



HAL
open science

Some approximation schemes in polynomial optimization

Roxana Hess

► **To cite this version:**

Roxana Hess. Some approximation schemes in polynomial optimization. Optimization and Control [math.OA]. Université Paul Sabatier - Toulouse III, 2017. English. NNT : 2017TOU30129 . tel-01809192v2

HAL Id: tel-01809192

<https://laas.hal.science/tel-01809192v2>

Submitted on 19 Oct 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université Fédérale



Toulouse Midi-Pyrénées

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ FÉDÉRALE TOULOUSE MIDI-PYRÉNÉES

Délivré par :

l'Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)

Présentée et soutenue le *28/09/2017* par :

ROXANA HESS

SOME APPROXIMATION SCHEMES IN POLYNOMIAL OPTIMIZATION

JURY

JÉRÔME BOLTE	PU Université Toulouse 1 Capitole	Examineur
STÉPHANE GAUBERT	DR INRIA Saclay-Île-de-France	Examineur
DIDIER HENRION	DR LAAS-CNRS, Toulouse	Directeur de thèse
ETIENNE DE KLERK	PU Tilburg University	Examineur
KRZYSZTOF KURDYKA	PU Université Savoie Mont-Blanc	Rapporteur
JEAN BERNARD LASSERRE	DR LAAS-CNRS, Toulouse	Directeur de thèse
MICHAEL OVERTON	PU New York University	Rapporteur
MARKUS SCHWEIGHOFER	PU Universität Konstanz	Examineur
LIEVEN VANDENBERGHE	PU Univ. of California, Los Angeles	Rapporteur

École doctorale et spécialité :

EDSYS : Automatique 4200046

Unité de Recherche :

Laboratoire d'analyse et d'architecture des systèmes

Directeur(s) de Thèse :

Didier HENRION et Jean Bernard LASSERRE

Rapporteurs :

Krzysztof KURDYKA, Michael OVERTON et Lieven VANDENBERGHE

Titre: QUELQUES SCHÉMAS D'APPROXIMATION EN OPTIMISATION POLYNOMIALE

Résumé: Cette thèse est dédiée à l'étude de la hiérarchie moments-sommes-de-carrés, une famille de problèmes de programmation semi-définie en optimisation polynomiale, couramment appelée hiérarchie de Lasserre. Nous examinons différents aspects de ses propriétés et applications. Comme application de la hiérarchie, nous approchons certains objets potentiellement compliqués, comme l'abscisse polynomiale et les plans d'expérience optimaux sur des domaines semi-algébriques. L'application de la hiérarchie de Lasserre produit des approximations par des polynômes de degré fixé et donc de complexité bornée. En ce qui concerne la complexité de la hiérarchie elle-même, nous en construisons une modification pour laquelle un taux de convergence amélioré peut être prouvé. Un concept essentiel de la hiérarchie est l'utilisation des modules quadratiques et de leurs duaux pour appréhender de manière flexible le cône des polynômes positifs et le cône des moments. Nous poursuivons cette idée pour construire des approximations étroites d'ensembles semi-algébriques à l'aide de séparateurs polynomiaux.

Mots-clés: optimisation non-convexe, optimisation non-lisse, approximations polynomiales, optimisation semi-algébrique, optimisation semi-définie positive

Title: SOME APPROXIMATION SCHEMES IN POLYNOMIAL OPTIMIZATION

Abstract: This thesis is dedicated to investigations of the moment-sums-of-squares hierarchy, a family of semidefinite programming problems in polynomial optimization, commonly called the Lasserre hierarchy. We examine different aspects of its properties and purposes. As applications of the hierarchy, we approximate some potentially complicated objects, namely the polynomial abscissa and optimal designs on semialgebraic domains. Applying the Lasserre hierarchy results in approximations by polynomials of fixed degree and hence bounded complexity. With regard to the complexity of the hierarchy itself, we construct a modification of it for which an improved convergence rate can be proved. An essential concept of the hierarchy is to use quadratic modules and their duals as a tractable characterization of the cone of positive polynomials and the moment cone, respectively. We exploit further this idea to construct tight approximations of semialgebraic sets with polynomial separators.

Key words: non-convex optimization, non-smooth optimization, polynomial approximations, semialgebraic optimization, semidefinite programming

Acknowledgments

First of all, I want to thank my supervisors Didier Henrion and Jean Bernard Lasserre for three pleasant years of PhD. In particular, I want to thank Didier for patiently explaining all sorts of mathematical objects to me, and also for the support in administrative questions and the French translation service. Great thanks to Jean Bernard for the enthusiasm and dedication with which he answers mathematical questions. Merci !

In autumn 2015, I was at CWI in Amsterdam for a stay abroad. Large parts of the paper on which Chapter 5 is based was written there. I want to thank Monique Laurent for her reliable and supporting way of supervising my stay, and Etienne de Klerk for the productive collaboration. Also, I thank CWI and, in particular, the team “Networks and Optimization” for their hospitality. The stay was funded by Université Paul Sabatier, École Doctorale Systèmes and École des Docteurs de l’Université Fédérale Toulouse Midi-Pyrénées.

I would like to thank Tien Son Pham for the nice collaboration on the polynomial abscissa, and Milan Korda, Silviu Filip, and Mioara Joldes for the subsequent discussions on uniform approximations. Contributions by Stéphane Gaubert to approximations with quadratic separators were very much appreciated. The paper on optimal designs, whose results are reported in Chapter 3, was written in collaboration with Fabrice Gamboa and Yohann de Castro. Thank you very much. The paper also profited from feedback from Henry Wynn, Pierre Maréchal, Luc Pronzato, Lieven Vandenberghe and Weng Kee Wong.

My sincere thanks to my fellow PhD-students, who kept me good company during these three years, especially Harmony Leduc and Clément Gazzino (who is also a great dictionary). Thank you to LAAS-CNRS and all the équipe MAC for providing a pleasant working environment (and reliably eating my cakes).

Special thanks to Tillmann Weißer, with whom it is always a pleasure to discuss measures, quadratic modules and GloptiPoly (und Gott und die Welt) and whose very constructive and thorough way of proofreading considerably improved the quality of this thesis.

Last, I want to cordially thank my family for always supporting me in every possible way. Danke!

Contents

Nomenclature	ix
Introduction	1
1 The Lasserre hierarchy	5
1.1 Problem formulation	5
1.2 Sums of squares and quadratic modules	6
1.3 Measures and moments	8
1.3.1 Moments and the moment cone	8
1.3.2 Moment and localizing matrices	9
1.4 Hierarchies of upper and lower bounds	9
1.4.1 Hierarchy of lower bounds	10
1.4.2 Hierarchy of upper bounds	11
2 Semidefinite approximations of the polynomial abscissa	15
2.1 The abscissa function	16
2.2 Upper abscissa approximation	17
2.2.1 Primal and dual formulation	17
2.2.2 Lasserre hierarchy	19
2.2.3 Examples	21
2.3 Lower abscissa approximation	24
2.3.1 Lower approximation via elementary symmetric functions	26
2.3.2 Lower approximation via Gauß-Lucas	31
2.4 Concluding remarks	37
3 Approximate polynomial optimal designs	39
3.1 Preliminaries	42
3.1.1 Polynomial optimal design	42
3.1.2 Approximations of the moment cone	43
3.2 Approximate Optimal Design	43
3.2.1 Relaxing the integrality constraints	44
3.2.2 Moment formulation	44
3.3 The ideal problem on moments and its approximation	45
3.3.1 The ideal problem on moments	45
3.3.2 Christoffel polynomials	49
3.3.3 The SDP relaxation scheme	50
3.3.4 Asymptotics	53
3.3.5 General regression polynomial bases	54
3.4 Recovering the measure	55
3.4.1 Via the Nie method	55
3.4.2 Via Christoffel-like polynomials	56
3.4.3 Calculating the corresponding weights	56
3.5 Examples	56
3.5.1 Univariate unit interval	57
3.5.2 Wynn's polygon	58

3.5.3	The 3-dimensional unit sphere	60
3.6	Concluding remark	61
4	Approximating semialgebraic sets with quadratic separators	63
4.1	Separators of arbitrary degree	64
4.2	Linear separators	65
4.2.1	Exactness	65
4.2.2	Examples	66
4.3	Quadratic separators	67
4.3.1	Exactness	68
4.3.2	Examples	68
4.4	Concluding remarks	69
5	Improved convergence rates for Lasserre-type hierarchies	71
5.1	A Lasserre-type hierarchy with error bound $O(\frac{1}{d^2})$	72
5.1.1	Background results	72
5.1.2	The hierarchy and main result	73
5.2	Background on the polynomial kernel method	75
5.2.1	Chebyshev polynomials	75
5.2.2	Jackson kernels	77
5.2.3	Jackson kernel approximation of the Dirac delta function	79
5.3	Convergence analysis	80
5.3.1	The univariate case	80
5.3.2	The multivariate case	81
5.4	Computing the parameter ϑ_d^h as a generalized eigenvalue problem	83
5.5	Numerical examples	84
Appendix	86
5.A	Proof of Lemma 5.4	86
5.B	Useful identities for the Chebyshev polynomials	88
5.C	Test functions	89
Bibliography		91

Nomenclature

x	Variable $x = (x_1, \dots, x_n)$, page 5.
x^α	Monomial $x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ to the power $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$, page 8.
$\mathbf{v}_d(x)$	Column vector of monomials up to degree d , page 8.
$\deg p$	Degree of the polynomial p , page 5.
$ \alpha $	Absolute value of $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$, $ \alpha = \alpha_1 + \cdots + \alpha_n$, page 8.
$s(d)$	Number of monomials of degree up to d , i.e., $\binom{n+d}{n}$, page 8.
$\lfloor a \rfloor, \lceil a \rceil$	Floor and ceiling function of $a \in \mathbb{R}$, i.e., $\max\{z \in \mathbb{Z} : z \leq a\}$ and $\min\{z \in \mathbb{Z} : z \geq a\}$, respectively, page 9.
\mathbf{K}	Basic closed semialgebraic set defined by $g_j \in \mathbb{R}[x]$, $j = 1, \dots, m$, of which one is of the form $R - \sum_{i=1}^n x_i^2$ for a constant $R \in \mathbb{N}$, page 5.
$\mathcal{C}(\mathbf{K})$	Space of continuous functions on \mathbf{K} , page 8.
$\mathcal{C}_+(\mathbf{K})$	Cone of non-negative continuous functions on \mathbf{K} , page 8.
$\mathcal{M}(\mathbf{K})$	Space of finite (signed) Borel measures supported on \mathbf{K} , page 8.
$\mathcal{M}_+(\mathbf{K})$	Cone of non-negative Borel measures supported on \mathbf{K} , page 8.
$\mathbb{R}[x]$	Vector space of real polynomials in the variable x , page 5.
$\mathbb{R}[x]_d$	Vector space of polynomials in $\mathbb{R}[x]$ of degree at most d , page 5.
$\mathcal{P}(\mathbf{K})$	Cone of polynomials in $\mathbb{R}[x]$ which are non-negative on \mathbf{K} , page 6.
$\mathcal{P}_d(\mathbf{K})$	Cone of polynomials non-negative on \mathbf{K} of degree at most d , page 7.
$\mathcal{M}(\mathbf{K})$	Moment cone of \mathbf{K} , page 8.
$\mathcal{M}_d(\mathbf{K})$	Moment cone of \mathbf{K} of moments up to order d , page 8.
$\Sigma[x]$	Cone of sums of squares of polynomials, page 6.
$\Sigma[x]_{2d}$	Cone of sums of squares of degree at most $2d$, page 6.
$\mathcal{Q}(g_1, \dots, g_m)$	Quadratic module generated by g_1, \dots, g_m , page 6.
$\mathcal{Q}_{2d}(g_1, \dots, g_m)$	Truncation of $\mathcal{Q}(g_1, \dots, g_m)$ at degree $2d$, page 7.
$\mathcal{M}(g_1, \dots, g_m)$	Dual cone of $\mathcal{Q}(g_1, \dots, g_m)$, page 9.
$\mathcal{M}_{2d}(g_1, \dots, g_m)$	Truncation of $\mathcal{M}(g_1, \dots, g_m)$ at degree $2d$, page 9.
$\mathbf{M}_d(\mathbf{y})$	Moment matrix of $\mathbf{y} = (y_\alpha)_{ \alpha \leq 2d}$ of size $s(d) \times s(d)$, page 9.
$\mathbf{M}_d(g, \mathbf{y})$	Localizing matrix w.r.t. g and \mathbf{y} of size $s(d) \times s(d)$, page 9.

Introduction

Polynomial optimization addresses the problem of optimizing a polynomial function over a given semialgebraic set, i.e., a set defined by finitely many polynomial inequalities. If not indicated otherwise, we consider minimization problems. Evidently, we can pass to maximization problems by simply negating the objective function.

In the literature several methods to find local minima have been proposed. However, when considering non-linear non-convex functions, searching for global optima is difficult in general. For instance, problems like the well-known NP-hard maximum stable set and maximum cut problems from graph theory are contained in this class of problems.

In the situations we encounter, we are confronted with highly non-convex problems, as we seek to optimize polynomials of arbitrary degree.

Lasserre [Las01] developed a hierarchy of semidefinite programs which allows to numerically approximate from below the global minimum of a polynomial on a compact basic semialgebraic set. This is in contrast with methods computing local minima, as these are upper bounds without any indication on the distance to the global optimum. Conversely, the lower bounds computed via the Lasserre hierarchy are proved to converge to the global minimum.

The hierarchy is constructed by first rephrasing the global optimization problem in terms of non-negative polynomials, and then replacing the non-negativity constraint by a certificate of positivity such that, after truncation, we obtain a problem which is *tractable*, i.e., a problem which can be solved numerically in polynomial time. More precisely, the obtained problems are semidefinite programs. The resulting sequence of solutions converges to the global optimum from below.

The convergence of the solutions is ensured by Putinar's Positivstellensatz, which can be considered as a key result in the construction of the hierarchy. It provides a representation of positive polynomials which is used as the above mentioned certificate of positivity.

The approach has a dual point of view since positive polynomials are dual to moment sequences, i.e., sequences of real numbers which have a representing measure. Thus, instead of rephrasing the optimization problem in terms of positive polynomials, we can also write it in terms of measures, or rather moments. By the dual facet of Putinar's Positivstellensatz, its feasible set can be described by infinitely many semidefinite constraints. When truncated, the problem therefore again leads to a hierarchy of semidefinite programs. Due to duality, the solutions are lower bounds of the infimum as well. They also converge to the global optimum from below.

Since by default we consider minimization problems, we call the hierarchy just described the *Lasserre hierarchy of lower bounds* since its solutions converge to the infimum from below. In contrast, in [Las11] Lasserre proposed a *hierarchy of upper bounds*, which is established using a characterization of non-negativity involving measures that allow sum-of-squares densities. The sequence of solutions to the so constructed hierarchy of semidefinite programs converges to the infimum from above.

Other hierarchies

Apart from Lasserre, several other authors have proposed hierarchies to approximate polynomial optimization problems. For further references and detailed information on the history of hierarchies using representations of non-negativity, the interested reader is referred to [Las15b, Section 6.5] and [Lau09], and also [Lau03].

Here we just want to mention that the idea to formulate global minimization of a polynomial as a convex optimization problem goes back to Shor [Sho87],[Sho98]. Nesterov [Nes00] then proposed exact semidefinite formulations for the univariate case. The multivariate case was treated by Parrilo [Par00],[Par03] and Lasserre [Las01]. An advantage of the investigation by Lasserre was the comprehensive treatment of the duality between positive polynomials and moments, which allowed to derive a mathematically sound proof of convergence of the hierarchy based on Putinar's Positivstellensatz.

Besides, in the meantime there has been developed a vast collection of modifications and related hierarchies which serve diverse purposes. In addition, extensions to the complex case and versions exploiting symmetry or sparsity have been proposed.

Contents of the thesis

The PhD thesis on hand assembles investigations of the Lasserre hierarchy. Applications of the hierarchy arising in different contexts are considered, the important notion of convergence rates is studied, and we extend ideas on which the hierarchy is based to tightly approximate semialgebraic sets with quadratic separators. The linking element between these investigations is that we always approximate potentially complicated objects by simpler objects, like polynomials of small degree. These are either computed using the hierarchy (or ideas of it) or the approximations are used to develop a modification of the hierarchy.

The following paragraphs shall give an overview of the topics treated in the manuscript and how they are integrated into the overall subject of the thesis.

As a start, **Chapter 1** is dedicated to a short introduction to the Lasserre hierarchy. We briefly examine sums of squares and quadratic modules in order to understand Putinar's Positivstellensatz which, as already mentioned, is essential to the construction of the hierarchy. As a preparation for the dual formulations, we continue with some basics on measures and moments. Then we are ready to build the Lasserre hierarchies, meaning the standard hierarchy of lower bounds using Putinar's Positivstellensatz and the hierarchy of upper bounds restricting the feasible set to measures which allow sum-of-squares densities.

As a first application of the Lasserre hierarchy, in **Chapter 2** we search to approximate the polynomial abscissa, which is relevant in systems control, for example when studying the stability of linear systems. The abscissa of a polynomial is the maximum real part of its roots and a function of very low regularity in general. Therefore, it is rather difficult to approximate. We present an optimization problem whose minimizer is the abscissa function such that by applying the Lasserre hierarchy we obtain polynomial upper approximations to the abscissa map. In order to compute approximations from below, we modify the problem in such a manner that the Lasserre hierarchy gives lower approximations. These are results published in [HHL16].

As an application of the hierarchy arising in a different context, we consider an optimization problem occurring in statistics. For regression problems the main objective is to estimate parameters appearing in an experimental setting. In order to calculate these parameters up to a given precision, it is necessary to run the experiment multiple times with different input data. However, since it may be costly or otherwise intricate, the experimenter usually wants to minimize the required number of runs of the experiment. Therefore, we need a rule to decide which input data should be chosen from the design space in order to obtain a well posed problem which leads to relevant output data. These rules are called optimal designs and computing them is the subject to **Chapter 3**. Optimal designs on semialgebraic design spaces can be identified by weighted sums of Dirac measures supported on the design space. These can be approximated using the Lasserre hierarchy. The result is a sequence of moments, from which we then recover

the representing measure and hence the optimal design. The chapter is based on [dCGH⁺17].

In **Chapter 4** our goal is to approximate a semialgebraic set by quadratic separators, i.e., by polynomials of degree 2 whose zero sublevel set intersected with the semialgebraic set is empty. Approximating semialgebraic sets, or rather their convex hulls, by linear separators is an idea introduced by Lasserre in [Las09], which is strongly related to the Lasserre hierarchy since it is based on replacing non-negativity constraints by certificates of positivity provided by Putinar's Positivstellensatz. Here we extend this concept to separators of degree 2, which results in a nested sequence of outer approximations converging to the semialgebraic set. These outer approximations are described by semidefinite constraints, hence membership can be tested by a semidefinite program.

An important question when studying semidefinite programming hierarchies is their rate of convergence. In practice, the Lasserre hierarchy gives results which converge quickly enough to the global optimum. However, the theoretical bounds which are known so far are not in line with the numerical experiments. Accordingly, and as another investigation of the Lasserre hierarchy, in **Chapter 5** we develop a modification of the hierarchy of upper bounds for which we can prove a better rate of convergence. For this, we approximate the Dirac delta function by measures which allow densities with Schmüdgen-type representations, instead of approximating it by measures which allow sum-of-square densities as it is done for the Lasserre hierarchy of upper bounds. This chapter reports results from [dKHL17].

Overview

We summarize the contributions assembled in this thesis and the links between them. As already mentioned, we examine the Lasserre hierarchy by approximating potentially complicated objects either by using the hierarchy or to develop a modified hierarchy.

Chapter 1 Introduction to the Lasserre hierarchies of lower bounds and upper bounds.

Chapter 2 Upper and lower approximations of the abscissa function by applying the Lasserre hierarchy of lower bounds to a suitable problem.

Chapter 3 Computing approximate optimal designs on semialgebraic design spaces using the Lasserre hierarchy of lower bounds.

Chapter 4 Outer approximations of semialgebraic sets with quadratic separators using ideas on which the Lasserre hierarchy is based.

Chapter 5 Approximations of the Dirac delta function with the objective to modify the Lasserre hierarchy of upper bounds in such a way that a better rate of convergence can be shown.

Contributions not included in the manuscript

Apart from the just presented topics, the candidate also worked on the Minkowski problem, which might be solvable using the Lasserre hierarchy. The Minkowski problem consists of reconstructing a convex body when its curvature measure is known; see e.g. [Car04]. It can be written as such an optimization problem that when applying the Lasserre hierarchy, the zero sublevel set of the minimizer seems to be an approximation of the convex body that we seek to recover. Unfortunately, the certificates are only of numerical nature and formal proofs are lacking. For this reason it is not included in the thesis manuscript.

Computer configuration

All examples are modeled by YALMIP [Löf04] or GloptiPoly3 [HLL09] and solved by MOSEK 8 [MOS15] or SeDuMi 1.3 [Stu99] under the MATLAB R2014a environment. We ran the experiments on an HP EliteBook 840 G1 with 16-GB RAM memory and an Intel Core i5-4300U processor under a Windows 7 Professional 64-bit operating system.

The Lasserre hierarchy

Contents

1.1	Problem formulation	5
1.2	Sums of squares and quadratic modules	6
1.3	Measures and moments	8
1.3.1	Moments and the moment cone	8
1.3.2	Moment and localizing matrices	9
1.4	Hierarchies of upper and lower bounds	9
1.4.1	Hierarchy of lower bounds	10
1.4.2	Hierarchy of upper bounds	11

The objective of this introductory chapter is to explain the construction of the Lasserre hierarchy in a brief and compact manner. For a more elaborate and detailed investigation on the hierarchy the interested reader is referred to the books [Las10] and [Las15b], and to the survey [Lau09]. The chapter is mainly based on these references, and Section 1.4.2 also on [Las11].

The chapter is organized as follows. As a start, the precise problem is stated and we explain the first step of developing the hierarchy, meaning we rephrase the problem in terms of positive polynomials. In order to understand Putinar's Positivstellensatz (Theorem 1.1), which is essential for the subsequent considerations, we continue with some basics on sums of squares and quadratic modules (Section 1.2) and, as a preparation for the dual formulations, on measures and moments (Section 1.3). Then we are ready to build the Lasserre hierarchy (Section 1.4), or more precisely, the hierarchies, as we first explain how to construct the standard hierarchy of lower bounds and then the hierarchy of upper bounds.

As a summary, the respective optimization problems are grouped in Figure 1.1 and Figure 1.2 in order to illustrate the process of constructing the hierarchy.

1.1 Problem formulation

Denote by $\mathbb{R}[x]$ the infinite dimensional vector space of real polynomials in the variables $x = (x_1, \dots, x_n)$, and for $d \in \mathbb{N}$ define $\mathbb{R}[x]_d := \{p \in \mathbb{R}[x] : \deg p \leq d\}$, where $\deg p$ denotes the degree of p .

We consider the basic closed semialgebraic set

$$\mathbf{K} := \{\mathbf{x} \in \mathbb{R}^n : g_j(\mathbf{x}) \geq 0, j = 1, \dots, m\} \quad (1.1)$$

for given polynomials $g_j \in \mathbb{R}[x]$, $j = 1, \dots, m$, of which one is of the form $R - \sum_{i=1}^n x_i^2$ for a constant $R \in \mathbb{N}$. Note that the latter assumption implies that \mathbf{K} is compact.

Remark. The condition that one of the g_j is of the form $R - \sum_{i=1}^n x_i^2$ is not much stronger than compactness; see the discussion following Assumption 2.14 in [Las15b] for details.

We are interested in the global polynomial optimization problem

$$\varrho^{\min} = \inf_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}) \text{ s.t. } \mathbf{x} \in \mathbf{K}, \quad (1.2)$$

where $f \in \mathbb{R}[x]$ is a polynomial. Since the problem is non-convex in general and we are looking for a global minimum, this is a non-trivial problem. Solving it is the concern of the Lasserre hierarchy.

Denoting by $\mathcal{P}(\mathbf{K}) := \{p \in \mathbb{R}[x] : p(\mathbf{x}) \geq 0 \forall \mathbf{x} \in \mathbf{K}\}$ the cone of polynomials which are non-negative on \mathbf{K} , we can write (1.2) as the convex optimization problem

$$\varrho^{\min} = \sup_{\lambda \in \mathbb{R}} \lambda \text{ s.t. } f - \lambda \in \mathcal{P}(\mathbf{K}), \quad (1.3)$$

meaning we are searching the largest lower bound on f .

Although linear, the problem remains difficult, since the cone $\mathcal{P}(\mathbf{K})$ is very hard to characterize. Finding tractable subcones of $\mathcal{P}(\mathbf{K})$ is the subject of the next section.

1.2 Sums of squares and quadratic modules

As already mentioned, in this section we want to approximate the cone $\mathcal{P}(\mathbf{K})$ of polynomials non-negative on \mathbf{K} by tractable subcones. For this we introduce the notion of sums of squares (sometimes abbreviated by ‘‘SOS’’). We call a polynomial $p \in \mathbb{R}[x]$ a *sum of squares of polynomials*, if it can be written in the form

$$p(x) = \sum_{j=1}^s p_j(x)^2$$

for some polynomials $p_j \in \mathbb{R}[x]$, $j = 1, \dots, s$. We denote by $\Sigma[x] = \{p \in \mathbb{R}[x] : p \text{ is SOS}\}$ the cone of SOS polynomials and by $\Sigma[x]_{2d} := \Sigma[x] \cap \mathbb{R}[x]_{2d}$ the cone of SOS polynomials of degree at most $2d$.

Obviously, every SOS polynomial is non-negative on \mathbb{R}^n . However, the converse is only true for $n = 1$, for polynomials of degree 2, and for polynomials of degree 4 in two variables. This was proved by Hilbert in the 19th century.

Although SOS polynomials and non-negative polynomials are not the same in general, sums of squares can be used as a certificate of non-negativity. To find a certificate for non-negativity on \mathbf{K} including sums of squares, we introduce the notion of quadratic modules.

Given polynomials $g_1, \dots, g_m \in \mathbb{R}[x]$, we call

$$\mathcal{Q}(g_1, \dots, g_m) := \left\{ \sigma_0 + \sum_{j=1}^m \sigma_j g_j : \sigma_j \in \Sigma[x], j = 0, \dots, m \right\}$$

the *quadratic module generated by g_1, \dots, g_m* . A quadratic module is called *archimedean*, if

$$\forall p \in \mathbb{R}[x] \exists R \in \mathbb{N} \text{ such that } R \pm p \in \mathcal{Q}(g_1, \dots, g_m).$$

This is equivalent to assuming that there exists a constant $R \in \mathbb{N}$ such that $R - \sum_{i=1}^n x_i^2 \in \mathcal{Q}(g_1, \dots, g_m)$ by [Lau09, Theorem 3.17].

It is easy to verify that $\mathcal{Q}(g_1, \dots, g_m) \subseteq \mathcal{P}(\mathbf{K})$. The converse is not true in general, but we have the following important result.

Theorem 1.1 (Putinar’s Positivstellensatz [Put93]). *Let $\mathcal{Q}(g_1, \dots, g_m)$ be archimedean. If $p \in \mathbb{R}[x]$ is strictly positive on \mathbf{K} , then $p \in \mathcal{Q}(g_1, \dots, g_m)$.*

Remark. Putinar's Positivstellensatz is the reason why we assume in the definition of \mathbf{K} that one of the g_j is of the form $R - \sum_{i=1}^n x_i^2$. Because this implies that $\mathcal{Q}(g_1, \dots, g_m)$ is archimedean, we will therefore be able to apply this powerful theorem.

To summarize, we have the following inclusions:

$$\{p \in \mathbb{R}[x] : p(\mathbf{x}) > 0 \ \forall \mathbf{x} \in \mathbf{K}\} \subseteq \mathcal{Q}(g_1, \dots, g_m) \subseteq \mathcal{P}(\mathbf{K}). \quad (1.4)$$

Thus, membership of the quadratic module $\mathcal{Q}(g_1, \dots, g_m)$ is a certificate of positivity and we can strengthen problem (1.3) to

$$\varrho^{\text{sos}} = \sup_{\lambda \in \mathbb{R}} \lambda \text{ s.t. } f - \lambda \in \mathcal{Q}(g_1, \dots, g_m). \quad (1.5)$$

Since problem (1.3) is equivalent to $\varrho^{\min} = \sup_{\lambda \in \mathbb{R}} \lambda$ s.t. $f - \lambda > 0$ on \mathbf{K} , meaning we can replace $\mathcal{P}(\mathbf{K})$ by the set of polynomials which are strictly positive on \mathbf{K} , equation (1.4) implies $\varrho^{\min} = \varrho^{\text{sos}}$.

Problem (1.5) is still an infinite dimensional optimization problem as we do not impose a bound on the degree of the sum of squares in the quadratic module, but contrary to the cone $\mathcal{P}(\mathbf{K})$ the quadratic module $\mathcal{Q}(g_1, \dots, g_m)$ can be characterized by semidefinite constraints. To finally obtain a finite dimensional problem, in Section 1.4.1 we will consider the *truncated quadratic module*¹

$$\mathcal{Q}_{2d}(g_1, \dots, g_m) := \left\{ \sigma_0 + \sum_{j=1}^m \sigma_j g_j : \sigma_j \in \Sigma[x]_{2(d-d_j)}, j = 0, \dots, m \right\},$$

where we have set $d_0 = 0$ and $d_j := \lceil \deg(g_j)/2 \rceil$, $j = 1, \dots, m$. Now, we have a cone which is finite dimensional and tractable. To be more precise, $\mathcal{Q}_{2d}(g_1, \dots, g_m)$ is *semidefinite representable*, by which we mean that it is the projection of a spectrahedron, i.e., the projection of a set described by a linear matrix inequality constraint.

To see this, note that by, e.g., [Las15b, Proposition 2.1] a polynomial $p \in \mathbb{R}[x]_{2d}$ is SOS if and only if there exists a real symmetric and positive semidefinite matrix Q of size $\binom{n+d}{n} \times \binom{n+d}{n}$ such that $p(\mathbf{x}) = \mathbf{v}_d(\mathbf{x})^T Q \mathbf{v}_d(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n$, where $\mathbf{v}_d(x)$ denotes the column vector of monomials up to order d . See [Las15b, Section 2.4.2] for details on how to test membership in quadratic modules via semidefinite programs.

Remark. Although finite dimensional, the truncation $\mathcal{P}_d(\mathbf{K}) := \mathcal{P}(\mathbf{K}) \cap \mathbb{R}[x]_d$ is not suitable for our purposes due to the lack of a tractable representation. Note in particular, that $\mathcal{P}_d(\mathbf{K})$ is not semidefinite representable for $n = 2$, $d \geq 6$, and $n \geq 3$, $d \geq 4$, respectively, as recently proved by Scheiderer in [Sch16, Corollary 4.24].

The above truncations provide a convergent nested sequence of subcones of $\mathcal{P}(\mathbf{K})$,

$$\mathcal{Q}_{2d}(g_1, \dots, g_m) \subseteq \mathcal{Q}_{2(d+1)}(g_1, \dots, g_m) \subseteq \dots \subseteq \mathcal{Q}(g_1, \dots, g_m) \subseteq \mathcal{P}(\mathbf{K}). \quad (1.6)$$

Note, however, that for a polynomial $p \in \mathcal{Q}(g_1, \dots, g_m)$ the degree $2d$ for which $p \in \mathcal{Q}_{2d}(g_1, \dots, g_m)$ may be arbitrarily large regardless of the size of $\deg p$.

In the following section we consider the duals to the just considered objects.

¹Technically, one should also consider truncated quadratic modules of odd degrees, since $\deg g_j$ might be odd. In order to avoid confusion with other popular notation, like $\mathcal{Q}_d = \left\{ \sigma_0 + \sum_{j=1}^m \sigma_j g_j : \sigma_j \in \Sigma[x], \deg(\sigma_j g_j) \leq 2d, j = 0, \dots, m \right\}$, we ignore them here.

1.3 Measures and moments

Let $\mathcal{C}(\mathbf{K})$ denote the space of continuous functions from \mathbf{K} to \mathbb{R} . Its topological dual is isometrically isomorphic to the vector space $\mathcal{M}(\mathbf{K})$ of finite signed Borel measures supported on \mathbf{K} . We denote the associated positive cones by $\mathcal{C}_+(\mathbf{K})$ and $\mathcal{M}_+(\mathbf{K})$, meaning the cone of non-negative continuous functions defined on \mathbf{K} and the cone of non-negative Borel measures supported on \mathbf{K} , respectively.

The duality pairing of a function $f \in \mathcal{C}(\mathbf{K})$ and a measure $\mu \in \mathcal{M}(\mathbf{K})$ is defined by $\langle f, \mu \rangle := \int_{\mathbf{K}} f d\mu$.

The monomials $x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ with $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ form a basis of the vector space $\mathbb{R}[x]$. We use the multi-index notation $x^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ to denote these monomials. In the same way, for a given $d \in \mathbb{N}$ the vector space $\mathbb{R}[x]_d$ has dimension $s(d) := \binom{n+d}{n}$ with basis $(x^\alpha)_{|\alpha| \leq d}$, where $|\alpha| := \alpha_1 + \cdots + \alpha_n$. We write

$$\mathbf{v}_d(x) := \left(\underbrace{1}_{\text{degree 0}}, \underbrace{x_1, \dots, x_n}_{\text{degree 1}}, \underbrace{x_1^2, x_1 x_2, \dots, x_1 x_n, x_2^2, \dots, x_n^2}_{\text{degree 2}}, \dots, \underbrace{x_1^d, \dots, x_n^d}_{\text{degree d}} \right)^T$$

for the column vector of the monomials ordered according to their degree, and where monomials of the same degree are ordered with respect to the lexicographic ordering.

1.3.1 Moments and the moment cone

For a finite Borel measure $\mu \in \mathcal{M}_+(\mathbf{K})$ we call

$$y_\alpha = \int_{\mathbf{K}} x^\alpha d\mu, \quad \alpha \in \mathbb{N}^n,$$

the *moments* of μ . Accordingly, we call the sequence $\mathbf{y} = (y_\alpha)_{\alpha \in \mathbb{N}^n}$ the moment sequence of μ . Conversely, we say that a sequence $\mathbf{y} = (y_\alpha)_{\alpha \in \mathbb{N}^n} \subseteq \mathbb{R}$ has a *representing measure*, if there exists a measure μ such that \mathbf{y} is the moment sequence of μ . We denote by $\mathcal{M}(\mathbf{K})$ the convex cone of all sequences \mathbf{y} which have a representing measure supported on \mathbf{K} . We call it the *moment cone* of \mathbf{K} ,

$$\mathcal{M}(\mathbf{K}) := \{ \mathbf{y} \in \mathbb{R}^{\mathbb{N}^n} : \exists \mu \in \mathcal{M}_+(\mathbf{K}), y_\alpha = \int_{\mathbf{K}} x^\alpha d\mu \forall \alpha \in \mathbb{N}^n \}.$$

Given a sequence $\mathbf{y} = (y_\alpha)_{\alpha \in \mathbb{N}^n} \subseteq \mathbb{R}$ we define the so-called *Riesz linear functional* $L_{\mathbf{y}} : \mathbb{R}[x] \rightarrow \mathbb{R}$ which maps a polynomial $p = \sum_{\alpha \in \mathbb{N}^n} p_\alpha x^\alpha$ (with $p_\alpha \in \mathbb{R}$) to

$$L_{\mathbf{y}}(p) = \sum_{\alpha \in \mathbb{N}^n} p_\alpha y_\alpha.$$

The next result shows that the cones $\mathcal{P}(\mathbf{K})$ and $\mathcal{M}(\mathbf{K})$ are duals of each other with duality pairing $L_{\mathbf{y}}(p)$.

Theorem 1.2 (Riesz-Haviland). *A sequence $\mathbf{y} = (y_\alpha)_{\alpha \in \mathbb{N}^n}$ has a representing measure μ supported on \mathbf{K} if and only if $L_{\mathbf{y}}(p) \geq 0$ for all polynomials $p \in \mathcal{P}(\mathbf{K})$.*

Now, let us consider truncated sequences $\mathbf{y} = (y_\alpha)_{|\alpha| \leq d}$. We define the *truncated moment cone* of \mathbf{K} of moments up to order d by

$$\mathcal{M}_d(\mathbf{K}) := \{ \mathbf{y} \in \mathbb{R}^{s(d)} : \exists \mu \in \mathcal{M}_+(\mathbf{K}), y_\alpha = \int_{\mathbf{K}} x^\alpha d\mu \forall |\alpha| \leq d \}. \quad (1.7)$$

Since \mathbf{K} is compact, by [Lau09, Theorem 5.13], the cone $\mathcal{M}_d(\mathbf{K})$ is dual to the convex cone $\mathcal{P}_d(\mathbf{K})$ of polynomials $p \in \mathbb{R}[x]_d$ of degree at most d which are non-negative on \mathbf{K} .

1.3.2 Moment and localizing matrices

The *moment matrix of order d* of a truncated sequence $\mathbf{y} = (y_\alpha)_{|\alpha| \leq 2d}$ is the $s(d) \times s(d)$ -matrix $\mathbf{M}_d(\mathbf{y})$ with rows and columns indexed by $\alpha \in \mathbb{N}^n$, $|\alpha| \leq d$, and such that

$$\mathbf{M}_d(\mathbf{y})(\alpha, \beta) = L_{\mathbf{y}}(x^\alpha x^\beta) = y_{\alpha+\beta} \quad \forall |\alpha|, |\beta| \leq d.$$

It is symmetric and linear in \mathbf{y} , and if \mathbf{y} has a representing measure, then $\mathbf{M}_d(\mathbf{y}) \succcurlyeq 0$ for all $d \in \mathbb{N}$.

Similarly, we define the *localizing matrix of order d* with respect to a polynomial $g = \sum_{|\alpha| \leq r} g_\alpha x^\alpha \in \mathbb{R}[x]_r$ of degree r and a sequence $\mathbf{y} = (y_\alpha)_{|\alpha| \leq 2d+r}$ as the $s(d) \times s(d)$ -matrix $\mathbf{M}_d(g\mathbf{y})$ with rows and columns indexed by α and constructed by

$$\mathbf{M}_d(g\mathbf{y})(\alpha, \beta) = L_{\mathbf{y}}(g(x) x^\alpha x^\beta) = \sum_{\gamma \in \mathbb{N}^n} g_\gamma y_{\gamma+\alpha+\beta} \quad \forall |\alpha|, |\beta| \leq d.$$

If \mathbf{y} has a representing measure μ whose support is contained in the set $\{x \in \mathbb{R}^n : g(x) \geq 0\}$ for a $g \in \mathbb{R}[x]$, then $\mathbf{M}_{d-d_g}(g\mathbf{y}) \succcurlyeq 0$ for all $d \in \mathbb{N}$ with $d_g := \lceil \deg(g)/2 \rceil := \min\{z \in \mathbb{Z} : z \geq \deg(g)/2\}$.

Now we can define the dual to the quadratic module $\mathcal{Q}(g_1, \dots, g_m)$. For this we set $d_j := \lceil \deg(g_j)/2 \rceil$, $j = 1, \dots, m$, as half the degrees of the polynomials g_1, \dots, g_m from the definition of \mathbf{K} . Then the dual cone is given by

$$\mathcal{M}(g_1, \dots, g_m) := \{\mathbf{y} \in \mathbb{R}^{\mathbb{N}^n} : \mathbf{M}_d(\mathbf{y}) \succcurlyeq 0, \mathbf{M}_{d-d_j}(g_j \mathbf{y}) \succcurlyeq 0 \quad \forall d \in \mathbb{N}, j = 1, \dots, m\}.$$

The inclusion $\mathcal{M}(\mathbf{K}) \subseteq \mathcal{M}(g_1, \dots, g_m)$ is clear from the above considerations. The converse is given by the dual version of Putinar's Positivstellensatz.

Theorem 1.3 (Dual facet of Putinar's Positivstellensatz). *Let $\mathcal{Q}(g_1, \dots, g_m)$ be archimedean. Then $\mathcal{M}(\mathbf{K}) = \mathcal{M}(g_1, \dots, g_m)$.*

As for the quadratic modules, we also consider truncations of $\mathcal{M}(g_1, \dots, g_m)$, that is,

$$\mathcal{M}_{2d}(g_1, \dots, g_m) := \{\mathbf{y} \in \mathbb{R}^{s(2d)} : \mathbf{M}(\mathbf{y})_d \succcurlyeq 0, \mathbf{M}(g_j \mathbf{y})_{d-d_j} \succcurlyeq 0, j = 1, \dots, m\}.$$

These cones are dual to $\mathcal{Q}_{2d}(g_1, \dots, g_m)$, and accordingly give a nested sequence of supercones² of $\mathcal{M}(\mathbf{K})$,

$$\mathcal{M}(\mathbf{K}) = \mathcal{M}(g_1, \dots, g_m) \subseteq \dots \subseteq \mathcal{M}_{2(d+1)}(g_1, \dots, g_m) \subseteq \mathcal{M}_{2d}(g_1, \dots, g_m). \quad (1.8)$$

1.4 Hierarchies of upper and lower bounds

The objective of the Lasserre hierarchy is to numerically approximate problem (1.2). This is done by calculating lower, respectively upper, bounds of the optimal value.

Finding lower bounds for a global infimum is generally a difficult task. The method by Lasserre achieves it by relaxing the problem formulated in terms of measures (to which we refer

²Note the slight abuse of notation. Of course, as a matter of principle, the cone $\mathcal{M}(g_1, \dots, g_m) \subseteq \mathbb{R}^{\mathbb{N}^n}$ is not a subcone of $\mathcal{M}_{2d}(g_1, \dots, g_m) \subseteq \mathbb{R}^{s(2d)}$, but we can embed the former into the latter by considering the truncations of the moment sequences.

as the *primal problem*³), and accordingly strengthening the problem formulated in terms of sums of squares. We obtain a hierarchy of finite dimensional semidefinite programming (SDP) problems, i.e., semidefinite programs. These can be implemented and solved numerically leading to a sequence of lower bounds converging to the optimal value. Describing this procedure is the subject of the first subsection.

In the second subsection we concentrate on the hierarchy of upper bounds. It consists of the inverse procedure of the hierarchy of lower bounds, meaning we strengthen the primal and relax the dual problem.

1.4.1 Hierarchy of lower bounds

As already mentioned, we will relax the primal and strengthen the dual in order to obtain sequences of lower bounds converging to the infimum in problem (1.2). As a preparation, we summarize the cones just seen in the previous two sections and their relations to each other.

$$\begin{aligned} \mathcal{Q}(g_1, \dots, g_m)_{2d} &\subseteq \mathcal{Q}(g_1, \dots, g_m) \subseteq \mathcal{P}(\mathbf{K}) \subseteq \mathcal{C}_+(\mathbf{K}) \\ \mathcal{M}(g_1, \dots, g_m)_{2d} &\supseteq \mathcal{M}(g_1, \dots, g_m) = \mathcal{M}(\mathbf{K}) \quad \text{“}\supseteq\text{”} \quad \mathcal{M}_+(\mathbf{K}) \end{aligned} \quad (1.9)$$

The objects which are aligned are duals of each other. To make sense of the “ \supseteq ” one needs to identify the measures of $\mathcal{M}_+(\mathbf{K})$ with their moments.

We have seen at the beginning of the chapter that problem (1.2) can be written in terms of non-negative polynomials (1.3), and that this is equivalent to the SOS formulation (1.5). Furthermore, we observe that problem (1.2) can also be written in terms of measures:

$$\varrho^{\min} = \inf_{\mu \in \mathcal{M}_+(\mathbf{K})} \langle f, \mu \rangle \text{ s.t. } \langle 1, \mu \rangle = 1. \quad (1.10)$$

This problem is linear in μ , however, it is infinite dimensional. As an intuition, one can imagine to search for a Dirac measure at a global minimizer \mathbf{x}^* .

We reformulate (1.10) in terms of moments which yields the primal problem associated to (1.3):

$$\varrho^{\min} = \inf_{\mathbf{y} \in \mathcal{M}(\mathbf{K})} L_{\mathbf{y}}(f) \text{ s.t. } y_0 = 1. \quad (1.11)$$

Accordingly, we obtain an infinite dimensional linear programming (LP) problem. In order to be able to truncate it and obtain a finite dimensional problem as proposed by Lasserre [Las10], we relax it to the following infinite dimensional SDP problem whose dual (or rather predual) is (1.5):

$$\varrho^{\text{mom}} = \inf_{\mathbf{y} \in \mathcal{M}(g_1, \dots, g_m)} L_{\mathbf{y}}(f) \text{ s.t. } y_0 = 1, \quad (1.12)$$

(“mom” for “moments”). We call this a relaxation of (1.11), because generally $\mathcal{M}(g_1, \dots, g_m) \subseteq \mathcal{M}(\mathbf{K})$. But since we assume that the g_j are such that the quadratic module $\mathcal{Q}(g_1, \dots, g_m)$ is archimedean, we know by the dual version of Putinar’s Positivstellensatz, Theorem 1.3, that the cones are actually equal. Hence, (1.11) and (1.12) are equivalent, $\varrho^{\min} = \varrho^{\text{mom}}$.

Now, instead of taking the moment and localizing matrices of all orders, we only consider the moment and localizing matrices of a certain order d . This leads to a hierarchy of finite dimensional SDP problems given by

$$\varrho_d^{\text{mom}} = \inf_{\mathbf{y} \in \mathcal{M}_{2d}(g_1, \dots, g_m)} L_{\mathbf{y}}(f) \text{ s.t. } y_0 = 1 \quad (1.13)$$

³Consequently, we call the SOS problem the *dual problem*, although it would be more precise to call it the *predual*, since the cone $\mathcal{C}_+(\mathbf{K})$ is a strict subset of the topological dual of the cone $\mathcal{M}_+(\mathbf{K})$.

with $d \in \mathbb{N}$ and $d \geq \max\{\lceil \deg(f)/2 \rceil, \lceil \deg(g_j)/2 \rceil, j = 1, \dots, m\}$. Since the constraint $R - \sum_{i=1}^n x_i^2 \geq 0$ is present in the description of \mathbf{K} , the infimum is attained by [Lau09, Proposition 6.2]. The dual is the truncation of problem (1.5), that is,

$$\varrho_d^{\text{sos}} = \sup_{\lambda \in \mathbb{R}} \lambda \text{ s.t. } f - \lambda \in \mathcal{Q}_{2d}(g_1, \dots, g_m). \quad (1.14)$$

By weak duality we have $\varrho_d^{\text{sos}} \leq \varrho_d^{\text{mom}}$, but the duality gap is not necessarily zero. However, if \mathbf{K} has non-empty interior, i.e., there exists a full dimensional ball contained in \mathbf{K} , by [Lau09, Theorem 6.1] strong duality holds for all $d \geq \max\{\lceil \deg(f)/2 \rceil, \lceil \deg(g_j)/2 \rceil, j = 1, \dots, m\}$, meaning $\varrho_d^{\text{sos}} = \varrho_d^{\text{mom}}$. Moreover, (1.14) attains its supremum for d sufficiently large in this case.

Due to (1.6), problem (1.14) is a strengthening of problem (1.5), so we get the sequence

$$\varrho_d^{\text{sos}} \leq \varrho_{d+1}^{\text{sos}} \leq \varrho^{\text{sos}} = \varrho^{\text{min}}.$$

On the contrary, problem (1.13) is a relaxation of the infinite dimensional problem (1.12), since the set of feasible sequences \mathbf{y} of the former is a subset of the set of feasible sequences of the latter due to the consideration of moment and localizing matrices only up to order d , compare (1.8). Because we are minimizing, we therefore have

$$\varrho_d^{\text{mom}} \leq \varrho_{d+1}^{\text{mom}} \leq \varrho^{\text{mom}} = \varrho^{\text{min}}.$$

This is consistent with the fact that in optimization a relaxation of the primal problem induces a strengthening of its dual, which becomes clear when considering that a sequence of supercones on the primal side corresponds to a sequence of subcones on the dual side, as in (1.9) in our case. Thus, in both cases we get an ascending sequence of lower bounds on the infimum. Since we assume $\mathcal{Q}(g_1, \dots, g_m)$ to be archimedean, these sequences converge to the infimum by [Lau09, Theorem 6.8], i.e.,

$$\lim_{d \rightarrow \infty} \varrho_d^{\text{mom}} = \varrho^{\text{min}} = \lim_{d \rightarrow \infty} \varrho_d^{\text{sos}}. \quad (1.15)$$

One may wonder under which conditions the moment relaxation is exact for a finite d , meaning $\varrho_d^{\text{mom}} = \varrho^{\text{min}}$. If the optimal value ϱ_d^{mom} of problem (1.13) is attained at some optimal solution \mathbf{y}^* , then for $d_{\mathbf{K}} := \max\{\lceil \deg(g_j)/2 \rceil, j = 1, \dots, m\}$ the condition

$$\text{rank } \mathbf{M}_d(\mathbf{y}^*) = \text{rank } \mathbf{M}_{d-d_{\mathbf{K}}}(\mathbf{y}^*) \quad (1.16)$$

is sufficient for $\varrho_d^{\text{mom}} = \varrho^{\text{min}}$ by [Las15b, Theorem 6.6].

In the same way, finite convergence takes place, when at every global minimizer the gradients are linearly independent and strict complementarity and second-order sufficiency hold. This is a result due to Nie, see [Las15b, Theorem 6.5] for details.

With the problems (1.13) and (1.14) we have finally found semidefinite programs whose optimal values converge from below to the optimal value of our original problem (1.2). Therefore, we are now able to compute lower bounds numerically.

