flat data.* Memoirs of the American Mathematical Society, vol. 119, no. 568, pages 1–62, 1996. (Cited on pages 46, 74, 82 and 136.)

[De Lathauwer 2000] L. De Lathauwer, B. De Moor and J. Vanderwalle. *On the best rank-1 and rank-(R1,R2,...,RN) approximations of highwe-order tensors.* SIAM Journal on Matrix Analysis and Applications, vol. 21, no. 4, pages 1324–1342, 2000. (Cited on page 101.)

[Demmel 2007] J. Demmel, J. Nie and V. Powers. *Representations of positive polynomials on noncompact semialgebraic sets via KKT ideals.* Journal of Pure and Applied Algebra, vol. 209, no. 1, pages 189 – 200, 2007. (Cited on pages 21, 45, 47, 50, 51 and 135.)

[Elkadi 2007] M. Elkadi and B. Mourrain. Introduction à la résolution des systèmes d'équations algébriques, volume 59 of *Mathématiques et Applications.* Springer-Verlag, 2007. (Cited on pages 12 and 80.)

[Floudas 1999] C. A. Floudas, P. M. Pardalos, C. S. Adjiman, W. R. Esposito, Z. H. Gumus, S. T. Harding, J. L. Klepeis, C. A. Meyer and C. A. Schweiger. Handbook of test problems in local and global optimization. Kluwer Academic Publishers., 1999. (Cited on pages 95, 96 and 97.)

[Galeani 2014] S. Galeani, D. Henrion, A. Jacquemard and L. Zaccarian. Design of marx generators as a structured eigenvalue assignment. 2014. (Cited on pages 111, 112, 116 and 117.)

[Giesbrecht 2009] M. Giesbrecht, G. Labahn and W.-S. Lee. *Symbolic-numeric Sparse Interpolation of Multivariate Polynomials.* J. Symb. Comput., vol. 44, no. 8, pages 943–959, August 2009. (Cited on page 66.)

[Greuet 2011] A. Greuet and M. Safey El Din. *Deciding reachability of the infimum of a multivariate polynomial.* In Proceedings of the 36th international symposium on Symbolic and algebraic computation, ISSAC '11, pages 131–138, New York, NY, USA, 2011. ACM. (Cited on page 46.)

[Guo 2010] F. Guo, M. Safey El Din and L. Zhi. *Global optimization of polynomials using generalized critical values and sums of squares.* In Proceedings of the 2010 International Symposium on Symbolic and Algebraic Computation, ISSAC '10, pages 107–114, New York, NY, USA, 2010. ACM. (Cited on page 46.)

[Ha 2008] H. V. Ha and T.S. Pham. *Global optimization of polynomials using the truncated tangency variety.* SIAM Journal on Optimization, vol. 19, no. 2, pages 941–951, 2008. (Cited on pages 46 and 47.)

[Ha 2010] H. V. Ha and T.S. Pham. *Representation of positive polynomials and optimization on noncompact semialgebraic sets.* SIAM Journal on Optimization, vol. 20, no. 6, pages 3082–3103, 2010. (Cited on pages 21, 25, 45, 60 and 135.)

[Henrion 2002] D. Henrion and J.B. Lasserre. *Gloptipoly:Global Optimization over Polynomials with Matlab and Sedumi.* Rapport technique Report No. 02057, LAAS-CNRS, Toulose, France, 2002. (Cited on page 82.)

[Henrion 2005] D. Henrion and J.B. Lasserre. Positive polynomials in control, chapitre Detecting Global Optimality and Extracting Solutions in GloptiPoly., pages 293–310. Lectures Notes in Control and Information Sciences. Springer, 2005. (Cited on pages 46 and 66.)

[John 1948] F. John. *Extremum problems with inequalities as side conditions.* In Studies and Essays, Courant Anniversary Volume, pages 187–204. Wiley (Interscience), New York, 1948. (Cited on pages 23, 24 and 46.)

[Joyce 2008] P. Joyce and P. Marjoram. *Approximate sufficient statistics and bayesian computation.* Statistical applications in Genetics and molecular biology, vol. 7, no. 1, pages 1544–1615, 2008. (Cited on page 107.)

