Conic optimization with applications in finance and approximation theory

Kirschner, Felix

DOI:
10.26116/x0f5-4k25

Publication date:
2023

Document Version
Publisher's PDF, also known as Version of record

Conic Optimization with Applications in Finance and Approximation Theory

FELIX C. KIRSCHNER

Conic Optimization with Applications in Finance and Approximation Theory

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan Tilburg University op gezag van de rector magnificus, prof. dr. W.B.H.J. van de Donk, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in de Aula van de Universiteit op

dinsdag 4 juli 2023 om 13.30 uur

door

FELIX CONSTANTIN KIRSCHNER

geboren te Hof, Duitsland.

PROMOTOR:	prof. dr. Etienne de Klerk
COPROMOTOR:	dr. Juan Vera Lizcano
PROMOTIECOMMISSIE:	dr. Georgina Hall (INSEAD) dr. ir. Pieter Kleer (Tilburg University) prof. dr. Monique Laurent (Tilburg University) prof. dr. Renata Sotirov (Tilburg University) prof. dr. Frank Vallentin (University of Cologne)

This project has received funding from the European Union's Framework Programme for Research and Innovation Horizon 2020 under the Marie Skłodowska-Curie grant agreement No. 813211 (POEMA).

Conic Optimization with Applications in Finance and Approximation Theory

Acknowledgments

First and foremost, I would like to express my gratitude to my supervisor, Etienne de Klerk. Your guidance was excellent, and the occasional 'have faith, Felix' kept me going. I greatly enjoyed working under your supervision and remember fondly the many 'off-topic' conversations we had in many places of this world. I learned a great deal from you, honestly. Thank you.

I also want to thank especially my former colleague Daniel Brosch, for countless virtual coffee breaks during the pandemic, your motivation, encouragement and help. For your academic career as well as all your personal goals I wish you all the best.

To my co-authors Didier Henrion, Milan Korda, Jean-Beanard Lasserre and Viktor Magron many thanks for making our collaboration such an enjoyable project.

To my co-supervisor Juan Vera as well as my PhD committee consisting of Georgina Hall, Pieter Kleer, Monique Laurent, Renata Sotirov and Frank Vallentin, thank you for agreeing to serve on the committee as well as the valuable feedback on the thesis.

The CWI crowd shall be mentioned as well: Lucas, Andries, Luis and Sven thank you for all the fun we had in the reading group, at conferences and workshops. It made my PhD so much more enjoyable.

Arne, the time we spent doing exercises, studying and preparing talks together during the last year of our masters really motivated me to pursue a PhD. For that and the many dinners in the last 4 years, during which we often discussed our work: Thank you.

To my parents Bettina and Thomas, and my sisters Nathalie and Livia as well as Betty and Klaus thank you for all the support and your faith in me, I am glad I have you.

Finally, I would like to thank my friend Julius Jedryka for designing the cover of this thesis. You did a great job.

Contents

List of Figures **ix**

List of Tables **xi**

List of Algorithms **xiii**

Notation **xv**

1 Introduction **1**
1.1 Conic optimization 2
1.1.1 Duality for conic programs 3
1.1.2 Special cases of finite dimensional conic programming . . . 4
1.1.3 Special cases of infinite dimensional conic programming . . 7
1.2 The moment-SOS hierarchy 10
1.3 Thesis overview 14
1.4 Contributions to the Literature 16

2 A predictor-corrector method for semidefinite programming using the factor width cone **17**
2.1 Introduction 17
2.1.1 Iterative approximation scheme 18
2.1.2 The factor width cone 20
2.1.3 A predictor-corrector method 21
2.2 Barrier functionals for $\mathbb{S}^n_+$ and $\mathrm{FW}_n(k)$ 25
2.3 Relations of the barrier functions 28
2.3.1 Relation of the Newton decrements 29
2.4 Complexity analysis 33
2.5 Discussion and future prospects 38

3 Convergence rates of RLT and Lasserre-type hierarchies 41
3.1 Introduction . . . 41
3.2 A linear relaxation hierarchy over the simplex . . . 43
3.2.1 Convergence analysis . . . 47
3.3 Moment-SOS hierarchy over the sphere . . . 50
3.4 Limiting behavior of the hierarchies of linear operators . . . 52
3.4.1 The simplex case . . . 53
3.4.2 The sphere case . . . 56
3.5 Concluding remarks . . . 57

4 Computing bounds for option prices 63
4.1 Introduction . . . 63
4.1.1 Prior work . . . 64
4.1.2 Contribution of this chapter . . . 66
4.1.3 Outline of the chapter . . . 66
4.2 Bounds on options via the GMP formulation . . . 67
4.2.1 Problem statement . . . 67
4.2.2 Existence of an optimal solution . . . 69
4.3 Bounding the support . . . 73
4.3.1 Atomic representation . . . 73
4.4 Examples for outer range . . . 74
4.4.1 Univariate case . . . 75
4.4.2 Explicit examples with two assets . . . 77
4.5 Relaxations of the non-compact case . . . 84
4.5.1 Lasserre hierarchy of inner range . . . 85
4.5.2 Univariate example . . . 87
4.6 Conclusion and further research . . . 88

5 Construction of approximation kernels via semidefinite programming 91
5.1 Chebyshev polynomials . . . 94
5.2 The kernel polynomial method . . . 97
5.2.1 Trigonometric polynomials . . . 97
5.2.2 Constructing optimal kernels . . . 99
5.2.3 Uniform convergence in terms of σ_r in the multivariate case 100
5.2.4 Positivstellensatz for the multivariate case . . . 105
5.3 Reformulations of the SDP . . . 107
5.3.1 Existence of a real solution . . . 107

5.3.2 Sampling based formulation . . . 108
5.3.3 Equating coefficients via trigonometric identities . . . 109
5.4 Symmetry reduction . . . 111
5.4.1 Symmetry adapted basis . . . 111
5.4.2 Construction of the symmetry adapted SDP . . . 114
5.5 Comparison to products of univariate minimum resolution kernels . 116
5.6 Numerical computations . . . 118
5.6.1 Decoupling the degrees . . . 120
5.7 Concluding remarks . . . 122

6 Conclusion and outlook 125

A Interior Point Methods 129
A.1 The central path . . . 131
A.2 Newton's method . . . 132
A.3 Short-step method . . . 135
A.4 Predictor-Corrector method . . . 138

B Symmetry reduction 143
B.1 Groups and their representations . . . 143
B.1.1 Representation theory . . . 144
B.2 Symmetric polynomials . . . 146
B.2.1 Block-diagonalization . . . 146
B.3 Specht modules . . . 149
B.3.1 Partitions and cycle types . . . 149

C Additional notes on approximation kernels 157
C.1 A note on Jackson kernels . . . 157
C.1.1 Jackson kernels . . . 161
C.1.2 Jackson's original work . . . 162
C.2 Plots and tables . . . 164

Bibliography 177

List of Figures

2.1 Visualization of the surjection from $\mathbb{S}_+^{(n,k)}$ to $\mathrm{FW}_n(k)$ 30

4.1 Visualization of the segmentation of the interval $[0,B]$ 75

4.2 Example of how the support might be split 79

5.1 Plot of a continuous function f and its approximation via the truncation of the series (5.2) at degree $r = 64$ visualizing Gibbs oscillations. 93

5.2 Plot of a continuous function f and its approximation via the minimum resolution kernel defined via (5.12) for degree $r = 64$ visualizing absence of Gibbs oscillations. 101

5.3 Comparison of plot of the function $(x,y) \mapsto \sin(2\pi x)|y|$ and its degree $r = 50$ approximation via the minimum resolution kernel. . 117

5.4 Comparison of uniform approximation errors of several approximations of the function $q(\mathbf{x}) := x_2 \sin(2\pi x_1)$. We plotted the errors for the kernel with minimal resolution σ_r and for the product of two univariate degree $r/2$ kernels, i.e., $K^{\mathrm{KPM}}_{r/2}$ as in (5.30). 119

5.5 Comparison of uniform approximation errors of several approximations of the peaks function $p(\mathbf{x}) := 3(1-x_1)^2 \exp(-x_1^2-(x_2+1)^2) - 10(x_1/5-x_1^3-x_2^5)\exp(-x_1^2-x_2^2) - (1/3)\exp(-(x_1+1)^2-x_2^2)$. We plotted the errors for the kernel with minimal resolution σ_r and for the product of two univariate degree $r/2$ kernels, i.e., $K^{\mathrm{KPM}}_{r/2}$ as in (5.30). 119

5.6 Plot $\sigma^2_{r,r'}$ vs r' for $r = 3,\ldots,10$ and $r' = r, r+1,\ldots,20$ and $n = 2$. . 121

5.7 Plot $\sigma^2_{r,r'}$ vs r' for $r = 2,\ldots,10$ and $r' = r, r+1,\ldots,20$ and $n = 3$. 121

5.8 Plot σ^2_r vs r for $n = 1,2,3,4$ ($\otimes, \square, \triangle, \diamond$ resp.) 122

A.1 Visualization of predictor-corrector method. Initial feasible solution close to central path (red) is given by $x^{(0)}$. Algorithm performs predictor step returning y_0. Corrector steps are taken until point close enough to central path ($x^{(1)}$) is found. Next predictor step returns new point y_0. Corrector steps are taken until $x^{(2)}$ is found, which is close enough to central path to perform next predictor step. After one corrector step the final point $x^{(4)}$ is ε-close to x^*. . . 139

C.1 Approximation of the function $(x,y) \mapsto \sin(2\pi x)y$ via the kernel with respect to σ_r for $r = 10, 20, \ldots, 50$ as well as a plot of the function itself. 173

C.2 Approximation of the *peaks* function $p(x,y) := 3(1-x)^2 \exp(-x^2-(y+1)^2) - 10(x/5 - x^3 - y^5)\exp(-x^2-y^2) - (1/3)\exp(-(x+1)^2 - y^2)$ via the kernel with respect to σ_r for $r = 10, 20, \ldots, 50$ as well as a plot of the function itself in graph C.2f. 174

C.3 Convolution of Dirac-δ at $(0,0)$ with optimal kernel with respect to σ_r for different values of r 175

List of Tables

4.1 Prices of European call options on the Microsoft stock from July '98 with strikes k_i 77
4.2 Strikes and corresponding prices for European call options 81
4.3 Optimal lower and upper bounds w.r.t. the data given in Table 4.2 . 81
4.4 Optimal lower and upper bounds for a currency basket option with different strikes for level $r = 1$ 82
4.5 Revisiting an example from Boyle and Lin, computing bounds on prices of a basket options given means and covariance of the underlying assets for different strikes. 84
4.6 Strikes and corresponding prices for European call options observed on March 1st 2022, all prices in USD. 84
4.7 Bounds for basket options on tech firms subject to observable data of Table 4.6 85
4.8 Optimal solutions for the level-r relaxation of the measure-based Lasserre hierarchy applied to the ε_r relaxation given in (4.15) for Laguerre basis with varying ε_r for $r = 2, \ldots, 7$. The ε_r are the smallest possible such that the resulting SDP still has a feasible solution. 89