See Figure 1.1 at the end of the chapter for an overview of the Lasserre hierarchy of lower bounds.

1.4.2 Hierarchy of upper bounds

In this section we consider the counterpart of the hierarchy of lower bounds. Instead of relaxing the primal and strengthening the dual problem, we strengthen the primal and relax the dual which leads to a sequence of upper bounds converging to the optimal value ϱ^{min} .

For simplicity, in this section we additionally assume $\mathbf{K} \subseteq [-1, 1]^n$. We need the following preliminary result:

Theorem 1.4. [Las11, Theorem 3.2] *Let $\mathbf{K} \subseteq [-1, 1]^n$ be compact and let $\nu \in \mathcal{M}_+(\mathbf{K})$ be arbitrary and fixed and such that $\text{supp}(\nu) = \mathbf{K}$. Let $\mathbf{y}^\nu = (y_\alpha^\nu)_{\alpha \in \mathbb{N}^n}$ be its moment vector. Then $p \in \mathbb{R}[x]$ is non-negative on \mathbf{K} if and only if*

$$\int_{\mathbf{K}} p g^2 d\nu \geq 0 \quad \forall g \in \mathbb{R}[x],$$

or, equivalently, if and only if $\mathbf{M}_d(p \mathbf{y}^\nu) \succeq 0$ for all $d \in \mathbb{N}$.

Now, let us start by building the hierarchy on the SOS side, i.e., we relax (1.3) to a finite dimensional SDP problem. For this, fix a finite Borel measure ν with $\text{supp}(\nu) = \mathbf{K}$. We call $\mathbf{y}^\nu = (y_\alpha^\nu)_{\alpha \in \mathbb{N}^n}$ the moment vector of ν and consider the problem

$$\vartheta_d^\lambda = \sup_{\lambda \in \mathbb{R}} \lambda \text{ s.t. } \mathbf{M}_d(f \mathbf{y}^\nu) - \lambda \mathbf{M}_d(\mathbf{y}^\nu) \succeq 0. \quad (1.17)$$

This is a relaxation of problem (1.3), since by Theorem 1.4 its constraint $\mathbf{M}_d((f - \lambda) \mathbf{y}^\nu) \succeq 0$ is only a necessary condition for $f - \lambda$ to be non-negative on \mathbf{K} . To obtain a sufficient constraint, we would need to consider the localizing matrices of all orders $d \in \mathbb{N}$. Hence, the set of feasible solutions of (1.17) is larger than that of (1.3), which means that the supremum over the former is greater than the supremum over the latter. Therefore, we get a sequence of upper bounds for the optimal value, $\vartheta_d^\lambda \geq \vartheta_{d+1}^\lambda \geq \varrho^{\min}$.

By [Las11, Theorem 4.1], problem (1.17) has an optimal solution and the sequence $(\vartheta_d^\lambda)_{d \in \mathbb{N}}$ is monotone non-increasing and converges to ϱ^{\min} from above, $\vartheta_d^\lambda \searrow \varrho^{\min}$.

The dual to problem (1.17) is given by

$$\vartheta_d^\sigma = \inf_{\sigma \in \Sigma[x]_{2d}} \int_{\mathbf{K}} f \sigma d\nu \text{ s.t. } \int_{\mathbf{K}} \sigma d\nu = 1. \quad (1.18)$$

Here, we optimize over all probability measures which have a density $\sigma \in \Sigma[x]_{2d}$ with respect to ν . Since these measures form a subset of $\mathcal{M}_+(\mathbf{K})$, problem (1.18) is a strengthening of (1.10). Thus, consistent with the SOS side, we obtain a sequence of upper bounds, $\vartheta_d^\sigma \geq \vartheta_{d+1}^\sigma \geq \varrho^{\min}$.

As an intuition, one can imagine to approximate the Dirac measure at a global minimizer \mathbf{x}^* (which would be the solution to the infinite dimensional problem (1.10)) by measures which admit sum-of-squares densities of increasing degree.

In general, it is not clear, whether $(\vartheta_d^\sigma)_{d \in \mathbb{N}}$ converges to ϱ^{\min} , but in case \mathbf{K} has a non-empty interior, we know by [Las11, Theorem 4.2] that the duality gap between (1.17) and (1.18) is zero and that (1.18) has an optimal solution. However, generally the convergence is only asymptotic and not finite.

Remark. In order to implement problem (1.18), respectively (1.17), we would need the moments $y_\alpha^\nu := \int_{\mathbf{K}} x^\alpha d\nu$ explicitly. So we can only solve it numerically, when \mathbf{K} and ν are such that the moments can be computed for all $\alpha \in \mathbb{N}^n$.

See Figure 1.2 at the end of the chapter for an overview of the Lasserre hierarchy of upper bounds.

$$\begin{aligned} \varrho^{\min} &= \inf_{\mathbf{x}} f(\mathbf{x}) \text{ s.t. } \mathbf{x} \in \mathbf{K}, \\ &\text{with} \\ \mathbf{K} &:= \{\mathbf{x} \in \mathbb{R}^n : g_j(\mathbf{x}) \geq 0, j = 1, \dots, m\}, \\ f, g_j &\in \mathbb{R}[x], 1 \leq j \leq m, \exists j : g_j(x) = R - \sum_{i=1}^n x_i^2 \text{ for some } R \in \mathbb{N}. \end{aligned}$$

PRIMAL FORMULATION

DUAL FORMULATION

$$\begin{array}{ccc} \varrho^{\min} = \inf_{\mu} \int_{\mathbf{K}} f(x) \mu(dx) & & \varrho^{\min} = \sup_{\lambda \in \mathbb{R}} \lambda \\ \text{s.t. } \mu(\mathbf{K}) = 1, \mu \in \mathcal{M}_+(\mathbf{K}) & & \text{s.t. } f - \lambda \in \mathcal{C}_+(\mathbf{K}) \\ \downarrow \text{moments} & & \downarrow \text{polynomials} \\ \varrho^{\min} = \inf_{\mathbf{y} \in \mathbb{R}^n} \sum_{\alpha \in \mathbb{N}^n} f_{\alpha} y_{\alpha} & & \varrho^{\min} = \sup_{\lambda \in \mathbb{R}} \lambda \\ \text{s.t. } y_0 = 1, \mathbf{y} \text{ has a representing} & & \text{s.t. } f - \lambda \geq 0 \text{ on } \mathbf{K} \\ \text{measure supported on } \mathbf{K} & & \downarrow \text{Putinar's} \\ & & \text{Positivstellensatz} \\ \downarrow \text{Dual facet of Putinar's} & & \downarrow \\ \text{Positivstellensatz} & & \varrho^{\text{SOS}} = \sup_{\lambda, \sigma_0, \dots, \sigma_m} \lambda \\ \varrho^{\text{mom}} = \inf_{\mathbf{y} \in \mathbb{R}^n} \sum_{\alpha \in \mathbb{N}^n} f_{\alpha} y_{\alpha} & & \text{s.t. } f - \lambda = \sigma_0 + \sum_{j=1}^m \sigma_j g_j, \\ \text{s.t. } y_0 = 1, \mathbf{M}(\mathbf{y}) \succeq 0, & & \lambda \in \mathbb{R}, \sigma_j \in \Sigma[x], 0 \leq j \leq m \\ \mathbf{M}(g_j \mathbf{y}) \succeq 0, 1 \leq j \leq m & & \downarrow \text{truncation} \\ \downarrow \text{truncation}^4 & & \varrho_d^{\text{SOS}} = \sup_{\lambda, \sigma_0, \dots, \sigma_m} \lambda \\ \varrho_d^{\text{mom}} = \inf_{\mathbf{y} \in \mathbb{R}^{s(2d)}} \sum_{\alpha \in \mathbb{N}^n} f_{\alpha} y_{\alpha} & & \text{s.t. } f - \lambda = \sigma_0 + \sum_{j=1}^m \sigma_j g_j, \lambda \in \mathbb{R} \\ \text{s.t. } y_0 = 1, \mathbf{M}_d(\mathbf{y}) \succeq 0, & & \sigma_j \in \Sigma[x]_{2(d-d_j)}, 0 \leq j \leq m \\ \mathbf{M}_{d-d_j}(g_j \mathbf{y}) \succeq 0, 1 \leq j \leq m & & \\ \\ \varrho_d^{\text{mom}} \leq \varrho_{d+1}^{\text{mom}} \leq \varrho^{\text{mom}} = \varrho^{\min} & & \varrho_d^{\text{SOS}} \leq \varrho_{d+1}^{\text{SOS}} \leq \varrho^{\text{SOS}} = \varrho^{\min} \end{array}$$

Figure 1.1 – Summary of the Lasserre hierarchy of lower bounds

Compare with equations (1.10) - (1.13) for the moment side and with equations (1.3), (1.5) and (1.14) for the SOS side.

⁴By $\mathbf{M}(\mathbf{y})$, and $\mathbf{M}(g_j \mathbf{y})$ respectively, we mean the infinite moment and localizing matrices, respectively.

$$\begin{aligned}
\varrho^{\min} &= \inf_{\mathbf{x}} f(\mathbf{x}) \text{ s.t. } \mathbf{x} \in \mathbf{K}, \\
&\text{with} \\
\mathbf{K} &:= \{\mathbf{x} \in \mathbb{R}^n : g_j(\mathbf{x}) \geq 0, j = 1, \dots, m\} \subseteq [-1, 1]^n, \\
f, g_j &\in \mathbb{R}[x], 1 \leq j \leq m, \exists j : g_j(x) = R - \sum_{i=1}^n x_i^2 \text{ for some } R \in \mathbb{N}.
\end{aligned}$$

PRIMAL FORMULATION

DUAL FORMULATION

$$\begin{aligned}
\varrho^{\min} &= \inf_{\mu} \int_{\mathbf{K}} f(x) \mu(dx) \\
&\text{s.t. } \mu(\mathbf{K}) = 1, \mu \in \mathcal{M}_+(\mathbf{K})
\end{aligned}$$

measures with
sum-of-squares-densities
with respect to a fixed
measure $\nu \in \mathcal{M}_+(\mathbf{K})$ with
 $\text{supp}(\nu) = \mathbf{K}$

$$\begin{aligned}
\vartheta^\sigma &= \inf_{\sigma \in \Sigma[x]} \int_{\mathbf{K}} f \sigma d\nu \\
&\text{s.t. } \int_{\mathbf{K}} \sigma d\nu = 1
\end{aligned}$$

truncation

$$\begin{aligned}
\vartheta_d^\sigma &= \inf_{\sigma \in \Sigma[x]_{2d}} \int_{\mathbf{K}} f \sigma d\nu \\
&\text{s.t. } \int_{\mathbf{K}} \sigma d\nu = 1
\end{aligned}$$

$$\vartheta_d^\sigma \geq \vartheta_{d+1}^\sigma \geq \vartheta^\sigma \geq \varrho^{\min}$$

$$\begin{aligned}
\varrho^{\min} &= \sup_{\lambda \in \mathbb{R}} \lambda \\
&\text{s.t. } f - \lambda \in \mathcal{C}_+(\mathbf{K})
\end{aligned}$$

polynomials

$$\begin{aligned}
\varrho^{\min} &= \sup_{\lambda \in \mathbb{R}} \lambda \\
&\text{s.t. } f - \lambda \geq 0 \text{ on } \mathbf{K}
\end{aligned}$$

Theorem 1.4

$$\begin{aligned}
\vartheta^\lambda &= \sup_{\lambda \in \mathbb{R}} \lambda \\
&\text{s.t. } \mathbf{M}(f \mathbf{y}^\nu) - \lambda \mathbf{M}(\mathbf{y}^\nu) \succeq 0
\end{aligned}$$

truncation⁵

$$\begin{aligned}
\vartheta_d^\lambda &= \sup_{\lambda \in \mathbb{R}} \lambda \\
&\text{s.t. } \mathbf{M}_d(f \mathbf{y}^\nu) - \lambda \mathbf{M}_d(\mathbf{y}^\nu) \succeq 0
\end{aligned}$$

$$\vartheta_d^\lambda \geq \vartheta_{d+1}^\lambda \geq \vartheta^\lambda = \varrho^{\min}$$

Figure 1.2 – Summary of the Lasserre hierarchy of upper bounds

Compare with equations (1.10) and (1.18) for the moment side and with equations (1.3) and (1.17) for the SOS side.

⁵Again, $\mathbf{M}(\mathbf{y})$ and $\mathbf{M}(f\mathbf{y})$ stand for the infinite moment and localizing matrices, respectively.

Semidefinite approximations of the polynomial abscissa

Contents

2.1	The abscissa function	16
2.2	Upper abscissa approximation	17
2.2.1	Primal and dual formulation	17
2.2.2	Lasserre hierarchy	19
2.2.3	Examples	21
2.3	Lower abscissa approximation	24
2.3.1	Lower approximation via elementary symmetric functions	26
2.3.2	Lower approximation via Gauß-Lucas	31
2.4	Concluding remarks	37

In this chapter, as a first application of the Lasserre hierarchy of lower bounds, we want to approximate a function of very low regularity, namely the polynomial abscissa.

Given a univariate polynomial p whose coefficients depend polynomially on parameters, its abscissa maps the parameter vector to the maximum real part of the roots of p for this parameter. When studying linear differential equations, the abscissa of the characteristic polynomial of the equation is used as a measure of the decay or growth rate of the solution. In linear systems control, the abscissa function is typically parametrized by a small number of real parameters (the controller coefficients), and it should be minimized so as to ensure a sufficiently fast decay rate of closed-loop trajectories.

As a function of the polynomial coefficients (expressed in some basis), the abscissa is a Hölder continuous function (with exponent equal to the reciprocal of the polynomial degree), but it is not locally Lipschitz in general. As a consequence of this low regularity, numerical optimization of the polynomial abscissa is typically a challenge.

Here, we use the Lasserre hierarchy of lower bounds to approximate the abscissa by polynomials of fixed degree. Hence, we get approximations of controlled complexity. When their degree tends to infinity, the polynomial approximations converge in L^1 norm to the abscissa, either from above or from below. The chapter reports results from [HHLP16].

The outline of the chapter is as follows. First, we give a short overview of research concerning the abscissa, and motivate our investigations. Then, after introducing in Section 2.1 the abscissa function and some relevant notation, we address in Section 2.2 the problem of finding an upper approximation of the abscissa. In Section 2.3, we address the more difficult problem of approximating the abscissa from below, first by using elementary symmetric functions, and second by using the Gauß-Lucas theorem, inspired by [BLO04]. Explicit numerical examples illustrate our findings throughout the text. Finally, we complete the chapter with a short conclusion and some comments on uniform approximations of the abscissa.

State of the art

For a recent survey on the abscissa function and its applications in systems control, see [Cro10]. A detailed variational analysis of the abscissa was first carried out in [BO01] and [BLO04]. These ideas were exploited in a systems control setup in [BHLO06], using randomized techniques of non-convex non-smooth local optimization, however without rigorous convergence guarantees.

In the space of controller parameters, the zero sublevel set of the abscissa function of the characteristic polynomial of a linear system is the so-called stability region, and it is typically non-convex and non-smooth; see [HPAŠ03], where this set is approximated with simpler sets such as balls or ellipsoids. In [HL12], ellipsoidal approximations of the stability region were generalized to polynomial sublevel set approximations, obtained by replacing negativity of the abscissa function with positive definiteness of the Hermite matrix of the characteristic polynomial.

2.1 The abscissa function

Let $\mathbf{Q} \subseteq \mathbb{R}^n$ be a compact semialgebraic set on which a Borel measure with support \mathbf{Q} can be defined and whose moments are easy to compute. For simplicity, we choose $\mathbf{Q} = [-1, 1]^n = \{q \in \mathbb{R}^n : 1 - q_1^2 \geq 0, \dots, 1 - q_n^2 \geq 0\}$.

Consider the monic non-constant polynomial $p \in \mathbb{R}[s]$ defined by

$$p : s \mapsto p(q, s) := \sum_{k=0}^{\ell} p_k(q) s^k$$

with $s \in \mathbb{C}$ complex, $q = (q_1, \dots, q_n) \in \mathbf{Q}$, and given polynomials $p_k \in \mathbb{R}[q]$ for $k = 0, 1, \dots, \ell$ with $p_\ell(q) \equiv 1$ and $\ell > 0$. Hence, we have a polynomial whose coefficients depend polynomially on the parameter q .

Denote by $s_r(q)$, $r = 1, \dots, \ell$, the roots of $p(q, \cdot)$ and by $a : \mathbf{Q} \rightarrow \mathbb{R}$ (or a_p if it is necessary to clarify the dependence on the polynomial) the abscissa map of p , i.e., the function which maps $q \in \mathbf{Q}$ to the maximal real part of the roots of $p(q, \cdot)$,

$$a : q \mapsto a(q) := \max_{r=1, \dots, \ell} \Re(s_r(q)).$$

Equivalently, with $i = \sqrt{-1}$ and $s = u + iz$ write

$$p(q, s) = p_{\Re}(q, u, z) + i p_{\Im}(q, u, z)$$

for two real polynomials $p_{\Re}, p_{\Im} \in \mathbb{R}[q, u, z]$ of degree ℓ in the variable u (resp. z). Then

$$a : q \mapsto a(q) = \max\{u \in \mathbb{R} : \exists z \in \mathbb{R} : p_{\Re}(q, u, z) = p_{\Im}(q, u, z) = 0\}.$$

We observe that the function $a : \mathbf{Q} \rightarrow \mathbb{R}$ is semialgebraic and we define the basic closed semialgebraic set

$$\mathbf{Z} := \{(q, u, z) \in \mathbb{R}^n \times \mathbb{R}^2 : q \in \mathbf{Q}, p_{\Re}(q, u, z) = p_{\Im}(q, u, z) = 0\}.$$

Remark. Set \mathbf{Z} is compact, since \mathbf{Q} is compact and p is monic in s .

Now we can write the abscissa map as

$$a : q \mapsto a(q) = \max\{u \in \mathbb{R} : \exists z \in \mathbb{R} : (q, u, z) \in \mathbf{Z}\}.$$

Since p is monic, its abscissa a is continuous, though in general not Lipschitz continuous. For example, for $n = 1$ and $p(q, s) = s^6 + q$ the map $a(q)$ is only Hölder continuous with exponent $\frac{1}{6}$ near the origin. To be precise, a is always Hölder continuous by the Łojasiewicz inequality [BCR98], since \mathbf{Q} is compact.

Stability regions for linear systems

For continuous-time dynamical systems described by linear differential equations, stability analysis amounts to studying the location of the roots of the characteristic polynomial obtained by applying the Laplace transform. A polynomial is then called stable, if all its roots lie in the open left part of the complex plane, i.e., if its abscissa is negative. In systems control, the characteristic polynomial depends on parameters, which are typically controller coefficients that must be chosen so that the polynomial is stable. For a polynomial with parameterized coefficients, as we consider in the chapter on hand, the stability region is then the set of parameters for which the abscissa is negative, that is, the zero sublevel set of the abscissa, in our notation $\{q \in \mathbf{Q} : a(q) < 0\}$.

2.2 Upper abscissa approximation

Our principal goal is to numerically approximate the abscissa function from above. Thus, we formulate a linear programming problem whose solution is the abscissa and apply the procedure described in Section 1.4.1 in order to obtain a hierarchy of semidefinite programs whose solution converges to the abscissa from above.

2.2.1 Primal and dual formulation

Given a polynomial p defined as above, any function v admissible for the following infinite-dimensional linear programming problem gives an upper approximation of the abscissa function a_p on \mathbf{Q} :

$$\begin{aligned} \varrho &= \inf_{v \in \mathcal{C}(\mathbf{Q})} \int_{\mathbf{Q}} v(q) dq \\ \text{s.t. } & v(q) - u \geq 0 \text{ for all } (q, u, z) \in \mathbf{Z}. \end{aligned} \tag{2.1}$$

We can consider problem (2.1) as a generalization of problem (1.3), since although the optimization variable $v \in \mathcal{C}(\mathbf{Q})$ is a continuous function, the problem is linear in v and the constraints are with respect to the cone $\mathcal{C}_+(\mathbf{Z})$, where \mathbf{Z} is a compact basic closed semialgebraic set. Note in passing that, in contrast to (1.3), here we are minimizing instead of maximizing, so the Lasserre hierarchy of lower bounds described in Section 1.4.1 will give upper bounds instead of lower bounds.

Importantly, there is a substantial difference from the situation encountered in problem (1.3). Here, we are not interested in the value ϱ , but in the function $v: q \mapsto v(q)$ solving LP problem (2.1). It is v , which is an approximation to the abscissa function, and since the constraint in (2.1) implies $v(q) \geq a(q)$ for all $q \in \mathbf{Q}$, it is an upper approximation. On the contrary, the value ϱ just bounds $\int_{\mathbf{Q}} v(q) dq$ from above.

Remark. Since the continuous functions defined on the compact set \mathbf{Q} can be uniformly approximated by polynomials by the Stone-Weierstraß theorem [Zor04, §16.4.3], we can replace $\mathcal{C}(\mathbf{Q})$ in problem (2.1) with the ring of polynomials $\mathbb{R}[q]$.

The primal LP problem corresponding to problem (2.1) can be constructed as described in [Bar02, Chapter IV] and reads

$$\begin{aligned} \varrho^* &= \sup_{\mu \in \mathcal{M}_+(\mathbf{Z})} \int_{\mathbf{Z}} u \, d\mu(q, u, z) \\ \text{s.t. } & \int_{\mathbf{Z}} q^\alpha \, d\mu = \int_{\mathbf{Q}} q^\alpha \, dq \quad \text{for all } \alpha \in \mathbb{N}^n. \end{aligned} \quad (2.2)$$

This problem can be seen as a generalization of (1.10) in the sense that we have infinitely many moment constraints in contrast to only one.

Remark 2.2.1. The constraint $\int_{\mathbf{Z}} q^\alpha \, d\mu = \int_{\mathbf{Q}} q^\alpha \, dq$ for all $\alpha \in \mathbb{N}^n$ implies that the marginal of μ on \mathbf{Q} is the Lebesgue measure on \mathbf{Q} , i.e., for every $g \in \mathcal{C}(\mathbf{Q})$ it holds that

$$\int_{\mathbf{Z}} g(q) \, d\mu(q, u, z) = \int_{\mathbf{Q}} g(q) \, dq.$$

In particular this implies that the mass $\|\mu\| = \text{vol } \mathbf{Q}$, where $\text{vol}(\cdot)$ denotes the volume or Lebesgue measure.

Lemma 2.1. *The supremum in LP problem (2.2) is attained, and there is no duality gap between LP problem (2.1) and LP problem (2.2), i.e., $\varrho = \varrho^*$.*

Proof. The set of feasible solutions for the dual LP problem (2.2) is a bounded subset of $\mathcal{M}_+(\mathbf{Z})$ with \mathbf{Z} compact and therefore it is weak-star compact. Since the objective function is linear, its supremum on this weak-star compact set is attained. For elementary background on weak-star topology, see, e.g., [Bar02, Chapter IV].

To prove that there is no duality gap, we apply [Bar02, Theorem IV.7.2]. For this purpose we introduce the notation used in [Bar02] in this context. There, the primal and the dual are written in the following canonical form:

$$\begin{aligned} \varrho^* &= \sup_{x \in E_1} \langle x, c \rangle_1 & \varrho &= \inf_{y \in F_2} \langle b, y \rangle_2 \\ \text{s.t. } & Ax = b, \, x \in E_1^+ & \text{s.t. } & A^*y - c \in F_1^+ \end{aligned}$$

So we set $E_1 := \mathcal{M}(\mathbf{Z})$ with its cone $E_1^+ := \mathcal{M}_+(\mathbf{Z})$. Then their (pre-)duals are $F_1 := \mathcal{C}(\mathbf{Z})$ and $F_1^+ := \mathcal{C}_+(\mathbf{Z})$, respectively. Similarly, we define $E_2 := \mathcal{M}(\mathbf{Q})$ and $F_2 := \mathcal{C}(\mathbf{Q})$.

Setting $x := \mu \in E_1$, $c := u \in F_1$, $b \in E_2$ the Lebesgue measure on \mathbf{Q} , and $y := v \in F_2$, the linear operator $A: E_1 \rightarrow E_2$ is given by $x \mapsto \pi_{\mathbf{Q}}x$, where $\pi_{\mathbf{Q}}$ denotes the projection onto \mathbf{Q} , i.e., $Ax(B) = x(B \times \mathbb{R}^2)$ for all $B \in \mathcal{B}(\mathbf{Q})$, the Borel sigma algebra on \mathbf{Q} .

According to [Bar02, Theorem IV.7.2] the duality gap is zero if the cone $\{(Ax, \langle x, c \rangle_1) : x \in E_1^+\}$ is closed in $E_2 \times \mathbb{R}$. This holds in our setup since $x \mapsto Ax$ and $x \mapsto \langle x, c \rangle_1$ are continuous linear maps and $E_1^+ = \mathcal{M}_+(\mathbf{Z})$ is weak-star closed due to the compactness of \mathbf{Z} . So if for some $\hat{x} \in E_2$, $Ax_n \rightarrow \hat{x}$ as $n \rightarrow \infty$, then from the definition of the mapping A and as $(x_n) \subset E_1^+$, one has $\|x_n\| \rightarrow \|\hat{x}\|$ as $n \rightarrow \infty$ (see Remark 2.2.1). Therefore the sequence $(x_n) \subset E_1^+$ is bounded and by Banach-Alaoglu's theorem [Ash00, Bar02], it contains a subsequence $(x_{n_k}) \subset E_1^+$ that converges to some $x \in E_1^+$ for the weak-star topology. By continuity of the mappings A and c , the result follows. \square

Remark 2.2.2. The infimum in LP problem (2.1) is not necessarily attained, since the set of feasible solutions is not compact. The reason for this is that we minimize over the L^1 norm of v , which implies that the limit of an optimizing sequence $(v_l)_{l \in \mathbb{N}}$ does not necessarily need to be

continuous, because on sets with Lebesgue measure zero it does not need to match the abscissa. Nor is the infimum attained when we replace $\mathcal{C}(\mathbf{Q})$ with $\mathbb{R}[q]$, since a is non-Lipschitz, so in particular not a polynomial.

However, the infimum is attained if we replace $\mathcal{C}(\mathbf{Q})$ with $\mathbb{R}[q]_d$ for d finite. Then, with $M := \min_{q \in \mathbf{Q}} a(q) > -\infty$ and $\tilde{v}(q) := v(q) - M$ we can rewrite LP problem (2.1) as the equivalent problem

$$\inf_{\tilde{v} \in \mathbb{R}[q]_d} \int_{\mathbf{Q}} \tilde{v}(q) dq \quad \text{s.t.} \quad \tilde{v}(q) + M - u \geq 0 \text{ on } \mathbf{Z}.$$

Now, any feasible \tilde{v} is non-negative on \mathbf{Q} , so $\int_{\mathbf{Q}} \tilde{v}(q) dq = \|\tilde{v}\|_{L^1} \geq 0$ is a norm because \mathbf{Q} has non-empty interior by assumption, and for every $R \in \mathbb{R}$ the set $\{\tilde{v} \in \mathbb{R}[q]_d : R \geq \int_{\mathbf{Q}} \tilde{v}(q) dq \text{ and } \tilde{v}(q) + M - u \geq 0 \text{ on } \mathbf{Z}\}$ is closed and bounded in the strong topology, thus compact. Besides, due to the continuity of a , there always exists an $R < \infty$ such that the mentioned set is not empty, hence the infimum is attained.

2.2.2 Lasserre hierarchy

Now, we build the SDP hierarchy for problem (2.1) as described in Section 1.4.1 and examine its convergence properties. For this, we exchange non-negativity on \mathbf{Z} for membership of the truncated quadratic module generated by the polynomials defining \mathbf{Z} , as we did for the hierarchy (1.14). To be more precise, we exchange $\mathcal{C}_+(\mathbf{Z})$ (respectively $\mathcal{P}(\mathbf{Z})$) for the quadratic module $\mathcal{Q}_{2d}(1 - q_1^2, \dots, 1 - q_n^2, \pm p_{\Re}, \pm p_{\Im})$ for $d \in \mathbb{N}$.

So, let $d_0 \in \mathbb{N}$ be sufficiently large, more precisely let d_0 be greater or equal than half the degree of p , $d_0 \geq \lceil \ell/2 \rceil$. The hierarchy of convex semidefinite programs for LP problem (2.1) indexed by the parameter $d \in \mathbb{N}$, $d \geq d_0$, reads

$$\begin{aligned} \varrho_d = \inf_{v_d, \sigma_0, \sigma_j, \tau_{\Re}, \tau_{\Im}} \int_{\mathbf{Q}} v_d(q) dq \\ \text{s.t. } v_d(q) - u = \sigma_0(q, u, z) + \sum_{j=1}^n \sigma_j(q, u, z)(1 - q_j^2) \\ + \tau_{\Re}(q, u, z)p_{\Re}(q, u, z) + \tau_{\Im}(q, u, z)p_{\Im}(q, u, z) \end{aligned} \quad (2.3)$$

for all $(q, u, z) \in \mathbb{R}^n \times \mathbb{R}^2$ and with $v_d \in \mathbb{R}[q]_{2d}$, $\sigma_0 \in \Sigma[q, u, z]_{2d}$, $\sigma_j \in \Sigma[q, u, z]_{2d-2}$ for $j = 1, \dots, n$, and $\tau_{\Re}, \tau_{\Im} \in \mathbb{R}[q, u, z]_{2d-\ell}$.

We have seen in the previous chapter that this is a strengthening of (2.1), so since we are minimizing (in contrast to Section 1.4.1), we get $\varrho_d \geq \varrho$.

Furthermore, the quadratic module generated by the polynomials $1 - q_1^2, \dots, 1 - q_n^2, \pm p_{\Re}, \pm p_{\Im}$ is archimedean by [Lau09, Lemma 3.17], since it contains the polynomial $f(q, u, z) := \sum_{j=1}^n (1 - q_j^2) - p_{\Re}^2 - p_{\Im}^2$ and the set $\{(q, u, z) \in \mathbb{R}^n \times \mathbb{R}^2 : f(q, u, z) \geq 0\}$ is compact. By the same argument as for (1.15), this implies that the hierarchy converges, i.e., $\lim_{d \rightarrow \infty} \varrho_d = \varrho$.

But as already mentioned, the value ϱ_d is not our main concern. What we actually want to show, is that the solution v_d converges to the abscissa function. Before approaching this task, we need to consider that there might be no optimal solution.

Example 2.2.1. The infimum in SDP problem (2.3) is not necessarily attained, e.g., consider the polynomial $p(q, s) = s^2$. Then $p_{\Re}(q, u, z) = u^2 - z^2$, $p_{\Im}(q, u, z) = 2uz$ and $\mathbf{Z} = \mathbf{Q} \times \{(0, 0)\}$. Obviously, the optimal solution to LP problem (2.1) is $v \equiv 0$. For SDP problem (2.3) we would want

$$\begin{aligned} v(q) - u = \sigma_0(q, u, z) + \sigma_1(q, u, z)(1 - q^2) \\ + \tau_{\Re}(q, u, z)(u^2 - z^2) + 2\tau_{\Im}(q, u, z)uz, \end{aligned}$$

meaning $0 \equiv v = u + \sigma_0 + \sigma_1(1 - q^2) + \tau_{\mathfrak{R}}u^2 - \tau_{\mathfrak{R}}z^2 + 2\tau_{\mathfrak{S}}uz$ with σ_0, σ_1 sums of squares. This is impossible, since it would require the construction of the term $-u$, which in this case is only possible as a summand of σ_0 . Then, however, we would always also produce a constant positive term. Practically this means that the multipliers $\sigma_0, \sigma_1, \tau_{\mathfrak{R}}, \tau_{\mathfrak{S}}$ blow up.

Hence, an optimal solution might not exist, but we always have a near optimal solution. This means we should allow solutions v_d with $\int_{\mathbf{Q}} v_d(q) dq \leq \varrho_d + \frac{1}{d}$, e.g., in the above example we would search for $v \equiv \varepsilon$ for an $\varepsilon > 0$ sufficiently small.

Remark. The existence of an optimal solution depends on further conditions, such as the ideal generated by the polynomials $1 - q_j^2$, $p_{\mathfrak{R}}$, and $p_{\mathfrak{S}}$ being radical, and goes beyond the scope of this thesis. The interested reader is referred to the proof of [HL12, Lemma 1] for further details.

In the following theorem we prove that the associated sequence of solutions converges.

Theorem 2.2. *Let $v_d \in \mathbb{R}[q]_{2d}$ be a near optimal solution for SDP problem (2.3), i.e., $\int_{\mathbf{Q}} v_d(q) dq \leq \varrho_d + \frac{1}{d}$, and consider the associated sequence $(v_d)_{d \geq d_0} \subseteq L^1(\mathbf{Q})$. Then this sequence converges to the abscissa a in L^1 norm on \mathbf{Q} as d tends to infinity, i.e., $\lim_{d \rightarrow \infty} \int_{\mathbf{Q}} |v_d(q) - a(q)| dq = 0$.*

Proof. Recall that $\varrho^* = \varrho$ according to Lemma 2.1. First we show that $\varrho = \int_{\mathbf{Q}} a(q) dq$. For every $(q, u, z) \in \mathbf{Z}$ we have $u \leq a(q)$ and since $\int_{\mathbf{Z}} q^\alpha d\mu = \int_{\mathbf{Q}} q^\alpha dq$ for all $\alpha \in \mathbb{N}^n$ which means that the marginal of μ on \mathbf{Q} is the Lebesgue measure on \mathbf{Q} (see Remark 2.2.1), it follows that for every feasible solution $\mu \in \mathcal{M}_+(\mathbf{Z})$

$$\int_{\mathbf{Z}} u d\mu(q, u, z) \leq \int_{\mathbf{Z}} a(q) d\mu(q, u, z) = \int_{\mathbf{Q}} a(q) dq.$$

Hence $\varrho \leq \int_{\mathbf{Q}} a(q) dq$. On the other hand, for every $q \in \mathbf{Q}$ there exists $(q, u_q, z_q) \in \mathbf{Z}$ such that $a(q) = u_q$. Let $\hat{\mu}$ be the Borel measure concentrated on (q, u_q, z_q) for all $q \in \mathbf{Q}$, i.e., for \mathbf{A} in the Borel sigma algebra of \mathbf{Z} it holds that

$$\hat{\mu}(\mathbf{A}) := \mathbf{1}_{\mathbf{A}}(q, u_q, z_q).$$

Then $\hat{\mu}$ is feasible for problem (2.2) with value

$$\int_{\mathbf{Z}} u d\hat{\mu}(q, u, z) = \int_{\mathbf{Q}} a(q) dq,$$

which proves that $\varrho \geq \int_{\mathbf{Q}} a(q) dq$, hence $\varrho = \int_{\mathbf{Q}} a(q) dq$.

Next we show convergence in L^1 . Since the abscissa a is continuous on the compact set \mathbf{Q} , by the Stone-Weierstraß theorem [Zor04, §16.4.3] it holds that for every $\varepsilon > 0$ there exists a polynomial $h_\varepsilon \in \mathbb{R}[q]$ such that

$$\sup_{q \in \mathbf{Q}} |h_\varepsilon(q) - a(q)| < \frac{\varepsilon}{2}.$$

Hence, the polynomial $v_\varepsilon := h_\varepsilon + \varepsilon$ satisfies $v_\varepsilon - a > 0$ on \mathbf{Q} and we have $v_\varepsilon(x) - u > 0$ on \mathbf{Z} . Since the corresponding quadratic module is archimedean (see discussion following (2.3)), by Putinar's Positivstellensatz, Theorem 1.1, there exist $\sigma_0^\varepsilon, \sigma_j^\varepsilon \in \Sigma[q, u, z]$, $\tau_{\mathfrak{R}}^\varepsilon, \tau_{\mathfrak{S}}^\varepsilon \in \mathbb{R}[q, u, z]$ such that for all $(q, u, z) \in \mathbb{R}^n \times \mathbb{R}^2$ we can write

$$\begin{aligned} v_\varepsilon(q) - u &= \sigma_0^\varepsilon(q, u, z) + \sum_{j=1}^n \sigma_j^\varepsilon(q, u, z)(1 - q_j^2) \\ &\quad + \tau_{\mathfrak{R}}^\varepsilon(q, u, z)p_{\mathfrak{R}}(q, u, z) + \tau_{\mathfrak{S}}^\varepsilon(q, u, z)p_{\mathfrak{S}}(q, u, z). \end{aligned}$$

Therefore, for $d \geq d_\varepsilon := \lceil \deg(v_\varepsilon)/2 \rceil$ the tuple $(v_\varepsilon, \sigma_0^\varepsilon, \sigma_j^\varepsilon, \tau_{\Re}^\varepsilon, \tau_{\Im}^\varepsilon)$ is a feasible solution for SDP problem (2.3) satisfying

$$0 \leq \int_{\mathbf{Q}} (v_\varepsilon(q) - a(q)) dq \leq \frac{3\varepsilon}{2} \int_{\mathbf{Q}} dq.$$

Together with $\int_{\mathbf{Q}} a(q) dq = \varrho \leq \varrho_d$ which is due to the first part of the proof and ϱ_d being a strengthening of ϱ , it follows that whenever $d \geq d_\varepsilon$,

$$0 \leq \varrho_d - \int_{\mathbf{Q}} a(q) dq \leq \int_{\mathbf{Q}} (v_\varepsilon(q) - a(q)) dq \leq \frac{3\varepsilon}{2} \int_{\mathbf{Q}} dq.$$

As $\varepsilon > 0$ was arbitrary, we obtain $\lim_{d \rightarrow \infty} \varrho_d = \int_{\mathbf{Q}} a(q) dq$ and since $a \leq v_d$ for all d , this is the same as convergence in L^1 :

$$\begin{aligned} 0 &\leq \lim_{d \rightarrow \infty} \|v_d - a\|_1 = \lim_{d \rightarrow \infty} \int_{\mathbf{Q}} |v_d(q) - a(q)| dq \\ &= \lim_{d \rightarrow \infty} \int_{\mathbf{Q}} (v_d(q) - a(q)) dq \leq \lim_{d \rightarrow \infty} \left(\varrho_d + \frac{1}{d} \right) - \int_{\mathbf{Q}} a(q) dq = 0. \end{aligned}$$

□

As mentioned in Section 2.1, the abscissa function is important when studying stability of linear differential equations, and the zero sublevel set of the abscissa of the characteristic polynomial $\{q \in \mathbf{Q} : a(q) < 0\}$ is called the stability region. The following statement on polynomial inner approximations of this set follows immediately from the L^1 convergence result of Theorem 2.2; see also [HL12].

Corollary 2.3. *Let $v_d \in \mathbb{R}[q]_{2d}$ denote, as in Theorem 2.2, a near optimal solution for SDP problem (2.3). Then $\{q \in \mathbf{Q} : v_d(q) < 0\} \subseteq \{q \in \mathbf{Q} : a(q) < 0\}$ and $\lim_{d \rightarrow \infty} \text{vol} \{q \in \mathbf{Q} : v_d(q) < 0\} = \text{vol} \{q \in \mathbf{Q} : a(q) < 0\}$.*

2.2.3 Examples

As stated in Corollary 2.3, while approximating the abscissa function from above we also get an inner approximation of the stability region. The authors of [HL12] surveyed a different approach. They described the stability region via the eigenvalues of the Hermite matrix of the polynomial and approximated it using an SDP hierarchy. In the following examples we compare the two different methods and highlight the specific advantages of our abscissa approximation.

Example 2.2.2 (The damped oscillator [Cro10]). Consider the second degree polynomial depending on $n = 1$ parameter $q \in \mathbf{Q} = [-1, 1]$,

$$p : s \mapsto p(q, s) = s^2 + 2qs + 1 - 2q.$$

Then $\mathbf{Z} = \{(q, u, z) \in [-1, 1] \times \mathbb{R}^2 : u^2 - z^2 + 2qu + 1 - 2q = 2uz + 2qz = 0\}$ and the corresponding hierarchy of SDP problems (2.3) reads

$$\begin{aligned} \varrho_d &= \inf_{v_d, \sigma_0, \sigma_1, \tau_{\Re}, \tau_{\Im}} \int_{-1}^1 v_d(q) dq \\ \text{s.t. } v_d(q) - u &= \sigma_0(q, u, z) + \sigma_1(q, u, z)(1 - q^2) \\ &\quad + \tau_{\Re}(q, u, z)(u^2 - z^2 + 2qu + 1 - 2q) + \tau_{\Im}(q, u, z)(2uz + 2qz) \end{aligned}$$

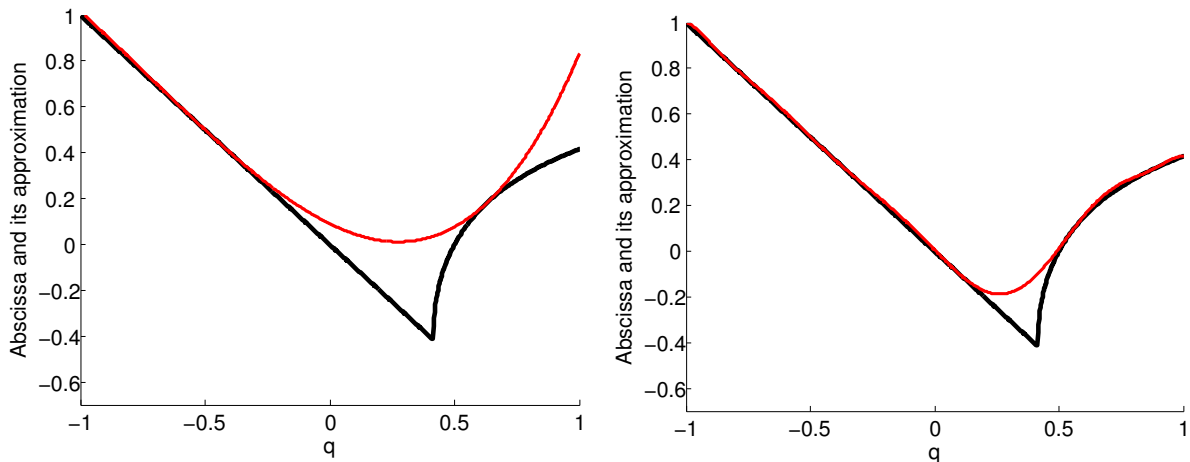


Figure 2.1 – Abscissa $a(q)$ (black) and its polynomial upper approximations $v_d(q)$ of degree $2d = 4$ (left, red) and $2d = 10$ (right, red) for Example 2.2.2. The quality of the approximation deteriorates near the minimum, where the abscissa is not Lipschitz.

for all $(q, u, z) \in \mathbb{R}^3$ and with $v_d \in \mathbb{R}[q]_{2d}$, $\sigma_0 \in \Sigma[q, u, z]_{2d}$, $\sigma_1 \in \Sigma[q, u, z]_{2d-2}$, and $\tau_{\mathbb{R}}, \tau_{\mathbb{C}} \in \mathbb{R}[q, u, z]_{2d-2}$. Apart from that, we only need the moments of the Lebesgue measure on $[-1, 1]$ for a successful implementation. These are readily given by

$$y_\alpha = \int_{-1}^1 q^\alpha dq = \frac{1 - (-1)^{\alpha+1}}{\alpha + 1},$$

meaning that $\int_{-1}^1 v_d(q) dq = \sum_{\alpha=1}^d v_{d\alpha} y_\alpha$ with $v_{d\alpha}$ denoting the coefficient of the monomial q^α of v_d . See Figure 2.1 for the graphs of the degrees 4 (i.e., $d = 2$) and 10 (i.e., $d = 5$) polynomial upper approximations of the abscissa.

For the Hermite approximation we compute the Hermite matrix H of p (see [HPAŠ03] for details)

$$H(q) = \begin{pmatrix} 4q - 8q^2 & 0 \\ 0 & 4q \end{pmatrix}$$

and write the hierarchy of optimization problems as presented in [HL12]:

$$\begin{aligned} & \max_{g_d, \sigma_0, \sigma_1, \tau} \int_{-1}^1 g_d(q) dq \\ & \text{s.t. } \mathbf{u}^T H(q) \mathbf{u} - g_d(q) = \sigma_0(q, \mathbf{u}) + \sigma_1(q, \mathbf{u})(1 - q^2) + \tau(q, \mathbf{u})(1 - \mathbf{u}^T \mathbf{u}) \end{aligned}$$

for all $(q, \mathbf{u}) \in [-1, 1] \times \mathbb{R}^2$ and with $g_d \in \mathbb{R}[q]_{2d}$, $\sigma_0 \in \Sigma[q, \mathbf{u}]_{2d}$, $\sigma_1 \in \Sigma[q, \mathbf{u}]_{2d-2}$, and $\tau \in \mathbb{R}[q, \mathbf{u}]_{2d-2}$. Already for degree $2d = 6$ we observe a close match between the genuine stability region, which is $\{q \in [-1, 1] : a(q) < 0\} = (0, \frac{1}{2})$, the Hermite inner approximation $\{q \in [-1, 1] : -g_3(q) < 0\}$, and the polynomial upper approximation $\{q \in [-1, 1] : v_5(q) < 0\}$. These three intervals are visually indistinguishable, so we do not represent them graphically.

Example 2.2.3. Consider the polynomial

$$p : s \mapsto p(q, s) = s^3 + \frac{1}{2}s^2 + q^2s + (q - \frac{1}{2})q(q + \frac{1}{2})$$

for $q \in \mathbf{Q} = [-1, 1]$. The abscissa function $a(q)$ of p is not differentiable at three points and therefore it is rather hard to approximate in their neighborhoods. In Figure 2.2 we see the abscissa and its polynomial upper approximations of degrees 6 ($d = 3$) and 12 ($d = 6$). Comparing

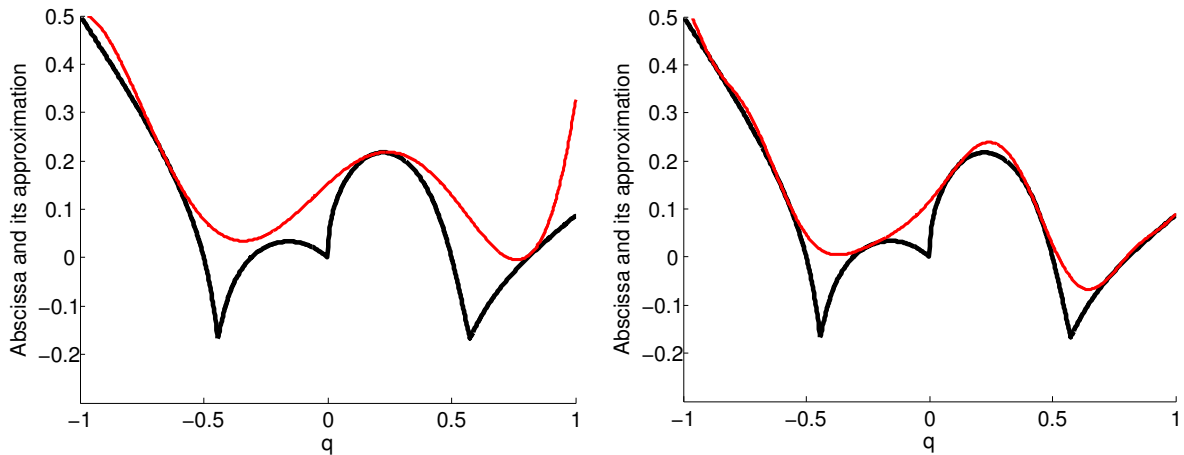


Figure 2.2 – Abscissa $a(q)$ (black) and its polynomial upper approximations $v_d(q)$ of degree $2d = 6$ (left, red) and $2d = 12$ (right, red) for Example 2.2.3. The quality of the approximation deteriorates near the points of non-differentiability of the abscissa.

the genuine stability region $\{q \in [-1, 1] : a(q) < 0\}$, the polynomial inner approximation $\{q \in [-1, 1] : v_6(q) < 0\}$, and the Hermite inner approximation $\{q \in [-1, 1] : -g_5(q) < 0\}$, we observe, maybe surprisingly, that the approximations are very similar and miss the same parts of the stability region. These are not reproduced graphically.

Remark. The approach via the Hermite matrix does not tell us anything about the abscissa function itself but just approximates its zero sublevel set, the so-called stability region defined in Section 2.1. As an illustration consider a polynomial of the form $p(q, s) = s^2 + p_0(q)$ for $n = 1$. Then $p(q, \cdot)$ has either 0 as a multiple root, two real roots (of which one is positive), or only imaginary roots. Since these are all possible cases it follows that the stability region of p is empty and its Hermite matrix $H(q)$ is zero. Therefore its eigenvalues and their approximation g_d are also zero for every d . In contrast, the upper abscissa approximation v_d always gives a suitable approximation for the abscissa function.

On the other hand, practical experiments (not reported here) reveal that computing the abscissa approximation is typically more challenging numerically than computing the Hermite approximation. For instance, computing the upper abscissa approximation may fail for polynomials with large coefficients, while the Hermite approximation continues to provide a proper inner approximation of the stability region.

Example 2.2.4. Consider the polynomial

$$p : s \mapsto p(q, s) = s^3 + (q_1 + \frac{3}{2})s^2 + q_1^2s + q_1q_2$$

depending on $n = 2$ parameters $q \in \mathbf{Q} = [-1, 1]^2$. Then $\mathbf{Z} = \{(q, u, z) \in [-1, 1]^2 \times \mathbb{R}^2 : u^3 - 3uz^2 + (q_1 + \frac{3}{2})u^2 - (q_1 + \frac{3}{2})z^2 + q_1^2u + q_1q_2 = -z^3 + 3u^2z + 2(q_1 + \frac{3}{2})uz + q_1^2z = 0\}$. See Figure 2.3 for the plot of the abscissa of p .

