Detecting optimality and extracting optimal solutions in polynomial optimization based on the Lasserre relaxation

Doctoral thesis for obtaining the academic degree Doctor of Natural Sciences

submitted by

López Quijorna, María

at the

Universität Konstanz

Faculty of Sciences
Department of Mathematics and Statistics

Day of the oral examination: 8th February, 2018
1. Referee: Prof. Dr. Markus Schweighofer
2. Referee: Prof. Dr. Didier Henrion

Summary

Abstract:
A basic closed semialgebraic subset of $\mathbb{R}^n$ is defined by simultaneous polynomial inequalities: $p_1 \geq 0, \ldots, p_m \geq 0$ where $p_1, \ldots, p_m \in \mathbb{R}[X_1, \ldots, X_n]$. We consider Lasserre's relaxation hierarchy to solve the problem of minimizing a polynomial over such a set, that is a polynomial optimization problem. These relaxations give an increasing sequence of lower bounds of the infimum. In this work we provide a new certificate for the optimal value of a Lasserre relaxation to be the optimal value of the polynomial optimization problem. This certificate is that a modified version of an optimal solution of the Lasserre relaxation is a generalized Hankel matrix. This certificate is more general than the already known certificate of an optimal solution being flat. In case that this modified matrix has a generalized Hankel form we will extract the potential minimizers with a truncated version of the Gelfand-Naimark-Segal construction on the optimal solution of the Lasserre relaxation. We prove also that the operators of this truncated construction commute if and only if the matrix of this modified optimal solution is a generalized Hankel matrix. This generalization of flatness will bring us to prove a result of Curto and Fialkow on the existence of quadrature rule if the optimal solution is flat and a result of Xu and Mysovskikh on the existence of a Gaussian quadrature rule if the modified optimal solution is a generalized Hankel matrix. We give a numerical linear algebraic algorithm for detecting optimality and extracting solutions of a polynomial optimization problem.

Finally, we provide an experimental algorithm that in many cases gives us an upper bound of the global infimum of a real polynomial on $\mathbb{R}^n$. The algorithm that we present involves to solve a series of semidefinite programs whose feasible set is included in the feasible set of a moment relaxation. Our additional constraint try to provoke a flatness condition, like used by Curto and Fialkow, for the computed moments. At the end we present numerical results of the application of the algorithm to nonnegative polynomials which are not sums of squares. We also provide numerical results for the application of a version of this algorithm based on the method proposed by Nie, Demmel and Sturmfels for the problem of minimizing a polynomial over its gradient variety.

Zusammenfassung:

Eine basis-abgeschlossene semialgebraische Menge im $\mathbb{R}^n$ ist durch ein System polynomieller Ungleichungen $p_1 \geq 0, \ldots, p_m \geq 0$ definiert, wobei $p_1, \ldots, p_m \in \mathbb{R}[X_1, \ldots, X_n]$. Wir betrachten das polynomielle Optimierungsproblem, ein Polynom über dieser Menge zu minimieren. Wir wollen eine Lösung mit Hilfe der Lasserre-Relaxierungs-Hierarchie finden. Diese besteht aus einer Abfolge einfacherer Optimierungsprobleme. Die Folge deren Optimalwerte ist wachsend und liegt unter dem Optimalwert des Ursprungsproblems. In dieser Arbeit präsentieren wir ein neues Zertifikat dafür, dass der Optimalwert einer Lasserre-Relaxierung dem Optimalwert des Ursprungsproblems entspricht. Dieses Zertifikat besteht darin, dass eine modifizierte Version einer Optimalstelle in der Lasserrerelaxierung die Form einer verallgemeinerten Hankelmatrix hat, und ist allgemeiner als ein bekanntes Kriterium von Curto und Fialkow, welches eine flache Optimalstelle in der Lasserrerelaxierung benötigt. Ist eine Optimalstelle der Lasserrerelaxierung eine verallgemeinerte Hankelmatrix, konstruieren wir aus dieser potenzielle Minimierer des polynomiellen Optimierungsproblems. Dazu benutzen wir eine trunkierte Version der Gelfand-Naimark-Segal-Konstruktion, welche zulässige Punkte in der Lassserrerelaxierung (Zustände) in Operatoren (für jede Variable ein Operator) übersetzt. Wir beweisen, dass diese Operatoren kommutieren, wenn die Modifikation des Minimierers von verallgemeinerter-Hankel-Form ist. Mit dieser Technik können wir einen Satz von Curto und Fialkow über die Existenz einer Quadraturformel für flache Minimierer und einen weiteren Satz von Xu und Mysovskikh über die Existenz einer Gaussquadraturformel für Minimierer, deren Modifikation verallgemeinerte-Hankel-Gestalt hat, beweisen. Insbesondere stellen wir einen Numerische-Lineare-Algebra-Algorithmus bereit, der verifizieren kann, ob der Optimalwert der Lasserrerelaxierung dem Optimalwert des polynomiellen Optimierungsproblems entspricht und eine Minimalstelle dessen liefert.

Zuletzt stellen wir einen experimentellen Algorithmus vor, welcher uns in vielen Fällen eine obere Schranke für das globale Infimum eines reellen Polynoms auf $\mathbb{R}^n$ gibt. Dieser beinhalt das Lösen einer Reihe von semidefiniten Programmen, deren zulässiger Bereich ein Teil des zulässigen Bereiches einer Momentenrelaxierung ist. Unsere zusätzlich gestellten Bedingungen zielen darauf ab, eine Flachheitsbedingung, wie sie auch von Curto und Fialkow verwendet wird, für die berechneten Momente zu erzwingen. Zum Schluss präsentieren wir numerische Ergebnisse des Algorithmus. Dazu minimieren wir nichtnegative Polynome, die jedoch keine Quadratsummendarstellung besitzen. Zudem arbeiten wir Techniken von Nie, Demmel und Sturmfels in unseren Algorithmus ein, um ein Polynom über der durch seinen Gradienten gebenen Varietät zu minimieren und geben numerische Beispiele.

Resumen:
Un subconjunto semialgebráico básico y cerrado de $\mathbb{R}^n$ es aquel que está definido simultáneamente por un número finito de desigualdades polinomiales de la forma: $p_1 \geq 0, \ldots, p_m \geq 0$ donde $p_1, \ldots, p_m \in \mathbb{R}[X_1, \ldots, X_n]$. Para resolver el problema de minimización de un polinomio $f \in \mathbb{R}[X_1, \ldots, X_n]$ sobre éste utilizamos la relajación de los momentos de Lasserre. Con esta jerarquía de relajaciones obtenemos una sucesión creciente de cotas inferiores del ínfimo. En este trabajo proporcionamos un nuevo criterio para certificar que el valor óptimo de una relajación de los momentos de cierto grado es igual que el valor del óptimo de nuestro problema de optimización polinomial. Este criterio se basa en que la matrix asociada a una versión modificada de la solución optima de la relajación de los momentos de Lasserre sea de forma hankel generalizada. Este criterio es más general que el conocido criterio de comprobar que la matriz asociada a una solución óptima de la relajacion de los momentos sea flat. En caso de que esta matriz sea de forma hankel generalizada extraeremos minimizadores potenciales a través de la construción de una versión truncada de la conocida construcción de Gelfand, Naimark y Segal, asociada a la solución óptima de la relajacion de Lasserre. También probaremos que los operadores truncados de esta construcción conmutan si y solo si la versión modificada de la solución óptima es una matriz hankel generalizada. Esta generalización del concepto de flat nos llevará a volver a probar un resultado de Curto y Fialkow sobre la existencia de una cuadratura para la solución óptima si y solo si la solución óptima es una matriz flat y también nos llevará a volver a probar un resultado de Xu y Mysovskikh sobre la existencia de una cuadratura gaussiana si la versión modificada de la solución óptima es una matriz hankel generalizada. En conclusión, proporcionaremos un algoritmo algebraico lineal que detecta optimalidad y extrae minimizadores de nuestro problema de optimización de polinomios.

Finalmente, proporcionamos un algoritmo experimental que en muchos casos nos dará una cota superior del ínfimo de un polinomio con coeficientes reales sobre $\mathbb{R}^n$. El algoritmo que presentamos supone resolver una serie de programas semidefinidos positivos cuyo conjuncto factible está contenido en el conjunto factible de la relajación de los momentos. La restricción adicional en este conjunto intenta provocar una solución flat. Al final mostraremos ejemplos numéricos de la aplicación de este algoritmo sobre polinomios no negativos que no son suma de cuadrados. También proporcionaremos ejemplos numéricos de la aplicacion de una versión de este algoritmo utilizando ideas del método propuesto por Nie, Demmel y Sturmmfels por el problema de minimizar un polinomio sobre su variedad gradiente.

Acknowledgments

Before starting I would like to express my gratitude to some people who made this work possible.

My greatest and sincere thanks goes to my advisor, Markus Schweighofer, for all the time, effort and support he gave me, for all the fruitful discussions and for providing a nice working environment. I will be forever thankful for this enormous opportunity.

My sincere thanks to my professors in the University Complutense in Madrid, in particular thanks to Antonio Díaz-Cano Ocaña and to María Emilia Alonso García for supervising my master's thesis.

Last but not least, I am very grateful to my family for all their love, patience and support.

Contents

CHAPTER 1

Notation

Throughout this paper, we suppose $n \in \mathbb{N} = \{1, 2, \ldots\}$ and abbreviate $(X_1, \ldots, X_n)$ by $\underline{X}$. We let $\mathbb{R}[\underline{X}]$ denote the ring of real polynomials in n indeterminates. We denote $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. For $\alpha \in \mathbb{N}_0^n$, we use the standard notation :

$$|\alpha| := \alpha_1 + \cdots + \alpha_n \text{ and } \underline{X}^\alpha := X_1^{\alpha_1} \cdots X_n^{\alpha_n}$$

For a polynomial $p \in \mathbb{R}[\underline{X}]$ we denote $p = \sum_\alpha p_\alpha \underline{X}^\alpha$ $(a_\alpha \in \mathbb{R})$. For $d \in \mathbb{N}_0$, by the notation $\mathbb{R}[\underline{X}]_d := \{\sum_{|\alpha| \leq d} a_\alpha \underline{X}^\alpha \mid a_\alpha \in \mathbb{R}\}$ we will refer to the vector space of polynomials with degree less or equal to d. Polynomials all of whose monomials have exactly the same degree $d \in \mathbb{N}_0$ are called d-forms. They form a finite dimensional vector space that we will denote by:

$$\mathbb{R}[\underline{X}]_{=d} := \left\{ \sum_{|\alpha|=d} a_\alpha \underline{X}^\alpha \mid a_\alpha \in \mathbb{R} \right\}$$

so that

$$\mathbb{R}[\underline{x}]_d = \mathbb{R}[\underline{X}]_{=0} \oplus \cdots \oplus \mathbb{R}[\underline{X}]_{=d}.$$

We will denote by $s_k := \dim \mathbb{R}[\underline{X}]_k$ and by $r_k := \dim \mathbb{R}[\underline{X}]_{=k}$. For $d \in \mathbb{N}_0$ we denote $\mathbb{R}[\underline{X}]_d^*$ the dual space of $\mathbb{R}[\underline{X}]_d$ i.e. the set of linear forms from $\mathbb{R}[\underline{X}]_d$ to $\mathbb{R}$ and for $L \in \mathbb{R}[\underline{X}]_{2d}^*$ we denote by $L' := L_{|\mathbb{R}[\underline{X}]_{2d-2}}$ the restriction of the linear form L to the space $\mathbb{R}[\underline{X}]_{2d-2}$. For $d \in \mathbb{N}_0$ and $a \in \mathbb{R}^n$ we denote $\mathrm{ev}_a \in \mathbb{R}[\underline{X}]_d^*$ the linear form such that for all $p \in \mathbb{R}[\underline{X}]_d$, $\mathrm{ev}_a(p) = p(a)$. Given a ring R and $A, B \in R$ we denote $[A, B]$ the commutator of A and B, that is to say:

$$[A, B] := AB - BA$$

CHAPTER 2

Introduction

Let polynomials $f, p_1, \ldots, p_m \in \mathbb{R}[\underline{X}]$ with $m \in \mathbb{N}_0$ be given. A polynomial optimization problem consists of finding the infimum of f over the so called basic closed semialgebraic set S, defined by:

$$S := \{x \in \mathbb{R}^n \mid p_1(x) \geq 0, \ldots, p_m(x) \geq 0\} \tag{2.1}$$

and also, if it is possible, a polynomial optimization problem involves extracting optimal points or minimizers i.e. elements in the set:

$$S^* := \{x^* \in S \mid \forall x \in S, f(x^*) \leq f(x)\}$$

So from now on we will denote by (P), to refer to the above defined polynomial optimization problem, that is to say:

$$(P) \text{ minimize } f(x) \text{ subject to } x \in S \tag{2.2}$$

The optimal value of (P), i.e. the infimum of $f(x)$ where x ranges over all feasible solutions S will be denoted by P^*, that is to say:

$$P^* := \inf\{ f(x) \mid x \in S\} \in \{-\infty\} \cup \mathbb{R} \cup \{\infty\}$$

Note that $P^* = +\infty$ if $S = \emptyset$ and $P^* = -\infty$ if and only if f is unbounded from below on S, for example if $S = \mathbb{R}^n$ and f is of odd degree.

In 2001, Lasserre presented a remarkable method to tackle this problem. This method constructs a hierarchy of semidefinite programs $(P_k)_{k\in\mathbb{N}}$. Every semidefinite program (P_k) in this method is the program that results after linearizing an equivalent formulation of the problem (P). This equivalent formulation of (P) is created by adding infinitely many redundant inequalites of the form $p \geq 0$ for all $p \in \sum \mathbb{R}[\underline{X}]^2 p_i \cap \mathbb{R}[\underline{X}]_k$. The set of this redundant inequalities builds a *cone*, which is a set containing 0, closed under addition and under multiplication by positive scalars. The cone generated by this redundant inequalities is called the *truncated quadratic module generated by the polynomials* $p_1, \dots, p_m$. More precise, let $p_1, \dots, p_m \in \mathbb{R}[\underline{X}]_k$ and $k \in \mathbb{N}_0 \cup \{\infty\}$. We define and denote the k-truncated quadratic module generated by $p_1, \dots, p_m$ as:

$$M_k(p_1, \dots, p_m) := \left(\mathbb{R}[\underline{X}]_k \cap \sum \mathbb{R}[\underline{X}]^2\right) + \left(\mathbb{R}[\underline{X}]_k \cap \sum \mathbb{R}[\underline{X}]^2 p_1\right) + \dots + \left(\mathbb{R}[\underline{X}]_k \cap \sum \mathbb{R}[\underline{X}]^2 p_m\right) \subseteq \mathbb{R}[\underline{X}]_k$$

and *the moment relaxation of degree k* is the following semidefinite program:

$$(P_k): \begin{cases} \text{minimize} & L(f) \\ \text{subject to:} & L \in \mathbb{R}[\underline{X}]_k^* \\ & L(1) = 1 \text{ and} \\ & L(M_k(p_1, ..., p_m)) \subseteq \mathbb{R}_{\geq 0} \end{cases}$$

the optimal value of (P_k) i.e., the infimum over all $L(f)$ where L ranges over all optimal solutions of (P_k) is denoted by $P_k^* \in \{-\infty\} \cup \mathbb{R} \cup \{\infty\}$. These relaxations give an increasing sequence $(P_k^*)_{k\in\mathbb{N}}$ of lower bounds of the infimum and Lasserre proved that this sequence converge asymptotically to the infimum providing that some archimedean property in the quadratic module $M(p_1, \dots, p_m)$ holds. However in many cases it is possible to have *optimality*; this is to achieve $P^* = P_k^*$ for some $k \in \mathbb{N}$. One of the main goals of this work is to look for a criterion to guarantee optimality from an optimal solution of the moment relaxation and if this is the case to extract minimizers from it.

Let $L \in \mathbb{R}[\underline{X}]_k^*$ be a feasible solution of the moment relaxation of degree k, and set $d := \lfloor \frac{k}{2} \rfloor$ and define:

$$V_d := (1, X_1, X_2, \dots, X_N, X_1^2, X_1X_2, \dots, X_1X_n, X_2^2, X_2X_3, \dots, X_n^2, \dots, \dots, X_n^d)^T$$

as a basis for the vector space of polynomials in n variables of degree at most d. We call and denote *the moment matrix associated to the linear form L* to the matrix indexed by this basis:

$$M_L := \begin{pmatrix} L(1) & L(X_1) & L(X_2) & \cdots & L(X_n^d) \\ L(X_1) & L(X_1^2) & L(X_1X_2) & \cdots & L(X_1X_n^d) \\ L(X_2) & L(X_1X_2) & L(X_2^2) & \cdots & L(X_2X_n^d) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ L(X_n^d) & L(X_1X_n^d) & L(X_2X_n^d) & \cdots & L(X_n^{2d}) \end{pmatrix} \in \operatorname{Sym} \mathbb{R}^{s_d \times s_d}$$

It is not difficult to prove that $a^T M_L a \geq 0$ for all $a \in \mathbb{R}^{1\times s_d}$, that is to say M_L is *positive semidefinite*, in symbols $M_L \succeq 0$. Every matrix indexed by a basis of monomials of $\mathbb{R}[\underline{X}]_d$ is said to be *generalized Hankel (in n variables) of degree d* therefore M_L is obviously a generalized Hankel matrix of degree d. In the rest of the script, we will refer just to a generalized Hankel matrix if the degree and number of variables are known.

In this work we will study the following topics:

- In Chapter 5, we will give a new proof of a useful result that was proven in 1959 by Smul'jan in [36, Proposition 2.2] , which says that there exists matrices: $W \in \mathbb{R}^{r_{d-1} \times s_d}$ and $X \in \mathbb{R}^{s_d \times s_d}$ such that:

$$M_L := \left(\begin{array}{c|c} M_{L'} & M_{L'}W \\ \hline W^T M_{L'} & W^T M_{L'} W + XX^T \end{array} \right)$$

where $L' := L_{|\mathbb{R}[\underline{X}]^*_{2d-2}}$ and the following matrix, that we call *the modified moment matrix of* L, is well defined:

$$\widetilde{M_L} := \left(\begin{array}{c|c} M_{L'} & M_{L'}W \\ \hline W^T M_{L'} & W^T M_{L'} W \end{array} \right)$$

- In 2000, Curto and Fialkow in [4, Corollary 5.14] proved that if $M_L = \widetilde{M_L}$, or in other words if L is *flat*, then there exists a *quadrature rule for* L with $N := \operatorname{rank} M_L$ *nodes*: that is to say there exist $a_1, \ldots, a_N \in \mathbb{R}^n$ and $\lambda_1 > 0, \ldots, \lambda_N > 0$ such that:

$$L(p) = \sum_{i=1}^{N} \lambda_i p(a_i) \text{ for all } p \in \mathbb{R}[\underline{X}]_k.$$

Using theory of polynomial optimization, this result of Curto and Fialkow implies that if the optimal solution L is flat and if it holds that the nodes of the quadrature rule are contained in S, then this nodes are also minimizers and we have reached optimality that is $P^* = P_k^*$.

In Chapter 6, we prove that if $\widetilde{M_L}$ is a generalized Hankel matrix, a condition that is more general than flatness, then there exists a *gaussian quadrature rule for* L, that is to say there exists $N := \operatorname{rank} M_{L'}$ points $a_1, \ldots, a_N \in \mathbb{R}^n$ and $\lambda_1 > 0, \ldots, \lambda_N > 0$ such that:

$$L(p) = \sum_{i=1}^{N} \lambda_i p(a_i) \text{ for all } p \in \mathbb{R}[\underline{X}]_{2d-1}.$$

and if $a_1, \ldots, a_n \in S$ and $f \in \mathbb{R}[\underline{X}]_{2d-1}$ we will be able to claim optimality:

$$P^* = P_k^* = f(a_i) \text{ for all } i = 1, \ldots, N.$$

We will give examples of polynomial optimization problems where despite of $M_L \neq \widetilde{M_L}$ it holds $\widetilde{M_L}$ is generalized Hankel and we can still claim optimality: see Examples 6.3.21, 7.2.1 and 7.2.4. Notice that if $\widetilde{M_L}$ is generalized Hankel, in order to claim optimality we need to check if the potential minimizers $a_1, \ldots, a_N$ belongs to S. Curto and Fialkow in [5, Theorem 1.6] showed that if $\operatorname{rank} M_L = \operatorname{rank} M_{L_{|\mathbb{R}[\underline{X}]_{k-s}}}$ for a certain $s \in \mathbb{N}_{\geq 1}$ depending on the degree of the polynomials $p_1, \ldots, p_m$, the points $a_1, \ldots, a_n \in S$. In Chapter 6, we prove that if the polynomials defining the set S are linear then $a_1, \ldots, a_N \in S$ and therefore they are minimizers of (P).

- In Section 6.1 we define the *GNS-truncated construction*. This construction was already done by Putinar in 1997, see [29]. In this section we construct for every variable an operator, *the GNS truncated multiplication operator*, associated to the linear form that gives us information about the nodes of a potential quadrature rule for L. Suppose that L is *positive definite* that is $L(\sum \mathbb{R}[\underline{X}]_d^2 \setminus \{0\}) \subseteq \mathbb{R}_{>0}$ then the commutativity of the GNS truncated multiplication operators is equivalent to the existence of a gaussian quadrature rule for L. This is a classical result proved by Xu in 1994, Mysovskikh in 1976 and by Putinar in 1997. In Section 5.6 we generalized this result for *positive semidefinite linear forms* that is $L(\sum \mathbb{R}[\underline{X}]_d^2) \subseteq \mathbb{R}_{\geq 0}$ and also we show the equivalence between the commutativity of the multiplication operators of L and $\widetilde{M_L}$ been generalized Hankel, a fact that seems not have been noticed so far.

- In Chapter 8, we consider a polynomial optimization problem without constraints and we provide an experimental algorithm based on moments and semidefinite programming that in many cases gives us an upper bound of the global infimum of a real polynomial on $\mathbb{R}^n$.

In Chapter 3, we provide all the necessary preliminaries from convexity theory and semidefinite programming to understand this work. In Chapter 4, we provide tools from polynomial optimization and it contains only few new results. In Chapter 4 we reformulate the problem we want to solve. The last chapter can be read independently of the others.

CHAPTER 3

Preliminaries

3.1 Cones and duality

Definition 3.1.1. Let V be a real vector space. A set $K \subseteq V$ is called a *(convex) cone* in V if $0 \in K$, $K + K \subseteq K$ and $\mathbb{R}_{\geq 0}K \subseteq V$. A cone $K \subseteq V$ is called *proper* if $K \neq V$.

Definition 3.1.2. A subset $M \subseteq \mathbb{R}[\underline{X}]$ is called a *quadratic module (in* $\mathbb{R}[\underline{X}]$*)* if $1 \in M$, $M + M \subseteq M$ and $\mathbb{R}[\underline{X}]^2 M \subseteq M$.

Definition and Notation 3.1.3. Let $t \in \mathbb{N}_0$ and $A \in \operatorname{Sym}\mathbb{R}^{t\times t}$, the notation $A \succeq 0$ (respectively $A \succ 0$) means that A is positive semidefinite (respectively A is positive definite), i.e. $a^T Aa \geq 0$ (respectively $a^T Aa > 0$) for all $a \in \mathbb{R}^t$.

Example 3.1.4. Very useful cones in semidefinite optimization are the following:

(i) The so called *non negative orthant*:

$$\mathbb{R}^n_{\geq 0} := \{x \in \mathbb{R}^n \mid x_1 \geq 0, \dots, x_n \geq 0\} \subseteq \mathbb{R}^n$$

(ii) Let $t \in \mathbb{N}_0$. The cone of positive semidefinite matrix in the real vector space $\operatorname{Sym}\mathbb{R}^{t\times t}$, that is $\operatorname{Sym}\mathbb{R}^{t\times t}_{\succeq 0}$.

(iii) The *sums of squares cone* in $\mathbb{R}[\underline{X}]$:

$$\sum \mathbb{R}[\underline{X}]^2 := \left\{ \sum_{i=1}^{m} p_i^2 \mid m \in \mathbb{N}_0, p_i \in \mathbb{R}[\underline{X}] \text{ for all } i \in \{1, \dots, m\} \right\} \subseteq \mathbb{R}[\underline{X}]$$

(iv) The cone of *nonnegative polynomials* in $\mathbb{R}[\underline{X}]$:

$$P_n := \{p \in \mathbb{R}[\underline{X}] \mid p(x) \geq 0 \text{ for all } x \in \mathbb{R}^n\} \subseteq \mathbb{R}[\underline{X}] \tag{3.1}$$

(v) Every quadratic module $M \subseteq \mathbb{R}[\underline{X}]$ is a cone in $\mathbb{R}[\underline{X}]$.

Let us give further charaterizations of positive semidefiniteness.

Proposition 3.1.5. Let $A \in \operatorname{Sym} \mathbb{R}^{t\times t}$. The following are equivalent:

(i) $A \succeq 0$.

(ii) All eigenvalues of A are nonnegative.

(iii) All coefficients of $\det(A + T \cdot I_t) \in \mathbb{R}[T]$ are nonnegative.

(iv) All the 2^t principal minors are nonnegative.

(v) There exists $B \in \mathbb{R}^{t\times t}$ such that $A = B^T B$.

(vi) A is a non negative linear combination or *conic combination*, of matrices of the form xx^T, where $x \in \mathbb{R}^t$.

Proof. For a proof we refer the reader to [22, Proposition 0.2.1]. □

Definition and Notation 3.1.6. Let V be a real vector space, and $K \subseteq V$ a cone. We use the following notation to refer us to the dual of V and K:

$$V^* := \{L \mid L : V \longrightarrow \mathbb{R} \text{ linear } \}$$

$$K^* := \{L \mid L : V \longrightarrow \mathbb{R},\, L(K) \subseteq \mathbb{R}_{\geq 0} \text{ and } L \text{ is linear } \}$$

Example 3.1.7. It holds that (see [1, Chapter IV, Theorem 4.2], for a proof):

(i) $(\mathbb{R}^n_{\geq 0})^* = \{x \in \mathbb{R}^n \mid \langle x, y\rangle \geq 0 \text{ for all } y \in \mathbb{R}^n_{\geq 0}\}$ where $\langle\ ,\ \rangle$ denotes the euclidean inner product of $\mathbb{R}^n$. Observe that by abuse of terminology we identify every $x \in (\mathbb{R}^n_{\geq 0})^*$ with the linear form $\langle \cdot, x\rangle$ defined as:

$$\langle \cdot, x\rangle : \mathbb{R}^n \longrightarrow \mathbb{R} : y \mapsto \langle y, x\rangle$$

Consequently:

$$(\mathbb{R}^n_{\geq 0})^* = \mathbb{R}^n_{\geq 0}.$$

(ii) Let $t \in \mathbb{N}_0$, it holds that:

$$(\operatorname{Sym} \mathbb{R}^{t\times t}_{\succeq 0})^* = \{A \in \operatorname{Sym} \mathbb{R}^{t\times t} \mid \operatorname{tr}(AB) \geq 0 \text{ for all } B \in \operatorname{Sym} \mathbb{R}^{t\times t}_{\succeq 0}\}$$

Note that $(A, B) \mapsto \operatorname{tr}(AB)$ for $A, B \in \operatorname{Sym} \mathbb{R}^{t\times t}$ defines an inner product [1] in the space of symmetric matrix. Observe that, in the same way as before, we made abuse of terminology by identifying every $A \in (\operatorname{Sym} \mathbb{R}^{t\times t}_{\succeq 0})^*$ with the linear form $\operatorname{tr}(\cdot, A)$ defined as:

$$\operatorname{tr}(\cdot, A) : \operatorname{Sym} \mathbb{R}^{t\times t} \longrightarrow \mathbb{R} : B \mapsto \operatorname{tr}(BA)$$

On other side, for all $A \in \operatorname{Sym} \mathbb{R}^{t\times t}$, it holds:

$$A \in \operatorname{Sym} \mathbb{R}^{t\times t}_{\succeq 0} \iff \operatorname{tr}(AB) \geq 0 \text{ for all } B \in \operatorname{Sym} \mathbb{R}^{t\times t}_{\succeq 0}.$$

See [22, Proposition 10.1.2] for a proof of the last fact. Consequently:

$$(\operatorname{Sym} \mathbb{R}^{t\times t}_{\succeq 0})^* = \operatorname{Sym} \mathbb{R}^{t\times t}_{\succeq 0}.$$

[1] Let $t, s \in \mathbb{N}_0$ and $A, B \in \mathbb{R}^{s\times t}$ for $(A, B) \mapsto \operatorname{tr}(A^T B)$ defined the so called Frobenius inner product.

Remark 3.1.8. Let V be a real vector space and $K \subseteq V$ a cone. Then K^* is always a closed cone, regardless if the cone K is closed or not.

Definition 3.1.9. Let V be a real vector space. A set $H \subseteq V$ is called a halfspace if there exists $L \in V^* \setminus \{0\}$ such that $H = L^{-1}(\mathbb{R}_{\geq 0})$.

Definition 3.1.10. Let K be a cone in a real vector space V and $u \in V$. Then u is called a *unit* of K if for every $v \in V$ there is a $N \in \mathbb{N}$ with $Nu + v \in K$.

Example 3.1.11.

(i) The units for a cone in $\mathbb{R}^n$ are exactly its interior points, see [20, Remark 7.1.8]. In particular, let $t \in \mathbb{N}_0$, the units for the cone $\operatorname{Sym} \mathbb{R}^{t\times t}_{\succeq 0}$ in the real vector space $\operatorname{Sym} \mathbb{R}^{t\times t}$ are exactly the elements of $\operatorname{Sym} \mathbb{R}^{t\times t}_{\succ 0}$.

(ii) The sums of squares cone $\sum \mathbb{R}[\underline{X}]^2$ in the vector space $\mathbb{R}[\underline{X}]$ does not have a unit.

Isolation Theorem for cones 3.1.12. Every proper cone in a real vector space with a unit is contained in a halfspace.

Proof. For a proof of this Theorem we refer the reader to [20, Theorem 7.1.13]. □

Separation Theorem for closed cones 3.1.13. Let V be a real vector space of finite dimension, $K \subseteq V$ a closed cone and $v \in V \setminus K$. Then there is a halfspace H with $K \subseteq H$ and $v \notin H$.

Proof. For a proof we refer the reader to [1, Chapter III, Theorem 1.3] □

Definition 3.1.14. Let $L \in \mathbb{R}[\underline{X}]^*_d$. A quadrature rule for L on $U \subseteq \mathbb{R}[\underline{X}]_d$ is a function $w : N \to \mathbb{R}_{>0}$ defined on a finite set $N \subseteq \mathbb{R}^n$, such that:

$$L(p) = \sum_{x\in N} w(x)p(x) \tag{3.2}$$

for all $p \in U$. A quadrature rule for L is a quadrature for L on $\mathbb{R}[\underline{X}]_d$. We call the elements of N the nodes of the quadrature rule.

Carathéodory's Theorem 3.1.15. Let V be a real vector space, $m \in \mathbb{N}_0$, $v_1, \dots, v_m \in V$ and $\lambda_1, \dots, \lambda_m \in \mathbb{R}_{\geq 0}$. Then there is $s \in \mathbb{N}_0$, $i_1, \dots, i_s \in \{1, \dots, m\}$ and $\mu_1, \dots, \mu_s \in \mathbb{R}_{\geq 0}$ such that $v_{i_1}, \dots, v_{i_s}$ is linearly independent and:

$$\sum_{i=1}^{m} \lambda_i v_i = \sum_{j=1}^{s} \mu_j v_{i_j}$$

Remark 3.1.16.

(i) Let $L \in \mathbb{R}[\underline{X}]^*_d$ and suppose that L has a quadrature rule. Then by Carathéodory Theorem 3.1.15, L has a quadrature rule with at most $\dim \mathbb{R}[\underline{X}]_d$ many nodes.

(ii) Let $p \in \mathbb{R}[\underline{X}]$ such that:

$$p = \sum_{i=1}^{m} u_j^2 \text{ where } u_j \in \mathbb{R}[\underline{X}] \text{ for all } j \in \{1, \ldots, m\}$$

Then $\deg p$ is even and $\deg u_j \leq \frac{d}{2}$ for all $j \in \{1, \ldots, m\}$, where $d := \deg p$, (see [17, Lemma 3.1] or Remark 4.2 for a proof of this last fact). Moreover, by Carathéodory Theorem 3.1.15, p can be written as a sum of at most $\dim \mathbb{R}[\underline{X}]_d$ squares of polynomials.

The following Theorem 3.1.17 is a little exercise but we will include the proof in order to have an idea how the separation theorem and the existence of quadrature rule connect with real algebraic geometry.

Theorem 3.1.17. Let $m \in \mathbb{N}_0$ and $p, p_1, \ldots, p_m \in \mathbb{R}[\underline{X}]_1$ and set:

$$S := \{x \in \mathbb{R}^n \mid p_1(x) \geq 0, \ldots, p_m(x) \geq 0\} \neq \emptyset$$

and

$$K := \mathbb{R}_{\geq 0} \cdot 1 + \mathbb{R}_{\geq 0} p_1 + \ldots + \mathbb{R}_{\geq 0} p_m \subseteq \mathbb{R}[\underline{X}]_1$$

then:

$$p \geq 0 \text{ on } S \iff p \in K$$

Proof. If $p \in K$ then obviously $p \geq 0$ on S. Then, let us prove the converse by contraposition. For this purpose suppose $p \notin K$, we want to show that there exists $x \in S$ such that $p(x) < 0$. It can be proved easily that K is a closed convex cone. The separation Theorem for closed convex cones in finite dimensional vector spaces tells us that there exits $L \in \mathbb{R}[\underline{X}]_1^*$ such that $L(K) \subseteq \mathbb{R}_{\geq 0}$ and $L(p) < 0$. In case $L(1) \neq 0$ then L has a quadrature rule with one node in S, more precise:

$$L = \frac{1}{L(1)} \operatorname{ev}_{(L(X_1), \ldots, L(X_n))}$$

In case $L(1) = 0$, take $x \in S$ and define $\overline{L} := L + \epsilon \operatorname{ev}_x$ where $\epsilon > 0$ is sufficiently small such that $\overline{L}(K) \geq 0$ and $\overline{L}(p) < 0$ and continue the argument as in the previous case. □

Theorem 3.1.18. Let $\ell \in \mathbb{R}[\underline{X}]_1$ and $L \in \operatorname{Sym} \mathbb{R}[\underline{X}]_1^{t \times t}$ for $t \in \mathbb{N}_0$. Let us set:

$$S := \{x \in \mathbb{R}^n \mid L(x) \succeq 0\}$$

suppose that S has non empty interior and set:

$$K := \{a + \operatorname{tr}(AL) \mid a \in \mathbb{R}_{\geq 0}, A \in \operatorname{Sym} \mathbb{R}_{\succeq 0}^{t \times t}\} \subseteq \mathbb{R}[\underline{X}]_1$$

Then:

$$\ell \geq 0 \text{ on } S \iff \ell \in K$$

Proof. For a proof see [34, Proposition 3.2.1]. □

3.2 Linear programming (LP)

A polynomial optimization problem with objective polynomial and feasible set defined by linear polynomials is called a *linear program* and its feasible set is a closed convex set called *polyhedron*. In 1979, Leonid Khachiyan used the so called "Ellipsoid method", developed by Shor, Yudin and Nemirovsky in 1970, to demonstrate that linear programming, in the Turing Machine model has a polynomial-time algorithm.

Notation 3.2.1. We write a primal linear program, in standard form, as:

$$(P) : \begin{cases} & \text{minimize } \ell(x) \\ \text{over:} & x \in \mathbb{R}^n \\ \text{subjet to:} & \ell_1(x) \geq 0, \ldots, \ell_m(x) \geq 0 \end{cases} \tag{3.3}$$

where $m, n \in \mathbb{N}_0$ and $\ell, \ell_1, \ldots, \ell_m \in \mathbb{R}[\underline{X}]_1$. The optimal value of (P) i.e., the infimum over all $x \in \mathbb{R}^n$ where x ranges over all feasible solutions of (P) is denoted by $P^* \in \{-\infty\} \cup \mathbb{R} \cup \{\infty\}$.

Notation 3.2.2. We write the dual program of the linear program (P) (3.3) , in standard form, as:

$$(D) : \begin{cases} & \text{maximize } \mu \\ \text{over:} & \mu, \lambda_1, \ldots, \lambda_m \in \mathbb{R} \\ \text{subjet to:} & \ell = \mu + \lambda_1 \ell_1 + \ldots, \lambda_m \ell_m \text{ and} \\ & \lambda_1 \geq 0, \ldots, \lambda_m \geq 0 \end{cases} \tag{3.4}$$

where $m, n \in \mathbb{N}_0$ and $\ell, \ell_1, \ldots, \ell_m \in \mathbb{R}[\underline{X}]_1$. The optimal value of (D) i.e., the maximum over all $(\mu, \lambda_1, \ldots, \lambda_m) \in \mathbb{R}^{m+1}$, where $(\mu, \lambda_1, \ldots, \lambda_m)$ ranges over all feasible solutions of (D), is denoted by $D^* \in \{-\infty\} \cup \mathbb{R} \cup \{\infty\}$.