[Lasserre. 2001] J.B. Lasserre. *Global optimization with polynomials and the problem of moments.* SIAM J. Optim., vol. 11, pages 796–817, 2001. (Cited on pages 31, 32, 41, 42, 43, 45, 65, 82, 95, 96, 97 and 134.)

[Lasserre 2008] J.B. Lasserre, M. Laurent and P. Rostalski. *Semidefinite characterization and computation of real radical ideals.* Foundations of Computational Mathematics, vol. 8, no. 5, pages 607–647, 2008. (Cited on pages 53 and 66.)

[Lasserre 2009a] J.B Lasserre. Moments, positive polynomials and their applications. Imperial College Press, 2009. (Cited on pages 24, 46, 57, 65, 66, 82, 93, 94, 95, 96, 97, 119 and 133.)

[Lasserre 2009b] J.B. Lasserre, M. Laurent and P. Rostalski. *A unified approach for real and complex zeros of zero-dimensional ideals.* In M. Putinar and S. Sullivant, editeurs, Emerging Applications of Algebraic Geometry., volume 149, pages 125–156. Springer, 2009. (Cited on pages 46, 47, 62, 66 and 71.)

[Lasserre 2012] J.-B. Lasserre, M. Laurent, B. Mourrain, P. Rostalski and P. Trébuchet. *Moment Matrices, Border Bases and Real Radical Computation.* Journal of Symbolic Computation, 2012. (Cited on pages 16, 46, 47, 57, 62, 65, 71, 72 and 82.)

[Laurent 2007] M. Laurent. *Semidefinite representations for finite varieties.* Math. Progr, vol. 109, pages 1–26, 2007. (Cited on pages 45, 46, 66 and 135.)

[Laurent 2009a] M. Laurent. Sums of squares, moment matrices and optimization over polynomials, volume 149 of *IMA Volumes in Mathematics and its Applications*, pages 157–270. Springer, 2009. (Cited on pages 39, 41, 45, 46, 61, 82, 88 and 135.)

[Laurent 2009b] M. Laurent and B. Mourrain. *A generalized flat extension theorem for moment matrices.* Arch. Math. (Basel), vol. 93, no. 1, pages 87–98, July 2009. (Cited on pages 46, 66, 74 and 136.)

[Ma 2013] Y. Ma, Ch. Wang and Zhi L. A certificate for semidefinite relaxations in computing positive dimensional real varieties. http://arxiv.org/abs/1212.4924, 2013. (Cited on page 62.)

[Mangasarian 1967] O.L. Mangasarian and S. Fromovitz. *The Fritz John Necessary Optimality Conditions in the Presence of Equality and Inequality Constraints.* Journal of Mathematical Analysis and Applications, vol. 17, pages 37–47, 1967. (Cited on pages 23, 24 and 46.)

[Marshall 2003] M. Marshall. *Optimization of Polynomial Functions.* Canad. Math. Bull., vol. 46, pages 575–587, 2003. (Cited on page 45.)

[Marshall 2009] M. Marshall. *Representations of non-negative polynomials, degree bounds and applications to optimization.* Can. J. Math., vol. 61, no. 1, pages 205–221, 2009. (Cited on pages 45 and 135.)

[Mourrain 1999] B. Mourrain. *A new criterion for normal form algorithms.* In M. Fossorier, H. Imai, Shu Lin and A. Poli, editeurs, Proc. AAECC, volume 1719 of *LNCS*, pages 430–443. Springer, Berlin, 1999. (Cited on page 69.)

[Mourrain 2005] B. Mourrain and P. Trébuchet. *Generalized normal forms and polynomials system solving.* In M. Kauers, editeur, ISSAC: Proceedings of the ACM SIGSAM International Symposium on Symbolic and Algebraic Computation, pages 253–260, 2005. (Cited on pages 65, 66, 68 and 69.)

[Mourrain 2008] B. Mourrain and Ph. Trébuchet. *Stable normal forms for polynomial system solving.* Theoretical Computer Science, vol. 409, no. 2, pages 229–240, 2008. (Cited on pages 66 and 118.)