5.1 Computational results for σ_r^2 and $\sigma_{r,\mathrm{KPM}}^2$ for different values of n and r obtained by solving (5.29). 123

C.1 Comparison of coefficients for $A_4(t)$ and $B_4(t)$. 162
C.2 Comparison for size of SDP (5.29) and (5.26) for different values of n and r. 166
C.3 Coefficients of minimum resolution kernel for $r = 5$, $n = 2$. 167
C.4 Coefficients of minimum resolution kernel for $r = 10$, $n = 2$. 167
C.5 Coefficients of minimum resolution kernel for $r = 15$, $n = 2$ 167

C.6 Coefficients of minimum resolution kernel for $r = 20$, $n = 2$ 168
C.7 Coefficients of minimum resolution kernel for $r = 25$, $n = 2$ 169
C.8 Coefficients of minimum resolution kernel for $r = 5$, $n = 3$. 170
C.9 Coefficients of minimum resolution kernel for $r = 10$, $n = 3$. 170
C.10 Coefficients of minimum resolution kernel for $r = 15$, $n = 3$. 171
C.11 Coefficients of minimum resolution kernel for $r = 5$, $n = 4$. 172
C.12 Coefficients of minimum resolution kernel for $r = 10$, $n = 4$. 172

List of Algorithms

2.1 Predictor-Corrector SDP algorithm using $\mathrm{FW}_n(k)$ 24
2.2 Subroutine `Predictor_Step`$(\mathcal{A}, A_0, \sigma)$ 24
2.3 Subroutine `Corrector_Step`$(\mathcal{A}, f, x^{(0)})$ 24

Notation

Set notation

$\emptyset$	the empty set
$A \cup B$	the union of the sets A and B
$A \cap B$	the intersection of the sets A and B
$A \subseteq B$	A is a subset of B
$A \setminus B$	Set difference, i.e., $\{a \in A : a \notin B\}$
$\mathrm{cl}(A)$	the closure of A
$\mathrm{int}(A)$	the interior of A

Special sets

$[n]$	the set of integers from 1 to n
$\mathbb{N}$	the natural numbers
$\mathbb{N}^n$	the set of n-tupels of the natural numbers
$\mathbb{Z}$	the integer number
$\mathbb{R}$	the real numbers
$\mathbb{R}_+$	the non-negative reals
$\mathbb{R}^n$	the n-dimensional vectors of reals
$\mathbb{R}^n_+$	the n-dimensional vectors of non-negative reals
$\mathbb{C}$	the complex numbers
$\mathbb{S}^n$	the set of $n \times n$ real symmetric matrices
$\mathbb{S}^{(n,k)}$	the direct product of $\binom{n}{k}$ copies of $\mathbb{S}^k$
$\mathbb{S}^n_+$	the set of $n \times n$ positive semidefinite matrices
$\mathbb{S}^{(n,k)}_+$	the direct product of $\binom{n}{k}$ copies of $\mathbb{S}^k_+$
$\mathrm{FW}_n(k)$	the set of $n \times n$ matrices of factor width $\leq k$

Δ_{n-1}	the standard probability simplex $\{\mathbf{x} \in \mathbb{R}^n_+ : x_1 + \cdots + x_n = 1\}$
$\mathcal{S}^{n-1}$	the unit sphere $\{\mathbf{x} \in \mathbb{R}^n : x_1^2 + \cdots + x_n^2 = 1\}$
$\mathcal{C}(\mathbf{K})$	the set of continuous functions over $\mathbf{K}$
$\mathcal{C}(\mathbf{K})_+$	the cone of continuous functions non-negative over $\mathbf{K}$
$\mathcal{M}(\mathbf{K})$	the set of finite Borel measures supported on $\mathbf{K}$
$\mathcal{M}(\mathbf{K})_+$	the cone of positive finite Borel measures supported on $\mathbf{K}$
$L^1_{2\pi}$	Lebesgue-integrable periodic functions over $[-\pi, \pi]$

Functions

$\delta_{\alpha,\beta}$	Kronecka-δ function
$\Delta(f, x)$	Newton decrement of f at x

Inner Products, dualities, and Norms

$\langle \cdot, \cdot \rangle_X$	bilinear function for dual system (X, X^*)
$\langle \cdot, \cdot \rangle_x$	local inner product at x
$\| \cdot \|_x$	norm induced by $\langle \cdot, \cdot \rangle_x$
$\langle \cdot, \cdot \rangle_{(n,k)}$	trace inner product on $\mathbb{S}^{(n,k)}$
$\| \cdot \|_{(n,k)}$	norm induced by $\langle \cdot, \cdot \rangle_{(n,k)}$
$\langle \cdot, \cdot \rangle_\mu$	inner product induced by measure μ

Polynomials and moments

α	exponent vector $\alpha = (\alpha_1, \ldots \alpha_n) \in \mathbb{N}^n$
$\|\alpha\|$	degree induced by α, i.e., $\alpha_1 + \ldots + \alpha_n$
$\mathbb{N}^n_r$	set of n-tuples α such that $\|\alpha\| \leq r$
$\mathbf{x}^\alpha$	the monomial $x_1^{\alpha_1} \cdot \ldots \cdot x_n^{\alpha_n}$
$\mathbb{R}[\mathbf{x}]$	the polynomial ring in n variables
$\mathbb{R}[\mathbf{x}]_r$	the polynomial ring in n variables of degree less than r
$\Sigma[\mathbf{x}]$	the set of sums of squares of polynomials
$\Sigma[\mathbf{x}]_r$	the set of sums of squares of polynomials up to degree $2r$
$[\mathbf{x}]_r$	monomial basis vector up to degree r

$L_{\mathbf{y}}$	Riesz linear functional
$M(\mathbf{y})$	moment matrix with respect to $\mathbf{y}$
$M_r(\mathbf{y})$	truncated moment matrix with respect to $\mathbf{y}$
$M_r(g * \mathbf{y})$	localizing matrix with respect to polynomial g and sequence $\mathbf{y}$

Linear Algebra

I	identity matrix (of fitting size)
I_n	$n \times n$ identity matrix
J	all-ones matrix (of fitting size)
J_n	$n \times n$ all-ones matrix
e	all-ones vector
e_i	i-th standard unit vector
Tr	the trace operator
P_L	orthogonal projection operator onto subspace L
$A \circ B$	Hadamard product
span(A)	real span of elements of A
A^*	adjoint operator of A
$v^\top$	transpose of vector v (or matrix)
$\mathcal{A}$	linear operator $\mathcal{A}(X) = (\langle A_1, X\rangle, \ldots, \langle A_m, X\rangle)$

1 Introduction

In this thesis we study conic optimization problems and consider applications in finance and approximation theory. An optimization problem consists of the task of finding an element x in a set $\mathcal{X}$, called the *feasible* set, such that x minimizes a cost function $f : \mathcal{X} \to \mathbb{R}$, i.e., such that $f(x) \leq f(y)$ for all elements $y \in \mathcal{X}$. Generic optimization problems do not provide enough structure to permit a useful analysis. To get a grip on things we begin by making the assumption that the cost function is linear, and the feasible set $\mathcal{X}$ is *convex*. A convex set $\mathcal{X}$ is characterized by the fact that if $x, y \in \mathcal{X}$, then all points on the line-segment joining x and y also lie in $\mathcal{X}$, i.e., $\lambda x + (1-\lambda)y \in \mathcal{X}$ for all $\lambda \in [0,1]$. We call optimization problems with convex feasible set and linear cost function *convex optimization problems*. We will, however, consider a more general setting, called *conic optimization*, which provides enough structure to leverage a useful tool called *duality theory*. Conic optimization problems are defined over *(convex) cones*, whose definition we state below.

Definition 1.1. Let $\mathcal{X}$ be a real topological vector space of arbitrary dimension. A subset $\mathcal{K} \subset \mathcal{X}$ is called a *(convex) cone* if

- for all $x, y \in \mathcal{K}$ and $\alpha, \beta \in \mathbb{R}_+$ one has

$$\alpha x + \beta y \in \mathcal{K},$$

that is, $\mathcal{K}$ is *closed under non-negative linear combinations*.

If $\mathcal{K}$ additionally satisfies

- $\mathcal{K} \cap (-\mathcal{K}) = \{0\}$, that is, $\mathcal{K}$ is *pointed*;
- $\mathcal{K} = \mathrm{cl}(\mathcal{K})$, that is, $\mathcal{K}$ is *closed*;
- $\mathrm{int}(\mathcal{K}) \neq \emptyset$, that is, $\mathcal{K}$ has a *non-empty interior*;

we refer to $\mathcal{K}$ as a *proper* convex cone.

1.1 Conic optimization

Let us now introduce the concept of conic programming and the associated duality theory in a setting that allows for infinite dimensional cones. For a comprehensive survey on the topic we refer to [Sha01] and [Bar02]. Let X be a vector space of arbitrary dimension over $\mathbb{R}$ and let X^* be its algebraic dual space, i.e., the set of all linear functionals on X. With the pair X, X^* we associate a bilinear functional

$$\begin{aligned} \langle \cdot, \cdot \rangle_X &: X \times X^* \to \mathbb{R} \\ &(x, x^*) \mapsto \langle x, x^* \rangle_X = x^*(x). \end{aligned} \tag{1.1}$$

When X is finite dimensional, X^* can be identified with X, since any finite dimensional space X is isomorphic to its algebraic dual space. In this case any non-degenerate bilinear form $\langle \cdot, \cdot \rangle_X$ will define an inner product on X. The infinite dimensional case is more subtle. If X is infinite dimensional its algebraic dual space is *too large* to develop a useful notion of duality. In this case we endow X with a topology and define the space X^* to be the subset of the algebraic dual space of X that defines *continuous* linear functionals on X with respect to the chosen topology. We want $(X, X^*, \langle \cdot, \cdot \rangle_X)$ to be a *dual system*. For this we need the bilinear form (1.1) to be non-degenerate, which we can achieve by requiring the chosen topology to be Hausdorff and locally convex. We thereby ensure that there are enough (continuous) linear functionals to separate points by the Hahn-Banach theorem [Meg98, cf. §1.9.5]. Similarly, a topology on X^* shall be chosen such that the set of continuous linear functionals on X^* is given by $\{\langle \cdot, x \rangle_X : x \in X\}$. For more information on these duality considerations we refer to [Bar02, Chapters II and III]. Let $\mathcal{K} \subseteq X$ be a closed convex cone, denote by $\mathcal{A} : X \to \mathbb{R}^m$ a (continuous)

linear operator for some $m \in \mathbb{N}$. Further, fix $c \in X^*$ and $b \in \mathbb{R}^m$. We define a *primal* conic optimization problem as

$$\begin{aligned} \text{val(P)} = \inf \quad & \langle c, x\rangle_X \\ \text{subject to:} \quad & \mathcal{A}(x) = b \\ & x \in \mathcal{K}. \end{aligned} \tag{P}$$

The set $\{x \in X : \mathcal{A}(x) = b\}$ is an affine subspace of X. Therefore, a conic optimization problem consists of minimizing a linear cost function over the intersection of an affine subspace with a convex cone. Let $\mathcal{A}^* : \mathbb{R}^m \to X^*$ be the adjoint mapping of $\mathcal{A}$ so that for $x \in X$ and $y \in \mathbb{R}^m$ we have $\langle y, \mathcal{A}(x)\rangle = \langle \mathcal{A}^*(y), x\rangle_X$ for some reference inner product $\langle \cdot, \cdot \rangle$ on $\mathbb{R}^m$.