Remark 3.2.3. Note that the dual program is also a polynomial optimization problem with linear objective polynomial and linear constraints, that is to say, it can be written as a linear program in primal form. Since making equal coefficients with the same degree on both side of the equality:

$$\ell = \mu + \lambda_1 \ell_1 + \ldots, \lambda_m \ell_m$$

we obtain a system of $n + 1$ linear equations which we can split in $2n + 2$ linear inequalities.

Remark 3.2.4. The feasible set of a linear program can also be seen as the intersection of the cone $\mathbb{R}^n_{\geq 0}$ with an affine subspace in $\mathbb{R}^n$, despite that with the definition 3.2.1 we can not see it directly. This equivalent formulation of a linear program motivates the definition of the dual of the linear program since if we build the dual problem of the dual we get the primal problem. This is a consequence of the fact that the nonnegative orthant is selfdual i.e, $(\mathbb{R}^n_{\geq 0})^* = \mathbb{R}^n_{\geq 0}$. For more information about the geometric interpretation of the primal dual pair we refer the reader to [19, page 46].

Proposition (weak duality of linear program) 3.2.5. Let (P) be the linear program defined in (3.3) and (D) its dual defined in (3.4), then:

$$P^* \geq D^*$$

Proof. Let $x \in \mathbb{R}^n$ be a feasible solution of (P) and $(\mu, \lambda_1, \ldots, \lambda_m)$ be a feasible solution of (D), then:

$$\ell - \mu = \lambda_1 \ell_1 + \ldots + \lambda_m \ell_m$$

and $\ell(x) - \mu \geq 0$ implying weak duality. □

Proposition (strong duality of linear program) 3.2.6. Let (P) be the linear program defined in (3.3) and (D) its dual defined in (3.4). Suppose that $P^* \in \mathbb{R}$, then (P) and (D) possess an optimal solution with the same optimal value. In particular:

$$P^* = D^*$$

Proof. Let us set:

$$S := \{x \in \mathbb{R}^n \mid \ell_1(x) \geq 0, \ldots, \ell_m(x) \geq 0\}$$

Since $\ell \in \mathbb{R}[\underline{X}]_1$ is linear, $\ell(S)$ is closed, since moreover $P^* \in \mathbb{R}$ then there exits $x \in S$ such that $P^* = \ell(x)$, hence let us set for this $x \in S$, $\mu := P^* = \ell(x)$. Due to the definition of minimum it holds:

$$\ell - P^* \geq 0 \text{ on } S$$

By the non-trivial implication of the Theorem 3.1.17:

$$\ell - P^* \in \mathbb{R}_{\geq 0} + \mathbb{R}_{\geq 0}\ell_1 + \ldots + \mathbb{R}_{\geq 0}\ell_m$$

Hence there exist $(\nu, \lambda_1, \ldots, \lambda_m) \in \mathbb{R}^{m+1}$ such that $\ell - P^* = \nu + \lambda_1 \ell_1 + \ldots + \lambda_m \ell_m$. Evaluating this equality at $x \in S$, we get that $0 = \nu + \lambda_1 \ell_1(x) + \ldots + \lambda_m \ell_m(x)$. As $\nu \geq 0$ and $\lambda_i \ell_i(x) \geq 0$ for all $i \in \{1, \ldots, m\}$, then $\nu = 0$. Consequently $(\mu, \lambda_1, \ldots, \lambda_m) \in \mathbb{R}^m$ is an optimal solution of the dual problem (D) with the same optimal value $D^* = P^*$. □

3.3 Semidefinite programming (SDP)

A *semidefinite program* is a program of the form:

$$\text{(3.5)} \qquad \begin{cases} & \text{minimize } \ell(x) \\ \text{over:} & x \in \mathbb{R}^n \\ \text{subjet to:} & L(x) \succeq 0 \end{cases}$$

where $\ell \in \mathbb{R}[\underline{X}]_1$ and $L \in \operatorname{Sym} \mathbb{R}[\underline{X}]_1^{t \times t}$ for $t \in \mathbb{N}_0$. The feasible set of a semidefinite program, i.e:

$$\{x \in \mathbb{R}^n \mid L(x) \succeq 0\}$$

is given by a *linear matrix inequality* and a set of this form is called a *spectrahedron* and it is a closed semialgebraic subset of $\mathbb{R}^n$. Note that a semidefinite program with a diagonal linear matrix inequality is a linear program, consequently semidefinite programs are generalizations of linear programs.

Notation 3.3.1. We write a primal semidefinite program, in standard form, as:

$$(P):\begin{cases} & \text{minimize } \ell(x) \\ \text{over:} & x \in \mathbb{R}^n \\ \text{subjet to:} & L(x) \succeq 0 \end{cases} \tag{3.6}$$

where $\ell \in \mathbb{R}[\underline{X}]_1$ and $L \in \operatorname{Sym} \mathbb{R}[\underline{X}]_1^{t\times t}$ for $t \in \mathbb{N}_0$. The optimal value of (P) i.e., the infimum over all $x \in \mathbb{R}^n$ where x ranges over all feasible solutions of (P) is denoted by $P^* \in \{-\infty\} \cup \mathbb{R} \cup \{\infty\}$.

Notation 3.3.2. We write the dual program of the linear program (P) (3.6) , in standard form, as:

$$(D):\begin{cases} & \text{maximize } \mu \\ \text{over:} & \mu \in \mathbb{R}, A \in \operatorname{Sym} \mathbb{R}^{t\times t} \\ \text{subjet to:} & \ell = \mu + \operatorname{tr}(AL) \text{ and} \\ & A \succeq 0 \end{cases} \tag{3.7}$$

The optimal value of (D) i.e., the maximum over all $(\mu, A) \in \mathbb{R} \times \operatorname{Sym} \mathbb{R}^{t\times t}$, where (μ, A) ranges over all feasible solutions of (D), is denoted by $D^* \in \{-\infty\} \cup \mathbb{R} \cup \{\infty\}$.

Remark 3.3.3. Note that the dual problem (D) is also a semidefinite program. Indeed, by making equal coefficients with the same degree on both side of the equality:

$$\ell = \mu + \operatorname{tr}(AL)$$

we obtain a system of $n+1$ linear equations which can split in $2n+2$ linear inequalities, let us denote the linear polynomials of this inequalities as $r_1, \ldots, r_{2n+2}$, where the unknowns are the entries of the matrix $A \in \operatorname{Sym} \mathbb{R}^{t\times t}$ and therefore its feasible set would be:

$$\left\{ x \in \mathbb{R}^{\frac{t(t+1)}{2}} \mid \left(\begin{array}{ccc|c} r_1(x) & & & \\ & \ddots & & \\ & & r_{2n+2}(x) & \\ \hline & & & A \end{array} \right) \succeq 0 \right\}$$

Remark 3.3.4. The feasible set of a semidefinite program can also be seen as the intersection of the cone positive semidefinite matrices with an affine subspace. In the same way as in linear programming, if we build the dual problem of the dual problem we get the primal problem, what motivates the definition of the dual problem. This is a consequence of the fact that the cone of positive semidefinite matrix of finite dimension is selfdual, that is:

$$(\operatorname{Sym} \mathbb{R}^{t\times t}_{\geq 0})^* := \operatorname{Sym} \mathbb{R}^{t\times t}_{\geq 0}$$

Proposition (weak duality of semidefinite optimization) 3.3.5. Let (P) be a semidefinite program as in (3.6) and (D) its dual as in (3.7). Then:

$$P^* \geq D^*$$

Proof. Let $x_0 \in \mathbb{R}^n$ be a feasible solution of (P) and (μ, A) be a feasible solution for (D). Then it holds:

$$\ell - \mu = \operatorname{tr}(AL)$$

By the trivial implication of Proposition 3.1.18 we have that:

$$\ell - \mu \geq 0 \text{ on } \{x \in \mathbb{R}^n \mid L(x) \succeq 0\}$$

In particular $x_0 \in \{x \in \mathbb{R}^n \mid L(x) \succeq 0\}$, which implies $\ell(x_0) - \mu \geq 0$ and therefore $P^* \geq D^*$. □

Proposition (strong duality of semidefinite optimization) 3.3.6. Let (P) be a semidefinite program as in 3.3.1 and let (D) its dual as in 3.3.2. Suppose that $P^* \in \mathbb{R}$ and that the feasible set of (P) i.e.:

$$\{x \in \mathbb{R}^n \mid L(x) \succeq 0\}$$

has non empty interior. Then (D) possesses an optimal solution and moreover

$$P^* = D^*$$

Proof. We will prove that there exists $A \in \operatorname{Sym} \mathbb{R}^{t \times t}_{\geq 0}$ such that $\ell = P^* + \operatorname{tr}(AL)$ and by weak duality we will be able to conclude the proposition. Since $\ell - P^* \geq 0$ on S, then by the non trivial implication of Proposition 3.1.18 it holds that there exist $a \in \mathbb{R}_{\geq 0}$ and $A \in \operatorname{Sym} \mathbb{R}^{t \times t}_{\geq 0}$ such that $\ell - P^* = a + \operatorname{tr}(AL)$.

$$\inf(\{\ell(x) - P^* \mid x \in S\}) \overset{P^* \in \mathbb{R}}{=} \inf(\{\ell(x) \mid x \in S\}) - P^* = 0$$

From there it follows that:

$$\inf\{a + \operatorname{tr}(AL(x)) \mid x \in S\} = 0$$

Since $a \geq 0$ and $\operatorname{tr}(AL(x)) \geq 0$ for all $x \in S$ it follows that $a = 0$. □

CHAPTER 4

The Lasserre relaxation

In this chapter we outline Lasserre's [16] approach to solve a polynomial optimization problem (P) (2.2). An introduction to this method can be found for example in [17] and in [33] and references therein. In order to understand this method let us give first the following definition.

Definition 4.0.1. Consider the *optimization problem*:

$$(Q) \text{ minimize } f(x) \text{ subject to } x \in S$$

where f is the objective function and S the feasible set, such that $f(S) \subseteq \mathbb{R}$. We say that the following optimization problem (Q_1):

$$(Q_1) \text{ minimize } f_1(y) \text{ subject to } y \in S_1$$

where $f_1(S_1) \subseteq \mathbb{R}$, is a *relaxation* of the optimization problem (Q) if for every $x \in S$ there exists $y \in S_1$ such that $f_1(y) = f(x)$. If we denote the optimal value of (Q) (respectively the optimal value of (Q_1)) i.e., the infimum over all $x \in S$ (respectively the infimum over all $y \in S_1$) by $Q^* \in \{-\infty\} \cup \mathbb{R} \cup \{\infty\}$ (respectively $Q_1^* \in \{-\infty\} \cup \mathbb{R} \cup \{\infty\}$) then obviously it holds:

$$Q_1^* \leq Q^*$$

The idea of building a relaxation of a optimization problem is that one expect that the relaxation is simpler to solve than the original optimization problem in a way that we can get an optimal solution that is not very far from the optimal solution of the original optimization problem. A first very naive idea of a relaxation for the problem (P) (2.2) is to linearize the problem, that is to say, to substitute every monomial $\underline{X}^\alpha$ appearing in (P) for a new variable y_α. This amounts to solve a linear program with the same or more number of variables than the original. However in practice, the lower bound that we get from this linear program is far from the infimum, so the following idea could be to add more redundant constraints before the linearization. Lasserre's method constructs a hierarchy of semidefinite programming relaxations of the problem (P) (2.2). In each relaxation of degree k we build a feasible convex set, obtained through the linearization of an equivalent polynomial optimization problem of (P). The relaxation consists in adding infinitely many redundant inequalites of the form:

$$p \geq 0 \text{ for all } p \in \sum \mathbb{R}[\underline{X}]^2 p_i \cap \mathbb{R}[\underline{X}]_k$$

(with the notation $\sum \mathbb{R}[\underline{X}]^2 p_i$ we mean the set of all finite sums of elements of the form $p^2 p_i$, for $p \in \mathbb{R}[\underline{X}]$). The set of this redundant inequalities builds a convex cone in $\mathbb{R}[\underline{X}]_k$. The cone generated for this redundant inequalities is called *truncated quadratic module generated by the polynomials* $p_1, \ldots, p_m$, as we see in Definition 4.1.1. These relaxations give us an increasing sequence of lower bounds of the infimum P^*, as you can see in Proposition 4.4.4. Lasserre proved that this sequence converge asymptotically to the infimum if we assume some archimedean property of the cone generetated by the redundant inequalities, see Theorem 4.5.4.

4.1 The Moment relaxation with linear forms

Definition 4.1.1. Let $p_1, \ldots, p_m \in \mathbb{R}[\underline{X}]_k$ for $k \in \mathbb{N}_0 \cup \{\infty\}$. We define the k-truncated quadratic module M, generated by $p_1, \ldots, p_m$ as:

$$\begin{aligned} M_k(p_1, \ldots, p_m) := \left(\mathbb{R}[\underline{X}]_k \cap \sum \mathbb{R}[\underline{X}]^2\right) + \left(\mathbb{R}[\underline{X}]_k \cap \sum \mathbb{R}[\underline{X}]^2 p_1\right) \\ + \cdots + \left(\mathbb{R}[\underline{X}]_k \cap \sum \mathbb{R}[\underline{X}]^2 p_m\right) \subseteq \mathbb{R}[\underline{X}]_k \end{aligned} \tag{4.1}$$

where here $\mathbb{R}[\underline{X}]_\infty := \mathbb{R}[\underline{X}]$ and $M(p_1, \ldots, p_m) := M_\infty(p_1, \ldots, p_m)$, to refer to the (smallest) quadratic module generated by the polynomials $p_1, \ldots, p_m \in \mathbb{R}[\underline{X}]$.

Definition 4.1.2. Let (P) be the polynomial optimization problem given in (2.2) and let $k \in \mathbb{N}_0 \cup \{\infty\}$ such that $f, p_1, \ldots, p_m \in \mathbb{R}[\underline{X}]_k$. The Moment relaxation (or Lasserre relaxation) of (P) of degree k is the following program:

$$(P_k): \begin{cases} \text{minimize} & L(f) \\ \text{subject to:} & L \in \mathbb{R}[\underline{X}]_k^* \\ & L(1) = 1 \text{ and} \\ & L(M_k(p_1, ..., p_m)) \subseteq \mathbb{R}_{\geq 0} \end{cases}$$

the optimal value of (P_k) i.e., the infimum over all $L(f)$ where L ranges over all optimal solutions of (P_k) is denoted by $P_k^* \in \{-\infty\} \cup \mathbb{R} \cup \{\infty\}$.

Remark 4.1.3. In the situation of Definition 4.1.2, (P_k) is a semidefinite program as we will see in 4.2.5.

Remark 4.1.4. Note that for $p \in \mathbb{R}[\underline{X}]_k$:

$$\mathbb{R}[\underline{X}]_k \cap \sum \mathbb{R}[\underline{X}]^2 p = \left\{ \sum_{i=1}^{l} h_i^2 p \mid l \in \mathbb{N}_0,\ h_i \in \mathbb{R}[\underline{X}],\ 2\deg(h_i) \leq k - \deg(p) \right\} \tag{4.2}$$

For a proof see [33, Page 5].

Remark 4.1.5. Suppose that we are in the conditions of Definition 4.1.2 and let us set $k := \min\{\deg f, \deg p_1, \ldots, \deg p_m\}$, it could happen:

$$M_{k+2i}(p_1, \ldots, p_m) = M_{k+2i+1}(p_1, \ldots, p_m) \text{ for } i \in \mathbb{N}_0$$

for example if k is even and $\deg p_i$ is even for all $i \in \{1, \ldots, m\}$ as a consequence of 4.2. In that case is $(P_{k+2i}) = (P_{k+2i+1})$ for $i \in \mathbb{N}_0$. However in practical examples, it could be useful to consider a relaxation of odd degree, since we can attain optimality in a relaxation of less degree, and in this way we avoid to work with more number of variables when we apply the linearization. See Example 7.2.2.

4.2 The Moment relaxation with matrices

For $d \in \mathbb{N}_0$ let us define:

$$V_d := (1, X_1, X_2, \ldots, X_n, X_1^2, X_1X_2, \ldots, X_1X_n, X_2^2, X_2X_3, \ldots, X_n^2, \ldots, \ldots, X_n^d)^T \tag{4.3}$$

as a basis for the vector space of polynomials in n variables of degree at most d . Then

$$V_d V_d^T = \begin{pmatrix} 1 & X_1 & X_2 & \cdots & X_n^d \\ X_1 & X_1^2 & X_1X_2 & \cdots & X_1X_n^d \\ X_2 & X_1X_2 & X_2^2 & \cdots & X_2X_n^d \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ X_n^d & X_1X_n^d & X_2X_n^d & \cdots & X_n^{2d} \end{pmatrix} \in \operatorname{Sym} \mathbb{R}[\underline{X}]_{2d}^{s_d \times s_d} \tag{4.4}$$

Let us map every monomial $\underline{X}^\alpha \in \mathbb{R}[\underline{X}]_{2d}$ in the matrix (4.4) to a new variable Y_α. Then the resulting matrix has the following form:

$$M_d := \begin{pmatrix} Y_{(0,\ldots,0)} & Y_{(1,\ldots,0)} & Y_{(0,1,\ldots,0)} & \cdots & Y_{(0,\ldots,1)} \\ Y_{(1,\ldots,0)} & Y_{(2,\ldots,0)} & Y_{(1,1,\ldots,0)} & \cdots & Y_{(1,\ldots,d)} \\ Y_{(0,1,\ldots,0)} & Y_{(1,1,\ldots,0)} & Y_{(0,2,\ldots,0)} & \cdots & Y_{(0,1,\ldots,d)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ Y_{(0,\ldots,d)} & Y_{(1,\ldots,d)} & Y_{(0,1,\ldots,d)} & \cdots & Y_{(0,\ldots,2d)} \end{pmatrix} \in \operatorname{Sym} \mathbb{R}[\underline{Y}]_1^{s_d \times s_d} \tag{4.5}$$

Definition 4.2.1. Every matrix $M \in \mathbb{R}^{s_d \times s_d}$ with the same shape than the matrix (4.5) is called a generalized Hankel matrix of order d. We denote the linear space of generalized Hankel matrix of order d by:

$$H_d := \{\, M_d(y) \mid y \in \mathbb{R}^{s_{2d}} \}$$

Definition and Notation 4.2.2. Let $k \in \mathbb{N}_0$ such that $p \in \mathbb{R}[\underline{X}]_k$ we denote:

$$d_p := \lfloor \frac{k - \deg p}{2} \rfloor$$

and consider the following symmetric matrix:

$$pV_{d_p}^t V_{d_p} = \begin{pmatrix} p & pX_1 & pX_2 & \cdots & pX_n^{d_p} \\ pX_1 & pX_1^2 & pX_1X_2 & \cdots & pX_1X_n^{d_p} \\ pX_2 & pX_2X_1 & pX_2^2 & \cdots & pX_2X_n^{d_p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ pX_n^{d_p} & pX_1X_n^{d_p} & pX_n^{d_p}X_2 & \cdots & pX_n^{2d_p} \end{pmatrix} \in \mathbb{R}[\underline{X}]_k^{s_{d_p}\times s_{d_p}} \tag{4.6}$$

For $p \in \mathbb{R}[\underline{X}]_k$ the *localizing matrix of p of degree k* is the matrix resulting from mapping every monomial $\underline{X}^\alpha$ such that $|\alpha| \leq k$ in (8.8) for a new variable Y_α. We denote this matrix by $M_{k,p} \in \operatorname{Sym}\mathbb{R}[\underline{Y}]_1^{s_{d_p}\times s_{d_p}}$.

Remark 4.2.3. Note that for $p \in \mathbb{R}[\underline{X}]_k$:

$$\mathbb{R}[\underline{X}]_k \cap \sum \mathbb{R}[\underline{X}]^2 p = \left\{ V_{d_p}^T G V_{d_p} p \mid G \in \operatorname{Sym}\mathbb{R}_{\geq 0}^{s_{d_p}\times s_{d_p}} \right\} \tag{4.7}$$

For a proof see [33, Lemma 24].

Lemma 4.2.4. Let $k \in \mathbb{N}$, $p \in \mathbb{R}[\underline{X}]_k \setminus \{0\}$. Let $L \in \mathbb{R}[\underline{X}]_k^*$. Then it holds:

$$L(\sum \mathbb{R}[\underline{X}]_k \cap \mathbb{R}[\underline{X}]^2 p) \subseteq \mathbb{R}_{\geq 0} \iff M_{k,p}(y) \succeq 0$$

where $y_\alpha := L(\underline{X}^\alpha)$ for all $\alpha \in \mathbb{N}^n$ with $|\alpha| \leq k$.

Proof. Let us set $d := \lfloor \frac{k-\deg(p)}{2} \rfloor$ and define the matrices $A_\alpha \in \mathbb{R}^{s_d\times s_d}$ for $|\alpha| \leq k$, as the matrices such that:

$$pV_dV_d^T = \sum_{|\alpha|\leq k} \underline{X}^\alpha A_\alpha \in \mathbb{R}[\underline{X}]_k^{s_d\times s_d}.$$

and $y_\alpha := L(\underline{X}^\alpha)$ for $|\alpha| \leq k$.

$$\begin{aligned} L(\sum \mathbb{R}[\underline{X}]_k \cap \mathbb{R}[\underline{X}]^2 p) \subseteq \mathbb{R}_{\geq 0} &\overset{4.2}{\iff} \forall h \in \mathbb{R}[\underline{X}]_d,\ L(h^2 p) \geq 0 \\ &\iff \forall H \in \mathbb{R}^{s_d},\ L((H^T V_d)(V_d^T H)p) \geq 0 \\ &\iff \forall H \in \mathbb{R}^{s_d},\ L(H^T pV_dV_d^T H) \geq 0 \\ &\iff \forall H \in \mathbb{R}^{s_d},\ L(H^T(\sum_{|\alpha|\leq k} \underline{X}^\alpha A_\alpha)H) \geq 0 \\ &\iff \forall H \in \mathbb{R}^{s_d},\ L(\sum_{|\alpha|\leq k} \underline{X}^\alpha H^T A_\alpha H) \geq 0 \\ &\overset{\text{L is linear}}{\iff} \forall H \in \mathbb{R}^{s_d},\ \sum_{|\alpha|\leq k} L(\underline{X}^\alpha) H^T A_\alpha H \geq 0 \\ &\iff \forall H \in \mathbb{R}^{s_d},\ H^T(\sum_\alpha y_\alpha A_\alpha)H \geq 0 \\ &\iff \sum_{|\alpha|\leq k} y_\alpha A_\alpha \succeq 0 \iff M_{k,p}(y) \succeq 0 \end{aligned}$$

□

Proposition 4.2.5. Let (P) be the polynomial optimization problem given in (2.2) and let $k \in \mathbb{N}_0 \cup \{\infty\}$ such that $f, p_1, \ldots, p_m \in \mathbb{R}[\underline{X}]_k$. The Moment relaxation (or Lasserre relaxation) of (P) of degree k is the following semidefinite program:

$$(P_k): \begin{cases} \text{minimize} & \sum_{|\alpha| \leq k} f_\alpha y_\alpha \\ \text{subject to:} & y_{(0,\ldots,0)} = 1 \\ & M_{k,1}(y) \succeq 0 \text{ and} \\ & M_{k,p_i}(y) \succeq 0 \text{ for all } i \in \{1, \ldots, m\} \end{cases}$$

the optimal value of (P_k) is the infimum over all

$$y = (y_{(0,\ldots,0)}, \ldots, y_{(0,\ldots,k)}) \in \mathbb{R}^{s_k}$$

that ranges over all feasible solutions of (P_k).

Proof. (P_k) is clearly a semidefinite program, see Definition 3.5. On other side, due to Lemma 4.2.4 the following definition of a Moment relaxation using matrices is equivalent to the definition given in 4.1.2 with linear forms. □

Corollary and Notation 4.2.6. Let $d \in \mathbb{N}_0$. The correspondence:

$$\begin{pmatrix} \mathbb{R}[\underline{X}]_{2d} & \to & \mathbb{R} \\ \underline{X}^\alpha & \mapsto & y_\alpha \end{pmatrix} \underset{\longleftarrow\!\shortmid M_d(y)}{\overset{L \mapsto (L(\underline{X}^{\alpha+\beta}))_{|\alpha|,|\beta| \leq d}}{}}$$

defines a bijection between the linear forms $L \in \mathbb{R}[\underline{X}]_{2d}^*$ such that $L(\sum \mathbb{R}[\underline{X}]_d^2]) \subseteq \mathbb{R}_{\geq 0}$ and the set of positive semidefinite generalized Hankel matrices of order d i.e. $H_d \cap \mathbb{R}_{\succeq 0}^{s_d \times s_d}$. Let $L \in \mathbb{R}[\underline{X}]_{2d}^*$ such that $L(\sum \mathbb{R}[\underline{X}]_d^2]) \subseteq \mathbb{R}_{\geq 0}$ we denote:

$$M_L := (L(\underline{X}^{\alpha+\beta}))_{|\alpha|,|\beta| \leq d}$$

and let $M_d(y) \succeq 0$ for $y \in \mathbb{R}^{s_d}$ we denote:

$$L_{M_d(y)} : \mathbb{R}[\underline{X}]_{2d} \longrightarrow \mathbb{R}, \underline{X}^\alpha \mapsto y_\alpha.$$

Proof. The well-definedness of both maps follows from Lemma 4.2.4. To prove that the correspondence is injective, take $L \in \mathbb{R}[\underline{X}]_{2d}^*$ such that $L(\sum \mathbb{R}[\underline{X}]_d^2) \subseteq \mathbb{R}_{\geq 0}$ then:

$$L_{M_L} : \mathbb{R}[\underline{X}]_{2d} \longrightarrow \mathbb{R}, \underline{X}^\alpha \mapsto L(\underline{X}^\alpha)$$

since $M_L = (L(\underline{X}^{\alpha+\beta}))_{|\alpha|,|\beta| \leq d} = M_d(L(0), \ldots, L(X_n^{2d}))$. Then $L_{M_L} = L$. On the other side, in order to prove the correspondence it is surjective, take $y \in \mathbb{R}^{s_d}$ such that $M_d(y) \succeq 0$ then:

$$M_{L_{M_d(y)}} = (L_{M_d(y)}(\underline{X}^{\alpha+\beta}))_{|\alpha|,|\beta| \leq d} = (y_{\alpha+\beta})_{|\alpha|,|\beta| \leq d} = M_d(y).$$

□

Definition 4.2.7. Given $d \in \mathbb{N}_0 \cup \{\infty\}$ and $L \in \mathbb{R}[\underline{X}]_{2d}^*$ we say L is positive semidefinite (respectively positive definite) if $L(p^2) \geq 0$ for all $p \in \mathbb{R}[\underline{X}]_d$ (respectively $L(p^2) > 0$ for all $p \in \mathbb{R}[\underline{X}]_d \backslash \{0\}$).

Remark 4.2.8. Note that by Proposition 4.2.4 we can conclude for $L \in \mathbb{R}[\underline{X}]_{2d}^*$ that: L is positive semidefinite (respectively positive definite) if and only if $M_L \succeq 0$ (respectively if and only if $M_L \succ 0$).

Notation 4.2.9. We will consider the following isomorphism of vector spaces:

$$\text{poly} : \mathbb{R}^{s_d} \longrightarrow \mathbb{R}[\underline{X}]_d, a \mapsto a^T V_d$$

Proposition 4.2.10. Let $d \in \mathbb{N}_0$ and $L \in \mathbb{R}[\underline{X}]_{2d}^*$ then:

$$L(pq) = P^T M_L Q \text{ for all } p, q \in \mathbb{R}[\underline{X}]_d$$

where $P := \text{poly}^{-1}(p)$ and $Q := \text{poly}^{-1}(q)$.

Proof. As usual let us set the matrices $A_\alpha \in \mathbb{R}^{s_d \times s_d}$ for $|\alpha| \leq 2d$ as the matrices such that:

$$V_d V_d^T = \sum_{|\alpha| \leq k} \underline{X}^\alpha A_\alpha \in \mathbb{R}[\underline{X}]_k^{s_d \times s_d}.$$

Then using the lineary of L it holds:

$$L(pq) = L(P^T V_d V_d^T Q) = L(P^T \sum_{|\alpha| \leq 2d} \underline{X}^\alpha A_\alpha Q) =$$

$$\sum_{|\alpha| \leq 2d} L(\underline{X}^\alpha) P^T A_\alpha Q = P^T (\sum_{|\alpha| \leq 2d} L(\underline{X}^\alpha) A_\alpha) Q = P^T M_L Q$$

□

4.3 Sums of squares (SOS) relaxations

In this section, we outline the sums of squares approach to solve the polynomial optimization problem (P) (2.2). For this note, that (P) admits this equivalent formulation:

$$(P) \text{ maximize } \mu \text{ subject to } \mu \in \mathbb{R} \text{ and } f(x) - \mu \geq 0 \text{ for all } x \in S$$

It is well known that to check nonnegative is a hard problem, but we can relax this condition by try to find a sums of squares representation or try to find a representation of $f(x) - \mu$ in the truncated quadratic module, which is much easier task and amounts to solve a semidefinite program.

Definition 4.3.1. Let (P) be the polynomial optimization problem given in (2.2) and let $k \in \mathbb{N}_0 \cup \{\infty\}$ such that $f, p_1, \ldots, p_m \in \mathbb{R}[\underline{X}]_k$. The *sums of squares relaxation* of the polynomial optimization problem (P) defined in (2.2) , is the following program:

$$(D_k): \begin{cases} \text{maximize} & \mu \\ \text{subject to:} & \mu \in \mathbb{R}, \text{ and} \\ & f - \mu \in M_k(p_1, \dots, p_m) \end{cases}$$

The optimal value of (D_k) i.e., the maximum over all $\mu \in \mathbb{R}$, where $\mu \in \mathbb{R}$ ranges in all feasible solutions of (D_k), is denoted by $D_k^* \in \{-\infty\} \cup \mathbb{R} \cup \{\infty\}$.

Remark 4.3.2. Suppose we are in the same conditions as in Definition 4.3.1. (D_k) is the dual of the moment relaxation (P_k) 4.2.5, and hence a semidefinite program. Indeed, to find the maximum $\mu \in \mathbb{R}$ such that:

$$f - \mu \in M_k(p_1, \dots, p_m)$$

is accordingly to Remark 4.7 and the definition of quadratic module 4.1.1, equivalent to find the maximum $\mu \in \mathbb{R}$ and symmetric matrices $G_i \in \operatorname{Sym} \mathbb{R}_{\geq 0}^{s_{d_{p_i}} \times s_{d_{p_i}}}$ for all $i \in \{0, \dots, m\}$ such that:

$$f - \mu = \sum_{i=0}^{m} V_{d_{p_i}}^t G_i V_{d_{p_i}} p_i \in \mathbb{R}[\underline{X}]_k \tag{4.8}$$

where we denote $p_0 := 1$. Let us set the matrices $A_{\alpha,i} \in \mathbb{R}_{\geq 0}^{s_{d_{p_i}} \times s_{d_{p_i}}}$ for all $|\alpha| \leq k$ and $i \in \{0, \dots, m\}$ such that:

$$p_i V_{d_{p_i}} V_{d_{p_i}}^T = \sum_{|\alpha| \leq k} \underline{X}^\alpha A_{\alpha,i}$$

The equation (4.8), can be rewritten as:

$$f - \mu = \sum_{i=0}^{m} \operatorname{tr}(V_{d_{p_i}}^t G_i V_{d_{p_i}} p_i) = \sum_{i=0}^{m} \operatorname{tr}(G_i p_i V_{d_{p_i}} V_{d_{p_i}}^t) = \sum_{i=0}^{m} \operatorname{tr}(G_i \sum_{|\alpha| \leq k} \underline{X}^\alpha A_{\alpha,i}) \in \mathbb{R}[\underline{X}]_k$$

To make the condition:

$$f - \mu = \sum_{i=0}^{m} \operatorname{tr}(G_i \sum_{|\alpha| \leq k} \underline{X}^\alpha A_{\alpha,i}) \in \mathbb{R}[\underline{X}]_k \tag{4.9}$$

holds for some $\mu \in \mathbb{R}$ and some matrices $G_i \in \operatorname{Sym} \mathbb{R}_{\geq 0}^{s_{d_{p_i}} \times s_{d_{p_i}}}$, we can build a linear system by making equal coefficients with the same degree on both sides of the equality. That is to say, if we substitute every monomial in (4.9) $\underline{X}^\alpha$ for a new variable Y_α, we get the equivalent linear condition to (4.9):

$$\sum_{\alpha} f_\alpha Y_\alpha - \mu = \sum_{i=0}^{m} \operatorname{tr}(G_i \sum_{|\alpha| \leq k} Y_\alpha A_{\alpha,i}) \in \mathbb{R}[\underline{Y}]_1$$

Taking into account that by construction $\sum_{|\alpha|\leq k} Y_\alpha A_{\alpha,i} = M_{k,p_i}$ for all $i \in \{1,\ldots,m\}$ finally we get that the sums squares relaxation (D_k) defined in 4.3.1 is equivalent to the semidefinite program:

$$(D_k) : \begin{cases} \text{maximize} & \mu \\ \text{subject to:} & \mu \in \mathbb{R}, \\ & G_i \in \operatorname{Sym} \mathbb{R}_{\geq 0}^{s_{d_{p_i}} \times s_{d_{p_i}}} \text{ for } i = 0,\ldots,m \text{ and} \\ & \sum_\alpha f_\alpha Y_\alpha - \mu = \sum_{i=0}^m \operatorname{tr}(G_i M_{k,p_i}) \end{cases}$$

One can see that this program is the dual of the moment relaxation of degree k (P_k), formulated in 4.2.5.

4.4 Properties of Lasserre relaxation

Proposition 4.4.1. Let (P) be the polynomial optimization problem defined in (2.2), then it holds:

$$P^* \geq P_k^* \geq D_k^*$$

for all $k \in \mathbb{N}_0 \cup \{\infty\}$.

Proof. $P^* \geq P_k^*$ since if $x \in S$ then $\operatorname{ev}_x \in \mathbb{R}[\underline{X}]_k^*$ is a feasible solution of (P_k). The last inequality is consequence of the duality of semidefinite progamming 3.3.5, since as we have seen in Remark 4.3.2 (D_k) is the dual program of (P_k). However, one can prove it directly very easily. Indeed, let $L \in \mathbb{R}[\underline{X}]_k^*$ be a feasible solution of (P_k) and $\mu \in \mathbb{R}$ such that $f - \mu \in M_k(p_1,\ldots,p_m)$ then:

$$0 \leq L(f-\mu) = L(f) - \mu$$

□

Theorem 4.4.2. Let (P) be the polynomial optimization problem defined in (2.2). Suppose that S has nonempty interior. Then $M_k(p_1,\ldots,p_m)$ is closed in the vector space $\mathbb{R}[\underline{X}]_k$. Moreover for all $f \in \mathbb{R}[\underline{X}]$ the following statements are equivalent:

(i) $L(f) \geq 0$ for all $L \in \mathbb{R}[\underline{X}]_k^*$ with $L(M_k(p_1,\ldots,p_m)) \subseteq \mathbb{R}_{\geq 0}$ and $L(1) = 1$.

(ii) $f \in M_k(p_1,\ldots,p_m)$.

Proof. Different proofs of the fact that $M_k(p_1,\ldots,p_m)$ is closed providing that S has non empty interior, can be found. Between others, we refer the reader to [27, Proposition 3.7], [33, Theorem 19] and [35, Proposition 3.5]. For the proof $(i) \implies (ii)$ we will use contraposition. Suppose that $f \notin M_k(p_1,\ldots,p_m)$. Since $M_k(p_1,\ldots,p_m)$ is closed we can use the separation Theorem for closed convex cones 3.1.13. Hence there exists a linear form $L \in \mathbb{R}[\underline{X}]_k^*$ such that:

$$\begin{cases} L(M_k(p_1,\ldots,p_m)) \subseteq \mathbb{R}_{\geq 0} \\ \text{and } L(f) < 0 \end{cases}$$

In case $L(1) \neq 0$ then $L(1) > 0$, and we can consider the linear form:

$$\hat{L} := \frac{1}{L(1)} L \in \mathbb{R}[\underline{X}]_k^*.$$

In case $L(1) = 0$, take x in the interior of S and consider the linear form:

$$\hat{L} := \frac{L + \epsilon\, \mathrm{ev}_x}{L(1) + \epsilon} \in \mathbb{R}[\underline{X}]_k^*$$

where $\epsilon > 0$ is taken sufficiently small such that $L(f) + \epsilon f(x) < 0$. In both cases we have found $\hat{L} \in \mathbb{R}[\underline{X}]_k^*$ such that:

$$\begin{cases} \hat{L}(M_k(p_1, \dots, p_m)) \subseteq \mathbb{R}_{\geq 0} \\ \hat{L}(f) < 0 \text{ and} \\ \hat{L}(1) = 1 \end{cases}$$

$(ii) \implies (i)$ is trivial.