[Mourrain 2012] B. Mourrain and Ph. Trébuchet. *Border basis representation of a general quotient algebra.* In Joris van der Hoeven, editeur, ISSAC 2012, pages 265–272, July 2012. (Cited on page 118.)

[Nesterov 1994] Y. Nesterov and A. Nemirovski. Interior-point polynomial algorithms in convex programming. SIAM, Philaldelphia, 1994. (Cited on page 65.)

[Nesterov 2000] Y. Nesterov. *Squared functional systems and optimization problems.* In H. Frenk, K. Roos, T. Terlaky and S. Zhang, editeurs, High performance optimization, chapitre 17, pages 405–440. Kluwer academic publishers, Dordrecht, The Netherlands, 2000. (Cited on page 31.)

[Nie 2006] J. Nie, J. Demmel and B. Sturmfels. *Minimizing Polynomials via Sum of Squares over Gradient Ideal.* Math. Program., vol. 106, no. 3, pages 587–606, 2006. (Cited on pages 45, 47, 58, 59 and 135.)

[Nie 2011] J. Nie. *An exact Jacobian SDP relaxation for polynomial optimization.* Mathematical Programming, pages 1–31, 2011. (Cited on pages 21, 24, 45, 47, 60 and 135.)

[Nie 2012] J. Nie. *Certifying convergence of Lasserre's hierarchy via flat truncation.* Mathematical Programming, pages 1–26, 2012. (Cited on pages 66 and 75.)

[Nie 2013a] J. Nie. *Polynomials optimization with real variety.* SIAM Journal On Optimization, vol. 23, no. 3, pages 1634–1646, 2013. (Cited on pages 45, 47, 61 and 135.)

[Nie 2013b] J. Nie and L. Wang. Semidefinite relaxations for best rank-1 tensor approximations. http://arxiv.org/abs/1308.6562v2, 2013. (Cited on pages 98 and 102.)

[Ottaviani 2013] G. Ottaviani, P-J. Spaenlehauer and B. Sturmfels. Exact solutions in structured low-rank approximation. http://http://arxiv.org/abs/1311.2376v2, 2013. (Cited on pages 98, 102 and 105.)

[Pagès 2011] L. Pagès. *Links between root developmental traits and foraging perfomance.* Plant, Cell and Enviroment, vol. 34, no. 10, pages 1749–1760, 2011. (Cited on page 107.)

[Parrilo 2003] P.A. Parrilo and B. Sturmfels. *Minimizing polynomial functions.* In Proceedings of the DIMACS Workshop on Algorithmic and Quantitative Aspects of Real Algebraic Geometry in Mathematics and Computer Science, pages 83–100. American Mathematical Society, 2003. (Cited on page 46.)

[Putinar 1993] M. Putinar. *Positive polynomials on compact semi-algebraic sets.* Indiana Univ. Math J., vol. 42, pages 969–984, 1993. (Cited on page 35.)

[Riener 2013] C. Riener, T. Theobald, L. J. Andrén and J. B. Lasserre. *Exploiting Symmetries in SDP-Relaxations for Polynomial Optimization.* Math. Oper. Res., vol. 38, no. 1, pages 122–141, February 2013. (Cited on page 65.)

[Rostalki 2009] P. Rostalki. *Algebraic moments,real root finding and related topics.* PhD thesis, ETH Zurich, 2009. (Cited on page 62.)

[Safey El Din 2008] M. Safey El Din. *Computing the global optimum of a multivariate polynomial over the reals.* In Proceedings of the twenty-first international symposium on Symbolic and algebraic computation, ISSAC '08, pages 71–78, New York, NY, USA, 2008. ACM. (Cited on page 46.)

[Schmüdgen 1991] K. Schmüdgen. *The K-moment problem for compact semi-algebraic sets.* Math Ann., vol. 289, pages 529–543, 1991. (Cited on page 33.)

[Schweighofer 2006] M. Schweighofer. *Global optimization of polynomials using gradient tentacles and sums of squares.* SIAM Journal on Optimization, vol. 17, no. 3, pages 920–942, 2006. (Cited on page 46.)