In Figure 2.4 we represent the graphs of the abscissa a and its polynomial approximations v_3 and v_5 . In Figure 2.5 we represent the stability region, i.e., the zero sublevel set of the abscissa $\{q \in [-1, 1]^2 : a(q) < 0\}$ (blue region), the degree 8 ($d = 4$) Hermite sublevel set $\{q \in [-1, 1]^2 : -g_4(q) < 0\}$ (green region, left), and the degree 10 ($d = 5$) polynomial sublevel set $\{q \in [-1, 1]^2 : v_5(q) < 0\}$ (red region, right).

Remark. In the examples we always chose lower degrees for the Hermite approximation than for the upper abscissa approximation. The Hermite approximation converges relatively fast, making

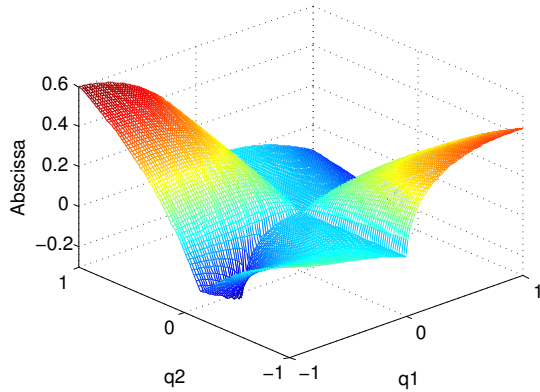


Figure 2.3 – Abscissa function $a(q)$ for Example 2.2.4.

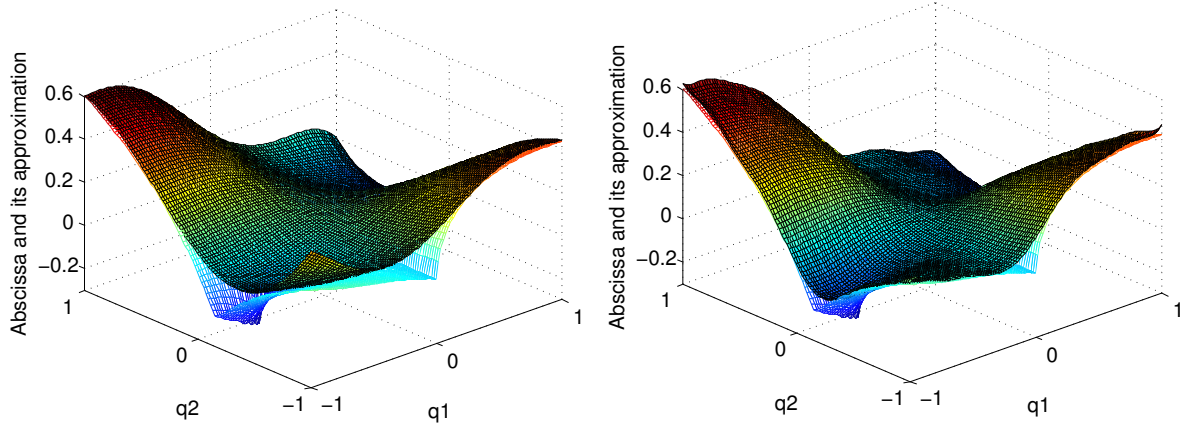


Figure 2.4 – Abscissa $a(q)$ (light, below) and its polynomial upper approximations $v_d(q)$ of degrees $2d = 6$ (dark, left) and $2d = 10$ (dark, right) for Example 2.2.4. See Figure 2.3 for the graph of a only. We observe that the approximation deteriorates near the regions of non-differentiability of the abscissa.

it unnecessary to consider higher degrees, especially since they require much more time. On the contrary, the upper abscissa approximation usually needs higher degrees to provide a useful approximation, but it is faster to compute.

2.3 Lower abscissa approximation

At first thought, finding a lower approximation for the abscissa map might sound like a straightforward task, since one is tempted to just solve the analogue of LP problem (2.1),

$$\begin{aligned} & \sup_{w \in \mathcal{L}(\mathbf{Q})} \int_{\mathbf{Q}} w(q) dq \\ & \text{s.t. } u - w(q) \geq 0 \text{ for all } (q, u, z) \in \mathbf{Z}. \end{aligned} \quad (2.4)$$

This, indeed, gives a valid lower bound on the abscissa function, although in general a very bad one since it is approximating not the abscissa but the minimal real part of the roots of p . To understand the reason we recall that

$$\mathbf{Z} = \{(q, u, z) \in \mathbb{R}^n \times \mathbb{R}^2 : q \in \mathbf{Q}, p_{\Re}(q, u, z) = p_{\Im}(q, u, z) = 0\}$$

and therefore this set contains all roots of p and not only those with maximal real part.

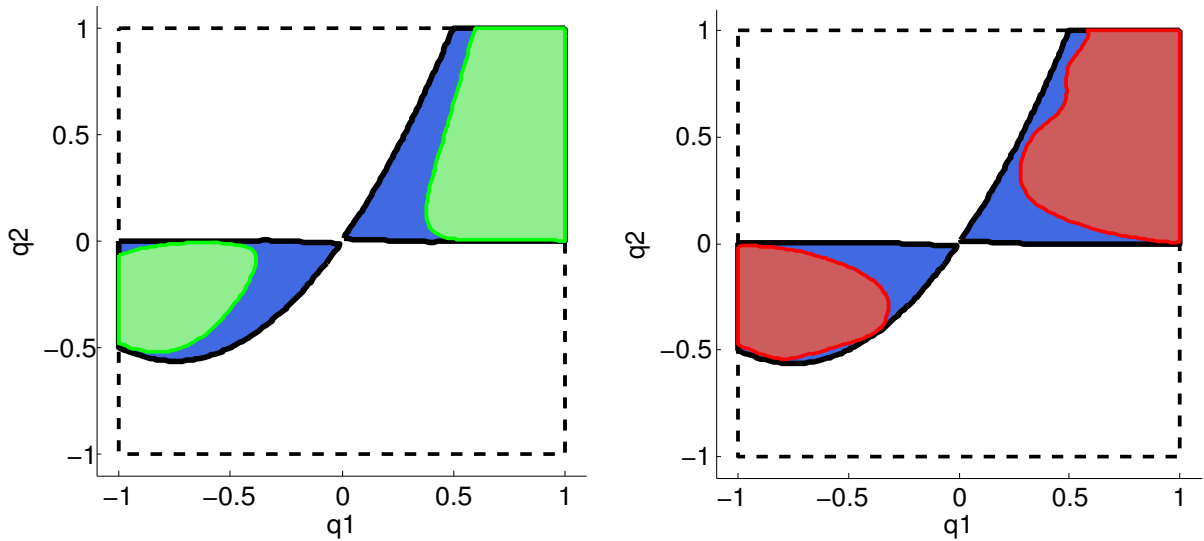


Figure 2.5 – Stability region (blue region) and its inner approximations with degree $2d = 8$ Hermite (green region, left) and degree $2d = 10$ upper polynomial approximation (red region, right) for Example 2.2.4 (compare with Figure 2.4).

Example 2.3.1. On the left of Figure 2.6 we show the degree 12 ($d = 6$) solution to the SDP hierarchy corresponding to LP problem (2.4) for the polynomial $p(q, s) = s^2 + 2qs + 1 - 2q$ of Example 2.2.2, which gives a tight lower approximation to the abscissa only in the left part of the domain, corresponding to a pair of complex conjugate roots. On the right of Figure 2.6 we show the degree $2d = 12$ solution to the SDP hierarchy corresponding to LP problem (2.4) for the polynomial $p(q, s) = s^3 + \frac{1}{2}s^2 + q^2s + (q - \frac{1}{2})q(q + \frac{1}{2})$ of Example 2.2.3. The lower approximation is nowhere tight, due to the presence of roots with real parts smaller than the abscissa.

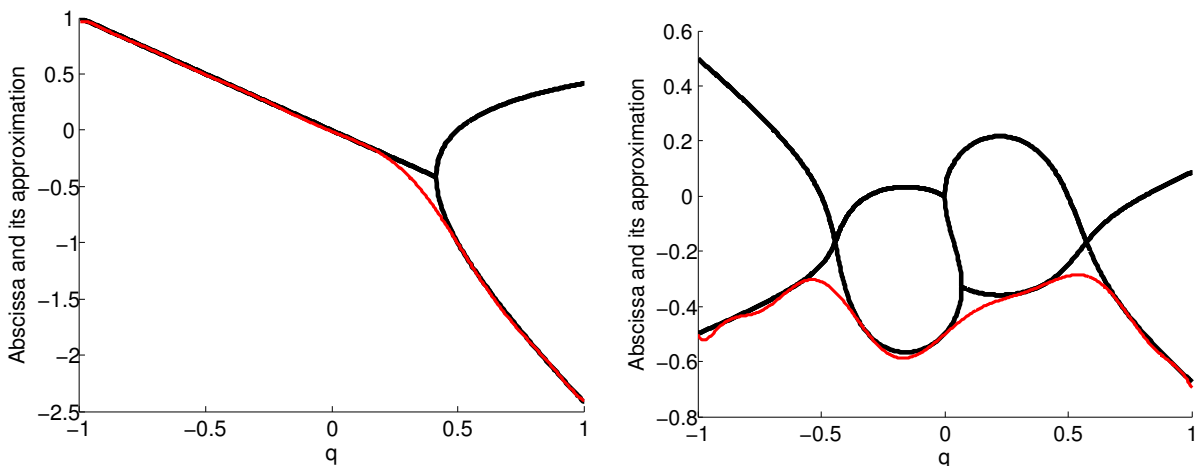


Figure 2.6 – Real parts of the roots (black) and degree $2d = 12$ polynomial lower approximations (red) for the second degree polynomial (left) and third degree polynomial (right) of Example 2.3.1.

To find a tighter approximation for the abscissa map from below we pursue two different approaches:

- First, we reformulate the set \mathbf{Z} with the help of elementary symmetric functions, in order to have access to the roots directly. This is a very neat way with options for variation, such as approximating the second largest real part of the roots from above or below, but it also

includes many additional variables and it is therefore not very efficient when implemented. However, it can be useful for small problems.

- Second, we restrict LP problem (2.4) further using the Gauß-Lucas theorem, i.e., instead of \mathbf{Z} we use a subset of \mathbf{Z} which contains only the roots with the abscissa as its real parts. This approach is much more complicated, and relies on assumptions and one needs to solve two optimization problems in order to get the lower approximation. Nevertheless, the implementation is much faster, so it can be used for bigger problems.

2.3.1 Lower approximation via elementary symmetric functions

2.3.1.1 Problem formulation

Let us derive another description of the set of roots of p which allows us to pick single roots according to the size of their real part. For this purpose let us recall the definition of our polynomial:

$$p : s \mapsto p(q, s) := \sum_{k=0}^{\ell} p_k(q) s^k \quad \text{with} \quad p_{\ell}(q) \equiv 1.$$

Following the notation of the previous sections, we denote the roots of $p(q, \cdot)$ by $s_r(q)$, $r = 1, \dots, \ell$, and split them up into their real and imaginary parts, $s_r(q) = u_r(q) + i z_r(q)$ with $u_r(q), z_r(q) \in \mathbb{R}$. To simplify notation we omit the dependence on q whenever it is clear and write only s_r , u_r , and z_r .

Now we write the coefficients of the polynomial as elementary symmetric functions (in the following sometimes abbreviated as ESF) of its roots:

$$p_{\ell-k}(q) = (-1)^k \sum_{1 \leq l_1 < l_2 < \dots < l_k \leq \ell} s_{l_1} s_{l_2} \dots s_{l_k}, \quad k = 1, \dots, \ell.$$

This allows us to define the set of roots of p in the following way, where we can order the roots according to the size of their real part:

$$\mathbf{Z}'_o := \left\{ (q, u_1, \dots, u_{\ell}, z_1, \dots, z_{\ell}) \in \mathbf{Q} \times \mathbb{R}^{\ell} \times \mathbb{R}^{\ell} : u_r \leq u_{\ell}, r = 1, \dots, \ell - 1, \right. \\ \left. p_{\ell-k}(q) = (-1)^k \sum_{1 \leq l_1 < l_2 < \dots < l_k \leq \ell} s_{l_1} s_{l_2} \dots s_{l_k}, \quad k = 1, \dots, \ell \right\}.$$

To avoid the complex variables s_{l_k} in the description of the set, we could replace them by $s_{l_k} = u_{l_k} + i z_{l_k}$ and split the sum $\sum_{1 \leq l_1 < \dots < l_k \leq \ell} s_{l_1} s_{l_2} \dots s_{l_k}$ into its real and imaginary parts. The latter would be zero, since all $p_{\ell-k}(q)$ are real. In what follows we omit this procedure, since it would only complicate notation.

For illustrative reasons let us fix q for a moment. Then the set \mathbf{Z}'_o contains only one element $(q, u_1, \dots, u_{\ell}, z_1, \dots, z_{\ell})$. From this it holds that $u_{\ell} = a(q)$ and the points (q, u_r, z_r) , $r = 1, \dots, \ell$, are exactly the elements of \mathbf{Z} .

Remark. One could order the roots further by adding more conditions, such as $u_r \leq u_{\ell-1}$, $r = 1, \dots, \ell - 2$. Then one could also access the root with the second largest real part. Of course, this would imply another $\ell - 2$ constraints in an implementation and therefore this would slow down further the solution process.

In theory, ℓ variables suffice to characterize the roots of a real polynomial via the elementary symmetric functions, but since we need all variables u_r explicitly in order to identify the maximal

one, we can only eliminate $\lfloor \ell/2 \rfloor = \max\{c \in \mathbf{Z} : c \leq \ell/2\}$ variables. We set

$$z_{r-1} = -z_r, \quad r = 2, 4, \dots, \ell \quad \text{if } \ell \text{ is even,}$$

meaning we decide which roots will be pairs in case they are complex. If ℓ is odd, we cannot assign the pairs of conjugate roots as easily as we did in the even case, because it is necessary to keep the variable z_ℓ , since we defined u_ℓ as the abscissa. Furthermore, we need to consider that we do not know whether s_ℓ is real. In fact, $s_\ell(q)$ can be real for some q and complex for others. So we set

$$z_{r-1} = -z_r, \quad r = 2, 4, \dots, \ell - 3 \text{ and } z_{\ell-2} = -z_{\ell-1} - z_\ell \quad \text{if } \ell \text{ is odd,}$$

where the latter assignment comes from the imaginary part of the constraint $p_{\ell-1}(q) = -s_1 - s_2 - \dots - s_\ell$ which reads $0 = -z_1 - z_2 - \dots - z_\ell$, and means that at least one of the roots $s_{\ell-2}, s_{\ell-1}, s_\ell$ is real, but we do not specify which.

Remark. Even though we know for ℓ odd that one root must be real, we cannot eliminate $\lfloor \ell/2 \rfloor$ variables, since it might happen that $s_\ell(q)$ is the single real root for some q while it is complex for other q .

Now we can write the set of roots with fewer variables and fewer constraints. As above, we keep the variables s_r in the description of the set for readability reasons but remark that with the reduced amount of z variables the constraints $0 = \Im(\sum_{1 \leq l_1 < \dots < l_k \leq \ell} s_{l_1} s_{l_2} \dots s_{l_k})$ for $k = 1, \dots, \lfloor \ell/2 \rfloor$ are superfluous. We have

$$\begin{aligned} \mathbf{Z}_o := \{ & (q, u_1, \dots, u_\ell, z_2, z_4, \dots, z_{2\lfloor (\ell-1)/2 \rfloor}, z_\ell) \in \mathbf{Q} \times \mathbb{R}^\ell \times \mathbb{R}^{\lfloor \ell/2 \rfloor} : \\ & u_r \leq u_\ell, \quad r = 1, \dots, \ell - 1, \\ & p_{\ell-k}(q) = (-1)^k \sum_{1 \leq l_1 < l_2 < \dots < l_k \leq \ell} s_{l_1} s_{l_2} \dots s_{l_k}, \quad k = 1, \dots, \ell \}. \end{aligned}$$

Example 2.3.2. For $\ell = 3$ the set \mathbf{Z}_o is given by

$$\begin{aligned} \mathbf{Z}_o = \{ & (q, u_1, u_2, u_3, z_2, z_3) \in \mathbf{Q} \times \mathbb{R}^3 \times \mathbb{R}^2 : u_1 \leq u_3, \quad u_2 \leq u_3, \\ & -p_2(q) = u_1 + u_2 + u_3, \\ & p_1(q) = u_1 u_2 + u_1 u_3 + u_2 u_3 + z_2^2 + z_2 z_3 + z_3^2, \\ & -p_0(q) = u_1 u_2 u_3 + (-u_1 + u_2 + u_3) z_2 z_3 + u_2 z_3^2 + u_3 z_2^2, \\ & 0 = (u_1 - u_2) z_2 + (u_1 - u_3) z_3, \\ & 0 = (u_1 - u_2) u_3 z_2 + (u_1 - u_3) u_2 z_3 + z_2^2 z_3 + z_2 z_3^2 \}. \end{aligned}$$

To clarify the formula also for ℓ even, we write \mathbf{Z}_o explicitly for $\ell = 4$,

$$\begin{aligned} \mathbf{Z}_o = \{ & (q, u_1, u_2, u_3, u_4, z_2, z_4) \in \mathbf{Q} \times \mathbb{R}^4 \times \mathbb{R}^2 : u_1 \leq u_4, \quad u_2 \leq u_4, \quad u_3 \leq u_4, \\ & -p_3(q) = u_1 + u_2 + u_3 + u_4, \\ & p_2(q) = u_1 u_2 + u_1 u_3 + u_1 u_4 + u_2 u_3 + u_2 u_4 + u_3 u_4 + z_2^2 + z_4^2, \\ & -p_1(q) = u_1 u_2 (u_3 + u_4) + (u_1 + u_2) u_3 u_4 + (u_1 + u_2) z_4^2 + (u_3 + u_4) z_2^2, \\ & p_0(q) = u_1 u_2 u_3 u_4 + (u_1 - u_2) (u_4 - u_3) z_2 z_4 + u_1 u_2 z_4^2 + u_3 u_4 z_2^2 + z_2^2 z_4^2, \\ & 0 = (u_1 - u_2) (u_3 + u_4) z_2 + (u_1 + u_2) (u_3 - u_4) z_4, \\ & 0 = (u_1 - u_2) (u_3 u_4 + z_4^2) z_2 + (u_3 - u_4) (u_1 u_2 + z_2^2) z_4 \}. \end{aligned}$$

Here we have set $z_1 = -z_2$ and $z_3 = -z_4$, so the constraint $0 = \Im(\sum_{1 \leq l_1 < \dots < l_k \leq \ell} s_{l_1} s_{l_2} \dots s_{l_k})$ for

$k = 1$ is obviously superfluous, because it reduces to $0 = 0$. The second superfluous constraint is the one for $k = 2$, that is, $0 = (u_1 - u_2)z_2 + (u_3 - u_4)z_4$, since we have $u_1 = u_2$, respectively, $u_3 = u_4$, in the case s_2 , respectively, s_4 , is complex.

Finally, we can reformulate LP problem (2.4) in such a way that it provides a proper approximation of the abscissa function from below:

$$\begin{aligned} \vartheta &= \sup_{w \in \mathcal{C}(\mathbf{Q})} \int_{\mathbf{Q}} w(q) dq \\ \text{s.t. } & u_\ell - w(q) \geq 0 \text{ for all } (q, u_1, \dots, u_\ell, z_2, z_4, \dots, z_{2\lfloor(\ell-1)/2\rfloor}, z_\ell) \in \mathbf{Z}_o. \end{aligned} \quad (2.5)$$

With the notation of Section 2.2.1 its dual LP problem reads

$$\begin{aligned} \vartheta^* &= \inf_{\mu \in \mathcal{M}_+(\mathbf{Z}_o)} \int_{\mathbf{Z}_o} u_\ell d\mu(q, u_1, \dots, u_\ell, z_2, z_4, \dots, z_{2\lfloor(\ell-1)/2\rfloor}, z_\ell) \\ \text{s.t. } & \int_{\mathbf{Z}_o} q^\alpha d\mu = \int_{\mathbf{Q}} q^\alpha dq \text{ for all } \alpha \in \mathbb{N}^n. \end{aligned} \quad (2.6)$$

In analogy with the upper approximation we have no duality gap and the infimum is attained.

Lemma 2.4. *The infimum in LP problem (2.6) is attained, and there is no duality gap between LP problem (2.5) and LP problem (2.6), i.e., $\vartheta = \vartheta^*$.*

Since \mathbf{Z}_o is compact, the proof is identical to that of Lemma 2.1.

Remark 2.3.1. For the same reasons as for the upper approximation (2.1), the supremum in (2.5) is not attained for $\mathcal{C}(\mathbf{Q})$ or $\mathbb{R}[q]$, but it is attained for $\mathbb{R}[q]_d$ with d finite. See Remark 2.2.2 with $M := \min_{q \in \mathbf{Q}} a(q) - N$ for $N \in \mathbb{N}$ sufficiently large, and $R := \int_{\mathbf{Q}} (a(q) - M) dq$.

2.3.1.2 Lasserre hierarchy

Similarly to the upper bound, we apply the method described in Section 1.4.1, meaning we exchange non-negativity on \mathbf{Z}_o by membership of the truncated quadratic module generated by the polynomials describing \mathbf{Z}_o .

Let $d_0 \in \mathbb{N}$ be sufficiently large, $d_0 \geq \lceil \ell/2 \rceil$. Then for $d \in \mathbb{N}$, $d \geq d_0$ the corresponding hierarchy of SDP problems reads

$$\begin{aligned} \vartheta_d &= \sup_{w_d, \sigma_0, \sigma_j, \sigma_{u_r}, \tau_{\Re, k}, \tau_{\Im, k}} \int_{\mathbf{Q}} w_d(q) dq \\ \text{s.t. } & u_\ell - w_d(q) = \sigma_0 + \sum_{j=1}^n \sigma_j (1 - q_j^2) + \sum_{r=1}^{\ell-1} \sigma_{u_r} (u_\ell - u_r) \\ & + \sum_{k=1}^{\ell} \tau_{\Re, k} \left((-1)^k p_{\ell-k}(q) - \Re \left(\sum_{1 \leq l_1 < l_2 < \dots < l_k \leq \ell} s_{l_1} s_{l_2} \dots s_{l_k} \right) \right) \\ & + \sum_{k=\lceil \ell/2 \rceil}^{\ell} \tau_{\Im, k} \Im \left(\sum_{1 \leq l_1 < l_2 < \dots < l_k \leq \ell} s_{l_1} s_{l_2} \dots s_{l_k} \right) \end{aligned} \quad (2.7)$$

for all $(q, u_1, \dots, u_\ell, z_2, z_4, \dots, z_{2\lfloor(\ell-1)/2\rfloor}, z_\ell) \in \mathbb{R}^n \times \mathbb{R}^\ell \times \mathbb{R}^{\lceil \ell/2 \rceil}$ and with $w_d \in \mathbb{R}[q]_{2d}$, $\sigma_0, \sigma_{u_r} \in \Sigma[q, u_1, \dots, u_\ell, z_2, z_4, \dots, z_\ell]_{2d}$ for $r = 1, \dots, \ell - 1$, $\sigma_j \in \Sigma[q, u_1, \dots, u_\ell, z_2, z_4, \dots, z_\ell]_{2d-2}$ for $k = 1, \dots, n$, $\tau_{\Re, k} \in \mathbb{R}[q, u_1, \dots, u_\ell, z_2, z_4, \dots, z_\ell]_{2d-k}$ for $k = 1, \dots, \ell$, and $\tau_{\Im, k}$ for $k = \lceil \ell/2 \rceil, \dots, \ell$.

Since we are maximizing, SDP problem (2.7) is a strengthening of LP problem (2.5), meaning $\vartheta_d \leq \vartheta$. Furthermore, the quadratic module corresponding to \mathbf{Z}_o is archimedean, i.e., $\lim_{d \rightarrow \infty} \vartheta_d = \vartheta$.

We conclude the section with the following result.

Theorem 2.5. *Let $w_d \in \mathbb{R}[q]_{2d}$ be a near optimal solution for SDP problem (2.7), i.e., $\int_{\mathbf{Q}} w_d(q) dq \geq \vartheta_d - \frac{1}{d}$, and consider the associated sequence $(w_d)_{d \geq d_0} \subseteq L^1(\mathbf{Q})$. Then this sequence converges to a in L^1 norm on \mathbf{Q} .*

Unsurprisingly, one can prove this result in exactly the same way as Theorem 2.2, so we do not detail the proof here. We remark that the first part of the proof can be shortened, since $\int_{\mathbf{Z}_o} u_\ell d\mu(q, u_1, \dots, u_\ell, z_2, z_4, \dots, z_\ell) = \int_{\mathbf{Z}_o} a(q) d\mu(q, u_1, \dots, u_\ell, z_2, z_4, \dots, z_\ell)$.

2.3.1.3 Examples

Just as the upper abscissa approximation automatically approximates the stability region from inside, the lower approximation gives, as a side effect, an outer approximation. In this section we will examine the same examples as for the upper approximation.

Example 2.3.3. As in Example 2.2.2, consider the polynomial

$$p: s \mapsto p(q, s) = s^2 + 2qs + 1 - 2q.$$

We have $z_1 = -z_2$, so $\mathbf{Z}_o := \{(q, u_1, u_2, z_2) \in \mathbf{Q} \times \mathbb{R}^3 : u_1 \leq u_2, -2q = u_1 + u_2, 1 - 2q = u_1 u_2 + z_2^2, 0 = (u_1 - u_2)z_2\}$. In Figure 2.7 we see the graphs of the degrees 6 ($d = 3$) and 10 ($d = 5$) polynomial lower approximations obtained by solving SDP problem (2.7). Due to the rather large number of variables and constraints, computing the degree 10 solution is already relatively expensive, with a few seconds of CPU time.

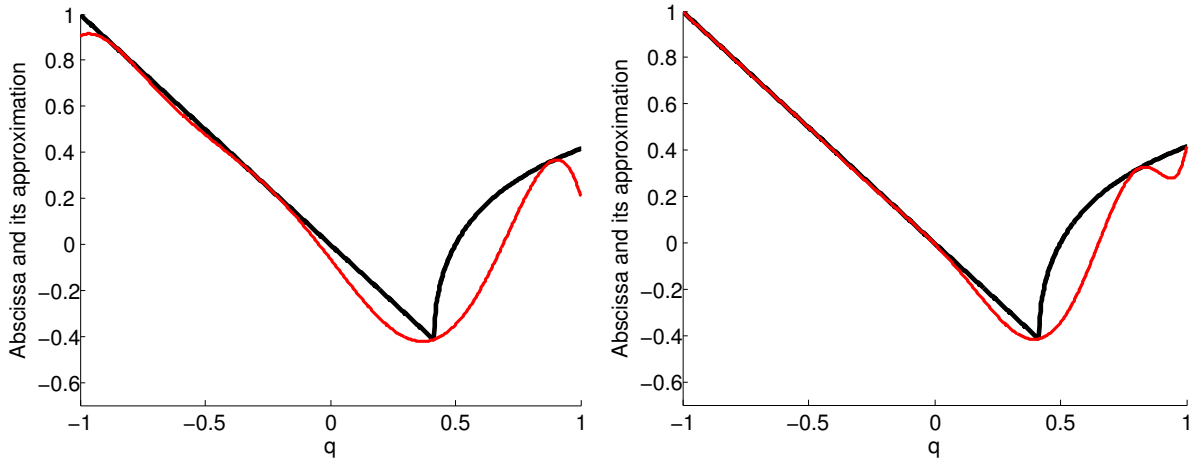


Figure 2.7 – Abscissa $a(q)$ (black) and its polynomial lower ESF approximations $w_d(q)$ of degree $2d = 6$ (red, left) and $2d = 10$ (red, right) for Example 2.3.3. The quality of the approximation deteriorates near the minimum, where the abscissa is not Lipschitz. Compare with Figure 2.1.

Example 2.3.4. As in Example 2.2.3, consider the polynomial

$$p: s \mapsto p(q, s) = s^3 + \frac{1}{2}s^2 + q^2s + (q - \frac{1}{2})q(q + \frac{1}{2}).$$

With $z_1 = -z_2 - z_3$ we calculate \mathbf{Z}_o as in Example 2.3.2. In Figure 2.8 we see the graphs of the degrees 6 ($d = 3$) and 10 ($d = 5$) polynomial lower approximations obtained by solving SDP problem (2.7). The computation time to get the degree 10 solution is nearly 7 minutes, which is arguably unreasonable given the quality of the approximation.

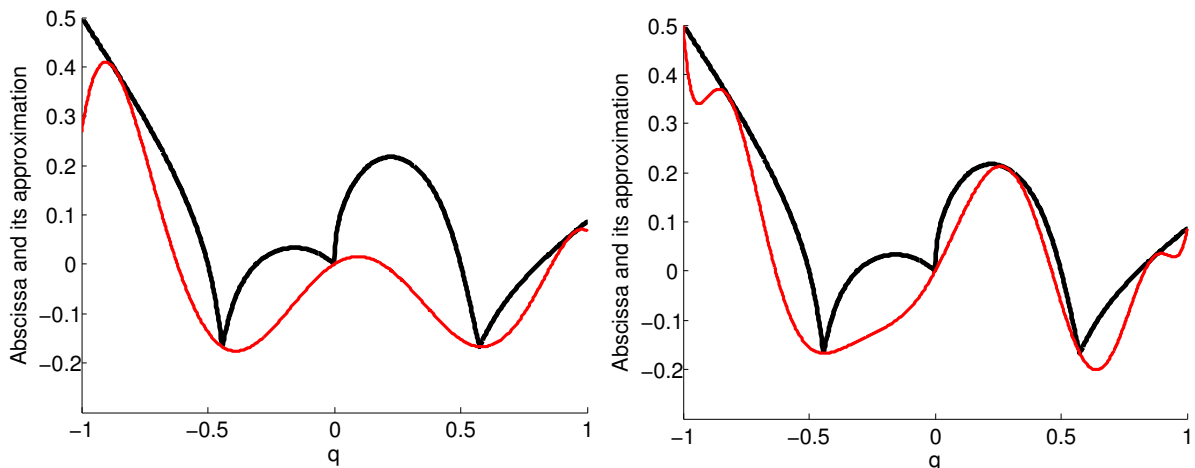


Figure 2.8 – Abscissa $a(q)$ (black) and its polynomial lower ESF approximations $w_d(q)$ of degree $2d = 6$ (red, left) and $2d = 10$ (red, right) for Example 2.3.4. The quality of the approximation deteriorates near the minimum, where the abscissa is not differentiable. Compare with Figure 2.2.

Remark. As for the upper abscissa approximation, we observe practically that the implementation for the lower approximation is rather sensitive to polynomials with large coefficients.

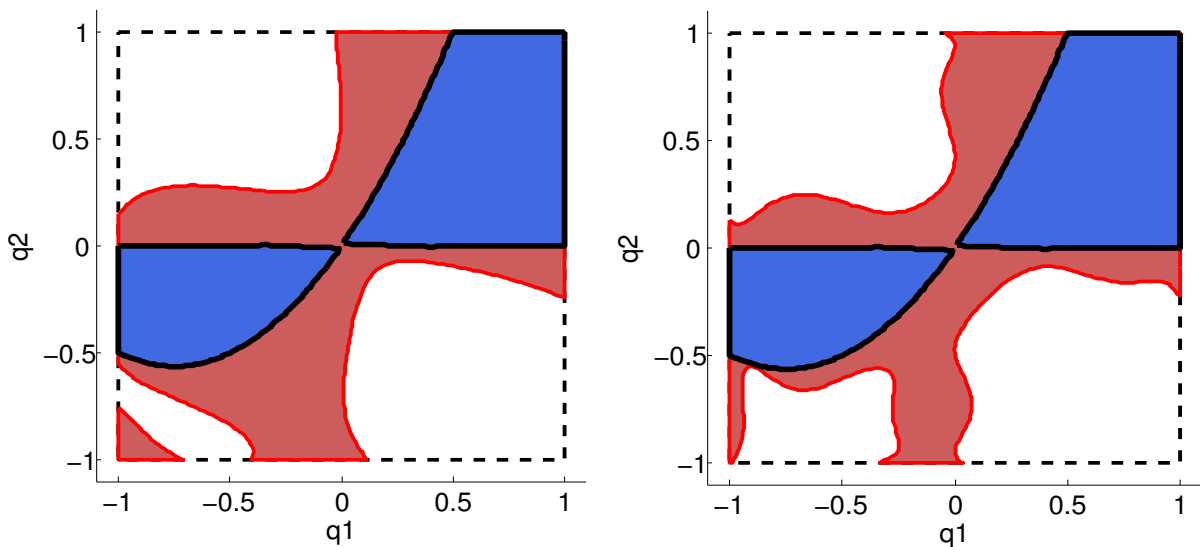


Figure 2.9 – Stability region (blue region) and its degree $2d = 6$ outer ESF approximation (red region, left) and degree $2d = 8$ outer ESF approximation (red region, right) for Example 2.3.5. Compare with Figure 2.5.

Example 2.3.5. As in Example 2.2.4, consider the polynomial

$$p : s \mapsto p(q, s) = s^3 + (q_1 + \frac{3}{2})s^2 + q_1^2 s + q_1 q_2.$$

Since we have degree $\ell = 3$, the set \mathbf{Z}_o is again given in Example 2.3.2. In Figure 2.9 we see the outer approximation of degrees 6 ($d = 3$) and 8 ($d = 4$) obtained by solving SDP problem (2.7). In the lower half of the picture we notice that the approximation of the stability region is rather bad near $q_1 = 0$, even for degree $2d = 8$. This is due to a being zero and non-smooth for $q_1 = 0$, meaning $a(0, q_2) = 0$ for all $q_2 \in [-1, 1]$, which makes the abscissa especially hard to approximate in this region. This phenomenon also prevents w_4 from getting closer to a for

$q_2 > 0$ than we observe in the picture. The degree 6 solution is computed in around 5 seconds, the degree 8 solution takes more than 4 minutes.

2.3.2 Lower approximation via Gauß-Lucas

2.3.2.1 Problem formulation

As indicated above, we want to find a semialgebraic subset of \mathbf{Z} which contains only those roots of p whose real part is maximal. In contrast to the approach of Section 2.3.1, we will not redefine \mathbf{Z} but formulate further constraints.

In order to do this we must distinguish between the roots of $p(q, \cdot)$ according to the size of their real parts. For this purpose we use the following result (in the following sometimes abbreviated as GL).

Theorem 2.6 (Gauß-Lucas). *The critical points of a non-constant polynomial lie in the convex hull of its roots.*

We refer to [BLO04] for further information and a proof. Let us denote the derivative of $p(q, s)$ with respect to s by $p'(q, s)$. By Theorem 2.6, the roots of $p'(q, \cdot)$ are contained in the convex hull of the roots of $p(q, \cdot)$. It follows readily that the abscissa $a_{p'}$ of p' lies below the abscissa a_p of p ,

$$a_{p'}(q) \leq a_p(q) \text{ for all } q \in \mathbf{Q}.$$

However, p may have some roots with real part strictly smaller than a_p and strictly bigger than $a_{p'}$, meaning that the root whose real part is the abscissa is not the only one whose real part lies above $a_{p'}$. Of course, this cannot happen for polynomials $\mathbb{R} \rightarrow \mathbb{R}$ because of monotonicity, and neither can it for polynomials $\mathbb{C} \rightarrow \mathbb{C}$ of degree 2. But, for example, for $n = 1$ (i.e. $\mathbf{Q} \subseteq \mathbb{R}$) the polynomial $p(q, \cdot) : \mathbb{C} \rightarrow \mathbb{C}$, $p(q, s) = s^4 + (q^2 + 1)s + q$ has two roots with different real parts greater than $a_{p'}$ for $q \in [-1, -0.4]$.

To prevent the lower abscissa approximation from converging to the real part of a root smaller than the abscissa, we make the following assumption.

Assumption 1. *None of the real parts of any root of p that differ from the abscissa coincide with $a_{p'}$ or lie strictly between a_p and $a_{p'}$, i.e., $u \notin [a_{p'}(q), a_p(q)[$ for all $(q, u, z) \in \mathbf{Z}$.*

Remark. Unfortunately, we do not know how restrictive this assumption is. For $n = 1$ it was rather difficult to find examples that violate it.

Now let $\hat{v} \in \mathcal{C}(\mathbf{Q})$ be a near optimal solution to LP problem (2.1) for the polynomial p' , meaning $\int_{\mathbf{Q}} \hat{v}(q) dq \leq \varrho + \varepsilon$ for an $\varepsilon > 0$. Then, \hat{v} is an upper approximation of the abscissa $a_{p'}$ of p' . We define the following subset:

$$\hat{\mathbf{Z}} := \{(q, u, z) \in \mathbf{Z} : u - \hat{v}(q) \geq 0\}.$$

In order to see where we are going, let us suppose for a moment that \hat{v} is actually an optimal solution. Then, under Assumption 1, the set $\hat{\mathbf{Z}}$ would contain exactly the points $(q, a_p(q), z_q)$ (with z_q denoting the imaginary part of the root of $p(q, \cdot)$) with maximal real part. Hence, the solution to the following LP problem would give a lower approximation of the abscissa function a_p of p :

$$\begin{aligned} & \sup_{w \in \mathcal{C}(\mathbf{Q})} \int_{\mathbf{Q}} w(q) dq \\ \text{s.t. } & u - w(q) \geq 0 \text{ for all } (q, u, z) \in \hat{\mathbf{Z}}. \end{aligned} \tag{2.8}$$

Since \hat{v} might not be optimal, the projection of $\hat{\mathbf{Z}}$ onto \mathbf{Q} might not be \mathbf{Q} as required, but differ from it on a set of volume ε . As a consequence, w might not be a valid lower bound of the abscissa on this set.

Taking this into account, we build an SDP hierarchy for LP problem (2.8) in the next section. The issue is that we have to consider the hierarchy for the upper approximation of $a_{p'}$ first and the solution to it might interfere with a_p .

2.3.2.2 Lasserre hierarchy

For $d'_0 \in \mathbb{N}$ sufficiently large we denote by $\hat{v}_{d'}$, $d' \geq d'_0$, the solutions to the SDP problems (2.3) for the polynomial p' . Thus, the $\hat{v}_{d'}$ are polynomials in $\mathbb{R}[q]_{2d'}$ and by Theorem 2.2 the sequence $(\hat{v}_{d'})_{d' \in \mathbb{N}}$ converges to $a_{p'}$ from above in L^1 norm.

Next, we want to describe the set $\hat{\mathbf{Z}}$ via the polynomials $\hat{v}_{d'}$ in order to have an implementable problem, i.e., we define

$$\hat{\mathbf{Z}}_{d'} := \{(q, u, z) \in \mathbf{Z} : u - \hat{v}_{d'}(q) \geq 0\}.$$

Of course, the set $\hat{\mathbf{Z}}_{d'}$ is highly dependent on the quality of $\hat{v}_{d'}$ and hence on the choice of d' . Evidently, $\hat{\mathbf{Z}}_{d'}$ is a subset of $\hat{\mathbf{Z}}$, possibly strictly. To ensure that $\hat{\mathbf{Z}}_{d'}$ contains all roots of p with the abscissa as their real parts we need $\hat{v}_{d'} \leq a_p$. However, in practice this is impossible in some cases.

Example 2.3.6. The abscissa a_p of $p(q, s) = (s^3 + q)^2$ and the abscissa $a_{p'}$ of p' coincide and have a point of non-differentiability at $q = 0$. As another example consider the polynomial $p(q, s) = s^4 + qs$ for which both a_p and $a_{p'}$ are not differentiable at $q = 0$ and $a_p(0) = a_{p'}(0) = 0$.

For these examples we cannot achieve $\hat{v}_{d'} \leq a_p$ with d' finite, since $\hat{v}_{d'}$ is a polynomial and therefore differentiable everywhere.

As a consequence, we formulate another assumption. In general, the points that may cause problems are the ones where a_p and $a_{p'}$ coincide, i.e., the points of the set

$$\mathbf{D} := \{q \in \mathbf{Q} : a_p(q) = a_{p'}(q)\}.$$

On this set the polynomial $\hat{v}_{d'}$ should approximate $a_{p'}$ perfectly for a finite d' , meaning $\hat{v}_{d'}(q) = a_{p'}(q)$ for all $q \in \mathbf{D}$. Calling a solution $\hat{v}_{d'}$ near optimal if it satisfies $\int_{\mathbf{Q}} \hat{v}_{d'}(q) dq \leq \varrho_{d'} + \frac{1}{d'}$, we assume as follows:

Assumption 2. *There is a near optimal solution $\hat{v}_{d'}$ to SDP problem (2.3) for the polynomial p' with d' finite such that $\hat{v}_{d'}$ and $a_{p'}$ coincide on \mathbf{D} .*

Remark. A sufficient condition for a violation of Assumption 2 is the existence of a value of q for which $a_{p'}$ is not differentiable and $a_p(q) = a_{p'}(q)$. This is the case for the examples given above. Note also that they are of degenerate nature.

To face another issue, we denote the projection of $\hat{\mathbf{Z}}_{d'}$ onto the set \mathbf{Q} by $\pi_{\mathbf{Q}}(\hat{\mathbf{Z}}_{d'})$, i.e.,

$$\pi_{\mathbf{Q}}(\hat{\mathbf{Z}}_{d'}) = \{q \in \mathbf{Q} : \exists u, z \in \mathbb{R} : (q, u, z) \in \hat{\mathbf{Z}}_{d'}\}.$$

Since $\hat{v}_{d'}$ converges to a in L^1 , but not necessarily uniformly, it might have spikes or similar irregularities, meaning that the set $\mathbf{Q} \setminus \pi_{\mathbf{Q}}(\hat{\mathbf{Z}}_{d'})$ is not empty. However, the L^1 convergence of $\hat{v}_{d'}$, or more precisely the convergence in measure, implies that there is a subsequence $(\hat{v}_{d'_l})_{l \in \mathbb{N}}$ which converges to $a_{p'}$ almost uniformly (see, e.g., [Ash00, Theorem 2.5.3]). In other words, for all $\delta > 0$, there exists a set \mathbf{A}_δ in the Borel sigma algebra of \mathbf{Q} such that $\int_{\mathbf{A}_\delta} dq < \delta$ and $\hat{v}_{d'_l}$

converges uniformly on \mathbf{A}_δ^C to $a_{p'}$ when $l \rightarrow \infty$, where \mathbf{A}_δ^C is the set-theoretic complement of \mathbf{A}_δ in \mathbf{Q} . With this notation we have

$$\pi_{\mathbf{Q}}(\hat{\mathbf{Z}}_{d'}) \subseteq \mathbf{A}_\delta^C \subseteq \mathbf{Q}.$$

Lemma 2.7. *Let Assumption 2 hold. Then, for every $\delta > 0$ there exists a finite $d' \in \mathbb{N}$ and a set \mathbf{A}_δ in the Borel sigma algebra of \mathbf{Q} with $\int_{\mathbf{A}_\delta} dq < \delta$ such that $\hat{v}_{d'} \leq a_p$ on \mathbf{A}_δ^C .*

Proof. Fix $\delta > 0$. As discussed above there exists a set \mathbf{A}_δ in the Borel sigma algebra of \mathbf{Q} such that $\int_{\mathbf{A}_\delta} dq < \delta$ and $\hat{v}_{d'_l}$ converges uniformly to $a_{p'}$ on \mathbf{A}_δ^C as $l \rightarrow \infty$. Obviously we want

$$0 \leq a_p(q) - \hat{v}_{d'}(q) = a_p(q) - a_{p'}(q) + a_{p'}(q) - \hat{v}_{d'}(q) \quad (2.9)$$

for every $q \in \mathbf{A}_\delta^C \subseteq \mathbf{Q}$. By Theorem 2.6, we have $a_p(q) - a_{p'}(q) \geq 0$ for all $q \in \mathbf{Q}$. On the contrary, the difference $a_{p'}(q) - \hat{v}_{d'}(q)$ is negative by construction, but due to Theorem 2.2 we find a subsequence $\hat{v}_{d'_l}$ converging uniformly to $a_{p'}$ on \mathbf{A}_δ^C . Hence, there is a finite d'_{l^*} such that (2.9) is fulfilled for all $q \in \{q \in \mathbf{A}_\delta^C : a_p(q) > a_{p'}(q)\}$. Because of Assumption 2 there is also a finite d^* such that $a_{p'}(q) - \hat{v}_{d^*}(q)$ vanishes on $\{q \in \mathbf{A}_\delta^C : a_p(q) = a_{p'}(q)\} \subseteq \mathbf{D}$. Taking $d' = d'_{l^*} \geq d^*$ with $l^* \geq l^*$ completes the proof. \square

Remark. Choosing d' according to Lemma 2.7 implies $\pi_{\mathbf{Q}}(\hat{\mathbf{Z}}_{d'}) = \mathbf{A}_\delta^C$.

Under Assumption 1 and Assumption 2 and with an appropriate choice of d' (depending on δ) the solution to the following LP problem gives a lower approximation of the abscissa function a_p of p on the set $\mathbf{A}_\delta^C \subseteq \mathbf{Q}$:

$$\begin{aligned} \vartheta_{d'} &= \sup_{w \in \mathcal{C}(\mathbf{Q})} \int_{\pi_{\mathbf{Q}}(\hat{\mathbf{Z}}_{d'})} w(q) dq \\ \text{s.t. } & u - w(q) \geq 0 \text{ for all } (q, u, z) \in \hat{\mathbf{Z}}_{d'}. \end{aligned} \quad (2.10)$$

Remark. Note that under Assumption 1, LP problem (2.10) always provides a proper approximation for the abscissa a_p from below on $\pi_{\mathbf{Q}}(\hat{\mathbf{Z}}_{d'})$, but this might not be very useful, since for bad $\hat{v}_{d'}$ this set may have big holes or even be empty. To achieve suitable results on \mathbf{A}_δ^C we need Assumption 2 and an appropriate d' , meaning a sufficiently good $\hat{v}_{d'}$ ensuring $\pi_{\mathbf{Q}}(\hat{\mathbf{Z}}_{d'}) = \mathbf{A}_\delta^C$.

In analogy with (2.2), the dual LP problem reads

$$\begin{aligned} \vartheta_{d'}^* &= \inf_{\mu \in \mathcal{M}^+(\hat{\mathbf{Z}}_{d'})} \int_{\hat{\mathbf{Z}}_{d'}} u d\mu(q, u, z) \\ \text{s.t. } & \int_{\hat{\mathbf{Z}}_{d'}} q^\alpha d\mu = \int_{\pi_{\mathbf{Q}}(\hat{\mathbf{Z}}_{d'})} q^\alpha dq \text{ for all } \alpha \in \mathbb{N}^n \end{aligned} \quad (2.11)$$

with the notation of Section 2.2.1.

Lemma 2.8. *The infimum in LP problem (2.11) is attained, and there is no duality gap between LP problem (2.10) and LP problem (2.11), i.e., $\vartheta_{d'} = \vartheta_{d'}^*$.*

Since $\hat{\mathbf{Z}}_{d'}$ is a compact subset of \mathbf{Z} , we can mimic the proof of Lemma 2.1 in order to obtain a proof of Lemma 2.8.

Remark. As in Remark 2.2.2, the supremum in LP problem (2.10) is not attained for $\mathcal{C}(\mathbf{Q})$ or $\mathbb{R}[q]$, but it is attained for $\mathbb{R}[q]_d$ with d finite. To adjust the proof of Remark 2.2.2, set $M := \min_{q \in \mathbf{Q}} a(q) - N$ for an $N \in \mathbb{N}$ sufficiently large, and $R := \int_{\mathbf{Q}} (a(q) - M) dq$ as in Remark 2.3.1.

Finally, for d' as in Lemma 2.7 and $d_0 \geq d'$ sufficiently large, e.g. $d_0 \geq \lceil \ell/2 \rceil$, we can write an SDP hierarchy indexed by $d \in \mathbb{N}$, $d \geq d_0$:

$$\begin{aligned} \vartheta_{d',d} = & \sup_{w_d, \sigma_0, \sigma_j, \sigma_{\hat{v}}, \tau_{\mathfrak{R}}, \tau_{\mathfrak{S}}} \int_{\pi_{\mathbf{Q}}(\hat{\mathbf{Z}}_{d'})} w_d(q) dq & (2.12) \\ \text{s.t. } u - w_d(q) = & \sigma_0(q, u, z) + \sum_{j=1}^n \sigma_j(q, u, z)(1 - q_j^2) + \sigma_{\hat{v}}(q, u, z)(u - \hat{v}_{d'}(q)) \\ & + \tau_{\mathfrak{R}}(q, u, z)p_{\mathfrak{R}}(q, u, z) + \tau_{\mathfrak{S}}(q, u, z)p_{\mathfrak{S}}(q, u, z) \end{aligned}$$

for all $(q, u, z) \in \mathbb{R}^n \times \mathbb{R}^2$ and with $w_d \in \mathbb{R}[q]_{2d}$, $\sigma_0 \in \Sigma[q, u, z]_{2d}$, $\sigma_j \in \Sigma[q, u, z]_{2d-2}$ for $j = 1, \dots, n$, $\sigma_{\hat{v}} \in \Sigma[q, u, z]_{2(d-d')}$, and $\tau_{\mathfrak{R}}, \tau_{\mathfrak{S}} \in \mathbb{R}[q, u, z]_{2d-\ell}$.

As in Section 2.2.2, SDP problem (2.12) is a strengthening of LP problem (2.10), meaning $\vartheta_{d',d} \leq \vartheta_{d'}$. Besides, the archimedean quadratic module corresponding to the set \mathbf{Z} is contained in the quadratic module corresponding to $\hat{\mathbf{Z}}_{d'}$. Hence, the latter is also archimedean, i.e., $\lim_{d \rightarrow \infty} \vartheta_{d',d} = \vartheta_{d'} = \vartheta_{d'}^*$.