□

Corollary 4.4.3. Let (P) be the polynomial optimization problem defined in (2.2). Suppose that S has nonempty interior. Then it holds $P_k^* = D_k^*$, and in case $P_k^* = D_k^* \neq \infty$ then D_k^* has an optimal solution.

Proof. By definition of infimum, $0 \leq L(f) - P_k^*$ for all $L \in \mathbb{R}[\underline{X}]_k^*$, feasible solution of (P_k). Then by Theorem 4.4.2 we have that $f - P_k^* \in M_k(p_1, \dots, p_m)$. This last fact together with weak duality yields to $P_k^* = D_k^*$. If moreover $P_k^* = D_k^* \neq \infty$ then we can take $\mu := P_k^*$ as a optimal solution of (D_k), defined in 4.3.1. □

Proposition 4.4.4. Let (P) be the polynomial optimization problem given in (2.2) with $f, p_1, \dots, p_m \in \mathbb{R}[\underline{X}]_k$. Then the following holds:

(i)

$$P^* \geq P_\infty^* \geq \cdots \geq P_{k+1}^* \geq P_k^* \text{ and}$$
$$P^* \geq D_\infty^* \geq \cdots \geq D_{k+1}^* \geq D_k^*$$

(ii) Let $L \in \mathbb{R}[\underline{X}]_k^*$ with $L(1) = 1$. Suppose L has a quadrature rule with nodes in S, then L is a feasible solution of (P_k) with $L(f) \geq P^*$.

(iii) Suppose (P_k) has an optimal solution L^*, which has a quadrature rule on $\mathbb{R}[\underline{X}]_l$ for some $l \in \{1, \dots, k\}$ with $f \in \mathbb{R}[\underline{X}]_l$ and with nodes in S. Then $L^*(f) = P^*$, moreover we have $P^* = P_{k+m}^*$ for $m \geq 0$ and the nodes of the quadrature rule are global minimizers of (P).

(iv) In the situation of (iii), suppose moreover that (P) has a unique global minimizer x^*, then $L^*(f) = f(x^*)$ and $x^* = (L^*(X_1), \dots, L^*(X_n))$.

Proof. (i) $P^* \geq P^*_\infty$ since if x is a feasible solution for (P) then $\mathrm{ev}_x \in \mathbb{R}[\underline{X}]^*$ is a feasible solution of P_∞ with the same value, that is $f(x) = \mathrm{ev}_x(f)$. It remains to prove $P^*_l \geq P^*_k$ for $l \in \mathbb{N}_{\geq k} \cup \{\infty\}$. For this, let $L \in \mathbb{R}[\underline{X}]^*_l$ be a feasible solution of (P_l). Because of Remark 4.2:

$$M_k(p_1, \ldots, p_m) \subseteq M_l(p_1, \ldots, p_m) \tag{4.10}$$

then $L_{|\mathbb{R}[\underline{X}]_k}$ is a feasible solution of (P_k) with the same value at the polynomial f. On other side, in Proposition 4.4.1 we already proved weak duality i.e, $P^*_\infty \geq D^*_\infty$ together with the inequality $P^* \geq P^*_\infty$ implies $P^* \geq D^*_\infty$. It remains to prove $D^*_l \geq D^*_k$ for $l \in \mathbb{N}_{\geq k} \cup \{\infty\}$. For this let $\mu \in \mathbb{R}$ such that $f - \mu \in M_k(p_1, \ldots, p_m)$. Since the inclusion 4.10 holds, then $f - \mu \in M_l(p_1, \ldots, p_m)$ and hence $D^*_l \geq D^*_k$.
(ii) Suppose L has a quadrature rule with nodes $a_1, \ldots, a_N \in S$ and with positive weights $\lambda_1 > 0, \ldots, \lambda_N > 0$. From $L(1) = 1$ we get $\sum_{i=1}^N \lambda_i = 1$ and since the nodes are in S it holds $L(M_k(p_1, \ldots, p_m)) \subseteq \mathbb{R}_{\geq 0}$. Hence L is a feasible solution of (P_k). Moreover the following holds:

$$P^* = L(1)P^* = \sum_{i=1}^N \lambda_i P^* \leq \sum_{i=1}^N \lambda_i f(a_i) = L(f)$$

where the above inequality follows from the fact that $P^* \leq f(x)$ for all $x \in S$.
(iii) Suppose L^* is an optimal solution of (P_k) then $L^*(f) = P^*_k \leq P^*$ using (i) and on the other side since $L^*(1) = 1$ and L^* has a quadrature rule on $\mathbb{R}[\underline{X}]_l$ with nodes in S and $f \in \mathbb{R}[\underline{X}]_l$, there exist $a_1, \ldots, a_N \in S$ nodes, and $\lambda_1 > 0, \ldots, \lambda_N > 0$ weights, such that:

$$P^*_k = L^*(f) = \sum_{i=1}^N \lambda_i f(a_i) \geq \sum_{i=1}^N \lambda_i P^* = P^* \tag{4.11}$$

Therefore $L^*(f) = P^*$, and since $P^*_k = P^*$ we get equality everywhere in (i) and we can conclude that $P^* = P^*_{k+m}$ for $m \geq 0$. It remains to show that the nodes are global minimimizers of (P), but this is true since in (4.11) we have equality everywhere, and if we factor out we get $\sum_{i=1}^N \lambda_i(f(a_i) - P^*) = 0$, as $\lambda_i > 0$ and $f(a_i) - P^* \geq 0$ for all $i \in \{1, \ldots, N\}$, implying $f(a_i) = P^*$ for all $i \in \{1, \ldots, N\}$.

(iv) Using (iii) we have that $L^*(f) = P^* = f(x^*)$, and continuing with the same notation as in the proof of (iii) we got by uniqueness of the minimizer x^*, that $a_i = x^*$ for all $i \in \{1, \ldots, N\}$. This implies that $L^* = \mathrm{ev}_{x^*}$ on $\mathbb{R}[\underline{X}]^*_l$, and evaluating in the polynomials $X_1, \ldots, X_N \in \mathbb{R}[\underline{X}]_1$ we got that:

$$L^*(X_i) = \mathrm{ev}_{x^*}(X_i) = x^*_i \text{ for all } i \in \{1, \ldots, N\}.$$

That is to say, $x^* = (L^*(X_1), \ldots, L^*(X_n))$. □

4.5 Asymptotic convergence

Definition 4.5.1. Let $M \subseteq \mathbb{R}[\underline{X}]$ a quadratic module. M is called *archimedean* if 1 is a unit of M.

Remark 4.5.2. Let $p_1, \ldots, p_m \in \mathbb{R}[\underline{X}]$ and suppose that $M := M(p_1, \ldots, p_m) \subseteq \mathbb{R}[\underline{X}]$ is an archimedean quadratic module then the basic closed semialgebraic set:

$$\{x \in \mathbb{R}^n \mid p_1(x) \geq 0, \ldots, p_m(x) \geq 0\}$$

is compact. Indeed, since $M(p_1, \ldots, p_m)$ is archimedean there exits $N \in \mathbb{N}_0$ such that $g := N - X_1^2 + \ldots + X_n^2 \in M(p_1, \ldots, p_m)$. And then:

$$\{x \in \mathbb{R}^n \mid p_1(x) \geq 0, \ldots, p_m(x) \geq 0\} =$$
$$\{x \in \mathbb{R}^n \mid p(x) \geq 0 \text{ for all } p \in M\} \subseteq \{x \in \mathbb{R}^n \mid g(x) \geq 0\}$$

Theorem of Putinar (1993) 4.5.3. Let $M \subseteq \mathbb{R}[\underline{X}]$ an archimedean quadratic module and consider the set:

$$S_M := \{x \in \mathbb{R}^n \mid p(x) \geq 0 \text{ for all } p \in M\}$$

and let $f \in \mathbb{R}[\underline{X}]$. Suppose $f > 0$ on S then it holds $f \in M$.

Proof. The original proof can be found in [28, Theorem 1.3]. For an elementary proof see [33, Page 8]. □

The following Theorem 4.5.4, is a corollary of Theorem 4.5.3. The Theorem 4.5.4 shows the asymptotic convergence of the hierarchy of moment relaxations 4.3.1 and sums of squares relaxations 4.1.2 to the minimum P^*, providing that the quadratic module generated by the polynomials $p_1, \ldots, p_m \in \mathbb{R}[\underline{X}]$ defining the basic closed semialgebraic set in our polynomial optimization problem (P) 2.2 is archimedean.

Theorem of Lasserre (2001) 4.5.4. Let (P) be the polynomial optimization problem defined in (2.2). Suppose that the quadratic module $M(p_1, \ldots, p_m)$ is archimedean and $S \neq \emptyset$. Then $(P_k^*)_{k \geq \mathscr{M}}$ and $(D_k^*)_{k \geq \mathscr{M}}$, where $\mathscr{M} := \max\{\deg f, \deg p_1, \ldots, \deg p_m\}$, are monotonous increasing sequences that converge to P^*.

Proof. We have already proved in Proposition 4.4.4 (i) that $(P_k^*)_{k \geq \mathscr{M}}$ and $(D_k^*)_{k \geq \mathscr{M}}$ are monotonous increasing sequences. Now, following the notation of the Theorem 4.5.3, notice that since $M(p_1, \ldots, p_m)$ is archimedean then $S_{M(p_1, \ldots, p_m)}$ is compact accordingly to Remark 4.5.2. Therefore P^* is attained for some $x \in S^*$. For every $\epsilon > 0$ we have that $f - P^* + \epsilon > 0$ on S_M. Hence by Putinar Positivstellensatz 4.5.3 there exists $k_\epsilon \in \mathbb{N}_0$ such that $f - P^* + \epsilon \in M_{k_\epsilon}(p_1, \ldots, p_m)$. Therefore $P^* - \epsilon$ is a feasible solution for (D_{k_ϵ}), defined in 4.3.1, and then $P^* - \epsilon \leq D_{k_\epsilon}^*$. Making ϵ tends to 0 and using the inequalities $P^* \geq P_\infty^* \geq D_\infty^*$, proved in Proposition 4.4.4 (i) and in Proposition 4.4.1, we can conclude that both sequences converge to P^*. □

CHAPTER 5

Reformulation of the problem

The aim of this work is finding optimality conditions in an optimal solution of the moment relaxation of the problem (P)(2.2). In other words, we look for *conditions* on an optimal solution of the moment relaxation of certain degree $2d+2$ of the problem (P), let us write it as $L \in \mathbb{R}[\underline{X}]^*_{2d+2}$, such that:

$$L(f) = P^*$$

Moreover we are interested in finding minimizers from this optimal solution. As we saw in Proposition 4.4.4 (iii), this reduces to find a quadrature rule representation for this linear form on some space $U \subseteq \mathbb{R}[\underline{X}]_{2d+2}$, where the nodes will be the minimizers in case they are contained in the basic closed semialgebraic set S. We can now reformulate our problem as:

Given $d \in \mathbb{N}_0$ and $L \in \mathbb{R}[\underline{X}]^*_{2d+2}$ such that $L(\sum \mathbb{R}[\underline{X}]^2_{d+1}) \subseteq \mathbb{R}_{\geq 0}$, we would like to obtain for all $p \in U$:

- Nodes $x_1, \ldots, x_r \in \mathbb{R}^n$ and and weights $\lambda_1 > 0, \ldots, \lambda_r > 0$ such that:

$$L(p) = \sum_{i=1}^{r} \lambda_i p(x_i)$$

in other words:

- $x_{1,1}, \ldots, x_{1,n}, \ldots, x_{r,1}, \ldots, x_{r,n} \in \mathbb{R}$ and $a_1, \ldots, a_r \in \mathbb{R}$ such that:

$$L(p) = \sum_{i=1}^{r} a_i^2 p(x_{i,1}, \ldots, x_{i,n})$$

again with other words:

- $x_{1,1}, \ldots, x_{1,n}, \ldots, x_{r,1}, \ldots, x_{r,n} \in \mathbb{R}$ and $a_1, \ldots, a_r \in \mathbb{R}$ such that:

$$L(p) = \left\langle \begin{pmatrix} p(x_{1,1}, \ldots, x_{1,n}) & & \\ & \ddots & \\ & & p(x_{r,1}, \ldots, x_{r,n}) \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_r \end{pmatrix}, \begin{pmatrix} a_1 \\ \vdots \\ a_r \end{pmatrix} \right\rangle$$

again written differently:

- $x_{1,1}, \dots, x_{1,n}, \dots, x_{r,1}, \dots, x_{r,n} \in \mathbb{R}$ and $a \in \mathbb{R}^r$ such that:

$$L(p) = \left\langle p \left[\begin{pmatrix} x_{1,1} & & \\ & \ddots & \\ & & x_{r,1} \end{pmatrix}, \dots, \begin{pmatrix} x_{1,n} & & \\ & \ddots & \\ & & x_{r,n} \end{pmatrix} \right] \begin{pmatrix} a_1 \\ \vdots \\ a_r \end{pmatrix}, \begin{pmatrix} a_1 \\ \vdots \\ a_r \end{pmatrix} \right\rangle$$

again with less words:

- Diagonal matrices $D_1, \dots, D_n \in \mathbb{R}^{r \times r}$ and $a \in \mathbb{R}^r$ such that:

$$L(p) = \langle p(D_1, \dots, D_n)a, a \rangle$$

Reminder 5.0.1. Let $r, n \in \mathbb{N}$ and $M_1, \dots, M_n \in \mathbb{R}^{r \times r}$ symmetric commuting matrices. Then there exist an orthogonal matrix $P \in \mathbb{R}^{r \times r}$ such that $P^T M_i P$ is a diagonal matrix for all $i \in \{1, \dots, n\}$.

Using the Theorem 5.0.1 we can continue with our reformulation of the problem:

Given $d \in \mathbb{N}_0$ and $L \in \mathbb{R}[\underline{X}]^*_{2d+2}$ such that $L(\sum \mathbb{R}[\underline{X}]^2_{d+1}) \subseteq \mathbb{R}_{\geq 0}$, we would like to find for all $p \in U \subseteq \mathbb{R}[\underline{X}]_{2d+2}$:

- Commuting symmetric matrices $M_1, \dots, M_n \in \mathbb{R}^{r \times r}$ and a vector $a \in \mathbb{R}^r$ such that:

$$L(p) = \langle p(M_1, \dots, M_n)a, a \rangle \tag{5.1}$$

We end the reformulation of the problem once and for all with the language of endomorphisms, instead of matrices. That is to say: given $d \in \mathbb{N}_0$ and $L \in \mathbb{R}[\underline{X}]^*_{2d+2}$ such that $L(\sum \mathbb{R}[\underline{X}]^2_{d+1}) \subseteq \mathbb{R}_{\geq 0}$, we would like to find for all $p \in U \subseteq \mathbb{R}[\underline{X}]_{2d+2}$:

- A finite dimensional euclidean vector space V, commuting self-adjoint endomorphisms $M_1, \dots, M_n$ of V and $a \in V$ such that:

$$L(p) = \langle p(M_1, \dots, M_n)a, a \rangle \tag{5.2}$$

Remark 5.0.2. Gelfand, Naimark and Segal gave a solution in the case that the vector space V is infinite dimensional and the linear form is stricly positive in the sums of squares, that is to say, in the case we are given a linear form $L \in \mathbb{R}[\underline{X}]^*$ such that $L(p^2) > 0$ for all $p \neq 0$. The solution was given by defining the inner product:

$$\langle p, q \rangle := L(pq) \tag{5.3}$$

and defining the self adjoint operators M_i, for all $i \in \{1, \dots, n\}$, on the infinite dimensional vector space $\mathbb{R}[\underline{X}]$, in the following way:

$$M_i : \mathbb{R}[\underline{X}] \longrightarrow \mathbb{R}[\underline{X}], p \mapsto X_i p$$

Taking $a := 1 \in \mathbb{R}[\underline{X}]$ we have the searched equality (5.2). For more information about this construction we refer the reader to [14] and [21].

Remark 5.0.3. Note that in the reformulation of the problem in case $S \neq \mathbb{R}^n$ we will consider a linear form which is a feasible solution of a Lasserre relaxation, therefore it will hold:

$$L(M_{2d+2}(p_1, \dots, p_m)) \subseteq \mathbb{R}_{\geq 0}$$

This inclusion could increases in principle the likelihood that the nodes of the quadrature rule are contained in S.

For $d \in \mathbb{N}_0$, finding necessary and sufficient conditions for a linear form $L \in \mathbb{R}[\underline{X}]^*_{2d+2}$ to have a quadrature rule, it is not only useful in finding optimality in polynomial optimization but also in at least two other different topics: in *real algebraic geometry* and in the *truncated moment problem*. Despite that this work is concerned in finding optimality conditions in polynomial optimization, for convenience of the reader, we include a few ideas of the relation with other topics.

On one side, *real algebraic geometry* studies the sets in $\mathbb{R}^n$ defined by polynomial inequalities, and one of the questions that comes out naturally is try to find conditions which guarantee us that a polynomial $p \in \mathbb{R}[\underline{X}]_{2d+2}$ such that $p(x) \geq 0$ for all $x \in S$ belongs to $M_{2d+2}(p_1, \dots, p_m)$. A very easy example where one can see the connection between this question and the existence of quadrature rule is the proof of Theorem 3.1.17. In this theorem we could prove that in case the polynomials $p, p_1, \dots, p_m$ are linear and $S \neq \emptyset$ $p \in M_1(p_1, \dots, p_m)$, by using the existence of a quadrature rule on S and the separation theorem for closed convex cones. Another example that shows the relation between the existence of quadrature rules with real algebraic geometry is the Corollary 5.0.5, where we can deduce a part[1] of Hilbert's result of 1888 with the use of a recent Theorem of Curto and Yoo 5.0.4 together with the separation theorem for closed cones.

Curto and Yoo Theorem (2016) 5.0.4. Let $L \in \mathbb{R}[X_1, X_2]^*_4$ such that $L(\sum \mathbb{R}[X_1, X_2]^2_2) \subseteq \mathbb{R}_{\geq 0}$ then L has a quadrature rule with at most 6 nodes.

Proof. For the proof we refer the reader to the paper [7, Theorem 2.1] for the non singular case, that is $L(\sum \mathbb{R}[X_1, X_2]^2_2) \subseteq \mathbb{R}_{>0}$, and the paper [6] for the singular case. □

Corollary 5.0.5. Let $p \in \mathbb{R}[X_1, X_2]_4$ such that $p(x) \geq 0$ for all $x \in \mathbb{R}^2$ then it holds $p \in \sum \mathbb{R}[X_1, X_2]^2_2$.

Proof. We will prove it by contraposition. Suppose $p \notin \sum \mathbb{R}[X_1, X_2]^2_2$. Since $\sum \mathbb{R}[X_1, X_2]^2_2$ is a closed cone, by the separation theorem for closed cones in a finite dimensional vector space 3.1.13 there exists $L \in \mathbb{R}[X_1, X_2]^*_4$ such that:

$$\begin{cases} L(\sum \mathbb{R}[X_1, X_2]^2_2) \subseteq \mathbb{R}_{\geq 0} \\ L(p) < 0 \end{cases}$$

now by the Theorem of Curto and Yoo 5.0.4 L has a quadrature rule, and therefore there exists a node $x \in \mathbb{R}^2$ such that $p(x) < 0$. □

[1] In 1888 Hilbert showed that a polynomial of two variables of degree 4 is always a sum of three squares, and his main effort went into show that three squares suffice.

On the other side, let $K \subseteq \mathbb{R}^n$ be a closed set, the *K-truncated moment poblem* is the following question: Let $N \subseteq \mathbb{N}_0^n$ be finite set and A be the real linear span of the monomials $\underline{X}^\alpha$ such that $\alpha \in N$ and let $L \in A^*$, does there exist a positive Borel measure μ on R^n with supp $\mu \subseteq K$ such that:

$$L(\underline{X}^\alpha) = \int \underline{X}^\alpha d\mu?$$

The existence of quadrature rule connect directly with the truncated moment problem through the following Theorem of Bayer and Teichmann 5.0.6.

Theorem 5.0.6. Let $d \in \mathbb{N}_0$ and $L \in \mathbb{R}[\underline{X}]_d^*$. Then L is integration with respect to a positive Borel measure μ on $\mathbb{R}^n$, i.e:

$$L(p) = \int p \, d\mu \text{ for all } p \in \mathbb{R}[\underline{X}]_d$$

if and only if there exists a quadrature rule for L.

Proof. For a proof of this Theorem see [2, Corollary 1] or [17, Theorem 5.9, Remark 5.10]. □

CHAPTER 6

The truncated GNS-construction

Throughout this chapter will assume we are given a linear form:

$$L \in \mathbb{R}[\underline{X}]^*_{2d+2} \text{ with } d \in \mathbb{N}_0 \cup \{\infty\}$$
$$\text{such that: } L(\sum \mathbb{R}[\underline{X}]^2_{d+1}) \subseteq \mathbb{R}_{\geq 0}$$

unless L is defined explicitly in another way.

In this chapter, motivated by the Gelfand-Neimark-Segal construction 5.0.2, we will explain how we can define the *finite* dimensional euclidean vector space V and the *truncated* multiplication operators associated to the positive semidefinite linear form L, in a way such that the equation (5.2) holds.

Since L is positive semidefinite, $L(p^2) = 0$ does not imply $p = 0$ for $p \in \mathbb{R}[\underline{X}]_{d+1}$ in general. In other words (5.3) does not define an inner product if L is not positive definite. By grouping together the polynomials with this property we will be able to define an inner product, on a quotient space. As a consequence, we will obtain a finite dimensional euclidean vector space. To define the *truncated* multiplication operators, we will need to do the orthogonal projection with respect to this euclidean product on the class of polynomials with one degree less, in such a way that when we do the multiplication for the variable X_i we are not out of our ambient space. This construction was already done in [29].

6.1 Introduction in the truncated GNS-construction

Definition and Notation 6.1.1. We define and denote the *truncated GNS kernel* of L:

$$U_L := \{p \in \mathbb{R}[\underline{X}]_{d+1} \mid L(pq) = 0 \text{ for all } q \in \mathbb{R}[\underline{X}]_{d+1}\}$$

Proposition 6.1.2. The truncated GNS kernel of L is a vector subspace in $\mathbb{R}[\underline{X}]_{d+1}$. Moreover:

$$U_L = \{p \in \mathbb{R}[\underline{X}]_{d+1} \mid L(p^2) = 0\} \tag{6.1}$$

Proof. The fact that U_L is a vector subspace follows directly from the linearity of L. Let us prove the equality (6.1). For this let us denote $A := \{p \in \mathbb{R}[\underline{X}]_{d+1} \mid L(p^2) = 0\}$. The inclusion $U_L \subseteq A$ is trivial. For the other inclusion we will demonstrate first, due to L is positive semidefinite and linear, that the Cauchy-Schwarz inequality holds:

$$L(pq)^2 \leq L(p^2)L(q^2) \tag{6.2}$$

Indeed, for all $t \in \mathbb{R}$ and $p, q \in \mathbb{R}[\underline{X}]_{d+1}$ we have:

$$0 \leq L((p+tq)^2) = L(p^2) + 2tL(pq) + t^2L(q^2)$$

Therefore the polynomial $r := L(p^2) + 2XL(pq) + X^2L(q^2) \in \mathbb{R}[X]_2$ is non negative, i.e. $r(x) \geq 0$ for all $x \in \mathbb{R}$. In the case $L(q^2) \neq 0$, the discriminant of r has to be less or equal to zero i.e., $4L(pq)^2 - 4L(p^2)L(q^2) \leq 0$ and we get the desired inequality (6.2). In the case $L(q^2) = 0$, then $L(pq) = 0$ and trivially we get also the inequality (6.2). As a consequence if $p \in A$ then $L(p^2) = 0$, and this implies due to (6.2) that $L(pq) = 0$ for all $q \in \mathbb{R}[\underline{X}]_d$, and therefore $p \in U_L$. □

Definition, Notation and Proposition 6.1.3. We define and denote the *GNS representation space* of L, as the following quotient of vector spaces:

$$V_L := \frac{\mathbb{R}[\underline{X}]_{d+1}}{U_L} \tag{6.3}$$

For every $p \in \mathbb{R}[\underline{X}]_{d+1}$ we will write $\overline{p}^L$ to refer to the class of p in V_L. We define and denote the *GNS inner product* of L, in the following way:

$$\langle \overline{p}^L, \overline{q}^L \rangle_L := L(pq) \tag{6.4}$$

for every $p, q \in \mathbb{R}[\underline{X}]_{d+1}$. $(V_L, \langle\, .\, ,\, .\, \rangle_L)$, defines an euclidean vector space.

Proof. Let us prove first that $\langle\, .\, ,\, .\, \rangle_L$ is well defined. To do this take $p_1, q_1, p_2, q_2 \in \mathbb{R}[\underline{X}]_{d+1}$ with $\overline{p_1}^L = \overline{p_2}^L$ and $\overline{q_1}^L = \overline{q_2}^L$ then:

$$\begin{aligned}
\langle \overline{p_1}^L, \overline{q_1}^L \rangle_L = \langle \overline{p_2}^L, \overline{q_2}^L \rangle_L &\iff L(p_1q_1) = L(p_2q_2) \iff L(p_1q_1) - L(p_2q_2) = 0 \\
&\iff L(p_1q_1) + L(-p_2q_1) - L(-p_2q_1) - L(p_2q_2) = 0 \\
&\iff L((p_1 - p_2)q_1) - L(p_2(q_2 - q_1)) = 0
\end{aligned}$$

The last equality holds since $p_1 - p_2, q_2 - q_1 \in U_L$. The bilinearity and symmetry is trivial. $\langle\, .\, ,\, .\, \rangle_L$ is positive semidefinite since $L(\sum \mathbb{R}[\underline{X}]_{d+1}^2) \subseteq \mathbb{R}_{\geq 0}$. It remains to prove that $\langle\, .\, ,\, .\, \rangle_L$ is even positive definite. Indeed, for all $p \in \mathbb{R}[\underline{X}]_{d+1}$ with $\langle \overline{p}^L, \overline{p}^L \rangle_L = 0$ then $L(p^2) = 0$ and then $p \in U_L$ as we have shown in (6.1). □

Definition and Notation 6.1.4. Let $i \in \{1,\dots,n\}$. We define and denote the *i-th truncated GNS multiplication operator* of L as the following map between euclidean vector subspaces of V_L:

$$M_{L,i} : \Pi_L(V_L) \longrightarrow \Pi_L(V_L),\ \overline{p}^L \mapsto \Pi_L(\overline{pX_i}^L) \text{ for } p \in \mathbb{R}[\underline{X}]_d \tag{6.5}$$

where Π_L is the orthogonal projection map of V_L into the vector subspace:

$$\{\ \overline{p}^L \mid p \in \mathbb{R}[\underline{X}]_d\}$$

with respect to the inner product $\langle\ .\ ,\ .\ \rangle_L$. We will call and denote the vector subspace:

$$T_L := \Pi_L(V_L) = \{\ \overline{p}^L \mid p \in \mathbb{R}[\underline{X}]_d\} \subseteq V_L \tag{6.6}$$

the *GNS-truncation* of L.

Proposition 6.1.5. The i-th truncated GNS multiplication operator of L is a self-adjoint endomorphism of T_L for all $i \in \{1,\dots,n\}$.

Proof. Let $i \in \{1,\dots,n\}$. We demonstrate first that the i-th truncated GNS multiplication operator of L is well defined. $M_{L,i}$ is well defined if and only if $M_{L,i}(\overline{p}^L) = \overline{0}^L$ for all $p \in U_L \cap \mathbb{R}[\underline{X}]_d$ if and only if $\Pi_L(\overline{X_ip}^L) = \overline{0}^L$ for all $p \in U_L \cap \mathbb{R}[\underline{X}]_d$. Since $\Pi_L(\overline{X_ip}^L) \in T_L$ we can choose $q \in \mathbb{R}[\underline{X}]_d$ such that $\overline{q}^L = \Pi_L(\overline{X_ip}^L)$ and then:

$$\begin{aligned} L(q^2) = \langle \overline{q}^L, \overline{q}^L\rangle_L = \langle \Pi_L(\overline{X_ip}^L), \Pi_L(\overline{X_ip}^L)\rangle_L &\overset{\Pi_L\circ\Pi_L=\Pi_L}{=} \langle \Pi_L(\overline{X_ip}^L), \overline{X_ip}^L\rangle_L \\ = \langle \overline{q}^L, \overline{X_ip}^L\rangle_L = L(q(X_ip)) = L((qX_i)p) &\overset{p\in U_L}{=} 0 \end{aligned}$$

Therefore $\Pi_L(\overline{X_ip}^L) = \overline{0}^L$ for all $p \in U_L$. Let us see now that $M_{L,i}$ is a self-adjoint endomorphism. For this purpose, take $p, q \in \mathbb{R}[\underline{X}]_d$ then:

$$\begin{aligned} \langle M_{L,i}(\overline{p}^L), \overline{q}^L\rangle_L = \langle \Pi_L(\overline{X_ip}^L), \overline{q}^L\rangle_L = \langle \overline{X_ip}^L, \Pi_L(\overline{q}^L)\rangle_L = \langle \overline{X_ip}^L, \overline{q}^L\rangle_L = L((X_ip)q) \\ = L(p(X_iq)) = \langle \overline{p}^L, \overline{X_iq}^L\rangle_L = \langle \Pi_L(\overline{p}^L), \overline{X_iq}^L\rangle_L = \langle \overline{p}^L, \Pi_L(\overline{X_iq}^L)\rangle_L = \langle \overline{p}^L, M_{L,i}(\overline{q}^L)\rangle_L \end{aligned}$$

□

Remark 6.1.6. The GNS truncated construction for $L \in \mathbb{R}[\underline{X}]^*$ with $L(\sum \mathbb{R}[\underline{X}]^2) \subseteq \mathbb{R}_{\geq 0}$ is the same as the original (5.0.2) modulo U_L. The truncated GNS representation space of L and the GNS truncation of L are $\frac{\mathbb{R}[\underline{X}]}{U_L}$ where:

$$U_L = \{p \in \mathbb{R}[\underline{X}] \mid L(p^2) \geq 0\}$$

Since $\frac{\mathbb{R}[\underline{X}]}{U_L}$ is a commutative ring then the truncated GNS multiplication operators of L commute, . One can easily prove that U_L is an ideal. Indeed it is clear that if $p, q \in U_L$ then $L((p+q)^2) = 0$, and if $p \in U_L$ and $q \in \mathbb{R}[\underline{X}]$ then $L(p^2q^2) = L(p(pq^2)) \overset{p\in U_L}{=} 0$ implies $pq \in U_L$.

6.2 A Smul'jan result: the modified moment matrix

Lemma and Definition 6.2.1. Remember we denote $L' := L_{\mathbb{R}[\underline{X}]_{2d}}$. Let B_L be the transformation matrix of the following bilinear form with respect to the standard monomial basis:

$$\mathbb{R}[\underline{X}]_{d+1} \times \mathbb{R}[\underline{X}]_d \longrightarrow \mathbb{R}, (p,q) \longmapsto L(pq)$$

Then it holds $\operatorname{rank} M_{L'} = \operatorname{rank} B_L$ and we define and denote the *modified moment matrix of* L as:

$$\widetilde{M_L} := \left(\begin{array}{c|c} M_{L'} & M_{L'}W \\ \hline W^T M_{L'} & W^T M_{L'} W \end{array}\right)$$

where W is a matrix such that $M_{L'}W = C_L$ and C_L is the submatrix of B_L remaining from eliminating the columns corresponding to the matrix $M_{L'}$. $\widetilde{M_L}$ is well defined since it does not depend from the choice of such W and it is positive semidefinite. Moreover, for every $S \in \operatorname{Sym} \mathbb{R}^{r_{d+1} \times r_{d+1}}$, it holds that the following matrix is positive semidefinite:

$$\left(\begin{array}{c|c} M_{L'} & M_{L'}W \\ \hline W^T M_{L'} & S \end{array}\right) \succeq 0$$

if and only if there exists $X \in \mathbb{R}^{r_{d+1} \times r_{d+1}}$ such that:

$$S = W^T M_{L'} W + XX^T$$

Proof. Notice that $M_{L'}$ is the transformation matrix of the linear map:

$$\varphi := \mathbb{R}[\underline{X}]_d \longrightarrow \mathbb{R}[\underline{X}]_d^*, p \mapsto (q \mapsto L(pq)) \text{ for } p, q \in \mathbb{R}[\underline{X}]_d$$

with respect to the standard monomial basis and in the same way B_L is the transformation matrix of the linear map:

$$\psi := \mathbb{R}[\underline{X}]_{d+1} \longrightarrow \mathbb{R}[\underline{X}]_d^*, p \mapsto (q \mapsto L(pq)) \text{ for } p \in \mathbb{R}[\underline{X}]_{d+1}, q \in \mathbb{R}[\underline{X}]_d$$

with respect to the standard monomial basis. Note that to prove $\operatorname{rank} M_{L'} = \operatorname{rank} B_L$ it is the same than to prove $\varphi(\mathbb{R}[\underline{X}]_d) = \psi(\mathbb{R}[\underline{X}]_{d+1})$. It is obvious that $\varphi(\mathbb{R}[\underline{X}]_d) \subseteq \psi(\mathbb{R}[\underline{X}]_{d+1})$. For the other inclusion we take $\Lambda \in \psi(\mathbb{R}[\underline{X}]_{d+1})$ then there exists $p \in \mathbb{R}[\underline{X}]_{d+1}$ such that $\Lambda(q) = \psi(p)(q) = L(pq)$ for all $q \in \mathbb{R}[\underline{X}]_d$. We look for a $g \in \mathbb{R}[\underline{X}]_d$ such that $\Lambda(q) = L(gq)$ for all $q \in \mathbb{R}[\underline{X}]_d$, because then $\varphi(g)(q) = L(gq) = \Lambda(q)$ for all $q \in \mathbb{R}[\underline{X}]_d$ implying $\varphi(g) = \Lambda$ and then we could conclude $\Lambda \in \varphi(\mathbb{R}[\underline{X}]_d)$. In other words, we want to show that there exists $g \in \mathbb{R}[\underline{X}]_d$ such that : $L(pq) = L(gq)$ for every $q \in \mathbb{R}[\underline{X}]_d$. With more different words, our aim is to find $g \in \mathbb{R}[\underline{X}]_d$ such that:

$$\langle \overline{p}^L, \overline{q}^L \rangle_L = \langle \overline{g}^L, \overline{q}^L \rangle_L \text{ for every } q \in \mathbb{R}[\underline{X}]_d$$

For this aim we define the following linear form:

$$\Lambda_p := \frac{\mathbb{R}[\underline{X}]_d}{U_L \cap \mathbb{R}[\underline{X}]_d} \longrightarrow \mathbb{R}, \overline{q}^{L'} \mapsto L(pq) \text{ for every } q \in \mathbb{R}[\underline{X}]_d$$

$\Lambda_p \in (\frac{\mathbb{R}[\underline{X}]_d}{U_L \cap \mathbb{R}[\underline{X}]_d})^*$. Since $\frac{\mathbb{R}[\underline{X}]_d}{U_L}$ is a finite dimensional euclidean vector space is also in particular a Hilbert space and then by the Fréchet-Riesz Representation Theorem there exists $\overline{g} \in \frac{\mathbb{R}[\underline{X}]_d}{U_L \cap \mathbb{R}[\underline{X}]_d}$ with $g \in \mathbb{R}[\underline{X}]_d$, such that:

$$\Lambda_p(\overline{q}) = \langle \overline{q}, \overline{g} \rangle_L \text{ for every } q \in \mathbb{R}[\underline{X}]_d$$