[Shafarevich 1974] Shafarevich. Basic algebraic geometry. Springer-Verlag, 1974. (Cited on page 48.)

[Shor 1987] N.Z. Shor. *Class of global minimum bounds of polynomial functions.* Cybernetics, vol. 23, pages 731–734, 1987. (Cited on page 31.)

[Shor 1998] N.Z. Shor. Nondifferentiable optimization and polynomial optimization. Springer, 1998. (Cited on page 31.)

[Zaccarian 2009] L. Zaccarian, S. Galeani, M. Francaviglia, C. T. Abdallah, E Schamiloglu and Buchenauer C. J. *A control theory approach on the design of Marx generator network.* IEEE Pulsed Power Conference, 2009. (Cited on pages 111 and 112.)

Résumé étendu

Dans le premier chapitre nous introduisons les définitions et théorèmes principaux sur les idéaux et les variétés, nous introduisons aussi l'espace dual et définissons un élément important dans notre étude, les matrices et opérateurs de Hankel et les propriétés de l'anneau quotient qui s'obtient comme quotient par le noyau de notre opérateur d'Hankel.

Dans le deuxième chapitre, nous introduisons notre problème d'optimisation polynomial :

$$\begin{aligned} f^* = inf_{\mathbf{x}\in\mathbb{R}^n} \quad & f(\mathbf{x}) \\ s.t. \quad & g_1^0(\mathbf{x}) = \cdots = g_{n_1}^0(\mathbf{x}) = 0 \\ & g_1(\mathbf{x}) \geq 0, ..., g_{n_2}(\mathbf{x}) \geq 0 \end{aligned} \tag{25}$$

qui consiste à minimiser une fonction polynomiale à coefficients réels dans l'ensemble semialgébrique

$$S := \mathcal{S}(\mathbf{g}) = \{\mathbf{x} \in \mathbb{R}^n \mid g_1^0(\mathbf{x}) = 0, \ldots, g_{n_1}^0(\mathbf{x}) = 0 g_1(\mathbf{x}) \geq 0, \ldots, g_{n_2}(\mathbf{x}) \geq 0\} \tag{26}$$

Nous étudions les différentes variétés associées aux points critiques: la variété gradiente, la variété de Karush-Kuhn-Tucker (KKT) et enfin la variété de Fritz-John (FJ). Pour cette dernière variété, il n'y a pas beaucoup de travaux (voir [Lasserre 2009a]) et grace à elle nous allons pouvoir traiter le cas où le minimum de notre fonction n'est pas atteint en un point de la variété KKT bien que le minimum de notre fonction soit atteint en un point de S (voir exemple 2.4.1). Nous montrons que tout point de S qui minimise notre fonction f est la projection d'un point réel de la variété de FJ. Nous montrons que la variété de FJ est l'union de la variété des points singuliers et de la variété de KKT. Comme chercher le minimum de la fonction f sur la variété des points singuliers est un problème du même type que chercher le minimimun de la fonction f sur la variété de KKT, nous pouvons réduire notre étude à ce dernier type de problèmes sur la variété de KKT.

Dans le troisième chapitre, nous présentons d'abord les définitions de polynômes positifs et sommes de carrés (SOS), de module quadratique $\mathcal{Q}(\mathbf{g})$ et de preordering $\mathcal{P}(\mathbf{g})$ et les théorèmes de Putinar et Schmudgen qui relient tous ces définitions. Puis nous relions notre problème (25) avec la théorie des matrices de moments et avec l'ensemble des polynômes positifs. D'un côté nous pouvons regarder notre problème comme un problème de maximisation :

$$\boxed{f^* = \sup \rho \ t.q. \ f(\mathbf{x}) - \rho \geq 0 \ dans \ S}$$

Ce problème est très difficil à resoudre. Nous avons essayer de le résoudre en utilisant un autre type de problème pour $S = \mathbb{R}^n$

$$\boxed{f^{sos} = \sup \rho \ t.q. \ f(\mathbf{x}) - \rho \ est \ sos}$$

Ceci nous rammènne à un problème de programmation semidéfinie (SDP). Dans le cas où $S = S(\mathbf{g})$ est un ensemble semialgébrique (25)

$$\boxed{f_{\mathbf{g}}^{sos} = \sup \rho \; t.q \; f - \rho \in \mathcal{P}(\mathbf{g})}$$