1.1.1 Duality for conic programs

Duality theory is a useful tool in optimization. Every conic optimization problem has an associated *dual* problem. To define it we need to introduce the concept of the *dual cone* $\mathcal{K}^* \subseteq X^*$ of a convex cone $\mathcal{K} \subseteq X$, which is defined as

$$\mathcal{K}^* := \{x^* \in X^* : \langle x^*, x\rangle_X \geq 0 \, \forall x \in X\}.$$

Clearly, $\mathcal{K}^{**} := (\mathcal{K}^*)^* \supseteq \mathcal{K}$. If $\mathcal{K}$ is closed, then $\mathcal{K}^{**} = \mathcal{K}$, see [Bar02, §IV.5.3]. We call the following the *dual* problem of (P):

$$\begin{aligned} \text{val(D)} = \sup \quad & \langle y, b\rangle \\ \text{subject to:} \quad & c - \mathcal{A}^*(y) \in \mathcal{K}^* \\ & y \in \mathbb{R}^m. \end{aligned} \tag{D}$$

The problems (P) and (D) share an intricate relationship. Any feasible solution to the primal (resp. dual) problem provides an upper (resp. lower) bound for the dual (resp. primal). This relation is referred to as *weak duality*.

Theorem 1.2 (Weak duality)**.** *Let x, y be feasible solutions to problems* (P) *and* (D)*, respectively. Then,*

$$\langle c, x\rangle_X - \langle y, b\rangle = \langle c - \mathcal{A}^*(y), x\rangle_X \geq 0.$$

Proof. Inequality follows directly from that fact that $x \in \mathcal{K}$ and $c - \mathcal{A}^*(y) \in \mathcal{K}^*$. □

We call the difference $\text{val(P)} - \text{val(D)}$ the *duality gap*. A pair of feasible solutions (x, y) such that $\langle c - \mathcal{A}^*(y), x\rangle_X = 0$ provides a *certificate for optimality*.

Corollary 1.3. *The optimal values* val(P) *and* val(D) *coincide and both problems have an optimal solution if and only if there exists a feasible pair* (x, y) *such that the following complementarity condition is satisfied*

$$\langle c - \mathcal{A}^*(y), x \rangle_X = 0.$$

Often, we are interested in the cases in which *strong duality* holds, meaning val(D) = val(P). Results that guarantee this equality are sometimes referred to as *zero duality gap* theorems; we will refer to them as *strong duality results*. One such theorem is stated below.

Theorem 1.4 (cf. Theorem 7.2 [Bar02])**.** *Consider the primal dual pair* (P)*,* (D)*. Suppose the cone*

$$\hat{\mathcal{A}}(\mathcal{K}) = \{(\mathcal{A}(x), \langle c, x \rangle_X) : x \in \mathcal{K}\}$$

is closed in $X \oplus \mathbb{R}$ *and that there exists a primal feasible solution* x*, i.e.,* $x \in \mathcal{K}$ *and* $\mathcal{A}(x) = b$*. Then* val(P) = val(D)*. Moreover, if* val(P) $> -\infty$*, then there exists a primal optimal solution.*

There are also other sufficient conditions for strong duality. Conditions based on strict feasibility, i.e., the existence of a feasible point in the interior of the respective cones are known as *Slater-type conditions*.

Theorem 1.5 (Strong duality (cf. Theorem 2.8. [Sha01]))**.** *Suppose* val(P) $> -\infty$*, and there exists a* strictly *feasible solution* x *for the primal* (P)*, i.e.,* $x \in \text{int}(\mathcal{K})$ *and* $\mathcal{A}(x) = b$*, then*

- *the dual optimal value is attained*
- val(P) = val(D).

Similarly, if val(D) $< \infty$*, and there exists a* strictly *feasible solution* y *for the dual* (D)*, i.e.,* $y \in \mathbb{R}^m$ *such that* $c - \mathcal{A}^*(y) \in \text{int}(\mathcal{K}^*)$ *then*

- *the primal optimal value is attained*
- *and* val(P) = val(D).

1.1.2 Special cases of finite dimensional conic programming

We now present two special classes of finite dimensional conic optimization problem which will play an important part in this thesis.

Linear Programming (LP)

The simplest and also most well-known case of conic programming is *linear programming* (LP). To formulate LP in the terms presented above, we fix $m, n \in \mathbb{N}$ and set $X, X^* = \mathbb{R}^n$. The bilinear form $\langle \cdot, \cdot \rangle_X$ is given by the standard Euclidean inner product, i.e., for $x, y \in \mathbb{R}^n : \langle x, y \rangle = \sum_{i=1}^n x_i y_i = x^\top y$. The cone $\mathcal{K}$ is given by the non-negative orthant $\mathbb{R}^n_+$. This cone is proper and self-dual, i.e.,

$$\mathbb{R}^n_+ = \left(\mathbb{R}^n_+\right)^* = \left(\mathbb{R}^n_+\right)^{**}.$$

A linear operator $\mathcal{A} : \mathbb{R}^n \to \mathbb{R}^m$ in this setting can be interpreted as a matrix $A \in \mathbb{R}^{m \times n}$ and application of $\mathcal{A}$ to x is now standard matrix vector multiplication $\mathcal{A}(x) = Ax$.

A linear program in standard form can be formulated as follows. Let $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$ be given. The linear program reads

$$\begin{aligned} \inf \quad & \langle c, x \rangle = c^\top x \\ \text{subject to:} \quad & Ax = b \\ & x \in \mathbb{R}^n_+, \end{aligned} \tag{1.2}$$

with its corresponding dual given by

$$\begin{aligned} \sup \quad & \langle y, b \rangle = y^\top b \\ \text{subject to:} \quad & c - y^\top A \in \mathbb{R}^n_+ \\ & y \in \mathbb{R}^m. \end{aligned} \tag{1.3}$$

In practice linear programming problem can be solved by state-of-the-art solvers with millions of variables and constraints [LY15, §1.3]. Possible algorithms are the simplex algorithm [Dan51] or interior point methods [Kar84, NN94, Ren01].

Semidefinite programming (SDP)

Let $\mathbb{S}^n$ be the space of real symmetric matrices. An inner product on this space is given by the trace inner product, i.e., for $A, B \in \mathbb{S}^n$ we define

$$\langle A, B \rangle := \operatorname{Tr}(AB) = \sum_{i,j=1}^n A_{i,j} B_{i,j}.$$

A matrix $X \in \mathbb{S}^n$ is called positive semidefinite, denoted by $X \succeq 0$, if $x^\top X x \geq 0 \ \forall x \in \mathbb{R}^n$. The set of positive semidefinite matrices is a proper and self-dual cone.

Let $\mathcal{A} : \mathbb{S}^n \to \mathbb{R}^m$ be a linear operator defined via $X \mapsto (\langle A_1, X\rangle, \ldots, \langle A_m, X\rangle)$, where $A_i \in \mathbb{S}^n$ for $i \in [m]$. A semidefinite program in standard form is for $C \in \mathbb{S}^n$ and $b \in \mathbb{R}^m$, given by

$$\begin{aligned} \text{val(P)} = \inf \quad & \langle C, X\rangle \\ \text{subject to:} \quad & \mathcal{A}(X) = b \\ & X \in \mathbb{S}^n_+, \end{aligned} \tag{1.4}$$

with its dual defined as

$$\begin{aligned} \text{val(D)} = \sup \quad & \langle y, b\rangle = y^\top b \\ \text{subject to:} \quad & C - \sum_{i=1}^m y_i A_i \in \mathbb{S}^n_+ \\ & y \in \mathbb{R}^m. \end{aligned} \tag{1.5}$$

If all matrices appearing in the input data are diagonal, the program is equivalent to an LP. SDP can therefore be seen as a generalization of LP. For more background info, see [BV96, BV04].

Easy and difficult cones

Special cases of conic programming like LP and SDP can be solved up to a fixed approximation error ε in polynomial time in the Turing model [dKV16] (under some assumptions). What this means is that the number of elementary operations needed to obtain an ε-optimal solution can be bounded by a function that depends polynomially on the size of the encoding of the problem. An ε-optimal solution is a feasible point with objective value at most ε worse than the true optimum. The most common class of algorithms used to solve LPs and SDPs in practice are *interior point methods*, see [Kar84, NN94] for historical references, [Ren01] for a comprehensive introduction and [Gon12] for a survey. In recent years research has been aimed at the analysis of interior point methods applied to problems where some underlying structure present, see [ZL21, GS19, CN21, DSdSG$^+$22, JV22]. We will deal with interior point methods in greater detail in Chapter 2 and give some background in the Appendix A.

However, not all conic programs can be solved in polynomial time. In fact, conic programming can capture NP-hard problems. Consider a graph $G = (V, E)$. The stability number of G, denoted by $\alpha(G)$, asks for the size of the largest set $S \subseteq V$ such that for no two vertices in S share an edge. Computing the stability number of a graph G is notoriously NP-hard problem and can be formulated as

a conic optimization problem as was shown in [dKP02]. Consider the copositive cone $\mathcal{COP}_n$ defined as

$$\mathcal{COP}_n = \left\{A \in \mathbb{S}^n : x^T Ax \geq 0 \text{ for all } x \in \mathbb{R}^n_+\right\}.$$

Let A_G be the adjacency matrix of G. Then the stability number of G is the solution to the following problem:

$$\alpha(G) = \min\{\lambda : \lambda(I + A_G) - J \in \mathcal{CP}_n\},$$

where $n = |V|$. So even though this problem is a convex conic optimization problem, it is difficult and there are no known polynomial-time algorithms for it. This demonstrates that the complexity of non-convex problems can be transferred into a (convex) cone constraint.