Therefore $L(pq) = \Lambda_p(\overline{q}) = \langle \overline{q}, \overline{g} \rangle = L(qg)$ for every $q \in \mathbb{R}[\underline{X}]_d$. Then $\operatorname{rank} M_{L'} = \operatorname{rank} B_L$ and therefore there exists W (may not be unique) such that $M_{L'}W = C_L$. Now, we claim that the modified moment matrix $\widetilde{M_L}$, does not depend on the choice of a matrix W with the property $M_{L'}W = C_L$. Indeed, assume there are matrices $W_1, W_2 \in \mathbb{R}^{s_d \times r_{d+1}}$ such that $C_L = M_{L'}W_1 = M_{L'}W_2$. Let us denote:

$$W_1 := (P_1 \dots P_{r_{d+1}}) \text{ and } W_2 := (Q_1 \dots Q_{r_{d+1}})$$

where $P_1, \dots, P_{r_{d+1}}$ and $Q_1, \dots, Q_{r_{d+1}}$ are the respective column vectors of the matrices W_1 and W_2 and define $p_i := \operatorname{poly}(P_i) \in \mathbb{R}[\underline{X}]_{s_d}$ and $q_i := \operatorname{poly}(Q_i) \in \mathbb{R}[\underline{X}]_{s_d}$ for $i \in \{1, \dots, r_{d+1}\}$. We have the following matrix equality:

$$M_{L'}(P_1 \dots P_{r_{d+1}}) = M_{L'}(Q_1 \dots Q_{r_{d+1}})$$

Let $i \in \{1, \dots, r_{d+1}\}$ then:

$$M_{L'}P_i = M_{L'}Q_i \iff M_{L'}(P_i - Q_i) = 0 \overset{M_{L'} \succeq 0}{\iff} (P_i - Q_i)^T M_{L'}(P_i - Q_i) = 0 \underset{(6.1)}{\overset{4.2.10}{\iff}} p_i - q_i \in U_{L'}$$

This implies $L'(p_i p_j) = L'(q_i q_j)$ and again due to 4.2.10 we get that:

$$P_i^T M_{L'} P_j = Q_i^T M_{L'} Q_j \text{ for all } i, j \in \{1, \dots, r\}$$

Then we have got that:

$$W_1^T M_{L'} W_1 = \begin{pmatrix} P_1^T M_{L'} P_1 & \dots & P_1^T M_{L'} P_r \\ \vdots & \ddots & \vdots \\ P_r^T M_{L'} P_1 & \dots & P_r^T M_{L'} P_r \end{pmatrix} =$$

$$\begin{pmatrix} Q_1^T M_{L'} Q_1 & \dots & Q_1^T M_{L'} Q_r \\ \vdots & \ddots & \vdots \\ Q_r^T M_{L'} Q_1 & \dots & Q_r^T M_{L'} Q_r \end{pmatrix} = W_2^T M_{L'} W_2$$

Therefore $W_1^T M_{L'} W_1 = W_2^T M_{L'} W_2$, and we can conclude that $\widetilde{M_L}$ is well defined. On the other side, choose W such that $M_{L'}W = C_L$ and let $S \in \operatorname{Sym} \mathbb{R}^{r_{d+1} \times r_{d+1}}$ such that:

$$\left(\begin{array}{c|c} M_{L'} & M_{L'}W \\ \hline W^T M_{L'} & S \end{array}\right) \succeq 0$$

Then for every $u \in \mathbb{R}^{r_{d+1}}$ it holds:

$$\left(-u^T W^T, u^T\right) \left(\begin{array}{c|c} M_{L'} & M_{L'}W \\ \hline W^T M_{L'} & S \end{array}\right) \begin{pmatrix} -Wu \\ u \end{pmatrix} \geq 0$$

or what is the same, for every $u \in \mathbb{R}^{r_{d+1}}$ it holds that:

$$u^T(-W^T M_{L'} W + S)u \geq 0$$

This last fact is equivalent to $S - W^T M_{L'} W \succeq 0$ and by the characterization (iii) in 3.1.5 of positive semidefinite matrix, we can conclude that $S = W^T M_{L'} W + XX^T$ for some matrix $X \in \mathbb{R}^{r_{d+1} \times r_{d+1}}$. Conversely, suppose $S = W^T M_{L'} W + XX^T$ for some matrix $X \in \mathbb{R}^{r_{d+1} \times r_{d+1}}$ and since $M_{L'}$ is positive semidefinite write $M_{L'} = CC^T$ for some $C \in \mathbb{R}^{s_d \times s_d}$, then we have the following factorization:

$$\begin{aligned}\left(\begin{array}{c|c} M_{L'} & M_{L'}W \\ \hline W^T M_{L'} & S \end{array}\right) &= \left(\begin{array}{c|c} C & 0 \\ \hline W^T C & X \end{array}\right) \left(\begin{array}{c|c} C & 0 \\ \hline W^T C & X \end{array}\right)^T \\ &= \left(\begin{array}{c|c} CC^T & CC^T W \\ \hline W^T CC^T & W^T CC^T W^T + XX^T \end{array}\right)\end{aligned} \tag{6.7}$$

and therefore once again with the factorization (iii) of positive semidefinite matrix in 3.1.5 we can conclude:

$$\left(\begin{array}{c|c} M_{L'} & M_{L'}W \\ \hline W^T M_{L'} & S \end{array}\right) \succeq 0$$

Note that by taking $X := 0$ in the factorization (6.7), we obtain $\widetilde{M_L}$ positive semidefinite. □

Remark and definition 6.2.2. Let $t \in \mathbb{N}$ such that $d > t$ and denote $L_t := L_{|\mathbb{R}[\underline{X}]_{2d+2-2t}}$. Following the proof of Lemma 6.2.1 one can show that there exist matrices $W \in \mathbb{R}^{s_{d-t+1} \times s}$ and $X \in \mathbb{R}^{s \times s}$ where $s := |\{\underline{X}^\alpha \mid |\alpha| \in [d+2-t, d+1]\}|$ such that:

$$M_L = \left(\begin{array}{c|c} M_{L_t} & M_{L_t}W \\ \hline W^T M_{L_t} & W^T M_{L_t} W + XX^T \end{array}\right) \succeq 0$$

We define and denote *the modified moment matrix of L of degree t* as:

$$\widetilde{M_L}^t := \left(\begin{array}{c|c} M_{L_t} & M_{L_t}W \\ \hline W^T M_{L_t} & W^T M_{L_t} W \end{array}\right)$$

$\widetilde{M_L}^t$ is well defined since it does not depend on the choice of the matrix $W \in \mathbb{R}^{s_{d-t+1} \times s}$. Note that for $t = 1$ we are in the case of Lemma 6.2.1 and it holds $\widetilde{M_L} = \widetilde{M_L}^1$.

6.3 Gaussian quadrature rule

In this section we will prove the existence of a quadrature rule representation for the positive semidefinite linear form $L \in \mathbb{R}[\underline{X}]_{2d+2}^*$ on the set:

$$\{\sum_{i=1}^{s} p_i q_i \mid s \in \mathbb{N}, p_i \in \mathbb{R}[\underline{X}]_{d+1} \text{ and } q_i \in \mathbb{R}[\underline{X}]_d + U_L\} \supseteq \mathbb{R}[\underline{X}]_{2d+1}$$

by providing that the truncated GNS multiplication operators commute. We will also demonstrate that this condition is strictly more general than the very well known condition of being flat (see definition in 6.3.18), condition that for its part ensure the existence of a quadrature rule representation for L on the whole space in contrast with the quadrature rule in a space that contains $\mathbb{R}[\underline{X}]_{2d+1}$ that we get in case the truncated GNS multiplication operators commute.

Proposition 6.3.1. The vector space T_L, defined on (6.6), and the vector space $\frac{\mathbb{R}[\underline{X}]_d}{U_L \cap \mathbb{R}[\underline{X}]_d}$ are canonically isomorphic.

Proof. Let us consider the following linear map between euclidean vector spaces:

$$\sigma_L : T_L \longrightarrow V_{L'} : \overline{p}^L \longmapsto \overline{p}^{L'} \text{ for every } p \in \mathbb{R}[\underline{X}]_d \tag{6.8}$$

where remember we denoted $L' := L_{|\mathbb{R}[\underline{X}]_{2d}}$. It is well defined since for every $\overline{p}^L, \overline{q}^L \in T_L$ such that $\overline{p}^L = \overline{q}^L$ we can assume without loss of generality that $p, q \in \mathbb{R}[\underline{X}]_d$, and therefore:

$$\overline{p}^L = \overline{q}^L \Leftrightarrow L((p-q)^2) = 0 \Leftrightarrow L'((p-q)^2) = 0 \Leftrightarrow \overline{p}^{L'} = \overline{q}^{L'} \Leftrightarrow \sigma_L(\overline{p}^L) = \sigma_L(\overline{q}^L)$$

σ_L is also a linear isometry, since for every $p, q \in \mathbb{R}[\underline{X}]_d$ we have:

$$\langle \overline{p}^L, \overline{q}^L \rangle_L = L(pq) = L^{'}(pq) = \langle \overline{p}^{L'}, \overline{q}^{L'} \rangle_{L'} = \langle \sigma_L(\overline{p}^L), \sigma_L(\overline{q}^L) \rangle_{L'}$$

Then σ_L is immediately injective. On the other side, σ_L is surjective since for every $\overline{p}^{L'} \in V_{L'}$ with $p \in \mathbb{R}[\underline{X}]_d$, it holds that $\sigma_L(\overline{p}^L) = \overline{p}^{L'}$. Thence σ_L is an isomorphism between vector spaces. □

Notation 6.3.2. For a linear form $\Lambda \in \mathbb{R}[\underline{X}]_{2d+2}^*$ such that $\Lambda(\sum \mathbb{R}[\underline{X}]_{d+1}^2) \subseteq \mathbb{R}_{\geq 0}$ we will detone by σ_Λ the following isomorphism of euclidean vector spaces already defined in (6.8):

$$\sigma_\Lambda : T_\Lambda \longmapsto V_{\Lambda'}, \overline{p}^\Lambda \mapsto \overline{p}^{\Lambda'}, \text{ for } p \in \mathbb{R}[\underline{X}]_d \tag{6.9}$$

Remark 6.3.3. For $v_1, \ldots, v_r \in \mathbb{R}[\underline{X}]_d$, we have $\overline{v_1}^L, \ldots, \overline{v_r}^L$ is an orthonormal basis of T_L if and only if $\overline{v_1}^{L'}, \ldots, \overline{v_r}^{L'}$ is an orthonormal basis of $V_{L'}$.

Definition 6.3.4. Let $I \subseteq \mathbb{R}[\underline{X}]$ be an ideal. I is said to be radical when $I = \mathfrak{I}(V_\mathbb{C}(I))$.

The following Theorem and Lemma, are some algebraic preliminaries that we will use to prove Proposition 6.3.8. The proofs of the algebraic preliminaries can be seen for example in [17] and [8].

Theorem 6.3.5. An ideal $I \subseteq \mathbb{R}[\underline{X}]$ is zero dimensional (i.e. $|V_\mathbb{C}(I)| < \infty$) if and only if the vector space $\mathbb{R}[\underline{X}]/I$ is finite dimensional. Moreover $|V_\mathbb{C}(I)| \leq \dim(\mathbb{R}[\underline{X}]/I)$, with equality if and only if the ideal I is radical.

Proof. [17, Theorem 2.6 in page 15]. □

Lemma 6.3.6. Let $I \subseteq \mathbb{R}[\underline{X}]$ be an ideal. I is radical if and only if

$$\text{For all } g \in \mathbb{R}[\underline{X}] \text{ such that } g^2 \in I \Longrightarrow g \in I \tag{6.10}$$

Proof. There is a proof in [17, Lemma 2.2 in page 13]. □

Proposition 6.3.7. Let $\Lambda \in \mathbb{R}[\underline{X}]^*$ such that $\Lambda(\sum \mathbb{R}[\underline{X}]^2) \subseteq \mathbb{R}_{\geq 0}$. Then U_Λ is a radical ideal.

Proof. In 6.1.6 we saw that U_Λ is an ideal, let us prove that it is radical ideal. Let $g \in \mathbb{R}[\underline{X}]$ such that $g^2 \in U_\Lambda$. In particular $\Lambda(g^2 1) = 0$ and this implies according to (6.1) $g \in U_\Lambda$. □

Proposition 6.3.8. Let $\Lambda = \sum_{i=1}^N \lambda_i \operatorname{ev}_{a_i} \in \mathbb{R}[\underline{X}]^*$, with $N \in \mathbb{N}$, $\lambda_1 > 0, \dots, \lambda_N > 0$, and $a_1, \dots, a_N \in \mathbb{R}^n$ then:

$$\dim\left(\frac{\mathbb{R}[\underline{X}]}{U_\Lambda}\right) = |\{a_1, \dots, a_N\}|$$

Proof. We have the following equalities:

$$\begin{aligned} U_\Lambda &= \{p \in \mathbb{R}[\underline{X}] \mid \sum_{i=1}^N \lambda_i p^2(a_i) = 0\} = \{p \in \mathbb{R}[\underline{X}] \mid p^2(a_i) = 0 \text{ for all } i \in \{1, \dots, n\}\} \\ &= \{p \in \mathbb{R}[\underline{X}] \mid p(a_i) = 0 \text{ for all } i \in \{1, \dots, n\}\} = \mathfrak{I}(\{a_1, \dots, a_N\}) \end{aligned}$$

and since $\{a_1, \dots, a_N\} \subseteq \mathbb{R}^n$ is an algebraic set, by the ideal-variety correspondence (see [8]), it holds:

$$V_{\mathbb{C}}(\mathfrak{I}(\{a_1, \dots, a_N\})) = \{a_1, \dots, a_N\}$$

which is the same as $V_{\mathbb{C}}(U_\Lambda) = \{a_1, \dots, a_N\}$. Notice that by Theorem 6.3.5 is enough to prove that U_Λ is radical to finish the proof. In fact by Proposition 6.3.7 U_Λ is radical. Applying Theorem 6.3.5 we have the result. □

Let us review some well known bounds on the number of nodes of quadrature rules for L on $\mathbb{R}[\underline{X}]_{2d+2}$ and on $\mathbb{R}[\underline{X}]_{2d+1}$ (see [4] and [29]).

Proposition 6.3.9. Then the number of nodes N, of a quadrature rule for L satisfies:

$$\operatorname{rank} M_L \leq N \leq |V_{\mathbb{R}}(U_L)| \tag{6.11}$$

Proof. Let $L = \sum_{i=1}^N \lambda_i \operatorname{ev}_{a_i} \in \mathbb{R}[\underline{X}]_{2d+2}^*$ for $a_1, \dots, a_N \in \mathbb{R}^n$ pairwise different points and $\lambda_1, \dots, \lambda_N > 0$ weights and define $\Lambda := \sum_{i=1}^N \lambda_i \operatorname{ev}_{a_i} \in \mathbb{R}[\underline{X}]^*$. Let us consider the following canonical map:

$$\frac{\mathbb{R}[\underline{X}]_{d+1}}{U_L} \hookrightarrow \frac{\mathbb{R}[\underline{X}]}{U_\Lambda}$$

By Proposition 6.3.8 we have that:

$$\operatorname{rank} M_L = \dim\left(\frac{\mathbb{R}[\underline{X}]_{d+1}}{U_L}\right) \leq \dim\left(\frac{\mathbb{R}[\underline{X}]}{U_\Lambda}\right) = N$$

On the other side, it holds that $\{a_1, \dots, a_N\} \subseteq V_{\mathbb{C}}(U_L) \cap \mathbb{R}^n = V_{\mathbb{R}}(U_L)$, since for all $p \in U_L$ we have $L(p^2) = 0$ and then $p(a_i) = 0$ for all $i \in \{1, \dots, N\}$. this implies $N \leq |V_{\mathbb{C}}(U_L)|$. □

Remark 6.3.10. Note that the bound in the number of nodes given in (6.11) could be in some cases a criterion to discard the existence of a quadrature rule for a positive semidefinite linear form. That is to say, if:

$$\operatorname{rank}(M_L) > V_{\mathbb{R}}(U_L)$$

L has not a quadrature rule. For example, if we take the positive semidefinite linear form in one variable: $\ell \in \mathbb{R}[\underline{X}]_4^*$, with moment matrix:

$$M_\ell = \begin{array}{c} \\ 1 \\ X \\ X^2 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & X & X^2 \end{array} \\ \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix} \end{array}$$

One can easily compute that $U_\ell = \{a - aX \mid a \in \mathbb{R}\}$ and therefore:

$$V_{\mathbb{C}}(U_\ell) = \{x \in \mathbb{R} \mid a - ax = 0 \text{ for all } a \in \mathbb{R}\} = \{1\}$$

and in conclusion we have that $\operatorname{rank} M_\ell = 2 > |V_{\mathbb{R}}(U_\ell)| = 1$. Therefore $\ell \in \mathbb{R}[X]_4^*$ has not a quadrature rule by Proposition 6.11. However $\ell(p) = p(1)$ for all $p \in \mathbb{R}[X]_3$. See also Example 6.3.21. This fact motivates the following Proposition.

Proposition 6.3.11. The number of nodes N, of a quadrature rule for L on $\mathbb{R}[\underline{X}]_{2d+1}$ satisfies:

$$N \geq \dim(T_L)$$

Proof. Assume that L has a quadrature rule on $\mathbb{R}[\underline{X}]_{2d+1}$ such that:

$$L(p) = \sum_{i=1}^{N} \lambda_i p(a_i) \text{ for every } p \in \mathbb{R}[\underline{X}]_{2d+1}$$

where we can assume without loss of generality that the points $a_1, \dots, a_N \in \mathbb{R}^n$ are pairwise different and $\lambda_1, \dots, \lambda_N > 0$ with $N < \infty$ for $N \in \mathbb{N}$. Let us set:

$$\Lambda := \sum_{i=1}^{N} \lambda_i \operatorname{ev}_{a_i} \in \mathbb{R}[\underline{X}]^*.$$

Then, the following canonical linear map between euclidean vector spaces is an isometry:

$$\sigma_1 : T_L \longrightarrow \frac{\mathbb{R}[\underline{X}]}{U_\Lambda}, \overline{p}^L \mapsto \overline{p}^\Lambda \text{ for every } p \in \mathbb{R}[\underline{X}]_d \tag{6.12}$$

Indeed, it is clearly well defined since $U_L \cap \mathbb{R}[\underline{X}]_d \subseteq U_\Lambda$. It holds also that σ_1 is a linear isometry since, for all $p, q \in \mathbb{R}[\underline{X}]_d$:

$$\langle \overline{p}^L, \overline{q}^L \rangle_L = L(pq) = \Lambda(pq) = \left\langle \overline{p}^\Lambda, \overline{q}^\Lambda \right\rangle_\Lambda = \langle \sigma_1(\overline{p}^L), \sigma_1(\overline{q}^L) \rangle_\Lambda$$

Since σ_1 is a linear isometry is immediately injective, and then:

$$\dim(T_L) \leq \dim \left(\frac{\mathbb{R}[\underline{X}]}{U_\Lambda} \right) \tag{6.13}$$

And now we can apply the Proposition 6.3.8, to conclude the proof. □

Definition 6.3.12. We call a quadrature rule for L on $\mathbb{R}[\underline{X}]_{2d+1}$ with minimal number of nodes, that is to say with $\dim(T_L)$ nodes a Gaussian quadrature rule.

Lemma 6.3.13. Assume that the truncated multiplication operators commute. Then for all $p \in \mathbb{R}[\underline{X}]_{d+1}$ we have the following equality:

$$p(M_{L,1}, \dots, M_{L,n})(\overline{1}^L) = \Pi_L(\overline{p}^L) \tag{6.14}$$

Proof. Let $p = \underline{X}^\alpha$ for $\alpha \in \mathbb{N}^n$ with $|\alpha| \leq d+1$. We continue the proof by induction on $|\alpha|$:

- For $|\alpha| = 0$, we have that $\underline{X}^\alpha = 1$ then:

$$1(M_{L,1}, \dots, M_{L,n})(\overline{1}^L) = \mathrm{Id}_{\Pi_L(V_L)}(\overline{1}^L) = \overline{1}^L = \Pi_L(\overline{1}^L)$$

- Let assume the statement is true for $|\alpha| = d$. Let us show it is also true for $|\alpha| = d+1$. Let $p = X_i q$ for some $i \in \{1, \dots, n\}$ and $q = \underline{X}^\beta$ with $|\beta| = d$, then $\Pi_L(\overline{q}^L) = \overline{q}^L$ since $\overline{q}^L \in T_L$, and it holds:

$$p(M_{L,1}, \dots, M_{L,n})(\overline{1}^L) = (M_{L,i} \circ q(M_{L,1}, \dots, M_{L,n}))(\overline{1}^L) =$$
$$M_{L,i}(q(M_{L,1}, \dots, M_{L,n})(\overline{1}^L)) = M_{L,i}(\overline{q}^L) = \Pi_L(\overline{X_i q}^L) = \Pi_L(\overline{p}^L)$$

since we have proved (6.14) for monomials then by the linearity of the orthogonal projection (6.14) is also true for polynomials. □

Theorem 6.3.14. Assume the truncated multiplication operators of L commute, and consider the set:

$$G_L := \{\sum_{i=1}^{s} p_i q_i \mid s \in \mathbb{N}_0, p_i \in \mathbb{R}[\underline{X}]_{d+1} \text{ and } q_i \in \mathbb{R}[\underline{X}]_d + U_L\} \tag{6.15}$$

then there exists a quadrature rule for L on G_L with $\dim(T_L)$ many nodes.

Proof. Since the truncated multiplication operators of L commute by Reminder 5.0.1 there exists an orthonormal basis $v := \{v_1, \dots, v_N\}$ of T_L consisting of common eigenvectors of the GNS truncated multiplication operators of L. That is to say, there exist $a_1, \dots, a_N \in \mathbb{R}^n$ such that:

$$M_{L,i}v_j = a_{j,i}v_j \text{ for all } i \in \{1,\dots,n\} \text{ and } j \in \{1,\dots,N\}$$

where $N := \dim(T_L)$. Since it always holds $\overline{1}^L \in T_L$, we can write:

$$\overline{1}^L = b_1v_1 + \cdots + b_Nv_N \tag{6.16}$$

for some $b_1,\dots,b_N \in \mathbb{R}$. Let us define $\lambda_i := b_i^2$ for all $i \in \{1,\dots,N\}$. Let $g = pq$ such that $p \in \mathbb{R}[\underline{X}]_{d+1}$ and $q \in \mathbb{R}[\underline{X}]_d + U_L$, then using Lemma 6.3.13 we have the two equalities:

$$\Pi_L(\overline{p}^L) = p(M_{L,1},\dots,M_{L,n})(\overline{1}^L) \text{ and } \overline{q}^L = q(M_{L,1},\dots,M_{L,n})(\overline{1}^L) \tag{6.17}$$

Using this equalities (6.17), using that the orthogonal projection Π_L is selfadjoint, using that $\{v_1,\dots,v_N\}$ is an orthonormal basis of T_L consisting of common eigenvectors of the GNS truncated multiplication operators of L and also using the equation (6.16), with the same idea as we got the reformulation of the problem in (5.2) we have:

$$L(g) = L(pq) = \langle \overline{p}^L, \overline{q}^L \rangle_L \overset{\overline{q}^L \in T_L}{=} \langle \overline{p}^L, \Pi_L(\overline{q}^L) \rangle_L =$$

$$\langle \Pi_L(\overline{p}^L), \overline{q}^L \rangle_L = \langle p(M_{L,1},\dots,M_{L,n})(\overline{1}^L), q(M_{L,1},\dots,M_{L,n})(\overline{1}^L) \rangle_L =$$

$$\langle p(M_{L,1},\dots,M_{L,n})(\sum_{i=1}^N b_iv_i), q(M_{L,1},\dots,M_{L,n})(\sum_{j=1}^N b_jv_j) \rangle_L =$$

$$\sum_{i=1}^N \sum_{j=1}^N b_ib_j \langle p(M_{L,1},\dots,M_{L,n})(v_i), q(M_{L,1},\dots,M_{L,n})(v_j) \rangle_L =$$

$$\sum_{i=1}^N \sum_{j=1}^N b_ib_j \langle p(a_{i,1},\dots,a_{i,n})v_i, q(a_{j,1},\dots,a_{j,n})v_j \rangle_L =$$

$$\sum_{i=1}^N \sum_{j=1}^N b_ib_jp(a_i)q(a_j)\langle v_i, v_j \rangle_L = \sum_{j=1}^N b_j^2 p(a_j)q(a_j) = \sum_{j=1}^N \lambda_j p(a_j)q(a_j)$$

Then by linearity $L(p) = \sum_{i=1}^N \lambda_i p(a_i)$ for all $p \in G_L$. It remains to prove that the nodes of the quadrature rule for L that we got, $a_1,\dots,a_N \in \mathbb{R}^n$ are pairwise different, but this is true since $N = \dim T_L$ is the minimal possible number of nodes for a quadrature rule on $\mathbb{R}[\underline{X}]_{2d+1}$ as we proved in Proposition 6.3.11. □

Remark 6.3.15. In the same conditions of Theorem 6.3.14, since $\mathbb{R}[\underline{X}]_{2d+1} \subseteq G_L$, we got in particular a Gaussian quadrature rule for the linear form L.

Corollary 6.3.16. Let $n = 1$, i.e. $L \in \mathbb{R}[X]^*_{2d+2}$ with $L(\sum \mathbb{R}[X]^2_{d+1}) \geq 0$. Then L has a quadrature rule on G_L (6.15).

Proof. L has one truncated GNS multiplication operator, therefore the hypothesis of Theorem 6.3.14 holds and there is a quadrature rule on G_L for L. □

Proposition 6.3.17. The following assertions are equivalent:

(i) $\mathbb{R}[\underline{X}]_{d+1} = \mathbb{R}[\underline{X}]_d + U_L$

(ii) $T_L = V_L$

(iii) For all $\alpha \in \mathbb{N}_0^n$ with $|\alpha| = d+1$, there exists $p \in \mathbb{R}[\underline{X}]_d$ such that $\underline{X}^\alpha - p \in U_L$

(iv) The canonical map:

$$V_{L'} = \mathbb{R}[\underline{X}]_d / U_{L'} \hookrightarrow \mathbb{R}[\underline{X}]_{d+1} / U_L = V_L \tag{6.18}$$

is an isomorphism.

(v) $\dim(V_{L'}) = \dim(V_L)$

(vi) The moment matrices $(L(\underline{X}^{\alpha+\beta}))_{|\alpha|,|\beta| \leq d}$ and $(L(\underline{X}^{\alpha+\beta}))_{|\alpha|,|\beta| \leq d+1}$ have the same rank.

(vii) $M_L = \widetilde{M_L}$.

Proof. Note that the map (6.18) it is well defined since $\mathbb{R}[\underline{X}]_d \cap U_L = U_{L'}$ and one can see inmediately that:

$$(i) \iff (ii) \iff (iii) \iff (iv) \iff (v).$$

Let us show $(v) \iff (vi)$: $(L(\underline{X}^{\alpha+\beta}))_{|\alpha|,|\beta| \leq d+1}$ is the transformation matrix (or the associated matrix) of the bilinear form:

$$\mathbb{R}[\underline{X}]_{d+1} \times \mathbb{R}[\underline{X}]_{d+1} \longrightarrow \mathbb{R}, (p,q) \mapsto L(pq)$$

with respect to the the standard monomial basis, and therefore it is also the transformation matrix (or the associated matrix) of the linear map:

$$\mathbb{R}[\underline{X}]_{d+1} \longrightarrow \mathbb{R}[\underline{X}]^*_{d+1}, p \mapsto (q \mapsto L(pq)) \tag{6.19}$$

with respect to the corresponding dual basis of the standard monomial basis. The kernel of this linear map (6.19) is U_L, in consequence:

$$\operatorname{rank}((L(\underline{X}^{\alpha+\beta}))_{|\alpha|,|\beta| \leq d+1}) = \dim \mathbb{R}[\underline{X}]_{d+1} - U_L = \dim V_L$$

reasoning in the same way:

$$\operatorname{rank}((L(\underline{X}^{\alpha+\beta}))_{|\alpha|,|\beta|\leq d}) = \dim V_{L'}$$

Finally $(vi) \iff (vii)$:

$$\operatorname{rank}((L(\underline{X}^{\alpha+\beta}))_{|\alpha|,|\beta|\leq d}) = \operatorname{rank}((L(\underline{X}^{\alpha+\beta}))_{|\alpha|,|\beta|\leq d+1}) \iff$$

$$\operatorname{rank}(M_{L'}) = \operatorname{rank}(M_L) \iff$$

$$\operatorname{rank}\left(\begin{array}{c|c} M_{L'} & M_{L'}W \\ \hline W^T M_{L'} & W^T M_{L'} W \end{array}\right) = \operatorname{rank}(M_L), \text{ for all } W \in \mathbb{R}^{s_d \times s_{d+1}} \iff$$

$$\widetilde{M_L} = M_L$$

□

Definition 6.3.18. The linear form L is called *flat* if the equivalent conditions (i), (ii), (iii), (iv), (v), (vi) and (vii) in the Proposition 6.3.17 are satisfied.

Proposition 6.3.19. Suppose L is flat then the truncated GNS operators of L commute.

Proof. Asumme L is flat, and let $i, j \in \{1, \dots, n\}$ and $p \in \mathbb{R}[\underline{X}]_d$. We want to prove:

$$M_{L,i} \circ M_{L,j}(\overline{p}^L) = M_{L,j} \circ M_{L,i}(\overline{p}^L)$$

Since L is flat, by the characterization (iii) of 6.3.17, we can write: $X_i p = p_1 + q_1$ and $X_j p = p_2 + q_2$ with $p_1, p_2 \in \mathbb{R}[\underline{X}]_d$ and $q_1, q_2 \in U_L$. Then, the following holds:

$$M_{L,j}(\overline{p}^L) = \Pi_L(\overline{X_j p}^L) = \Pi_L(\overline{p_2 + q_2}^L) = \Pi_L(\overline{p_2}^L) + \Pi_L(\overline{q_2}^L) = \Pi_L(\overline{p_2}^L) = \overline{p_2}^L.$$

In the same way we get $M_{L,i}(\overline{p}^L) = \overline{p_1}^L$. Therefore:

$$M_{L,i} \circ M_{L,j}(\overline{p}^L) = M_{L,j} \circ M_{L,i}(\overline{p}^L) \iff$$

$$M_{L,i}(\overline{p_2}^L) = M_{L,j}(\overline{p_1}^L) \iff$$

$$\Pi_L(\overline{X_i p_2}^L) = \Pi_L(\overline{X_j p_1}^L)$$

Let $f \in \mathbb{R}[\underline{X}]_{d+1}$ and write $\Pi_L(f) := \overline{g}$ for some $g \in \mathbb{R}[\underline{X}]_d$. Then it holds:

$$\begin{aligned} \left\langle \Pi_L(\overline{X_i p_2 - X_j p_1}^L), \overline{f}^L \right\rangle_L = \left\langle \overline{X_i p_2}^L, \Pi_L(\overline{f}^L) \right\rangle_L - \left\langle \overline{X_j p_1}^L, \Pi_L(\overline{f}^L) \right\rangle_L = \\ \left\langle \overline{X_i p_2}^L, \overline{g}^L \right\rangle - \left\langle \overline{X_j p_1}^L, \overline{g}^L \right\rangle_L = L(X_i p_2 g) - L(X_j p_1 g) = \\ L(X_i p_2 g) - L(X_j p_1 g) = L(X_i g(X_j p - q_2)) - L(X_j g(X_i p - q_1)) = \\ L(X_i g X_j p) - L(X_i g q_2) - L(X_j q X_i p) - L(X_j g q_1) = 0 \end{aligned}$$

□

Here we show some examples of positive semidefinite linear forms such that the converse of Proposition 6.3.19 does not hold.

Example 6.3.20. The truncated GNS multiplication operators of the following linear form:

$$L : \mathbb{R}[X_1, X_2]_4 \to \mathbb{R}, p \mapsto \frac{1}{4}(p(0,0) + p(1,0) + p(-1,0) + p(0,1))$$

commute but L is not flat. Indeed, let us do the truncated GNS-construction of L.

$$M_L = \begin{array}{c} \\ 1 \\ X_1 \\ X_2 \\ X_1^2 \\ X_1X_2 \\ X_2^2 \end{array} \begin{array}{c} \begin{array}{cccccc} 1 & X_1 & X_2 & X_1^2 & X_1X_2 & X_2^2 \end{array} \\ \left(\begin{array}{ccc|ccc} 1 & 0 & \frac{1}{4} & \frac{1}{2} & 0 & \frac{1}{4} \\ 0 & \frac{1}{2} & 0 & 0 & 0 & 0 \\ \frac{1}{4} & 0 & \frac{1}{4} & 0 & 0 & \frac{1}{4} \\ \hline \frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{4} & 0 & \frac{1}{4} & 0 & 0 & \frac{1}{4} \end{array}\right) \end{array} := \left(\begin{array}{c|c} M_{L'} & B_L \\ \hline B_L^T & C_L \end{array}\right)$$

is the associated moment matrix of the linear form L and a basis of the truncated GNS-kernel of L is $\langle X_1X_2, X_2^2 - X_2 \rangle$. That is, the rank of M_L is 4. And since in the kernel there is no polynomials of degree less or equal than 1, we get that the unique element in the kernel of L' is the polynomial 0, then the truncated GNS space T_L is full dimensional:

$$T_L \overset{6.3.1}{\cong} \frac{\mathbb{R}[X_1, X_2]_1}{U_{L'}} \cong \mathbb{R}[X_1, X_2]_1$$

This implies the dimension of the GNS-truncated space is 3 and therefore L is not flat by (vi) in 6.3.17. We can also verify that L is not flat by computing $\widetilde{M_L}$. Indeed, in this case $M_{L'}$ is invertible and W_M is uniquely defined by $W_M = M_{L'}^{-1}B_M$, then $\widetilde{M_L}$ reads:

$$\widetilde{M_L} = \left(\begin{array}{ccc|ccc} 1 & 0 & \frac{1}{4} & \frac{1}{2} & 0 & \frac{1}{4} \\ 0 & \frac{1}{2} & 0 & 0 & 0 & 0 \\ \frac{1}{4} & 0 & \frac{1}{4} & 0 & 0 & \frac{1}{4} \\ \hline \frac{1}{2} & 0 & 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{4} & 0 & \frac{1}{4} & 0 & 0 & \frac{1}{4} \end{array}\right)$$

Since $\widetilde{M_L} \neq M_L$ then L is not flat by the characterization (vii) in 6.3.17.