Quand nous bornons le degré de nos polynômes, la formulation ci-dessus nous conduit à un problème SDP semidéfinie de dimension finie. Nous considerons pour t tel que $2t \geq max(deg(f), deg(g_1^0), \ldots, deg(g_{n_1}^0), deg(g_1^+, \ldots, deg(g_{n_2}^+))$ le problème tronqué :

$$\boxed{f_{t,\mathbf{g}}^{sos} = \sup \rho \; t.q. \; f - \rho \in \mathcal{P}_t(\mathbf{g})}$$

Nous pouvons alors construire une série de problèmes (croissant en le degré t) dont la série de maximums converge vers le minimum de notre problème de départ

$$\boxed{f_{t,\mathbf{g}}^{sos} \leq f_{t1,\mathbf{g}}^{SOS} \leq f_{\mathbf{g}}^{sos} \leq f^*}$$

D'un autre côté, nous pouvons regarder notre problème comme un problème de minimisation:

$$\boxed{f^* = \inf_\mu \int_S f(x)\mu(dx)}$$

où l'infimum est pris sur toutes les mesures de probabilité dans $\mathbb{R}^n$ supportée pour S. Comme $\int f(x)\mu(dx) = \sum_\alpha f_\alpha \int x^\alpha \mu(dx) = f^T y$, ou $y = \int x^\alpha \mu(dx)$ est le vecteur de moments associé à μ, la formulation ci-dessus est équivalente à :

$$\boxed{f^* = \inf f^T y \; t.q. \; y_0 = 1, \; y \; a \; une \; mesure \; representative \; en \; S}.$$

Le problème ci-dessus est aussi tres difficile a resoudre. Nous considerons une borne inférieure en imposant des conditions sur la matrice des moments:

$$\boxed{f_{\mathbf{g}}^{\mu} = \inf_{y \in \mathbb{R}^{\mathbb{N}^n}} f^T y \; s.t. \; y_0 = 1, \; H(y) \succcurlyeq 0, \; H(g_j y) \succcurlyeq 0}$$

Comme travailler avec des formes linéaires est la même chose que travailler avec des vecteur de moments nous étudions le problème suivant :

$$\boxed{f_{\mathbf{g}}^{\mu} = \inf_{\Lambda \in \mathbb{R}[\mathbf{x}]^*} \Lambda(f) \; sT. \; \Lambda(1) = 1, \; \Lambda(p) \geq 0 \; \forall p \in \mathcal{P}(\mathbf{g})}$$

Dans les deux derniers problèmes, nous travaillons avec des matrices de moments infinies et des matrices de Hankel infinies (respectivement). En 2001, Lasserre (voir [Lasserre. 2001]) a eu l'idée de tronquer ces problèmes pour trouver des problèmes de dimension finie que nous pouvons résoudre avec des méthodes SDP.

$$\boxed{f_{t,\mathbf{g}}^{\mu} = \inf_{y \in \mathbb{R}_{2t}^{\mathbb{N}}} f^T y \; s.t. \; y_0 = 1, \; H_t(y) \succcurlyeq 0, \; H_{t-d_{g_j}}(g_j y) \succcurlyeq 0 (j = 1...., n_2)}$$

ou

$$\boxed{f_{t,\mathbf{g}}^{\mu} = \inf_{\Lambda \in \mathbb{R}[\mathbf{x}]_{2t}^*} \Lambda(f) \; t.q. \; \Lambda(1) = 1, \; \Lambda(p) \geq 0 \; \forall p \in \mathcal{P}_t(\mathbf{g})}$$

où $t \geq max(d_f, d_S)$, et

$$\mathcal{L}_{t,\mathbf{g}} := \{\Lambda \in \mathbb{R}[\mathbf{x}]^*_{2t} \mid \Lambda(p) \geq 0,\ \forall p \in \mathcal{P}_t(\mathbf{g}), \Lambda(1) = 1\}.$$