Remark. In practice one can assume that \mathbf{A}_{δ} is empty and substitute $\pi_{\mathbf{Q}}(\hat{\mathbf{Z}}_{d'})$ by \mathbf{Q} .

The associated sequence converges as follows:

Theorem 2.9. *Let Assumption 1 and Assumption 2 hold and let \mathbf{A}_{δ}^C and d' be as in Lemma 2.7. Let $w_d \in \mathbb{R}[q]_{2d}$ be a near optimal solution for SDP problem (2.12), i.e., $\int_{\mathbf{Q}} w_d(q) dq \geq \vartheta_{d,d'} - \frac{1}{d}$. Consider the associated sequence $(w_d)_{d \geq d_0} \subseteq L^1(\mathbf{Q})$. Then w_d is a valid lower bound of a_p on \mathbf{A}_{δ}^C and it converges to a_p in L^1 norm on \mathbf{A}_{δ}^C .*

The proof of this result is very similar to the proof of Theorem 2.2, so we omit it. Note that by Lemma 2.7 every feasible solution to SDP problem (2.12) is a valid lower bound of a_p on \mathbf{A}_{δ}^C and that we have $\pi_{\mathbf{Q}}(\hat{\mathbf{Z}}_{d'}) = \mathbf{A}_{\delta}^C$ due to our choice of d' . As for the proof of Theorem 2.5, the first part can be shortened, since for every $(q, u, z) \in \hat{\mathbf{Z}}_{d'}$ it holds that $u = a(q)$.

2.3.2.3 Examples

Example 2.3.7. As in Example 2.2.2 and Example 2.3.3 consider

$$p : s \mapsto p(q, s) = s^2 + 2qs + 1 - 2q.$$

Assumption 1 is naturally fulfilled, since p is of degree 2. In the same way, Assumption 2 is fulfilled, since $a_{p'}(q) = -q$ is polynomial. We have $\hat{\mathbf{Z}}_{d'} = \{(q, u, z) \in [-1, 1] \times \mathbb{R}^2 : u - \hat{v}_{d'}(q) \geq 0, u^2 - z^2 + 2qu + 1 - 2q = 2uz + 2qz = 0\}$ and the corresponding SDP problem (2.3) reads

$$\begin{aligned} \vartheta_{d',d} = & \sup_{w_d, \sigma_0, \sigma_1, \sigma_{\hat{v}}, \tau_{\mathfrak{R}}, \tau_{\mathfrak{S}}} \int_{-1}^1 w_d(q) dq \\ \text{s.t. } u - w_d(q) = & \sigma_0(q, u, z) + \sigma_1(q, u, z)(1 - q^2) + \sigma_{\hat{v}}(q, u, z)(u - \hat{v}_{d'}(q)) \\ & + \tau_{\mathfrak{R}}(q, u, z)(u^2 - z^2 + 2qu + 1 - 2q) + \tau_{\mathfrak{S}}(q, u, z)(2uz + 2qz) \end{aligned}$$

for all $(q, u, z) \in \mathbb{R}^3$ and with $w_d \in \mathbb{R}[q]_{2d}$, $\sigma_0 \in \Sigma[q, u, z]_{2d}$, $\sigma_1 \in \Sigma[q, u, z]_{2d-2}$, $\sigma_{\hat{v}} \in \Sigma[q, u, z]_{2(d-d')}$, and $\tau_{\mathfrak{R}}, \tau_{\mathfrak{S}} \in \mathbb{R}[q, u, z]_{2d-2}$. Due to the simplicity of $a_{p'}$ it suffices to choose $d' = 1$, meaning $\deg \hat{v}_{d'} = 2$. We see the degrees 6 ($d = 3$) and 12 ($d = 6$) polynomial lower approximations in Figure 2.10. They are both computed in less than a second.

Example 2.3.8. As in Example 2.2.3 and Example 2.3.4 consider

$$p : s \mapsto p(q, s) = s^3 + \frac{1}{2}s^2 + q^2s + (q - \frac{1}{2})q(q + \frac{1}{2}).$$

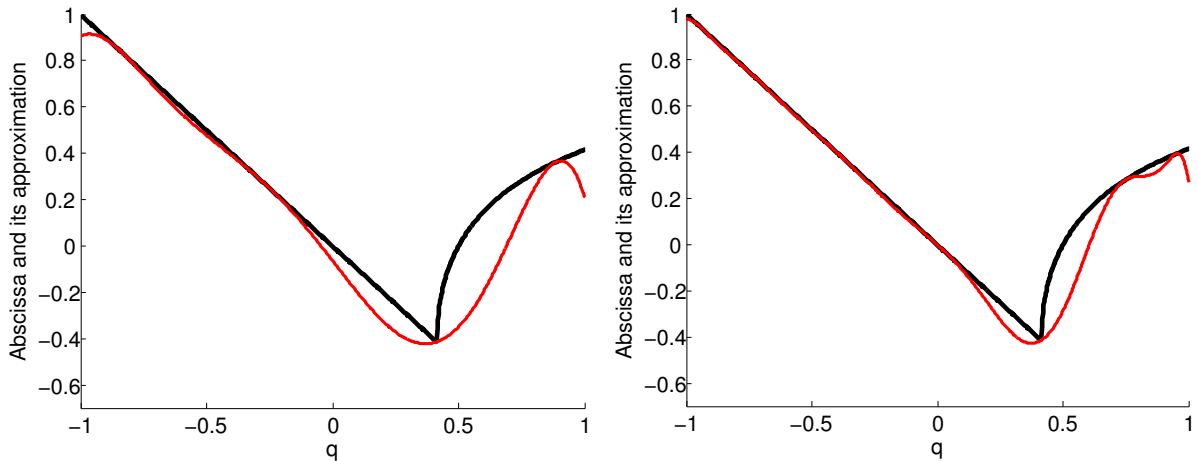


Figure 2.10 – Abscissa $a_p(q)$ (black) and its polynomial lower GL approximations $w_d(q)$ of degree $2d = 6$ (red, left) and $2d = 12$ (red, right) for Example 2.3.7. The quality of the approximation deteriorates near the minimum, where the abscissa is not Lipschitz. Compare with Figure 2.1 and Figure 2.7 .

The abscissa $a_{p'}$ of p' is not differentiable in two points, hence it is not a polynomial and it cannot be described perfectly by $\hat{v}_{d'}$ for finite d' . Let us choose $d' = 4$ and $d = 3$ (resp., $d = 6$). We observe in Figure 2.11 that w_3 (resp., w_6) is not everywhere a valid lower bound. Indeed, the set $\mathbf{D} = \{q \in \mathbf{Q} : a_p(q) = a_{p'}(q)\}$ contains three points and for two of these (near $q = -0.5$ and $q = 0$), the approximation \hat{v}_4 of degree 8 is not tight enough to ensure $\pi_{\mathbf{Q}}(\mathbf{Z}_{r,4}) = \mathbf{Q}$. Consequently, Assumption 2 is violated. The CPU time for the degree 12 solution is less than 4 seconds.

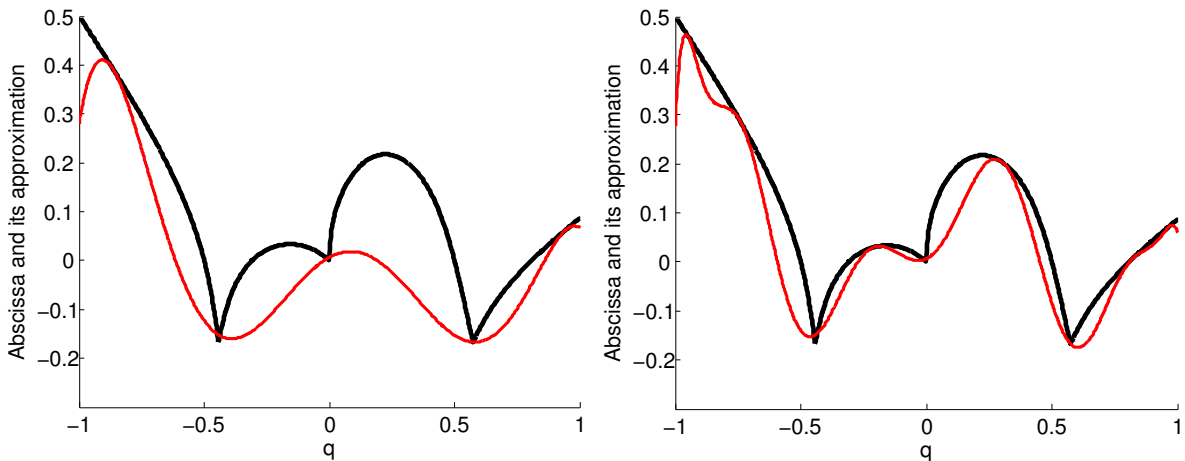


Figure 2.11 – Abscissa $a_p(q)$ (black) and its polynomial lower GL approximations $w_d(q)$ of degree $2d = 6$ (red, left) and $2d = 12$ (red, right) for Example 2.3.8. We observe that the approximations are not valid near $q = -0.5$ and $q = 0$, as Assumption 2 is violated.

Example 2.3.9. In order to discuss another example for which \mathbf{D} is a non-empty interval, consider the polynomial

$$p : s \mapsto p(q, s) = s^2 + (20q^2 - 1)s + q + \frac{1}{2}.$$

Here $a_{p'}(q) = -10q^2 + \frac{1}{2}$ is a quadratic polynomial. Thus, Assumption 2 is fulfilled, in particular $\hat{v}_1 = a_{p'}$, and the lower approximations are valid; see Figure 2.12.

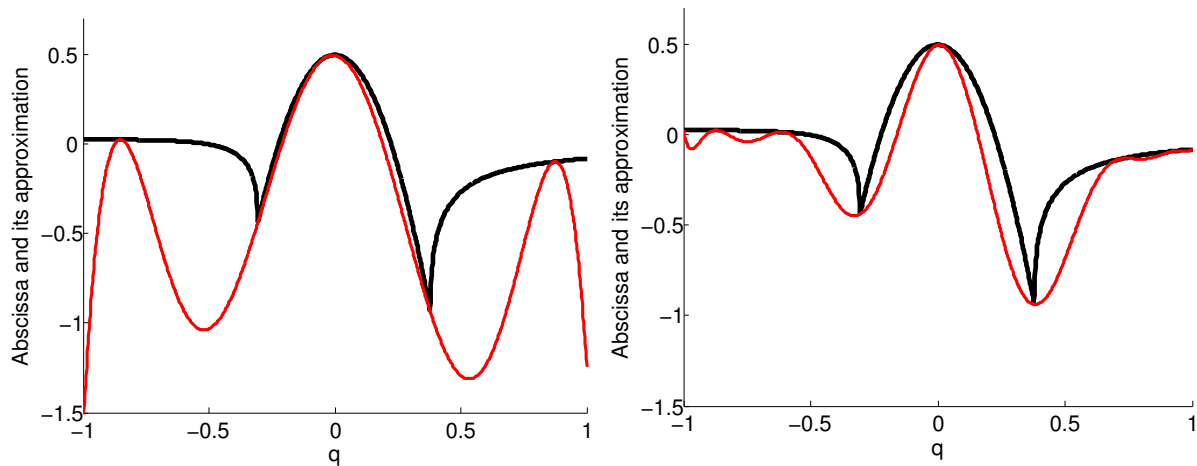


Figure 2.12 – Abscissa $a_p(q)$ (black) and its polynomial lower GL approximations $w_d(q)$ of degree $2d = 6$ (red, left) and $2d = 12$ (red, right) for Example 2.3.9 ($d' = 1$).

Example 2.3.10. As in Example 2.2.4 and Example 2.3.5 consider the polynomial

$$p : s \mapsto p(q, s) = s^3 + (q_1 + \frac{3}{2})s^2 + q_1^2 s + q_1 q_2.$$

We have $\hat{\mathbf{Z}}_{d'} := \{(q, u, z) \in \mathbf{Z} : u - \hat{v}_{d'}(q) \geq 0\}$ with \mathbf{Z} given in Example 2.2.4. In Figure 2.13 we see the outer approximations of degree $2d = 8$ (resp., $2d = 12$) of the stability region obtained for the choice $d' = 4$. A careful examination reveals that Assumption 2 is slightly violated here, yet this has no effect on the validity of the zero sublevel set approximation. Computing the degree 12 approximation takes nearly 10 minutes. For comparison, the computation of the degree 10 solution takes less than 2 minutes and the degree 8 solution around 20 seconds. Therefore, it is much faster than the computation via the elementary symmetric functions.

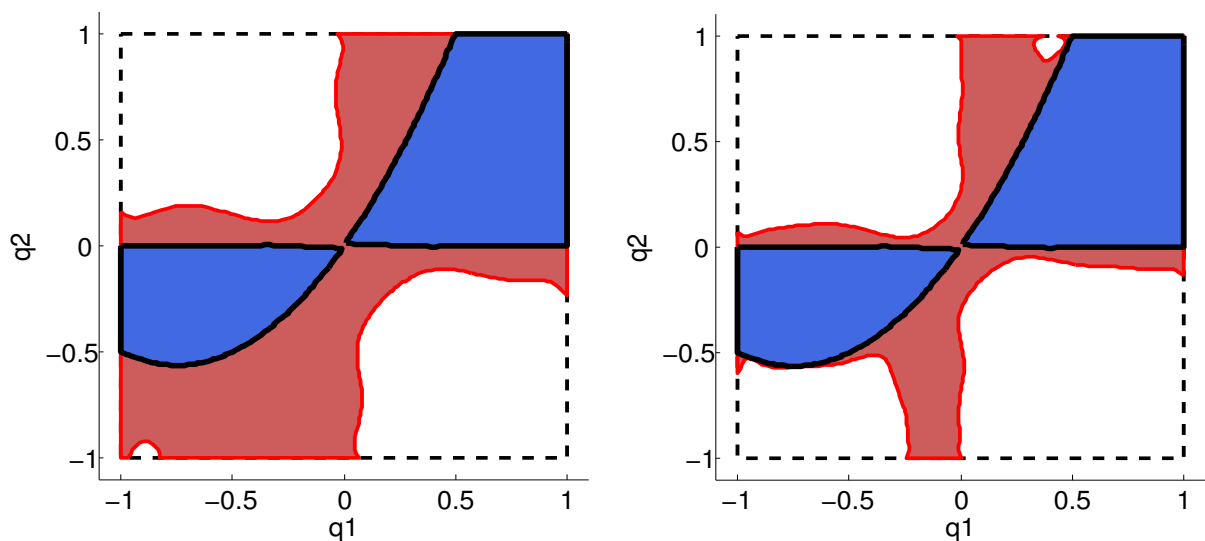


Figure 2.13 – Stability region (blue region) and its degree $2d = 8$ outer GL approximation (red region, left) and degree $2d = 12$ outer GL approximation (red region, right) for Example 2.3.10 ($d' = 4$). Compare with Figure 2.9.

2.4 Concluding remarks

As mentioned in the introduction, this thesis consists of investigations of the Lasserre hierarchy by approximating potentially complicated objects with simple objects, namely, polynomials of given degrees. The complicated object of interest here was the polynomial abscissa, which has low regularity, while being ubiquitous in linear systems control.

Note that we focused exclusively on the polynomial abscissa, but our techniques readily extend to the polynomial radius (defined as the maximum modulus of the roots) or to any semialgebraic function of the polynomial roots. By semialgebraic function, we mean any function whose graph can be described by finitely many intersections and unions of polynomial sublevel sets or level sets (see, e.g., [BCR98]).

In Section 2.2 we described how to construct polynomial upper approximations to the abscissa with guarantees of L^1 convergence (or equivalently almost uniform convergence) on compact sets via the Lasserre hierarchy. Constructing polynomial lower approximations with similar convergence guarantees has proved to be much more challenging. We proposed a first approach in Section 2.3.1 using elementary symmetric functions which is quite general but also computationally challenging due to the introduction of many lifting variables. This motivated the study of a second approach in Section 2.3.2 using the Gauß-Lucas theorem, which is less computationally demanding but unfortunately much more involved and subject to working assumptions.

As illustrated by our numerical examples, a shortcoming of our methods is that they can be computationally demanding and not applicable when the degree of the polynomial p and/or the number of parameters in q is large. Moreover, it is unknown which level in the semidefinite programming hierarchy can guarantee an a priori given precision level on the abscissa. In terms of complexity, whereas interior-point algorithms can provably solve semidefinite programs at given accuracy in polynomial time, it turns out that the number of variables is exponential in the degree of the polynomial p and the number of parameters q , and only the first levels of the semidefinite programming hierarchy can be solved in practice in a reasonable amount of time.

Uniform approximations

An interesting theoretical question that deserves careful investigation is whether our L^1 convergence guarantees can be strengthened to L^∞ , i.e., to uniform convergence, since we know that the polynomial abscissa is continuous and hence that it can be uniformly approximated by polynomials on compact sets.

A natural idea to achieve this would be to replace the L^1 norm of $v(q) - a(q)$ in problem (2.1) by its L^∞ norm. However, this would not lead to a useful result, since the infimum $\|v(q) - a(q)\|_\infty = \sup_{q \in \mathbf{Q}} |v(q) - a(q)|$ with respect to the constraint $v(q) - u \geq 0$ for all $(q, u, z) \in \mathbf{Z}$ would be the maximal distance between two real parts of the roots as a consequence of the presence of the real part of all roots in the set \mathbf{Z} . Thus, in general the optimal value is not zero, and therefore, v would not be a proper approximation to the abscissa.

This problem does not occur when using the formulation via elementary symmetric functions, since there we can explicitly access the real part of the root corresponding to the abscissa. Indeed, we obtain uniform approximations when optimizing the L^∞ norm of $v(q) - a(q)$ instead of its L^1 norm in problem (2.5).

As is typical for best uniform approximations, the solution oscillates by the Chebyshev Equioscillation theorem [Dav63, Theorem 7.6.2]. Due to this and the large amount of variables, this formulation has practical drawbacks. The approach can be improved significantly, when considering rational approximations instead of polynomial approximations. Rational minimax

approximations are an active field of research [FNTB17] and may be used for uniform one-sided approximations of the abscissa.

Approximate optimal designs for multivariate polynomial regression

Contents

3.1 Preliminaries	42
3.1.1 Polynomial optimal design	42
3.1.2 Approximations of the moment cone	43
3.2 Approximate Optimal Design	43
3.2.1 Relaxing the integrality constraints	44
3.2.2 Moment formulation	44
3.3 The ideal problem on moments and its approximation	45
3.3.1 The ideal problem on moments	45
3.3.2 Christoffel polynomials	49
3.3.3 The SDP relaxation scheme	50
3.3.4 Asymptotics	53
3.3.5 General regression polynomial bases	54
3.4 Recovering the measure	55
3.4.1 Via the Nie method	55
3.4.2 Via Christoffel-like polynomials	56
3.4.3 Calculating the corresponding weights	56
3.5 Examples	56
3.5.1 Univariate unit interval	57
3.5.2 Wynn's polygon	58
3.5.3 The 3-dimensional unit sphere	60
3.6 Concluding remark	61

As a second application of the hierarchy of lower bounds, we consider the optimal design problem. It arises in statistics when regarding regression problems, where the experimenter wants to estimate parameters up to a certain precision. For this, it is necessary to repeat the experiment multiple times with different input data. Optimal designs are rules to decide which input data need to be chosen in order to reduce the necessary runs of the experiment. The problem shall be rendered more precisely below. Before, we shortly explain the objective of this chapter in more detail.

We introduce a general method to compute so-called approximate optimal designs—in the sense of Kiefer's ϕ_q -criteria—on a large variety of design spaces, namely semialgebraic sets. We apply the Lasserre hierarchy to solve numerically and approximately the optimal design problem. The theoretical guarantees are given by Theorem 3.3 and Theorem 3.4. These theorems demonstrate the convergence of our procedure towards the approximate optimal designs as the order of the hierarchy increases. Furthermore, they give a characterization of finite order

convergence of the hierarchy. In particular, our method recovers the optimal design when finite convergence of the hierarchy occurs. To recover the geometry of the design we use SDP duality theory and Christoffel-like polynomials (to be defined in Remark 3.3.1) involved in the optimality conditions.

We have run several numerical experiments for which finite convergence holds, leading to a surprisingly fast and reliable method to compute optimal designs. As illustrated by our examples, in polynomial regression models with degree order higher than one we obtain designs with points in the interior of the domain. This contrasts with the classical use of ellipsoids for linear regressions where points are obtained on the boundary.

The chapter is based on [dCGH⁺17] and it is organized as follows. After a short introduction to convex design theory and a paragraph on the state of the art, we consider in Section 3.1 polynomial optimal designs and approximations of the moment cone which slightly differ from those considered in Chapter 1. In Section 3.2 we introduce approximate optimal designs and propose a two step procedure to solve the approximate design problem. Solving the first step is subject to Section 3.3. There, we find a sequence of moments associated with the optimal design measure. Recovering this measure (step two of the procedure) is discussed in Section 3.4. We finish the chapter with some illustrating examples and a short conclusion.

Convex design theory

The optimum experimental designs are computational and theoretical objects that aim to minimize the uncertainty contained in the best linear unbiased estimators in regression problems. In this frame, the experimenter models the responses $\mathbf{z}_1, \dots, \mathbf{z}_N$ of a random *experiment* whose inputs are represented by a vector $\xi_i \in \mathbb{R}^n$ with respect to known *regression functions* f_1, \dots, f_p , namely,

$$\mathbf{z}_i = \sum_{j=1}^p \theta_j f_j(\xi_i) + \varepsilon_i, \quad i = 1, \dots, N,$$

where $\theta_1, \dots, \theta_p$ are unknown parameters that the experimenter wants to estimate, ε_i is some noise, and the inputs ξ_i are chosen by the experimenter in a *design space* $\mathcal{X} \subseteq \mathbb{R}^n$. Assume that the inputs ξ_i for $i = 1, \dots, N$ are chosen within a set of distinct points $\mathbf{x}_1, \dots, \mathbf{x}_\ell$ with $\ell \leq N$, and let n_k denote the number of times the particular point \mathbf{x}_k occurs among ξ_1, \dots, ξ_N . This would be summarized by

$$\zeta := \begin{pmatrix} \mathbf{x}_1 & \cdots & \mathbf{x}_\ell \\ \frac{n_1}{N} & \cdots & \frac{n_\ell}{N} \end{pmatrix}, \quad (3.1)$$

whose first row gives distinct points in the design space \mathcal{X} where the input parameters have to be taken and the second row indicates to the experimenter which proportion of experiments (frequencies) have to be done at these points. The goal of the design of experiment theory is then to assess which input parameters and frequencies the experimenter has to consider. For a given ζ the standard analysis of the Gaussian linear model shows that the minimal covariance matrix (with respect to Löwner ordering) of unbiased estimators can be expressed in terms of the Moore-Penrose pseudo-inverse of the *information matrix* which is defined by

$$\mathcal{J}(\zeta) := \sum_{i=1}^{\ell} w_i \mathfrak{F}(\mathbf{x}_i) \mathfrak{F}^T(\mathbf{x}_i), \quad (3.2)$$

where $\mathfrak{F} := (f_1, \dots, f_p)$ is the column vector of regression functions and $w_i := \frac{n_i}{N}$ is the weight corresponding to the point \mathbf{x}_i . One major aspect of designs in experiment theory seeks to maximize a suitable functional of the information matrix over the set of all possible ζ . Notice

that the Löwner ordering is partial and, in general, there is no greatest element among all possible information matrices $\mathfrak{J}(\zeta)$. The standard approach is then to consider some statistical criteria, namely *Kiefer's ϕ_q -criteria* [Kie74], in order to describe and construct the “*optimal designs*” with respect to those criteria. Observe that the information matrix belongs to \mathbb{S}_p^+ , the space of symmetric positive semidefinite matrices of size p . For all $q \in [-\infty, 1]$ define the function

$$\phi_q : \mathbb{S}_p^+ \rightarrow \mathbb{R}, \quad M \mapsto \phi_q(M),$$

where for positive definite matrices M

$$\phi_q(M) := \begin{cases} (\frac{1}{p}\text{trace}(M^q))^{1/q} & \text{if } q \neq -\infty, 0 \\ \det(M)^{1/p} & \text{if } q = 0 \\ \lambda_{\min}(M) & \text{if } q = -\infty \end{cases}$$

and for singular positive semidefinite matrices M

$$\phi_q(M) := \begin{cases} (\frac{1}{p}\text{trace}(M^q))^{1/q} & \text{if } q \in (0, 1] \\ 0 & \text{if } q \in [-\infty, 0]. \end{cases}$$

We recall that $\text{trace}(M)$, $\det(M)$ and $\lambda_{\min}(M)$ denote the trace, determinant and least eigenvalue of the symmetric positive semidefinite matrix M , respectively. Those criteria are meant to be real valued, positively homogeneous, non-constant, upper semi-continuous, isotonic (with respect to the Löwner ordering) and concave functions.

In particular, we search for solutions ζ^* to the following optimization problem

$$\max \phi_q(\mathfrak{J}(\zeta)), \tag{3.3}$$

where the maximum is taken over all ζ of the form (3.1). Standard criteria are given by the parameters $q = 0, -1, 1, -\infty$ and are referred to *D*-, *A*-, *T*- or *E*-optimum designs, respectively.

State of the art

Optimal design is at the heart of statistical planning for inference in the linear model, see for example [BHH78]. While the case of discrete input factors is generally tackled by algebraic and combinatoric arguments (e.g., [Bai08]), the one of continuous input factors often leads to an optimization problem. In general, the continuous factors are generated by a vector \mathfrak{F} of linearly independent regular functions on the *design space* \mathcal{X} .

One way to handle the problem is to focus only on \mathcal{X} ignoring the function \mathfrak{F} and to try to draw the design points *filling* the set \mathcal{X} in the best way. This is generally done by optimizing a cost function on \mathcal{X}^N that reflects the way the design points are positioned between each other and/or how they fill the space. Generic examples are the so-called MaxMin or MinMax criteria (see for example [PM12, WPN97]) and the minimum discrepancy designs; see for example [LQX05].

Another point of view—which is the one developed here—relies on the *maximization* of the information matrix. Of course, as explained before, the set of information matrices is a partially ordered set with respect to the Löwner ordering, and so the optimization cannot be performed directly on this matrix but on a real function on it. A pioneer paper adopting this point of view is the one of Elfving [Elf52]. In the early 60's, in a series of papers, Kiefer and Wolfowitz throw new light on this kind of methods for experimental design by introducing the equivalence principle and proposing algorithms to solve the optimization problem for some cases; see [Kie74] and references therein. Following the early works of Karlin and Studden [KS66a, KS66b], the case of polynomial regression on a compact interval on \mathbb{R} has been widely studied. In this framework,

the theory is almost complete and many things can be said about the optimal solutions for the design problem; see for instance [DS93]. Roughly speaking, the optimal design points are related to the zeros of orthogonal polynomials built on an equilibrium measure. We refer to the inspiring book of Dette and Studden [DS97] and references therein for a complete overview on the subject.

In the one dimensional frame, other systems of functions \mathfrak{F} —trigonometric functions or T -systems, see [KN77] for a definition—are studied in the same way in [DS97], [LS85] and [IS01] (see also the recent paper [KW17] for another perspective on the subject). In the multidimensional case, even for polynomial systems, very few cases of explicit solutions are known. Using tensoring arguments the case of a rectangle is treated in [DS97, Sch12]. Particular models of degree two are studied in [DG14, PW94].

Apart from these particular cases, the construction of the optimal design relies on numerical optimization procedures. The case of the determinant—which corresponds to the choice $q = 0$, i.e., the D -optimality—is studied for example in [Wyn70] and [VBW98]. Another criterion based on matrix conditioning—referred to as G -optimality—is developed in [MYZ15]. In the latter paper, the construction of an optimal design is performed in two steps. In the first step one only deals with an optimization problem on the set of all possible information matrices, while in the second step, one wishes to identify a possible probability distribution associated with the optimal information matrix. General optimization algorithms are discussed in [Fed10] and [ADT07]. A general optimization frame on measure sets including gradient descent methods is considered in [MZ04]. In the frame of fixed given support points, efficient SDP based algorithms are proposed and studied in [Sag11] and [SH15]. Let us mention the paper [VBW98] which is one of the original motivations to develop SDP solvers, especially for Max-Det-Problems—corresponding to D -optimal design—and the so-called problem of analytical centering.

3.1 Preliminaries

3.1.1 Polynomial optimal design

We will restrict our attention to *polynomial* optimal design problems, i.e., we consider polynomial regression functions and semialgebraic design spaces. To be more precise, we assume that the regression functions are multivariate polynomials of degree smaller than d , that is, $\mathfrak{F} = (f_1, \dots, f_p) \in (\mathbb{R}[x]_d)^p$. Moreover, we consider that the design space $\mathcal{X} \subseteq \mathbb{R}^n$ is a given basic closed semialgebraic set

$$\mathcal{X} := \{\mathbf{x} \in \mathbb{R}^n : g_j(\mathbf{x}) \geq 0, j = 1, \dots, m\} \quad (3.4)$$

for given polynomials $g_j \in \mathbb{R}[x]$, $j = 1, \dots, m$. Assume one of the polynomial inequalities $g_j(x) \geq 0$ is of the form $R - \sum_{i=1}^n x_i^2 \geq 0$ for a sufficiently large constant R . Note that this implies that \mathcal{X} is compact.

We remark that the set \mathcal{X} is of the same form as the set \mathbf{K} defined in (1.1). However, as design spaces are traditionally called \mathcal{X} , we follow this notation.

Notice further that these assumptions cover a large class of problems in optimal design theory, see for instance [DS97, Chapter 5]. In particular, observe that the design space \mathcal{X} defined by (3.4) is not necessarily convex.

Since we consider polynomial regression functions $\mathfrak{F} = (f_1, \dots, f_p)$, there exists a unique matrix \mathfrak{A} of size $p \times \binom{n+d}{n}$ such that

$$\mathfrak{F}(x) = \mathfrak{A} \mathbf{v}_d(x), \quad (3.5)$$

where $\mathbf{v}_d(x)$ is the vector of monomials up to order d as defined in Section 1.3.

3.1.2 Approximations of the moment cone

We recall the notation $s(d) = \binom{n+d}{n}$ and the definition of the truncated moment cone $\mathcal{M}_d(\mathcal{X})$ defined in (1.7),

$$\mathcal{M}_d(\mathcal{X}) := \left\{ \mathbf{y} \in \mathbb{R}^{s(d)} : \exists \mu \in \mathcal{M}_+(\mathcal{X}), y_\alpha = \int_{\mathcal{X}} x^\alpha d\mu \ \forall |\alpha| \leq d \right\}.$$

In Chapter 1 we approximated the moment cone $\mathcal{M}(\mathcal{X})$ by the semidefinite representable cones $\mathcal{M}_{2d}(g_1, \dots, g_m)$. Now, we consider projections of these.

As in Chapter 1, let $d_j := \lceil \deg g_j / 2 \rceil$, $j = 1, \dots, m$, denote half the degree of the g_j . For $\delta \in \mathbb{N}$ we approximate the truncated moment cone $\mathcal{M}_{2d}(\mathcal{X})$ by the cones

$$\begin{aligned} \mathcal{M}_{2(d+\delta)}^{2d}(g_1, \dots, g_m) &:= \left\{ \mathbf{y} \in \mathbb{R}^{s(2d)} : \exists \mathbf{y}_\delta \in \mathbb{R}^{s(2(d+\delta))}, \mathbf{y} = (y_{\delta, \alpha})_{|\alpha| \leq 2d}, \right. \\ &\quad \left. \mathbf{M}_{d+\delta}(\mathbf{y}_\delta) \succcurlyeq 0, \mathbf{M}_{d+\delta-d_j}(g_j \mathbf{y}_\delta) \succcurlyeq 0, j = 1, \dots, m \right\}. \end{aligned} \quad (3.6)$$

These are projections of the cones $\mathcal{M}_{2(d+\delta)}(g_1, \dots, g_m)$, and evidently they are semidefinite representable. Since $\mathcal{M}_{2d}(\mathcal{X})$ is contained in every $\mathcal{M}_{2(d+\delta)}^{2d}(g_1, \dots, g_m)$, $\delta \in \mathbb{N}$, the latter are outer approximations of the truncated moment cone. Moreover, they form a nested sequence, so we can build the hierarchy

$$\mathcal{M}_{2d}(\mathcal{X}) \subseteq \dots \subseteq \mathcal{M}_{2(d+2)}^{2d}(g_1, \dots, g_m) \subseteq \mathcal{M}_{2(d+1)}^{2d}(g_1, \dots, g_m) \subseteq \mathcal{M}_{2d}^{2d}(g_1, \dots, g_m). \quad (3.7)$$

This hierarchy actually converges, meaning $\mathcal{M}_{2d}(\mathcal{X}) = \overline{\bigcap_{\delta=0}^{\infty} \mathcal{M}_{2(d+\delta)}^{2d}(g_1, \dots, g_m)}$, where $\overline{\mathbf{A}}$ denotes the topological closure of the set \mathbf{A} . Note in passing that $\mathcal{M}_{2d}^{2d}(g_1, \dots, g_m) = \mathcal{M}_{2d}(g_1, \dots, g_m)$.

The topological dual of $\mathcal{M}_{2(d+\delta)}^{2d}(g_1, \dots, g_m)$ is given by the following quadratic module:

$$\begin{aligned} \mathcal{Q}_{2(d+\delta)}^{2d}(g_1, \dots, g_m) &:= \left\{ h = \sigma_0 + \sum_{j=1}^m g_j \sigma_j : \deg(h) \leq 2d, \right. \\ &\quad \left. \sigma_0 \in \Sigma[x]_{2(d+\delta)}, \sigma_j \in \Sigma[x]_{2(d+\delta-d_j)}, j = 1, \dots, m \right\}. \end{aligned} \quad (3.8)$$

In line with (3.7) we have the inclusions

$$\mathcal{P}_{2d}(\mathcal{X}) \supseteq \dots \supseteq \mathcal{Q}_{2(d+2)}^{2d}(g_1, \dots, g_m) \supseteq \mathcal{Q}_{2(d+1)}^{2d}(g_1, \dots, g_m) \supseteq \mathcal{Q}_{2d}^{2d}(g_1, \dots, g_m),$$

and $\mathcal{Q}_{2d}^{2d}(g_1, \dots, g_m) = \mathcal{Q}_{2d}(g_1, \dots, g_m)$.

3.2 Approximate Optimal Design

For $i = 1, \dots, p$ write $f_i(x) := \sum_{|\alpha| \leq d} a_{i,\alpha} x^\alpha$ for appropriate $a_{i,\alpha} \in \mathbb{R}$. Then $\mathfrak{A} = (a_{i,\alpha})$ where \mathfrak{A} is defined by (3.5). For $\mu \in \mathcal{M}_+(\mathcal{X})$ with moment sequence \mathbf{y} define the information matrix

$$\mathfrak{I}_d(\mathbf{y}) := \left(\int_{\mathcal{X}} f_i f_j d\mu \right)_{1 \leq i, j \leq p} = \left(\sum_{|\alpha|, |\beta| \leq d} a_{i,\alpha} a_{j,\beta} y_{\alpha+\beta} \right)_{1 \leq i, j \leq p} = \sum_{|\gamma| \leq 2d} A_\gamma y_\gamma,$$

where we have set $A_\gamma := \left(\sum_{\alpha+\beta=\gamma} a_{i,\alpha} a_{j,\beta} \right)_{1 \leq i, j \leq p}$ for $|\gamma| \leq 2d$. Observe that

$$\mathfrak{I}_d(\mathbf{y}) = \mathfrak{A} M_d(\mathbf{y}) \mathfrak{A}^T. \quad (3.9)$$

If \mathbf{y} is the moment sequence of $\mu = \sum_{i=1}^{\ell} w_i \delta_{\mathbf{x}_i}$, where $\delta_{\mathbf{x}}$ denotes the Dirac measure at the point $\mathbf{x} \in \mathcal{X}$ and the w_i are again the weights corresponding to the points \mathbf{x}_i , then $\mathfrak{I}_d(\mathbf{y}) = \sum_{i=1}^{\ell} w_i \mathfrak{F}(\mathbf{x}_i) \mathfrak{F}^T(\mathbf{x}_i)$ as in (3.2).

Consider the optimization problem

$$\begin{aligned} & \max \phi_q(M) \\ & \text{s.t. } M = \sum_{|\gamma| \leq 2d} A_\gamma y_\gamma \succcurlyeq 0, \quad y_\gamma = \sum_{i=1}^{\ell} \frac{n_i}{N} \mathbf{x}_i^\gamma, \quad \sum_{i=1}^{\ell} n_i = N, \\ & \mathbf{x}_i \in \mathcal{X}, \quad n_i \in \mathbb{N}, \quad i = 1, \dots, \ell, \end{aligned} \quad (3.10)$$

where the maximization is with respect to \mathbf{x}_i and n_i , $i = 1, \dots, \ell$, subject to the constraint that the information matrix M is positive semidefinite. By construction, it is equivalent to the original design problem (3.3). In this form, problem (3.10) is difficult because of the integrality constraints on the n_i and the nonlinear relation between \mathbf{y} , \mathbf{x}_i and n_i . We will address these difficulties in the sequel by first relaxing the integrality constraints.

3.2.1 Relaxing the integrality constraints

In problem (3.10), the set of admissible frequencies $w_i = \frac{n_i}{N}$ is discrete, which makes it a potentially difficult combinatorial optimization problem. A popular solution is then to consider “*approximate*” designs defined by

$$\zeta := \begin{pmatrix} \mathbf{x}_1 & \cdots & \mathbf{x}_\ell \\ w_1 & \cdots & w_\ell \end{pmatrix}, \quad (3.11)$$

where the frequencies w_i belong to the unit simplex $\mathcal{W} := \{\mathbf{w} \in \mathbb{R}^\ell : 0 \leq w_i \leq 1, \sum_{i=1}^{\ell} w_i = 1\}$. Accordingly, any solution to problem (3.3), where the maximum is taken over all matrices of type (3.11), is called “*approximate optimal design*”. This yields the following relaxation of problem (3.10):

$$\begin{aligned} & \max \phi_q(M) \\ & \text{s.t. } M = \sum_{|\gamma| \leq 2d} A_\gamma y_\gamma \succcurlyeq 0, \quad y_\gamma = \sum_{i=1}^{\ell} w_i \mathbf{x}_i^\gamma, \\ & \mathbf{x}_i \in \mathcal{X}, \quad \mathbf{w} \in \mathcal{W}, \end{aligned} \quad (3.12)$$

where the maximization is with respect to \mathbf{x}_i and w_i , $i = 1, \dots, \ell$, subject to the constraint that the information matrix M is positive semidefinite. In this problem the nonlinear relation between \mathbf{y} , \mathbf{x}_i and w_i is still an issue.

3.2.2 Moment formulation

Let us introduce a two-step-procedure to solve the approximate optimal design problem (3.12). For this, we first reformulate our problem again.

By Carathéodory’s theorem, the subset of moment sequences in the truncated moment cone $\mathcal{M}_{2d}(\mathcal{X})$, which we recalled in Section 3.1.2, with $y_0 = 1$, is exactly the set

$$\left\{ \mathbf{y} \in \mathcal{M}_{2d}(\mathcal{X}) : y_0 = 1 \right\} = \left\{ \mathbf{y} \in \mathbb{R}^{s(2d)} : y_\alpha = \int_{\mathcal{X}} x^\alpha d\mu \quad \forall |\alpha| \leq 2d, \right. \\ \left. \mu = \sum_{i=1}^{\ell} w_i \delta_{\mathbf{x}_i}, \quad \mathbf{x}_i \in \mathcal{X}, \quad \mathbf{w} \in \mathcal{W} \right\},$$

where $\ell \leq s(2d)$, see the so-called Tchakaloff theorem [Las10, Theorem B12]. Hence, problem (3.12) is equivalent to

$$\begin{aligned} \max \quad & \phi_q(M) \\ \text{s.t.} \quad & M = \sum_{|\gamma| \leq 2d} A_\gamma y_\gamma \succcurlyeq 0, \\ & y \in \mathcal{M}_{2d}(\mathcal{X}), \quad y_0 = 1, \end{aligned} \tag{3.13}$$

where the maximization is now with respect to the sequence \mathbf{y} . The moment problem (3.13) is finite-dimensional and convex, yet the constraint $\mathbf{y} \in \mathcal{M}_{2d}(\mathcal{X})$ is difficult to handle. We will show that by approximating the truncated moment cone $\mathcal{M}_{2d}(\mathcal{X})$ by the nested sequence of semidefinite representable cones introduced in (3.7), we obtain a hierarchy of semidefinite programs converging to the optimal solution of problem (3.13). Since semidefinite programming problems can be solved efficiently, we can compute a numerical solution to problem (3.11).

This describes step one of our procedure. The result of it is a sequence \mathbf{y}^* of moments. Consequently, in a second step, we need to find a representing atomic measure μ^* of \mathbf{y}^* in order to identify the approximate optimal design ζ^* .

The next section is dedicated to consider step one in more detail. Step 2 is then the subject of Section 3.4.

3.3 The ideal problem on moments and its approximation

For notational simplicity, let us use the standard monomial basis of $\mathbb{R}[x]_d$ for the regression functions, meaning $\mathfrak{F} = (f_1, \dots, f_p) := (x^\alpha)_{|\alpha| \leq d}$, i.e., $p = s(d)$. This case corresponds to $\mathfrak{A} = \text{Id}$ the identity matrix in (3.5). Note that this is not a restriction, since one can get the results for other choices of \mathfrak{F} by simply performing a change of basis. Indeed, in view of (3.9) one should substitute $M_d(\mathbf{y})$ by $\mathfrak{A}M_d(\mathbf{y})\mathfrak{A}^T$ to get the statement of our results in whole generality; see Section 3.3.5 for a statement of the results in this case. Different polynomial bases can be considered and, for instance, one may consult the standard framework described in the book [DS97, Chapter 5.8].

For the sake of conciseness, we do not expose the notion of incomplete q -way m -th degree polynomial regression here, but the reader may note that the strategy developed in this chapter can handle such a framework.

Before stating the main results, we recall the gradients of the Kiefer's ϕ_q -criteria in Table 3.1.

3.3.1 The ideal problem on moments

The ideal formulation (3.13) of our approximate optimal design problem reads

$$\begin{aligned} \rho = \max_{\mathbf{y}} \quad & \phi_q(M_d(\mathbf{y})) \\ \text{s.t.} \quad & \mathbf{y} \in \mathcal{M}_{2d}(\mathcal{X}), \quad y_0 = 1. \end{aligned} \tag{3.14}$$

Name	q	$\phi_q(M)$	$\nabla\phi_q(M)$
D-optimal	0	$\det(M)^{1/p}$	$\det(M)^{1/p}M^{-1/p}$
A-optimal	-1	$p(\text{trace}(M^{-1}))^{-1}$	$p(\text{trace}(M^{-1})M)^{-2}$
T-optimal	1	$\frac{1}{p} \text{trace}(M)$	$\frac{1}{p} \text{Id}$
E-optimal	$-\infty$	$\lambda_{\min}(M)$	$\Pi_{\min}(M)$
generic case	$q \neq 0, -\infty$	$\left[\frac{\text{trace}(M^q)}{p}\right]^{1/q}$	$\left[\frac{\text{trace}(M^q)}{p}\right]^{1/q-1} \frac{M^{q-1}}{p}$

Table 3.1 – Gradients of the Kiefer’s ϕ_q -criteria. We recall that $\Pi_{\min}(M) = \mathbf{u}\mathbf{u}^T / \|\mathbf{u}\|_2^2$ is only defined when the least eigenvalue of M has multiplicity one and \mathbf{u} denotes a nonzero eigenvector associated to this least eigenvalue. If the least eigenvalue has multiplicity greater than 2, then the subgradient $\partial\phi_q(M)$ of $\lambda_{\min}(M)$ is the set of all projectors on subspaces of the eigenspace associated to $\lambda_{\min}(M)$; see for example [Lew96]. Notice further that ϕ_q is upper semi-continuous and is a positively homogeneous function

For this we have the following standard result.

Theorem 3.1 (Equivalence theorem). *Let $q \in (-\infty, 1)$ and let $\mathcal{X} \subseteq \mathbb{R}^n$ be a compact semialgebraic set as defined in (3.4) and with nonempty interior. Then problem (3.14) is a convex optimization problem with a unique optimal solution $\mathbf{y}^* \in \mathcal{M}_{2d}(\mathcal{X})$. Denote by p_d^* the polynomial*

$$p_d^*(x) := \mathbf{v}_d(x)^T \mathbf{M}_d(\mathbf{y}^*)^{q-1} \mathbf{v}_d(x) = \|\mathbf{M}_d(\mathbf{y}^*)^{\frac{q-1}{2}} \mathbf{v}_d(x)\|_2^2. \quad (3.15)$$

Then \mathbf{y}^* is the vector of moments up to order $2d$ of a discrete measure μ^* supported on at least $\binom{n+d}{n}$ and at most $\binom{n+2d}{n}$ points in the set

$$\Omega := \{\mathbf{x} \in \mathcal{X} : \text{trace}(\mathbf{M}_d(\mathbf{y}^*)^q) - p_d^*(\mathbf{x}) = 0\}.$$

In particular, the following statements are equivalent:

- $\mathbf{y}^* \in \mathcal{M}_{2d}(\mathcal{X})$ is the unique solution to problem (3.14);
- $\mathbf{y}^* \in \{\mathbf{y} \in \mathcal{M}_{2d}(\mathcal{X}) : y_0 = 1\}$ and $p^* := \text{trace}(\mathbf{M}_d(\mathbf{y}^*)^q) - p_d^* \geq 0$ on \mathcal{X} .

Proof. A general equivalence theorem for concave functionals of the information matrix is stated and proved in [Kie74, Theorem 1]. The case of ϕ_q -criteria is tackled in [Puk06] and [DS97, Theorem 5.4.7]. In order to be self-contained and because the proof of Theorem 3.3 follows the same road map, we recall a sketch of the proof.

First, let us prove that problem (3.14) has an optimal solution. The feasible set is nonempty with finite associated objective value (take as feasible point the vector $\mathbf{y} \in \mathcal{M}_{2d}(\mathcal{X})$ associated with the Lebesgue measure on the compact set \mathcal{X} , scaled to be a probability measure). Moreover, as \mathcal{X} is compact with nonempty interior, it follows that $\mathcal{M}_{2d}(\mathcal{X})$ is closed (as the dual of $\mathcal{P}_{2d}(\mathcal{X})$).

In addition, the feasible set $\{\mathbf{y} \in \mathcal{M}_{2d}(\mathcal{X}) : y_0 = 1\}$ of problem (3.14) is compact. Indeed, there exists $M > 1$ such that $\int_{\mathcal{X}} x_i^{2d} d\mu < M$ for every probability measure μ on \mathcal{X} and every $i = 1, \dots, n$. Hence, $\max\{y_0, \max_i \{L_y(x_i^{2d})\}\} < M$, which by, e.g., [LN07] implies that $|y_\alpha| \leq M$ for every $|\alpha| \leq 2d$, which in turn implies that the feasible set of (3.14) is compact.

Next, as the function ϕ_q is upper semi-continuous, the supremum in (3.14) is attained at some optimal solution $\mathbf{y}^* \in \mathcal{M}_{2d}(\mathcal{X})$. Moreover, as the feasible set is convex and ϕ_q is strictly concave (see, e.g., [Puk06, Chapter 6.13]), \mathbf{y}^* is the unique optimal solution.

Now, we examine the properties of the polynomial p^* and show the equivalence statement. For this we notice that there exists a strictly feasible solution because the cone $\mathcal{M}_{2d}(\mathcal{X})$ has nonempty interior by [Las15a, Lemma 2.6]. Hence, Slater's condition¹ holds for (3.14). Further, by an argument in [Puk06, Chapter 7.13], the matrix $\mathbf{M}_d(\mathbf{y}^*)$ is non-singular. Therefore, ϕ_q is differentiable at \mathbf{y}^* . Since additionally Slater's condition is fulfilled and ϕ_q is concave, this implies that the *Karush-Kuhn-Tucker* (KKT) optimality conditions² at \mathbf{y}^* are necessary and sufficient for \mathbf{y}^* to be an optimal solution.

The KKT-optimality conditions at \mathbf{y}^* read

$$\lambda^* e_0 - \nabla \phi_q(\mathbf{M}_d(\mathbf{y}^*)) = \hat{\mathbf{p}}^* \quad \text{with } \hat{p}^*(x) := \langle \hat{\mathbf{p}}^*, \mathbf{v}_{2d}(x) \rangle \in \mathcal{P}_{2d}(\mathcal{X}),$$

where $\hat{\mathbf{p}}^* \in \mathbb{R}^{s(2d)}$, $e_0 = (1, 0, \dots, 0)$, and λ^* is the dual variable associated with the constraint $y_0 = 1$. The complementarity condition reads $\langle \mathbf{y}, \hat{\mathbf{p}}^* \rangle = 0$.