1.1.3 Special cases of infinite dimensional conic programming

A special class of infinite dimensional conic optimization problems, called generalized moment problems (GMPs) will play an integral part of this thesis. Let $\mathbf{K} \subset \mathbb{R}^n$ be a compact set and denote by $\mathcal{M}(\mathbf{K})$ the (infinite dimensional) vector space of signed finite Borel measure whose support is contained in $\mathbf{K}$. We equip this space with the norm of total variation, defined as

$$\|\mu\|_{\text{TV}} = \int_{\mathbf{K}} \mathrm{d}\mu^+ + \int_{\mathbf{K}} \mathrm{d}\mu^-,$$

where $\mu = \mu^+ - \mu^-$ is the Jordan decomposition of μ (cf. [Bil86, Chapter 6 §32]). Since $\mathbf{K}$ is compact, the (algebraic) dual space of $\mathcal{M}(\mathbf{K})$ is the space of continuous functions on $\mathbf{K}$, denoted by $\mathcal{C}(\mathbf{K})$, which we endow with the supremum norm $\|\cdot\|_\infty$, defined for $f \in \mathcal{C}(\mathbf{K})$ as $\|f\|_\infty := \max_{\mathbf{x}\in\mathbf{K}} |f(\mathbf{x})|$. These two spaces form a dual system with the bilinear form $\langle\cdot,\cdot\rangle : \mathcal{C}(\mathbf{K}) \times \mathcal{M}(\mathbf{K}) \to \mathbb{R}$ defined by

$$\mathcal{C}(\mathbf{K}) \times \mathcal{M}(\mathbf{K}) \ni (f, \mu) \mapsto \langle f, \mu\rangle = \int_{\mathbf{K}} f(\mathbf{x})\mathrm{d}\mu(\mathbf{x}).$$

We are interested in optimizing over the convex cone of positive finite Borel measures supported on $\mathbf{K}$, i.e., Radon measures, which we denote by $\mathcal{M}(\mathbf{K})_+$. The dual cone of $\mathcal{M}(\mathbf{K})_+$ with respect to the bilinear form defined above is given by the set

$$\mathcal{M}(\mathbf{K})_+^* = \left\{f \in \mathcal{C}(\mathbf{K}) : \int_{\mathbf{K}} f(\mathbf{x})\mathrm{d}\mu(\mathbf{x}) \geq 0 \ \forall \mu \in \mathcal{M}(\mathbf{K})_+\right\},$$

which is the set of continuous functions that are non-negative on $\mathbf{K}$. To see this, note that since the Dirac-$\delta_{\mathbf{x}}$ lies in $\mathcal{M}(\mathbf{K})_+$ for every $\mathbf{x} \in \mathbf{K}$ it follows that all $f \in \mathcal{M}(\mathbf{K})_+^*$ satisfy $f(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in \mathbf{K}$. On the other hand, if f is a non-negative continuous function on $\mathbf{K}$, we find $f^* = \min_{\mathbf{x}\in\mathbf{K}} f(\mathbf{x}) \geq 0$ and hence $\int_{\mathbf{K}} f(\mathbf{x})\mathrm{d}\mu(\mathbf{x}) \geq f^* \int_{\mathbf{K}} \mathrm{d}\mu(\mathbf{x}) \geq 0$ for all $\mu \in \mathcal{M}(\mathbf{K})_+$. Hence, we denote by $\mathcal{C}(\mathbf{K})_+ = \{f \in \mathcal{C}(\mathbf{K}) : f(\mathbf{x}) \geq 0 \ \forall \mathbf{x} \in \mathbf{K}\}$ the dual cone of $\mathcal{M}(\mathbf{K})_+$. We denote a GMP as follows for input data $f_0, f_1, \ldots, f_m \in \mathcal{C}(\mathbf{K})$ and $b \in \mathbb{R}^m$

$$\begin{aligned} \inf \quad & \langle f, \mu \rangle \\ \text{subject to:} \quad & \langle f_i, \mu \rangle = b_i, \ i \in [m] \\ & \mu \in \mathcal{M}(\mathbf{K})_+. \end{aligned} \tag{1.6}$$

This is a conic program in terms of Section 1.1 when setting $X = \mathcal{M}(\mathbf{K}), X^* = \mathcal{C}(\mathbf{K})$ and $\mathcal{K} = \mathcal{M}(\mathbf{K})_+$ as well as defining $\mathcal{A}$ via $\mathcal{A}(\mu) := (\langle f_1, \mu \rangle, \ldots, \langle f_m, \mu \rangle)$. The dual is given by

$$\begin{aligned} \sup \quad & y^\top b \\ \text{subject to:} \quad & f_0(\mathbf{x}) - \sum_{i=1}^m y_i f_i(\mathbf{x}) \in \mathcal{C}(\mathbf{K})_+ \\ & y \in \mathbb{R}^m. \end{aligned} \tag{1.7}$$

Checking non-negativity of a function f on a set $\mathbf{K}$ is as difficult as minimizing f over $\mathbf{K}$:

$$f(\mathbf{x}) \in \mathcal{C}(\mathbf{K})_+ \Leftrightarrow \min_{\mathbf{x}\in\mathbf{K}} f(\mathbf{x}) \geq 0.$$

Many NP-hard problems can be modeled as a minimization problem over a basic semi-algebraic set $\mathbf{K}$, see, e.g., [MK87]. Hence, checking membership to $\mathcal{C}(\mathbf{K})_+$ is difficult, in the sense that no polynomial-time membership oracle for this cone exists, unless P = NP. Therefore, solving the GMP in general is numerically intractable. Nevertheless, it is a powerful modeling tool, allowing us to find equivalent convex formulations of non-convex problems. This convexification comes at the price of increasing the dimension of the convex problem to infinity. However, we will introduce finite dimensional approximation hierarchies for these kinds of problems in Section 1.2. But before, we present a few applications of the GMP.

Polynomial and rational optimization

Consider the problem of finding the minimum of a rational polynomial over a compact set $\mathbf{K} \in \mathbb{R}^n$

$$p^* = \min_{\mathbf{x}\in\mathbf{K}} \frac{p(\mathbf{x})}{q(\mathbf{x})}, \tag{1.8}$$

where $p, q \in \mathbb{R}[\mathbf{x}]$ are relatively prime and $q(\mathbf{x}) > 0$ over $\mathbf{K}$. In fact, if q changes signs over $\mathbf{K}$, one can show $p^* = -\infty$, see [JdK06]. Let us now show that (1.8) is equivalent to the following GMP

$$\text{val} = \inf_{\mu \in \mathcal{M}(\mathbf{K})_+} \left\{ \int_{\mathbf{K}} p(\mathbf{x}) \mathrm{d}\mu(\mathbf{x}) \, : \, \int_{\mathbf{K}} q(\mathbf{x}) \mathrm{d}\mu(\mathbf{x}) = 1 \right\}. \tag{1.9}$$

Let $\mathbf{x}^*$ be such that $\frac{p(\mathbf{x}^*)}{q(\mathbf{x}^*)} = p^*$ and define $\mu^* = \frac{1}{q(\mathbf{x}^*)} \delta_{\mathbf{x}^*}$. Then μ^* is feasible for (1.9) and leads to the objective value

$$\int_{\mathbf{K}} p(\mathbf{x}) \mathrm{d}\mu^*(\mathbf{x}) = \frac{p(\mathbf{x}^*)}{q(\mathbf{x}^*)} = p^*.$$

Thus, $\text{val} \leq p^*$. To prove $\text{val} \geq p^*$ note that

$$\frac{p(\mathbf{x})}{q(\mathbf{x})} \geq p^* \; \Rightarrow p(\mathbf{x}) \geq p^* q(\mathbf{x}) \; \forall \mathbf{x} \in \mathbf{K}.$$

Therefore, for every positive finite Borel measure $\mu \in \mathcal{M}(\mathbf{K})_+$ feasible for problem (1.9) we find

$$\int_{\mathbf{K}} p(\mathbf{x}) \mathrm{d}\mu(\mathbf{x}) \geq p^* \int_{\mathbf{K}} q(\mathbf{x}) \mathrm{d}\mu(\mathbf{x}) = p^*.$$

Note that if one sets $q \equiv 1$ on $\mathbf{K}$, problem (1.8) reduces to a polynomial optimization problem, i.e., finding the minimum of an n-variate polynomial over a compact set.

We will discuss in greater detail in Chapter 3 the generalized moment problem defined for $\mathbf{K}$ being the standard probability simplex $\Delta_{n-1} = \{\mathbf{x} \in \mathbb{R}^n : x_1 + \cdots + x_n = 1\}$ and the $n-1$ dimensional unit sphere $\mathcal{S}^{n-1} = \{\mathbf{x} \in \mathbb{R}^n : x_1^2 + \ldots + x_n^2 = 1\}$. We take a moment to discuss two important polynomial optimization problems defined over these sets. For one, the problem of computing the stability number $\alpha(G)$ of a graph $G = (V, E)$ can be formulated as a quadratic polynomial optimization problem over the simplex. Indeed, for a graph G with adjacency matrix A_G, Motzkin and Strauss [MS65] showed that

$$\frac{1}{\alpha(G)} = \min_{\mathbf{x} \in \Delta_{n-1}} \mathbf{x}^\top (A_G + I) \mathbf{x},$$

where I is the identity matrix, which is a quadratic polynomial optimization problem over the simplex.

Similarly, deciding convexity of a homogeneous polynomial f of degree 4 or higher is known to be NP-hard [AOPT13]. It can be modeled as polynomial optimization problem over the sphere. A homogeneous polynomial f is convex if and only if

$$\min_{(\mathbf{x},\mathbf{y})\in\mathcal{S}^{2n-1}} \mathbf{y}^\top \nabla f(\mathbf{x})\mathbf{y} \geq 0,$$

which in turn be cast as a GMP over the sphere. For polynomials of degree less than 4 the situation is much simpler. Linear functions are trivially convex. Homogeneous quadratic polynomials of the form $\mathbf{x}^\top Q\mathbf{x}$ for $Q \in \mathbb{R}^{n\times n}$ are convex if and only if $Q \succeq 0$ and polynomials of degree 3 are never convex. Both these examples illustrate why studying the GMP over simple sets like the simplex and the sphere is of interest.