Donc cette dernière formulation est égale à:

$$\boxed{f^{\mu}_{t,\mathbf{g}} = \inf_{\Lambda \in \mathbb{R}[\mathbf{x}]^*_{2t}} \Lambda(f)\ t.q.\ \Lambda \in \mathcal{L}_t(\mathbf{g})}$$

Lasserre montre que nous pouvons construire une séquence de problèmes d'optimisation convexes tronqués, dont les minimums convergent vers le minimimun de notre problème du départ.

$$\boxed{f^{\mu}_{t,\mathbf{g}} \leq f^{\mu}_{t+1,\mathbf{g}} \leq \cdots \leq f^{\mu} \leq f^*}$$

Dans le quatrième chapitre, nous étudions quand la séquence ci-dessus converge en un nombre fini de pas, c'est-à-dire, quand s'il y a une convergence exacte. Il y a de nombreuses études sur ce sujet (voir par exemple [Nie 2006, Laurent 2007, Laurent 2009a, Nie 2013a, Nie 2011, Marshall 2009, Demmel 2007, Ha 2010]. Dans une première partie nous étendons les résultats sur la représentation de polynômes qui sont positifs aux les points critiques (voir théorème principal 4.1.9). Cette représentation ne va dépendre que de la variable $\mathbf{x}$ et donc ce théorème généralise le résultat de [Demmel 2007]. Nous utilisons ces résultats sur polynômes positifs pour montrer les deux résultats principaux suivants :

- Si l'ensemble de minimiseurs KKT est vide, notre séquence de problèmes est vide aussi et vice versa (voir Proposition 4.2.1).
- Nous pouvons construire des séquences exactes de problèmes qui dépendent seulement de la variable $\mathbf{x}$. De plus l'exactitude de la séquence ne va dépendre que de points réels et l'idéal minimiseur peut être construit à partir du noyau de la matrice des moments associée à la forme linéaire même si l'idéal des minimiseurs n'est pas zéro dimensionnel. Sous certaines conditions de regularité nous montrons qu'il n'y pas d'écart de dualité, c'est-à-dire que le minimum de μ-problème et le maximum de sos-problème sont les mêmes et égaux au minimimun de notre problème initial (voir Théorème 4.2.10)

Enfin nous montrons des conséquences de ces résultats sûr des cas particuliers: optimisation globale, le cas dont $\mathbf{g}$ est régulair, le cas où la variété réelle des $(\mathbf{g}^0)$ est de dimension finie, le cas où la variété réelle de $(\mathbf{g}^0)$ (quand $\mathbf{g} = (\mathbf{g}^0)$) est lisse et équidimensionnelle et le cas où nous voulons calculer juste les points réels de S (en minimisant $f = 0$).

Dans le cinquième chapitre, nous proposons un nouvel algorithme pour résoudre notre problème quand le nombre de points qui minimisent notre fonction est

fini. On utilise les bases de bord que nous présentons dans le début du chapitre avec des exemples, définitions et théorèmes qui nous permettent de connaître ces bases en plus détail. Les bases de bord sont une généralisation des bases de groebner mais avec une meilleure stabilité. Grâce à l'usage des bases de bord nous pouvons réduire la taille de nos matrices de moments et le nombre de paramètres associés au notre problème (SDP). La complexité de ces SDP problèmes est $\mathcal{O}((p\,s^{3.5}c\,p^2s^{2.5}c\,p^3s^{0.5})\log(\varepsilon^{-1}))$ où $\varepsilon > 0$ est la précision d'approximation, s est la taille of the moment matrices, p est le nombre de paramètres et c est le nombre de contraintes. La solution de la SDP est une forme linéaire qui est optimale (elle minimise notre function f). Pour vérifier si le minimimun de notre problème initiale est atteint, notre matrice de moments (Hankel) solution du probléme SDP doit vérifier une généralisation du critère de l'extension plate de Curto Fialko (voir [Laurent 2009b, Curto 1996]) . Nous proposons un nouvel algorithme qui vérifie ce critère grâce a l'usage de polynômes orthogonaux. À la sortie de cet algorithme si le critère est vérifié nous obtenons une base et les relations dans le noyau de notre matrice de moments. Avec ces deux éléments nous pouvons construire les matrices de multiplication associées aux différentes variables et obtenir les points qui annulent le noyau, c'est-à-dire les points qui minimisent notre fonction f. À la fin du chapitre nous donnons des exemples qui montrent comment notre algorithme marche. Le premier montre plus en détail les différentes étapes de notre algorithme. Dans le cas où l'idéal n'est pas zéro dimensionnel, il n'y a pas de critère pour arrêter notre algorithme, mais nous savons que nous pouvons récupérer les points qui minimisent notre fonction f en regardant le noyau de la matrice de moments comme le montrent les deux derniers exemples du chapitre.