Writing \mathbf{B}_α , $|\alpha| \leq 2d$, for the real symmetric matrices satisfying

$$\sum_{|\alpha| \leq 2d} \mathbf{B}_\alpha x^\alpha = \mathbf{v}_d(x) \mathbf{v}_d(x)^T,$$

and $\langle A, B \rangle = \text{trace}(AB)$ for the dot product of two real symmetric matrices A and B , this can be expressed as

$$(1_{\alpha=0} \lambda^* - \langle \nabla \phi_q(\mathbf{M}_d(\mathbf{y}^*)), \mathbf{B}_\alpha \rangle)_{|\alpha| \leq 2d} = \hat{\mathbf{p}}^*, \quad \hat{\mathbf{p}}^* \in \mathcal{P}_{2d}(\mathcal{X}). \quad (3.16)$$

Taking the dot product of (3.16) with \mathbf{y}^* and invoking the complementarity condition yields

$$\begin{aligned} \lambda^* &= \lambda^* y_0^* \stackrel{(3.16)}{=} \left\langle \nabla \phi_q(\mathbf{M}_d(\mathbf{y}^*)), \sum_{|\alpha| \leq 2d} y_\alpha^* \mathbf{B}_\alpha \right\rangle \\ &= \langle \nabla \phi_q(\mathbf{M}_d(\mathbf{y}^*)), \mathbf{M}_d(\mathbf{y}^*) \rangle \stackrel{\text{Euler}}{=} \phi_q(\mathbf{M}_d(\mathbf{y}^*)), \end{aligned} \quad (3.17)$$

where the last equality holds by Euler's homogeneous function theorem for the positively homogeneous function ϕ_q .

Similarly, taking the dot product of (3.16) with $\mathbf{v}_{2d}(\mathbf{x})$ gives for all $\mathbf{x} \in \mathcal{X}$:

$$\begin{aligned} 0 &\leq \hat{p}^*(\mathbf{x}) \stackrel{(3.16)}{=} \lambda^* - \left\langle \nabla \phi_q(\mathbf{M}_d(\mathbf{y}^*)), \sum_{|\alpha| \leq 2d} \mathbf{B}_\alpha \mathbf{x}^\alpha \right\rangle \\ &= \lambda^* - \langle \nabla \phi_q(\mathbf{M}_d(\mathbf{y}^*)), \mathbf{v}_d(\mathbf{x}) \mathbf{v}_d(\mathbf{x})^T \rangle. \end{aligned} \quad (3.18)$$

For $q \neq 0$ let $c^* := s(d) \left(\frac{1}{s(d)} \text{trace}(\mathbf{M}_d(\mathbf{y}^*)^q) \right)^{1-1/q}$. As $\mathbf{M}_d(\mathbf{y}^*)$ is positive semidefinite and non-singular, we have $c^* > 0$. If $q = 0$, let $c^* := 1$ and replace $\phi_0(\mathbf{M}_d(\mathbf{y}^*))$ by $\log \det \mathbf{M}_d(\mathbf{y}^*)$, for which the gradient is $\mathbf{M}_d(\mathbf{y}^*)^{-1}$.

¹For the optimization problem $\max \{f(\mathbf{x}) : \mathbf{A}\mathbf{x} = \mathbf{b}; \mathbf{x} \in \mathbf{C}\}$, where $A \in \mathbb{R}^{m \times n}$ and $\mathbf{C} \subseteq \mathbb{R}^n$ is a nonempty closed convex cone, Slater's condition holds, if there exists a feasible solution \mathbf{x} in the interior of \mathbf{C} .

²For the optimization problem $\max \{f(\mathbf{x}) : \mathbf{A}\mathbf{x} = \mathbf{b}; \mathbf{x} \in \mathbf{C}\}$, where f is differentiable, $A \in \mathbb{R}^{m \times n}$ and $\mathbf{C} \subseteq \mathbb{R}^n$ is a nonempty closed convex cone, the KKT-optimality conditions at a feasible point \mathbf{x} state that there exist $\lambda^* \in \mathbb{R}^m$ and $\mathbf{u}^* \in \mathbf{C}^*$ such that $A^T \lambda^* - \nabla f(\mathbf{x}) = \mathbf{u}^*$ and $\langle \mathbf{x}, \mathbf{u}^* \rangle = 0$.

Using Table 3.1, we find that $c^* \nabla \phi_q(\mathbf{M}_d(\mathbf{y}^*)) = \mathbf{M}_d(\mathbf{y}^*)^{q-1}$. It follows that

$$c^* \lambda^* \stackrel{(3.17)}{=} c^* \langle \nabla \phi_q(\mathbf{M}_d(\mathbf{y}^*)), \mathbf{M}_d(\mathbf{y}^*) \rangle = \text{trace}(\mathbf{M}_d(\mathbf{y}^*)^q)$$

and $c^* \langle \nabla \phi_q(\mathbf{M}_d(\mathbf{y}^*)), \mathbf{v}_d(x) \mathbf{v}_d(x)^T \rangle \stackrel{(3.15)}{=} p_d^*(x)$.

Therefore, equation (3.18) is equivalent to $p^* := c^* \hat{p}^* = c^* \lambda^* - p_d^* \in \mathcal{P}_{2d}(\mathcal{X})$. To summarize,

$$p^*(x) = \text{trace}(\mathbf{M}_d(\mathbf{y}^*)^q) - p_d^*(x) \in \mathcal{P}_{2d}(\mathcal{X}).$$

Since the KKT-conditions are necessary and sufficient, the equivalence statement follows.

Finally, we investigate the measure μ^* associated with \mathbf{y}^* . Multiplying the complementarity condition $\langle \mathbf{y}^*, \hat{\mathbf{p}}^* \rangle = 0$ with c^* , we have

$$\int_{\mathcal{X}} \underbrace{p^*(x)}_{\geq 0 \text{ on } \mathcal{X}} \mu^*(dx) = 0.$$

Hence, the support of μ^* is included in the set $\Omega = \{\mathbf{x} \in \mathcal{X} : p^*(\mathbf{x}) = 0\}$.

The measure μ^* can be chosen to be an atomic measure supported on $\ell \leq s(2d)$ points by Tchakaloff's theorem,³ [Las10, Theorem B.12]. If $\ell < s(d)$, then $\text{rank } \mathbf{M}_d(\mathbf{y}^*) < s(d)$ in contradiction to $\mathbf{M}_d(\mathbf{y}^*)$ being non-singular. Therefore, $s(d) \leq \ell \leq s(2d)$. \square

Remark 3.3.1 (On the optimal dual polynomial). The polynomial p_d^* contains all the information concerning the optimal design. Indeed, its level set Ω supports the optimal design points. The polynomial is related to the so-called Christoffel function (see Section 3.3.2). For this reason, in the sequel, the polynomial p_d^* defined in (3.15) will be called a *Christoffel-like* polynomial.

Notice further that

$$\mathcal{X} \subseteq \{p_d^* \leq \text{trace}(\mathbf{M}_d(\mathbf{y}^*)^q)\}.$$

Hence, the optimal design problem related to ϕ_q is similar to the standard problem of computational geometry consisting in minimizing the volume of a polynomial level set containing \mathcal{X} (Löwner-John's ellipsoid theorem). Here, the volume functional is replaced by $\phi_q(M)$ for the polynomial $\|\mathbf{M}^{\frac{q-1}{2}} \mathbf{v}_d(\mathbf{x})\|_2^2$. We refer to [Las15a] for a discussion and generalizations of Löwner-John's ellipsoid theorem for general homogenous polynomials on non-convex domains.

Remark 3.3.2 (Equivalence theorem for T -optimality). Theorem 3.1 also holds for $q = 1$. This is the T -optimal design case for which the objective function is linear. Hence, in this case \mathbf{y}^* is not unique. Further, note that the polynomial p_d^* can be explicitly written as it does not depend on \mathbf{y}^* . Namely,

$$p_d^*(x) = \|\mathbf{v}_d(x)\|_2^2.$$

Thus, the support of any solution is included in the level set

$$\Omega = \arg \max_{\mathbf{x} \in \mathcal{X}} p_d^*(\mathbf{x}).$$

It follows that the set of solutions is exactly the set of probability measures supported by Ω .

Remark 3.3.3 (Equivalence theorem for E -optimality). Theorem 3.1 also holds for $q = -\infty$. This is the E -optimal design case, in which the objective function is not differentiable at points for which the least eigenvalue has multiplicity greater than 2. We get that \mathbf{y}^* is the vector of

³Tchakaloff's theorem states that for every finite Borel probability measure on \mathcal{X} and every $r \in \mathbb{N}$, there exists an atomic measure μ_r supported on $\ell \leq s(r)$ points such that all moments of μ_r and μ^* agree up to order r .

moments up to order $2d$ of a discrete measure μ^* supported on at most $s(2d)$ points in the set

$$\Omega := \left\{ \mathbf{x} \in \mathcal{X} : \lambda_{\min}(\mathbf{M}_d(\mathbf{y}^*)) \|\mathbf{u}\|_2^2 - \left(\sum_{|\alpha| \leq 2d} u_\alpha x^\alpha \right)^2 = 0 \right\},$$

where $\mathbf{u} = (u_\alpha)_{|\alpha| \leq 2d}$ is a nonzero eigenvector of $\mathbf{M}_d(\mathbf{y}^*)$ associated to $\lambda_{\min}(\mathbf{M}_d(\mathbf{y}^*))$. In particular, the following statements are equivalent

- $\mathbf{y}^* \in \mathcal{M}_{2d}(\mathcal{X})$ is a solution to problem (3.14);
- $\mathbf{y}^* \in \{\mathbf{y} \in \mathcal{M}_{2d}(\mathcal{X}) : y_0 = 1\}$ and $\lambda_{\min}(\mathbf{M}_d(\mathbf{y}^*)) \|\mathbf{u}\|_2^2 - \left(\sum_\alpha u_\alpha x^\alpha \right)^2 \geq 0$ on \mathcal{X} .

Furthermore, if the least eigenvalue of $\mathbf{M}_d(\mathbf{y}^*)$ has multiplicity one then $\mathbf{y}^* \in \mathcal{M}_{2d}(\mathcal{X})$ is unique.

3.3.2 Christoffel polynomials

In the case of D -optimality it turns out that the unique optimal solution $\mathbf{y}^* \in \mathcal{M}_{2d}(\mathcal{X})$ of problem (3.13) can be characterized in terms of the *Christoffel polynomial* of degree $2d$ associated with an optimal measure μ whose moments up to order $2d$ coincide with \mathbf{y}^* .

Definition 3.2 (Christoffel polynomial). Let $\mathbf{y} \in \mathbb{R}^{s(2d)}$ be such that $\mathbf{M}_d(\mathbf{y}) \succ 0$. Then there exists a family of orthonormal polynomials $(P_\alpha)_{|\alpha| \leq d} \subseteq \mathbb{R}[x]_d$ satisfying

$$L_{\mathbf{y}}(P_\alpha P_\beta) = \delta_{\alpha=\beta} \quad \text{and} \quad L_{\mathbf{y}}(x^\alpha P_\beta) = 0 \quad \forall \alpha \prec \beta,$$

where monomials are ordered with respect to the lexicographical ordering on \mathbb{N}^n . We call the polynomial

$$p_d(\mathbf{x}) := \sum_{|\alpha| \leq d} P_\alpha(\mathbf{x})^2,$$

the *Christoffel polynomial* (of degree d) associated with \mathbf{y} .

The Christoffel polynomial⁴ can be expressed in different ways. For instance, via the inverse of the moment matrix by

$$p_d(x) = \mathbf{v}_d(x)^T \mathbf{M}_d(\mathbf{y})^{-1} \mathbf{v}_d(x),$$

or via its extremal property

$$\frac{1}{p_d(\mathbf{x})} = \min_{P \in \mathbb{R}[x]_d} \left\{ \int P(x)^2 d\mu(x) : P(\mathbf{x}) = 1 \right\} \quad \forall \mathbf{x} \in \mathbb{R}^n,$$

if \mathbf{y} has a representing measure μ . If \mathbf{y} does not have a representing measure μ just replace $\int P(x)^2 d\mu(x)$ with $L_{\mathbf{y}}(P^2) (= P^T \mathbf{M}_d(\mathbf{y}) P)$. For more details the interested reader is referred to [LP16] and the references therein. Notice also that there is a regain of interest in the asymptotic study of the Christoffel function as it relies on eigenvalue marginal distributions of invariant random matrix ensembles; see for example [Led04].

Remark (Equivalence theorem for D -optimality). In the case of D -optimal designs, observe that

$$t^* := \max_{\mathbf{x} \in \mathcal{X}} p_d^*(\mathbf{x}) = \text{trace}(\text{Id}) = s(d),$$

where p_d^* is given by (3.15) for $q = 0$. Furthermore, note that p_d^* is the Christoffel polynomial of degree $2d$ of the D -optimal measure μ^* .

⁴Actually, what is referred to as the *Christoffel function* in the literature is its reciprocal $\mathbf{x} \mapsto \frac{1}{p_d(\mathbf{x})}$.

3.3.3 The SDP relaxation scheme

Let $\mathcal{X} \subseteq \mathbb{R}^n$ be as defined in (3.4), assumed to be compact. So with no loss of generality (and possibly after scaling), assume that $g_1(x) = 1 - \|x\|^2 \geq 0$ is one of the constraints defining \mathcal{X} .

Since the ideal moment problem (3.14) involves the moment cone $\mathcal{M}_{2d}(\mathcal{X})$, which is not semidefinite representable, we use the hierarchy (3.7) of outer approximations of the moment cone to relax problem (3.14) to a semidefinite program. So, for a fixed integer $\delta \geq 1$ we consider the problem

$$\begin{aligned} \rho_\delta = \max_{\mathbf{y}} \phi_q(\mathbf{M}_d(\mathbf{y})) \\ \text{s.t. } \mathbf{y} \in \mathcal{M}_{2(d+\delta)}^{2d}(g_1, \dots, g_m), \quad y_0 = 1. \end{aligned} \quad (3.19)$$

Since problem (3.19) is a relaxation of the ideal problem (3.14), necessarily $\rho_\delta \geq \rho$ for all δ . In analogy with Theorem 3.1, we have the following result characterizing the solutions of the SDP relaxation (3.19) by means of sum-of-squares polynomials.

Theorem 3.3 (Equivalence theorem for SDP relaxations). *Let $q \in (-\infty, 1)$ and let $\mathcal{X} \subseteq \mathbb{R}^n$ be a compact semialgebraic set as defined in (3.4) and with non-empty interior. Then,*

- a) SDP problem (3.19) has a unique optimal solution $\mathbf{y}^* \in \mathbb{R}^{s(2d)}$.
- b) Let p_d^* be as defined in (3.15), associated with \mathbf{y}^* . Then $p^* := \text{trace}(\mathbf{M}_d(\mathbf{y}^*)^q) - p_d^* \geq 0$ on \mathcal{X} and $L_{\mathbf{y}^*}(p^*) = 0$.

In particular, the following statements are equivalent:

- $\mathbf{y}^* \in \mathcal{M}_{2(d+\delta)}^{2d}(g_1, \dots, g_m)$ is the unique solution to problem (3.19);
- $\mathbf{y}^* \in \left\{ \mathbf{y} \in \mathcal{M}_{2(d+\delta)}^{2d}(g_1, \dots, g_m) : y_0 = 1 \right\}$ and the polynomial $p^* = \text{trace}(\mathbf{M}_d(\mathbf{y}^*)^q) - p_d^*$ belongs to the quadratic module $\mathcal{Q}_{2(d+\delta)}^{2d}(g_1, \dots, g_m)$.

Proof. We follow the proof of Theorem 3.1.

- a) Let us prove that problem (3.19) has an optimal solution. The feasible set is nonempty with finite associated objective value (take as feasible point the vector $\tilde{\mathbf{y}}$ associated with the Lebesgue measure on the compact set \mathcal{X} , scaled to be a probability measure).

Next, let $\mathbf{y} \in \mathbb{R}^{s(2d)}$ be an arbitrary feasible solution and $\mathbf{y}_\delta \in \mathcal{M}_{2(d+\delta)}(g_1, \dots, g_m)$ an arbitrary lifting of \mathbf{y} —recall the definition of $\mathcal{M}_{2(d+\delta)}^{2d}(g_1, \dots, g_m)$ given in (3.6). As $g_1(x) = 1 - \|x\|^2$ and $\mathbf{M}_{d+\delta-1}(g_1 \mathbf{y}_\delta) \succcurlyeq 0$, we have $L_{\mathbf{y}_\delta}(x_i^{2(d+\delta)}) \leq 1$, $i = 1, \dots, n$, and so by [LN07],

$$|y_{\delta, \alpha}| \leq \max_{\substack{=1 \\ \alpha}} \{y_{\delta, 0}, \max_i \{L_{\mathbf{y}_\delta}(x_i^{2(d+\delta)})\}\} \leq 1 \quad \forall |\alpha| \leq 2(d+\delta). \quad (3.20)$$

This implies that the set of feasible liftings \mathbf{y}_δ is compact, and therefore, the feasible set of (3.19) is also compact. As the function ϕ_q is upper semi-continuous, the supremum in (3.19) is attained at some optimal solution $\mathbf{y}^* \in \mathbb{R}^{s(2d)}$. It is unique due to convexity of the feasible set and strict concavity of the objective function ϕ_q ; e.g., see [Puk06, Chapter 6.13].

b) Let $\mathbf{B}_\alpha, \tilde{\mathbf{B}}_\alpha$ and $\mathbf{C}_{j\alpha}$ be real symmetric matrices such that

$$\begin{aligned} \sum_{|\alpha| \leq 2d} \mathbf{B}_\alpha x^\alpha &= \mathbf{v}_d(x) \mathbf{v}_d(x)^T \\ \sum_{|\alpha| \leq 2(d+\delta)} \tilde{\mathbf{B}}_\alpha x^\alpha &= \mathbf{v}(x)_{d+\delta} \mathbf{v}_{d+\delta}(x)^T \\ \sum_{|\alpha| \leq 2(d+\delta)} \mathbf{C}_{j\alpha} x^\alpha &= g_j(x) \mathbf{v}_{d+\delta-d_j}(x) \mathbf{v}_{d+\delta-d_j}(x)^T, \quad j = 1, \dots, m. \end{aligned}$$

First, we notice that there exists a strictly feasible solution to (3.19) because the cone $\mathcal{M}_{2(d+\delta)}^{2d}(g_1, \dots, g_m)$ has nonempty interior as a supercone of $\mathcal{M}_{2d}(\mathcal{X})$, which has nonempty interior by [Las15a, Lemma 2.6]. Hence, Slater's condition holds for (3.19). Further, by an argument in [Puk06, Chapter 7.13]) the matrix $\mathbf{M}_d(\mathbf{y}^*)$ is non-singular. Therefore, ϕ_q is differentiable at \mathbf{y}^* . Since additionally Slater's condition is fulfilled and ϕ_q is concave, this implies that the *Karush-Kuhn-Tucker* (KKT) optimality conditions at \mathbf{y}^* are necessary and sufficient for \mathbf{y}^* to be an optimal solution.

The KKT-optimality conditions at \mathbf{y}^* read

$$\lambda^* e_0 - \nabla \phi_q(\mathbf{M}_d(\mathbf{y}^*)) = \hat{\mathbf{p}}^* \quad \text{with } \hat{\mathbf{p}}^*(x) := \langle \hat{\mathbf{p}}^*, \mathbf{v}_{2d}(x) \rangle \in \mathcal{Q}_{2(d+\delta)}^{2d}(g_1, \dots, g_m),$$

where $\hat{\mathbf{p}}^* \in \mathbb{R}^{s(2d)}$, $e_0 = (1, 0, \dots, 0)$, and λ^* is the dual variable associated with the constraint $y_0 = 1$. The complementarity condition reads $\langle \mathbf{y}, \hat{\mathbf{p}}^* \rangle = 0$.

Recalling the definition (3.8) of the quadratic module $\mathcal{Q}_{2(d+\delta)}^{2d}(g_1, \dots, g_m)$, we can express the membership $\hat{\mathbf{p}}^*(x) \in \mathcal{Q}_{2(d+\delta)}^{2d}(g_1, \dots, g_m)$ more explicitly in terms of some “dual variables” $\Lambda_j \succcurlyeq 0$, $j = 0, \dots, m$,

$$1_{\alpha=0} \lambda^* - \langle \nabla \phi_q(\mathbf{M}_d(\mathbf{y}^*)), \mathbf{B}_\alpha \rangle = \langle \Lambda_0, \tilde{\mathbf{B}}_\alpha \rangle + \sum_{j=1}^m \langle \Lambda_j, \mathbf{C}_\alpha^j \rangle, \quad |\alpha| \leq 2(d+\delta). \quad (3.21)$$

Then, for a lifting $\mathbf{y}_\delta^* \in \mathcal{M}_{2(d+\delta)}(g_1, \dots, g_m)$ of \mathbf{y}^* the complementarity condition $\langle \mathbf{y}^*, \hat{\mathbf{p}}^* \rangle = 0$ reads

$$\langle \mathbf{M}_{d+\delta}(\mathbf{y}_\delta^*), \Lambda_0 \rangle = 0, \quad \langle \mathbf{M}_{d+\delta-d_j}(\mathbf{y}_\delta^* g_j), \Lambda_j \rangle = 0, \quad j = 1, \dots, m. \quad (3.22)$$

Multiplying by $y_{\delta,\alpha}^*$, summing up, and using the complementarity conditions (3.22) yields

$$\lambda^* - \langle \nabla \phi_q(\mathbf{M}_d(\mathbf{y}^*)), \mathbf{M}_d(\mathbf{y}^*) \rangle = \underbrace{\langle \Lambda_0, \mathbf{M}_{d+\delta}(\mathbf{y}_\delta^*) \rangle}_{=0} + \sum_{j=1}^m \underbrace{\langle \Lambda_j, \mathbf{M}_{d+\delta-d_j}(g_j \mathbf{y}_\delta^*) \rangle}_{=0}. \quad (3.23)$$

We deduce that $\lambda^* = \langle \nabla \phi_q(\mathbf{M}_d(\mathbf{y}^*)), \mathbf{M}_d(\mathbf{y}^*) \rangle = \phi_q(\mathbf{M}_d(\mathbf{y}^*))$, where the latter equality follows from the Euler formula for homogeneous functions.

Similarly, multiplying (3.21) by x^α and summing up gives the degree $2d$ polynomial

$$\begin{aligned} &\lambda^* - \langle \nabla \phi_q(\mathbf{M}_d(\mathbf{y}^*)), \mathbf{v}_d(x)^T \mathbf{v}_d(x) \rangle \\ &= \left\langle \Lambda_0, \sum_{|\alpha| \leq 2(d+\delta)} \tilde{\mathbf{B}}_\alpha x^\alpha \right\rangle + \sum_{j=1}^m \left\langle \Lambda_j, \sum_{|\alpha| \leq 2(d+\delta-d_j)} \mathbf{C}_\alpha^j x^\alpha \right\rangle \\ &= \underbrace{\left\langle \Lambda_0, \mathbf{v}(x)_{d+\delta} \mathbf{v}_{d+\delta}(x)^T \right\rangle}_{=: \sigma_0(x)} + \sum_{j=1}^m g_j(x) \underbrace{\left\langle \Lambda_j, \mathbf{v}_{d+\delta-d_j}(x) \mathbf{v}_{d+\delta-d_j}(x)^T \right\rangle}_{=: \sigma_j(x)} \end{aligned}$$

$$= \sigma_0(x) + \sum_{j=1}^n \sigma_j(x) g_j(x) = \hat{p}^*(x) \in \mathcal{Q}_{2(d+\delta)}^{2d}(g_1, \dots, g_m). \quad (3.24)$$

Note that $\sigma_0 \in \Sigma[x]_{2(d+\delta)}$ and $\sigma_j \in \Sigma[x]_{2(d+\delta-d_j)}$, $j = 1, \dots, m$, by definition.

For $q \neq 0$ let $c^* := s(d) \left(\frac{1}{s(d)} \text{trace}(\mathbf{M}_d(\mathbf{y}^*)^q) \right)^{1-1/q}$. As $\mathbf{M}_d(\mathbf{y}^*)$ is positive semidefinite and non-singular, we have $c^* > 0$. If $q = 0$, let $c^* := 1$ and replace $\phi_0(\mathbf{M}_d(\mathbf{y}^*))$ by $\log \det \mathbf{M}_d(\mathbf{y}^*)$, for which the gradient is $\mathbf{M}_d(\mathbf{y}^*)^{-1}$.

Using Table 3.1, we find that $c^* \nabla \phi_q(\mathbf{M}_d(\mathbf{y}^*)) = \mathbf{M}_d(\mathbf{y}^*)^{q-1}$. It follows that

$$\begin{aligned} c^* \lambda^* &\stackrel{(3.23)}{=} c^* \langle \nabla \phi_q(\mathbf{M}_d(\mathbf{y}^*)), \mathbf{M}_d(\mathbf{y}^*) \rangle = \text{trace}(\mathbf{M}_d(\mathbf{y}^*)^q) \\ c^* \langle \nabla \phi_q(\mathbf{M}_d(\mathbf{y}^*)), \mathbf{v}_d(x) \mathbf{v}_d(x)^T \rangle &\stackrel{(3.15)}{=} p_d^*(x). \end{aligned}$$

Therefore, equation (3.24) implies $p^* := c^* \hat{p}^* = c^* \lambda^* - p_d^* \in \mathcal{Q}_{2(d+\delta)}^{2d}(g_1, \dots, g_m)$. To summarize,

$$p^*(x) = \text{trace}(\mathbf{M}_d(\mathbf{y}^*)^q) - p_d^*(x) \in \mathcal{Q}_{2(d+\delta)}^{2d}(g_1, \dots, g_m).$$

We remark that all elements of $\mathcal{Q}_{2(d+\delta)}^{2d}(g_1, \dots, g_m)$ are non-negative on \mathcal{X} and that (3.23) implies $L_{\mathbf{y}^*}(p^*) = 0$. Hence, we have shown b).

The equivalence statement follows from the argument in b). \square

Remark 3.3.4 (Finite convergence). If the optimal solution \mathbf{y}^* of problem (3.19) is coming from a measure μ^* on \mathcal{X} , that is $\mathbf{y}^* \in \mathcal{M}_{2d}(\mathcal{X})$, then $\rho_\delta = \rho$ and \mathbf{y}^* is the unique optimal solution to problem (3.14). In addition, by Theorem 3.1, μ^* can be chosen to be atomic and supported on at least $s(d)$ and at most $s(2d)$ ‘‘contact points’’ on the level set $\Omega := \{\mathbf{x} \in \mathcal{X} : \text{trace}(\mathbf{M}_d(\mathbf{y}^*)^q) - p_d^*(\mathbf{x}) = 0\}$.

Remark 3.3.5 (SDP relaxation for T -optimality). In this case, recall that \mathbf{y}^* is not unique and recall that the polynomial p_d^* can be explicitly written as $p_d^*(x) = \|\mathbf{v}_d(x)\|_2^2$. The above proof can be extended to the case $q = 1$ and one derives that any solution \mathbf{y}^* satisfies $\text{trace}(\mathbf{M}_d(\mathbf{y}^*)) - \|\mathbf{v}_d(\mathbf{x})\|_2^2 \geq 0$ for all $\mathbf{x} \in \mathcal{X}$ and $L_{\mathbf{y}^*}(p_d^*) = \text{trace}(\mathbf{M}_d(\mathbf{y}^*))$. In particular, the following statements are equivalent:

- $\mathbf{y}^* \in \mathcal{M}_{2(d+\delta)}^{2d}(g_1, \dots, g_m)$ is a solution to problem (3.19) (for $q = 1$);
- $\mathbf{y}^* \in \left\{ \mathbf{y} \in \mathcal{M}_{2(d+\delta)}^{2d}(g_1, \dots, g_m) : y_0 = 1 \right\}$ and $p^*(x) := \text{trace}(\mathbf{M}_d(\mathbf{y}^*)) - \|\mathbf{v}_d(x)\|_2^2$ belongs to the quadratic module $\mathcal{Q}_{2(d+\delta)}^{2d}(g_1, \dots, g_m)$.

Remark 3.3.6 (SDP relaxation for E -optimality). Theorem 3.3 holds also for $q = -\infty$. This is the E -optimal design case, in which the objective function is not differentiable at points for which the least eigenvalue has multiplicity greater than 2. We get that \mathbf{y}^* satisfies $\lambda_{\min}(\mathbf{M}_d(\mathbf{y}^*)) - \left(\sum_{\alpha} u_{\alpha} \mathbf{x}^{\alpha} \right)^2 \geq 0$ for all $\mathbf{x} \in \mathcal{X}$ and $L_{\mathbf{y}^*} \left(\left(\sum_{\alpha} u_{\alpha} \mathbf{x}^{\alpha} \right)^2 \right) = \lambda_{\min}(\mathbf{M}_d(\mathbf{y}^*))$, where $\mathbf{u} = (u_{\alpha})_{|\alpha| \leq 2d}$ is a nonzero eigenvector of $\mathbf{M}_d(\mathbf{y}^*)$ associated to $\lambda_{\min}(\mathbf{M}_d(\mathbf{y}^*))$. In particular, the following statements are equivalent:

- $\mathbf{y}^* \in \mathcal{M}_{2(d+\delta)}^{2d}(g_1, \dots, g_m)$ is a solution to problem (3.19) (for $q = -\infty$);
- $\mathbf{y}^* \in \left\{ \mathbf{y} \in \mathcal{M}_{2(d+\delta)}^{2d}(g_1, \dots, g_m) : y_0 = 1 \right\}$ and $p^*(x) = \lambda_{\min}(\mathbf{M}_d(\mathbf{y}^*)) \|\mathbf{u}\|_2^2 - \left(\sum_{\alpha} u_{\alpha} \mathbf{x}^{\alpha} \right)^2$ belongs to the quadratic module $\mathcal{Q}_{2(d+\delta)}^{2d}(g_1, \dots, g_m)$.

Furthermore, if the least eigenvalue of $\mathbf{M}_d(\mathbf{y}^*)$ has multiplicity one, then \mathbf{y}^* is unique.

3.3.4 Asymptotics

We now analyze what happens when δ tends to infinity.

Theorem 3.4. *Let $q \in (-\infty, 1)$ and $d \in \mathbb{N}$. For every $\delta = 0, 1, 2, \dots$ let $\mathbf{y}_{d,\delta}^*$ be an optimal solution to (3.19) and $p_{d,\delta}^* \in \mathbb{R}[x]_{2d}$ the Christoffel-like polynomial associated with $\mathbf{y}_{d,\delta}^*$ as in Theorem 3.3. Then,*

- a) $\rho_\delta \rightarrow \rho$ as $\delta \rightarrow \infty$, where ρ is the supremum in (3.14).
- b) For every $|\alpha| \leq 2d$ we have $\lim_{\delta \rightarrow \infty} y_{d,\delta,\alpha}^* = y_\alpha^*$, where $\mathbf{y}^* \in \mathcal{M}_{2d}(\mathcal{X})$ is the unique optimal solution to (3.14).
- c) $p_{d,\delta}^* \rightarrow p_d^*$ as $\delta \rightarrow \infty$, where p_d^* is the Christoffel-like polynomial associated with \mathbf{y}^* defined in (3.15).
- d) If the dual polynomial $p^* := \text{trace}(\mathbf{M}_d(\mathbf{y}^*)^q) - p_d^*$ to problem (3.14) belongs to the quadratic module $\mathcal{Q}_{2(d+\delta)}^{2d}(g_1, \dots, g_m)$ for some δ , then finite convergence takes place, that is, $\mathbf{y}_{d,\delta}^*$ is the unique optimal solution to problem (3.14) and $\mathbf{y}_{d,\delta}^*$ has a representing measure, namely, the target measure μ^* .

Proof. We prove the four claims consecutively.

- a) For every δ complete the lifted finite sequence $\mathbf{y}_\delta^* \in \mathbb{R}^{s(2(d+\delta))}$ of $\mathbf{y}_{d,\delta}^*$ with zeros to make it an infinite sequence $\mathbf{y}_\delta^* = (y_{\delta,\alpha}^*)_{\alpha \in \mathbb{N}^n}$. Therefore, every such \mathbf{y}_δ^* can be identified with an element of ℓ_∞ , the Banach space of bounded sequences equipped with the supremum norm. Moreover, inequality (3.20) holds for every \mathbf{y}_δ^* . Thus, denoting by \mathcal{B} the unit ball of ℓ_∞ , which is compact in the $\sigma(\ell_\infty, \ell_1)$ weak- \star topology on ℓ_∞ , we have $\mathbf{y}_\delta^* \in \mathcal{B}$. By Banach-Alaoglu's theorem, there is an element $\hat{\mathbf{y}} \in \mathcal{B}$ and a converging subsequence $(\delta_k)_{k \in \mathbb{N}}$ such that

$$\lim_{k \rightarrow \infty} y_{\delta_k, \alpha}^* = \hat{y}_\alpha \quad \forall \alpha \in \mathbb{N}^n. \quad (3.25)$$

Let $s \in \mathbb{N}$ be arbitrary, but fixed. By the convergence (3.25), we also have

$$\lim_{k \rightarrow \infty} \mathbf{M}_s(\mathbf{y}_{\delta_k}^*) = \mathbf{M}_s(\hat{\mathbf{y}}) \succcurlyeq 0, \quad \lim_{k \rightarrow \infty} \mathbf{M}_s(g_j \mathbf{y}_{\delta_k}^*) = \mathbf{M}_s(g_j \hat{\mathbf{y}}) \succcurlyeq 0, \quad j = 1, \dots, m.$$

Next, by Putinar's Positivstellensatz (Theorem 1.1), $\hat{\mathbf{y}}$ is the sequence of moments of some measure $\hat{\mu} \in \mathcal{M}_+(\mathcal{X})$, and so $\hat{\mathbf{y}}_d := (\hat{y}_\alpha)_{|\alpha| \leq 2d}$ is a feasible solution to (3.14), meaning $\phi_q(\mathbf{M}_d(\hat{\mathbf{y}}_d)) \leq \rho$. On the other hand, as (3.19) is a relaxation of (3.14), we have $\rho_{\delta_k} \geq \rho$ for all δ_k . So the convergence (3.25) yields

$$\phi_q(\mathbf{M}_d(\hat{\mathbf{y}}_d)) = \lim_{k \rightarrow \infty} \rho_{\delta_k} \geq \rho,$$

which proves that $\hat{\mathbf{y}}_d$ is an optimal solution to (3.14), and $\lim_{\delta \rightarrow \infty} \rho_\delta = \rho$.

- b) As the optimal solution to (3.14) is unique, we have $\mathbf{y}^* = \hat{\mathbf{y}}_d$ with $\hat{\mathbf{y}}_d$ defined in the proof of a) and the whole sequence $(\mathbf{y}_{d,\delta}^*)_{\delta \in \mathbb{N}}$ converges to \mathbf{y}^* , that is, for $|\alpha| \leq 2d$ fixed,

$$\lim_{\delta \rightarrow \infty} y_{d,\delta,\alpha}^* = \lim_{\delta \rightarrow \infty} y_{\delta,\alpha}^* = \hat{y}_\alpha = y_\alpha^*. \quad (3.26)$$

- c) It suffices to observe that the coefficients of the Christoffel-like polynomial $p_{d,\delta}^*$ are continuous functions of the moments $(y_{d,\delta,\alpha}^*)_{|\alpha| \leq 2d} = (y_{\delta,\alpha}^*)_{|\alpha| \leq 2d}$. Therefore, by the convergence (3.26) one has $p_{d,\delta}^* \rightarrow p_d^*$, where $p_d^* \in \mathbb{R}[x]_{2d}$ as in Theorem 3.1.

- d) The last point follows directly observing that in this case the two programs (3.14) and (3.19) satisfy the same KKT conditions. \square

In order to illustrate what happens for general ϕ_q -criteria, we give a corollary describing the relation between the solutions of program (3.14) and program (3.19) for T -optimal designs ($q = 1$).

Corollary 3.5. *Let $q = 1$ and $\delta \geq 1$. Denote by \mathbf{y}^* (resp. $\mathbf{y}_{d,\delta}^*$) a solution to problem (3.14) (resp. program (3.19)). Only one of the following two cases can occur.*

- *Either*

$$\begin{aligned} & \left\{ \text{trace}(\mathbf{M}_d(\mathbf{y}^*)) - \|\mathbf{v}_d(x)\|_2^2 \in \mathcal{Q}_{2(d+\delta)}^{2d}(g_1, \dots, g_m) \right\} \\ & \Leftrightarrow \left\{ \text{trace}(\mathbf{M}_d(\mathbf{y}_{d,\delta}^*)) - \|\mathbf{v}_d(x)\|_2^2 \text{ has a root in } \mathcal{X} \right\} \\ & \Leftrightarrow \left\{ \text{trace}(\mathbf{M}_d(\mathbf{y}_{d,\delta}^*)) = \text{trace}(\mathbf{M}_d(\mathbf{y}^*)) \right\}, \end{aligned}$$

- *or*

$$\begin{aligned} & \left\{ \text{trace}(\mathbf{M}_d(\mathbf{y}^*)) - \|\mathbf{v}_d(x)\|_2^2 \notin \mathcal{Q}_{2(d+\delta)}^{2d}(g_1, \dots, g_m) \right\} \\ & \Leftrightarrow \left\{ \text{trace}(\mathbf{M}_d(\mathbf{y}_{d,\delta}^*)) - \|\mathbf{v}_d(x)\|_2^2 > 0 \text{ on } \mathcal{X} \right\} \\ & \Leftrightarrow \left\{ \text{trace}(\mathbf{M}_d(\mathbf{y}_{d,\delta}^*)) > \text{trace}(\mathbf{M}_d(\mathbf{y}^*)) \right\}. \end{aligned}$$

Recall that $\text{trace}(\mathbf{M}_d(\mathbf{y}^*)) = \max_{\mathbf{x} \in \mathcal{X}} \|\mathbf{v}_d(\mathbf{x})\|_2^2$ for $q = 1$.

3.3.5 General regression polynomial bases

We return to the general case described by a matrix \mathfrak{A} of size $p \times \binom{n+d}{n}$ such that the regression polynomials satisfy $\mathfrak{F}(x) = \mathfrak{A} \mathbf{v}_d(x)$ on \mathcal{X} . Now, the objective function becomes $\phi_q(\mathfrak{A} \mathbf{M}_d(\mathbf{y}) \mathfrak{A}^T)$ at point \mathbf{y} . Note that the constraints on \mathbf{y} are unchanged, i.e., $\mathbf{y} \in \mathcal{M}_{2d}(\mathcal{X})$, $y_0 = 1$ in the ideal problem and $\mathbf{y} \in \mathcal{M}_{2(d+\delta)}^{2d}(g_1, \dots, g_m)$, $y_0 = 1$ in the SDP relaxation scheme. We recall the notation $\mathfrak{J}_d(\mathbf{y}) := \mathfrak{A} \mathbf{M}_d(\mathbf{y}) \mathfrak{A}^T$, and we observe that the KKT conditions at \mathbf{y} are given by

$$\phi_q(\mathfrak{J}_d(\mathbf{y})) - \underbrace{\mathfrak{F}(x)^T \nabla \phi_q(\mathfrak{J}_d(\mathbf{y})) \mathfrak{F}(x)}_{p_d^*(x)} = p^*(x),$$

with $p^* \in \mathcal{P}_{2d}(\mathcal{X})$ in the ideal problem and $p^* \in \mathcal{Q}_{2(d+\delta)}^{2d}(g_1, \dots, g_m)$ in the SDP relaxation scheme. Our analysis leads to the following equivalence results.

Proposition 3.6. *Let $q \in (-\infty, 1)$ and let $\mathcal{X} \subseteq \mathbb{R}^n$ be a compact semialgebraic set as defined in (3.4) and with nonempty interior. Problem (3.13) is a convex optimization problem with an optimal solution $\mathbf{y}^* \in \mathcal{M}_{2d}(\mathcal{X})$. Denote by p_d^* the polynomial*

$$p_d^*(x) := \mathfrak{F}(x)^T \mathfrak{J}_d(\mathbf{y}^*)^{q-1} \mathfrak{F}(x) = \|\mathfrak{J}_d(\mathbf{y}^*)^{\frac{q-1}{2}} \mathfrak{F}(x)\|_2^2. \quad (3.27)$$

Then \mathbf{y}^* is the vector of moments up to order $2d$ of a discrete measure μ^* supported on at least $s(d)$ and at most $s(2d)$ points in the set

$$\Omega := \{\mathbf{x} \in \mathcal{X} : \text{trace}(\mathfrak{J}_d(\mathbf{y}^*)^q) - p_d^*(\mathbf{x}) = 0\}.$$

In particular, the following statements are equivalent:

- $\mathbf{y}^* \in \mathcal{M}_{2d}(\mathcal{X})$ is the solution to problem (3.13);
- $\mathbf{y}^* \in \{\mathbf{y} \in \mathcal{M}_{2d}(\mathcal{X}) : y_0 = 1\}$ and $p^* := \text{trace}(\mathfrak{J}_d(\mathbf{y})^q) - p_d^*(x) \geq 0$ on \mathcal{X} .

Furthermore, if \mathfrak{A} has full column rank then \mathbf{y}^* is unique.

The SDP relaxation is given by the program

$$\begin{aligned} \rho_\delta = \max_{\mathbf{y}} \quad & \phi_q(\mathfrak{J}_d(\mathbf{y})) \\ \text{s.t.} \quad & \mathbf{y} \in \mathcal{M}_{2(d+\delta)}^{2d}(g_1, \dots, g_m), \quad y_0 = 1, \end{aligned} \quad (3.28)$$

for which the following result holds.

Proposition 3.7. *Let $q \in (-\infty, 1)$ and let $\mathcal{X} \subseteq \mathbb{R}^n$ be a compact semialgebraic set as defined in (3.4) and with nonempty interior. Then,*

- a) SDP problem (3.28) has an optimal solution $\mathbf{y}^* \in \mathbb{R}^{s(2d)}$.
- b) Let p_d^* be as defined in (3.27), associated with \mathbf{y}^* . Then $p^* := \text{trace}(\mathfrak{J}_d(\mathbf{y}^*)^q) - p_d^* \geq 0$ on \mathcal{X} and $L_{\mathbf{y}^*}(p^*) = 0$.

In particular, the following statements are equivalent:

- $\mathbf{y}^* \in \mathcal{M}_{2(d+\delta)}^{2d}(g_1, \dots, g_m)$ is a solution to problem (3.28);
- $\mathbf{y}^* \in \{\mathbf{y} \in \mathcal{M}_{2(d+\delta)}^{2d}(g_1, \dots, g_m) : y_0 = 1\}$ and the polynomial $p^* = \text{trace}(\mathfrak{J}_d(\mathbf{y}^*)^q) - p_d^*$ belongs to the quadratic module $\mathcal{Q}_{2(d+\delta)}^{2d}(g_1, \dots, g_m)$.

Furthermore, if \mathfrak{A} has full column rank then \mathbf{y}^* is unique.

3.4 Recovering the measure

By solving step one as explained in Section 3.3, we obtain a solution \mathbf{y}^* to SDP problem (3.19). As $\mathbf{y}^* \in \mathcal{M}_{2(d+\delta)}^{2d}(g_1, \dots, g_m)$, it is likely that it comes from a measure. If this is the case, by Tchakaloff's theorem (see page 48), there exists an atomic measure supported on at most $s(2d)$ points having these moments. For computing the atomic measure, we propose two approaches: A first one which follows a procedure by Nie [Nie14], and a second one which uses properties of the Christoffel-like polynomial associated with \mathbf{y}^* . The approaches have the benefit that they can numerically certify finite convergence of the hierarchy.

3.4.1 Via the Nie method

This approach to recover a measure from its moments is based on a formulation proposed by Nie in [Nie14]. Let $\mathbf{y}^* \in \mathbb{R}^{s(2d)}$ be a finite sequence of moments. For $r \in \mathbb{N}$ consider the SDP problem

$$\begin{aligned} \min_{\mathbf{y}_r} \quad & L_{\mathbf{y}_r}(f_r) \\ \text{s.t.} \quad & \mathbf{y}_r \in \mathcal{M}_{2(d+r)}(g_1, \dots, g_m), \\ & y_{r,\alpha} = y_\alpha^*, \quad |\alpha| \leq 2d, \end{aligned} \quad (3.29)$$

where $f_r \in \mathbb{R}[x]_{2(d+r)}$ is any polynomial strictly positive on \mathcal{X} , and again $d_j = \lceil \deg g_j / 2 \rceil$, $j = 1, \dots, m$. We check whether the optimal solution \mathbf{y}_r^* to (3.29) satisfies the rank condition (1.16), i.e.,

$$\text{rank } \mathbf{M}_{d+r}(\mathbf{y}_r^*) = \text{rank } \mathbf{M}_{d+r-d_{\mathcal{X}}}(\mathbf{y}_r^*), \quad (3.30)$$

where $d_{\mathcal{X}} := \max_j d_j$. As mentioned in Section 1.4.1, if (3.30) holds, then \mathbf{y}_r^* is the moment sequence of a measure supported on \mathcal{X} . If the test is passed, then we stop, otherwise we increase r by one and repeat the procedure. As $\mathbf{y}^* \in \mathcal{M}_{2d}(\mathcal{X})$, the rank condition (3.30) is satisfied for a sufficiently large value of r .

We extract the support points $\mathbf{x}_1, \dots, \mathbf{x}_\ell \in \mathcal{X}$ of the representing atomic measure of \mathbf{y}_r^* , and \mathbf{y}^* respectively, as described in [Las10, Section 4.3].

Our experience has revealed that in most cases the polynomial $f_r(x) = \sum_{|\alpha| \leq d+r} x^{2\alpha} = \|\mathbf{v}_{d+r}(x)\|_2^2$ is a good choice. In problem (3.29) this corresponds to minimizing the trace of $\mathbf{M}_{d+r}(\mathbf{y})$ and hence leads to an optimal solution \mathbf{y} with low rank matrix $\mathbf{M}_{d+r}(\mathbf{y})$.

3.4.2 Via Christoffel-like polynomials

Another possibility to recover the atomic representing measure of an optimal solution \mathbf{y}^* to (3.19) is to find the zeros of the polynomial $p^*(x) = \text{trace}(\mathbf{M}_d(\mathbf{y}^*)^q) - p_d^*(x)$, where p_d^* is the Christoffel-like polynomial associated with \mathbf{y}^* defined in (3.15). In other words, we compute the set $\Omega = \{\mathbf{x} \in \mathcal{X} : \text{trace}(\mathbf{M}_d(\mathbf{y}^*)^q) - p_d^*(\mathbf{x}) = 0\}$, which due to Theorem 3.3 is the support of the atomic representing measure. To that end we minimize p^* on \mathcal{X} . As the polynomial p^* is non-negative on \mathcal{X} , the minimizers are exactly Ω . For minimizing p^* , we use the Lasserre hierarchy of lower bounds, that is, we solve the semidefinite program

$$\begin{aligned} \min_{\mathbf{y}_r} \quad & L_{\mathbf{y}_r}(p^*) \\ \text{s.t.} \quad & \mathbf{y}_r \in \mathcal{M}_{2(d+r)}(g_1, \dots, g_m), \quad y_{r,0} = 1. \end{aligned} \quad (3.31)$$

By Theorem 3.3, the polynomial p^* belongs to the quadratic module $\mathcal{Q}_{2(d+\delta)}^{2d}(g_1, \dots, g_m)$. Thus, the value of problem (3.31) is zero, i.e., optimal, for all $r \geq \delta$.

When condition (3.30) is fulfilled, the optimal solution \mathbf{y}_r^* comes from a measure. We extract the support points $\mathbf{x}_1, \dots, \mathbf{x}_\ell \in \mathcal{X}$ of the representing atomic measure of \mathbf{y}_r^* , and \mathbf{y}^* respectively, as described in [Las10, Section 4.3].

Alternatively, instead of solving problem (3.31), we can solve the semidefinite program

$$\begin{aligned} \min_{\mathbf{y}_r} \quad & \text{trace}(\mathbf{M}_{d+r}(\mathbf{y}_r)) \\ \text{s.t.} \quad & L_{\mathbf{y}_r}(p^*) = 0, \quad y_{r,0} = 1, \\ & \mathbf{y}_r \in \mathcal{M}_{2(d+r)}(g_1, \dots, g_m), \end{aligned} \quad (3.32)$$

which also searches for a moment sequence of a measure supported on the zero level set of p^* . Again, if condition (3.30) is holds, the finite support can be extracted.

3.4.3 Calculating the corresponding weights

After recovering the support $\{\mathbf{x}_1, \dots, \mathbf{x}_\ell\}$ of the atomic representing measure by one of the previously presented methods, we might be interested in also computing the corresponding weights w_1, \dots, w_ℓ . These can be calculated easily by solving the following linear system of equations: $\sum_{i=1}^{\ell} w_i \mathbf{x}_i^\alpha = y_\alpha^*$ for all $|\alpha| \leq 2d$, i.e., $\int_{\mathcal{X}} x^\alpha \mu^*(dx) = y_\alpha^*$.

3.5 Examples

We illustrate the procedure on three examples: a univariate one, a polygon in the plane and one example on the three-dimensional sphere. We concentrate on D -optimal designs ($q = 0$) and T -optimal designs ($q = 1$). We do not report computation times, since they are negligible for our small examples.

3.5.1 Univariate unit interval

We consider as design space the interval $\mathcal{X} = [-1, 1]$ and on it the polynomial measurements $\sum_{j=0}^d \theta_j x^j$ with unknown parameters $\theta \in \mathbb{R}^{d+1}$.