Let us compute the truncated GNS multiplication operators of L. First note that:

$$T_L \cong \frac{\mathbb{R}[X_1, X_2]_1}{U_{L'}} = \left\langle \overline{1}^{L'}, \overline{X_1}^{L'}, \overline{X_2}^{L'} \right\rangle$$

Therefore by Remark 6.3.3 the truncated GNS space of L is:

$$T_L = \left\langle \overline{1}^{L}, \overline{X_1}^{L}, \overline{X_2}^{L} \right\rangle$$

With the Gram-Schmidt orthonormalization process we get the following orthonormal basis with respect to the GNS inner product of L:

$$\underline{v} := \left\{ \overline{1}^L, \overline{\sqrt{2}X_1}^L, \overline{-\frac{\sqrt{3}}{3} + \frac{4\sqrt{3}}{3}X_2}^L \right\}$$

The matrices of the GNS-multiplication operators with respect to this orthonormal basis are:

$$A_1 := M(M_{L,X_1}, \underline{v}) = \begin{pmatrix} 0 & \frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & 0 & -\frac{\sqrt{6}}{6} \\ 0 & -\frac{\sqrt{6}}{6} & 0 \end{pmatrix}$$

$$A_2 := M(M_{L,X_2}, \underline{v}) = \begin{pmatrix} \frac{1}{4} & 0 & \frac{\sqrt{3}}{4} \\ 0 & 0 & 0 \\ \frac{\sqrt{3}}{4} & 0 & \frac{3}{4} \end{pmatrix}$$

It is easy to check that the truncated GNS multiplication operators of L commute, that is $M_{L,X_1} \circ M_{L,X_2} - M_{L,X_2} \circ M_{L,X_1} = 0$. Now since M_{L,X_1} and M_{L,X_2} commute we can do the simultaneous diagonalization on both of them, applying Theorem 5.0.1, in order to find an orthonormal basis of the GNS truncation of L consisting of common eigenvectors of M_{L,X_1} and M_{L,X_2}. To do this we follow the same idea as in [24, Algorithm 4.1, Step 1] and compute for a matrix:

$$A = r_1 A_1 + r_2 A_2 \text{ such that, } r_1^2 + r_2^2 = 1$$

a matrix P orthogonal such that $P^T A P$ is a diagonal matrix. In this case, we get for:

$$P = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{6}}{4} & -\frac{\sqrt{6}}{4} \\ 0 & \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{3}}{2} & \frac{\sqrt{2}}{4} & \frac{\sqrt{2}}{4} \end{pmatrix}$$

$$P^T A_1 P = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -\frac{\sqrt{6}}{3} & 0 \\ 0 & 0 & \frac{\sqrt{6}}{3} \end{pmatrix} \text{ and } P^T A_2 P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

By an inspection over the proof of the Theorem 6.3.14 we can obtain the weights $\lambda_1 > 0$, $\lambda_2 > 0$, and $\lambda_3 > 0$ through the following operations:

$$P^T \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ -\frac{\sqrt{6}}{4} \\ -\frac{\sqrt{6}}{4} \end{pmatrix}$$

then $\lambda_1 = (\frac{1}{2})^2$ and $\lambda_2 = \lambda_3 = (-\frac{\sqrt{6}}{4})^2$. Therefore we get the following quadrature rule:

$$L = \frac{1}{4}\operatorname{ev}(0,1) + \frac{3}{8}\operatorname{ev}(-\frac{\sqrt{6}}{3}, 0) + \frac{3}{8}\operatorname{ev}(\frac{\sqrt{6}}{3}, 0) \text{ on } \mathbb{R}[X_1, X_2]_3 + U_L$$

Example 6.3.21. Consider the following polynomial optimization problem, taken from [11, Problem 4.7]:

$$\begin{aligned} \text{minimize} \quad & f(\underline{x}) = -12x_1 - 7x_2 + x_2^2 \\ \text{subject to:} \quad & \underline{x} = (x_1, x_2) \in \mathbb{R}^2 \\ & -2x_1^4 + 2 - x_2 = 0 \\ & 0 \leq x_1 \leq 2 \\ & 0 \leq x_2 \leq 3 \end{aligned}$$

We get the optimal value $P_4^* = -16.7389$ associated to the following optimal solution $y \in \mathbb{R}^{s_4}$, in the Lasserre relaxation of degree 4:

$$\mathbf{M} := M_2(y) = \begin{array}{c} \\ 1 \\ X_1 \\ X_2 \\ X_1^2 \\ X_1X_2 \\ X_2^2 \end{array} \begin{array}{c} \begin{array}{cccccc} 1 & X_1 & X_2 & X_1^2 & X_1X_2 & X_2^2 \end{array} \\ \left(\begin{array}{ccc|ccc} 1.0000 & 0.7175 & 1.4698 & 0.5149 & 1.0547 & 2.1604 \\ 0.7175 & 0.5149 & 1.0547 & 0.3694 & 0.7568 & 1.5502 \\ 1.4698 & 1.0547 & 2.1604 & 0.7568 & 1.5502 & 3.1755 \\ \hline 0.5149 & 0.3694 & 0.7568 & 0.2651 & 0.5430 & 1.1123 \\ 1.0547 & 0.7568 & 1.5502 & 0.5430 & 1.1123 & 2.2785 \\ 2.1604 & 1.5502 & 3.1755 & 1.1123 & 2.2785 & 8.7737 \end{array}\right) \end{array} \tag{6.20}$$

Let us build the truncated GNS construction associated to this optimal solution. Setting $\alpha := \mathbf{M}(1,2)$ and $\beta := \mathbf{M}(1,3)$, the truncated GNS kernel of $L_{\mathbf{M}}$ is:

$$U_{L_{\mathbf{M}}} = \langle -\alpha + X_1, -\beta + X_2, -\alpha^2 + X_1^2, -\alpha\beta + X_1X_2 \rangle$$

the truncated GNS representation space is:

$$V_{L_{\mathbf{M}}} = \langle 1, X_2^2 \rangle$$

we have that:

$$U_{L_{\mathbf{M}}} \cap \mathbb{R}[X_1, X_2]_1 = \langle -\alpha + X_1, -\beta + X_2 \rangle$$

We need to add the polynomial 1 to $U_{\mathbf{M}} \cap \mathbb{R}[X_1, X_2]_1$ to get basis of $\mathbb{R}[X_1, X_2]_1$ therefore we have that:

$$\frac{\mathbb{R}[X_1, X_2]_1}{U_{L_{\mathbf{M}}} \cap \mathbb{R}[X_1, X_2]_1} = \left\langle \overline{1}^{L_{\mathbf{M}'}} \right\rangle$$

Thence by Remark 6.3.3 we get that:

$$T_{\mathbf{M}} = \left\langle \overline{1}^{L_{\mathbf{M}}} \right\rangle$$

Since $v := \{ \overline{1}^{\mathbf{M}} \}$ is also an orthonormal basis with respect to the GNS product of L we can directly compute the matrices of truncated GNS multiplication operators of $\mathbf{M}$:

$$M(M_{\mathbf{M},X_1}, v) = \text{poly}^{-1}(X_1 1)\mathbf{M}\,\text{poly}^{-1}(1) = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix} \mathbf{M} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = (\alpha)$$

$$M(M_{\mathbf{M},X_2}, v) = \text{poly}^{-1}(X_2 1)\mathbf{M}\,\text{poly}^{-1}(1) = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} \mathbf{M} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = (\beta)$$

Despite of the optimal solution is not flat, note that 6.3.17 (v) does not hold:

$$1 = \dim T_{L_\mathbf{M}} \overset{6.3.1}{=} \dim V_{L_{\mathbf{M}'}} \neq \dim V_\mathbf{M} = 2$$

one can easily see that the GNS truncated multiplation operators of L commute. Then accordingly with Theorem 6.3.14, $\mathbf{M}$ admits the following quadrature rule:

$$L_\mathbf{M} = \text{ev}(\alpha, \beta) \text{ on } \mathbb{R}[X_1, X_2]_3 + U_{L_\mathbf{M}}$$

Moreover we have reach optimality. Indeed, $f \in \mathbb{R}[X_1, X_2]_3$ and (α, β) belongs to the basic closed semialgebraci set:

$$\{(x_1, x_2) \in \mathbb{R}^2 \mid -2x_1^4 + 2 - x_2 = 0,\ 0 \leq x_1 \leq 2,\ 0 \leq x_2 \leq 3\}$$

Then accordingly to Proposition 4.4.4 (iii) we obtain $P_4^* = P^* = -16.7389$. Note that L does not admit a quadrature rule on the whole space. Indeed, suppose $\mathbf{M}$ admits a quadrature rule with N nodes, then according with Proposition 6.3.9:

$$2 = \text{rank}\,\mathbf{M} \leq N \leq |V_\mathbb{C}(\mathbf{U_M})|$$

But we can easily see that $V_\mathbb{C}(\mathbf{U_M}) = \{(\alpha, \beta)\}$ and

$$\text{rank}\,\mathbf{M} = 2 > |V_\mathbb{C}(\mathbf{U_M})| = 1$$

prevents to $\mathbf{M}$ to have a quadrature rule on the whole space.

6.4 The Theorem of Curto and Fialkow for flat moment matrices

The following Corollary 6.4.1 is a very well known result of Curto and Fialkow, [4, corollary 5.14], written in terms of quadrature rules instead of nonnegative measures.

As it is known and as we have already mentioned in Theorem 5.0.6, if a linear form defined in a finite dimensional space is integration with respect to a measure then the linear form is integration with respect to another measure which is finitely atomic (and whose support is contained in the support of the first one). In [2] and in [17, Theorem 5.9] there is a nice proof of this result. The result of Curto and Fialkow uses tools of functional analysis like the Spectral theorem and the Riesz representation theorem. Monique Laurent gave also a more elementary proof (see [18, corollary 1.4]) that uses a corollary of the Hilbert Nullstellensatz and elementary linear algebra. The main contribution of this proof is that it does not need to find a flat extension of the linear form since the truncated GNS multiplication operators commute and we can apply directly the Theorem 5.0.1, and despite it uses the Hilbert Nullstellensatz in the proof of Theorem 6.3.5, we do not need to apply the Hilbert Nullstellensatz to show that the nodes are in $\mathbb{R}^n$, since the nodes are real because their coordinates are the eigenvalues of a symmetric matrix.

Corollary 6.4.1. Suppose L is flat then L has a quadrature rule with $\operatorname{rank}(M_L)$ many nodes, the minimal number of nodes.

Proof. If L is flat by Proposition 6.3.19 the truncated GNS multiplication operators of L commute and applying Theorem 6.3.14 then L has a quadrature rule on G_L (6.15), with:

$$\dim(T_L) \overset{L \text{ is flat}}{=} \dim(V_L) = \operatorname{rank}(M_L)$$

many nodes. Since L is flat, by the characterization (i) of Proposition 6.3.17:

$$\mathbb{R}[\underline{X}]_{d+1} = \mathbb{R}[\underline{X}]_d + U_L$$

Hence one can easily see that $G_L = \mathbb{R}[\underline{X}]_{2d+2}$. As a conclusion we get a quadrature rule for L with $\operatorname{rank}(M_L)$ many nodes. $\square$

6.5 The main Theorem

In this section we will demonstrate that the commutativity of the truncated GNS multiplication operators of L is equivalent to the matrix $W_L^T M_{L'} W_L$ being generalized Hankel.

Main Theorem 6.5.1. The following assertions are equivalent:

(i) The truncated multiplication operators $M_{L,1}, \ldots, M_{L,n}$ pairwise commute.

(ii) There exists a unique $\hat{L} \in \mathbb{R}[\underline{X}]_{2d+2}^*$ such that:

$$\begin{cases} \hat{L} = L \text{ on } \mathbb{R}[\underline{X}]_{2d+1} \\ \hat{L}(\sum \mathbb{R}[\underline{X}]_{d+1}^2) \subseteq \mathbb{R}_{\geq 0} \text{ and} \\ \hat{L} \text{ is flat.} \end{cases}$$

(iii) $\widetilde{M_L}$ is a generalized Hankel matrix

Proof. (i) $\implies$ (ii). By the Theorem 6.3.14 there exist $a_1, \ldots, a_N \in \mathbb{R}^n$ pairwise different nodes and $\lambda_1 > 0, \ldots, \lambda_N > 0$ weights, where $N := \dim(T_L)$ such that:

$$L(p) = \sum_{i=1}^{N} \lambda_i p(a_i) \text{ for all } p \in G_L$$

where G_L was defined in (6.15). Let us define, $\hat{L} := \sum_{i=1}^{N} \lambda_i \operatorname{ev}_{a_i} \in \mathbb{R}[\underline{X}]^*_{2d+2}$. We have shown in Theorem 6.3.14 that:

$$\begin{cases} \hat{L} = L \text{ on } \mathbb{R}[\underline{X}]_{2d+1} \\ \hat{L}(\sum \mathbb{R}[\underline{X}]^2_{d+1}) \subseteq \mathbb{R}_{\geq 0} \text{ and} \end{cases}$$

So it remains to show that $\tilde{L}$ is flat and according to the characterization (v) of 6.3.17 and 6.3.1, this is equivalent to prove:

$$\dim V_{\hat{L}} = \dim T_{\hat{L}}$$

or equivalently using Proposition 6.3.1, it remains to show:

$$\dim\left(\frac{\mathbb{R}[\underline{X}]_{d+1}}{U_{\hat{L}}}\right) = \dim\left(\frac{\mathbb{R}[\underline{X}]_d}{U_{\hat{L}} \cap \mathbb{R}[\underline{X}]_d}\right)$$

Since $U_{\hat{L}} \cap \mathbb{R}[\underline{X}]_d = U_L \cap \mathbb{R}[\underline{X}]_d$ and using again Proposition 6.3.1, we have the following:

$$\dim\left(\frac{\mathbb{R}[\underline{X}]_d}{U_{\hat{L}} \cap \mathbb{R}[\underline{X}]_d}\right) = \dim\left(\frac{\mathbb{R}[\underline{X}]_d}{U_L \cap \mathbb{R}[\underline{X}]_d}\right) \overset{6.3.1}{=} \dim(T_L) = N$$

On the other side, we will prove $\dim(\frac{\mathbb{R}[\underline{X}]_{d+1}}{U_{\hat{L}}}) = N$. For this purpose, let us consider the following linear map, between euclidean vector spaces:

$$\frac{\mathbb{R}[\underline{X}]_{d+1}}{U_{\hat{L}}} \hookrightarrow \frac{\mathbb{R}[\underline{X}]}{U_\Lambda}, \overline{p}^{\hat{L}} \mapsto \overline{p}^{\Lambda} \tag{6.21}$$

where $\Lambda := \sum_{i=1}^{N} \lambda_i \operatorname{ev}_{a_i} \in \mathbb{R}[\underline{X}]^*$. Notice that the canonical map (6.21) is well defined since $U_{\hat{L}} = U_\Lambda \cap \mathbb{R}[\underline{X}]_{d+1}$ and therefore it is injective. Then

$$\dim\left(\frac{\mathbb{R}[\underline{X}]_{d+1}}{U_{\hat{L}}}\right) \leq \dim\left(\frac{\mathbb{R}[\underline{X}]}{U_\Lambda}\right) \overset{6.3.8}{=} N$$

It remains to show $N \leq \dim(\frac{\mathbb{R}[\underline{X}]_{d+1}}{U_{\hat{L}}})$. But this is true, since:

$$N = \dim(T_L) \overset{6.3.1}{=} \dim\left(\frac{\mathbb{R}[\underline{X}]_d}{U_L \cap \mathbb{R}[\underline{X}]_d}\right) = \dim\left(\frac{\mathbb{R}[\underline{X}]_d}{U_{\hat{L}} \cap \mathbb{R}[\underline{X}]_d}\right) \leq \dim\left(\frac{\mathbb{R}[\underline{X}]_{d+1}}{U_{\hat{L}}}\right)$$

(ii) $\Longrightarrow$ (i). Since $\hat{L}$ is flat, then by Proposition 6.3.19 we know that the truncated GNS multiplication operators of $\hat{L}$ pairwise commute. Then by applying again Theorem 6.3.14 there exists $a_1, \dots, a_N \in \mathbb{R}^n$, pairwise different nodes, and $\lambda_1 > 0, \dots, \lambda_N > 0$ weights, with $N = \dim(T_{\hat{L}})$ such that if we set $\Lambda := \sum_{i=1}^{N} \lambda_i \operatorname{ev}_{a_i} \in \mathbb{R}[\underline{X}]^*$, we get $\Lambda(p) = \hat{L}(p) = L(p)$ for all $p \in \mathbb{R}[\underline{X}]_{2d+1}$, and $U_L \subseteq U_{\hat{L}} \subseteq U_\Lambda$. Indeed notice that $U_L \subseteq U_{\hat{L}}$ since for $p \in U_L$, $L(pq) = \hat{L}(pq) = 0$ for all $q \in \mathbb{R}[\underline{X}]_d$, and since $\hat{L}$ is flat this implies $p \in U_{\hat{L}}$. Obviously $M_{\Lambda,i}$ pairwise commute for all $i \in \{1, \dots, n\}$, since they are the original GNS operators modulo U_Λ defined in 6.1.6. In order to prove that $M_{L,i}$ pairwise commute for all $i \in \{1, \dots, n\}$, let us first consider the linear isometry σ_1 (6.12) of the Proposition 6.3.11. We proved already that σ_1 is a linear isometry then:

$$\dim(T_L) \leq \dim(\frac{\mathbb{R}[\underline{X}]}{U_\Lambda})$$

Therefore we have the following inequalities:

$$N = \dim(T_{\hat{L}}) = \dim\left(\frac{\mathbb{R}[\underline{X}]_d}{U_{\hat{L}} \cap \mathbb{R}[\underline{X}]_d}\right) = \dim\left(\frac{\mathbb{R}[\underline{X}]_d}{U_L \cap \mathbb{R}[\underline{X}]_d}\right) =$$
$$\dim(T_L) \leq \dim\left(\frac{\mathbb{R}[\underline{X}]}{U_\Lambda}\right) \overset{6.3.8}{=} N$$

then we have equality everywhere and $\dim(T_L) = \dim(\frac{\mathbb{R}[\underline{X}]}{U_\Lambda})$. Then σ_1 in this particular case is surjective and in conclusion is an isomorphism of vector spaces. This last fact will allow us to prove that the following diagram is commutative for all $i \in \{1, \dots, n\}$:

(6.22)
$$\begin{array}{ccc} T_L & \xrightarrow{M_{L,i}} & T_L \\ \downarrow{\scriptstyle \sigma_1} & & \uparrow{\scriptstyle \sigma_1^{-1}} \\ \frac{\mathbb{R}[\underline{X}]}{U_\Lambda} & \xrightarrow{M_{\Lambda,i}} & \frac{\mathbb{R}[\underline{X}]}{U_\Lambda} \end{array}$$

That is to say $M_{L,i} = \sigma_1^{-1} \circ M_{\Lambda,i} \circ \sigma_1$. To show this let $p, q \in \mathbb{R}[\underline{X}]_d$, then we have:

$$\left\langle M_{L,i}(\overline{p}^L), \overline{q}^L \right\rangle_L = \left\langle \Pi_L(\overline{X_i p}^L), \overline{q}^L \right\rangle_L \overset{\Pi_L \circ \Pi_{\underline{L}} = \Pi_L}{=} \left\langle \overline{X_i p}^L, \overline{q}^L \right\rangle_L = L(X_i pq) =$$
$$\Lambda(X_i pq) = \left\langle \overline{X_i p}^\Lambda, \overline{q}^\Lambda \right\rangle_\Lambda = \left\langle \sigma_1 \circ \sigma_1^{-1}(\overline{X_i p}^\Lambda), \overline{q}^\Lambda \right\rangle_\Lambda = \left\langle \sigma_1^{-1}(\overline{X_i p}^\Lambda), \sigma_1^{-1}(\overline{q}^\Lambda) \right\rangle_L =$$
$$\left\langle \sigma_1^{-1} \circ M_{\Lambda,i}(\overline{p}^\Lambda), \overline{q}^L \right\rangle_L = \left\langle \sigma_1^{-1} \circ M_{\Lambda,i} \circ \sigma_1(\overline{p}^L), \overline{q}^L \right\rangle_L$$

Finally we can conclude that the truncated GNS multiplication operators of L pairwise commute, using the commutativity of the GNS multiplication operators of Λ. Indeed:

$$M_{L,i} \circ M_{L,j} = \sigma_1^{-1} \circ M_{\Lambda,i} \circ \sigma_1 \circ \sigma_1^{-1} \circ M_{\Lambda,j} \circ \sigma_1 = \sigma_1^{-1} \circ M_{\Lambda,i} \circ M_{\Lambda,j} \circ \sigma_1 = \sigma_1^{-1} \circ M_{\Lambda,j} \circ M_{\Lambda,i} \circ \sigma_1 = \sigma_1^{-1} \circ M_{\Lambda,j} \circ \sigma_1 \circ \sigma_1^{-1} \circ M_{\Lambda,i} \circ \sigma_1 = M_{L,j} \circ M_{L,i} \tag{6.23}$$

(ii) $\Longrightarrow$ (iii). Due to $\hat{L} = L$ on $\mathbb{R}[\underline{X}]_{2d+1}$ the moment matrix $M_{\hat{L}}$ is of the form:

$$M_{\hat{L}} = \left(\begin{array}{c|c} M_{L'} & M_{L'}W \\ \hline W^T M_{L'} & W^T M_{L'} W + XX^T \end{array}\right) \succeq 0 \tag{6.24}$$

for some matrices $W \in \mathbb{R}^{s_d \times r_{d+1}}$ and $X \in \mathbb{R}^{r_{d+1} \times r_{d+1}}$ which make the upper left block of $M_{\hat{L}}$ the to be the same as the upper left block of M_L. Hence $M_{\hat{L}} = \widetilde{M_L}$, and since $M_{\hat{L}}$ is obviously generalized Hankel, $\widetilde{M_L}$ is generalized Hankel as well.

(iii) $\Longrightarrow$ (ii). Define the following linear form $\hat{L} \in \mathbb{R}[\underline{X}]_{2d+2}^*$ such that for all $p, q \in \mathbb{R}[\underline{X}]_{d+1}$:

$$\hat{L}(pq) := \operatorname{poly}^{-1}(p)^T \widetilde{M_L} \operatorname{poly}^{-1}(q)$$

It is obviously well defined since $\widetilde{M_L}$ is generalized Hankel, and it is positive semidefinite since $\widetilde{M_L}$ is always positive semidefinite, see Lemma 6.2.1 for a proof. $\hat{L}$ is flat since its moment matrix is $M_{\hat{L}} = \widetilde{M_L}$ which is flat.

Let us add the prove of the implication (iii) $\Longrightarrow$ (i) even if we already know that is true, since it helps us to understand better the Theorem.

(iii) $\Longrightarrow$ (i) Suppose $\widetilde{M_L}$ is a Generalized Hankel matrix, and set:

$$\hat{L} := L_{\widetilde{M_L}} \in \mathbb{R}[\underline{X}]_{2d+2}^*.$$

Note that $\widetilde{M_L}$ is the moment matrix of the linear form $\hat{L} \in \mathbb{R}[\underline{X}]_{2d+2}^*$ since by 4.2.6:

$$\widetilde{M_L} = M_{L_{\widetilde{M_L}}}.$$

Hence $\hat{L} = L$ on $\mathbb{R}[\underline{X}]_{2d+1}$ and $\hat{L}$ is flat and by Theorem 6.3.19 the truncated GNS multiplication operators of $\hat{L}$ commute. Now, in order to prove that the truncated GNS operators of L commute let us define:

$$\sigma := \sigma_{\hat{L}}^{-1} \circ \sigma_L$$

where $\sigma_{\hat{L}}$ and σ_L were defined in (6.9) and σ is an isomorphism of euclidean vector spaces. Note that since:

$$\frac{\mathbb{R}[\underline{X}]_d}{U_L \cap \mathbb{R}[\underline{X}]_d} = \frac{\mathbb{R}[\underline{X}]_d}{U_{\hat{L}} \cap \mathbb{R}[\underline{X}]_d}$$

σ is well defined. Now we will prove that the following diagram is commutative:

$$\begin{array}{ccc} T_L & \xrightarrow{M_{L,i}} & T_L \\ \Big\downarrow \sigma & & \Big\uparrow \sigma^{-1} \\ T_{\hat{L}} & \xrightarrow{M_{\hat{L},i}} & T_{\hat{L}} \end{array}$$

The diagram is commutative if and only if $M_{L,i} = \sigma \circ M_{\hat{L},i} \circ \sigma^{-1}$. To prove this equality let us take $p, q \in \mathbb{R}[\underline{X}]_d$. Then:

$$\left\langle M_{L,i}(\overline{p}^L), \overline{q}^L \right\rangle_L = \left\langle \Pi_L(\overline{X_i p}^L), \overline{q}^L \right\rangle_L \overset{\Pi_L \circ \Pi_{\underline{L}} = \Pi_L}{=} \left\langle \overline{X_i p}^L, \overline{q}^L \right\rangle_L = L(X_i pq) =$$

$$\hat{L}(X_i pq) = \left\langle \overline{X_i p}^{\hat{L}}, \overline{q}^{\hat{L}} \right\rangle_{\hat{L}} = \left\langle \overline{X_i p}^{\hat{L}}, \Pi_{\hat{L}}(\overline{q}^{\hat{L}}) \right\rangle_{\hat{L}} = \left\langle \Pi_{\hat{L}}(\overline{X_i p}^{\hat{L}}), \overline{q}^{\hat{L}}, \right\rangle_{\hat{L}} =$$

$$\left\langle \sigma \circ \sigma^{-1}(\Pi_{\hat{L}}(\overline{X_i p}^{\hat{L}})), \overline{q}^{\hat{L}} \right\rangle_{\hat{L}} = \left\langle \sigma^{-1}(\Pi_{\hat{L}}(\overline{X_i p}^{\hat{L}})), \sigma^{-1}(\overline{q}^{\hat{L}}) \right\rangle_L =$$

$$\left\langle \sigma^{-1} \circ M_{\hat{L},i}(\overline{p}^{\hat{L}}), \overline{q}^L \right\rangle_L = \left\langle \sigma^{-1} \circ M_{\hat{L},i} \circ \sigma(\overline{p}^L), \overline{q}^L \right\rangle_L$$

Finally we can conclude the truncated GNS multiplication operators of L commute using the commutativity of the truncated GNS multiplication operators of $\hat{L}$ in the identical way we already did in the implication (ii) $\implies$ (i) in (6.23).

Note that the uniqueness of the linear form $\hat{L}$ in the second statement is determined because $\hat{L}$ is flat and $\hat{L} = L$ on $\mathbb{R}[\underline{X}]_{2d+1}$. Indeed, the modified moment matrix of L is of the form:

$$\widetilde{M_L} = \left(\begin{array}{c|c} M_{L'} & M_{L'} \\ \hline W^T M_{L'} & W^T M_{L'} W + X^T X \end{array} \right)$$

for some matrices $W \in \mathbb{R}^{s_d \times r_{d+1}}$ and $X \in \mathbb{R}^{r_{d+1} \times r_{d+1}}$. Since $\hat{L} = L$ on $\mathbb{R}[\underline{X}]_{2d+1}$ and $\hat{L}$ is flat its moment matrix it is uniquely determined, that is to say has the form:

$$M_{\hat{L}} = \left(\begin{array}{c|c} M_{L'} & M_{L'} W \\ \hline W^T M_{L'} & W^T M_{L'} W \end{array} \right)$$

and therefore $\hat{L}$ is unique.

□

Remark 6.5.2. Note that the linear form $\hat{L}$ that we construct in the implication $(i) \implies (ii)$ using Theorem 6.3.14 is flat.

6.6 A result of Xu and Mysovskikh for positive semidefinite linear forms

The following result is a cororollary of Theorem 6.5.1 and give us a generalization of a classical Theorem from Mysovskikh [26], Dunkl and Xu [10, Theorem 3.8.7] and Putinar [29, pages 189-190]. They proved the equivalence between the existence of a minimal gaussian quadrature rule with the commutativity of the truncated GNS multiplication operators for a positive definite linear form on $\mathbb{R}[\underline{X}]$. The generalization here comes from the fact that the result holds also if the linear form is defined on $\mathbb{R}[\underline{X}]_{2d+2}$ for $d \in \mathbb{N}_0$ and it is positive semidefinite i.e. we do not assume $U_L = \{0\}$. We also provide a third equivalent condition in the result which is $\widetilde{M_L}$ is a generalized Hankel matrix, a fact which seems no to have been noticed so far.

Corollary 6.6.1. The following assertions are equivalent:

(i) The linear form L admits a Gaussian quadrature rule.

(ii) The truncated GNS multiplication operators of L commute.

(iii) $\widetilde{M_L}$ is a generalized Hankel matrix.

Proof. (i) $\Longrightarrow$ (ii). Assume that L admits a Gaussian quadrature rule, that is to say:

$$L(p) = \sum_{i=1}^{N} \lambda_i p(a_i) \text{ for all } p \in \mathbb{R}[\underline{X}]_{2d+1}$$

where $N := \dim(T_L)$, $\lambda_1 > 0, \ldots, \lambda_N > 0$ and the points $a_1, \ldots, a_N \in \mathbb{R}^n$ are pairwise different. Let us set $\Lambda := \sum_{i=1}^{N} \lambda_i \operatorname{ev}_{a_i} \in \mathbb{R}[\underline{X}]^*$. Using 6.3.8 we have the following:

$$\dim\left(\frac{\mathbb{R}[\underline{X}]}{U_\Lambda}\right) = N = \dim T_L$$

Let us consider the following canonical map σ_2:

$$\sigma_2 : \frac{\mathbb{R}[\underline{X}]_d}{U_L \cap \mathbb{R}[\underline{X}]_d} \hookrightarrow \frac{\mathbb{R}[\underline{X}]}{U_\Lambda}$$

By Proposition 6.3.1:

$$\dim\left(\frac{\mathbb{R}[\underline{X}]_d}{U_L \cap \mathbb{R}[\underline{X}]_d}\right) = \dim(T_L)$$

Therefore the canonical map σ_2, is an isomorphism of euclidean vector spaces. Now, define the following isomorphism of euclidean vector spaces:

$$\beta := \sigma_2 \circ \sigma_L$$

where σ_L was defined in (6.9). The proof continues in the same way we proved (ii) $\Longrightarrow$ (i) of Theorem 6.5.1 by proving the following diagram is commutative:

$$\begin{array}{ccc} T_L & \xrightarrow{M_{L,i}} & T_L \\ \downarrow \beta & & \uparrow \beta^{-1} \\ \frac{\mathbb{R}[\underline{X}]}{U_\Lambda} & \xrightarrow{M_{\Lambda,i}} & \frac{\mathbb{R}[\underline{X}]}{U_\Lambda} \end{array}$$

Notice that here β does the job of σ_1 in (ii) $\Longrightarrow$ (i) in Theorem 6.5.1.

$(ii) \Longrightarrow (i)$. It is a particular case of (i) $\Longrightarrow$ (ii) in 6.5.1.

$(ii) \iff (iii)$. It is (i) $\iff$ (iii) in 6.5.1. □

6.7 A result of Möller for positive semidefinite linear forms

The following result of Möller will give us a better lower bound in the number of nodes of a quadrature rule on $\mathbb{R}[\underline{X}]_{2d+1}$ than the very well-known bound given in Proposition 6.3.11. This bound, was already found for positive linear forms by Möller in 1975 and by Putinar in 1997 ([23],[29]). This result will show that the bound is also true for positive semidefinite linear forms and it uses the same ideas as in [29]. We include the proof for the convenience of the reader. This bound will help us in polynomial optimization problems in which we know the number of global minimizers in advance, to discard optimality if this bound is bigger than the number of global minimizers, see Example 6.7.3 below.

Theorem 6.7.1. The number of nodes N of a quadrature rule for L on $\mathbb{R}[\underline{X}]_{2d+1}$ satisfies:

$$N \geq \dim(T_L) + \frac{1}{2} \max_{1\leq j,k\leq n} (\operatorname{rank}[M_{L,j}, M_{L,k}]) \tag{6.25}$$

Proof. Assume L has a quadrature rule on $\mathbb{R}[\underline{X}]_{2d+1}$ with N nodes, that is to say, there exist pairwise different nodes $a_1, \ldots, a_N \in \mathbb{R}^n$ and $\lambda_1 > 0, \ldots, \lambda_N > 0$ weights, such that:

$$L(p) = \sum_{i=1}^{N} \lambda_i p(a_i) \text{ for all } p \in \mathbb{R}[\underline{X}]_{2d+1}.$$

Let us set $\Lambda := \sum_{i=1}^{N} \lambda_i \operatorname{ev}_{a_i} \in \mathbb{R}[\underline{X}]^*$. By using the proposition 6.3.8 we have that:

$$\dim\left(\frac{\mathbb{R}[\underline{X}]}{U_\Lambda}\right) = N < \infty.$$

Then we can choose an orthonormal basis of $\frac{\mathbb{R}[\underline{X}]}{U_\Lambda}$. Let us denote such a basis by:

$$\beta_\Lambda := \{\overline{\beta_1}^\Lambda, \ldots, \overline{\beta_N}^\Lambda\}$$

for $\beta_1, \ldots, \beta_N \in \mathbb{R}[\underline{X}]$ pairwise different. Then we have that the transformation matrix of the multiplication operators $M_{\Lambda,i}$ with respect to this orthonormal basis is:

$$(\Lambda(X_i\beta_k\beta_j))_{1\le k,j\le N}$$

The set $A := \{\, \overline{p}^{\Lambda} \mid p \in \mathbb{R}[\underline{X}]_d\}$ is a subspace of $\frac{\mathbb{R}[\underline{X}]}{U_\Lambda}$ so we can assume without loss of generality that $\beta_1, \ldots, \beta_r \in \mathbb{R}[\underline{X}]_d$ generate a basis of A, where $r := \dim A$. Then since $L = \Lambda$ on $\mathbb{R}[\underline{X}]_{2d+1}$, we obtain:

$$(M(M_{\Lambda,i}, \beta_\Lambda)) := (\Lambda(X_i\beta_k\beta_j))_{1\le k,j\le N} = \left(\begin{array}{c|c} (L(X_i\beta_k\beta_j))_{1\le k,j\le r} & B_i \\ \hline B_i^t & C_i \end{array}\right) \tag{6.26}$$

where $B_i \in \mathbb{R}^{r\times N-r}$ and $C_i \in \mathbb{R}^{N-r\times N-r}$ are symmetric matrices. We will show that:

$$\beta_L := \{\overline{\beta_1}^L, \ldots, \overline{\beta_r}^L\}$$

is an orthonormal basis of T_L. To this end, note that by Proposition 6.3.11, $\dim T_L \ge N$, and then by Proposition 6.3.1 it also holds that:

$$\dim\left(\frac{\mathbb{R}[\underline{X}]_d}{U_L \cap \mathbb{R}[\underline{X}]_d}\right) \ge N$$

This last inequality together with Proposition 6.3.8 implies that the following canonical map:

$$\sigma_2 : \frac{\mathbb{R}[\underline{X}]_d}{U_L \cap \mathbb{R}[\underline{X}]_d} \hookrightarrow \frac{\mathbb{R}[\underline{X}]}{U_\Lambda}$$

is an isomorphism of euclidean vector spaces. Then we have that:

$$\sigma := \sigma_2 \circ \sigma_L$$

is also an isomorphism of euclidean vector spaces, where σ_L was defined in (6.9). And then:

$$\sigma(T_L) = \{\sigma(\overline{p}^L) \mid \overline{p}^L \in T_L\} = \{\sigma(\overline{p}^L) \mid p \in \mathbb{R}[\underline{X}]_d\} = \{\, \overline{p}^{\Lambda} \mid p \in \mathbb{R}[\underline{X}]_d\} = A$$

Hence, we get:

$$\sigma(T_L) = \langle \overline{\beta_1}^{\Lambda}, \ldots, \overline{\beta_r}^{\Lambda}\rangle$$

And since we have chosen $\beta_i \in \mathbb{R}[\underline{X}]_d$ for all $i \in \{1, \ldots, r\}$, then we have $\sigma(\overline{\beta_i}^L) = \overline{\beta_i}^{\Lambda}$. Therefore $\beta_L := \{\overline{\beta_1}^L, \ldots, \overline{\beta_r}^L\}$ generate a basis of T_L. It remains to show that β_L is orthonormal. To see that β_L is orthonormal we use the fact that σ is an isometry and that $\sigma(T_L) = A$. Indeed for $1 \le i, j \le r$:

$$\delta_{ij} = \Lambda(\beta_i\beta_j) = \langle \overline{\beta_j}^{\Lambda}, \overline{\beta_i}^{\Lambda}\rangle_\Lambda = \langle \sigma^{-1}(\overline{\beta_j}^{\Lambda}), \sigma^{-1}(\overline{\beta_i}^{\Lambda})\rangle_L = \langle \overline{\beta_j}^L, \overline{\beta_i}^L\rangle_L$$

Therefore we have shown that:

$$M(M_{\Lambda,i},\beta_\Lambda) = \left(\begin{array}{c|c} M(M_{L,i},\beta_L) & B_i \\ \hline B_i^T & C_i \end{array}\right)$$

where :

$$(M(M_{L,i},\beta_L)) := (L(X_i\beta_k\beta_j))_{1\leq k,j\leq r}$$

Using the fact that the matrices $M(M_{\Lambda,i},\beta_\Lambda)$ commute, we have the following equality:

$$M(M_{L,j},\beta_L)M(M_{L,i},\beta_L) - M(M_{L,i},\beta_L)M(M_{L,j},\beta_L) = B_iB_j^T - B_jB_i^T$$

Therefore the following holds:

$$\operatorname{rank}(B_iB_j^T - B_jB_i^T) \leq 2\operatorname{rank}(B_iB_j^T) \leq 2\operatorname{rank}(B_i) \leq 2\min\{r, N-r\} \leq 2(N-r)$$

and then:

$$\operatorname{rank}[M(A_{L,j},\beta_L), M(A_{L,i},\beta_L)] \leq 2(N-r)$$

Since we have already proved $r = \dim T_L$, we can conclude the proof. □

Remark 6.7.2. Note that we can use the previous Theorem 6.7.1 to show in a different way (i) $\Longrightarrow$ (ii) in the Corollary 6.6.1. Indeed, let us suppose that L has a Gaussian quadrature rule that is to say with $N = \dim T_L$ nodes. Using the inequality of Theorem 6.7.1 we get that $\operatorname{rank}[M_{L,j}, M_{L,k}] = 0$ for $j,k \in \{1,\ldots,n\}$, therefore the truncated GNS multiplication operators of L commute.