1.2 The moment-SOS hierarchy

Since solving the GMP in full generality is out of reach, so-called *approximation hierarchies* are often used in practice, which provide approximate solutions. Continuing we will consider the following GMP

$$\begin{aligned} \text{val}_{\text{D}}^* = \sup \quad & y^\top b \\ \text{subject to:} \quad & f_0(\mathbf{x}) - \sum_{i=1}^m y_i f_i(\mathbf{x}) \in \mathcal{C}(\mathbf{K})_+ \\ & y \in \mathbb{R}^m, \end{aligned} \tag{1.10}$$

with polynomial data functions $f_i \in \mathbb{R}[\mathbf{x}]$, $i = 0, 1, \ldots, m$ and the set $\mathbf{K}$ will be assumed to be a compact semi-algebraic set defined by polynomials $g_1, \ldots, g_\ell \in \mathbb{R}[\mathbf{x}]$, i.e.,

$$\mathbf{K} = \{\mathbf{x} \in \mathbb{R}^n : g_1(\mathbf{x}) \geq 0, \ldots, g_\ell(\mathbf{x}) \geq 0\}. \tag{1.11}$$

A common tool to tackle intractable optimization problems is to relax (some of) the constraints which make the problem difficult. Solving such a relaxation is often much easier and leads to an *approximation* of the sought optimal value. One way of achieving this in the conic optimization setting is to replace the cone in question by a more tractable cone. The underlying idea of the approximation hierarchies that we will consider is to construct a set of polynomials non-negative on $\mathbf{K}$ over which we can optimize efficiently. The objects making this possible are

the so-called sums-of-squares (SOS) polynomials, which we denote by

$$\Sigma[\mathbf{x}] := \left\{ \sum_{i=1}^{k} q_i(\mathbf{x})^2 : k \in \mathbb{N},\, q_i \in \mathbb{R}[\mathbf{x}] \right\}.$$

The set $\Sigma[\mathbf{x}]$ is a proper convex cone. Trivially, every $p \in \Sigma[\mathbf{x}]$ is globally non-negative. The converse, however is not true: Not all polynomials that are globally non-negative can be expressed as a sum-of-squares as Hilbert proved in 1888, see [Rez96] for a survey on this matter.

Theorem 1.6. *Hilbert (1888) Every non-negative homogeneous polynomial is a sum-of-squares if and only if*

- $n = 2$ *(univariate non-homogeneous case);*
- $2r = 2$ *(quadratic forms);*
- $n = 3$ *and* $2r = 4$ *(ternary quartics);*

where n is the number of variables and $2r$ is the degree.

Even though SOS polynomials do not capture *all* non-negative polynomials, they provide a powerful tool because of their relation to the cone of positive semidefinite matrices.

Lemma 1.7. *Every SOS polynomial $p \in \Sigma[\boldsymbol{x}]$ of degree $2r$ can be written as*

$$p(\boldsymbol{x}) = [\boldsymbol{x}]_r^\top A [\boldsymbol{x}]_r$$

for a positive semidefinite matrix $A \succeq 0$, where

$$[\boldsymbol{x}]_r = [1, x_1, \ldots, x_n, x_1^2, x_1 x_2, \ldots, x_n^r] \tag{1.12}$$

is the monomial basis vector.

This lemma implies that the question whether a polynomial can be expressed as a sum-of-squares is equivalent to a semidefinite program. We call the following convex cone the *quadratic module* generated by $g_1, \ldots, g_\ell$:

$$\mathcal{Q}(\mathbf{g}) = \mathcal{Q}(g_1, \ldots, g_\ell) := \left\{ \sigma_0 + \sum_{i=1}^{\ell} \sigma_i(\mathbf{x}) g_i(\mathbf{x}) : \sigma_i \in \Sigma[\mathbf{x}],\, i = 0, 1, \ldots, \ell \right\}. \tag{1.13}$$

The crucial observation is that any $y \in \mathbb{R}^m$ for which

$$f_0(\mathbf{x}) - \sum_{i=1}^{m} y_i f_i(\mathbf{x}) \in \mathcal{Q}(\mathbf{g})$$

is a feasible solution to (1.10) and thereby provides a lower bound on the optimal value. In order to formulate a finite dimensional semidefinite program over this cone we need to consider a *truncated* version of it, restricting the degrees that the involved SOS polynomials may have. We refer to this as the truncated quadratic module

$$\mathcal{Q}_r(\mathbf{g}) := \left\{ \sigma_0 + \sum_{i=1}^{\ell} \sigma_i(\mathbf{x}) g_i(\mathbf{x}) : \sigma_i \in \Sigma_r[\mathbf{x}],\, i = 0, 1, \ldots, \ell \right\}, \tag{1.14}$$

where $\Sigma_r[\mathbf{x}]$ is the set of SOS polynomials of degree $\leq 2r$. Clearly, $\mathcal{Q}_r(\mathbf{g}) \subseteq \mathcal{Q}_{r+1}(\mathbf{g}) \subseteq \mathcal{C}(\mathbf{K})_+$, leading to a hierarchy of optimization problems. We call the following instance the r-th level of the *SOS hierarchy*

$$\begin{aligned} \mathrm{val}_{\mathrm{D}}^{(r)} = \sup \quad & y^\top b \\ \text{subject to:} \quad & f_0(\mathbf{x}) - \sum_{i=1}^{m} y_i f_i(\mathbf{x}) \in \mathcal{Q}_r(\mathbf{x}) \\ & y \in \mathbb{R}^m, \end{aligned} \tag{1.15}$$

which, for every $r \in \mathbb{N}$ is equivalent to a semidefinite program of size polynomial in n. The hierarchy provides a non-decreasing sequence $\{\mathrm{val}_{\mathrm{D}}^{(r)}\}_{r\in\mathbb{N}}$ upper bounded by $\mathrm{val}_{\mathrm{D}}^*$. This hierarchy was first proposed by Lasserre [Las01] and Parrilo [Par00] in the early 2000s with focus on polynomial optimization problems. We also refer to [Las08] for a treatment of the GMP. A considerable amount of research has been aimed at the analysis of these hierarchies. Using tools from algebraic geometry like *Putinar's Positivstellensatz* [Put93] it is possible under some mild assumption on $\mathbf{K}$ to prove that the hierarchy converges to the optimal value of (1.10), i.e.,

$$\lim_{r\to\infty} \mathrm{val}_{\mathrm{D}}^{(r)} = \mathrm{val}_{\mathrm{D}}^*.$$

Formulating problem (1.15) as an SDP and dualizing it gives little insight into the objects, i.e., measures, that we are approximating in the primal. We will approach the dual of (1.15) from another angle. Let $\mathbf{y} = \{y_\alpha\}_{\alpha\in\mathbb{N}^n}$ be an infinite real sequence and let $L_{\mathbf{y}} : \mathbb{R}[\mathbf{x}] \to \mathbb{R}$ be the Riesz linear functional defined by

$$f(\mathbf{x}) = \sum_{\alpha\in\mathbb{N}^n} f_\alpha \mathbf{x}^\alpha \mapsto L_{\mathbf{y}}(f) = \sum_{\alpha\in\mathbb{N}^n} f_\alpha y_\alpha,$$

where we define by abuse of notation $\mathbf{x}^\alpha := x_1^{\alpha_1} \cdot \ldots \cdot x_n^{\alpha_n}$. Therefore, if $\mathbf{y}$ is the moment sequence of a measure μ supported on a set $\mathbf{K}$, i.e.,

$$y_\alpha = \int_{\mathbf{K}} \mathbf{x}^\alpha \mathrm{d}\mu(\mathbf{x}) \text{ for } \alpha \in \mathbb{N}^n$$

then $L_\mathbf{y}$ coincides with the integration operator on polynomials with respect to μ, i.e., for a polynomial $f \in \mathbb{R}[\mathbf{x}]$ we find

$$L_\mathbf{y}(f) = \sum_{\alpha \in \mathbb{N}^n} f_\alpha y_\alpha = \int_{\mathbf{K}} f(\mathbf{x}) \mathrm{d}\mu(\mathbf{x}).$$

Given a finite sequence $\mathbf{y} = \{y_\alpha\}_{\alpha \in \mathbb{N}^n_{2r}}$, we associate the so-called *truncated moment matrix* $M_r(\mathbf{y})$ to $\mathbf{y}$, defined as $(M_r(\mathbf{y}))_{\alpha,\beta} = y_{\alpha+\beta}$ for $\alpha, \beta \in \mathbb{N}^n_r$. Such a matrix has dimensions $s(n,r) \times s(n,r)$, where

$$s(n,r) = \binom{n+r}{r}.$$

For $\mathbf{y}$ as above, given a polynomial $g \in \mathbb{R}[\mathbf{x}]$ of degree d, we define the *localizing matrix* $M_r(g \star \mathbf{y})$ associated to $\mathbf{y}$ and g as

$$(M_r(g \star \mathbf{y}))_{\alpha,\beta} = \sum_{\gamma \in \mathbb{N}^n_d} g_\gamma y_{\alpha+\beta+\gamma}, \text{ for } \alpha, \beta \in \mathbb{N}^n_r.$$

Define

$$r_{\min} := \max_{i \in [m], j \in [\ell]} \{\deg(f_0), \deg(f_i), \deg(g_j), \}.$$

For $r \in \mathbb{N}$, with $r \geq r_{\min}$, the level r of moment-SOS relaxation of is defined as

$$\begin{aligned} \inf_{\mathbf{y} \in \mathbb{R}^{s(n,2r)}} \; & L_\mathbf{y}(f_0) \\ \text{s.t. } & L_\mathbf{y}(f_i) = b_i, \text{ for } i \in [m] \\ & M_r(\mathbf{y}) \succeq 0 \\ & M_r(g_i \star \mathbf{y}) \succeq 0, \text{ for } i \in [\ell]. \end{aligned} \tag{1.16}$$

For each r this is a semidefinite optimization problem (SDP) whose size depends polynomially on n. The moment and localizing matrices depend linearly on $\mathbf{y}$ and the objective is linear in $\mathbf{y}$.

It is not difficult to see that the semidefinite programming formulation of (1.16) is the dual of (1.15). This primal-dual pair sequence is referred to as the *moment-SOS hierarchy*[1] in the literature and will play an integral part of this thesis.

[1] It used to be called the *Lasserre* hierarchy, yet in recent years the name moment-SOS hierarchy is used more frequently.

1.3 Thesis overview

We continue to give a brief overview of the contents of this thesis, chapter by chapter.