3.5.1.1 D-optimal design

To compute the D -optimal design we first solve problem (3.19). As this includes the projection of the cone $\mathcal{M}_{2d}(1 - \|x\|^2)$, this is equivalent to solving

$$\begin{aligned} \max_{\mathbf{y}_\delta} \quad & \log \det \mathbf{M}_d(\mathbf{y}_\delta) \\ \text{s.t.} \quad & \mathbf{M}_{d+\delta}(\mathbf{y}_\delta) \succcurlyeq 0, \\ & \mathbf{M}_{d+\delta-1}((1-x^2)\mathbf{y}_\delta) \succcurlyeq 0, \\ & y_{\delta,0} = 1 \end{aligned} \tag{3.33}$$

for $\mathbf{y}_\delta \in \mathbb{R}^{s(2(d+\delta))}$ and given regression order d and relaxation order $d + \delta$, and then taking the truncation $\mathbf{y}^* := (y_{\delta,\alpha}^*)_{|\alpha| \leq 2d}$ of an optimal solution \mathbf{y}_δ^* . For instance, for $d = 5$ and $\delta = 0$ we obtain the sequence $\mathbf{y}^* \approx (1, 0, 0.56, 0, 0.45, 0, 0.40, 0, 0.37, 0, 0.36)^T$.

To recover the corresponding atomic measure from the sequence \mathbf{y}^* we solve problem (3.29),

$$\begin{aligned} \min_{\mathbf{y}_r} \quad & \text{trace } \mathbf{M}_{d+r}(\mathbf{y}_r) \\ \text{s.t.} \quad & \mathbf{M}_{d+r}(\mathbf{y}_r) \succcurlyeq 0, \\ & \mathbf{M}_{d+r-1}((1-x^2)\mathbf{y}_r) \succcurlyeq 0, \\ & y_{r,\alpha} = y_\alpha^*, \quad |\alpha| \leq 2d. \end{aligned} \tag{3.34}$$

We find the atomic measure supported on the points $-1, -0.765, -0.285, 0.285, 0.765$ and 1 (for $d = 5, \delta=0, r = 1$). As a result, our optimal design is the weighted sum of the Dirac measures supported on these points. The points match with the known analytic solution to the problem, which are the critical points of the Legendre polynomial, see, e.g., [DS97, Theorem 5.5.3]. Calculating the corresponding weights as described in Section 3.4.3, we find $w_1 = \dots = w_6 \approx 0.166$.

Alternatively, we compute the roots of the polynomial $p^*(x) = 6 - p_5^*(x)$, where p_5^* is the Christoffel polynomial of degree $2d = 10$ on \mathcal{X} associated with \mathbf{y}^* . We find the same points as in the previous approach by solving problem (3.32). See Figure 3.1 for the graph of the polynomial p^* .

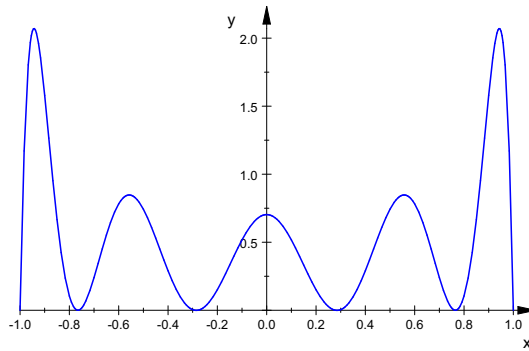


Figure 3.1 – Polynomial p^* for Example 3.5.1.1.

We observe that we get less points when using problem (3.31) to recover the support for this example. This may occur due to numerical issues.

3.5.1.2 T-optimal design

For the T -optimal design, instead of solving (3.33) we solve

$$\begin{aligned} \max_{\mathbf{y}_\delta} \quad & \text{trace } \mathbf{M}_d(\mathbf{y}_\delta) \\ \text{s.t.} \quad & \mathbf{M}_{d+\delta}(\mathbf{y}_\delta) \succcurlyeq 0, \\ & \mathbf{M}_{d+\delta-1}((1-x^2)\mathbf{y}_\delta) \succcurlyeq 0, \\ & y_{\delta,0} = 1 \end{aligned}$$

for $\mathbf{y}_\delta \in \mathbb{R}^{s(2(d+\delta))}$ and given d and δ . Then we take $\mathbf{y}^* := (y_{\delta,\alpha}^*)_{|\alpha| \leq 2d}$. For example, for $d = 3$ and $\delta = 0$ we get the sequence $\mathbf{y}^* \approx (1, 0, 1, 0, 1, 0, 1)^T$.

To recover the corresponding measure we solve problem (3.34) as before. In contrary to the D -optimal design, we get the points -1 and 1 with weights $w_1 = w_2 = 0.5$ for all $d \leq 7$, meaning our T -optimal design is the measure $\mu^* = 0.5 \delta_{-1} + 0.5 \delta_1$ independently of the regression order d . The measure μ^* is also the D -optimal design for regression order $d = 1$, but for the D -optimal design the support gets larger when d increases.

Note that we can also compute the support points via Christoffel-like polynomials. Evidently, this gives the same points. Generally, in the case of T -optimal design the method via Christoffel-like polynomials is numerically more stable than in the D -optimal design case, since $p_d^*(x) = \mathbf{v}_d(x)^T \mathbf{v}_d(x)$ for $q = 1$. Hence, p_d^* is independent of the moment matrix and it is not necessary to compute its inverse or the orthogonal polynomials.

3.5.2 Wynn's polygon

As a two-dimensional example we take the polygon given by the vertices $(-1, -1)$, $(-1, 1)$, $(1, -1)$ and $(2, 2)$, scaled to fit the unit circle, i.e., we consider the design space

$$\mathcal{X} = \{\mathbf{x} \in \mathbb{R}^2 : x_1, x_2 \geq -\frac{1}{4}\sqrt{2}, x_1 \leq \frac{1}{3}(x_2 + \sqrt{2}), x_2 \leq \frac{1}{3}(x_1 + \sqrt{2}), x_1^2 + x_2^2 \leq 1\}.$$

Note that we need the redundant constraint $x_1^2 + x_2^2 \leq 1$ in order to have an algebraic certificate of compactness.

3.5.2.1 D-optimal design

As before, in order to find the D -optimal design for the regression, we solve problems (3.19) and (3.29). Let us start by analyzing the results for $d = 1$ and $\delta = 3$. Solving (3.19) we obtain $\mathbf{y}^* \in \mathbb{R}^{45}$, which leads to 4 atoms when solving (3.29) with $r = 3$. For the latter the moment matrices of order 2 and 3 both have rank 4, so condition (3.30) is fulfilled. As expected, the 4 atoms are exactly the vertices of the polygon.

Again, by solving problem (3.32) instead of (3.29) we obtain the same atoms. As in the univariate example we get less points when using problem (3.31). To be precise, GloptiPoly is not able to extract any solutions for this example.

For increasing d , we get an optimal measure with a larger support. For $d = 2$ we recover 7 points, and 13 for $d = 3$. See Figure 3.2 for the polygon, the supporting points of the optimal measure and the $\binom{2+d}{2}$ -level set of the Christoffel polynomial p_d^* for different d . The latter demonstrates graphically that the set of zeros of $\binom{2+d}{d} - p_d^*$ intersected with \mathcal{X} are indeed the atoms of our representing measure. In Figure 3.3 we visualized the weights corresponding to each point of the support for the different d .

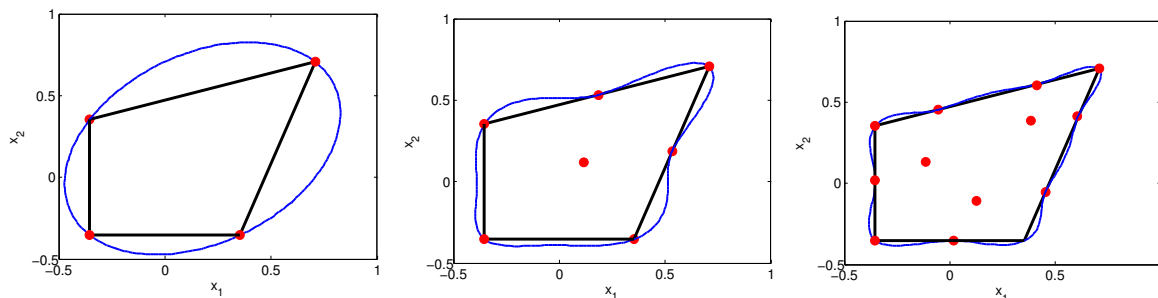


Figure 3.2 – The polygon (bold black) of Section 3.5.2, the support of the optimal design measure (red points) and the $s(d)$ -level set of the Christoffel polynomial (thin blue) for $d = 1$ (left), $d = 2$ (middle), $d = 3$ (right) and $\delta = 3$.

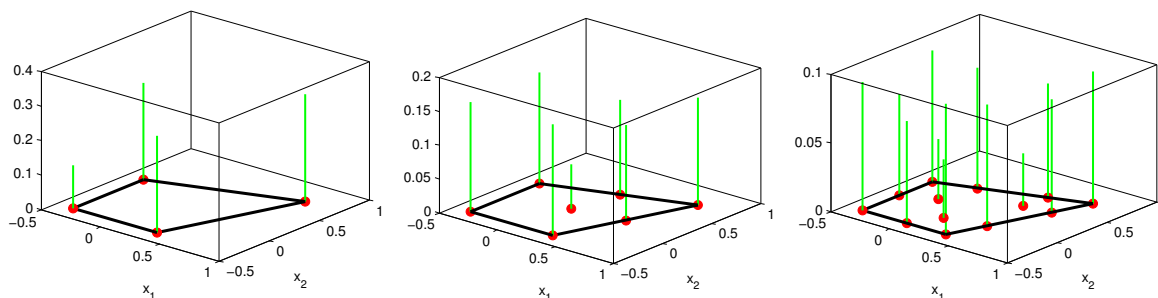


Figure 3.3 – The polygon (bold black) of Section 3.5.2 and the support of the optimal design measure (red points) with the corresponding weights (green bars) for $d = 1$ (left), $d = 2$ (middle), $d = 3$ (right) and $\delta = 3$.

3.5.2.2 T-optimal design

For the T -optimal design, meaning with the trace as objective function, we obtain a moment sequence $\mathbf{y}^* \in \mathbb{R}^{45}$ with a rank 1 moment matrix when solving problem (3.19) for $d = 1$ and $\delta = 3$. Hence, when recovering the corresponding measure, we obtain a Dirac measure supported on one point, namely the point $(0.7, 0.7)$, which is the right most vertex of the polygon. For $d = 2$ and $d = 3$ we also obtain moment matrices of rank one and recover the same support point. See Figure 3.4 for an illustration of the recovered point and the respective zero level set of the polynomial p^* .

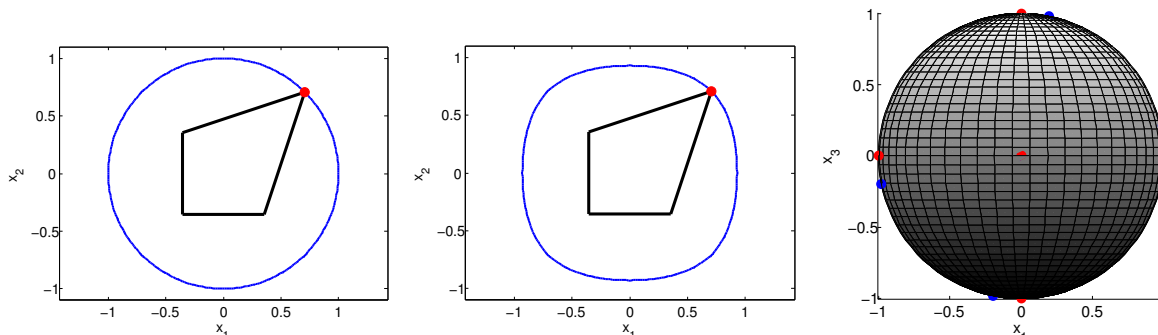


Figure 3.4 – The two pictures showing the polygon illustrate the recovered point (red) and the zero level set of the polynomial p^* (blue line) for the T -optimal design example of Section 3.5.2; we have $d = 1$ on the left and $d = 3$ in the middle. The picture on the right shows the support points recovered in Example 3.5.3.1 for $d = 1$ (red) and the points which are recovered when additionally fixing some moments as described in Example 3.5.3.3 (blue).

3.5.3 The 3-dimensional unit sphere

Last, let us consider the regression for the degree d polynomial measurements $\sum_{|\alpha| \leq d} \theta_\alpha x^\alpha$ on the unit sphere $\mathcal{X} = \{\mathbf{x} \in \mathbb{R}^3 : x_1^2 + x_2^2 + x_3^2 = 1\}$.

3.5.3.1 D-optimal design

Again, we first solve problem (3.19). For $d = 1$ and $\delta \geq 0$ we obtain the sequence $\mathbf{y}^* \in \mathbb{R}^{10}$ with $y_{000}^* = 1$, $y_{200}^* = y_{020}^* = y_{002}^* = 0.333$ and all other entries zero.

In the second step we solve problem (3.29) to recover the corresponding measure. For $r = 2$ the moment matrices of order 2 and 3 both have rank 6, meaning the rank condition (3.30) is fulfilled. We obtain the six atoms $\{(\pm 1, 0, 0), (0, \pm 1, 0), (0, 0, \pm 1)\} \subseteq \mathcal{X}$, on which the optimal measure $\mu \in \mathcal{M}_+(\mathcal{X})$ is uniformly supported.

For quadratic regressions, i.e., $d = 2$, we obtain an optimal measure supported on 14 atoms evenly distributed on the sphere. Choosing $d = 3$, meaning cubic regressions, we find a Dirac measure supported on 26 points, which again are evenly distributed on the sphere. See Figure 3.5 for an illustration of the supporting points of the optimal measures for $d = 1$, $d = 2$, $d = 3$ and $\delta = 0$.

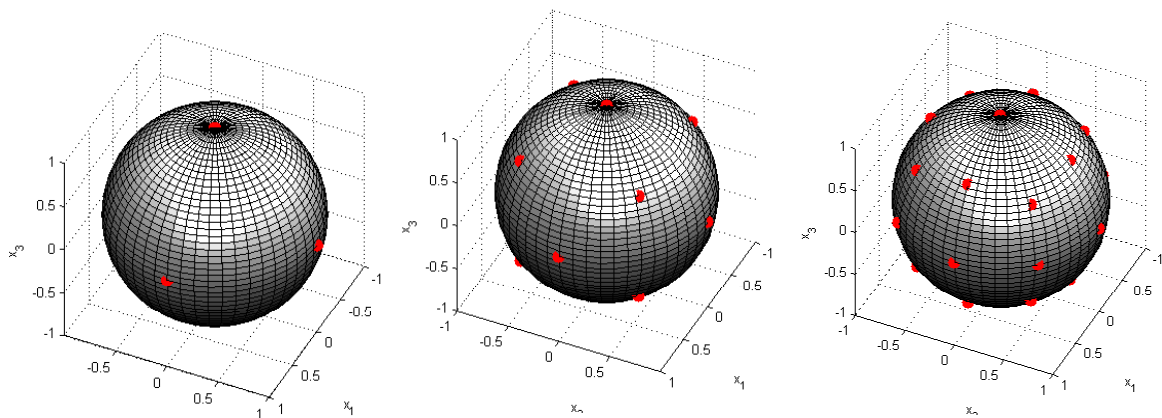


Figure 3.5 – The red points illustrate the support of the optimal design for $d = 1$ (left), $d = 2$ (middle), $d = 3$ (right) and $\delta = 0$ for Example 3.5.3.1.

Using the method via Christoffel-like polynomials gives again less points. No solution is extracted when solving problem (3.32) and we find only two supporting points for problem (3.31).

3.5.3.2 T-optimal design

For this example, the solutions for the D -optimal design and the T -optimal design coincide for regression order $d = 1$. As for the other examples, here again the T -optimal design gives the same solution for all d , so we obtain the same six support points as for $d = 1$ when choosing $d = 2$ or $d = 3$.

3.5.3.3 Fixing some moments

Our method has an additional nice feature. Indeed, in problem (3.19) one may easily include the additional constraint that some moments (y_α) , $\alpha \in \Gamma \subseteq \mathbb{N}_{2d}^n$, are fixed to some prescribed value. We illustrate this potential on one example. For instance, for $\Gamma = \{(020), (002), (110), (101)\}$ set $y_{020} := 2$, $y_{002} := 1$, $y_{110} := 0.01$ and $y_{101} := 0.95$. In order to obtain a feasible problem, we scale them with respect to the Gauss distribution.

For the D -optimal design case with $d = 1$ and $\delta = 0$ and after computing the support of the corresponding measure using the Nie method, we get 6 points as we obtain without fixing the moments. However, now four of the six points are shifted and the measure is no longer uniformly supported on these points. Still, each two opposite points have the same weight. See the picture on the right hand side of Figure 3.4 for an illustration of the position of the points with fixed moments with respect to the position of the support points without fixing the points.

3.6 Concluding remark

In this chapter, we gave a general method to compute optimal designs for multidimensional polynomial regression on semialgebraic sets. The Lasserre hierarchy of lower bounds was used in both steps of our method, both for the computation of the moment sequence of an optimal design and to recover the corresponding measure.

The considered method is highly versatile as it can be used for all classical functionals of the information matrix. Furthermore, it can easily be tailored to incorporate prior knowledge on some multidimensional moments of the targeted optimal measure (as proposed in [MZ04]). In future works, one may extend the method to multi-response polynomial regression problems or to general smooth parametric regression models by linearization.

Approximating semialgebraic sets with quadratic separators

Contents

4.1	Separators of arbitrary degree	64
4.2	Linear separators	65
4.2.1	Exactness	65
4.2.2	Examples	66
4.3	Quadratic separators	67
4.3.1	Exactness	68
4.3.2	Examples	68
4.4	Concluding remarks	69

In this chapter we exploit further a concept on which the Lasserre hierarchy is based. We use it to approximate a compact semialgebraic set \mathbf{K} with simple objects, that is, quadratic polynomials. In other words, we express the set \mathbf{K} described by polynomials of potentially large degree with (infinitely many) polynomials of degree 2.

We call a polynomial a *separator*, if its zero sublevel set intersected with \mathbf{K} is empty. So, the objective is to obtain approximations of \mathbf{K} by quadratic separators, meaning separators of degree 2.

In [Las09], Lasserre approximates the convex hull $\text{co}(\mathbf{K})$ of a semialgebraic set \mathbf{K} by intersections of zero superlevel sets of linear separators. Related to this concept is the notion of theta bodies treated in [GPT10] which are defined for sets \mathbf{K} which are algebraic varieties. These outer approximations of the convex hull motivated further investigations, namely to consider separators of higher degree, for example of degree 2. The idea is to obtain outer approximations of the set \mathbf{K} itself, in contrast to approximating the convex hull in the linear case.

The concept of approximating a semialgebraic set by polynomial separators is closely related to the Lasserre hierarchy. To see why, we first recall that $\mathcal{P}(\mathbf{K})$ denotes the cone of polynomials which are non-negative on \mathbf{K} , and observe that polynomial separators are elements of $\mathcal{P}(\mathbf{K})$. Now we are in the situation which we have already encountered in the first chapter, since we need a tractable characterization of the cone $\mathcal{P}(\mathbf{K})$. This is where we use the idea by Lasserre, and instead of taking $\mathcal{P}(\mathbf{K})$, we take the quadratic module $\mathcal{Q}(g_1, \dots, g_m)$ generated by the polynomials g_j defining \mathbf{K} , which we then truncate.

As a result, we get approximations which are described by semidefinite constraints of smaller size, since generally the original polynomials g_j are no longer part of the description. Hence, given a point $\mathbf{x} \in \mathbb{R}^n$, we can decide whether \mathbf{x} is an element of the approximation set solving a small semidefinite program.

The outline of the chapter is as follows. First, we introduce the general concept by considering separators of arbitrary degree. Then we present the existing results on approximations of the convex hull using linear separators, and continue with our results about quadratic separators

which show that these indeed approximate the set \mathbf{K} itself. After some illustrating examples, we conclude the chapter with some remarks on the approach, especially on the difficulties we encountered and which may require further investigations.

4.1 Separators of arbitrary degree

We consider the compact semialgebraic set

$$\mathbf{K} = \{\mathbf{x} \in \mathbb{R}^n : g_j(\mathbf{x}) \geq 0, j = 1, \dots, m\}$$

described by given polynomials $g_j \in \mathbb{R}[x]$, $j = 1, \dots, m$. As in (1.1), we assume that one of the g_j is of the form $R - \sum_{i=1}^n x_i^2$ for some $R > 0$. Additionally, we assume that \mathbf{K} has non-empty interior.

Remark. In Chapter 1 we have seen that the former assumption implies that the quadratic module $\mathcal{Q}(g_1, \dots, g_m)$ generated by the polynomials g_j is archimedean. Consequently, Putinar's Positivstellensatz is applicable.

In the chapter on hand, let us abbreviate the quadratic module $\mathcal{Q}(g_1, \dots, g_m)$ by \mathcal{Q} . We also omit the dependence on the g_j in the notation for the truncated quadratic module and write

$$\mathcal{Q}_{2r} := \mathcal{Q}_{2r}(g_1, \dots, g_m) = \left\{ \sigma_0 + \sum_{j=1}^m \sigma_j g_j : \sigma_j \in \Sigma[x]_{2(d-d_j)}, j = 0, \dots, m \right\},$$

where again $d_0 = 0$ and $d_j := \lceil \deg(g_j)/2 \rceil$, $j = 1, \dots, m$. For given $r, t \in \mathbb{N}$, we define the set

$$\mathbf{K}_{r,t} := \{\mathbf{x} \in \mathbb{R}^n : p(\mathbf{x}) \geq 0 \text{ for all } p \in \mathcal{Q}_{2r} \cap \mathbb{R}[x]_t\}. \quad (4.1)$$

Since $\mathcal{Q}_{2r} \cap \mathbb{R}[x]_t \subseteq \mathcal{P}(\mathbf{K})$, the zero sublevel set of a polynomial $p \in \mathcal{Q}_{2r} \cap \mathbb{R}[x]_t$ intersected with \mathbf{K} is empty, and therefore p is a separator of degree t . Hence, the set $\mathbf{K}_{r,t}$ is the intersection of zero superlevel sets of polynomial separators of degree t .

Remark. Obviously $\mathcal{Q}_t \subseteq \mathcal{Q}_{2r} \cap \mathbb{R}[x]_t$ for $2r \geq t$, but in general the inclusion is strict due to possible cancellations of the higher degree terms. This strict inclusion is the reason why the approach is successful even if there are g_j with $\deg g_j \geq t$.

For $\mathbf{x} \in \mathbf{K}$ we always have $p(\mathbf{x}) \geq 0$ for all $p \in \mathcal{Q}_{2r} \cap \mathbb{R}[x]_t$, so $\mathbf{K} \subseteq \mathbf{K}_{r,t}$ for all $r, t \in \mathbb{N}$. Hence, $\mathbf{K}_{r,t}$ is an outer approximation, or relaxation, of \mathbf{K} .

Furthermore, $\mathbf{K}_{r,t+1} \subseteq \mathbf{K}_{r,t}$ for all r and t , since the inclusion $\mathcal{Q}_{2r} \cap \mathbb{R}[x]_t \subseteq \mathcal{Q}_{2r} \cap \mathbb{R}[x]_{t+1}$ implies that an element of $\mathbf{K}_{r,t+1}$ fulfills the same constraints as $\mathbf{K}_{r,t}$, and potentially more. In the same way, it follows from the inclusion $\mathcal{Q}_{2r} \cap \mathbb{R}[x]_t \subseteq \mathcal{Q}_{2(r+1)} \cap \mathbb{R}[x]_t$ that $\mathbf{K}_{r+1,t} \subseteq \mathbf{K}_{r,t}$ for all r and t . Thus, the relaxations become tighter when r or t increases.

The moment side

As seen in the first chapter, the dual of the truncated quadratic module \mathcal{Q}_{2r} is the truncated moment cone

$$\mathcal{M}_{2r} := \mathcal{M}_{2r}(g_1, \dots, g_m) := \{\mathbf{y} \in \mathbb{R}^{s(2r)} : \mathbf{M}_r(\mathbf{y}) \succeq 0, \mathbf{M}_{r-d_j}(g_j \mathbf{y}) \succeq 0, j = 1, \dots, m\},$$

where $s(2r) := \binom{n+2r}{n}$, $d_j := \lceil \deg(g_j)/2 \rceil$, $j = 1, \dots, m$, and $r \geq \max\{d_1, \dots, d_m\}$. As on the SOS side, we also suppress the dependence on the g_j in the notation of the infinite dimensional moment cone and write $\mathcal{M} := \mathcal{M}(g_1, \dots, g_m)$.

In the following we show that we can express the set $\mathbf{K}_{r,t}$ in terms of moments:

Lemma 4.1. *Let $r, t \in \mathbb{N}$, and let $\mathbf{K}_{r,t}$ be the set defined in (4.1). Then*

$$\mathbf{K}_{r,t} = \{\mathbf{x} \in \mathbb{R}^n : \exists \mathbf{y} \in \mathcal{M}_{2r} \text{ such that } y_\alpha = \mathbf{x}^\alpha \text{ for all } |\alpha| \leq t\}. \quad (4.2)$$

Proof. To see that the right hand side of the equation is indeed the same set as the set defined in (4.1), let us denote the set described via polynomial separators (4.1) by \mathbf{A} and the description via moments (4.2) by \mathbf{B} .

“ $\mathbf{A} \subseteq \mathbf{B}$ ”: Let $\mathbf{x} \in \mathbf{A}$, i.e. $p(\mathbf{x}) \geq 0$ for all $p \in \mathcal{Q}_{2r} \cap \mathbb{R}[x]_t$. Define $y_\alpha^* := \mathbf{x}^\alpha$ for $|\alpha| \leq t$, and write $\mathbf{y}^* = (y_\alpha^*)_{|\alpha| \leq t}$. Then $0 \leq p(\mathbf{x}) = \sum_{|\alpha| \leq t} p_\alpha \mathbf{x}^\alpha = \sum_{|\alpha| \leq t} p_\alpha y_\alpha^* = \langle \mathbf{p}, \mathbf{y}^* \rangle$ for all $p \in \mathcal{Q}_{2r} \cap \mathbb{R}[x]_t$ with $\mathbf{p} = (p_\alpha)_\alpha$ denoting the vector of coefficients of p . So \mathbf{y}^* is an element of the dual cone of $\mathcal{Q}_{2r} \cap \mathbb{R}[x]_t$ which is the projection of \mathcal{M}_{2r} onto the first $s(t) = \binom{n+t}{n}$ entries. Hence, there exists a lifting $\mathbf{y} \in \mathcal{M}_{2r}$ of \mathbf{y}^* with $y_\alpha = y_\alpha^*$ for $|\alpha| \leq t$. Since by definition $y_\alpha^* = \mathbf{x}^\alpha$ for $|\alpha| \leq t$, we have $\mathbf{x} \in \mathbf{B}$.

“ $\mathbf{B} \subseteq \mathbf{A}$ ”: Let $\mathbf{x} \in \mathbf{B}$, i.e. there exists $\mathbf{y}^* \in \mathcal{M}_{2r}$ such that $y_\alpha^* = \mathbf{x}^\alpha$ for all $|\alpha| \leq t$. Let $p \in \mathcal{Q}_{2r} \cap \mathbb{R}[x]_t$. Then $p(\mathbf{x}) = \sum_{|\alpha| \leq t} p_\alpha \mathbf{x}^\alpha = \sum_{|\alpha| \leq t} p_\alpha y_\alpha^* = \langle \mathbf{p}, \mathbf{y}^* \rangle \geq 0$, since $p \in \mathcal{Q}_{2r}$ and $\mathbf{y}^* \in \mathcal{M}_{2r}$. Hence, $\mathbf{x} \in \mathbf{A}$. \square

4.2 Linear separators

The case of linear separators (treated in [Las09]) corresponds to approximating \mathbf{K} with sets of the form $\mathbf{K}_{r,1}$, i.e., we let $t = 1$ and we let r increase. It turns out that this generates relaxations of $\text{co}(\mathbf{K})$, the convex hull of \mathbf{K} , defined as the smallest convex set containing \mathbf{K} .

Definition 4.2. We call the sets $\mathbf{K}_{r,1}$ linear relaxations of \mathbf{K} .

The following theorem formally states the above mentioned, i.e. that the linear relaxations $\mathbf{K}_{r,1}$ form a converging sequence of outer approximations of $\text{co}(\mathbf{K})$. The theorem is a direct consequence of Putinar’s Positivstellensatz, Theorem 1.1.

Theorem 4.3. $\text{co}(\mathbf{K}) = \bigcap_{r \in \mathbb{N}} \mathbf{K}_{r,1}$.

Proof. We follow the proof of [Las09, Theorem 2], respectively [Sin10, Corollary 2.2.8].

“ \subseteq ”: Choose a point $\mathbf{x} \in \text{co}(\mathbf{K})$. By definition of the convex hull, it holds that $\mathbf{x} = \int x \mu(dx)$ for some probability measure $\mu \in \mathcal{M}_+(\mathbf{K})$. Let $\mathbf{y} = (y_\alpha)_\alpha \in \mathbb{R}^{\mathbb{N}^n}$ be the sequence of moments of μ , i.e. $y_\alpha = \int x^\alpha \mu(dx)$, $\alpha \in \mathbb{N}^n$. The sequence \mathbf{y} is well-defined because μ has compact support, and since we are considering probability measures, we have $y_0 = 1$. Furthermore, $\mathbf{M}_r(\mathbf{y}) \succcurlyeq 0$ and $\mathbf{M}_{r-d_j}(g_j \mathbf{y}) \succcurlyeq 0$ for all $j = 1, \dots, m$ and $r \in \mathbb{N}$ as \mathbf{y} is the moment sequence of a measure supported on \mathbf{K} . This means that $\mathbf{y} \in \mathcal{M}_{2r}$. Moreover, $\int x_i d\mu = x_i$, $i = 1, \dots, n$, by the definition of μ , which implies $\mathbf{x} \in \bigcap_{r \in \mathbb{N}} \mathbf{K}_{r,1}$.

“ \supseteq ” by contradiction: Assume there is $\mathbf{x} \in \bigcap_{r \in \mathbb{N}} \mathbf{K}_{r,1} \setminus \text{co}(\mathbf{K})$. Then, as $\text{co}(\mathbf{K})$ is convex and compact, by the separation theorem for convex sets (see e.g. [Bar02, Theorem III.1.3]) there exists a linear polynomial $p_{\mathbf{x}} \in \mathbb{R}[x]_1$ such that $p_{\mathbf{x}}(\mathbf{x}) < 0$ and $p_{\mathbf{x}}(\mathbf{z}) > 0$ for all $\mathbf{z} \in \text{co}(\mathbf{K})$. In particular, $p_{\mathbf{x}} > 0$ on \mathbf{K} , so Putinar’s Positivstellensatz implies $p_{\mathbf{x}} \in \mathcal{Q}$ which contradicts $\mathbf{x} \in \bigcap_{r \in \mathbb{N}} \mathbf{K}_{r,1}$, meaning $p(\mathbf{x}) \geq 0$ for all $p \in \mathcal{Q} \cap \mathbb{R}[x]_1$. \square

4.2.1 Exactness

An important question for practical purposes is, in which cases there exists a finite $r \in \mathbb{N}$ such that $\mathbf{K}_{r,1} = \text{co}(\mathbf{K})$. For this we introduce the notion of exactness.

Definition 4.4. We say that the linear relaxation $\mathbf{K}_{r,1}$ of order $r \in \mathbb{N}$ is *exact*, if $\mathbf{K}_{r,1} = \text{co}(\mathbf{K})$.

Remark. The existence of a finite r such that the relaxation of order r is exact implies that the set $\text{co}(\mathbf{K})$ is semidefinite representable. For details and further references consult [Las09], [NPS10] and [KS17]. Note additionally that by the recent result by Scheiderer [Sch16], there are indeed many convex semialgebraic sets which do not have a semidefinite representation.

In [Las09], Lasserre gives a sufficient condition for exactness, namely, if there exists $r^* \in \mathbb{N}$ such that \mathcal{Q}_{2r^*} contains all linear polynomials which are non-negative on \mathbf{K} , then the relaxation $\mathbf{K}_{r^*,1}$ is exact [Las09, Theorem 2b]. This condition is also necessary by [NPS10, Proposition 3.1].

As explained in [Las09, Examples 1 and 2], the condition is fulfilled for $r^* = 1$ for convex polytopes defined by linear inequalities, and also for sets \mathbf{K} described by g_j 's which are concave and quadratic. For convex sets $\mathbf{K} \subseteq \mathbb{R}^2$ for which the g_j 's are concave with degree 2 or 4, the condition is fulfilled for $r^* = 2$, [Las09, Example 2]. More generally, [Las09, Theorem 3] says that the above property holds with order $r^* = 1$, if $\mathbf{K} \subseteq \mathbb{R}^n$ is defined by two polynomials g_1 and g_2 of the form $g_i(x) = x^T A_i x + c_i$ for some real symmetric matrices A_i and $c_i \in \mathbb{R}$, $i = 1, 2$. Note that this includes sets \mathbf{K} which may be disconnected.

So, in these cases we know that the linear relaxation for the given order r^* is exact, meaning $\text{co}(\mathbf{K}) = \mathbf{K}_{r^*,1}$. In addition, very recently Kriel and Schweighofer showed the existence of exact linear relaxations for convex sets \mathbf{K} , where the defining polynomials g_1, \dots, g_m satisfy a certain concavity condition, see [KS17, Theorem 4.8].

Conversely, since the above mentioned sufficient condition is also necessary, we can produce planar examples for which $\text{co}(\mathbf{K}) \neq \mathbf{K}_{r,1}$ by choosing $\mathbf{K} = \{\mathbf{x} \in \mathbb{R}^2 : g(\mathbf{x}) \geq 0\}$ defined by one single polynomial g which does not have any linear term. Since this implies that the truncated quadratic module \mathcal{Q}_{2r} does not contain all affine functions which are non-negative on \mathbf{K} , the condition is violated. [Hen08] assembles examples of planar quartics which illustrate this approach. We cite two of them in Example 4.2.1 and Example 4.2.2.

Furthermore, in [NPS10, Theorem 3.5], it is shown that in case \mathbf{K} is convex and has a non-exposed face, then $\mathbf{K}_{r,1}$ cannot be exact for any finite $r \in \mathbb{N}$. This is illustrated in Example 4.2.3.

4.2.2 Examples

Example 4.2.1 (Water drop, [Hen08, Example 5.3]). Let $\mathbf{K} \subseteq \mathbb{R}^2$ be the set defined by the water drop quartic $g(x) = -x_1^2 - x_2^2 - (x_1^2 + x_2^2)^2$. The curve has a singular point at the origin. On the left hand side of Figure 4.1 we see that the sets $\mathbf{K}_{r,1}$ for $r = 2, 3$ and 4 are indeed strict supersets of $\text{co}(\mathbf{K})$. For increasing r the sets are no longer visually distinguishable from $\text{co}(\mathbf{K})$, but one can show that the linear relaxation is not exact for any finite r .

Example 4.2.2 (Folium, [Hen08, Example 5.5]). Let $g(x) = -x_1(x_1^2 - 2x_2^2) - (x_1^2 + x_2^2)^2$. This polynomial defines a folium which is a curve of genus zero with a triple singular point at the origin. See Figure 4.1 (right side) for an illustration of the sets $\mathbf{K}_{2,1}$ and $\mathbf{K}_{3,1}$. While it is obvious that the linear relaxation of order 2 is not exact, the one of order 3 might be, but we do not know.

Example 4.2.3 (Non-exposed face, [NPS10, Example 3.7]). Consider the convex set $\mathbf{K} \subseteq \mathbb{R}^2$ defined by $g_1(x) = x_2 - x_1^3$, $g_2(x) = x_1 + 1$, $g_3(x) = x_2$ and $g_4(x) = 1 - x_2$. In [NPS10], it is shown that the point $(0, 0)$ is a non-exposed face of \mathbf{K} . Therefore, there is no finite r such that $\mathbf{K}_{r,1}$ is exact. See Figure 4.2 for an illustration of $\mathbf{K}_{2,1}$. Already the set $\mathbf{K}_{3,1}$ (not represented)

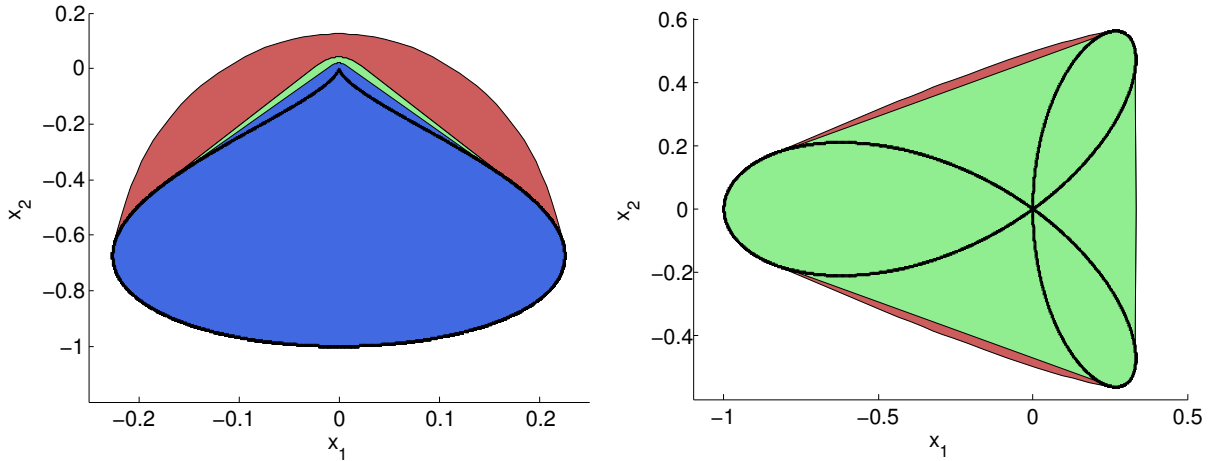
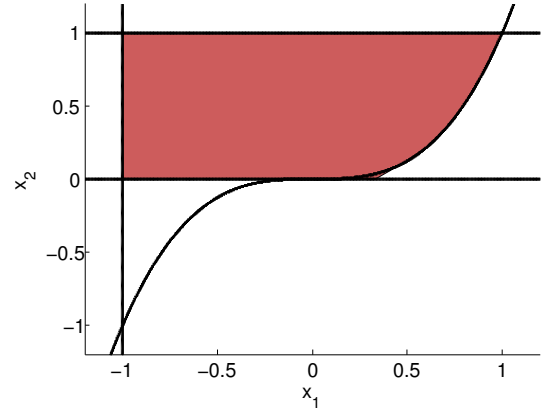


Figure 4.1 – The sets $\mathbf{K}_{2,1}$ (red), $\mathbf{K}_{3,1}$ (green) and $\mathbf{K}_{4,1}$ (blue, only left) for Example 4.2.1 on the left and Example 4.2.2 on the right. The black curves represent the zero level sets of the respective polynomials g .

is no longer visually distinguishable from \mathbf{K} . However, we know that it is a strict superset by [NPS10, Theorem 3.5].

Note that \mathbf{K} is nonetheless semidefinite representable, as is also shown in [NPS10]. However, its semidefinite representation cannot be found via the linear separators.

Figure 4.2 – The black curves represent the zero level sets of the polynomials g_1, g_2, g_3 and g_4 for Example 4.2.3. The red region corresponds to the set $\mathbf{K}_{2,1}$, meaning the approximation via linear separators of the smallest possible order. One can see that the approximation is not exact near the non-exposed face $(0, 0)$.



4.3 Quadratic separators

Let us denote by $d_{\mathbf{K}} := \max_{j=1, \dots, m} \deg g_j$ the maximum degree of the polynomials describing \mathbf{K} . Then we have $g_j \in \mathcal{Q} \cap \mathbb{R}[x]_{d_{\mathbf{K}}}$ for all $j = 1, \dots, m$, and hence $\mathbf{K} = \bigcap_{r \in \mathbb{N}} \mathbf{K}_{r, d_{\mathbf{K}}}$. More generally,

$$\mathbf{K} = \bigcap_{r \in \mathbb{N}} \mathbf{K}_{r, d_{\mathbf{K}}} \subseteq \dots \subseteq \bigcap_{r \in \mathbb{N}} \mathbf{K}_{r, 2} \subseteq \bigcap_{r \in \mathbb{N}} \mathbf{K}_{r, 1} = \text{co}(\mathbf{K}) \subseteq \bigcap_{r \in \mathbb{N}} \mathbf{K}_{r, 0} = \mathbb{R}^n, \quad (4.3)$$

since the relaxations $\mathbf{K}_{r,t}$ get tighter for increasing t , as explained in Section 4.1, and where the second equality is due to Theorem 4.3. In the following we consider approximations with quadratic separators.

Definition 4.5. We call the sets $\mathbf{K}_{r,2}$ quadratic relaxations of \mathbf{K} .

Now we show that there is actually a big jump from using linear separators ($t = 1$), which model the convex hull of \mathbf{K} as shown in Theorem 4.3, to using quadratic separators ($t = 2$). In

fact, quadratic relaxations are already enough to model exactly the set \mathbf{K} .

Theorem 4.6. $\mathbf{K} = \bigcap_{r \in \mathbb{N}} \mathbf{K}_{r,2}$.

Proof. The inclusion “ \subseteq ” holds due to (4.3). We prove “ \supseteq ” by contraposition, that is, we show $\hat{\mathbf{x}} \notin \mathbf{K} \Rightarrow \hat{\mathbf{x}} \notin \bigcap_{r \in \mathbb{N}} \mathbf{K}_{r,2}$.

Let $\hat{\mathbf{x}} \in \mathbb{R}^n \setminus \mathbf{K}$, i.e., there exists an open neighborhood of $\hat{\mathbf{x}}$ whose intersection with \mathbf{K} is empty. This means that there exists an $\varepsilon_{\hat{\mathbf{x}}} > 0$ such that $\|\mathbf{x} - \hat{\mathbf{x}}\|_2^2 \geq \varepsilon_{\hat{\mathbf{x}}}$ for all $\mathbf{x} \in \mathbf{K}$, i.e., for any $0 < \varepsilon < \varepsilon_{\hat{\mathbf{x}}}$ the polynomial

$$p(x) := \|x - \hat{\mathbf{x}}\|_2^2 - \varepsilon$$

is strictly positive on \mathbf{K} . By Putinar’s Positivstellensatz (Theorem 1.1), it follows that $p \in \mathcal{Q}$, and since $\deg p = 2$, we therefore have $p \in \mathcal{Q} \cap \mathbb{R}[x]_2$. Evaluating p in $\hat{\mathbf{x}}$ gives $p(\hat{\mathbf{x}}) = -\varepsilon < 0$. This implies $\hat{\mathbf{x}} \notin \bigcap_{r \in \mathbb{N}} \mathbf{K}_{r,2}$. \square

Remark 4.3.1. In the same way, one can show that, in case there exists an $r \in \mathbb{N}$ such that all quadratic polynomials which are positive on \mathbf{K} are contained in \mathcal{Q}_{2r} , then $\mathbf{K} = \mathbf{K}_{r,2}$. This would be the quadratic counterpart to [Las09, Theorem 2].

We conclude that

$$\mathbf{K} = \bigcap_{r \in \mathbb{N}} \mathbf{K}_{r,2} \subseteq \bigcap_{r \in \mathbb{N}} \mathbf{K}_{r,1} = \text{co}(\mathbf{K}) \subseteq \bigcap_{r \in \mathbb{N}} \mathbf{K}_{r,0} = \mathbb{R}^n.$$

4.3.1 Exactness

As in the linear case one may ask in which cases the relaxation is exact. So we define the notion of exactness also for quadratic separators.

Definition 4.7. We say that the quadratic relaxation $\mathbf{K}_{r,2}$ of order $r \in \mathbb{N}$ is *exact*, if $\mathbf{K}_{r,2} = \mathbf{K}$.

As already mentioned in Remark 4.3.1, a sufficient condition for the quadratic relaxation $\mathbf{K}_{r,2}$ to be exact is the set inclusion $\{p \in \mathbb{R}[x]_2 : p(\mathbf{x}) > 0 \forall \mathbf{x} \in \mathbf{K}\} \subseteq \mathcal{Q}_{2r}$. With regard to the proof of Theorem 4.6, we see that already the weaker condition

$$\{\|x - \hat{\mathbf{x}}\|_2^2 - \varepsilon_{\hat{\mathbf{x}}} : \hat{\mathbf{x}} \in \mathbb{R}^n, \varepsilon_{\hat{\mathbf{x}}} > 0 \text{ and } \|x - \hat{\mathbf{x}}\|_2^2 - \varepsilon_{\hat{\mathbf{x}}} > 0 \forall \mathbf{x} \in \mathbf{K}\} \subseteq \mathcal{Q}_{2r}$$

is sufficient. However, this condition is not necessary for exactness, since we do not need all the polynomials $\|x - \hat{\mathbf{x}}\|_2^2 - \varepsilon_{\hat{\mathbf{x}}}$ to describe \mathbf{K} . Instead, there might be different polynomials to replace them.

In the quadratic case it is much more difficult to find examples for $\mathbf{K} \subsetneq \mathbf{K}_{r,2}$ than in the linear case, because the quadratic module contains sums of squares, meaning it will always contain quadratic polynomials. In fact, we did not succeed in finding any.

4.3.2 Examples

To represent graphically the set $\mathbf{K}_{r,2}$ we use a membership oracle, meaning we randomly choose points \mathbf{x} in an appropriate box and solve the problem

$$\inf_{\mathbf{y} \in \mathcal{M}_{2r}} 0 \quad \text{s.t. } y_\alpha = \mathbf{x}^\alpha \text{ for all } |\alpha| \leq 2$$

of deciding whether \mathbf{x} is an element of $\mathbf{K}_{r,2}$. Since the constraint $\mathbf{y} \in \mathcal{M}_{2r}$ is equivalent to $\mathbf{M}_d(\mathbf{y}) \succeq 0$, $\mathbf{M}_{d-d_j}(g_j \mathbf{y}) \succeq 0$, $j = 1, \dots, m$, this is a semidefinite program, and therefore it can be solved numerically. If the semidefinite program has a solution, we plot a dot at location \mathbf{x} .

We consider the same examples as in the linear case:

Example 4.3.1 (Water drop). Once again, let us start with the water drop quartic given by $g(x) = -x_1^2 - x_2^3 - (x_1^2 + x_2^2)^2$. Since the degree of g is 4, the smallest possible relaxation order is $r = 2$. On the left hand side of Figure 4.3, we see the result of the membership oracle for 10 000 random points in the box $[-0.25, 0.25] \times [-1.2, 0.2]$, meaning an illustration for $\mathbf{K}_{2,2}$. Numerically, the approximation seems to be exact. This is due to the quadratic information that we take into account when choosing $t = 2$.

Example 4.3.2 (Folium). Let again $g(x) = -x_1(x_1^2 - 2x_2^2) - (x_1^2 + x_2^2)^2$ be a polynomial defining a folium. Numerical investigations allow the assumption that, again, the largest approximation $\mathbf{K}_{2,2}$ is already exact. We found this surprising, since one may expect difficulties near the singularity at the origin. In Figure 4.3 on the right, we see the result of the membership oracle for 10 000 points in the box $[-1.1, 0.5] \times [-0.6, 0.6]$.

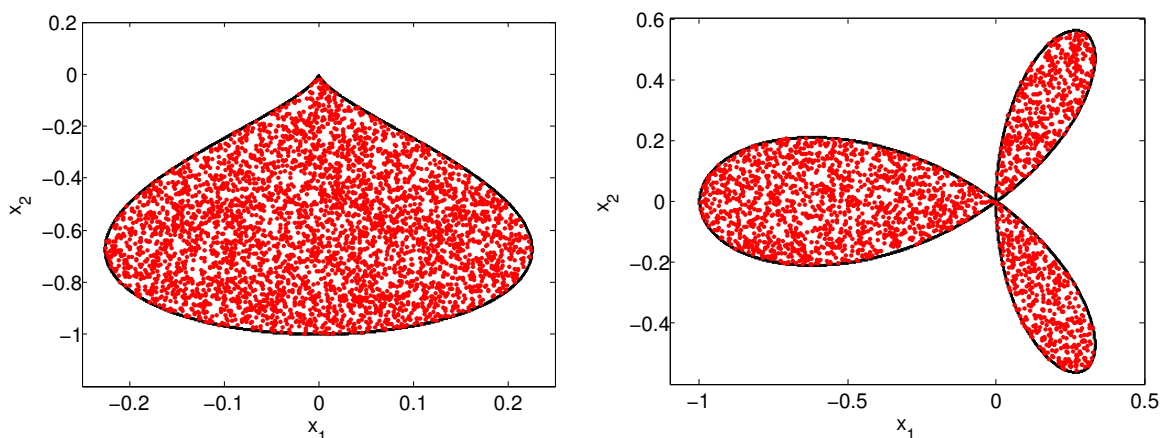
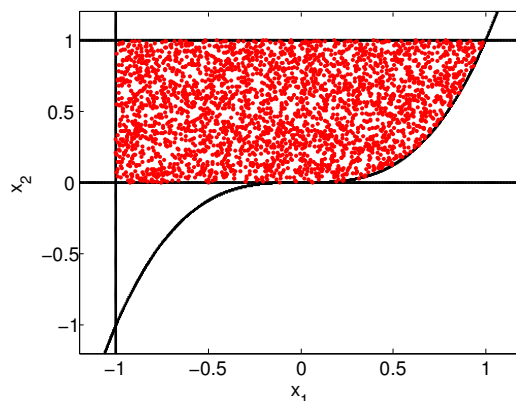


Figure 4.3 – The result of the membership oracle of $\mathbf{K}_{2,2}$ for 10 000 random points in a suitable box for Example 4.3.1 on the left and for Example 4.3.2 on the right.

Example 4.3.3 (Non-exposed face). Recall the convex set $\mathbf{K} \subseteq \mathbb{R}^2$ defined by $g_1(x) = x_2 - x_1^3$, $g_2(x) = x_1 + 1$, $g_3(x) = x_2$ and $g_4(x) = 1 - x_2$ which has a non-exposed face $(0, 0)$. See Figure 4.4.