Example 6.7.3. Let us consider the following polynomial optimization problem taken from [16]:

$$\begin{array}{ll} \text{minimize} & f(x) = x_1^2x_2^2(x_1^2+x_2^2-1) \\ \text{subject to} & x_1, x_2 \in \mathbb{R} \end{array}$$

We know that the minimizers of f occur in the real points common to the partial derivatives of f (the real gradient variety) and we can easily check that the zero sets of these derivatives intersect in 4 real points: $\left(\pm\frac{1}{\sqrt{3}}, \pm\frac{1}{\sqrt{3}}\right) \in \mathbb{R}^2$. Therefore we know in advance that (P) has at most 4 minimizers. On the other side, an optimal solution of the Moment relaxation of order 8 (P_8), that is $\mathbf{M} := M_{8,1}(y)$ read as:

$$\left(\begin{matrix}
1.00 & 0.00 & 0.00 & 62.12 & -0.00 & 62.12 & 0.00 & 0.00 & 0.00 & 0.00 \\
0.00 & 62.12 & -0.00 & 0.00 & 0.00 & 0.00 & 9666.23 & -0.00 & 8.33 & -0.00 \\
0.00 & -0.00 & 62.12 & 0.00 & 0.00 & 0.00 & -0.00 & 8.33 & -0.00 & 9666.23 \\
62.12 & 0.00 & 0.00 & 9666.23 & -0.00 & 8.33 & 0.00 & -0.00 & 0.00 & 0.00 \\
-0.00 & 0.00 & 0.00 & -0.00 & 8.33 & -0.00 & -0.00 & 0.00 & 0.00 & -0.00 \\
62.12 & 0.00 & 0.00 & 8.33 & -0.00 & 9666.23 & 0.00 & 0.00 & -0.00 & 0.00 \\
0.00 & 9666.23 & -0.00 & 0.00 & -0.00 & 0.00 & 3150633.17 & -0.00 & 2.27 & 0.00 \\
0.00 & -0.00 & 8.33 & -0.00 & 0.00 & 0.00 & -0.00 & 2.27 & 0.00 & 2.27 \\
0.00 & 8.33 & -0.00 & 0.00 & 0.00 & -0.00 & 2.27 & 0.00 & 2.27 & -0.00 \\
0.00 & -0.00 & 9666.23 & 0.00 & -0.00 & 0.00 & 0.00 & 2.27 & -0.00 & 3150630.69 \\
9666.23 & 0.00 & -0.00 & 3150633.17 & -0.00 & 2.27 & 0.42 & -0.00 & -0.00 & 0.00 \\
-0.00 & -0.00 & 0.00 & -0.00 & 2.27 & 0.00 & -0.00 & -0.00 & 0.00 & 0.00 \\
8.33 & 0.00 & 0.00 & 2.27 & 0.00 & 2.27 & -0.00 & 0.00 & 0.00 & -0.00 \\
-0.00 & 0.00 & -0.00 & 0.00 & 2.27 & -0.00 & 0.00 & 0.00 & -0.00 & -0.00 \\
9666.23 & -0.00 & 0.00 & 2.27 & -0.00 & 3150630.69 & 0.00 & -0.00 & -0.00 & 0.33
\end{matrix}\right.$$

$$\left.\begin{matrix}
9666.23 & -0.00 & 8.33 & -0.00 & 9666.23 \\
0.00 & -0.00 & 0.00 & 0.00 & -0.00 \\
-0.00 & 0.00 & 0.00 & -0.00 & 0.00 \\
3150633.17 & -0.00 & 2.27 & 0.00 & 2.27 \\
-0.00 & 2.27 & 0.00 & 2.27 & -0.00 \\
2.27 & 0.00 & 2.27 & -0.00 & 3150630.69 \\
0.42 & -0.00 & -0.00 & 0.00 & 0.00 \\
-0.00 & -0.00 & 0.00 & 0.00 & -0.00 \\
-0.00 & 0.00 & 0.00 & -0.00 & -0.00 \\
0.00 & 0.00 & -0.00 & -0.00 & 0.33 \\
2466755083.36 & -43.48 & 169698627.89 & -6.08 & 134568970.57 \\
-43.48 & 169698627.89 & -6.08 & 134568970.57 & 15.08 \\
169698627.89 & -6.08 & 134568970.57 & 15.08 & 169698562.66 \\
-6.08 & 134568970.57 & 15.08 & 169698562.66 & 25.61 \\
134568970.57 & 15.08 & 169698562.66 & 25.61 & 2466752654.76
\end{matrix}\right)$$

and the rank of the commutator of the truncated GNS multiplication operators is:

$$\operatorname{rank}[M_{\mathbf{M},X_1}, M_{\mathbf{M},X_1}] =$$

$$\operatorname{rank}\begin{pmatrix}
0 & 0.00 & 0.00 & -0.00 & -0.00 & 0.00 & 0.00 & -0.00 & 0.00 & -0.00 \\
-0.00 & 0 & -0.00 & 0.00 & 0.00 & 0.00 & -0.00 & 0.00 & 0.00 & -0.00 \\
-0.00 & 0.00 & 0 & 0.00 & 0.00 & 0.00 & 0.00 & -0.00 & 0.00 & 0.00 \\
0.00 & -0.00 & -0.00 & 0 & -0.00 & -0.00 & -0.00 & -0.00 & 0.00 & -0.00 \\
0.00 & -0.00 & -0.00 & 0.00 & 0 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\
-0.00 & -0.00 & -0.00 & 0.00 & -0.00 & 0 & -0.00 & -0.00 & 0.00 & -0.00 \\
-0.00 & 0.00 & -0.00 & 0.00 & -0.00 & 0.00 & 0 & -313.91 & -0.00 & 115.50 \\
0.00 & -0.00 & 0.00 & 0.00 & -0.00 & 0.00 & 313.91 & 0 & 0.18 & 0.00 \\
-0.00 & -0.00 & -0.00 & -0.00 & -0.00 & -0.00 & 0.00 & -0.18 & 0 & 0.05 \\
0.00 & 0.00 & -0.00 & 0.00 & -0.00 & 0.00 & -115.50 & -0.00 & -0.05 & 0
\end{pmatrix} = 4$$

If $\mathbf{M}$ had a quadrature rule on $\mathbb{R}[X_1, X_2]_7$ with N nodes, since $f \in \mathbb{R}[X_1, X_2]_7$ and by 4.4.4 (iii), the N nodes of the quadrature rule would be global minimizers of f and $P^* = P_8^*$, and according to Theorem 6.7.1:

$$N \geq \dim(T_L) + \frac{1}{2} \max_{1 \leq j,k \leq n} (\operatorname{rank}[M_{L,j}, M_{L,k}]) = 10 + \frac{1}{2} 4 = 12$$

Therefore the polynomial f would have at least 12 global minimizers and this is a contradiction with the fact that f has at most 4 global minimizers. Notice that then $\mathbf{M}$ does not have a quadrature rule on $\mathbb{R}[\underline{X}]_7$, and in particular it does not have a quadrature rule.

6.8 A generalized Smul'jan result for positive semidefinite linear forms

Theorem 6.8.1. Let $d \in \mathbb{N}_0$ and $L \in \mathbb{R}[\underline{X}]_{2d}^*$ and $L(\sum \mathbb{R}[\underline{X}]_d^2) \subseteq \mathbb{R}_{\geq 0}$ with $L(1) = 1$. Then the truncated multiplication operators commute if and only if there exist $a_1, ..., a_r \in \mathbb{R}^n$ nodes and $\lambda_1, ..., \lambda_r \in \mathbb{R}_{>0}$ weights with $\sum_{i=1}^r \lambda_i = 1$ where $r := \dim T_L$ and $L_\infty \in \mathbb{R}[\underline{X}]_{2d}^*$ such that:

$$L = L_0 + L_\infty$$

where:

$$\begin{cases} L_0 := \sum_{i=1}^r \lambda_i \operatorname{ev}_{a_i} \in \mathbb{R}[\underline{X}]_{2d}^* \text{ is flat} \\ L_\infty(G_L) = 0 \text{ and} \\ L_\infty(\sum \mathbb{R}[\underline{X}]_d^2) \subseteq \mathbb{R}_{\geq 0} \end{cases}$$

Moreover, $L(q) \geq L_0(q)$ for all $q \in W_d$, where:

$$W_d := \{\sigma + p \mid \sigma \in \sum \mathbb{R}[\underline{X}]_{=d}^2, p \in \mathbb{R}[\underline{X}]_{2d-1}\} \tag{6.27}$$

Proof. $\Rightarrow$. Suppose the truncated multiplication operators commute by Theorem 6.3.14 there are $a_1, \ldots, a_r \in \mathbb{R}^n$ nodes with $r := \dim T_L$ and $\lambda_1, \ldots, \lambda_r \in \mathbb{R}_{>0}$ weights such that:

$$L(p) = \sum_{i=1}^r \lambda_i p(a_i) \text{ for all } p \in G_L \text{ (6.15)}$$

and by Theorem 6.5.1 L_0 is flat. For every $\alpha \in \mathbb{R}^n$ with $|\alpha| \leq 2d$ set $\lambda_\alpha := L(\underline{X}^\alpha) - L_0(\underline{X}^\alpha)$ and define:

$$L_\infty := \sum_{\alpha:\, |\alpha|=2d} \lambda_\alpha \operatorname{d}_\alpha \in \mathbb{R}[\underline{X}]_{2d}^*$$

where for every $\alpha \in \mathbb{N}^n$, d_α is the following linear form:

$$\begin{aligned} \operatorname{d}_\alpha : \mathbb{R}[\underline{X}]_{2d} &\longrightarrow \mathbb{R} \\ \underline{X}^\beta &\longmapsto \begin{cases} 1 \text{ if } \alpha = \beta \\ 0 \text{ otherwise} \end{cases} \end{aligned}$$

or in other words, $(\operatorname{d}_\alpha)_{\alpha \in \mathbb{N}^n}$ is the dual basis of $(\underline{X}^\alpha)_{\alpha \in \mathbb{N}^n}$. It also holds $L = L_0 + L_\infty$. Indeed, for $|\beta| = 2d - 1$ we have that:

$$L_0(\underline{X}^\beta) + L_\infty(\underline{X}^\beta) = L_0(\underline{X}^\beta) + 0 = L(\underline{X}^\beta)$$

And for $|\beta| = 2d$ it holds:

$$\begin{aligned} L_0(\underline{X}^\beta) + L_\infty(\underline{X}^\beta) = L_0(\underline{X}^\beta) + \sum_{|\alpha|=2d} \lambda_\alpha \operatorname{d}_\alpha(\underline{X}^\beta) = \\ L_0(\underline{X}^\beta) + \lambda_\beta = L_0(\underline{X}^\beta) + (L(\underline{X}^\beta) - L_0(\underline{X}^\beta)) = L(\underline{X}^\beta) \end{aligned}$$

Hence by linearity we have $L = L_0 + L_\infty$. It also holds $L_\infty(G_L) = 0$. Indeed, for every $p \in \mathbb{R}[\underline{X}]_d$, $q \in \mathbb{R}[\underline{X}]_{d-1}$ and $g \in U_L$:

$$\begin{aligned} L(pq) = L(p(q+g)) = L_0(p(q+g)) + L_\infty(p(q+g)) = \\ L_0(pq) + L_0(pg) + L_\infty(pq) + L_\infty(pg) = L(pq) + L_\infty(pg) \end{aligned}$$

Then $L_\infty(p(g+q)) = 0$ and by linearity $L_\infty(G_L) = 0$. To finish the first implication it remains to show that $L_\infty(\sum \mathbb{R}[\underline{X}]_d^2) \subseteq \mathbb{R}_{\geq 0}$. By Lemma 6.2.1, we have the following equatity:

$$M_L = \left(\begin{array}{c|c} M_{L'} & M_{L'}W \\ \hline W^T M_{L'} & W^T M_{L'} W + XX^T \end{array} \right)$$

for some matrices $W \in \mathbb{R}^{s_{d-1}\times r_d}$ and $X \in \mathbb{R}^{r_d\times r_d}$. Now, since L_0 is flat as it was point out in Remark 6.5.2, it must be:

$$M_{L_0} = \left(\begin{array}{c|c} M_{L'} & M_{L'}W \\ \hline W^T M_{L'} & W^T M_{L'} W \end{array}\right)$$

and by linearity:

$$M_{L_\infty} = \left(\begin{array}{c|c} 0 & 0 \\ \hline 0 & XX^T \end{array}\right)$$

Then we obviously have $M_{L_\infty} \succeq 0$. Note that this last fact together with Remark 4.2.8 implies that L_∞ is a positive semidefinite linear form.
$\Leftarrow$. Suppose now that $L = L_0 + L_\infty$ where L_0 has a quadrature rule representation and $L_\infty(\mathbb{R}[\underline{X}]_{2d-1}) = 0$. Then $\widetilde{M_L} = M_{L_0}$ which implies that $\widetilde{M_L}$ is a generalized Hankel matrix and therefore by the Theorem 6.5.1 we have that the truncated multiplication operators commute.

Let us show now that:

$$L(q) \geq L_0(q) \text{ for all } q \in W_d \text{ (6.27)}.$$

For this purpose by linearity it remains to show that $L(\sigma) \geq L_0(q)$ for $\sigma \in \mathbb{R}[\underline{X}]^2_{=d}$. For this write $\sigma = s^2$ with $s \in \mathbb{R}[\underline{X}]_{=d}$. It holds:

$$\begin{aligned} L(\sigma) = L(s^2) = \left\langle \bar{s}^L, \bar{s}^L \right\rangle_L \geq \left\langle \pi_L(\bar{s}^L), \pi_L(\bar{s}^L) \right\rangle_L &\overset{(6.14)}{=} \\ \left\langle s(M_1,\dots,M_n)(\bar{1}^L), s(M_1,\dots,M_n)(\bar{1}^L) \right\rangle_L &\overset{(*)}{=} \\ \sum_{i=1}^{r} \lambda_i s(a_i)^2 = L_0(s^2) = L_0(\sigma) \end{aligned}$$

The equality $(*)$ follows in the same way as in the proof of the Theorem 6.3.14. □

Example 6.8.2. Let us consider again the linear form $L \in \mathbb{R}[X_1, X_2]_4^*$ of Example 6.3.21 with associated moment matrix:

$$M_L = \begin{array}{c} \\ 1 \\ X_1 \\ X_2 \\ X_1^2 \\ X_1X_2 \\ X_2^2 \end{array} \begin{array}{c} \begin{array}{cccccc} 1 & X_1 & X_2 & X_1^2 & X_1X_2 & X_2^2 \end{array} \\ \left(\begin{array}{ccc|ccc} 1.0000 & 0.7175 & 1.4698 & 0.5149 & 1.0547 & 2.1604 \\ 0.7175 & 0.5149 & 1.0547 & 0.3694 & 0.7568 & 1.5502 \\ 1.4698 & 1.0547 & 2.1604 & 0.7568 & 1.5502 & 3.1755 \\ \hline 0.5149 & 0.3694 & 0.7568 & 0.2651 & 0.5430 & 1.1123 \\ 1.0547 & 0.7568 & 1.5502 & 0.5430 & 1.1123 & 2.2785 \\ 2.1604 & 1.5502 & 3.1755 & 1.1123 & 2.2785 & 8.7737 \end{array}\right) \end{array}$$

In this case, following the notation as in Theorem 6.8.1 it holds:

$$L_0 = \operatorname{ev}(\alpha, \beta) \in \mathbb{R}[X_1, X_2]_4^*$$

with $\alpha := M_L(1,2)$ and $\beta := M_L(1,3)$. And, for

$$A := \begin{array}{c} \\ X_1^2 \\ X_1X_2 \\ X_2^2 \end{array} \begin{array}{c} \begin{array}{ccc} X_1^2 & X_1X_2 & X_2^2 \end{array} \\ \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 4.1057 \end{pmatrix} \end{array}$$

it also holds according to Theorem 6.8.1:

$$M_L - M_{L_0} = M_{L_\infty} = \left(\begin{array}{c|c} 0 & 0 \\ \hline 0 & A \end{array}\right)$$

Notice that $\lambda_{(0,4)} = 4.1057$ and $\lambda_\alpha = 0$ in the other cases, for all $\alpha \in \mathbb{N}^2$ such that $|\alpha| = 4$. If we denote by p_h the homogenization of a polynomial to the space. Then for every $p \in \mathbb{R}[X_1, X_2]_4$ we have:

$$L(p) = p(\alpha,\beta) + \sum_{|\alpha|=4, \alpha\in\mathbb{N}^2} \lambda_\alpha \, \mathrm{d}_\alpha(p) = p_h(1,\alpha,\beta) + \sum_{|\alpha|=4,\alpha\in\mathbb{N}^2} \lambda_\alpha \, \mathrm{d}_{(0,\alpha)}(p_h)$$

Using the following notation:

$$\lambda_{(0,0,4)} = \lambda_{(0,4)} \text{ and } \lambda_\alpha = 0 \text{ in other cases, for all } \alpha \in \mathbb{N}^3 \text{ such that } |\alpha| = 4$$

And consider that for $\xi := (0, 0, \sqrt{\sqrt{4.1057}}) \in \mathbb{R}^3$ we have that:

$$\xi^\alpha = \lambda_\alpha \text{ for all } \alpha \in \mathbb{N}^3 \text{ with } |\alpha| = 4 \tag{6.28}$$

And also taking in consideration that:

$$p_h(\xi_0, \xi_1, \xi_2) = \mathrm{ev}_\xi(p_h) = \sum_{|\alpha|=4,\alpha\in\mathbb{N}^3} \xi^\alpha \, \mathrm{d}_\alpha(p_h)$$

It holds that:

$$L(p) = p_h(1,\alpha,\beta) + \sum_{|\alpha|=4,\alpha\in\mathbb{N}^3} \xi^\alpha \, \mathrm{d}_\alpha(p_h) = p_h(1,\alpha,\beta) + p_h(\xi_0,\xi_1,\xi_2)$$

that is to say, in this case L has what in [3, Definition 4.4] was called a *generalized quadrature rule* with one *regular node* $(1,\alpha,\beta)$ and with one *node at infinity* (ξ_0, ξ_1, ξ_2). In this case the Theorem 6.8.1 help us to find a generalized quadrature rule, however this procedure is not always optimal to find a generalized quadrature rule since the nodes at infinity can be more than one and the system (6.28) that we need to solve to find the nodes at infinity can be more difficult to solve or may not have a solution.

CHAPTER 7

Algorithm for extracting minimizers in polynomial optimization problems

As an application of all the previous results in this section we find a stopping criterion for the moment relaxation hierarchy, in other words, we find a condition on L, the optimal solution of the moment relaxation (P_d), such that $L(f) = P_d^* = P^*$. In this this case we also find potential global minimizers. In the paper [12] and [9] of Lasserre and Henrion, the stopping criterion was L to be flat and in this algorithm the stopping criterium is $\widetilde{M_L}$ being generalized Hankel, and as we have seen in the Proposition 6.3.19 and in Example 6.3.21 this condition is more general. It is important to point out that despite this condition is more general than being flat we can not ensure optimality until we check that the candidates to minimizers are inside to the basic closed semialgebraic set S, condition that it is always possible to ensure if the set S is a set described by linear polynomials, as we proved in Corollary 7.1.4 (in particular if the set S is $\mathbb{R}^n$) or we have flat extension of some degree on the optimal solution, that is to say $\operatorname{rank} M_d(y) = \operatorname{rank} M_s(y)$ for sufficient small s, see [17, Theorem 6.18] or [5, Theorem 1.6] for a proof. At the end of this chapter we summarize all these results in an algorithm with examples and also we illustrate polynomial optimization problems where this new stopping criterion allow us to conclude optimality even in case where the optimal solution is not flat: see Examples 6.3.21, 7.2.1 and 7.2.4.

7.1 Optimality conditions

Theorem 7.1.1. Let $f, p_1, \ldots, p_m \in \mathbb{R}[\underline{X}]_{2d}$ and L be an optimal solution of (P_{2d}). Suppose that $\widetilde{M_L}$ is a generalized Hankel matrix. Then L has a quadrature rule on G_L. Moreover, suppose the nodes of the quadrature rule lie on S (2.1) and $f \in W_d$ (set defined in (6.27)), then $L(f) = P^*$ and the nodes are global minimizers.

Proof. Since $\widetilde{M_L}$ is a generalized Hankel matrix by Corollary 6.6.1 and Theorem 6.3.14 there exists nodes $a_1, \ldots, a_N \in \mathbb{R}^n$ and weights $\lambda_1 > 0, \ldots, \lambda_N > 0$, where $N := \dim T_L$

such that:

$$L(p) = \sum_{i=1}^{N} \lambda_i p(a_i) \text{ for all } p \in G_L$$

Moreover if the nodes of this quadrature rule are contained in S and $f \in W_d$, set defined in (6.27), it holds:

$$\sum_{i=1}^{N} \lambda_i f(a_i) \overset{4.4.4(ii)}{\geq} P^* \overset{4.4.4(i)}{\geq} P_k^* \overset{\text{Theorem 6.8.1}}{\geq} \sum_{i=1}^{N} \lambda_i f(a_i)$$

and we can conclude by the proof of the Proposition 4.4.4 (iii) that $a_1, \dots, a_N \in \mathbb{R}^n$ are global minimizers of the polynomial optimization problem (P). □

Corollary 7.1.2. Suppose we are in the conditions of Theorem 7.1.1 and moreover suppose that S has non empty interior then $f - P^* \in M_{2d}(p_1, \dots, p_m)$ and hence:

$$P^* = P_{2d}^* = D_{2d}^*$$

Proof. Since L is an optimal solution of the Lasserre relaxation (P_{2d}) then: $L(f) \leq \Lambda(f)$ for all $\Lambda \in \mathbb{R}[\underline{X}]_{2d}^*$ feasible solution of the Lasserre relaxation (P_{2d}). Then $0 = L(f - P^*) = L(f) - P^* \leq \Lambda(f) - P^*$ for all $\Lambda \in \mathbb{R}[\underline{X}]_{2d}^*$ feasible solution of the Lasserre relaxation (P_{2d}). Now by the implication $(i) \implies (ii)$ in Theorem 4.4.2, $f - P^* \in M_{2d}(p_1, \dots, p_m)$, and by Corollary 4.4.3 it holds $P^* = P_{2d}^* = D_{2d}^*$.

□

The following Lemma was already proved in [18, lemma 2.7]. We will use it to prove the Corollary 7.1.4.

Lemma 7.1.3. Let $L = \sum_{i=1}^{N} \lambda_i \operatorname{ev}_{a_i} \in \mathbb{R}[\underline{X}]_{2d}^*$ for $a_1, \dots, a_N \in \mathbb{R}^n$ and $\lambda_1 > 0, \dots, \lambda_N > 0$ such that $N = \dim T_L$. Then there exist interpolation polynomials $q_1, \dots, q_N \in \mathbb{R}[\underline{X}]_{d-1}$ at the points $a_1, \dots, a_N$.

Proof. Let us consider the linear isometry (6.12) defined in Proposition 6.3.11:

$$\sigma_1 : T_L \longrightarrow \frac{\mathbb{R}[\underline{X}]}{U_\Lambda}, \overline{p}^L \mapsto \overline{p}^\Lambda, \text{ for } p \in \mathbb{R}[\underline{X}]_{d-1}$$

It is moreover an isomorphism of euclidean vector spaces since $\dim(\frac{\mathbb{R}[\underline{X}]}{U_\Lambda}) = N$ by Proposition 6.3.8. It is very well known that there exits polynomials $h_1, \dots, h_N \in \mathbb{R}[\underline{X}]$ at the points $a_1, \dots, a_N$, such that $h_i(a_j) = \delta_{i,j}$ for $i, j \in \{1, \dots, N\}$. Define $\overline{q_j}^L := \sigma_1^{-1}(\overline{h_j}^\Lambda)$ for $q_j \in \mathbb{R}[\underline{X}]_{d-1}$ for all $j \in \{1, \dots, N\}$. Then:

$$0 \leq \sum_{i=1}^{N} \lambda_i q_j^2(a_i) = L(q_j^2) = \left\langle \overline{q_j}^L, \overline{q_j}^L \right\rangle_L = \left\langle \overline{h_j}^\Lambda, \overline{h_j}^\Lambda \right\rangle_\Lambda = \Lambda(h_j^2) = \lambda_j$$

and therefore $q_j(a_i) = \delta_{i,j}$ for all $i, j \in \{1, \dots, N\}$. □

Corollary 7.1.4. Suppose that in the polynomial optimization problem (P) it holds $p_1, \ldots, p_m \in \mathbb{R}[\underline{X}]_1$ and let L be an optimal solution of (P_{2d}) with $f \in W_d$, set defined in (6.27). Moreover suppose that $\widetilde{M_L}$ is a generalized Hankel matrix. Then L has a quadrature rule representation on G_L whose nodes are minimizers of (P) and $L(f) = P^*$.

Proof. From Theorem 7.1.1 there exists exists nodes $a_1, \ldots, a_N \in \mathbb{R}^n$ and weights $\lambda_1 > 0, \ldots, \lambda_N > 0$, where $N := \dim T_L$ such that:

$$L(p) = \sum_{i=1}^{N} \lambda_i p(a_i) \text{ for all } p \in G_L$$

To conclude the proof by Theorem 7.1.1 it is enough to show that the nodes $a_1, \ldots, a_N$ are contained in S. In Theorem 6.5.1 we proved that $\hat{L} := \sum_{i=1}^{N} \lambda_i \operatorname{ev}_{a_i} \in \mathbb{R}[\underline{X}]_{2d}^*$ is flat. Then by the Lemma 7.1.3 there are interpolation polynomials $q_1, \ldots, q_N$ at the points $a_1, \ldots, a_N$ having at most degree $d-1$. Since $\deg(q_i^2 p_j) \leq 2d-1$ then $q_i^2 p_j \in M_{2d}(p_1, \ldots, p_m)$ and therefore:

$$0 \leq L(q_i^2 p_j) = \hat{L}(q_i^2 p_j) = \lambda_i p_j(a_i)$$

This equality proves that $p_j(a_i) \geq 0$ for $j \in \{1, \ldots, m\}$ and $i \in \{1, \ldots, N\}$ so we can conclude $\{a_1, \ldots, a_N\} \subseteq S$. □

Remark 7.1.5. The above results: Theorem 7.1.1 and Corollary 7.1.4 can be written in terms of an optimal solution of a Moment relaxation of odd degree by taking as an optimal solution its restriction to one degree less.

Algorithm 1: Algorithm for extracting minimizers of (P)

Input: A polynomial optimization problem $(\mathbf{P})$ (2.2).
Output: The minimum $\mathbf{P}^*$ and minimizers $\mathbf{a}_1, \ldots, \mathbf{a}_r \subseteq \mathbf{S}$ of **(P)**.

1 $\mathbf{k} := \max\{\deg f, \deg p_1, \ldots, \deg p_m\}$;

2 Compute an optimal solution $\mathbf{M} := M_{\lfloor \frac{k}{2} \rfloor}(\mathbf{y})$ of the Moment relaxation $(\mathbf{P_k})$ and also compute $\mathbf{W_M} \in \mathbb{R}^{r_{d-1} \times s_d}$ matrix such that:

$$\mathbf{M} = \left(\begin{array}{c|c} A_M & A_M \mathbf{W_M} \\ \hline \mathbf{W}_\mathbf{M}^T A_M & C_M \end{array} \right)$$

;

3 **if** $W_M^T A_M W_M$ *is a Hankel matrix* **then**

4 go to **7**

5 **else**

6 **k:=k+1** and go to **2**.

7 **if** *(*k *even and* $f \in \mathbb{R}[\underline{X}]_{k-1}$*)* ***or*** *(*$k$ *odd and* $f \in \mathbb{R}[\underline{X}]_{k-2}$*)* **then**

8 go to **14**

9 **else**

10 **if** $C_M = W_M A_M W_M$ **then**

11 go to **14**

12 **else**

13 **k:=k+1** go to **2**

14 Compute the truncated multiplication operators of $\mathbf{M}$: $\mathbf{A_{1,M}}, \ldots, \mathbf{A_{n,M}}$ and go to **15**.;

15 Compute an orthonormal basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ of $\mathbf{T_M}$ of common eigenvectors of the truncated multiplication operators such that $\mathbf{A_{i,M}} \mathbf{v}_j = \mathbf{a}_{j,i} \mathbf{v}_j$ and go to **16**.;

16 **if** $a_1, \ldots, a_n \in S$ **then**

17 go to **20**

18 **else**

19 **k:=k+1** and go to **2**

20 We can conclude that the points $\{\mathbf{a}_1, \ldots, \mathbf{a}_r\} \subseteq \mathbf{S}$ are minimizers of $(\mathbf{P})$, and $\mathbf{P}^* = \mathbf{f}(\mathbf{a}_i)$ for all $i \in \{1, \ldots, r\}$

7.2 Numerical results for Algorithm 1

Example 7.2.1. Let us apply the algorithm to the following polynomial optimization problem, taken from [15, Table 1: Rosenbrock function] :

$$
\begin{array}{ll}
\text{minimize} & f(x) = 100(x_2 - x_1^2)^2 + 100(x_3 - x_2^2)^2 + (x_1 - 1)^2 + (x_2 - 1)^2 \\
\text{subject to} & x_1, x_2, x_3 \in \mathbb{R} \\
 & -2.048 \leq x_1 \leq 2.048 \\
 & -2.048 \leq x_2 \leq 2.048 \\
 & -2.048 \leq x_3 \leq 2.048
\end{array}
$$

We initialize $\mathbf{k} = 4$ and compute an optimal solution of the moment relaxation (P_4). In this case reads as:

$$
\mathbf{M} := M_{4,1}(y) = \begin{array}{c} \\ 1 \\ X_1 \\ X_2 \\ X_3 \\ X_1^2 \\ X_1X_2 \\ X_1X_3 \\ X_2^2 \\ X_2X_3 \\ X_3^2 \end{array}
\begin{array}{c}
\begin{array}{cccccccccc} 1 & X_1 & X_2 & X_3 & X_1^2 & X_1X_2 & X_1X_3 & X_2^2 & X_2X_3 & X_3^2 \end{array} \\
\left(\begin{array}{cccc|cccccc}
1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 \\
1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 \\
1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 \\
1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 \\
\hline
1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 \\
1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 \\
1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 \\
1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 \\
1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 \\
1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & \mathbf{5.6502}
\end{array}\right)
\end{array}
$$

We can calculate that:

$$
W_{\mathbf{M}}^T A_{\mathbf{M}} W_{\mathbf{M}} = \begin{pmatrix}
1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 \\
1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 \\
1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 \\
1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 \\
1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 \\
1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & \mathbf{1.0000}
\end{pmatrix}
$$

is a generalized Hankel matrix and $f \in W_d$, set defined in (6.27), note that in this case the optimal solution is not flat since $C_{\mathbf{M}} \neq W_{\mathbf{M}}^T A_{\mathbf{M}} W_{\mathbf{M}}$. Since the polynomials defined the basic closed semialgebraic set are of degree 1 we can conclude by Corollary 7.1.4 that $P^* = P_4^* = 3.2568 \cdot 10^{-9} \approx 0$. Finally we get that the matrices of the truncated GNS operators with respect to the orthonormal basis $v := \langle 1 \rangle_{\mathbf{M}}$ are:

$$
M(M_{\mathbf{M},X_1}, v) = (1)\,, M(M_{\mathbf{M},X_2}, v) = (1) \text{ and } M(M_{\mathbf{M},X_2}, v) = (1)
$$

The operators are in diagonal form so we have already an orthonormal basis of $T_{\mathbf{M}}$ of common eigenvectors of the truncated GNS operators of $\mathbf{M}$ $v := \langle 1 \rangle_{\mathbf{M}}$. Then a global minimizer is $(1, 1, 1) \in \mathbb{R}^n$, and:

$$
\widetilde{\mathbf{M}} = V_2(1,1,1)V_2^T(1,1,1).
$$

Note that the hierarchy of relaxations does not attain a flat optimal solution until the moment relaxation (P_5). In this case we get the following optimal solution in the moment relaxation (P_5):

$$
\mathbf{M} := M_{5,1}(y) = \begin{array}{c|cccc|cccccc}
 & 1 & X_1 & X_2 & X_3 & X_1^2 & X_1X_2 & X_1X_3 & X_2^2 & X_2X_3 & X_3^2 \\
\hline
1 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 \\
X_1 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 \\
X_2 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 \\
X_3 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 \\
\hline
X_1^2 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 \\
X_1X_2 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 \\
X_1X_3 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 \\
X_2^2 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 \\
X_2X_3 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 \\
X_3^2 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & \mathbf{1.0014}
\end{array}
$$

we can calculate that:

$$
W_{\mathbf{M}}^T A_{\mathbf{M}} W_{\mathbf{M}} = \begin{pmatrix}
1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 \\
1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 \\
1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 \\
1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 \\
1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 \\
1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 & \mathbf{1.0000}
\end{pmatrix}
$$

In this case, after rounding to the third decimal, we can consider $C_{\mathbf{M}} = W_{\mathbf{M}}^T A_{\mathbf{M}} W_{\mathbf{M}}$, i.e. M is flat, and continue with the algorithm and we get the same minimizer.

In the next Example 7.2.2 we will see that it could happen that the modified moment matrix is of generalized Hankel form and the candidates to minimizers are not inside the basic closed semialgebraic set so we can not conclude optimality until a relaxation of bigger degree.