Dans le dernier chapitre nous analysons le comportement pratique de notre algorithme. Nous decrivons une série d'expérimentations en comparant notre algorithme avec celui de la méthode de Lasserre implémente en Gloptipoly pour Matlab. Nous avons codé notre algorithme et celui de gloptipoly en C++ pour pouvoir comparer dans le même enviroment. Nous montrons des tableaux où nous pouvons voir la différence de taille de matrices, de nombre de paramètres et de temps d'exécution. Pour la résolution du SDP, on utilise différents logiciels: SDPA, CSDP, SDPA-GMP et MOSEK. Ce dernier réduit le temps d'exécution entre un 50% et un 80%. Nous montrons comment utiliser notre algorithme. À la fin, nous donnons trois applications de notre algorithme dans trois domaines différents:

- Le traitement des signaux et télécommunications : en calculant la meilleure approximation de range 1 et 2.
- La biologie : détermination des facteurs plus représentatifs dans la croissance des racines de plantes.
- L'électronique : le problème du générateur de Marx comme un problème d'assignation de valeurs propres.

Matrices de Moments, Géométrie algébrique réelle et Optimisation polynomiale

Le but de cette thése est de calculer l'optimum d'un polynôme sur un ensemble semi-algébrique et les points où cet optimum est atteint. Pour atteindre cet objectif, nous combinons des méthodes de base de bord avec la hiérarchie de relaxation convexe de Lasserre afin de réduire la taille des matrices de moments dans les problèmes de programmation semidéfinie positive (SDP). Afin de vérifier si le minimum est atteint, nous apportons un nouveau critère pour vérifier l'extension plate de Curto Fialkow utilisant des bases orthogonales. En combinant ces nouveaux résultats, nous fournissons un nouvel algorithme qui calcule l'optimum et les points minimiseurs. Nous décrivons plusieurs experimentations et des applications dans diffèrents domaines qui montrent les perfomances de l'algorithme. Au niveau théorique nous prouvons aussi la convergence finie d'une hiérarchie SDP construite à partir d'un idéal de Karush-Kuhn-Tucker et ses conséquences dans des cas particuliers. Nous étudions aussi le cas particulier où les minimiseurs ne sont pas des points de KKT en utilisant la variété de Fritz-John.

Mots clés: Optmisation polynomiale, Matrices de Moments, Method de relaxation, Base de Bord, extension plate, Programation SemiDéfinie, variete de Fritz-John

---

Moments Matrices, Real Algebraic Geometry and Polynomial Optimization

The objective of this thesis is to compute the optimum of a polynomial on a closed basic semialgebraic set and the points where this optimum is reached. To achieve this goal we combine border basis method with Lasserre's hierarchy in order to reduce the size of the moment matrices in the SemiDefinite Programming (SDP) problems. In order to verify if the minimum is reached we describe a new criterion to verify the flat extension condition using border basis. Combining these new results we provide a new algorithm which computes the optimum and the minimizers points. We show several experimentations and some applications in different domains which prove the perfomance of the algorithm. Theorethically we also prove the finite convergence of a SDP hierarchie contructed from a Karush-Kuhn-Tucker ideal and its consequences in particular cases. We also solve the particular case where the minimizers are not KKT points using Fritz-John Variety.

Keywords: Polynomial optmization, Moment matrices, relaxation method, border basis, flat extension, SemiDefinite Programming, Fritz-John variety

---