Figure 4.4 – The black curves represent the zero level sets of the polynomials g_1, g_2, g_3 and g_4 for Example 4.3.3. The red points illustrate the set $\mathbf{K}_{2,2}$, meaning the approximation via quadratic separators of the smallest possible order. Further numerical examination of the region around the non-exposed face $(0, 0)$ suggests that the approximation is exact.



4.4 Concluding remarks

In this chapter the complicated object to approximate was a semialgebraic set, and we wanted to describe it by polynomials of degree 2. For this, we extended the approach of approximating

the convex hull of semialgebraic sets by linear separators proposed by Lasserre in [Las09] to quadratic separators, which indeed gave an outer approximation of the semialgebraic set itself.

In the case of linear separators, the question of exactness has already been treated extensively in the literature, see for example [NPS10] and [KS17]. Finding sufficient conditions for exactness in the case of quadratic separators was a straightforward task. On the contrary, searching necessary conditions for exactness proved to be much more difficult.

We did not succeed in finding any examples for which the approximation $\mathbf{K}_{r,2}$ for the smallest possible value of r does not already seem to be exact. However, we only searched in dimension 2, and we only have numerical certificates for the exactness. Nonetheless, one may wonder, if this is generic and the approximation is always exact. However, this seems very unlikely.

Another problem that we encountered is how to graphically represent the set $\mathbf{K}_{r,2}$. Our method via the membership oracle has obvious drawbacks, as one needs to solve thousands of semidefinite programming problems to get a good idea of the geometry of the set. So, an interesting question would be, if there are more efficient or more illustrative ways to represent $\mathbf{K}_{r,2}$.

Improved convergence rates for Lasserre-type hierarchies

Contents

5.1	A Lasserre-type hierarchy with error bound $O(\frac{1}{d^2})$	72
5.1.1	Background results	72
5.1.2	The hierarchy and main result	73
5.2	Background on the polynomial kernel method	75
5.2.1	Chebyshev polynomials	75
5.2.2	Jackson kernels	77
5.2.3	Jackson kernel approximation of the Dirac delta function	79
5.3	Convergence analysis	80
5.3.1	The univariate case	80
5.3.2	The multivariate case	81
5.4	Computing the parameter ϑ_d^h as a generalized eigenvalue problem	83
5.5	Numerical examples	84
Appendix		86
5.A	Proof of Lemma 5.4	86
5.B	Useful identities for the Chebyshev polynomials	88
5.C	Test functions	89

This final chapter is dedicated to the study of convergence rates, an important notion when investigating hierarchies. The objective is to construct a modification of the hierarchy of upper bounds for which we can show an improved rate of convergence. For the construction we use an approximation of the Dirac delta function. In this sense, the chapter is of slightly different flavor than the preceding chapters, as we do not approximate an object *using* the Lasserre hierarchy (or ideas of it) but we approximate an object - the Dirac delta function - *in order to* develop a hierarchy of Lasserre-type.

Once more we consider problem (1.2) with $\mathbf{K} = [-1, 1]^n$, meaning

$$\varrho^{\min} = \inf_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}) \quad \text{s.t.} \quad \mathbf{x} \in [-1, 1]^n \quad (5.1)$$

for an n -variate polynomial $f \in \mathbb{R}[x]$. We are interested in the convergence rate for hierarchies of upper bounds which solve this problem. For the Lasserre hierarchy of upper bounds introduced in Section 1.4.2, i.e.,

$$\vartheta_d^{\text{mom}} = \inf_{\sigma \in \Sigma[x]_{2d}} \int_{\mathbf{K}} f \sigma \, d\nu \quad \text{s.t.} \quad \int_{\mathbf{K}} \sigma \, d\nu = 1,$$

an error bound in $O(\frac{1}{\sqrt{d}})$ can be shown when choosing $\nu(dx) = dx$, the Lebesgue measure. For this, sum-of-squares polynomials are used as density functions. As densities are positive on the

support of the corresponding measure, this is the most obvious choice. In order to increase the convergence rate, we consider more general densities than sums of squares, more precisely we take polynomials for which the non-negativity is only enforced on the support of the measure.

When considering densities of a beta distribution, an $O(\frac{1}{d})$ error bound can be proved. In this chapter, we show an error bound in $O(\frac{1}{d^2})$ for

$$\nu(dx) = \left(\prod_{i=1}^n \sqrt{1-x_i^2} \right)^{-1} dx$$

(the well-known measure in the study of orthogonal polynomials) and densities admitting a Schmüdgen-type representation with respect to $[-1, 1]^n$.

The approach can be interpreted as approximating the Dirac delta function at the global minimizer not by sums of squares as in the Lasserre hierarchy, but by the above-mentioned measures with densities admitting a Schmüdgen-type representation.

The convergence rate analysis relies on the theory of polynomial kernels and, in particular, on Jackson kernels. We also show that the resulting upper bounds may be computed as generalized eigenvalue problems, as is also the case for sum-of-squares densities.

The chapter reports results from [dKHL17].

Outline of the chapter

The chapter is organized as follows. In the first section, after recalling Lasserre-type hierarchies of upper bounds whose convergence rates have already been examined, we introduce the new hierarchy of upper bounds for which we want to show an error bound in $O(\frac{1}{d^2})$.

As a preparation for the convergence analysis, Section 5.2 contains some background information about the polynomial kernel method. Specifically, we introduce Chebyshev polynomials and Jackson kernels which we then use to construct suitable polynomial densities giving good approximations of the Dirac delta function at a global minimizer of f in the box.

The analysis of the upper bounds is finally carried out in Section 5.3, first for the univariate case, and then for the general multivariate case. In Section 5.4 we show how the new bounds can be computed as generalized eigenvalue problems and in Section 5.5 we conclude with some numerical examples illustrating the behavior of the bounds.

5.1 A Lasserre-type hierarchy with error bound $O(\frac{1}{d^2})$

As aforementioned, we focus on the question of finding a sequence of upper bounds converging to the global minimum and allowing a known estimate on the rate of convergence. The starting point is the same as for the construction of the standard Lasserre hierarchy, namely, the formulation of problem (5.1) as an optimization problem over measures (problem (1.10), recalled below). Then we restrict it to subclasses of measures that we are able to analyze. Sequences of upper bounds have been recently proposed and analyzed in [dKLS16, dKLLS15]. Here, we propose new bounds for which we can prove a sharper rate of convergence.

5.1.1 Background results

For convenience, we recall the formulation of problem (5.1) in terms of measures, i.e., problem (1.10):

$$q^{\min} = \inf_{\mu \in \mathcal{M}_+(\mathbf{K})} \langle f, \mu \rangle \text{ s.t. } \langle 1, \mu \rangle = 1.$$

In Section 1.4.2 we strengthened this problem to the hierarchy (1.18) by optimizing only over probability measures which have a sum-of-squares density $\sigma \in \Sigma[x]_{2d}$ with respect to a fixed measure $\nu \in \mathcal{M}_+(\mathbf{K})$. The infinite dimensional problem associated to the truncated problem (1.18) reads

$$\varrho^{\text{mom}} = \inf_{\sigma \in \Sigma[x]} \int_{\mathbf{K}} f \sigma d\nu \quad \text{s.t.} \quad \int_{\mathbf{K}} \sigma d\nu = 1 \quad (5.2)$$

for a fixed $\nu \in \mathcal{M}_+(\mathbf{K})$. By Theorem 1.4, we have $\varrho^{\min} = \varrho^{\text{mom}}$.

In the recent work [dKLLS15], it is shown that for a compact set $\mathbf{K} \subseteq [0, 1]^n$ one may obtain a similar result using density functions arising from (products of univariate) beta distributions. In particular, the following theorem is implicit in [dKLLS15].

Theorem 5.1. [dKLLS15] *Let $\mathbf{K} \subseteq [0, 1]^n$ be a compact set, let ν be an arbitrary finite Borel measure supported on \mathbf{K} , and let p be a continuous function on \mathbb{R}^n . Then, p is non-negative on \mathbf{K} if and only if*

$$\int_{\mathbf{K}} p h d\nu \geq 0$$

for all h of the form

$$h(x) = \frac{\prod_{i=1}^n x_i^{\beta_i} (1 - x_i)^{\eta_i}}{\int_{\mathbf{K}} \prod_{i=1}^n x_i^{\beta_i} (1 - x_i)^{\eta_i}}, \quad (5.3)$$

where the β_i and η_i are non-negative integers. Therefore, the minimum of p over \mathbf{K} can be expressed as

$$\varrho^{\min} = \inf_h \int_{\mathbf{K}} p h d\nu \quad \text{s.t.} \quad \int_{\mathbf{K}} h d\nu = 1, \quad (5.4)$$

where the infimum is taken over all beta-densities h of the form (5.3).

For the box $\mathbf{K} = [0, 1]^n$ and selecting for ν the Lebesgue measure, we obtain a hierarchy of upper bounds ϑ_d^H converging to ϱ^{\min} , where ϑ_d^H is the optimal value of the program (5.4) when the infimum is taken over all beta-densities h of the form (5.3) with degree d .

The rate of convergence of the upper bounds ϑ_d^{mom} (with $\nu(dx) = dx$) and ϑ_d^H have been investigated recently in [dKLS16] and [dKLLS15], respectively. It is shown in [dKLS16] that $\vartheta_d^{\text{mom}} - \varrho^{\min} = O(\frac{1}{\sqrt{d}})$ for a large class of compact sets \mathbf{K} (including all convex bodies and thus the box $[0, 1]^n$ or $[-1, 1]^n$) and the stronger rate $\vartheta_d^H - \varrho^{\min} = O(\frac{1}{d})$ is shown in [dKLLS15] for the box $\mathbf{K} = [0, 1]^n$. While the parameters ϑ_d^{mom} can be computed using semidefinite optimization (in fact, a generalized eigenvalue computation problem; see [Las11]), an advantage of the parameters ϑ_d^H is that their computation involves only elementary operations (see [dKLLS15]).

Another possibility for getting a hierarchy of upper bounds is grid search, where one takes the best function evaluation at all rational points in $\mathbf{K} = [0, 1]^n$ with given denominator d . It has been shown in [dKLLS15] that these bounds have a rate of convergence in $O(\frac{1}{d^2})$. However, the computation of the order d bound needs an exponential number d^n of function evaluations.

5.1.2 The hierarchy and main result

In [dKHL17], on which this chapter is based, we continued this line of research. For the box $\mathbf{K} = [-1, 1]^n$, our objective is to build a new hierarchy of measure-based upper bounds, for which we are able to show a sharper rate of convergence in $O(\frac{1}{d^2})$. We obtain these upper bounds by considering a specific Borel measure ν (specified below in (5.6)) and polynomial density functions with a so-called Schmüdgen-type SOS representation (as in (5.5) below).

We first recall the relevant result of Schmüdgen [Sch91], which gives SOS representations for positive polynomials on a basic closed semialgebraic set (see also, e.g., [PD01],[Lau09, Theorem 3.16], [Mar08]). Write $[m] := \{1, \dots, m\}$.

Theorem 5.2 (Schmüdgen [Sch91]). *Let $\mathbf{K} = \{x \in \mathbb{R}^n \mid g_1(x) \geq 0, \dots, g_m(x) \geq 0\}$ with $g_1, \dots, g_m \in \mathbb{R}[x]$ be compact. If $p \in \mathbb{R}[x]$ is positive on \mathbf{K} , then p can be written as $p = \sum_{I \subseteq [m]} \sigma_I \prod_{i \in I} g_i$, with $\sigma_I \in \Sigma[x]$ for all $I \subseteq [m]$.*

For the box $\mathbf{K} = [-1, 1]^n$, described by the polynomial inequalities $1 - x_1^2 \geq 0, \dots, 1 - x_n^2 \geq 0$, we consider polynomial densities that allow a Schmüdgen-type representation of bounded degree $2d$,

$$h(x) = \sum_{I \subseteq [n]} \sigma_I(x) \prod_{i \in I} (1 - x_i^2), \quad (5.5)$$

where the polynomials $\sigma_I \in \Sigma[x]_{2(d-|I|)}$ are sum-of-squares polynomials with degree at most $2(d - |I|)$ for $|I|$ denoting the cardinality of I (to ensure that the degree of h is at most $2d$). Further, we fix the following Borel measure ν on $[-1, 1]^n$ (which, as will be recalled below, is associated with some orthogonal polynomials),

$$\nu(dx) = \left(\prod_{i=1}^n \pi \sqrt{1 - x_i^2} \right)^{-1} dx. \quad (5.6)$$

For $d \in \mathbb{N}$, this leads to the following new hierarchy of upper bounds ϑ_d^h for ϱ^{\min} :

$$\vartheta_d^h := \inf_h \int_{[-1, 1]^n} f h d\nu \quad \text{s.t.} \quad \int_{[-1, 1]^n} h d\nu = 1, \quad (5.7)$$

where the infimum is taken over the polynomial densities h that allow a Schmüdgen-type representation (5.5) with $\sigma_I \in \Sigma[x]_{2(d-|I|)}$.

The convergence of the decreasing sequence of parameters $\vartheta_d^h \geq \vartheta_{d+1}^h \geq \varrho^{\min}$ to ϱ^{\min} follows as a direct consequence of the convergence of ϑ_d^{mom} (with ν as in (5.6)), since sums of squares allow a Schmüdgen-type representation, so $\vartheta_d^h \leq \vartheta_d^{\text{mom}}$ (with ν as in (5.6)). Note that we used the fact that $\mathbf{K} = [-1, 1]^n$ has a nonempty interior, since otherwise ϑ_d^{mom} might be unbounded. As a small remark, note also that, again due to the non-empty interior of \mathbf{K} , the program (5.7) has an optimal solution h^* for all d by [Las11, Theorem 4.2].

A main result in this chapter is to show that the bounds ϑ_d^h have a rate of convergence in $O(\frac{1}{d^2})$. Moreover, we will show that the parameter ϑ_d^h can be computed through generalized eigenvalue computations.

Theorem 5.3. *Let $f \in \mathbb{R}[x]$ be a polynomial and ϱ^{\min} be its minimum value over the box $[-1, 1]^n$. For any d large enough, the parameters ϑ_d^h defined in (5.7) satisfy*

$$\vartheta_d^h - \varrho^{\min} = O\left(\frac{1}{d^2}\right).$$

As already observed above this result compares favorably with the estimate $\vartheta_d^{\text{mom}} - \varrho^{\min} = O(\frac{1}{\sqrt{d}})$ shown in [dKLS16] for the bounds ϑ_d^{mom} based on using SOS densities and the Lebesgue measure. (Note however that the latter convergence rate holds for a larger class of sets \mathbf{K} that includes all convex bodies; see [dKLS16] for details.) The new result also improves the estimate $\vartheta_d^H - \varrho^{\min} = O(\frac{1}{d})$ shown in [dKLLS15] for the bounds ϑ_d^H obtained by using densities arising from beta distributions. The theorem follows directly from Theorem 5.12 which will be proven in Section 5.3.

We now illustrate the optimal densities appearing in the new bounds ϑ_d^h on an example.

Example 5.1.1. Consider the minimization of the Motzkin polynomial

$$f(x_1, x_2) = 64(x_1^4 x_2^2 + x_1^2 x_2^4) - 48x_1^2 x_2^2 + 1$$

over the hypercube $[-1, 1]^2$, which has four global minimizers at the points $(\pm\frac{1}{2}, \pm\frac{1}{2})$, and for which $\varrho^{\min} = 0$. Figure 5.1 shows the optimal density function h^* computed when solving the problem (5.7) for degrees $2d = 12$ and $2d = 16$, respectively. Note that the optimal density h^* shows four peaks at the four global minimizers of f in $[-1, 1]^2$. The corresponding upper bounds from (5.7) are $\vartheta_6^h = 0.8098$ and $\vartheta_8^h = 0.6949$.

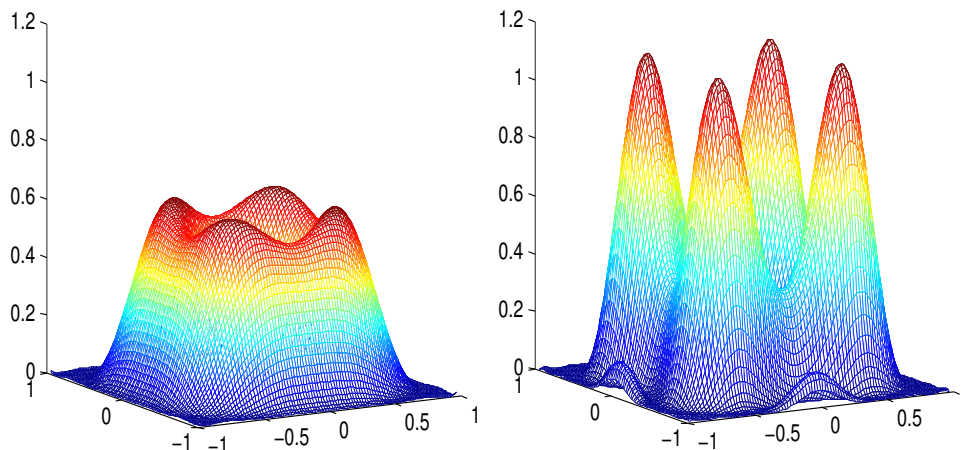


Figure 5.1 – Graphs of h^* with $\deg h^* = 12$ (left) and $\deg h^* = 16$ (right) on $[-1, 1]^2$ for the Motzkin polynomial.

In order to show the convergence rate in $O(\frac{1}{d^2})$ of Theorem 5.3 we need to exhibit a polynomial density function h_{2d} of degree at most $2d$ which admits an SOS representation of Schmüdgen-type and for which we are able to show that $\int_{[-1, 1]^n} f h_{2d} d\nu - \varrho^{\min} = O(\frac{1}{d^2})$.

The idea is to find such a polynomial density which approximates well the Dirac delta function at a global minimizer \mathbf{x}^* of f over $[-1, 1]^n$. For this we use the well-established polynomial kernel method (KPM) and, more specifically, we use the Jackson kernel, a well-known tool in approximation theory to yield best (uniform) polynomial approximations of continuous functions.

5.2 Background on the polynomial kernel method

Our goal is to approximate the Dirac delta function at a given point $\mathbf{x}^* \in \mathbb{R}^n$ as well as possible, using polynomial density functions of bounded degrees. This is a classical question in approximation theory. In this section we review how this may be done using the polynomial kernel method and, in particular, using Jackson kernels. This theory is usually developed using the Chebyshev polynomials, and we start by reviewing their properties.

We follow mainly the work [WWAF06] for our exposition and we refer to the handbook [AS72] for more background information.

5.2.1 Chebyshev polynomials

We will use the univariate polynomials $T_k(x)$ and $U_k(x)$, respectively, known as the Chebyshev polynomials of the first and second kind. They are defined as follows:

$$T_k(x) = \cos(k \arccos(x)), \quad U_k(x) = \frac{\sin((k+1) \arccos(x))}{\sin(\arccos(x))} \quad (5.8)$$

for $x \in [-1, 1]$, $k \in \mathbb{N}$. They satisfy the following recurrence relationships:

$$T_0(x) = 1, \quad T_{-1}(x) = T_1(x) = x, \quad T_{k+1}(x) = 2xT_k(x) - T_{k-1}(x), \quad (5.9)$$

$$U_0(x) = 1, \quad U_{-1}(x) = 0, \quad U_{k+1}(x) = 2xU_k(x) - U_{k-1}(x). \quad (5.10)$$

As a direct application one can verify that for $k \in \mathbb{N}$,

$$\begin{aligned} T_k(0) &= 0 \text{ for } k \text{ odd, } T_k(0) = (-1)^{k/2} \text{ for } k \text{ even, } T_k(1) = 1, \\ U_k(1) &= k + 1, \quad U_k(-1) = (-1)^k(k + 1). \end{aligned} \quad (5.11)$$

The Chebyshev polynomials have the extrema

$$\max_{\mathbf{x} \in [-1, 1]} |T_k(\mathbf{x})| = 1 \text{ and } \max_{\mathbf{x} \in [-1, 1]} |U_k(\mathbf{x})| = k + 1,$$

attained at $\mathbf{x} = \pm 1$ (see, e.g., [AS72, §22.14.4, 22.14.6]).

The Chebyshev polynomials are orthogonal for the following inner product on the space of integrable functions over $[-1, 1]$:

$$\langle f, g \rangle = \int_{-1}^1 \frac{f(x)g(x)}{\pi\sqrt{1-x^2}} dx, \quad (5.12)$$

and their orthogonality relationships read

$$\langle T_k, T_s \rangle = 0 \text{ for } k \neq s, \quad \langle T_0, T_0 \rangle = 1, \quad \langle T_k, T_k \rangle = \frac{1}{2} \text{ for } k \geq 1. \quad (5.13)$$

For any $d \in \mathbb{N}$ the Chebyshev polynomials T_k ($k \leq d$) form a basis of the space of univariate polynomials with degree at most d . One may write the Chebyshev polynomials in the standard monomial basis using the relations

$$\begin{aligned} T_k(x) &= \sum_{i=0}^k t_i^{(k)} x^i = \frac{k}{2} \sum_{s=0}^{\lfloor k/2 \rfloor} (-1)^s \frac{(k-s-1)!}{s!(k-2s)!} (2x)^{k-2s}, & k > 0, \\ U_{k-1}(x) &= \sum_{i=0}^{k-1} u_i^{(k)} x^i = \sum_{s=0}^{\lfloor \frac{k-1}{2} \rfloor} (-1)^s \frac{(k-s-1)!}{s!(k-1-2s)!} (2x)^{k-1-2s}, & k > 1; \end{aligned}$$

see, e.g. [AS72, Chap. 22]. From this, one may derive a bound on the largest coefficient in absolute value appearing in the above expansions of $T_k(x)$ and $U_{k-1}(x)$. A proof for the following result is given in Appendix 5.A.

Lemma 5.4. *For any fixed integer $k > 1$, one has*

$$\max_{0 \leq i \leq k-1} |u_i^{(k)}| \leq \max_{0 \leq i \leq k} |t_i^{(k)}| = 2^{k-1-2\psi(k)} \frac{k(k-\psi(k)-1)!}{\psi(k)!(k-2\psi(k))!}, \quad (5.14)$$

where $\psi(k) = 0$ for $k \leq 4$ and $\psi(k) = \left\lceil \frac{1}{8} (4k - 5 - \sqrt{8k^2 - 7}) \right\rceil$ for $k \geq 4$. Moreover, the right-hand side of (5.14) increases monotonically with increasing k .

In the multivariate case we use the following notation. We let $\nu(dx)$ denote the Lebesgue

measure on $[-1, 1]^n$ with the function $\prod_{i=1}^n \left(\pi \sqrt{1 - x_i^2} \right)^{-1}$ as the density function, i.e.,

$$\nu(dx) = \prod_{i=1}^n \left(\pi \sqrt{1 - x_i^2} \right)^{-1} dx, \quad (5.15)$$

and we consider the following inner product for two integrable functions f, g on the box $[-1, 1]^n$:

$$\langle f, g \rangle = \int_{[-1, 1]^n} f(x)g(x) \nu(dx)$$

(which coincides with (5.12) in the univariate case $n = 1$). For $\alpha \in \mathbb{N}^n$, we define the multivariate Chebyshev polynomial

$$T_\alpha(x) = \prod_{i=1}^n T_{\alpha_i}(x_i) \quad \text{for } x \in \mathbb{R}^n.$$

The multivariate Chebyshev polynomials satisfy the following orthogonality relationship:

$$\langle T_\alpha, T_\beta \rangle = \left(\frac{1}{2} \right)^{|\text{supp}(\alpha)|} \delta_{\alpha, \beta} \quad \text{for } \alpha, \beta \in \mathbb{N}^n, \quad (5.16)$$

where $\delta_{\alpha, \beta} \in \{0, 1\}$ is equal to 1 if and only if $\alpha = \beta$ and $\text{supp}(\alpha) := \{i \in \{1, \dots, n\} : \alpha_i \neq 0\}$. For any $d \in \mathbb{N}$, the set of Chebyshev polynomials $\{T_\alpha(x) : |\alpha| \leq d\}$ is a basis of the space of n -variate polynomials of degree at most d .

5.2.2 Jackson kernels

A classical problem in approximation theory is to find a best (uniform) approximation of a given continuous function $f : [-1, 1] \rightarrow \mathbb{R}$ by a polynomial of given maximum degree d . Following [WWAF06], a possible approach is to take the convolution $f_{\text{KPM}}^{(d)}$ of f with a kernel function of the form

$$K_d(x, y) = \frac{1}{\pi \sqrt{1 - x^2} \pi \sqrt{1 - y^2}} \left(g_0^d T_0(x)T_0(y) + 2 \sum_{k=1}^d g_k^d T_k(x)T_k(y) \right),$$

where $d \in \mathbb{N}$ and the coefficients g_k^d are selected so that the following properties hold:

1. The kernel is positive: $K_d(x, y) > 0$ for all $x, y \in [-1, 1]$.
2. The kernel is normalized: $g_0^d = 1$.
3. The second coefficients g_1^d tend to 1 as $d \rightarrow \infty$.

The function $f_{\text{KPM}}^{(d)}$ is then defined by

$$f_{\text{KPM}}^{(d)}(x) = \int_{-1}^1 \pi \sqrt{1 - y^2} K_d(x, y) f(y) dy. \quad (5.17)$$

As the first coefficient is $g_0^d = 1$, the kernel is normalized, $\int_{-1}^1 K_d(x, y) dy = T_0(x)(\pi \sqrt{1 - x^2})^{-1}$, and we have $\int_{-1}^1 f_{\text{KPM}}^{(d)}(x) dx = \int_{-1}^1 f(x) dx$. The positivity of the kernel K_d implies that the integral operator $f \mapsto f_{\text{KPM}}^{(d)}$ is a positive linear operator, i.e., a linear operator that maps the set of non-negative integrable functions on $[-1, 1]$ onto itself. Thus, the general (Korovkin) convergence theory of positive linear operators applies and one may conclude the

uniform convergence result

$$\lim_{d \rightarrow \infty} \|f - f_{\text{KPM}}^{(d)}\|_{\infty}^{\varepsilon} = 0$$

for any $0 < \varepsilon < 2$, where $\|f - f_{\text{KPM}}^{(d)}\|_{\infty}^{\varepsilon} = \max_{-1+\varepsilon \leq x \leq 1-\varepsilon} |f(x) - f_{\text{KPM}}^{(d)}(x)|$. (One needs to restrict the range to subintervals of $[-1, 1]$ because of the denominator in the kernel K_d .)

In what follows, we select the following parameters g_k^d for $k = 1, \dots, d$ which define the so-called *Jackson kernel*, again denoted by $K_d(x, y)$,

$$\begin{aligned} g_k^d &= \frac{1}{d+2} \left((d+2-k) \cos(k\theta_d) + \frac{\sin(k\theta_d)}{\sin\theta_d} \cos\theta_d \right) \\ &= \frac{1}{d+2} \left((d+2-k) T_k(\cos\theta_d) + U_{k-1}(\cos\theta_d) \cos\theta_d \right) \end{aligned} \quad (5.18)$$

with $\theta_d := \frac{\pi}{d+2}$. This choice of the parameters g_k^d minimizes the quantity $\int_{[-1,1]^2} K_d(x, y) (x-y)^2 dx dy$, which ensures that the corresponding Jackson kernel is maximally peaked at $x = y$ (see [WWAF06, §II.C.3]).

One may show that the Jackson kernel $K_d(x, y)$ is indeed positive on $[-1, 1]^2$; see [WWAF06, §II.C.2]. Moreover, $g_0^d = 1$ and, for $k = 1$, we have $g_1^d = \cos(\theta_d) = \cos(\frac{\pi}{d+2}) \rightarrow 1$ for $d \rightarrow \infty$, as required. This is true for all k , as follows from Lemma 5.5 below. Note that one has $|g_k^d| \leq 1$ for all k , since $|T_k(\cos\theta_d)| \leq 1$ and $|U_{k-1}(\cos\theta_d)| \leq k$.

For later use, we now give an estimate on the Jackson coefficients g_k^d , showing that $1 - g_k^d$ is on the order $O(\frac{1}{d^2})$.

Lemma 5.5. *Let $\ell \geq 1$ and $d \geq \ell$ be given integers, and set $\theta_d = \frac{\pi}{d+2}$. There exists a constant C_{ℓ} (depending only on ℓ) such that the following inequalities hold:*

$$|1 - g_k^d| \leq C_{\ell}(1 - \cos\theta_d) \leq \frac{C_{\ell}\pi^2}{2(d+2)^2} \quad \text{for all } 0 \leq k \leq \ell.$$

For the constant C_{ℓ} we may take $C_{\ell} = \ell^2(1 + 2c_{\ell})$, where

$$\begin{aligned} c_{\ell} &= 2^{\ell-1-2\psi(\ell)} \frac{\ell(\ell - \psi(\ell) - 1)!}{\psi(\ell)!(\ell - 2\psi(\ell))!} \\ \text{and } \psi(\ell) &= \begin{cases} 0 & \text{for } \ell \leq 4, \\ \left\lceil \frac{1}{8} (4\ell - 5 - \sqrt{8\ell^2 - 7}) \right\rceil & \text{for } \ell \geq 4. \end{cases} \end{aligned} \quad (5.19)$$

Proof. Define the polynomial

$$P_k(x) = 1 - \frac{d+2-k}{d+2} T_k(x) - \frac{1}{r+2} x U_{k-1}(x)$$

with degree k . Then, in view of relation (5.18), we have $1 - g_k^d = P_k(\cos\theta_d)$. Recall from relation (5.11) that $T_k(1) = 1$ and $U_{k-1}(1) = k$ for any $k \in \mathbb{N}$. This implies that $P_k(1) = 0$, and thus we can factor $P_k(x)$ as $P_k(x) = (1-x)Q_k(x)$ for some polynomial $Q_k(x)$ with degree $k-1$. If we write $P_k(x) = \sum_{i=0}^k p_i x^i$, then it follows that $Q_k(x) = \sum_{i=0}^{k-1} q_i x^i$ with scalars q_i given by

$$q_i = \sum_{j=0}^i p_j \quad \text{for } i = 0, 1, \dots, k-1. \quad (5.20)$$

It now suffices to observe that for any $0 \leq i \leq k$ and $k \leq \ell$, the p_i 's are bounded by a constant depending only on ℓ , which will imply that the same holds for the scalars q_i . For this, set

$T_k(x) = \sum_{i=0}^k t_i^{(k)} x^i$ and $U_{k-1}(x) = \sum_{i=0}^{k-1} u_i^{(k)} x^i$. Then the coefficients p_i of $P_k(x)$ can be expressed as

$$p_0 = 1 - \frac{d+2-k}{d+2} t_0^{(k)}, \quad p_i = \frac{d+2-k}{d+2} t_i^{(k)} - \frac{u_{i-1}^{(k)}}{d+2} \quad (1 \leq i \leq k).$$

For all $0 \leq k \leq \ell$ the coefficients of the Chebyshev polynomials T_k, U_{k-1} can be bounded by an absolute constant depending only on ℓ . Namely, by Lemma 5.4, $|t_i^{(k)}|, |u_i^{(k)}| \leq c_\ell$ for all $0 \leq i \leq k$ and $k \leq \ell$, where c_ℓ is as defined in (5.19).

As $k \leq \ell \leq d$, we have $d+2-k \leq d+2$, and thus $|p_i| \leq 1 + 2c_\ell$ for all $0 \leq i \leq k \leq \ell$. Moreover, using (5.20), $|q_i| \leq \ell(1 + 2c_\ell)$ for all $0 \leq i \leq k-1$.

Putting things together we can now derive $1 - g_k^d = (1 - \cos \theta_d) Q_k(\cos \theta_d)$, where $Q_k(\cos \theta_d) = \sum_{i=0}^{k-1} q_i(\cos \theta_d)^i$, so that $|Q_k(\cos \theta_d)| \leq \sum_{i=0}^{k-1} |q_i| \leq \ell^2(1 + 2c_\ell)$. This implies $|1 - g_k^d| \leq (1 - \cos \theta_d) C_\ell$, after setting $C_\ell = \ell^2(1 + 2c_\ell)$. Finally, combining this with the fact that $1 - \cos x \leq \frac{x^2}{2}$ for all $x \in [0, \pi]$, we obtain the desired inequality from the lemma statement. \square

5.2.3 Jackson kernel approximation of the Dirac delta function

If one approximates the Dirac delta function δ_{x^*} at a given point $x^* \in [-1, 1]$ by taking its convolution with the Jackson kernel $K_d(x, y)$, then the result is the function

$$\delta_{\text{KPM}}^{(d)}(x - x^*) = \frac{1}{\pi\sqrt{1-x^2}} \left(1 + 2 \sum_{k=1}^d g_k^d T_k(x) T_k(x^*) \right);$$

see [WWAF06, eq. (72)]. As mentioned in [WWAF06, eq. (75)–(76)], the function $\delta_{\text{KPM}}^{(d)}$ is a good approximation to the Gaussian density,

$$\delta_{\text{KPM}}^{(d)}(x - x^*) \approx \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - x^*)^2}{2\sigma^2}\right) \quad \text{with } \sigma^2 \simeq \left(\frac{\pi}{d+1}\right)^2 \left[1 - x^{*2} + \frac{3x^{*2} - 2}{d+1}\right]. \quad (5.21)$$

(Recall that the Dirac delta measure may be defined as a limit of the Gaussian measure when $\sigma \searrow 0$.) This approximation is illustrated in Figure 5.2 for several values of d .

By construction, the function $\delta_{\text{KPM}}^{(d)}(x - x^*)$ is non-negative over $[-1, 1]$ and we have the normalization $\int_{-1}^1 \delta_{\text{KPM}}^{(d)}(x - x^*) dx = \int_{-1}^1 \delta_{x^*}(x) dx = 1$ (cf. Section 5.2.2). Hence, it is a probability density function on $[-1, 1]$ for the Lebesgue measure. It is convenient to consider the univariate polynomial

$$h_d(x) = 1 + 2 \sum_{k=1}^d g_k^d T_k(x) T_k(x^*), \quad (5.22)$$

so that $\delta_{\text{KPM}}^{(d)}(x - x^*) = \frac{1}{\pi\sqrt{1-x^2}} h_d(x)$. The following facts follow directly, and we use them below for the convergence analysis of the new bounds ϑ_d^h .

Lemma 5.6. *For any $d \in \mathbb{N}$ the polynomial h_d from (5.22) is non-negative over $[-1, 1]$ and $\int_{-1}^1 h_d(x) \frac{dx}{\pi\sqrt{1-x^2}} = 1$. In other words, h_d is a probability density function for the measure $(\pi\sqrt{1-x^2})^{-1} dx$ on $[-1, 1]$.*

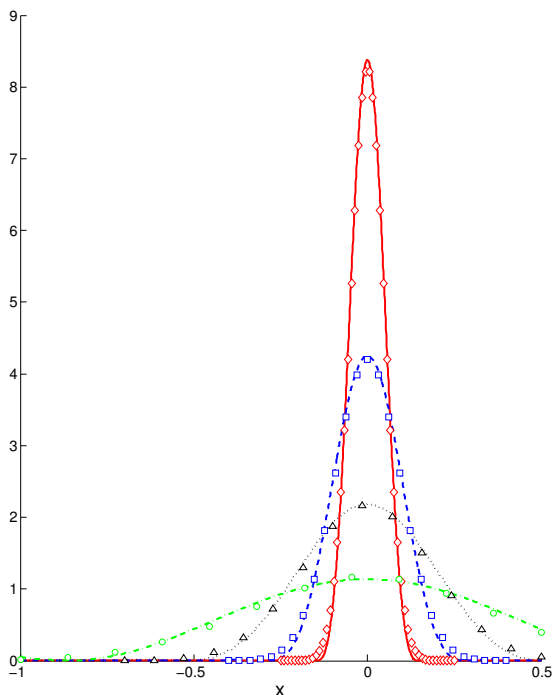


Figure 5.2 – The Jackson kernel approximation $\delta_{\text{KPM}}^{(d)}$ to the Dirac delta function at $\mathbf{x}^* = 0$ for $d = 8$ (green), 16 (black), 32 (blue), 64 (red). The corresponding scatterplots show the values of the Gaussian density function in (5.21) with $\mathbf{x}^* = 0$.

5.3 Convergence analysis

In this section we analyze the convergence rate of the new bounds ϑ_d^h and we show the result from Theorem 5.3. We first consider the univariate case in Section 5.3.1 (see Theorem 5.9), and then the general multivariate case in Section 5.3.2 (see Theorem 5.12). As we will see, the polynomial h_d arising from the Jackson kernel approximation of the Dirac delta function introduced above in relation (5.22), plays a key role in the convergence analysis.

5.3.1 The univariate case

We consider a univariate polynomial f and let \mathbf{x}^* be a global minimizer of f in $[-1, 1]$. As observed in Lemma 5.6 the polynomial h_d from (5.22) is a density function for the measure $\frac{dx}{\pi\sqrt{1-x^2}}$. The key observation now is that the polynomial h_d admits a Schmüdgen-type representation, of the form $\sigma(x) + \sigma_1(x)(1-x^2)$ with σ_0, σ_1 sums of squares, since it is non-negative over $[-1, 1]$. This fact allows us to use the polynomial h_{2d} to get feasible solutions for the program defining the bound ϑ_d^h . It follows from the following classical result (see, e.g., [PR00]) which characterizes univariate polynomials that are non-negative on $[-1, 1]$. (Note that this is a strengthening of Schmüdgen's theorem (Theorem 5.2) for the univariate case.)

Theorem 5.7 (Fekete, Markov-Lukàcz). *Let $p(x)$ be a univariate polynomial of degree ℓ . Then $p(x)$ is non-negative on the interval $[-1, 1]$ if and only if it can be written as*

$$p(x) = \sigma_0(x) + (1-x^2)\sigma_1(x)$$

for some sum-of-squares polynomials σ_0 of degree $2\lceil\ell/2\rceil$ and σ_1 of degree $2\lceil\ell/2\rceil - 2$.

We start the convergence analysis with the following technical lemma.

Lemma 5.8. *Let f be a polynomial of degree ℓ written in the Chebyshev basis as $f = \sum_{k=0}^{\ell} f_k T_k$, let \mathbf{x}^* be a global minimizer of f in $[-1, 1]$, and let h_d be the polynomial from (5.22). For any*

integer $d \geq \ell$ we have

$$\int_{-1}^1 f(x) h_d(x) \frac{dx}{\pi \sqrt{1-x^2}} - f(\mathbf{x}^*) \leq \frac{D_f}{(d+2)^2},$$

where $D_f = \frac{C_\ell \pi^2}{2} \sum_{k=1}^{\ell} |f_k|$ and C_ℓ is the constant from Lemma 5.5.

Proof. As $f = \sum_{k=0}^{\ell} f_k T_k$ and $h_d = 1 + 2 \sum_{k=1}^d g_k^d T_k(\mathbf{x}^*) T_k$, we use the orthogonality relationships (5.13) to obtain

$$\int_{-1}^1 f(x) h_d(x) \frac{dx}{\pi \sqrt{1-x^2}} = \sum_{k=0}^{\ell} f_k T_k(\mathbf{x}^*) g_k^d. \quad (5.23)$$

Combined with $f(\mathbf{x}^*) = \sum_{k=0}^{\ell} f_k T_k(\mathbf{x}^*)$, this gives

$$\int_{-1}^1 f(x) h_d(x) \frac{dx}{\pi \sqrt{1-x^2}} - f(\mathbf{x}^*) = \sum_{k=1}^d f_k T_k(\mathbf{x}^*) (g_k^d - 1). \quad (5.24)$$

Now we use the upper bound on $g_k^d - 1$ from Lemma 5.5 and the bound $|T_k(\mathbf{x}^*)| \leq 1$ to conclude the proof. \square

We can now conclude the convergence analysis of the bounds ϑ_d^h in the univariate case.

Theorem 5.9. *Let $f = \sum_{k=0}^{\ell} f_k T_k$ be a polynomial of degree ℓ . For any integer $d \geq \lceil \ell/2 \rceil$ we have*

$$\vartheta_d^h - \varrho^{\min} \leq \frac{D_f}{(2d+2)^2},$$

where $D_f = \frac{C_\ell \pi^2}{2} \sum_{k=1}^{\ell} |f_k|$ and C_ℓ is the constant from Lemma 5.5.

Proof. Using the degree bounds in Theorem 5.7 for the sum-of-squares polynomials entering the decomposition of the polynomial h_{2d} , we can conclude that h_{2d} is feasible for the program defining the parameter ϑ_d^h . Setting $D_f = (\sum_{k=1}^{\ell} |f_k|) \frac{C_\ell \pi^2}{2}$ and using Lemma 5.8, this implies $\vartheta_d^h - \varrho^{\min} \leq \frac{D_f}{(2d+2)^2}$. The result of the theorem now follows. \square

5.3.2 The multivariate case

We consider now a multivariate polynomial f and we let $\mathbf{x}^* = (x_1^*, \dots, x_n^*) \in [-1, 1]^n$ denote a global minimizer of f on $[-1, 1]^n$, i.e., $f(\mathbf{x}^*) = \varrho^{\min}$.

In order to obtain a feasible solution to the program defining the parameter ϑ_d^h we will consider suitable products of the univariate polynomials h_d from (5.22). Namely, for given integers $d_1, \dots, d_n \in \mathbb{N}$ we define the n -tuple $\mathbf{d} = (d_1, \dots, d_n)$ and the n -variate polynomial

$$H_{\mathbf{d}}(x_1, \dots, x_n) = \prod_{i=1}^n h_{d_i}(x_i). \quad (5.25)$$

We group in the next lemma some properties of the polynomial $H_{\mathbf{d}}$.

Lemma 5.10. *The polynomial $H_{\mathbf{d}}$ satisfies the following properties:*

- (a) $H_{\mathbf{d}}$ is non-negative on $[-1, 1]^n$.
- (b) $\int_{[-1, 1]^n} H_{\mathbf{d}}(x) \nu(dx) = 1$, where $d\nu$ is the measure from (5.6).
- (c) $H_{\mathbf{d}}$ has a Schmüdgen-type representation of the form $H_{\mathbf{d}}(x) = \sum_{I \subseteq [n]} \sigma_I(x) \prod_{i \in I} (1 - x_i^2)$, where each σ_I is a sum-of-squares polynomial of degree at most $2 \sum_{i=1}^n \lceil d_i/2 \rceil - 2|I|$.

Proof. (a) and (b) follow directly from the corresponding properties of the univariate polynomials h_{d_i} , and (c) follows using Theorem 5.7 applied to the polynomials h_{d_i} . \square

The next lemma is the analog of Lemma 5.8 for the multivariate case.

Lemma 5.11. *Let f be a multivariate polynomial of degree ℓ , written in the basis of multivariate Chebyshev polynomials as $f = \sum_{|\alpha| \leq \ell} f_\alpha T_\alpha$, and let \mathbf{x}^* be a global minimizer of f in $[-1, 1]^n$. Consider $\mathbf{d} = (d_1, \dots, d_n)$, where each d_i is an integer satisfying $d_i \geq \ell$, and the polynomial $H_{\mathbf{d}}$ from (5.25). We have*

$$\int_{[-1, 1]^n} f(x) H_{\mathbf{d}}(x) \nu(dx) - f(\mathbf{x}^*) \leq D_f \sum_{i=1}^n \frac{1}{(d_i + 2)^2},$$

where $D_f = \frac{C_\ell \pi^2}{2} \sum_{|\alpha| \leq \ell} |f_\alpha|$ and C_ℓ is the constant from Lemma 5.5.

Proof. As $H_{\mathbf{d}}(x) = \prod_{i=1}^n h_{d_i}(x_i) = \prod_{i=1}^n (1 + 2 \sum_{k_i=1}^{d_i} g_{k_i}^{d_i} T_{k_i}(x_i) T_{k_i}(\mathbf{x}_i^*))$ and $f = \sum_{|\alpha| \leq \ell} f_\alpha T_\alpha$, we can use the orthogonality relationships (5.16) among the multivariate Chebyshev polynomials to derive

$$\int_{[-1, 1]^n} f(x) H_{\mathbf{d}}(x) \nu(dx) = \sum_{|\alpha| \leq \ell} f_\alpha T_\alpha(\mathbf{x}^*) \prod_{i=1}^n g_{\alpha_i}^{d_i}.$$

Combined with $f(\mathbf{x}^*) = \sum_{|\alpha| \leq \ell} f_\alpha T_\alpha(\mathbf{x}^*)$, this gives

$$\int_{[-1, 1]^n} f(x) H_{\mathbf{d}}(x) \nu(dx) - f(\mathbf{x}^*) = \sum_{|\alpha| \leq \ell} f_\alpha T_\alpha(\mathbf{x}^*) \left(\prod_{i=1}^n g_{\alpha_i}^{d_i} - 1 \right).$$

Using the identity $\prod_{i=1}^n g_{\alpha_i}^{d_i} - 1 = \sum_{j=1}^n (g_{\alpha_j}^{d_j} - 1) \prod_{k=j+1}^n g_{\alpha_k}^{d_k}$ and the fact that $|g_{\alpha_k}^{d_k}| \leq 1$, we get $|\prod_{i=1}^n g_{\alpha_i}^{d_i} - 1| \leq \sum_{j=1}^n |g_{\alpha_j}^{d_j} - 1|$. Now use $|T_\alpha(\mathbf{x}^*)| \leq 1$ and the bound from Lemma 5.5 for each $|1 - g_{\alpha_j}^{d_j}|$ to conclude the proof. \square

We can now show our main result, which implies Theorem 5.3.

Theorem 5.12. *Let $f = \sum_{|\alpha| \leq \ell} f_\alpha T_\alpha$ be an n -variate polynomial of degree ℓ . For any integer $d \geq n \left\lceil \frac{\ell+2}{2} \right\rceil$, we have*

$$\vartheta_d^h - \varrho^{\min} \leq \frac{D_f n^3}{(2d+1)^2},$$

where $D_f = \frac{C_\ell \pi^2}{2} \sum_{|\alpha| \leq \ell} |f_\alpha|$ and C_ℓ is the constant from Lemma 5.5.

Proof. Write $2d - n = sn + n_0$, where $s, n_0 \in \mathbb{N}$ and $0 \leq n_0 < n$, and define the n -tuple $\mathbf{d}_2 = (d_1, \dots, d_n)$ setting $d_i = s + 1$ for $1 \leq i \leq n_0$ and $d_i = s$ for $n_0 + 1 \leq i \leq n$, so that $2d - n = d_1 + \dots + d_n$. Note that the condition $d \geq n \left\lceil \frac{\ell+2}{2} \right\rceil$ implies $s \geq \ell$, and thus $d_i \geq \ell$ for all i . Moreover, we have $2 \sum_{i=1}^n \lceil d_i/2 \rceil = 2n_0 \lceil (s+1)/2 \rceil + 2(n - n_0) \lceil s/2 \rceil$, which is equal to $2d - n + n_0$ for even s and to $2d - n_0$ for odd s , and thus always at most $2d$. Hence, the polynomial $H_{\mathbf{d}_2}$ from (5.25) has degree at most $2d$. By Lemma 5.10(b),(c), it follows that the polynomial $H_{\mathbf{d}_2}$ is feasible for the program defining the parameter ϑ_d^h . By Lemma 5.11 this implies that

$$\vartheta_d^h - \varrho^{\min} \leq \int_{[-1, 1]^n} f(x) H_{\mathbf{d}_2}(x) \nu(dx) - f(\mathbf{x}^*) \leq D_f \sum_{i=1}^n \frac{1}{(d_i + 2)^2}.$$

Finally, $\sum_{i=1}^n \frac{1}{(d_i+2)^2} = \frac{n_0}{(s+3)^2} + \frac{n-n_0}{(s+2)^2} \leq \frac{n}{(s+2)^2} = \frac{n^3}{(2d+n-n_0)^2} \leq \frac{n^3}{(2d+1)^2}$, since $n_0 \leq n - 1$. \square

5.4 Computing the parameter ϑ_d^h as a generalized eigenvalue problem

As the parameter ϑ_d^h is defined in terms of sum-of-squares polynomials (cf. equation (5.7)), it can be computed by means of a semidefinite program. As we now observe, ϑ_d^h can in fact be computed in a cheaper way as a generalized eigenvalue problem, since the program (5.7) has only one affine constraint. Showing how this can be done, is the concern of the section on hand.

Using the inner product from (5.12), the parameter ϑ_d^h can be rewritten as

$$\begin{aligned} \vartheta_d^h &= \min_{h \in \mathbb{R}[x]} \langle f, h \rangle \quad \text{s.t.} \quad \langle h, T_0 \rangle = 1, \\ h(x) &= \sum_{I \subseteq [n]} \sigma_I(x) \prod_{i \in I} (1 - x_i^2), \\ \sigma_I &\in \Sigma[x]_{2(d-|I|)}, \quad I \subseteq [n]. \end{aligned} \quad (5.26)$$

For convenience, we use below the following notation. For a set $I \subseteq [n]$ and an integer $d \in \mathbb{N}$ we let Λ_d^I denote the set of sequences $\beta \in \mathbb{N}^n$ with $|\beta| \leq \lfloor d - |I| \rfloor$.