Example 7.2.2. Let us consider the following polynomial optimization problem, defined on a non convex closed semialgebraic set, taken from [11, problem 4.6] :

$$
\begin{array}{ll}
\text{minimize} & f(x) = -x_1 - x_2 \\
\text{subject to} & x_1, x_2 \in \mathbb{R} \\
 & x_2 \leq 2x_1^4 - 8x_1^3 + 8x_1^2 + 2 \\
 & x_2 \leq 4x_1^4 - 32x_1^3 + 88x_1^2 - 96x_1 + 36 \\
 & 0 \leq x_1 \leq 3 \\
 & 0 \leq x_2 \leq 4
\end{array}
$$

We initialize $\mathbf{k} = 4$. An optimal solution of (P_4) reads as:

(7.1)

$$
\mathbf{M} := M_{4,1}(y) = \begin{array}{c|ccc|ccc}
 & 1 & X_1 & X_2 & X_1^2 & X_1X_2 & X_2^2 \\
\hline
1 & 1.0000 & 3.0000 & 4.0000 & 9.0000 & 12.0000 & 16.0000 \\
X_1 & 3.0000 & 9.0000 & 12.0000 & 27.0000 & 36.0000 & 48.0000 \\
X_2 & 4.0000 & 12.0000 & 16.0000 & 36.0000 & 48.0000 & 64.0000 \\
\hline
X_1^2 & 9.0000 & 27.0000 & 36.0000 & 107.6075 & 109.0814 & 176.3211 \\
X_1X_2 & 12.0000 & 36.0000 & 48.0000 & 109.0814 & 176.3211 & 194.9661 \\
X_2^2 & 16.0000 & 48.0000 & 64.0000 & 176.3211 & 194.9661 & 368.5439
\end{array}
$$

and

$$(7.2) \qquad \widetilde{\mathbf{M}} = \begin{array}{c} \\ 1 \\ X_1 \\ X_2 \\ X_1^2 \\ X_1X_2 \\ X_2^2 \end{array} \begin{array}{c} \begin{array}{cccccc} 1 & X_1 & X_2 & X_1^2 & X_1X_2 & X_2^2 \end{array} \\ \left(\begin{array}{ccc|ccc} 1.0000 & 3.0000 & 4.0000 & 9.0000 & 12.0000 & 16.0000 \\ 3.0000 & 9.0000 & 12.0000 & 27.0000 & 36.0000 & 48.0000 \\ 4.0000 & 12.0000 & 16.0000 & 36.0000 & 48.0000 & 64.0000 \\ \hline 9.0000 & 27.0000 & 36.0000 & 81.000 & 108.000 & 144.000 \\ 12.0000 & 36.0000 & 48.0000 & 108.000 & 144.000 & 192.00 \\ 16.0000 & 48.0000 & 64.0000 & 144.000 & 192.000 & 256.000 \end{array}\right) \end{array}$$

taking for example:

$$W_{\mathbf{M}} = \begin{pmatrix} 9 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 3 & 4 \end{pmatrix}$$

$\widetilde{\mathbf{M}}$ is a generalized Hankel matrix. The matrix of the truncated GNS multiplaction operators with respect to the orthonormal basis $v = \left\langle \overline{1}^{\mathbf{M}} \right\rangle$ are:

$$M(M_{\mathbf{M},1}, v) = \begin{pmatrix} 3 \end{pmatrix} \text{ and } M(M_{\mathbf{M},1}, v) = \begin{pmatrix} 4 \end{pmatrix}$$

Hence the candidate to minimizer is $(3,4)$, however it does not lie in S, then $(3,4)$ cannot be a minimizer and $f(3,4) = -7$ cannot be the minimum. Then we try with the next relaxation of order $\mathbf{k} = 5$. An optimal solution of the Moment relaxation (P_5) is the following:

$$(7.3) \qquad \mathbf{M} := M_{5,1}(y) = \begin{array}{c} \\ 1 \\ X_1 \\ X_2 \\ X_1^2 \\ X_1X_2 \\ X_2^2 \end{array} \begin{array}{c} \begin{array}{cccccc} 1 & X_1 & X_2 & X_1^2 & X_1X_2 & X_2^2 \end{array} \\ \left(\begin{array}{ccc|ccc} 1.00 & 2.67 & 4.00 & 8.00 & 10.67 & 16.00 \\ 2.67 & 8.00 & 10.67 & 24.00 & 32.00 & 42.67 \\ 4.00 & 10.67 & 16.00 & 32.00 & 42.67 & 64.00 \\ \hline 8.00 & 24.00 & 32.00 & 72.00 & 96.00 & 128.00 \\ 10.67 & 32.00 & 42.67 & 96.00 & 128.00 & 170.67 \\ 16.00 & 42.67 & 64.00 & 128.00 & 170.67 & 256.00 \end{array}\right) \end{array}$$

In this case $C_{\mathbf{M}} = W_{\mathbf{M}}^T A_{\mathbf{M}} W_{\mathbf{M}}$, therefore $\mathbf{M}$ is flat and in particular the operators commute by 6.3.19. After the simultaneous diagonalization of the truncated GNS operators we get that the candidates to minimizers are $(0,4) \notin S$ and $(3,4) \notin S$. Hence we try with a relaxation of order $\mathbf{k} = 6$. An optimal solution of the Moment relaxation (P_6), after rounding in the moments of biggest degree, reads as:

(7.4)
$$\mathbf{M} := M_{6,1}(y) = \begin{array}{c|cccccc|cccc}
 & 1 & X_1 & X_2 & X_1^2 & X_1X_2 & X_2^2 & X_1^3 & X_1^2X_2 & X_1X_2^2 & X_2^3 \\
\hline
1 & 1.00 & 2.67 & 4.00 & 8.00 & 10.67 & 16.00 & 24.00 & 32.00 & 42.67 & 64.00 \\
X_1 & 2.67 & 8.00 & 10.67 & 24.00 & 32.00 & 42.67 & 72.00 & 96.00 & 128.00 & 170.67 \\
X_2 & 4.00 & 10.67 & 16.00 & 32.00 & 42.67 & 64.00 & 96.00 & 128.00 & 170.67 & 256.00 \\
X_1^2 & 8.00 & 24.00 & 32.00 & 72.00 & 96.00 & 128.00 & 216.00 & 288.00 & 384.00 & 512.00 \\
X_1X_2 & 10.67 & 32.00 & 42.67 & 96.00 & 128.00 & 170.67 & 288.00 & 384.00 & 512.00 & 682.66 \\
X_2^2 & 16.00 & 42.67 & 64.00 & 128.00 & 170.67 & 256.00 & 384.00 & 512.00 & 682.66 & 10^3 \\
\hline
X_1^3 & 24.00 & 72.00 & 96.00 & 216.00 & 288.00 & 384.00 & 2\cdot 10^5 & 870.25 & 2\cdot 10^4 & 1583.15 \\
X_1^2X_2 & 32.00 & 96.00 & 128.00 & 288.00 & 384.00 & 512.00 & 870.25 & 2\cdot 10^4 & 10^3 & 2\cdot 10^4 \\
X_1X_2^2 & 42.67 & 128.00 & 170.67 & 384.00 & 512.00 & 682.66 & 2\cdot 10^4 & 10^3 & 2\cdot 10^4 & 3\cdot 10^3 \\
X_2^3 & 64.00 & 170.67 & 256.00 & 512.00 & 682.66 & 10^3 & 10^3 & 2\cdot 10^4 & 3\cdot 10^3 & 6\cdot 10^4
\end{array}$$

and

$$W_{\mathbf{M}}^T A_{\mathbf{M}} W_{\mathbf{M}} = \begin{array}{c|cccc}
 & X_1^3 & X_1^2X_2 & X_1X_2^2 & X_2^3 \\
\hline
X_1^3 & 648.00 & 863.99 & 1151.99 & 1535.99 \\
X_1^2X_2 & 863.99 & 1151.99 & 1535.99 & 2047.99 \\
X_1X_2^2 & 1151.99 & 1535.99 & 2047.99 & 2730.65 \\
X_2^3 & 1535.99 & 2047.99 & 2730.65 & 4095.99
\end{array}$$

is a generalized Hankel matrix. However we get the same candidates to minimizers as in the previous relaxation and they do not belong to S. Finally we increase to $\mathbf{k} = 7$, and we get after rounding, the following optimal solution of (P_7):

$$\mathbf{M} := M_{7,1}(y) = \begin{array}{c|cccccc|cccc}
 & 1 & X_1 & X_2 & X_1^2 & X_1X_2 & X_2^2 & X_1^3 & X_1^2X_2 & X_1X_2^2 & X_2^3 \\
\hline
1 & 1.00 & 2.33 & 3.18 & 5.43 & 7.40 & 10.10 & 12.64 & 17.25 & 23.53 & 32.11 \\
X_1 & 2.33 & 5.43 & 7.40 & 12.64 & 17.25 & 23.53 & 29.45 & 40.18 & 54.82 & 74.80 \\
X_2 & 3.18 & 7.40 & 10.10 & 17.25 & 23.53 & 32.11 & 40.18 & 54.82 & 74.80 & 102.07 \\
X_1^2 & 5.43 & 12.64 & 17.25 & 29.45 & 40.18 & 54.82 & 68.60 & 93.60 & 127.72 & 174.26 \\
X_1X_2 & 7.40 & 17.25 & 23.53 & 40.18 & 54.82 & 74.80 & 93.60 & 127.72 & 174.26 & 237.77 \\
X_2^2 & 10.10 & 23.53 & 32.11 & 54.82 & 74.80 & 102.07 & 127.72 & 174.26 & 237.77 & 324.42 \\
\hline
X_1^3 & 12.64 & 29.45 & 40.18 & 68.60 & 93.60 & 127.72 & 159.81 & 218.05 & 297.51 & 405.94 \\
X_1^2X_2 & 17.25 & 40.18 & 54.82 & 93.60 & 127.72 & 174.26 & 218.05 & 297.51 & 405.94 & 553.88 \\
X_1X_2^2 & 23.53 & 54.82 & 74.80 & 127.72 & 174.26 & 237.77 & 297.51 & 405.94 & 553.88 & 755.74 \\
X_2^3 & 32.11 & 74.80 & 102.07 & 174.26 & 237.77 & 324.42 & 405.94 & 553.88 & 755.74 & 1031.16
\end{array}$$

It holds that $\widetilde{\mathbf{M}} = \mathbf{M}$, therefore in particular $\widetilde{\mathbf{M}}$ is a generalized Hankel matrix and the truncated multiplication operators commute. The matrices of the truncated GNS multiplication operators with respect to the orthonormal basis $v := \left\langle \overline{1}^{\mathbf{M}} \right\rangle$ are:

$$M(M_{\mathbf{M},X_1}, v) = \begin{pmatrix} 2.3295 \end{pmatrix} \text{ and } M(M_{\mathbf{M},X_2}, v) = \begin{pmatrix} 3.1785 \end{pmatrix}$$

Since $(2.3295, 3.1785) \in S$ then it is also a minimizer and we proved optimality i.e. $P^* = P_7^* = -5.5080$.

Example 7.2.3. Let us considerer the following polynomial optimization problem taken from [16, example 5]:

$$\begin{aligned} \text{minimize} \quad & f(x) = -(x_1-1)^2 - (x_1-x_2)^2 - (x_2-3)^2 \\ \text{subject to} \quad & x_1, x_2, x_3 \in \mathbb{R} \\ & 1-(x_1-1)^2 \geq 0 \\ & 1-(x_1-x_2)^2 \geq 0 \\ & 1-(x_2-3)^2 \geq 0 \end{aligned}$$

For $\mathbf{k} = 2$ and $\mathbf{k} = 3$ in the algorithm, the modified moment matrix of the optimal solution of the Moment relaxation is generalized Hankel and we get as a potential minimizers, after the truncated GNS construction, $(1.56, 2.18) \in S$ in both relaxations, however $f \notin W_d$ (6.27), so we can not conclude $(1.56, 2.18)$ is a global minimum. For $\mathbf{k} = 4$, the optimal solution of the Moment relaxation (P_4) reads as:

$$\mathbf{M} := M_4(y) = \begin{array}{c} 1 \\ X_1 \\ X_2 \\ X_1^2 \\ X_1X_2 \\ X_2^2 \end{array} \overset{\begin{array}{cccccc} 1 & X_1 & X_2 & X_1^2 & X_1X_2 & X_2^2 \end{array}}{\left(\begin{array}{ccc|ccc} 1.0000 & 1.4241 & 2.1137 & 2.2723 & 3.0755 & 4.5683 \\ 1.4241 & 2.2723 & 3.0755 & 3.9688 & 4.9993 & 6.8330 \\ 2.1137 & 3.0755 & 4.5683 & 4.9993 & 6.8330 & 10.1595 \\ \hline 2.2723 & 3.9688 & 4.9993 & 7.3617 & 8.8468 & 11.3625 \\ 3.0755 & 4.9993 & 6.8330 & 8.8468 & 11.3625 & 15.7120 \\ 4.5683 & 6.8330 & 10.1595 & 11.3625 & 15.7120 & 23.3879 \end{array}\right)}$$

and we can verify $\widetilde{\mathbf{M}} = \mathbf{M}$. Hence in this case $\mathbf{M}$ is flat, then it is clear that $\widetilde{\mathbf{M}}$ is a generalized Hankel matrix implying that the truncated GNS multiplication operators of $\mathbf{M}$ commute. We proceed to do the truncated GNS construction and we get the following orthonormal basis of $W_{\mathbf{M}}$:

$$W_{\mathbf{M}} = \left\langle \overline{1}^{\mathbf{M}}, \overline{-2.08816 + 2.0234X_1}^{\mathbf{M}}, \overline{-6.0047 - 0.9291X_1 + 3.4669X_2}^{\mathbf{M}} \right\rangle$$

Denote $v := \left\langle \overline{1}^{\mathbf{M}}, \overline{-2.08816 + 2.0234X_1}^{\mathbf{M}}, \overline{-6.0047 - 0.9291X_1 + 3.4669X_2}^{\mathbf{M}} \right\rangle$ such a basis. Then the transformation matrices of the truncated GNS multiplication operators with respect to this basis are:

$$A_1 := M(M_{\mathbf{M},X_1}, v) = \begin{pmatrix} 1.4241 & 0.4942 & 0.0000 \\ 0.4942 & 1.5759 & 0.0000 \\ 0.0000 & 0.0000 & 2.0000 \end{pmatrix}$$

$$A_2 := M(M_{\mathbf{M},X_2}, v) = \begin{pmatrix} 2.1137 & 0.1324 & 0.2884 \\ 0.1324 & 2.1543 & 0.3361 \\ 0.2884 & 0.3361 & 2.7320 \end{pmatrix}$$

Again we follow the same idea as in [24, algorithm 4.1 Step 1] to apply simultaneous diagonalization to the matrices A_1 and A_2. For this we find the orthogonal matrix P that diagonalize a matrix of the following form:

$$A = r_1 A_1 + r_2 A_2 \text{ where } r_1^2 + r_2^2 = 1$$

For

$$P = \begin{pmatrix} 0.7589 & 0.5572 & 0.3371 \\ -0.6512 & 0.6493 & 0.3929 \\ 0.0000 & -0.5177 & 0.8556 \end{pmatrix}$$

we get the following diagonal matrices:

$$P^T A_1 P = \begin{pmatrix} 1.0000 & 0.0000 & -0.0000 \\ 0.0000 & 2.0000 & 0.0000 \\ -0.0000 & 0.0000 & 2.0000 \end{pmatrix}, P^T A_2 P = \begin{pmatrix} 2.0000 & -0.0000 & 0.0000 \\ -0.0000 & 2.0000 & -0.0000 \\ 0.0000 & -0.0000 & 3.0000 \end{pmatrix}$$

and with the operation:

$$P^T \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0.7589 \\ 0.5572 \\ 0.3371 \end{pmatrix}$$

we get the square roots of the weights of the quadrature formula. Then we have the following decomposition:

$$\mathbf{M} = \widetilde{\mathbf{M}} = 0.5759 V_2(1,2) V_2^T(1,2) + 0.3105 V_2(2,2) V_2^T(2,2) + 0.1137 V_2(2,3) V_2^T(2,3)$$

In this case the points $(1,2)$,$(2,2)$, and $(2,3)$ lie on S, as we already knew it since it holds the condition of the Theorem 1.6 in [5], and therefore they are global minimizers of (P), and the minimum is $P^* = P_4^* = -2$.

Example 7.2.4. Consider the following polynomial optimization problem:

$$\begin{aligned} \text{minimize} \quad & f(x) = x_1 x_2 \\ \text{subject to:} \quad & x_1, x_2 \in \mathbb{R} \\ & x_1^2 + x_2^2 - 1 = 0 \end{aligned}$$

In the moment relaxation of degree $\mathbf{k} = 4$ we get the following optimal solution $y \in \mathbb{R}^{s_2}$ with associated moment matrix:

$$\mathbf{M} := M_4(y) = \begin{array}{c} \\ 1 \\ X_1 \\ X_2 \\ X_1^2 \\ X_1X_2 \\ X_2^2 \end{array} \begin{array}{c} \begin{array}{cccccc} 1 & X_1 & X_2 & X_1^2 & X_1X_2 & X_2^2 \end{array} \\ \left(\begin{array}{ccc|ccc} 1.00 & -0.00 & 0.00 & 0.50 & -0.50 & 0.50 \\ -0.00 & 0.50 & -0.50 & -0.00 & 0.00 & -0.00 \\ 0.00 & -0.50 & 0.50 & 0.00 & -0.00 & 0.00 \\ \hline 0.50 & -0.00 & 0.00 & 0.25 & -0.25 & 0.25 \\ -0.50 & 0.00 & -0.00 & -0.25 & 0.25 & -0.25 \\ 0.50 & -0.00 & 0.00 & 0.25 & -0.25 & 0.25 \end{array}\right) \end{array}$$

and optimal value $P_4^* = -0.5$. This matrix is not flat, however $\widetilde{\mathbf{M}}$ is of generalized Hankel form since:

$$W_{\mathbf{M}}^T A_{\mathbf{M}} W_{\mathbf{M}} = \begin{array}{c} \\ X_1^2 \\ X_1X_2 \\ X_2^2 \end{array} \begin{array}{c} \begin{array}{ccc} X_1^2 & X_1X_2 & X_2^2 \end{array} \\ \begin{pmatrix} 0.25 & -0.25 & 0.25 \\ -0.25 & 0.25 & -0.25 \\ 0.25 & -0.25 & 0.25 \end{pmatrix} \end{array}$$

Then the GNS truncated multiplication operators commute. Taking: $v := \langle 1, 1.4142 \cdot X_1 \rangle$ as a orthonormal basis of the GNS truncated space, we get the following matrices of the GNS truncated multiplication operators:

$$A_1 := M(M_{\mathbf{M},X_1}, v) = \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 \end{pmatrix}$$

$$A_2 := M(M_{\mathbf{M},X_2}, v) = \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 \end{pmatrix}$$

After diagonalizing this matrix $A_1 = A_2$ we get as a potential minimizers the points:

$$(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}), (\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}) \in \mathbb{R}^2$$

Therefore, since the points are in the circumference and $f \in \mathbb{R}[X_1, X_2]_2$ we can conclude optimality, and extract two minimizers, i.e:

$$f(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}) = f(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}) = P_4^* = P^*$$

CHAPTER 8

An experimental approach for global polynomial optimization based on moments and semidefinite programming

In this chapter, we consider the following polynomial optimization problem, for $f \in \mathbb{R}[\underline{X}]$:

$$(P) \text{ minimize } f(x) \text{ subject to } x \in \mathbb{R}^n \tag{8.1}$$

Throughout this chapter, we suppose $f \in \mathbb{R}[\underline{X}]_k$ and we will denote $y \in \mathbb{R}^{s_k}$ an optimal solution of the moment relaxation of degree k for the problem (8.1), $d := \left\lfloor \frac{k}{2} \right\rfloor$ and $M := M_d(y)$.

8.1 Algorithm 2: based on moments and semidefinite programming

It is not always the case that M is flat or it is not always the more general case that $\widetilde{M}$ is a generalized Hankel matrix. In this situation, in order to find the minimum P^* and minimizers, we could try to increase k, solve again the moment relaxation and hope that we get an optimal solution with M a flat matrix or $\widetilde{M}$ generalized Hankel matrix, as we did in Algorithm 1. However the dimension of the problem could increase considerably and one frequently runs into numerical problems. In this section, we build the Algorithm 2. This algorithm tries to modify a little bit the optimal solution y to get a flat solution or a solution close to be flat, that is to say, we try to avoid to increase k. A first try to get a flat optimal solution of (P_k) would be to add linear constraints into the Moment relaxation in order to restrict our set of feasible solutions to a set of flat feasible solutions or at least "close" to be flat. Let me explain shortly why this is in principle, a hard problem. As we have mentioned before, a first approach could be try to describe the following program:

$$\begin{cases} \text{minimize} & \sum_{|\alpha|\le k} f_\alpha y_\alpha \\ \text{subject to:} & M_d(y) \succeq 0, \\ & y_{(0,\dots,0)} = 1 \text{ and} \\ & \operatorname{rank} M_d(y) = \operatorname{rank} M_{d-1}(y) \end{cases}$$

as a positive semidefinite program. However, $\operatorname{rank} M_d(y) = \operatorname{rank} M_{d-1}(y)$ if and only if $M_i \in \operatorname{span}\langle M_1, \dots, M_{s_{d-1}}\rangle$ for all $i \in \{s_{d-1}+1, \dots, s_d\}$. However we can not add the constraints:

$$M_i = \sum_{j=1}^{s_{d-1}} a_j^i M_j \text{ for some } a_1^i, \dots, a_{s_{d-1}}^i \in \mathbb{R} \quad \text{for all } i \in \{s_{d-1}+1, \dots, s_d\} \tag{8.2}$$

to our moment relaxation since this condition is not linear due to the a_i and the entries of the matrix are decision variables, and this can not be written, at least not in any obvious way, as a semidefinite program. In the same way, the program:

$$\begin{cases} \text{minimize} & \sum_{|\alpha|\le k} f_\alpha y_\alpha \\ \text{subject to:} & M_d(y) \succeq 0, \\ & y_{(0,\dots,0)} = 1 \text{ and} \\ & M_d(y) \text{ is a generalized Hankel matrix} \end{cases}$$

is not a positive semidefinite program due to that $W_{M_d(y)}^T M_{d-1}(y) W_{M_d(y)}$ is not a linear matrix since the entries of $M_{d-1}(y)$ and the entries of $W_{M_d(y)}$ are decision variables. Moreover to solve polynomial optimization problems without constraints already for degree 4 polynomials is NP hard [25], so it is reasonable to expect that to convert these programs into a semidefinite program is hard. Nevertheless, we can modify a little bit the optimal solution y into $y_0 \in \mathbb{R}^{s_k}$ in such a way that y_0 is feasible solution of (P_k) and $M_d(y_0)$ is approximately flat. Since y_0 is a feasible solution of (P_k) the inequality $\sum_{|\alpha|\le k} f_\alpha (y_0)_\alpha \ge P_k^*$ holds. Moreover if $M_d(y_0)$ is flat then by Proposition 4.4.4 (*ii*) together with Theorem 6.3.19 we know that $\sum_{|\alpha|\le 2d} f_\alpha (y_0)_\alpha \ge P^*$. More precise in this last case it holds that:

$$P^* \in [P_k^*, \sum_{|\alpha|\le k} f_\alpha (y_0)_\alpha]$$

Reminder 8.1.1. Let us consider the following polynomial optimization problem, called the *Least Squares Problem*:

$$(P_{A,b}) : \begin{cases} \text{minimize} & ||Ax - b||_2^2 \\ \text{subject to:} & x \in \mathbb{R}^n \end{cases} \tag{8.3}$$

where $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$ are given and $m > n$. Minimizers of this problem are called a *least squares approximate solution*. Suppose that the columns of A are linearly independent then we denote to the unique solution of the least squares problem as $x^*_{A,b}$ and the solution is given by the following formula:

$$x^*_{A,b} = (A^T A)^{-1} A^T b \tag{8.4}$$

For a proof of the Reminder 8.1.1 and more details about the topic we refer the reader to [32, Section 12.2] and references therein.

The first step in the algorithm is to build the closest matrix to $M := M_d(y)$, let us denote it by $B_M \in \mathbb{R}^{s_d \times s_d}$, whose first s_{d-1} columns are the same as the first s_{d-1} columns of $M_{d-1}(y)$ and whose rank is the same as the rank of $M_{d-1}(y)$. Hence to build the last r_d columns of B_M such that they are in the linear span of $M_{d-1}(y)$ we will use the least square problem. More precise, for $j \in \{s_{d-1}+1, \dots, s_d\}$ the column B_{M_j} is defined as the closest vector to M_j which lies in the real linear span $\langle M_1, \dots, M_{s_{d-1}} \rangle$. Note that, without loss of generality we can assume that the columns of $M_{d-1}(y)$ are linearly independent (otherwise we take a linearly independent subset of the columns of $M_{d-1}(y)$), then for all $j \in \{s_{d-1}+1, \dots, s_d\}$:

$$B_{M_j} = M_{d-1}(y) x^*_{M_{d-1}(y), M_j}$$

The matrix B_M holds the desired condition in the rank (8.2), however it is not necessarily positive semidefinite, not necessarily generalized Hankel and not even symmetric. Ultimately B_M is not the moment matrix of a feasible solution of (P_k). So now we look for $y_0 \in \mathbb{R}^{s_k}$, such that $E \in \mathbb{R}$ is the smallest possible in the following inequality:

$$||M_d(y_0) - B_M||^2 \leq E||M - B_M||^2 \tag{8.5}$$

Setting $A_M := M_d(y_0)$, we will solve the following program:

$$(P_M) : \begin{cases} \text{minimize} & E \\ \text{subject to:} & E \in \mathbb{R} \text{ and} \\ & ||A_M - B_M||^2 \leq E||M - B_M||^2 \end{cases}$$

Note that in the program (P_M) the decision variables are: $y_0 \in \mathbb{R}^{s_k}$ and $E \in \mathbb{R}$. The condition (8.5) attempts to simultaneously control the rank of A_M by minimizing the distance from A_M to B_M and at the same time we get a matrix with lower or equal rank than the original matrix M, since note the inequality (8.5) holds taking $E := 1$ and $A_M := M$ and there exists always a feasible solution. The Schur complement will be defined in 8.1.2 and it will enable us to show that this condition is equivalent to the positive semidefiniteness of a matrix with linear entries and therefore we can conclude that (P_M) is a positive semidefinite program, and in conclusion possible to solve efficiently.

Definition 8.1.2. Let us consider a matrix $X \in S\mathbb{R}^{m+l \times m+l}$ in block form:

$$X = \left(\begin{array}{c|c} A & B \\ \hline B^T & C \end{array} \right) \tag{8.6}$$

with $A \in \mathbb{R}^{m\times m}$, $B \in \mathbb{R}^{m\times l}$, $C \in \mathbb{R}^{l\times l}$. Assume A is non-singular matrix. Then the matrix $C - B^T A^{-1} B$ is called the *Schur complement of* A in X.

Lemma 8.1.3. Let $X \in \operatorname{Sym}\mathbb{R}^{m\times m}$ be in block form (8.6), where A is non-singular. Then,

$$X \succeq 0 \iff A \succeq 0 \text{ and } C - B^T A^{-1} B \succeq 0$$

Proof. For a proof of this lemma we refer the reader to [19, Lemma 1.7.6]. □

Therefore applying Lemma 8.1.3 we get that:

$$||A_M - B_M||^2 \leq E||M - B_M||^2 \iff$$

(8.7)

$$\left(\begin{array}{ccc|c} & & & (A_M)_{1,1} - (B_M)_{1,1} \\ & I_{s_d s_d} & & \vdots \\ & & & (A_M)_{s_d,s_d} - (B_M)_{s_d,s_d} \\ \hline (A_M)_{1,1} - (B_M)_{1,1} & \cdots & (A_M)_{s_d,s_d} - (B_M)_{s_d,s_d} & E||M - B_M||^2 \end{array}\right) \succeq 0$$

The matrix (8.7) is linear in the variables E and the entries of the matrix A_M, that is in $y_0 \in \mathbb{R}^{s_k}$. We could solve directly the semidefinite program (P_M) to get a lower bound of P^*. However in practice if instead we consider the following semidefinite program for $\lambda \in [0,1]$ fixed, we get better bounds:

$$(P_{M,\lambda}) : \begin{cases} \text{minimize} & E\lambda + (1-\lambda)\sum_{|\alpha|\leq k} f_\alpha y_{0,\alpha} \\ \text{subject to:} & E \in \mathbb{R}\,, \\ & ||M_k(y_0) - B||^2 \leq E||M - B_M||^2 \text{ and} \\ & M_k(y_0) \succeq 0 \end{cases}$$

The optimal value of $(P_{M,\lambda})$ that is to say, the infimum over all:

$$(y_{0,(0,\ldots,0)}, \ldots, y_{0,(0,\ldots,k)}) \in \mathbb{R}^{s_k} \text{ and } E \in \mathbb{R}$$

that ranges over all optimal solutions of $(P_{M,\lambda})$ is denoted by $P^*_{M,\lambda}$.

Algorithm 2: Given (P) (8.1), finding an upper bound U of P^*, i.e $U \geq P^*$ and if possible minimizers or potential minimizers

---

Input: A polynomial optimization problem $(\mathbf{P})$ (2.2) without constraints and an strategic $\lambda \in [0,1]$

Output: An upper bound $U \geq \mathbf{P}^*$ and if possible, points $\mathbf{a}_1, \ldots, \mathbf{a}_r$ such that $f(\mathbf{a_i}) = U$.

1 Compute a feasible solution of the Moment relaxation $(P_{\deg f})$ and denote it by $y \in \mathbb{R}^{s_{\deg f}}$. Set $d := \lfloor \frac{\deg f}{2} \rfloor$ and $\mathbf{M} := M_d(y)$. ;

2 Take a maximum linearly independent subset of the first s_{d-1} columns of $\mathbf{M}$ and denote it by $\{M_1, \ldots, M_l\}$ and set $C := (M_1 \cdots M_l) \in \mathbb{R}^{s_d \times l}$. ;

3 Define the folowing matrix using the definition of the solution of the least squares problem (8.4), $\mathbf{B_M} := (M_1 \cdots M_{s_{d-1}} | C x^*_{C,M_{s_{d-1}+1}} \cdots C x^*_{C,M_{s_d}})$.;

4 Finally define the following matrix:

$$\mathbf{T_M} := \left(\begin{array}{ccc|c} & & & (A_M)_{1,1} - (B_M)_{1,1} \\ & I_{s_d s_d} & & \vdots \\ & & & (A_M)_{s_d,s_d} - (B_M)_{s_d,s_d} \\ \hline (A_M)_{1,1} - (B_M)_{1,1} & \cdots & (A_M)_{s_d,s_d} - (B_M)_{s_d,s_d} & E||M - B_M||^2 \end{array} \right)$$

where the decision variables are $\mathbf{E}$ and the entries of the matrix $\mathbf{A_M} \in S\mathbb{R}^{s_d \times s_d}$. ;

5 Solve the following positive semidefinite program:

$$(P_{M,\lambda}) : \begin{cases} \text{minimize} & E\lambda + (1-\lambda) \sum_{|\alpha| \leq k} f_\alpha y_{0,\alpha} \\ \text{subject to:} & E \in \mathbb{R}\,, \\ & \mathbf{A_M} = M_d(y_0) \text{ and} \\ & \mathbf{T_M} \succeq 0 \\ & A_M \succeq 0 \end{cases}$$

Note that here the decision variables are E and the entries of the matrix $\mathbf{A_M} \in S\mathbb{R}^{s_d \times s_d}$ which are $y_0 \in \mathbb{R}^{s_k}$.

6 **if** *A_M is flat or $||\mathbf{A_M} - \mathbf{B_M}||$ is small enough* **then**

7 $\quad$ $\mathbf{U} := \sum_{|\alpha| \leq k} f_\alpha y_{0,\alpha}$. If possible extract $a_1, \ldots a_r \in \mathbb{R}^n$ such that $f(a_i) = \mathbf{U}$ as in Algorithm 1.

8 **else**

9 $\quad$ $\mathbf{M} := \mathbf{A_M}$ and go to **3**

Remark 8.1.4. We do not know if the algorithm ever terminates. In the examples we did, if the algorithm took too much time, we interrupted the algorithm even if the matrix $\mathbf{A_M}$ was not flat. Note also that since y_0 is a feasible solution of P_k then $U \geq P_k^*$ and in case $\mathbf{A_M}$ is flat we can even conclude by Proposition 4.4.4 (ii) and Theorem 6.3.19 that:

$$P^* \in [P_k^*, \mathbf{U}]$$

8.2 Numerical results for Algorithm 2

Example 8.2.1. The *Motzkin polynomial*: $X_1^4X_2^2 + X_1^2X_2^4 - 3X_1^2X_2^2 + 1$, is nonnegative but is not a sum of squares, see for example [22, Proposition 1.2.2] for a nice proof ot this last fact. Let us consider the following polynomial optimization problem:

$$\begin{array}{ll} \text{minimize} & f(x) = x_1^4x_2^2 + x_1^2x_2^4 - 3x_1^2x_2^2 + 1 \\ \text{subject to} & x_1, x_2 \in \mathbb{R} \end{array}$$

Since the Motzkin polynomial is nonnegative and $f(\pm 1, \pm 1) = 0$, is not difficult to conclude that $P^* = 0$. Let us use this polynomial to see how the Algorithm 2 behaves. An optimal solution $z \in \mathbb{R}^{s_6}$ of the Moment relaxation (P_6), has the following Moment matrix:

$M_3(z) =$

	1	X_1	X_2	X_1^2	X_1X_2	X_2^2	X_1^3	$X_1^2X_2$	$X_1X_2^2$	X_2^3
1	1.00	0.00	0.00	5300.32	0.00	5300.31	0.00	0.00	0.00	0.00
X_1	0.00	5300.32	0.00	0.00	0.00	0.00	57120303.73	0.00	2966.08	0.00
X_2	0.00	0.00	5300.31	0.00	0.00	0.00	0.00	2966.08	0.00	57120215.99
X_1^2	5300.32	0.00	0.00	57120303.73	0.00	2966.08	−2.14	0.00	0.00	0.00
X_1X_2	0.00	0.00	0.00	0.00	2966.08	0.00	0.00	0.00	0.00	0.00
X_2^2	5300.31	0.00	0.00	2966.08	0.00	57120215.99	0.00	0.00	0.00	−24.83
X_1^3	0.00	57120303.73	0.00	2.14	0.00	0.00	2272816741819.88	0.00	2491.80	0.00
$X_1^2X_2$	0.00	0.00	2966.08	0.00	0.00	0.00	0.00	2491.80	0.00	2491.80
$X_1x_2^2$	0.00	2966.08	0.00	0.00	0.00	0.00	2491.80	0.00	2491.80	0.00
X_23	0.00	0.00	57120215.99	0.00	0.00	−24.83	0.00	2491.80	0.00	2272810890195.67

This matrix is neither flat nor $\widetilde{M_3(z)}$ is a generalized Hankel matrix so we can not conclude optimality. In fact the optimal value is $P_6^* = -\infty << P^* = 0$ stricly smaller than the minimum. Since the entries of the matrix $M_3(z)$ are very far to each other and working with this matrix could bring us numerical problems, so instead we take a

	U	Flat	$\|\|A_M - B_M\|\|$	Potential Minimizers	Iterations	Time
$\lambda = \frac{1}{1000}$	-209.9332	No	$1.8995 \cdot 10^7$	-	73	$2^M 14^S$
$\lambda = \frac{1}{100}$	-48.50	No	$1.584 \cdot 10^5$	-	73	$1^M 42^S$
$\lambda = \frac{1}{60}$	0.00156	Yes	9.16135	$(\pm 1.0109, \pm 1.0109)$	66	$1^M 4^S$
$\lambda = \frac{1}{4}$	0.0650	No	$2.0449 \cdot 10^{-4}$	-	66	$1^M 15^S$
$\lambda = \frac{1}{2}$	0.2537	No	$9.0867 \cdot 10^{-4}$	-	66	$1^M 20^S$
$\lambda = \frac{3}{4}$	0.3543	Approximately	0.0015	$(\pm 0.9960, \pm 0.9960)$	74	$1^M 5^S$
$\lambda = 1$	0.2870	No	0.0013	-	115	2^M

Table 8.1: Data of the Algorithm 2 applyed to the Motzkin polynomial.

random feasible solution of (P_6), $y \in \mathbb{R}^{s_6}$ with moment matrix $M := M_3(y)$. Compute, as in the Algorithm 2, the matrices B_M and T_M and solve the semidefinite program $(P_{M,1/60})$. We get after 66 iterations the optimal solution $y_0 \in \mathbb{R}^{s_6}$ with Moment matrix:

$$A_M := M_3(y_0) = \begin{array}{c} \\ 1 \\ X_1 \\ X_2 \\ X_1^2 \\ X_1X_2 \\ X_2^2 \\ X_1^3 \\ X_1^2X_2 \\ X_1X_2^2 \\ X_2^3 \end{array} \begin{array}{c} \begin{array}{cccccccccc} 1 & X_1 & X_2 & X_1^2 & X_1X_2 & X_2^2 & X_1^3 & X_1^2X_2 & X_1X_2^2 & X_2^3 \end{array} \\ \begin{pmatrix} 1.00 & 0.00 & 0.00 & 1.02 & 0.00 & 1.02 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 1.02 & 0.00 & 0.00 & 0.00 & 0.00 & 1.04 & 0.00 & 1.04 & 0.00 \\ 0.00 & 0.00 & 1.02 & 0.00 & 0.00 & 0.00 & 0.00 & 1.04 & 0.00 & 1.04 \\ 1.02 & 0.00 & 0.00 & 1.04 & 0.00 & 1.04 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 & 1.04 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 1.02 & 0.00 & 0.00 & 1.04 & 0.00 & 1.04 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 1.04 & 0.00 & 0.00 & 0.00 & 0.00 & 1.07 & 0.00 & 1.07 & 0.00 \\ 0.00 & 0.00 & 1.04 & 0.00 & 0.00 & 0.00 & 0.00 & 1.07 & 0.00 & 1.07 \\ 0.00 & 1.04 & 0.00 & 0.00 & 0.00 & 0.00 & 1.07 & 0.00 & 1.07 & 0.00 \\ 0.00 & 0.00 & 1.04 & 0.00 & 0.00 & 0.00 & 0.00 & 1.07 & 0.00 & 1.07 \end{pmatrix} \end{array}$$

Moreover we get that $U := \sum_{|\alpha| \leq 6} f_\alpha y_{0,\alpha} = 0.00156$. It turns out that this matrix is flat and therefore is an upper bound of the minimum $P^* \leq U$. Moreover we can extract the points a_i such that $f(a_i) = U$ as in the Algorithm 1and we get the factorization:

$$\begin{aligned} A_M = & \frac{1}{4} \cdot V_3 V_3^T(-1.0109, -1.0109) + \frac{1}{4} \cdot V_3 V_3^T(1.0109, -1.0109) + \\ & \frac{1}{4} \cdot V_3 V_3^T(-1.0109, 1.0109) + \frac{1}{4} \cdot V_3 V_3^T(1.0109, 1.0109) \end{aligned}$$

Therefore $f(\pm 1.0109, \pm 1.0109) = U$, and this four points are close to the minimizers of the Motzkin polynomial that are $(\pm 1, \pm 1)$. The table 8.1 shows the different data that we get by solving the semidefinite program $(P_{M,\lambda})$ for different values of $\lambda \in [0, 1]$. Note that if the algorithm 2 did not seem to reach a flat A_M then we manually stopped the algorithm when took more than 3 minutes. Note also that for $\lambda = \frac{3}{4}$ we get an approximately flat solution and we extract potential minimizers by block-simultaneous diagonalization the truncated multiplication operators. The algorithm for the block simultaneous diagonalization is described in [24, Algorithm 4.1].