Chapter 2: A predictor corrector method for SDP using the factor width cone

The moment-SOS hierarchy provides a sequence of finite dimensional semdefinite programming approximations for the GMP. Even though the size of these SDPs depends polynomially on n, it grows rather quickly. Hence, often only the first few levels are actually computable. To address this problem we propose an interior point algorithm in Chapter 2 which is more suitable for parallelization that the ordinary interior point methods used to solve SDPs. The underlying idea of the presented algorithm is to replace the cone of positive semidefinite matrices by a more tractable cone, namely the cone of matrices of constant factor width which is defined as $\mathrm{FW}_n(k) = \left\{Y \in \mathbb{S}^n : Y = \sum_{i\in\mathbb{N}} x_i x_i^\top \text{ for } x_i \in \mathbb{R}^n, \mathrm{supp}(x_i) \leq k\,, \forall i\right\}$, see [BCPT05]. More precisely, instead of performing the subroutines of the algorithm in the cone $\mathbb{S}^n_+$ they are performed in the cone product $\mathbb{S}^k_+ \times \cdots \times \mathbb{S}^k_+$ consisting of $\binom{n}{k}$ cones for some $n \geq k \in \mathbb{N}$. Simply replacing the cone, however, does of course only lead to a relaxation of the problem. The algorithm we develop utilizes interim solutions to rescale the problem, an idea first proposed by [AH17]. This iterative rescaling leads to a convergent algorithm. We are thereby extending results from [RSS22].

Chapter 3: Convergence rates for RLT and Lasserre-type hierarchies

In Chapter 3 we consider the GMP over the standard probability simplex Δ_{n-1} and the $n-1$ dimensional unit sphere $\mathcal{S}^{n-1}$. We propose a relaxation hierarchy based on linear programming for the simplex case and conduct a convergence analysis proving a rate of convergence of $O(1/r)$, where r is the level of the hierarchy. The foundation of the analysis is a quantitative version of Pólya's Positivstellensatz due to Powers and Reznick [Rez95]. Following up, we apply the moment-SOS hierarchy due to Lasserre [Las08] as introduced in Section 1.2 to the sphere case and also analyze the convergence behavior of the hierarchy proving a rate of $O(1/r^2)$, where r is the level of the hierarchy. The main result we rely on in the latter analysis is a quantitative Positivstellensatz by Fang and Fawzi [FF21].

Chapter 4: Computing bounds for option prices

A European call option is a financial contract that gives the holder the right, but not the obligation, to buy an underlying asset (such as a stock, commodity, or currency) at a specified price on a specified date. Under the assumption of the absence of *arbitrage* opportunities the problem of finding upper and lower bounds on the value of such financial products can be formulated as a GMP over the non-negative orthant involving piecewise linear data functions. In Chapter 4, we revisit a method previously employed by Bertsimas and Popescu [BP02], which utilized semidefinite programming techniques to tackle this problem. We prove that under the assumption of finite d-th order moments that the underlying set can be compactified. We then apply the moment-SOS hierarchy due to Lasserre [Las08] in this setting and provide several numerical examples.

Chapter 5: Construction of approximation kernels via SDP

Chapter 5 is dedicated to the problem of uniform approximation of non-differentiable functions $f : [-1,1]^n \to \mathbb{R}$. It is well-known that a convergent sequence of approximating polynomials can be obtained from a hierarchy of so-called *polynomial approximation kernels* [WWAF06]. For a given level r of the hierarchy such a kernel can be obtained by solving a semidefinite program. The convolution of an optimal kernel with a non-differentiable function f results in a degree r polynomial which serves as an approximation of f. The kernels we look for ought to minimize oscillations such approximations often exhibit, and are known as kernels of the Jackson-type, due to their similarity to the approximation kernels first studied by Dunham Jackson [Jac11, Jac12]. We construct a hierarchy of semidefinite programs and prove the respective solutions constitute a sequence of approximation kernels which lead to a uniform approximation of f. The size of the involved SDPs grows quickly, and we use symmetry reduction techniques as reviewed in Appendix B to limit this growth.

Societal and scientific relevance of the thesis topics

Large-scale SDP is important in engineering design projects like optimal power flow in large electrical networks [ZJJ+20], in robotics [DT14], and the wing-design of aircraft like the Airbus A380 [SKL09]. The current state-of-the-art SDP software is often unable to solve these type of large-scale instances, and therefore

there is a need for novel algorithmic approaches as the one described in Chapter 2.

The GMP finds many applications including options pricing in finance. In Chapter 3 and Chapter 4 we consider LP and SDP-based approaches to solve such problems. The novel aspect of the approach in Chapter 4 is that it is data driven: optimal pricing without assuming some underlying model. This is desirable since the underlying assumptions in pricing models are often unrealistic or difficult to verify.

The use of positive kernels to approximate functions is ubiquitous in physics, as reviewed in the survey [WWAF06]. Our approach presented in Chapter 5 should enhance such applications in the case of several variables.

1.4 Contributions to the Literature

This thesis is based on the following articles:

[KdK22] Felix Kirschner and Etienne de Klerk. Convergence rates of RLT and Lasserre-type hierarchies for the generalized moment problem over the simplex and the sphere. *Optimization Letters*, 16:2191–2208, 2022

[HKdK$^+$23] Didier Henrion, Felix Kirschner, Etienne de Klerk, Milan Korda, Jean B. Lasserre, and Victor Magron. Revisiting semidefinite programming approaches to options pricing: complexity and computational perspectives. *Informs Journal on Computing*, 2023. Advance online publication

[KdKng] Felix Kirschner and Etienne de Klerk. Construction of multivariate approximation kernels via semidefinite programming. *SIAM Journal on Optimization*, forthcoming

[KdK23] Felix Kirschner and Etienne de Klerk. A predictor-corrector algorithm for semidefinite programming that uses the factor width cone. arXiv preprint, arXiv:2301.06368 [math.OC], 2023.

These articles are used in the chapters of this thesis as follows:

Chapter 2	Based on [KdK23]
Chapter 3	Based on [KdK22]
Chapter 4	Based on [HKdK$^+$23]
Chapter 5	Based on [KdKng].

2

A predictor-corrector method for semidefinite programming using the factor width cone

2.1 Introduction

Semidefinite programming problems (SDPs) are a generalization of linear programming problems (LPs). While capturing a much larger set of problems, SDPs are still being solvable up to fixed precision in polynomial time in terms of the input data [NN94]; see [dKV16] for the complexity in the Turing model of computation. Solving large SDPs in practice is, however, more complicated. While we are able to solve linear programs with millions of variables and constraints routinely ([LY15, §1.3]), SDPs become intractable already for a few tens of thousands of constraints and for $n \times n$ matrix variables of the order $n \approx 1,000$. The reason is that each iteration of a typical interior point algorithm for SDP requires $\mathcal{O}(n^3m + n^2m^2 + m^3)$ floating point operations, where n is the size of the matrix variable and m is the number of equality constraints; see e.g., [AHO98]. However, solving large instances of SDPs is of growing interest, due to applications in power flow problems on large power grids, SDP-based hierarchies for polynomial

and combinatorial problems, etc. (see [Las01, ZL21, ZJJ+20]). In the following we will revisit a relaxation of a given SDP, where the cone of positive semidefinite matrices is replaced by a more tractable cone, namely the cone of matrices of constant factor width [BCPT05]. The simplest examples of matrices of constant factor width are non-negative diagonal matrices (corresponding to linear programs), and scaled diagonally dominant matrices (corresponding to second order cone programming) [AM14]. We then review how iteratively rotating the cone and solving the given optimization problem over this new set leads to a non-increasing sequence of values lower bounded by the optimum of the sought SDP. This iterative procedure, due to [ADH17], does not lead to a convergent algorithm. However, its essence can be used to construct a convergent predictor-corrector interior point method, as was done in [RSS22]. This chapter is inspired by ideas from [AM14, AH17, ADH17, AM19, RSS22]. In particular, we will extend the results in [RSS22], and give a more concise complexity analysis in our extended setting. We refer those readers, who are not familiar with interior point methods and the predictor-corrector method in particular to Appendix A for a brief introduction and background results.

2.1.1 Iterative approximation scheme

Let a set $\{A_i \in \mathbb{S}^n : i \in [m]\}$ of symmetric data matrices be given and define the linear operator

$$\mathcal{A}(X) = (\langle A_1, X\rangle, \ldots, \langle A_m, X\rangle)^\top \in \mathbb{R}^m.$$

Further, define for $b \in \mathbb{R}^m$ the affine subspace

$$L = \{X \in \mathbb{S}^n : \mathcal{A}(X) = b\}. \tag{2.1}$$

Consider the following semidefinite program

$$v^*_{\text{SDP}} = \inf\left\{\langle A_0, X\rangle : \mathcal{A}(X) = b, X \in \mathbb{S}^n_+\right\}, \tag{2.2}$$

which we assume to be strictly feasible. Replacing the cone of positive semidefinite (psd) matrices in (2.2) by a cone $\mathcal{K} \subseteq \mathbb{S}^n_+$, which is more tractable, leads to the following program

$$v_{\mathcal{K}} = \inf\left\{\langle A_0, X\rangle : \mathcal{A}(X) = b, X \in \mathcal{K}, \text{ where } \mathcal{K} \subseteq \mathbb{S}^n_+\right\}. \tag{2.3}$$

Clearly, $v_{\mathcal{K}} \geq v^*_{\text{SDP}}$. The quality of the approximation depends on the chosen cone $\mathcal{K}$. In [AM14], while focusing on sums-of-squares optimization the authors consider the cones of *diagonally dominant* and *scaled diagonally dominant* matrices.