As is well known, one can express the condition that σ_I is a sum-of-squares polynomial as a semidefinite program. More precisely, writing $\sigma_I(x) = \sum_k p_k(x)^2$ for some $p_k \in \mathbb{R}[x]$ and using the Chebyshev basis to express the polynomials p_k , we obtain that σ_I is a sum-of-squares polynomial if and only if there exists a matrix variable M^I indexed by Λ_d^I which is positive semidefinite and satisfies

$$\sigma_I = \sum_{\beta, \gamma \in \Lambda_d^I} M_{\beta, \gamma}^I T_\beta T_\gamma. \quad (5.27)$$

For each $I \subseteq [n]$, we introduce the following matrices A^I and B^I , also indexed by the set Λ_d^I , and whose entries are

$$\begin{aligned} A_{\beta, \gamma}^I &= \langle f, T_\beta T_\gamma \prod_{i \in I} (1 - x_i^2) \rangle, \\ B_{\beta, \gamma}^I &= \langle T_0, T_\beta T_\gamma \prod_{i \in I} (1 - x_i^2) \rangle \end{aligned} \quad (5.28)$$

for $\beta, \gamma \in \Lambda_d^I$. We indicate in Appendix 5.B how to compute the matrices A^I and B^I .

We can now reformulate the parameter ϑ_d^h as follows.

Lemma 5.13. *Let A^I and B^I be the matrices defined as in (5.28) for each $I \subseteq [n]$. Then the parameter ϑ_d^h can be reformulated using the following semidefinite program in the matrix variables M^I , $I \subseteq [n]$:*

$$\begin{aligned} \vartheta_d^h &= \min_{M^I} \sum_{I \subseteq [n]} \text{Trace}(A^I M^I) \quad \text{s.t.} \quad \sum_{I \subseteq [n]} \text{Trace}(B^I M^I) = 1, \\ &M^I \succcurlyeq 0, \quad I \subseteq [n]. \end{aligned} \quad (5.29)$$

Proof. Using relation (5.27) we can express the polynomial variable h in (5.26) in terms of the matrix variables M^I and obtain

$$h = \sum_{I \subseteq [n]} \sum_{\beta, \gamma \in \Lambda_d^I} M_{\beta, \gamma}^I T_\beta T_\gamma \prod_{i \in I} (1 - x_i^2).$$

First, this permits to reformulate the objective function $\langle f, h \rangle$ in terms of the matrix variables

M^I in the following way:

$$\begin{aligned}\langle f, h \rangle &= \sum_I \sum_{\beta, \gamma} M_{\beta, \gamma}^I \langle f, T_\beta T_\gamma \prod_{i \in I} (1 - x_i^2) \rangle \\ &= \sum_I \sum_{\beta, \gamma} M_{\beta, \gamma}^I A_{\beta, \gamma}^I \\ &= \sum_I \text{Trace}(A^I M^I).\end{aligned}$$

Second we can reformulate the constraint $\langle T_0, h \rangle = 1$ using

$$\begin{aligned}\langle T_0, h \rangle &= \sum_I \sum_{\beta, \gamma} M_{\beta, \gamma}^I \langle T_0, T_\beta T_\gamma \prod_{i \in I} (1 - x_i^2) \rangle \\ &= \sum_I \sum_{\beta, \gamma} M_{\beta, \gamma}^I B_{\beta, \gamma}^I \\ &= \sum_I \text{Trace}(B^I M^I).\end{aligned}$$

From this it follows that the program (5.26) is indeed equivalent to the program (5.29). \square

The program (5.29) is a semidefinite program with only one constraint. Hence, as we show next, it is equivalent to a generalized eigenvalue problem.

Theorem 5.14. *For $I \subseteq [n]$ let A^I and B^I be the matrices from (5.28) and define the parameter*

$$\lambda^{(I)} = \max \left\{ \lambda \mid A^I - \lambda B^I \succcurlyeq 0 \right\} = \min \left\{ \lambda \mid A^I \mathbf{x} = \lambda B^I \mathbf{x} \text{ for some non-zero vector } \mathbf{x} \right\}.$$

One then has $\vartheta_d^h = \min_{I \subseteq [n]} \lambda^{(I)}$.

Proof. The dual semidefinite program of the program (5.29) is given by

$$\sup \lambda \quad \text{s.t.} \quad A^I - \lambda B^I \succcurlyeq 0, \quad I \subseteq [n]. \quad (5.30)$$

We first show that the primal problem (5.29) is strictly feasible. To see this, it suffices to show that $\text{Trace}(B^I) > 0$, since then one may set M^I equal to a suitable multiple of the identity matrix, and thus one gets a strictly feasible solution to (5.29). Indeed, the matrix B^I is positive semidefinite since, for any scalars g_β ,

$$\sum_{\beta, \gamma} g_\beta g_\gamma B_{\beta, \gamma}^I = \int_{[-1, 1]^n} \left(\sum_{\beta} g_\beta T_\beta \right)^2 \prod_{i \in I} (1 - x_i^2) \nu(dx) \geq 0.$$

Thus, $\text{Trace}(B^I) \geq 0$ and, moreover, $\text{Trace}(B^I) > 0$ since B^I is nonzero.

Moreover, the dual problem (5.30) is also feasible, since $\lambda = \varrho^{\min}$ is a feasible solution. This follows from the fact that the polynomial $f - \varrho^{\min}$ is non-negative over $[-1, 1]^n$, which implies that the matrix $A^I - \varrho^{\min} B^I$ is positive semidefinite. Indeed, using the same argument as above for showing that $B^I \succcurlyeq 0$, we have

$$\sum_{\beta, \gamma} g_\beta g_\gamma (A^I - \varrho^{\min} B^I)_{\beta, \gamma} = \int_{[-1, 1]^n} (f(x) - \varrho^{\min}) \left(\sum_{\beta} g_\beta T_\beta \right)^2 \prod_{i \in I} (1 - x_i^2) \nu(dx) \geq 0.$$

Since the primal problem is strictly feasible and the dual problem is feasible, there is no duality gap and the dual problem attains its supremum. The result follows. \square

5.5 Numerical examples

We examine the polynomial test functions which were also used in [dKLS16] and [dKLLS15], and are described in Appendix 5.C.

The numerical examples given here only serve to illustrate the observed convergence behavior of the sequence ϑ_d^h as compared to the theoretical convergence rate. In particular, the computational demands for computing ϑ_d^h for large d are such that it cannot compete in practice with the known iterative methods referenced at the beginning of the chapter.

For the polynomial test functions we list in Table 5.1 the values of ϑ_d^h for d up to $d = 24$, obtained by solving the generalized eigenvalue problem in Theorem 5.14 using the `eig` function of Matlab. Recall that for step d of the hierarchy the polynomial density function h is of Schmüdgen-type and has degree $2d$.

For the examples listed the computational time is negligible, and therefore not listed; recall that the computation of ϑ_d^h for even n requires the solution of 2^n generalized eigenvalue problems indexed by subsets $I \subset [n]$, where the order of the matrices equals $\binom{n+|d-|I||}{n}$; cf. Theorem 5.14.

Table 5.1 – The upper bounds ϑ_d^h for the test functions.

d	Booth	Matyas	Motzkin	Three-Hump	Styblinski-Tang		Rosenbrock	
					$n = 2$	$n = 3$	$n = 2$	$n = 3$
3	145.3633	4.1844	1.1002	24.6561	-27.4061		157.7604	
4	118.0554	3.9308	0.8764	15.5022	-34.5465	-40.1625	96.8502	318.0367
5	91.6631	3.8589	0.8306	9.9919	-40.0362	-47.6759	68.4239	245.9925
6	71.1906	3.8076	0.8098	6.5364	-47.4208	-55.4061	51.7554	187.2490
7	57.3843	3.0414	0.7309	4.5538	-51.2011	-64.0426	39.0613	142.8774
8	47.6354	2.4828	0.6949	3.3453	-56.0904	-70.2894	30.3855	111.0703
9	40.3097	2.0637	0.5706	2.5814	-58.8010	-76.0311	24.0043	88.3594
10	34.5306	1.7417	0.5221	2.0755	-61.8751	-80.5870	19.5646	71.5983
11	28.9754	1.4891	0.4825	1.7242	-63.9161	-85.4149	16.2071	59.0816
12	24.6380	1.2874	0.4081	1.4716	-65.5717	-88.5665	13.6595	49.5002
13	21.3151	1.1239	0.3830	1.2830	-67.2790		11.6835	
14	18.7250	0.9896	0.3457	1.1375	-68.2078		10.1194	
15	16.6595	0.8779	0.3016	1.0216	-69.5141		8.8667	
16	14.9582	0.7840	0.2866	0.9263	-70.3399		7.8468	
17	13.5114	0.7044	0.2590	0.8456	-71.0821		7.0070	
18	12.2479	0.6363	0.2306	0.7752	-71.8284		6.3083	
19	11.0441	0.5776	0.2215	0.7129	-72.2581		5.7198	
20	10.0214	0.5266	0.2005	0.6571	-72.8953		5.2215	
21	9.1504	0.4821	0.1815	0.6070	-73.3011		4.7941	
22	8.4017	0.4430	0.1754	0.5622	-73.6811		4.4266	
23	7.7490	0.4084	0.1597	0.5220	-74.0761		4.1070	
24	7.1710	0.3778	0.1462	0.4860	-74.3070		3.8283	

We note that the observed rate of convergence seems in line with the $O(\frac{1}{d^2})$ error bound.

As a second numerical experiment, we compare the upper bound ϑ_d^h to the upper bound ϑ_d^{mom} of the Lasserre hierarchy defined in (1.18) with $\nu(dx) = dx$; see Table 5.2. Recall that the bound ϑ_d^{mom} corresponds to using sum-of-squares density functions of degree at most $2d$. As shown in [dKLS16], the computation of ϑ_d^{mom} with ν the Lebesgue measure may be done by solving a single generalized eigenvalue problem with matrices of order $\binom{n+|d-|I||}{n}$. Thus, the computation of ϑ_d^{mom} is significantly cheaper than that of ϑ_d^h .

It is interesting to note that, in almost all cases, $\vartheta_d^h > \vartheta_d^{\text{mom}}$. Thus, even though the measure $\nu(dx)$ and the Schmüdgen-type densities are useful in getting improved error bounds, they mostly do not lead to improved upper bounds for these examples. This also suggests that it might be possible to improve the error result $\vartheta_d^{\text{mom}} - \varrho^{\min} = O(\frac{1}{\sqrt{d}})$ in [dKLS16], at least for the case $\mathbf{K} = [-1, 1]^n$. To illustrate this effect we graphically represent the results of Table 5.2 in Figure 5.3. Note that the bound $\frac{D_f n^3}{(2d+1)^2}$ of Theorem 5.12 would lie far above these graphs. To give

Table 5.2 – Comparison of the upper bounds ϑ_d^h and ϑ_d^{mom} for Booth, Matyas, Three-Hump Camel, and Motzkin functions.

d	Booth function		Matyas function		Three-Hump Camel function		Motzkin polynomial	
	ϑ_d^{mom}	ϑ_d^h	ϑ_d^{mom}	ϑ_d^h	ϑ_d^{mom}	ϑ_d^h	ϑ_d^{mom}	ϑ_d^h
3	118.383	145.3633	4.2817	4.1844	29.0005	24.6561	1.0614	1.1002
4	97.6473	118.0554	3.8942	3.9308	9.5806	15.5022	0.8294	0.8764
5	69.8174	91.6631	3.6894	3.8589	9.5806	9.9919	0.8010	0.8306
6	63.5454	71.1906	2.9956	3.8076	4.4398	6.5364	0.8010	0.8098
7	47.0467	57.3843	2.5469	3.0414	4.4398	4.5538	0.7088	0.7309
8	41.6727	47.6354	2.0430	2.4828	2.5503	3.3453	0.5655	0.6949
9	34.2140	40.3097	1.8335	2.0637	2.5503	2.5814	0.5655	0.5706
10	28.7248	34.5306	1.4784	1.7417	1.7127	2.0755	0.5078	0.5221
11	25.6050	28.9754	1.3764	1.4891	1.7127	1.7242	0.4060	0.4825
12	21.1869	24.6380	1.1178	1.2874	1.2775	1.4716	0.4060	0.4081
13	19.5588	21.3151	1.0686	1.1239	1.2775	1.2830	0.3759	0.3830
14	16.5854	18.7250	0.8742	0.9896	1.0185	1.1375	0.3004	0.3457
15	15.2815	16.6595	0.8524	0.8779	1.0185	1.0216	0.3004	0.3016
16	13.4626	14.9582	0.7020	0.7840	0.8434	0.9263	0.2819	0.2866
17	12.2075	13.5114	0.6952	0.7044	0.8434	0.8456	0.2300	0.2590
18	11.0959	12.2479	0.5760	0.6363	0.7113	0.7752	0.2300	0.2306
19	9.9938	11.0441	0.5760	0.5776	0.7113	0.7129	0.2185	0.2215
20	9.2373	10.0214	0.4815	0.5266	0.6064	0.6571	0.1817	0.2005

an idea for the value of the constants D_f we calculated them for the Booth, Matyas, Three-Hump Camel, and Motzkin functions: $D_{\text{Booth}} \approx 2.6 \cdot 10^5$, $D_{\text{Matyas}} \approx 9.9 \cdot 10^3$, $D_{\text{ThreeHump}} \approx 3.5 \cdot 10^7$, and $D_{\text{Motzkin}} \approx 1.1 \cdot 10^5$.

Finally, it is shown in [dKLS16] that one may obtain feasible points corresponding to bounds like ϑ_d^h through sampling from the probability distribution defined by the optimal density function. In particular, one may use the *method of conditional distributions* (see e.g., [Law07, Section 8.5.1]). For $\mathbf{K} = [0, 1]^n$, the procedure is described in detail in [dKLS16, Section 3].

Appendix

5.A Proof of Lemma 5.4

We give here a proof of Lemma 5.4. We repeat the statement of the Lemma for convenience.

Lemma 5.4. *For any fixed integer $k > 1$, one has*

$$\max_{0 \leq i \leq k-1} |u_i^{(k)}| \leq \max_{0 \leq i \leq k} |t_i^{(k)}| = 2^{k-1-2\psi(k)} \frac{k(k-\psi(k)-1)!}{\psi(k)!(k-2\psi(k))!}, \quad (5.14)$$

where $\psi(k) = 0$ for $k \leq 4$ and $\psi(k) = \left\lceil \frac{1}{8} (4k - 5 - \sqrt{8k^2 - 7}) \right\rceil$ for $k \geq 4$. Moreover, the right-hand side of the equation increases monotonically with increasing k .

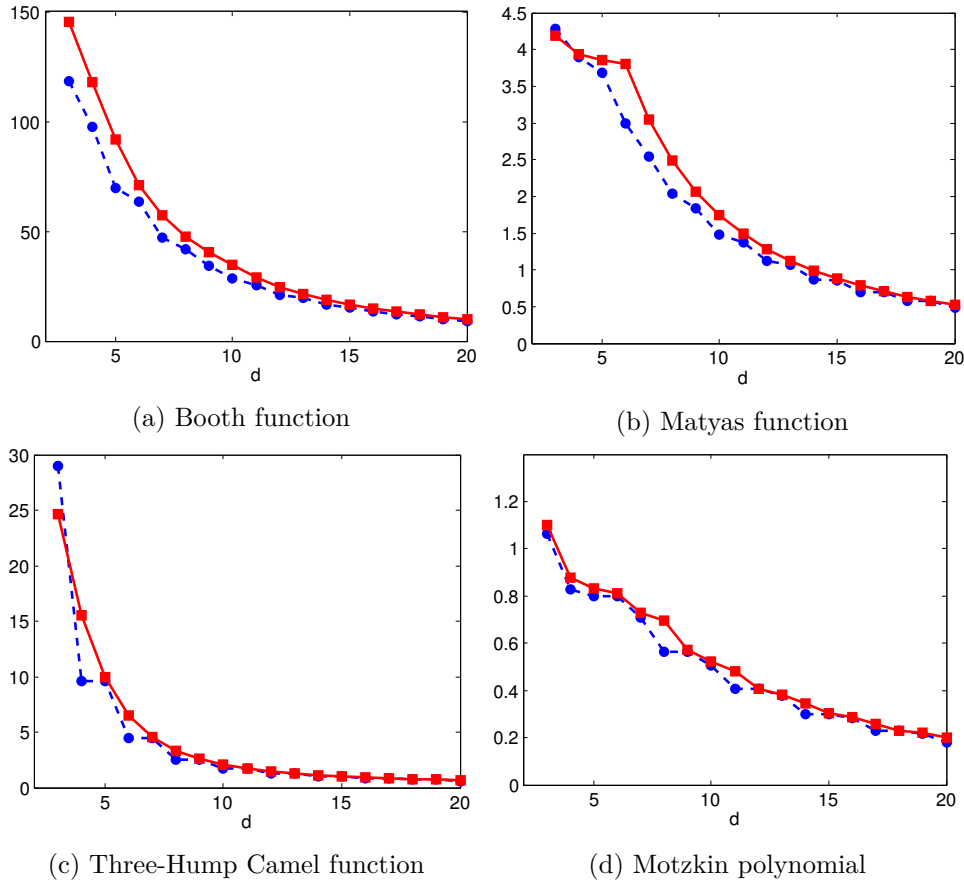


Figure 5.3 – Graphical representation of Table 5.2 to illustrate the comparison of the upper bounds ϑ_d^{mom} and ϑ_d^h . The values ϑ_d^{mom} are marked with circles connected by a dashed line and ϑ_d^h with squares connected by a solid line.

Proof. We recall the representation of the Chebyshev polynomials in the monomial basis:

$$T_k(x) = \sum_{i=0}^k t_i^{(k)} x^i = \frac{k}{2} \sum_{s=0}^{\lfloor k/2 \rfloor} (-1)^s \frac{(k-s-1)!}{s!(k-2s)!} (2x)^{k-2s}, \quad k > 0,$$

$$U_{k-1}(x) = \sum_{i=0}^{k-1} u_i^{(k)} x^i = \sum_{s=0}^{\lfloor \frac{k-1}{2} \rfloor} (-1)^s \frac{(k-s-1)!}{s!(k-1-2s)!} (2x)^{k-1-2s}, \quad k > 1.$$

So, concretely, the coefficients are given by

$$t_{k-2s}^{(k)} = (-1)^s \cdot 2^{k-1-2s} \cdot \frac{k(k-s-1)!}{s!(k-2s)!}, \quad k > 0, \quad 0 \leq s \leq \left\lfloor \frac{k}{2} \right\rfloor,$$

$$u_{k-1-2s}^{(k)} = (-1)^s \cdot 2^{k-1-2s} \cdot \frac{(k-s-1)!}{s!(k-1-2s)!}, \quad k > 1, \quad 0 \leq s \leq \left\lfloor \frac{k-1}{2} \right\rfloor.$$

It follows directly that $t_{k-2s}^{(k)} = \frac{k}{k-2s} u_{k-1-2s}^{(k)}$, and thus $|t_{k-2s}^{(k)}| > |u_{k-1-2s}^{(k)}|$ for $s < \frac{k}{2}$ and all $k > 1$ which implies the inequality on the left-hand side of (5.14).

Now we show that the value of $\max_{0 \leq s \leq \lfloor k/2 \rfloor} |t_{k-2s}^{(k)}|$ is attained for $s = \psi(k)$. For this we

examine the quotient

$$\frac{|t_{k-2(s+1)}^{(k)}|}{|t_{k-2s}^{(k)}|} = \frac{(k-2s)(k-2s-1)}{4(s+1)(k-s-1)} = \frac{k^2 - 4sk + 4s^2 + 2s - k}{4sk - 4s^2 - 8s + 4k - 4}. \quad (5.31)$$

Observe that this quotient is at most 1 if and only if $s_1 \leq s \leq s_2$, where we set $s_1 = \frac{1}{8}(4k - 5 - \sqrt{8k^2 - 7})$ and $s_2 = \frac{1}{8}(4k - 5 + \sqrt{8k^2 - 7})$. Hence, the function $s \mapsto |t_{k-2s}^{(k)}|$ is monotone increasing for $s \leq s_1$ and monotone decreasing for $s_1 \leq s \leq s_2$. Moreover, as $\lceil s_1 \rceil \leq s_1$, we deduce that $|t_{k-2\lceil s_1 \rceil}^{(k)}| \geq |t_{k-2\lfloor s_1 \rfloor}^{(k)}|$. Observe furthermore that $s_1 \geq 0$ if and only if $k \geq 4$, and $s_2 \geq \frac{k}{2}$ for all $k > 1$.

Therefore, in the case $k \geq 4$, $\max_{0 \leq s \leq \lfloor k/2 \rfloor} |t_{k-2s}^{(k)}|$ is attained at $\lceil s_1 \rceil = \psi(k)$, and thus it is equal to $|t_{k-2\psi(k)}^{(k)}|$. In the case $1 < k \leq 4$, $\max_{0 \leq s \leq \lfloor k/2 \rfloor} |t_{k-2s}^{(k)}|$ is attained at $s = 0$, and thus it is equal to $|t_k^{(k)}| = 2^{k-1}$.

Finally, we show that the rightmost term of (5.14) increases monotonically with k . We show the inequality $|t_{k-2\psi(k)}^{(k)}| \leq |t_{k+1-2\psi(k+1)}^{(k+1)}|$ for $k \geq 4$. For this we consider again the sequence of Chebyshev coefficients, but this time we are interested in the behavior for increasing k , i.e., in the map $k \mapsto |t_{k-2s}^{(k)}|$. So, for fixed s , we consider the quotient

$$\frac{|t_{k+1-2s}^{(k+1)}|}{|t_{k-2s}^{(k)}|} = \frac{2^{k-2s}(k+1)(k-s)!s!(k-2s)!}{2^{k-1-2s}k(k-s-1)!s!(k+1-2s)!} = 2 \cdot \frac{k+1}{k} \cdot \frac{k-s}{k+1-2s},$$

which is equal to 2 if $s = 0$, and at least 1 if $s > 0$ since every factor is at least 1. Thus, for $s = \psi(k)$, we obtain

$$|t_{k-2\psi(k)}^{(k)}| \leq |t_{k+1-2\psi(k+1)}^{(k+1)}|. \quad (5.32)$$

Consider the map $\phi: [4, \infty) \rightarrow \mathbb{R}$, $k \mapsto \phi(k) = \frac{1}{8}(4k - 5 - \sqrt{8k^2 - 7})$, so that $\psi(k) = \lceil \phi(k) \rceil$. The map ϕ is monotone increasing, since its derivative $\phi'(k) = \frac{1}{8}\left(4 - \frac{16k}{2\sqrt{8k^2 - 7}}\right) = \frac{\sqrt{8k^2 - 7} - 2k}{2\sqrt{8k^2 - 7}}$ is positive for all $k \geq 4$. Hence, we have $\psi(k) \leq \psi(k+1)$. Then, in view of (5.31) and the comment thereafter, we have $|t_{k+1-2s}^{(k+1)}| \leq |t_{k+1-2(s+1)}^{(k+1)}|$ for $s \leq \psi(k+1)$, and thus

$$|t_{k+1-2\psi(k)}^{(k+1)}| \leq |t_{k+1-2\psi(k+1)}^{(k+1)}|. \quad (5.33)$$

Combining (5.32) and (5.33), we obtain the desired inequality $|t_{k-2\psi(k)}^{(k)}| \leq |t_{k+1-2\psi(k+1)}^{(k+1)}|$. \square

5.B Useful identities for the Chebyshev polynomials

First, recall the notation $\nu(dx)$, denoting the Lebesgue measure with density function $\left(\prod_{i=1}^n \pi \sqrt{1 - x_i^2}\right)^{-1}$. In order to compute the matrices A^I and B^I we need to evaluate the following integrals:

$$\begin{aligned} \langle T_\alpha, T_\beta T_\gamma \prod_{i \in I} (1 - x_i^2) \rangle = \\ \prod_{i \in I} \int_{-1}^1 T_{\alpha_i}(x_i) T_{\beta_i}(x_i) T_{\gamma_i}(x_i) (1 - x_i^2) \nu(dx_i) \cdot \prod_{i \notin I} \int_{-1}^1 T_{\alpha_i}(x_i) T_{\beta_i}(x_i) T_{\gamma_i}(x_i) \nu(dx_i). \end{aligned}$$

Thus we can now assume that we are in the univariate case. Suppose we are given integers $a, b, c \geq 0$ and the goal is to evaluate the integrals

$$\int_{-1}^1 T_a(x)T_b(x)T_c(x) \nu(dx) \quad \text{and} \quad \int_{-1}^1 T_a(x)T_b(x)T_c(x)(1-x^2) \nu(dx).$$

We use the following identities for the (univariate) Chebyshev polynomials:

$$T_a T_b = \frac{1}{2}(T_{a+b} + T_{|a-b|}), \quad T_a T_b T_c = \frac{1}{4}(T_{a+b+c} + T_{|a+b-c|} + T_{|a-b|+c} + T_{||a-b|-c|}),$$

so that

$$\begin{aligned} T_a T_b T_c T_2 &= \frac{1}{8}(T_{a+b+c+2} + T_{|a+b+c-2|} + T_{|a+b-c|+2} + T_{||a+b-c|-2|} \\ &\quad + T_{|a-b|+c+2} + T_{||a-b|+c-2|} + T_{|a-b|-c|+2} + T_{||a-b|-c|-2|}). \end{aligned}$$

Using the orthogonality relation $\int_{-1}^1 T_a \nu(dx) = \delta_{0,a}$, we obtain that

$$\int_{-1}^1 T_a T_b T_c \nu(dx) = \frac{1}{4}(\delta_{0,a+b+c} + \delta_{0,a+b-c} + \delta_{0,|a-b|+c} + \delta_{0,|a-b|-c}).$$

Moreover, using the fact that $1-x^2 = \frac{1-T_2}{2}$, we get

$$\begin{aligned} \int_{-1}^1 T_a T_b T_c (1-x^2) \nu(dx) &= \frac{1}{2} \int_{-1}^1 T_a T_b T_c (1-T_2) \nu(dx) \\ &= \frac{1}{2} \int_{-1}^1 T_a T_b T_c d\nu(x) - \frac{1}{2} \int_{-1}^1 T_a T_b T_c T_2 \nu(dx), \end{aligned}$$

and thus

$$\begin{aligned} \int_{-1}^1 T_a T_b T_c (1-x^2) \nu(dx) &= \frac{1}{8}(\delta_{0,a+b+c} + \delta_{0,a+b-c} + \delta_{0,|a-b|+c} + \delta_{0,|a-b|-c}) \\ &\quad - \frac{1}{16}(\delta_{0,a+b+c-2} + \delta_{0,|a+b-c|-2} + \delta_{0,|a-b|+c-2} + \delta_{0,||a-b|-c|-2}). \end{aligned}$$

5.C Test functions

Here, we list the test functions used for the numerical examples in Section 5.5. Besides giving the functions in the monomial and Chebyshev basis, we state their minimal value ϱ^{\min} , the minimizer(s), their range on the unit box, and the number of variables n for which we used them. Note that we scaled the functions to fit the unit box.

Booth function ($n = 2$, $\varrho^{\min} = f(0.1, 0.3) = 0$, $f([-1, 1]^2) \approx [0, 2500]$):

$$\begin{aligned} f(x) &= (10x_1 + 20x_2 - 7)^2 + (20x_1 + 10x_2 - 5)^2 \\ &= 250(T_2(x_1) + T_2(x_2)) + 800T_1(x_1)T_1(x_2) - 340T_1(x_1) - 380T_1(x_2) + 574. \end{aligned}$$

Matyas function ($n = 2$, $\varrho^{\min} = f(0, 0) = 0$, $f([-1, 1]^2) \approx [0, 100]$):

$$f(x) = 26(x_1^2 + x_2^2) - 48x_1x_2 = 13(T_2(x_1) + T_2(x_2)) - 48T_1(x_1)T_1(x_2) + 26.$$

Motzkin polynomial ($n = 2$, $\varrho^{\min} = f(\pm\frac{1}{2}, \pm\frac{1}{2}) = 0$, $f([-1, 1]^2) \approx [0, 80]$):

$$\begin{aligned} f(x) &= 64(x_1^4 x_2^2 + x_1^2 x_2^4) - 48x_1^2 x_2^2 + 1 = 4(T_4(x_1) + T_4(x_1)T_2(x_2) \\ &\quad + T_2(x_1)T_4(x_2) + T_4(x_2)) + 20T_2(x_1)T_2(x_2) + 16(T_2(x_1) + T_2(x_2)) + 13. \end{aligned}$$

Three-Hump Camel function ($n = 2$, $\varrho^{\min} = f(0, 0) = 0$, $f([-1, 1]^2) \approx [0, 2000]$):

$$\begin{aligned} f(x) &= \frac{5^6}{6} x_1^6 - 5^4 \cdot 1.05 x_1^4 + 50 x_1^2 + 25 x_1 x_2 + 25 x_2^2 \\ &= \frac{5^6}{192} T_6(x_1) + \frac{1625}{4} T_4(x_1) + \frac{58725}{64} T_2(x_1) + 25 T_1(x_1)T_1(x_2) + \frac{25}{2} T_2(x_2) + \frac{14525}{24}. \end{aligned}$$

Styblinski-Tang function ($n = 2$ and $n = 3$, $\varrho^{\min} = -39.17 \cdot n$, $f([-1, 1]^2) \approx [-70, 200]$):

$$f(x) = \sum_{j=1}^n 312.5 x_j^4 - 200 x_j^2 + 12.5 x_j = \sum_{j=1}^n \left(\frac{625}{16} T_4(x_j) + \frac{225}{4} T_2(x_j) + \frac{25}{2} T_1(x_j) + \frac{275}{16} \right).$$

Rosenbrock function ($n = 2$ and $n = 3$, $\varrho^{\min} = 0$, $f([-1, 1]^2) \approx [0, 4000]$):

$$\begin{aligned} f(x) &= \sum_{j=1}^{n-1} 100(2.048 \cdot x_{j+1} - 2.048^2 \cdot x_j^2)^2 + (2.048 \cdot x_j - 1)^2 \\ &= \sum_{j=1}^{n-1} \left[12.5 \cdot 2.048^4 T_4(x_j) - 100 \cdot 2.048^3 T_2(x_j)T_1(x_{j+1}) \right. \\ &\quad \left. + (0.5 + 50 \cdot 2.048^2)2.048^2 T_2(x_j) + 50 \cdot 2.048^2 T_2(x_{j+1}) - 4.096 T_1(x_j) \right. \\ &\quad \left. - 100 \cdot 2.048^3 T_1(x_{j+1}) + 1 + 2.048^2(37.5 \cdot 2.048^2 + 50.5) \right]. \end{aligned}$$

Bibliography

- [ADT07] A. C. Atkinson, A. N. Donev, and R. D. Tobias. *Optimum experimental designs, with SAS*, volume 34 of *Oxford Statistical Science Series*. Oxford University Press, New York, USA, 2007. (Cited on page 42.)
- [AS72] M. Abramowitz and I. A. Stegun, editors. *Handbook of Mathematical Functions with formulas, graphs, and mathematical tables*. Applied Mathematics Series 55. National Bureau of Standards - U.S. Government Printing Office, Washington, D.C., USA, 10th edition, 1972. (Cited on pages 75 and 76.)
- [Ash00] R. B. Ash. *Probability and measure theory*. Academic Press, San Diego, USA, 2nd edition, 2000. (Cited on pages 18 and 32.)
- [Bai08] R. A. Bailey. *Design of comparative experiments*, volume 25 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, UK, 2008. (Cited on page 41.)
- [Bar02] A. Barvinok. *A course in convexity*. American Mathematical Society, Providence, USA, 2002. (Cited on pages 18 and 65.)
- [BCR98] J. Bochnak, M. Coste, and M. F. Roy. *Real algebraic geometry*. Springer, Berlin, Germany, 1998. (Cited on pages 17 and 37.)
- [Ber96] D. P. Bertsekas. *Constrained Optimization and Lagrange Multiplier Methods*. Athena Scientific, Belmont (MA), USA, 1996. (Not cited.)
- [BHH78] G. E. P. Box, W. G. Hunter, and J. S. Hunter. *Statistics for experimenters: An introduction to design, data analysis, and model building*. John Wiley & Sons, New York, USA, 1 edition, 1978. (Cited on page 41.)
- [BHLO06] J. V. Burke, D. Henrion, A. S. Lewis, and M. L. Overton. Stabilization via nonsmooth, nonconvex optimization. *IEEE Transactions on Automatic Control*, 51(11):1760–1769, 2006. (Cited on page 16.)
- [BLO04] J. V. Burke, A. S. Lewis, and M. L. Overton. Variational analysis of the abscissa mapping for polynomials via the Gauß-Lucas theorem. *Journal of Global Optimization*, 28:259–268, 2004. (Cited on pages 15, 16 and 31.)
- [BO01] J. V. Burke and M. L. Overton. Variational analysis of the abscissa mapping for polynomials. *SIAM Journal on Control and Optimization*, 39(6):1651–1676, 2001. (Cited on page 16.)
- [Car04] G. Carlier. On a theorem of Alexandrov. *Journal of Nonlinear and Convex Analysis*, 5(1):49–58, 2004. (Cited on page 3.)
- [Cla13] M. Claeys. *Mesures d’occupation et relaxations semi-définies pour la commande optimale*. PhD thesis, Institut National des Sciences Appliquées de Toulouse, France, 2013. (Not cited.)
- [Cro10] J. A. Cross. *Spectral abscissa optimization using polynomial stability conditions*. PhD thesis, University of Washington, Seattle, USA, 2010. (Cited on pages 16 and 21.)

- [Dav63] P. J. Davis. *Interpolation and Approximation*. Blaisdell Publishing Company, A Division of Ginn and Company, New York, USA, 1963. (Cited on page 37.)
- [dCGH⁺17] Y. de Castro, F. Gamboa, D. Henrion, R. Hess, and J. B. Lasserre. Approximate optimal designs for multivariate polynomial regression. Preprint at [arXiv:1703.01777](https://arxiv.org/abs/1703.01777), June 2017. (Cited on pages 3 and 40.)
- [DG14] H. Dette and Y. Grigoriev. E -optimal designs for second-order response surface models. *The Annals of Statistics*, 42(4):1635–1656, 2014. (Cited on page 42.)
- [dKHL17] E. de Klerk, R. Hess, and M. Laurent. Improved convergence rates for Lasserre-type hierarchies of upper bounds for box-constrained polynomial optimization. *SIAM Journal on Optimization*, 27(1):347–367, 2017. (Cited on pages 3, 72 and 73.)
- [dKLLS15] E. de Klerk, J. B. Lasserre, M. Laurent, and Z. Sun. Bound-constrained polynomial optimization using only elementary calculations. To appear in *Mathematics of Operations Research*, preprint at [arXiv: 1507.04404](https://arxiv.org/abs/1507.04404), July 2015. (Cited on pages 72, 73, 74 and 84.)
- [dKLS16] E. de Klerk, M. Laurent, and Z. Sun. Convergence analysis for Lasserre’s measure-based hierarchy of upper bounds for polynomial optimization. *Mathematical Programming Series A*, pages 1–30, 2016. (Cited on pages 72, 73, 74, 84, 85 and 86.)
- [DS93] H. Dette and W. J. Studden. Geometry of E -optimality. *The Annals of Statistics*, 21(1):416–433, 1993. (Cited on page 42.)
- [DS97] H. Dette and W. J. Studden. *The theory of canonical moments with applications in statistics, probability, and analysis*, volume 338 of *Wiley series in probability and statistics*. John Wiley & Sons, New York, USA, 1997. (Cited on pages 42, 45, 46 and 57.)
- [Elf52] G. Elfving. Optimum allocation in linear regression theory. *The Annals of Mathematical Statistics*, 23(2):255–262, 1952. (Cited on page 41.)
- [Fed10] V. Fedorov. Optimal experimental design. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(5):581–589, 2010. (Cited on page 42.)
- [Fle87] R. Fletcher. *Practical Methods of Optimization*. John Wiley & Sons, Inc., New York, USA, 2nd edition, 1987. (Not cited.)
- [FNTB17] S.-I. Filip, Y. Nakatsukasa, L. N. Trefethen, and B. Beckermann. Rational minimax approximation via adaptive barycentric representations. Preprint at [arXiv:1705.10132](https://arxiv.org/abs/1705.10132), May 2017. (Cited on page 38.)
- [GMW81] P. E. Gill, W. Murray, and M. H. Wright. *Practical Optimization*. Academic Press, New York, USA, 1981. (Not cited.)
- [GN11] J. Gouveia and T. Netzer. Positive polynomials and projections of spectrahedra. *SIAM Journal on Optimization*, 21(3):960–976, 2011. (Not cited.)
- [GPT10] J. Gouveia, P. A. Parrilo, and R. R. Thomas. Theta bodies for polynomial ideals. *SIAM Journal on Optimization*, 20(4):2097–2118, 2010. (Cited on page 63.)
- [Hen08] D. Henrion. On semidefinite representations of plane quartics. [arXiv:0809.1826](https://arxiv.org/abs/0809.1826), 2008. (Cited on page 66.)

- [HHLP16] R. Hess, D. Henrion, J. B. Lasserre, and T. S. Pham. Semidefinite approximations of the polynomial abscissa. *SIAM Journal on Control and Optimization*, 54(3):1633–1656, 2016. (Cited on pages 2 and 15.)
- [HL12] D. Henrion and J. B. Lasserre. Inner approximations for polynomial matrix inequalities and robust stability regions. *IEEE Transactions on Automatic Control*, 57(6):1456–1467, 2012. (Cited on pages 16, 20, 21 and 22.)
- [HLL09] D. Henrion, J. B. Lasserre, and J. Löfberg. GloptiPoly 3: moments, optimization and semidefinite programming. *Optimization Methods and Software*, 24(4–5):761–779, 2009. (Cited on page 4.)
- [HPAŠ03] D. Henrion, D. Peaucelle, D. Arzelier, and M. Šebek. Ellipsoidal approximation of the stability domain of a polynomial. *IEEE Transactions on Automatic Control*, 48(12):2255–2259, 2003. (Cited on pages 16 and 22.)
- [HR15] P. Hungerländer and F. Rendl. A feasible active set method for strictly convex quadratic problems with simple bounds. *SIAM Journal on Optimization*, 25(3):1633–1659, 2015. (Not cited.)
- [HZ06] W. W. Hager and H. Zhang. A new active set algorithm for box constrained optimization. *SIAM Journal on Optimization*, 17(2):526–557, 2006. (Not cited.)
- [IS01] L. Imhof and W. J. Studden. E -optimal designs for rational models. *The Annals of Statistics*, 29(3):763–783, 2001. (Cited on page 42.)
- [Kie74] J. Kiefer. General equivalence theory for optimum designs (approximate theory). *The Annals of Statistics*, 2(5):849–879, 1974. (Cited on pages 41 and 46.)
- [KN77] M. G. Krein and A. A. Nudelman. *The Markov moment problem and extremal problems*, volume 50 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, USA, 1977. (Cited on page 42.)
- [Kor16] M. Korda. *Moment-sum-of-squares hierarchies for set approximation and optimal control*. PhD thesis, École Polytechnique Fédérale de Lausanne, Switzerland, 2016. (Not cited.)
- [KS66a] S. Karlin and W. J. Studden. Optimal experimental designs. *The Annals of Mathematical Statistics*, 37(4):783–815, 1966. (Cited on page 41.)
- [KS66b] S. Karlin and W. J. Studden. *Tchebycheff systems: With applications in analysis and statistics*, volume 15 of *Pure and Applied Mathematics*. Interscience Publishers, New York, USA, 1966. (Cited on page 41.)
- [KS17] T.-L. Kriel and M. Schweighofer. On the exactness of Lasserre relaxations for compact convex basic closed semialgebraic sets. Preprint on [arXiv:1704.07231](https://arxiv.org/abs/1704.07231), April 2017. (Cited on pages 66 and 70.)
- [KW17] S. Kuriki and H. P. Wynn. Optimal experimental design that minimizes the width of simultaneous confidence bands. Preprint at [arXiv:1704.03995](https://arxiv.org/abs/1704.03995), April 2017. (Cited on page 42.)
- [Las01] J. B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, 11(3):796–817, 2001. (Cited on pages 1 and 2.)

- [Las09] J. B. Lasserre. Convex sets with semidefinite representations. *Mathematical Programming*, 149:457–477, 2009. (Cited on pages 3, 63, 65, 66, 68 and 70.)
- [Las10] J. B. Lasserre. *Moments, positive polynomials and their applications*. Imperial College Press, London, UK, 2010. (Cited on pages 5, 10, 45, 48 and 56.)
- [Las11] J. B. Lasserre. A new look at nonnegativity on closed sets and polynomial optimization. *SIAM Journal on Optimization*, 21(3):864–885, 2011. (Cited on pages 1, 5, 12, 73 and 74.)
- [Las15a] J. B. Lasserre. A generalization of Löwner-John’s ellipsoid theorem. *Mathematical Programming*, 152(1-2):559–591, 2015. (Cited on pages 47, 48 and 51.)
- [Las15b] J. B. Lasserre. *An introduction to polynomial and semi-algebraic optimization*. Cambridge University Press, Cambridge, UK, 2015. (Cited on pages 1, 5, 7 and 11.)
- [Lau03] M. Laurent. A comparison of the Sherali-Adams, Lovász-Schrijver and Lasserre relaxations for 0–1 programming. *Mathematics of Operations Research*, 28(3):470–496, 2003. (Cited on page 1.)
- [Lau09] M. Laurent. Sums of squares, moment matrices and polynomial optimization. In M. Putinar and S. Sullivan, editors, *Emerging applications of algebraic geometry*, volume 149 of *IMA Volumes in Mathematics and its Applications*, pages 157–270. Springer, New York, USA, 2009. (Cited on pages 1, 5, 6, 9, 11, 19 and 73.)
- [Law07] A. M. Law. *Simulation Modeling and Analysis*. Mc Graw-Hill, Boston, USA, 4th edition, 2007. (Cited on page 86.)
- [Led04] M. Ledoux. Differential operators and spectral distributions of invariant ensembles from the classical orthogonal polynomials. The continuous case. *Electronic Journal of Probability*, 9(7):177–208, 2004. (Cited on page 49.)
- [Lew96] A. S. Lewis. Convex analysis on the Hermitian matrices. *SIAM Journal on Optimization*, 6(1):164–177, 1996. (Cited on page 46.)
- [LN07] J. B. Lasserre and T. Netzer. SOS approximations of nonnegative polynomials via simple high degree perturbations. *Mathematische Zeitschrift*, 256(1):99–112, 2007. (Cited on pages 46 and 50.)
- [Löf04] J. Löfberg. YALMIP: A toolbox for modeling and optimization in MATLAB. *IEEE CACSD Conference*, 2004. users.isy.liu.se/johanl/yalmip/. (Cited on page 4.)
- [LP16] J. B. Lasserre and E. Pauwels. Sorting out typicality with the inverse moment matrix SOS polynomial. In *Advances in Neural Information Processing Systems*, pages 190–198, 2016. (Cited on page 49.)
- [LQX05] M.-Q. Liu, H. Qin, and M.-Y. Xie. Discrete discrepancy and its application in experimental design. In J. Fan and G. Li, editors, *Contemporary multivariate analysis and design of experiments*, volume 2 of *Series in Biostatistics*, pages 227–241. World Scientific Publishing, Singapore, 2005. (Cited on page 41.)
- [LS85] T.-S. Lau and W. J. Studden. Optimal designs for trigonometric and polynomial regression using canonical moments. *The Annals of Statistics*, 13(1):383–394, 1985. (Cited on page 42.)

- [Mar08] M. Marshall. *Positive Polynomials and Sums of Squares*. Mathematical Surveys and Monographs 146. American Mathematical Society, Providence, USA, 2008. (Cited on page 73.)
- [MOS15] MOSEK. *Mosek ApS*. Copenhagen, Denmark, 2015. www.mosek.com. (Cited on page 4.)
- [MYZ15] P. Maréchal, J. J. Ye, and J. Zhou. K -optimal design via semidefinite programming and entropy optimization. *Mathematics of Operations Research*, 40(2):495–512, 2015. (Cited on page 42.)
- [MZ04] I. Molchanov and S. Zuyev. Optimisation in space of measures and optimal design. *ESAIM: Probability and Statistics*, 8:12–24, 2004. (Cited on pages 42 and 61.)
- [Nes00] Y. Nesterov. Squared functional systems and optimization problems. In H. Frenk, K. Roos, T. Terlaky, and S. Zhang, editors, *High Performance Optimization*, volume 33 of *Applied Optimization*, pages 405–440. Springer, New York, USA, 2000. (Cited on page 2.)
- [Nie14] J. Nie. The \mathcal{A} -truncated K -moment problem. *Foundations of Computational Mathematics*, 14(6):1243–1276, 2014. (Cited on page 55.)
- [NPS10] T. Netzer, D. Plaumann, and M. Schweighofer. Exposed faces of semidefinitely representable sets. *SIAM Journal on Optimization*, 20(4):1944–1955, 2010. (Cited on pages 66, 67 and 70.)
- [Pál10] L. Pál. *Global optimization algorithms for bound constrained problems*. PhD thesis, University of Szeged, Hungary, 2010. Available at http://www2.sci.u-szeged.hu/fokozatok/PDF/Pal_Laszlo/Diszertacio_PalLaszlo.pdf. (Not cited.)
- [Par00] P. A. Parrilo. *Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization*. PhD thesis, California Institute of Technology, Pasadena, USA, 2000. (Cited on page 2.)
- [Par03] P. A. Parrilo. Semidefinite programming relaxations for semialgebraic problems. *Mathematical Programming B*, 96:293–320, 2003. (Cited on page 2.)
- [PD01] A. Prestel and C. N. Delzell. *Positive Polynomials - From Hilbert's 17th Problem to Real Algebra*. Springer Monographs in Mathematics. Springer, Berlin, Germany, 2001. (Cited on page 73.)
- [PH11] M.-J. Park and S.-P. Hong. Rank of Handelman hierarchy for max-cut. *Operations Research Letters*, 39(5):323–328, 2011. (Not cited.)
- [PH13] M.-J. Park and S.-P. Hong. Handelman rank of zero-diagonal quadratic programs over a hypercube and its applications. *Journal of Global Optimization*, 56(2):727–736, 2013. (Not cited.)
- [PM12] L. Pronzato and W. G. Müller. Design of computer experiments: space filling and beyond. *Statistics and Computing*, 22(3):681–701, 2012. (Cited on page 41.)
- [PR00] V. Powers and B. Reznick. Polynomials that are positive on an interval. *Transactions of the American Mathematical Society*, 352(10):4677–4692, 2000. (Cited on page 80.)

- [Puk06] F. Pukelsheim. *Optimal design of experiments*, volume 50 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics, Philadelphia, USA, 2006. (Cited on pages 46, 47, 50 and 51.)
- [Put93] M. Putinar. Positive polynomials on compact semi-algebraic sets. *Indiana University Mathematics Journal*, 42(3):969–984, 1993. (Cited on page 6.)
- [PW94] L. Pronzato and E. Walter. Minimal volume ellipsoids. *International Journal of Adaptive Control and Signal Processing*, 8(1):15–30, 1994. (Cited on page 42.)
- [Riv90] T. J. Rivlin. *Chebyshev polynomials: From Approximation Theory to Algebra and Number Theory*. Pure and Applied Mathematics. John Wiley & Sons, New York, USA, 2nd edition, 1990. (Not cited.)
- [Sag11] G. Sagnol. Computing optimal designs of multiresponse experiments reduces to second-order cone programming. *Journal of Statistical Planning and Inference*, 141(5):1684–1708, 2011. (Cited on page 42.)
- [Sch91] K. Schmüdgen. The K -moment problem for compact semi-algebraic sets. *Mathematische Annalen*, 289:203–206, 1991. (Cited on pages 73 and 74.)
- [Sch12] R. Schwabe. *Optimum designs for multi-factor models*, volume 113 of *Lecture Notes in Statistics*. Springer, New York, USA, 2012. (Cited on page 42.)
- [Sch16] C. Scheiderer. Semidefinitely representable convex sets. Preprint at [arXiv:1612.07048](https://arxiv.org/abs/1612.07048), December 2016. (Cited on pages 7 and 66.)
- [SH15] G. Sagnol and R. Harman. Computing exact D -optimal designs by mixed integer second-order cone programming. *The Annals of Statistics*, 43(5):2198–2224, 2015. (Cited on page 42.)
- [Sho87] N. Z. Shor. Quadratic optimization problems. *Soviet journal of computer and systems sciences*, 25:1–11, 1987. (Cited on page 2.)
- [Sho98] N. Z. Shor. *Nondifferentiable Optimization and Polynomial Problems*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998. (Cited on page 2.)
- [Sin10] R. Sinn. *Spectrahedra and a Relaxation of Convex Semialgebraic Sets*. PhD thesis, Universität Konstanz, Germany, 2010. (Cited on page 65.)
- [Stu99] J. F. Sturm. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software*, 11–12:625–653, 1999. Version 1.3 available from <http://sedumi.ie.lehigh.edu>. (Cited on page 4.)
- [VBW98] L. Vandenberghe, S. Boyd, and S.-P. Wu. Determinant maximization with linear matrix inequality constraints. *SIAM Journal on Matrix Analysis and Applications*, 19(2):499–533, 1998. (Cited on page 42.)
- [WPN97] E. Walter, L. Pronzato, and J. Norton. *Identification of Parametric Models from Experimental Data*. Springer, London, UK, 1997. (Cited on page 41.)
- [WWAF06] A. Weisse, G. Wellein, A. Alvermann, and H. Fehske. The kernel polynomial method. *Review of Modern Physics*, 78:275–306, 2006. (Cited on pages 75, 77, 78 and 79.)

-
- [Wyn70] H. P. Wynn. The sequential generation of D -optimum experimental designs. *The Annals of Mathematical Statistics*, 41(5):1655–1664, 1970. (Cited on page 42.)
- [Zor04] V. A. Zorich. *Mathematical analysis II*. Springer, Berlin, Germany, 2004. (Cited on pages 17 and 20.)