Example 8.2.2. Let us consider the again the polynomial optimization problem of Example 6.7.3:

$$\begin{aligned} \text{minimize} \quad & f(x) = x_1^2 x_2^2 (x_1^2 + x_2^2 - 1) \\ \text{subject to} \quad & x_1, x_2 \in \mathbb{R} \end{aligned}$$

Since it holds $-1 \leq X_1^2 + X_2^2 - 1$ it also holds that $-1 \leq X_1^2 X_2^2 (X_1^2 + X_2^2 - 1)$, or what is the same $f + 1$ is nonnegative. $f + 1$ is also not a sum of squares as one can easily check with the so called *Gram-matrix method*, see [17, Lemma 3.8]. In fact, as we said in 6.7.3 one can easily compute with Calculus techniques that $P^* = f(\pm\frac{1}{\sqrt{3}}, \pm\frac{1}{\sqrt{3}}) = \frac{1}{27}$. An optimal solution $z \in \mathbb{R}^{s_6}$ of the Moment relaxation (P_6), has the following Moment matrix:

$$
M_3(z) = \begin{array}{c} \\ 1 \\ X_1 \\ X_2 \\ X_1^2 \\ X_1X_2 \\ X_2^2 \\ X_1^3 \\ X_1^2X_2 \\ X_1X_2^2 \\ X_2^3 \end{array}
\begin{array}{cccccc}
1 & X_1 & X_2 & X_1^2 & X_1X_2 & X_2^2 \\
1.00 & 0.00 & 0.00 & 2270.73 & -0.00 & 2270.73 \\
0.00 & 2270.73 & -0.00 & 0.00 & 0.00 & 0.00 \\
0.00 & -0.00 & 2270.73 & 0.00 & 0.00 & 0.00 \\
2270.73 & 0.00 & 0.00 & 10462942.40 & 0.00 & 411.34 \\
-0.00 & 0.00 & 0.00 & 0.00 & 411.34 & -0.00 \\
2270.73 & 0.00 & 0.00 & 411.34 & -0.00 & 10462943.55 \\
0.00 & 10462942.40 & 0.00 & 2.13 & 0.00 & 0.00 \\
0.00 & 0.00 & 411.34 & 0.00 & 0.00 & 0.00 \\
0.00 & 411.34 & -0.00 & 0.00 & 0.00 & 0.00 \\
0.00 & -0.00 & 10462943.55 & 0.00 & 0.00 & 2.50
\end{array}
$$

$$
\begin{array}{cccc}
X_1^3 & X_1^2X_2 & X_1X_2^2 & X_2^3 \\
0.00 & 0.00 & 0.00 & 0.00 \\
10462942.40 & 0.00 & 411.34 & -0.00 \\
0.00 & 411.34 & -0.00 & 10462943.55 \\
2.13 & 0.00 & 0.00 & 0.00 \\
0.00 & 0.00 & 0.00 & 0.00 \\
0.00 & 0.00 & 0.00 & 2.50 \\
165266760920.34 & 0.00 & 116.88 & 0.00 \\
0.00 & 116.88 & 0.00 & 116.88 \\
116.88 & 0.00 & 116.88 & -0.00 \\
0.00 & 116.88 & -0.00 & 165266795981.98
\end{array}
\begin{array}{c} \\ 1 \\ X_1 \\ X_2 \\ X_1^2 \\ X_1X_2 \\ X_2^2 \\ X_1^3 \\ X_1^2X_2 \\ X_1X_2^2 \\ X_2^3 \end{array}
$$

This matrix is neither flat nor $\widetilde{M_3(z)}$ is a generalized Hankel matrix so we can not conclude optimality. In fact the value of this optimal solution is:

$$P_6^* = -177.5859 << P^* = -\frac{1}{27} \approx -0.0370$$

stricly smaller than the minimum. According to the Algorithm 2 we compute the matrices B_M, T_M associated to a feasible solution M and finally we solve the semidefinite program $(P_{M,\frac{1}{100}})$ and we get the following matrix:

$$
A_M := M_3(y_0) = \begin{array}{c} \\ 1 \\ X_1 \\ X_2 \\ X_1^2 \\ X_1X_2 \\ X_2^2 \\ X_1^3 \\ X_1^2X_2 \\ X_1X_2^2 \\ X_2^3 \end{array}
\begin{array}{cccccccccc}
1 & X_1 & X_2 & X_1^2 & X_1X_2 & X_2^2 & X_1^3 & X_1^2X_2 & X_1X_2^3 & X_2^3 \\
1.00 & -0.00 & -0.00 & 0.41 & -0.00 & 0.41 & 0.00 & 0.00 & 0.00 & 0.00 \\
-0.00 & 0.41 & -0.00 & 0.00 & 0.00 & 0.00 & 0.17 & -0.00 & 0.16 & -0.00 \\
-0.00 & -0.00 & 0.41 & 0.00 & 0.00 & 0.00 & -0.00 & 0.16 & -0.00 & 0.17 \\
0.41 & 0.00 & 0.00 & 0.17 & -0.00 & 0.16 & -0.00 & -0.00 & 0.00 & 0.00 \\
-0.00 & 0.00 & 0.00 & -0.00 & 0.16 & -0.00 & -0.00 & 0.00 & 0.00 & -0.00 \\
0.41 & 0.00 & 0.00 & 0.16 & -0.00 & 0.17 & 0.00 & 0.00 & -0.00 & -0.00 \\
0.00 & 0.17 & -0.00 & -0.00 & -0.00 & 0.00 & 0.07 & -0.00 & 0.07 & -0.00 \\
0.00 & -0.00 & 0.16 & -0.00 & 0.00 & 0.00 & -0.00 & 0.07 & -0.00 & 0.07 \\
0.00 & 0.16 & -0.00 & 0.00 & 0.00 & -0.00 & 0.07 & -0.00 & 0.07 & -0.00 \\
0.00 & -0.00 & 0.17 & 0.00 & -0.00 & -0.00 & -0.00 & 0.07 & -0.00 & 0.07
\end{array}
$$

	U	Flat	$\|\|A_M - B_M\|\|$	Potential Minimizers	Iterations	Time
$\lambda = \frac{1}{1000}$	-4.4169	No	$7.8542 \cdot 10^4$	-	76	$2^M 30^S$
$\lambda = \frac{1}{100}$	-0.0305	Aproximately	$7.928 \cdot 10^{-4}$	$(\pm 0.6354, \pm 0.6354)$	85	$3^M 3^S$
$\lambda = \frac{1}{60}$	-0.0255	Yes	$9.9708 \cdot 10^{-5}$	$(\pm 0.6566, \pm 0.6566)$	76	$2^M 50^S$
$\lambda = \frac{1}{4}$	0.0802	No	0.0627	-	50	$1^M 10^S$
$\lambda = \frac{1}{2}$	0.1634	Yes	$9.5965 \cdot 10^{-5}$	$(\pm 0.8233, \pm 0.8233)$	123	3^M
$\lambda = \frac{3}{4}$	0.2399	No	0.0430	-	50	$1^M 20^S$
$\lambda = 1$	0.1331	No	0.0103	-	84	$3^M 11^S$

Table 8.2: Data of the Algorithm 2 for the polynomial $x_1^2 x_2^2 (x_1^2 + x_2^2 - 1)$

We diagonalize this matrix to compute the rank and we get that the eigenvalues are:

$$\{1.3284, 0.5383, 0.58383, 0.1630, 0.0024, 0.0008, 0.0008, 0.0000, 0.0000, 0.0000\}$$

what implies, after rounding to the fourth decimal, that the rank $A_M = 6$ and therefore A_M is flat and since we got that $U = \sum_{|\alpha| \leq 6} f_\alpha y_{0,\alpha} = -0.0305$ then we can deduce:

$$P^* \in [-177.5859, -0.0305]$$

So we proceed to extract minimizers with the Algortihm defined as we implemented in Algorithm 1 and we get that:

$$\begin{aligned} A_M =& 0.2501 \cdot V_3 V_3^T(-0.6354, -0.6354) + 0.2499 \cdot V_3 V_3^T(-0.6354, 0.6354) \\ &+ 0.2499 \cdot V_3 V_3^T(-0.6354, 0.6354) + 0.2501 \cdot V_3 V_3^T(0.6354, 0.6354) \end{aligned}$$

In conclusion we get that:

$$f(\pm 0.6354, \pm 0.6354) = -0.0305 \geq P^* = -0.0370 = f(\pm 0.5774, \pm 0.5774)$$

The table 8.2 shows the different data that we get by solving the semidefinite program $(P)_{M,\lambda}$ for different values of $\lambda \in [0, 1]$.

8.3 Algorithm 3: based on a Nie, Demmel and Sturmfels method

In this section we modify the Algorithm 2 by following ideas of the paper of Nie, Demmel and Sturmfels [13]. In the Algorithm 2 the starting matrix $M := M_d(y)$ was a generalized Hankel matrix positive semidefinite associated to a feasible solution $y \in \mathbb{R}^{s_d}$ of the moment relaxation $(P_{\deg f})$. In particular we will start this algorithm with a matrix associated to and optimal solution of the *NDS relaxation*, see 8.3.2 for a definition.

For $p \in \mathbb{R}[\underline{X}]_k$ consider the following vector:

$$\begin{aligned} pV_{s_{k-\deg p}} := (&p1, pX_1, pX_2, \dots, pX_n, pX_1^2, pX_1X_2, \dots, pX_1X_n, pX_2^2, \\ &pX_2X_3, \dots, pX_2^2, \dots, pX_1^{s_{k-\deg p}}, \dots, pX_{n-1}X_n^{s_{k-\deg p}-1}, pX_n^{s_{k-\deg p}})^T \end{aligned} \tag{8.8}$$

Definition 8.3.1. For $p \in \mathbb{R}[\underline{X}]_k$ the *localizing vector of p of degree k* is the vector resulting from substituting every monomial $\underline{X}^\alpha$ such that $|\alpha| \leq k$ in (8.8) for a new variable Y_α. We denote this vector by $V_{k,p} \in \mathbb{R}[\underline{Y}]_1^{s_{k-\deg p}}$.

Providing that $P^* = f(x)$ for some $x \in \mathbb{R}^n$, we know that the local and global minimizers are contained in the real gradient variety:

$$\mathscr{V}_f := \{u \in \mathbb{R}^n \mid \nabla f(u) = 0\}$$

In the paper [13], it is considered to add to the problem 8.1 the redundant constraints:

$$\begin{cases} \text{minimize} & f(x) \\ \text{subject to:} & x \in \mathbb{R}^n, \\ & p(x)^2 \geq 0 \text{ for all } p \in \mathbb{R}[\underline{X}]_d \text{ and} \\ & p\frac{\partial f}{\partial X_i}(x) = 0 \text{ for all } p \in \mathbb{R}[\underline{X}]_{k-\deg f+1} \end{cases}$$

After linearization, or in other words, after substitute every monomial $\underline{X}^\alpha$ for a new variable Y_α, this turned out into the following hierarchy of semidefinite programs.

Definition and Notation 8.3.2. Let (P) be a polynomial optimization problem as in (8.1) such that there exists $x \in \mathbb{R}^n$ with $P^* = f(x)$. We call the NDS relaxation of (P) of degree k to the following semidefinite optimization program:

$$(P_{k,\text{NDS}}) \begin{cases} \text{minimize} & \sum_{|\alpha| \leq k} f_\alpha y_\alpha \\ \text{subject to:} & y_{(0,\dots,0)} = 1, \\ & M_{\lfloor \frac{\deg k}{2} \rfloor}(y) \succeq 0 \text{ and} \\ & V_{k,\frac{\partial f}{\partial X_i}}(y) = 0 \text{ for all } i \in \{1,\dots,n\} \end{cases}$$

the optimal value of $(P_{k,\text{NDS}})$ that is to say, the infimum over all

$$y = (y_{(0,\dots,0)}, \dots, y_{(0,\dots,k)}) \in \mathbb{R}^{s_k}$$

that ranges over all feasible solutions of $(P_{k,\text{NDS}})$ is denoted by $P^*_{k,\text{NDS}} \in \{-\infty\} \cup \mathbb{R} \cup \{\infty\}$.

For all the details about the convergence of the NDS relaxation hierarchy we refer to the reader to [13].

Algorithm 3: Finding an upper bound U for (P) (8.1), $U \geq P^*$ using the NDS relaxation and if possible potential minimizers or minimizers

Input: A polynomial optimization problem $(\mathbf{P})$ (2.2) without constraints and a strategic $\lambda \in [0,1]$.

Output: An upper bound $U \geq \mathbf{P}^*$ and if possible, points $\mathbf{a}_1, \ldots, \mathbf{a}_r \in \mathscr{V}_f$ such that $f(\mathbf{a_i}) = U$.

1 Compute an optimal solution of the NDS relaxation $(P_{\deg f, \text{NDS}})$. Denote it by $y \in \mathbb{R}^{s_{\deg f}}$ and set $k := \deg f$ and $d := \lfloor \frac{\deg f}{2} \rfloor$ and $\mathbf{M} := M_d(y), .$;

2 Take a maximum linearly independent subset of the first s_{d-1} columns of $\mathbf{M}$ and denote it by $\{M_1, \ldots, M_l\}$ and set $C := (M_1 \cdots M_l).$;

3 Define the folowing matrix using the definition of the solution of the least squares problem (8.4), $\mathbf{B_M} := (M_1 \cdots M_{s_{d-1}} | Cx^*_{C, M_{s_{d-1}+1}} \cdots Cx^*_{C, M_{s_d}})$.;

4 Finally define the following matrix:

$$\mathbf{T_M} := \left(\begin{array}{ccc|c} & & & (A_M)_{1,1} - (B_M)_{1,1} \\ & I_{s_d s_d} & & \vdots \\ & & & (A_M)_{s_d,s_d} - (B_M)_{s_d,s_d} \\ \hline (A_M)_{1,1} - (B_M)_{1,1} & \cdots & (A_M)_{s_d,s_d} - (B_M)_{s_d,s_d} & E||M - B_M||^2 \end{array}\right)$$

where the unknows are $\mathbf{E}$ and the entries of the matrix $\mathbf{A_M} \in S\mathbb{R}^{s_d \times s_d}$. ;

5 Solve the following positive semidefinite program:

$$(P_{M,\lambda,\text{NDS}}) \begin{cases} \text{minimize} & E\lambda + (1-\lambda)\sum_{|\alpha| \leq k} f_\alpha y_{0,\alpha} \\ \text{subject to:} & \mathbf{A_M} = M_d(y_0), \\ & E \in \mathbb{R}\,, \mathbf{T_M} \succeq 0, \\ & \mathbf{A_M} \succeq 0 \\ & V_{k, \frac{\partial f}{\partial x_i}}(y_0) = 0 \text{ for all } i \in \{1, \ldots, n\} \end{cases}$$

6 Note that here the decision variables are E and the entries of the matrix $\mathbf{A_M} \in S\mathbb{R}^{s_d \times s_d}$ which are $y_0 \in \mathbb{R}^{s_k}$. ;

7 **if** $\boldsymbol{A_M}$ *is flat or* $||\boldsymbol{A_M} - \boldsymbol{B_M}||$ *small enough* **then**

8 $\quad$ $\mathbf{U} := \sum_{|\alpha| \leq k} f_\alpha y_{0,\alpha}$ and extract minimizers if possible with algorithm 1.

9 **else**

10 $\quad$ $\mathbf{M} := \mathbf{A_M}$ and go to **3**

Remark 8.3.3. In Algorithm 3 since $y_0 \in \mathbb{R}^{s_k}$ is a feasible solution of $(P_{k,\text{NDS}})$ then it holds $U \geq P^*_{k,\text{NDS}}$. Moreover if $\mathbf{A_M}$ is flat then by 4.4.4 (ii) together with Theorem 6.3.19, there exits $a_1, \ldots, a_N \in \mathbb{R}^n$ and $\lambda_1 > 0, \ldots, \lambda_N > 0$ with $N = \operatorname{rank} \mathbf{A_M}$ such that:

$$A_M = \sum_{i=1}^{N} \lambda_i V_d V_d^T(a_i) \in S\mathbb{R}^{s_d \times s_d}$$

But in contrast with Algorithm 2 we need to check that $a_1, \ldots, a_N \in \mathscr{V}_f$ to apply 4.4.4 (ii) in order to conclude:

$$P^* \in [P^*_{k,\text{NDS}}, U]$$

8.4 Numerical results of Algorithm 3

Example 8.4.1. Let us consider the problem of minimizing the so called *Robinson polynomial*:

$$X_1^6 + X_2^6 + 1 - (X_1^4 X_2^2 + X_2^4 + X_1^2 + X_1^2 X_2^4 + X_2^2 + X_1^2) + 3X_1^2 X_2^2$$

a nonnegative polynomial which is not sum of squares (see [30, 31]):

$$\begin{array}{ll} \text{minimize} & f(x) = x_1^6 + x_2^6 + 1 - (x_1^4 x_2^2 + x_2^4 + x_1^4 + x_1^2 x_2^4 + x_2^2 + x_1^2) + 3x_1^2 x_2^2 \\ \text{subject to} & x_1, x_2 \in \mathbb{R}^2 \end{array}$$

This polynomial attains the minimum in the NDS relaxation of degree 8, since we get the following optimal solution $z \in \mathbb{R}^{s_8}$:

(8.9) $M_4(z) =$

	1	X_1	X_2	X_1^2	X_1X_2	X_2^2	X_1^3	$X_1^2X_2$	$X_1X_2^3$	X_2^3
1	1.00	−0.00	0.00	0.64	0.00	0.64	−0.00	0.00	0.00	0.00
X_1	−0.00	0.64	0.00	−0.00	0.00	0.00	0.64	0.00	0.29	0.00
X_2	0.00	0.00	0.64	0.00	0.00	0.00	0.00	0.29	0.00	0.64
X_1^2	0.64	−0.00	0.00	0.64	0.00	0.29	−0.00	0.00	0.00	0.00
X_1X_2	0.00	0.00	0.00	0.00	0.29	0.00	0.00	0.00	0.00	0.00
X_2^2	0.64	0.00	0.00	0.29	0.00	0.64	0.00	0.00	0.00	0.00
X_1^3	−0.00	0.64	0.00	−0.00	0.00	0.00	0.64	0.00	0.29	0.00
$X_1^2X_2$	0.00	0.00	0.29	0.00	0.00	0.00	0.00	0.29	0.00	0.29
$X_1X_2^2$	0.00	0.29	0.00	0.00	0.00	0.00	0.29	0.00	0.29	0.00
X_2^3	0.00	0.00	0.64	0.00	0.00	0.00	0.00	0.29	0.00	0.64
X_1^4	0.64	−0.00	0.00	0.64	0.00	0.29	−0.00	0.00	0.00	0.00
$X_1^3X_2$	0.00	0.00	0.00	0.00	0.29	0.00	0.00	0.00	0.00	0.00
$X_1^2X_2^2$	0.29	0.00	0.00	0.29	0.00	0.29	0.00	0.00	0.00	0.00
$X_1X_2^3$	0.00	0.00	0.00	0.00	0.29	0.00	0.00	0.00	0.00	0.00
X_1^4	0.64	0.00	0.00	0.29	0.00	0.64	0.00	0.00	0.00	0.00

X_1^4	$X_1^3X_2$	$X_1^2X_2^2$	$X_1X_2^3$	X_2^4	
0.64	0.00	0.29	0.00	0.64	1
−0.00	0.00	0.00	0.00	0.00	X_1
0.00	0.00	0.00	0.00	0.00	X_2
0.64	0.00	0.29	0.00	0.29	X_1^2
0.00	0.29	0.00	0.29	0.00	X_1X_2
0.29	0.00	0.29	0.00	0.64	X_2^2
−0.00	0.00	0.00	0.00	0.00	X_1^3
0.00	0.00	0.00	0.00	0.00	$X_1^2X_2$
0.00	0.00	0.00	0.00	0.00	$X_1X_2^2$
0.00	0.00	0.00	0.00	0.00	X_2^3
1.01	−0.00	0.65	−0.00	0.65	X_1^4
−0.00	0.65	−0.00	0.65	−0.00	$X_1^3X_2$
0.65	−0.00	0.65	−0.00	0.65	$X_1^2X_2^2$
−0.00	0.65	−0.00	0.65	−0.00	$X_1X_2^3$
0.65	−0.00	0.65	−0.00	1.01	X_2^4

with optimal value $P^*_{8,\text{NDS}} = 1.3558 \cdot 10^{-10} \approx 0$. We can conclude optimality since $\widetilde{M_4(y)}$ is a generalized Hankel matrix, or equivalently, the truncated multiplication operators commute and we can get with the Algorithm 1 8 potential minimizers:

$$(\pm 1, \pm 1), (\pm 1, 0), (0, \pm 1) \in \mathscr{V}_f.$$

Therefore with the NDS relaxation of degree 8 by Theorem 7.1.1 we have:

$$P^* = P^*_{8,\text{NDS}} = f(\pm 1, \pm 1) = f(\pm 1, 0) = f(0, \pm 1) = 0.$$

Let us now use the Algorithm 3 to see if we can get an upper bound for the minimum and potential miminimizers starting with a moment matrix of lower dimension which is an optimal solution of the NDS relaxation of degree 6. Namely, we start with the moment matrix:

$\mathbf{M} := M_3(y) =$

	1	X_1	X_2	X_1^2	X_1X_2	X_2^2	X_1^3	$X_1^2X_2$	$X_1X_2^3$	X_2^3
1	1.00	0.00	0.00	1.39	0.00	1.39	0.00	0.00	−0.00	0.00
X_1	0.00	1.39	0.00	0.00	0.00	−0.00	2.37	0.00	1.50	0.00
X_2	0.00	0.00	1.39	0.00	−0.00	0.00	0.00	1.50	0.00	2.37
X_1^2	1.39	0.00	0.00	2.37	0.00	1.50	−0.00	0.00	−0.00	0.00
X_1X_2	0.00	0.00	−0.00	0.00	1.50	0.00	0.00	−0.00	0.00	−0.00
X_2^2	1.39	−0.00	0.00	1.50	0.00	2.37	−0.00	0.00	−0.00	0.00
X_1^3	0.00	2.37	0.00	−0.00	0.00	−0.00	3946.00	−0.00	3945.46	−0.00
$X_1^2X_2$	0.00	0.00	1.50	0.00	−0.00	0.00	−0.00	3945.46	−0.00	3945.46
$X_1X_2^2$	−0.00	1.50	0.00	−0.00	0.00	−0.00	3945.46	−0.00	3945.46	−0.00
X_2^3	0.00	0.00	2.37	0.00	−0.00	0.00	−0.00	3945.46	−0.00	3946.00

with optimal value $P^*_{6,\text{NDS}} = -0.9333$. After applying the Algorithm (3) we get the following results, see Table 8.3. Here we obviously get that $P^*_{6,\text{NDS}} = -0.9333 \leq P^*_{M,\lambda,\text{NDS}}$, but we can no certify the belonging to any interval since unfortunately we do not get neither flat nor the truncated multiplication operators commute.

Example 8.4.2. Let us consider again the problem of minimizing the Motzkin polynomial:

$$\begin{array}{ll} \text{minimize} & f(x) = x_1^4 x_2^2 + x_1^2 x_2^4 - 3x_1^2 x_2^2 + 1 \\ \text{subject to} & x \in \mathbb{R}^2 \end{array}$$

	U	Flat	$\|\|A_M - B_M\|\|$	Potential Minimizers	Iterations	Time
$\lambda = \frac{1}{100}$	-0.9278	No	$7.329 \cdot 10^3$	-	10	33^S
$\lambda = \frac{1}{60}$	-0.9288	No	$7.1978 \cdot 10^3$	-	10	40^S
$\lambda = \frac{1}{30}$	-0.5709	No	$6.8149 \cdot 10^3$	-	10	50^S
$\lambda = \frac{1}{10}$	-0.0159	No	$6.1421 \cdot 10^3$	-	10	48^S
$\lambda = \frac{1}{5}$	0.0549	No	$5.9942 \cdot 10^3$	-	10	36^S
$\lambda = \frac{1}{4}$	0.0670	No	$6.0629\ddot{1}0^3$	-	500	40^M
$\lambda = \frac{3}{4}$	0.1018	No	$5.8919 \cdot 10^3$	-	10	38^S
$\lambda = 1$	1.060	No	$5.7389 \cdot 10^3$	-	10	36^S

Table 8.3: Data of the Algorithm 3 for the Robinson polynomial

As it was mentioned in [13] this polynomial attains the minimum in the relaxation $(P_{8,\text{NDS}})$. After applying the Algorithm 3 we get the following information, see Table 8.4. Notice that we start with an optimal solution, associated to the matrix M, of the relaxation $(P_{6,\text{NDS}})$. More precise we start the algorithm with the following matrix:

$$M := M_3(y) = \begin{array}{c|cccccc} & 1 & X_1 & X_2 & X_1^2 & X_1X_2 & X_2^2 \\ \hline 1 & 1.00 & -0.00 & -0.00 & 663.07 & -0.00 & 663.07 \\ X_1 & -0.00 & 663.07 & -0.00 & -0.00 & -0.00 & 0.00 \\ X_2 & -0.00 & -0.00 & 663.07 & -0.00 & 0.00 & -0.00 \\ X_1^2 & 663.07 & -0.00 & -0.00 & 891655.96 & -0.00 & 423.13 \\ X_1X_2 & -0.00 & -0.00 & 0.00 & -0.00 & 423.13 & -0.00 \\ X_2^2 & 663.07 & 0.00 & -0.00 & 423.13 & -0.00 & 891655.96 \\ X_1^3 & -0.00 & 891655.96 & -0.00 & -0.00 & 0.00 & -0.00 \\ X_1^2X_2 & -0.00 & -0.00 & 423.13 & 0.00 & -0.00 & -0.00 \\ X_1X_2^2 & 0.00 & 423.13 & -0.00 & -0.00 & -0.00 & 0.00 \\ X_2^3 & -0.00 & -0.00 & 891655.96 & -0.00 & 0.00 & -0.00 \end{array}$$

$$\begin{array}{cccc|c} X_1^3 & X_1^2X_2 & X_1X_2^2 & X_2^3 & \\ \hline -0.00 & -0.00 & 0.00 & -0.00 & 1 \\ 891655.96 & -0.00 & 423.13 & -0.00 & X_1 \\ -0.00 & 423.13 & -0.00 & 891655.96 & X_2 \\ -0.00 & 0.00 & -0.00 & -0.00 & X_1^2 \\ 0.00 & -0.00 & -0.00 & 0.00 & X_1X_2 \\ -0.00 & -0.00 & 0.00 & -0.00 & X_2^2 \\ 4116203733.28 & -0.00 & 423.13 & -0.00 & X_1^3 \\ -0.00 & 423.13 & -0.00 & 423.13 & X_1^2X_2 \\ 423.13 & -0.00 & 423.13 & -0.00 & X_1X_2^2 \\ -0.00 & 423.13 & -0.00 & 4116203751.60 & X_2^3 \end{array}$$

with associated optimal value $P^*_{6,\text{NDS}} = -422,13$.

As we can see in the Table 8.4, we get $P^*_{M,1,\text{NDS}} = 1$ after 20 iterations of the Algorithm 3. Namely we get the following flat optimal solution of $(P_{M,1,\text{NDS}})$:

	U	Flat	$\|\|A_M - B_M\|\|$	Potential Minimizers	Iterations	Time
$\lambda = \frac{1}{1000}$	-30.6672	No	$1.1486 \cdot 10^6$	-	50	$4^M 42^S$
$\lambda = \frac{1}{100}$	0.1688	No	$1.1486 \cdot 10^6$	-	10	50^S
$\lambda = \frac{1}{60}$	0.2458	No	$1.4886 \cdot 10^6$	-	10	40^S
$\lambda = \frac{1}{4}$	-0.8789	No	$1.4887 \cdot 10^6$	-	50	4^M
$\lambda = \frac{1}{2}$	0.4232	No	$1.1485 \cdot 10^6$	-	10	41^S
$\lambda = \frac{3}{4}$	1	Yes	$1.1629 \cdot 10^6$	$(0, \pm 44.9350), (\pm 44.9350, 0)$	50	$4^M 42^S$
$\lambda = 1$	1	Yes	$1.1629 \cdot 10^6$	$(0, \pm 44.9352), (\pm 44.9352, 0)$	20	$1^M 20^S$

Table 8.4: Data of the Algorithm 3 for the Motzkin polynomial

$A_M = M_3(y_0) =$

	1	X_1	X_2	X_1^2	X_1X_2	X_2^2	X_1^3	$X_1^2X_2$	$X_1X_2^2$	X_2^3
1	1.00	0.00	−0.00	1009.61	0.00	1009.61	0.02	0.00	−0.00	−0.04
X_1	0.00	1009.61	0.00	0.02	0.00	−0.00	2038565.01	−0.00	0.00	−0.00
X_2	−0.00	0.00	1009.61	0.00	−0.00	−0.04	−0.00	0.00	−0.00	2038564.29
X_1^2	1009.61	0.02	0.00	2038565.01	−0.00	0.00	−0.03	−0.00	0.00	0.00
X_1X_2	0.00	0.00	−0.00	−0.00	0.00	−0.00	−0.00	0.00	0.00	−0.00
X_2^2	1009.61	−0.00	−0.04	0.00	−0.00	2038564.29	0.00	0.00	−0.00	−0.30
X_1^3	0.02	2038565.01	−0.00	−0.03	−0.00	0.00	4116200366.80	0.00	0.00	−0.00
$X_1^2X_2$	0.00	−0.00	0.00	−0.00	0.00	0.00	0.00	0.00	−0.00	0.00
$X_1X_2^2$	−0.00	0.00	−0.00	0.00	0.00	−0.00	0.00	−0.00	0.00	0.00
X_2^3	−0.04	−0.00	2038564.29	0.00	−0.00	−0.30	−0.00	0.00	0.00	4116200384.54

Since A_M is flat, by Theorem 7.1.1, we can find the decomposition:

$$A_M = 0 \cdot V_3 V_3^T(0, -44.9351) + \frac{1}{2} \cdot V_3 V_3^T(-44.9351, 0) + 0 \cdot V_3 V_3^T(44.9351, 0) + \frac{1}{2} \cdot V_3 V_3^T(0, 44.9351)$$

Moreover since $(-44.9351, 0), (0, 44.9351) \in \mathscr{V}_f$ we can conclude by Theorem 7.1.1, that:

$$P^* \in [-422.13, 1]$$

Software

To find an optimal solution of the moment relaxation and for the big calculations we have used the following softwares:

- YALMIP: developed by J. Löfberg. It is a toolbox for Modeling and Optimization in MATLAB. Published in the Journal Proceedings of the CACSD Conference in 2004. For more information see: http://yalmip.github.io/.
- SEDUMI: developed by J. F. Sturm. It is a toolbox for optimization over symmetric cones. Published in the Journal Optimization Methods and Software in 1999. For more information see: http://sedumi.ie.lehigh.edu/.
- MATLAB and Statistics Toolbox Release 2016a, The MathWorks, Inc., Natick, Massachusetts, United States.

Bibliography

[1] A. Barvinok. *A course in convexity*, volume 54. Editorial Board. Graduate Studies in Mathematics, Providence, Rhode Island, 2002. 16, 17

[2] Christian Bayer and Josef Teichmann. The proof of Tchakaloff's Theorem. *Proc. Amer. Math. Soc.*, 134(10):3035–3040, May 2006. 38, 56

[3] C. Riener and M. Schweighofer. Optimization approaches to quadrature: new characterizations of Gaussian quadrature on the line and quadrature with few nodes on plane algebraic curves, on the plane and in higher dimensions. *arxiv.org/abs/1607.08404*. 68

[4] R. E. Curto and L. A. Fialkow. Solution of the truncated complex moment problem for flat data. *Mem. Amer. Math. Soc.*, 119(568), 1996. 13, 46, 55

[5] R. E. Curto and L. A. Fialkow. The truncated complex K-moment problem. *Trans. Amer. Math. Soc*, 352(6):2825–2855, 2000. 13, 69, 78

[6] R. E. Curto and L. A. Fialkow. Solution of the singular quartic moment problem. *Journal of Operator Theory*, 48(2):315–354, 2002. 37

[7] R. E. Curto and S. Yoo. Concrete solution to the nonsigular quartic binary moment problem. *Proc. Amer. Math. Soc*, 144(1):249–259, 2016. 37

[8] D. A. Cox, J. Little, D. O'Shea. *Ideals, Varieties and Algorithms. An introduction to Computational Algebraic Geometry and Commutative Algebra*. Springer Science, NY 10013, 2007. 45, 46

[9] J. B. Lasserre D. Henrion. Detecting global optimality and Extracting Solutions in Gloptipoly. *In: Henrion D., Garulli A. (eds) Positive Polynomials in Control. Lecture Notes in Control and Information Science,Springer, Berlin, Heidelberg*, 312:293–310, 2005. 69

[10] C. F. Dunkl and Y. Xu. *Orthogonal polynomials of several variables*. Cambridge University Press, CB2 8BS, 2014. 61

[11] C. A. Floudas and P.M Pardalos. *A collection of Test Problems for Constrained Global Optimization Algorithms*. Springer-Verlag, Druckhaus Beltz, Germany, 1990. 54, 74

[12] D. Henrion and J. B. Lasserre. Gloptipoly: Global optimization over polynomials with MATLAB and SEDUMI. *ACM Trans. Math. Software*, 29(2):165–194, 2003. 69

[13] J. Demmel J. Nie and B. Sturmfels. Minimizing Polynomials via Sum of Squares over the Gradient Ideal. *Math. Program., Ser.A*, 106:587–106, May 2006. 89, 90, 94

[14] K. R. Davidson. *C*-Algebras by Example*, volume 6. Field Inst. Monographs, 1996. 36

[15] E. Klerk, M. Laurent, and Z. Sun. Convergence analysis for Lasserre's measure-based hierarchy of upper bounds for polynomial optimization. *Math. Program.*, 162, 2017. 73

[16] J. B. Lasserre. Global optimization with polynomials and the problems of moments. *SIAM J.Optim*, 11:796–817, 2001. 23, 64, 77

[17] M. Laurent. Sums of squares, moment matrices and optimization over polynomials. *The IMA Volumes in Mathematics and its Applications*, 149:157–270, 2008. 18, 23, 38, 45, 46, 56, 69, 88

[18] M. Laurent. Revisiting two theorems of Curto and Fialkow on moment matrices. *Proc. Amer. Math. Soc*, 10:2965–2976, 2009. 56, 70

[19] M. Laurent and F. Vallentin. *Semidefinite Optimization*. Notes for Mastermath 2012, The Netherlands, March 2012. 19, 84

[20] M. Schweighofer. *Real Algebraic Geometry, Positivity and Convexity*. Lecture Notes in University of Konstanz, Germany, August 2017. 17

[21] M. Takesaki. *Theory of Operators Algebras I*, volume 1. Springer Science and Business Media, 2001. 36

[22] M. Marshall. *Positive Polynomials and Sums of Squares*. American Mathematical Society, Rhode Island 02904-2294, 2008. 16, 86

[23] H. M. Möller. Kubaturformeln mit minimaler Knotenzahl (in german). *Numer. Math.*, 25:185–200, 1976. 62

[24] M. Murota, Y. Kanno, M. Kojima, and S. Kojima. A numerical Algorithm for Block-Diagonal Decomposition of Matrix*-Algebras, Part I: Proposed Approach and Application to semidefinite programming. *Jpn.J.Ind.Appl.Math.*, 27:125–160, 2010. 53, 78, 87

[25] K. G. Murty and S. N. Kabadi. Some NP-Complete Problems in quadratic and nonlinear programming. *Math. Program.*, 39:117–129, May 1987. 82

[26] I.P. Mysovskikh. Interpolatory Cubature Formulas, Nauka, Moscow (in Russian). Interpolatorische Kubaturformel, Institut fur Geometrie und Praktische Mathematik der RWTH Aachen (in German). *IGPM*, 74, 1981. 61

[27] V. Powers and C. Scheiderer. The moment problem for non-compact semialgebraic sets. *Adv. Geom.*, 1(1):71–88, 2001. 30

[28] M. Putinar. Positive polynomials on compact semi-algebraic sets. *Indiana Univ. Math. J.*, 42(3):969–984, 1993. 33

[29] M. Putinar. A Dilation theory Approach to Cubature Formulas. *Expo. Math*, 15:183–192, 1997. 14, 39, 46, 61, 62

[30] B. Reznick. On Hilbert's construction of positive polynomials. *Contem. Mathematics Amer. Math. Soc.*, 253:251–272, 2000. 92

[31] R. M. Robinson. Some Definite Polynomials Which are not Sums of Squares of Real Polynomials. *Selected Questions of Algebra and Logic (Collection Dedicated to the Memory of A.I. Malcev) (Russian)*, pages 264–282, 1973. 92

[32] S. Boyd and L. Vandenberghe. *Vectors, matrices and Least Squares (working title)*. Draft, California, 2017. 83

[33] M. Schweighofer. Optimization of polynomials on compact semialgebraic sets. *SIAM J.Optim*, 15:805–825, 2005. 23, 24, 26, 30, 33

[34] M. Schweighofer and I. Klep. An Exact Duality Theory for Semidefinite Programming Based on Sums of Squares . *Math. Oper. Res.*, 38:569–590, 2013. 18

[35] S.Kuhlmann and M.Marshall. Positivity, sums of squares and the multi-dimensional moment problem. *Trans. Amer. Math. Soc*, 354(11):4285–4301, 2002. 30

[36] J. L. Smul'jan. An operator Hellinger integral (Russian). *Mat Sb*, 91:381–430, 1959. 13