Ahmadi and Hall developed the idea of replacing the psd cone by a simpler cone further in [ADH17], leveraging an optimal solution of the relaxation. Essentially, the idea is as follows. Define the feasible set for (2.2) as

$$\mathcal{F}_{\text{SDP}} = \{X \succeq 0 : \mathcal{A}(X) = b\}.$$

We will consider a sequence of strictly feasible points for (2.3), denoted by X_ℓ for $\ell = 0, 1 \ldots$. Since $X_\ell \succeq 0$, the matrix $X_\ell^{1/2}$ is well-defined. One can *update* the data matrices in the following way

$$A_i^{(\ell)} = X_\ell^{1/2} A_i X_\ell^{1/2} \quad (i \in \{0, 1, \ldots, m\},\, \ell = 0, 1, \ldots),$$

giving rise to a new linear operator

$$\mathcal{A}^{(\ell)}(X) = (\langle A_1^{(\ell)}, X\rangle, \ldots, \langle A_m^{(\ell)}, X\rangle)^\top \in \mathbb{R}^m.$$

We may also refer to this operation as *rescaling* with respect to X_ℓ. Via this rescaling one obtains the following sequence of reformulations of (2.2)

$$v^*_{\text{SDP}} = \inf\left\{\langle A_0^{(\ell)}, X\rangle : \mathcal{A}^{(\ell)}(X) = b, X \in \mathbb{S}^n_+\right\}, \tag{2.4}$$

whose feasible set we define as

$$\mathcal{F}_{\text{SDP}_\ell} = \left\{X \succeq 0 : \mathcal{A}^{(\ell)}(X) = b\right\}.$$

For each ℓ the identity matrix is feasible, i.e., we have $X = I \in \mathcal{F}_{\text{SDP}_\ell}$. To see this, note that for all $i \in [m]$ we have

$$\langle A_i^{(\ell)}, I\rangle = \langle (X_\ell)^{\frac{1}{2}} A_i (X_\ell)^{\frac{1}{2}}, I\rangle = \langle A_i, X_\ell\rangle = b_i.$$

Similarly, the identity leads to the same objective value in (2.4) as X_ℓ in (2.3). Let X_0 be an optimal solution to (2.3). Rescaling with respect to X_0 we find by the same reasoning that $v_{\mathcal{K}}^{(0)} \leq v_{\mathcal{K}}$, where

$$v_{\mathcal{K}}^{(\ell)} = \min\left\{\langle A_0^{(\ell)}, X\rangle : \mathcal{A}^{(\ell)}(X) = b, X \in \mathcal{K}\right\}. \tag{2.5}$$

Reiterating this procedure leads to a non-increasing sequence of values $\left\{v_{\mathcal{K}}^{(\ell)}\right\}_{\ell\in\mathbb{N}}$ lower bounded by v^*_{SDP}. Unfortunately, this procedure does not converge to the true optimum of (2.2) in general, as mentioned in [RSS22]. Indeed, it can happen that $\liminf_{\ell\to\infty} v_{\mathcal{K}}^{(\ell)} > v^*_{\text{SDP}}$. The rest of this chapter is devoted to the development and analysis of an algorithm, which converges to the optimal value v^*_{SDP}. We thereby generalize results from [RSS22].

2.1.2 The factor width cone

Fix $n \in \mathbb{N}$. The cone of $n \times n$ matrices of *factor width* k, denoted by $\mathrm{FW}_n(k)$, is defined as

$$\mathrm{FW}_n(k) = \left\{ Y \in \mathbb{S}^n : Y = \sum_{i \in \mathbb{N}} x_i x_i^\top \text{ for } x_i \in \mathbb{R}^n, \operatorname{supp}(x_i) \leq k \,, \forall i \right\}.$$

The notion of factor width was first used in [BCPT05], where the authors proved that $\mathrm{FW}_n(2)$ is the cone of scaled diagonally dominant matrices. Trivially, $\mathrm{FW}_n(1)$ is the cone of non-negative $n \times n$ diagonal matrices. Clearly, we have that

$$\mathrm{FW}_n(k) \subseteq \mathrm{FW}_n(k+1) \subseteq \mathbb{S}^n_+ \quad \forall k \in [n-1].$$

Moreover, $\mathrm{FW}_n(n) = \mathbb{S}^n_+$. It is easy to see these cones are proper for $k \geq 2$. As they define an inner approximation of the cone $\mathbb{S}^n_+$ we may use them in the aforementioned iterative scheme. Define

$$\mathbb{S}^{(n,k)} := \underbrace{\mathbb{S}^k \times \cdots \times \mathbb{S}^k}_{\binom{n}{k}\text{-times}} \text{ and } \mathbb{S}^{(n,k)}_+ := \underbrace{\mathbb{S}^k_+ \times \cdots \times \mathbb{S}^k_+}_{\binom{n}{k}\text{-times}}.$$

An optimization problem over the cone $\mathrm{FW}_n(k)$ may be formulated as an optimization problem over the cone product $\mathbb{S}^{(n,k)}_+$. To see this we need to consider principal submatrices. For a matrix $S \in \mathbb{R}^{n \times n}$ we define the principal submatrix $S_{J,J}$ for $J \subseteq [n]$ to be the restriction of S to rows and columns whose indices appear in J. Further, for a set $J = \{i_1, \ldots, i_{|J|}\} \subseteq [n]$ and a matrix $S \in \mathbb{R}^{|J| \times |J|}$ we define the $n \times n$ matrix $S_J^{\to n}$ as follows for $i, j \in [n]$

$$\left(S_J^{\to n}\right)_{i,j} := \begin{cases} S_{k,l} & \text{if } i = i_k, j = i_l \\ 0 & \text{otherwise.} \end{cases} \tag{2.6}$$

In other words, $S_J^{\to n}$ has S_J as principal sub-matrix indexed by J, and zeros elsewhere. Now, to write a program over $\mathrm{FW}_n(k)$ as an SDP note the following lemma.

Lemma 2.1. *For any $X \in \mathrm{FW}_n(k)$ we have that*

$$X = \sum_{|J|=k} Y_J^{\to n}$$

for suitable $Y_J \in \mathbb{S}^k_+$ and $J \subseteq [n], |J| = k$.

Proof. The proof is straight-forward, and omitted for the sake of brevity. □

Thus, we can write

$$\inf\{\langle A_0, X\rangle : \mathcal{A}(X) = b, X \in \mathrm{FW}_n(k)\} \tag{2.7}$$

as

$$\inf\left\{\sum_{|J|=k} \langle (A_0)_{J,J}, Y_J\rangle : \sum_{|J|=k} \langle (A_i)_{J,J}, Y_J\rangle = b_i, Y_J \in \mathbb{S}^k_+, \ \forall |J| = k\right\}. \tag{2.8}$$

It is straightforward to show that the dual cone of $\mathrm{FW}_n(k)$ is given by

$$\mathrm{FW}_n(k)^* = \{S \in \mathbb{S}^n : S_{J,J} \succeq 0 \text{ for } J \subseteq [n], |J| = k\}.$$

The dual cone has been studied in the context of semidefinite optimization in [BDMS22], where it was shown that the distance of $\mathrm{FW}_n(k)^*$ and $\mathbb{S}^n_+$ in the Frobenius norm can be upper bounded by $\frac{n-k}{n+k-2}$ for matrices of trace 1. For $k \geq 3n/4$ and $n \geq 97$ this bound can be improved to $O(n^{-3/2})$ (see [BDMS22]).

2.1.3 A predictor-corrector method

In this subsection we propose our algorithm which makes use of the rescaling introduced in Section 2.1.1. Our aim is to provide a comprehensible exposition, while the details are postponed to the second part of the chapter, beginning with Section 2.2.

Algorithm 2.1 is an adaption of the predictor-corrector method as described in [Ren01, § 2.2.4], see also Appendix A. Before describing the algorithm in detail, we fix some notation. Let

$$\mathcal{Y} = \left\{Y_J \in \mathbb{S}^k : J \subset [n], |J| = k\right\}$$

be a collection of $\binom{n}{k}$ matrices of size $k \times k$. We define the operator Ψ as

$$\Psi(\mathcal{Y}) = \sum_{|J|=k} Y_J^{\to n},$$

where we make use of the notation defined in (2.6). Hence, if $\mathcal{Y}$ is a collection of positive semidefinite $k \times k$ matrices, then $\Psi(\mathcal{Y}) \in \mathrm{FW}_n(k)$. Furthermore, let

$$\mathcal{Y}_0 = \{Y_J = \left(1/C^{n-1}_{k-1}\right) I_{k\times k} : J \subset [n], |J| = k\}, \tag{2.9}$$

where we denote for $n,k \in \mathbb{N}$ the binomial coefficient as $\binom{n}{k} =: C_k^n$, so that $\Psi(\mathcal{Y}_0) = I$. Now let X_ℓ be a strictly feasible solution to a problem of form (2.2) and rescale the data matrices with respect to X_ℓ. Recall the feasible set of the resulting SDP is given by

$$L_\ell = \{X \in \mathbb{S}^n : \mathcal{A}^{(\ell)}(X) = b\}. \tag{2.10}$$

Likewise, the feasible set of the factor width relaxation written over $\mathbb{S}_+^{(n,k)}$ (cf. (2.8)) can be written as

$$L_\ell^\Psi = \{\mathcal{Y} \in \mathbb{S}^{(n,k)} : (\mathcal{A}^{(\ell)} \circ \Psi)(\mathcal{Y}) = b\}. \tag{2.11}$$

Note that $I \in L_\ell$ and $\mathcal{Y}_0 \in L_\ell^\Psi$. We emphasize that, by definition, for any element $\mathcal{Y} \in L_\ell^\Psi$ we have $\Psi(\mathcal{Y}) \in L_\ell$.

Main method

The algorithm requires a striclty feasible starting point X_0 close to the central path, which is used in the first rescaling step. We also require an $\varepsilon > 0$, i.e., our desired accuracy as well as a parameter $\sigma \in (0,1)$ used in the predictor step. In the following let $f^{\mathrm{FW}(k)}$ be a self-concordant barrier function for $\mathbb{S}_+^{(n,k)}$ (we postpone its derivation to Section 2.2, for now we assume it exists and is efficiently computable). In the algorithm we denote the restriction of $f^{\mathrm{FW}(k)}$ to the subspace $\mathrm{null}(L_\ell^\Psi)$ by $f^{\mathrm{FW}(k)}_{|\mathrm{null}(L_\ell^\Psi)}$. The algorithm initializes $\ell = 0$. The outer while loop repeats until an ε-optimal solution is found. If after rescaling with respect to X_ℓ the Newton decrement (cf. Definition A.8) at $\mathcal{Y}_0$ satisfies

$$\Delta\left(f^{\mathrm{FW}(k)}_{|\mathrm{null}(L_\ell^\Psi)}, \mathcal{Y}_0\right) \leq 1/14,$$

the predictor subroutine is called. Here, the affine-scaling direction is projected onto the null space of L_ℓ^Ψ, call it $\mathcal{Z}$. Clearly, $\mathcal{Y}_0 + s\mathcal{Z} \in L_\ell^\Psi$ for all $s \in \mathbb{R}$. Then the subroutine computes

$$s^* = \sup\left\{s : \mathcal{Y}_0 - s\mathcal{Z} \in \mathbb{S}_+^{(n,k)}\right\},$$

which provides the necessary notion of distance to the boundary in terms of $\mathcal{Y}_0$ and $\mathcal{Z}$. The returned point $\mathcal{Y}_\ell := \mathcal{Y}_0 + \sigma s^* \mathcal{Z}$ is feasible and decreases the objective value, as shown in Section 2.4. If the Newton decrement is not small enough, the

corrector subroutine is called. Let $v_\ell = \langle A_0, X_\ell \rangle$, i.e., the objective value of the previous iteration, and define

$$L_\ell^\Psi(v_\ell) = \{\mathcal{Y} \in \mathbb{S}_+^{(n,k)} : \langle A_0, \Psi(\mathcal{Y}) \rangle = v_\ell, \mathcal{A}^{(\ell)}(\Psi(\mathcal{Y})) = b\}.$$

Let $x_0 := \mathcal{Y}_0$. Denote by $n_{|L_\ell^\Psi(v_\ell)}(x_i)$ the Newton step of $f^{\mathrm{FW}(k)}_{|L_\ell^\Psi(v_\ell)}$ at a point x_i. The corrector step now computes

$$x_{i+1} = \operatorname{argmin}_t f^{\mathrm{FW}(k)}\left(x_i + t n_{|L_\ell^\Psi(v_\ell)}(x_i)\right)$$

until x_{i+1} is close enough to the central path of the rescaled problem over $\mathbb{S}_+^{(n,k)}$ and returns $\mathcal{Y}_\ell := x_{i+1}$. We will prove in Section 2.3 how this leads to a decrease in distance to the central path of the original SDP. Note that multiple calls of the corrector step may be necessary as after rescaling the Newton decrement might not be small enough anymore. However, as we prove in Section 2.4, the maximum number of corrector steps can be bounded in terms of the problem data. Let $\mathcal{Y}_\ell$ be the point returned by one of the subroutines. We set

$$X_{\ell+1} = X_\ell^{1/2} \Psi(\mathcal{Y}_\ell) X_\ell^{1/2}.$$

Then

$$\langle A_i^{(\ell+1)}, I \rangle = \langle A_i^{(\ell)}, \Psi(\mathcal{Y}_\ell) \rangle = \langle A_i, X_{\ell+1} \rangle$$

for all $i = 0, 1, \ldots, m$.

Termination criterion

In the predictor as well as in the corrector subroutine we solve a linear system for $y \in \mathbb{R}^m$. The solution of this linear system may be interpreted as a dual feasible solution provided the current iterate is sufficiently close to the central path. Hence, we can approximate the duality gap of our problem by calculating the difference

$$\langle A_0, X_\ell \rangle - y^\top b \geq 0,$$

where y is calculated in every subroutine call. We may use this as a termination criterion. Once the duality gap falls below the required $\varepsilon > 0$ chosen beforehand, we terminate with an ε-optimal solution.

Algorithm 2.1 Predictor-Corrector SDP algorithm using $\mathrm{FW}_n(k)$

Require: $\varepsilon > 0$, $\sigma \in (0,1)$, X_0 strictly feasible
  $\ell \leftarrow 0$
  while Duality gap $> \varepsilon$ **do**
    $A_i^{(\ell)} \leftarrow (X_\ell)^{1/2} A_i (X_\ell)^{1/2}$, for $i = 0, 1, ..., m$
    if $\Delta\left(f^{\mathrm{FW}(k)}_{|\mathrm{null}(L_\ell^\Psi)}, \mathcal{Y}_0\right) \leq \frac{1}{14}$ **then**
      $\mathcal{Y}_\ell \leftarrow$ `Predictor_Step`$(\mathcal{A}^{(\ell)}, A_0^{(\ell)}, \sigma)$
    else
      $\mathcal{Y}_\ell \leftarrow$ `Corrector_Step`$(\mathcal{A}_0^{(\ell)} \circ \Psi, f^{\mathrm{FW}(k)}, \mathcal{Y}_0)$
    $X_{\ell+1} \leftarrow (X_\ell)^{1/2} \Psi(\mathcal{Y}_\ell) (X_\ell)^{1/2}$
    $\ell \leftarrow \ell + 1$
  return X_ℓ

Algorithm 2.2 Subroutine `Predictor_Step`$(\mathcal{A}, A_0, \sigma)$

Require: $\mathcal{A}, A_0, \sigma \in (0,1)$
  Solve for y: $\mathcal{A}A_0 = \mathcal{A}\mathcal{A}^* y$
  $\mathcal{Z} = \Psi^\dagger(\mathcal{A}^* y - A_0)$
  $s^* \leftarrow \sup\{s : \mathcal{Y}_0 - s\mathcal{Z} \in \mathrm{FW}_n(k)\}$
  $\mathcal{Y} \leftarrow \mathcal{Y}_0 - \sigma s^* \mathcal{Z}$
  return $\mathcal{Y}$

Algorithm 2.3 Subroutine `Corrector_Step`$(\mathcal{A}, f, x^{(0)})$

Require: $\mathcal{A}, f, x^{(0)} : \Delta\left(f|_L, x^{(0)}\right) > \frac{1}{14}$, $(L = \mathrm{null}(\mathcal{A}))$
  $j \leftarrow 0$
  while $\Delta\left(f_{|_L}, x^{(j)}\right) > \frac{1}{14}$ **do**
    Solve for y: $\mathcal{A}H(x^{(j)})^{-1}\mathcal{A}^* y = \mathcal{A}H(x^{(j)})^{-1} g(x^{(j)})$
    $n_{|_L}(x^{(j)}) \leftarrow H(x^{(j)})^{-1}\left(\mathcal{A}^* y - g(x^{(j)})\right)$
    $x^{(j+1)} \leftarrow \mathrm{argmin}_t\, f\left(x^{(j)} + t n_{|_L}(x^{(j)})\right)$
    $j \leftarrow j + 1$
  return $x^{(j-1)}$

2.2 Barrier functionals for $\mathbb{S}^n_+$ and $\mathrm{FW}_n(k)$

In this section we derive the self-concordant barrier functional for the cone $\mathbb{S}^{(n,k)}_+$ which is used in the algorithm. Note that the ordinary self-concordant barrier for $\mathbb{S}^n_+$ is given by $f^{\mathrm{SDP}}(X) = -\log(\det(X))$. We will emphasize parallels to the work of Roig-Solvas and Sznaier [RSS22].

In order to construct a self-concordant barrier function for our underlying set, we introduce the notions of hypergraphs and edge colorings as well as a well-known result about these objects.

Definition 2.2. A hypergraph $\mathcal{H} = (V, E)$ consists of a set $V = \{1, \dots, n\}$ of vertices and a set of hyperedges $E \subseteq \{J \subseteq V : |J| \geq 2\}$, which are subsets of the vertex set V. If all elements in E contain exactly k vertices, we call the corresponding hypergraph *k-uniform*. If $\mathcal{H}$ is k-uniform and E consists of *all* subsets of V of size k, it is referred to as the *complete* k-uniform hypergraph.

Definition 2.3. Let $\mathcal{H} = (V, E)$ be a hypergraph. A proper hyperedge coloring with m colors is a partition of the hyperedge set E into m disjoint sets, say $E = \cup_{i \in [m]} S_i$ such that $S_i \cap S_j = \emptyset$ if $i \neq j$, i.e., two hyperedges that share a vertex are not in the same set. In other words, a proper hyperedge coloring assigns a color to every hyperedge such that, if a vertex appears in two different hyperedges, they have different colors.

Theorem 2.4 (Baranyai's theorem [Bar75])**.** *Let $k, n \in \mathbb{N}$ such that $k|n$ and let K^n_k the complete k-uniform hypergraph on n vertices. Then there exists a proper hyperedge coloring using C^{n-1}_{k-1} colors.*

Remark 2.5. *For the case $k = 2$ the above theorem reduces to the statement that the complete graph has a proper edge coloring using $n-1$ colors. This fact was used by Roig-Solvas and Sznaier in their analysis [RSS22] as they only considered the case $k \leq 2$.*

In (2.8) we wrote a program over $\mathrm{FW}_n(k)$ as an equivalent program over the cone product $\mathbb{S}^{(n,k)}_+$. The algorithm uses a self-concordant barrier function over said cone product. The mapping Ψ from $\mathbb{S}^{(n,k)}_+$ to $\mathrm{FW}_n(k)$ is surjective, but not bijective, since multiple elements in the former may give rise to the same element in the latter set.

Assumption 2.6. *Throughout we will assume $k|n$ for some $n \in \mathbb{N}$ and $2 \leq k \in \mathbb{N}$.*

In the following we will let $\mathcal{J} = \{J \subset [n] : |J| = k\}$ and

$$\mathcal{Y} = \{Y_J : J \in \mathcal{J}\}$$

be a collection of $\binom{n}{k}$ matrices of size $k \times k$. We recall the operator Ψ is defined as

$$\Psi(\mathcal{Y}) = \sum_{J \in \mathcal{J}} Y_J^{\to n}.$$

The following generalizes Lemma 4.4 in [RSS22], where a similar result is proved for $k = 2$. It will be crucial in our analysis as it allows us to compare the values taken by the barrier functionals on $\mathbb{S}_+^{(n,k)}$ and $\mathbb{S}_+^n$ at $\mathcal{Y}$ and $\Psi(\mathcal{Y})$, respectively.

Lemma 2.7. *Let*

$$f^{\mathrm{FW}(k)}(\mathcal{Y}) = -\sum_{J \in \mathcal{J}} \log(\det(Y_J)),\ \mathcal{Y} \in \operatorname{int}\left(\mathbb{S}_+^{(n,k)}\right).$$

The barrier $f^{\mathrm{FW}(k)}(\mathcal{Y})$ is self-concordant on $\operatorname{int}\left(\mathbb{S}_+^{(n,k)}\right)$. Furthermore, if $X = \Psi(\mathcal{Y})$ then

$$\begin{aligned} f^{\mathrm{FW}(k)}(\mathcal{Y}) &\geq -C_{k-1}^{n-1} \log(\det(X)) + nC_{k-1}^{n-1} \log\left(C_{k-1}^{n-1}\right) \\ &=: C_{k-1}^{n-1} f^{\mathrm{SDP}}(X) + nC_{k-1}^{n-1} \log\left(C_{k-1}^{n-1}\right). \end{aligned}$$

Let us emphasize here that $f^{\mathrm{FW}(k)}$ is a self-concordant barrier for $\mathbb{S}_+^{(n,k)}$ not $\mathrm{FW}_n(k)$. Before proving Lemma 2.7, we need an auxiliary result which extends Lemma A.1 from [RSS22] to general values of k such that $k|n$. To prove it we will make use of Theorem 2.4.

Lemma 2.8. *Consider the set $\mathcal{Y} = \{Y_J : J \in \mathcal{J}\}$ consisting of positive definite $k \times k$ matrices and let $X = \Psi(\mathcal{Y}) \in \mathrm{FW}_n(k)$. Then there exists a set of C_{k-1}^{n-1} matrices $Z_i \succ 0$ of size $n \times n$ such that $X = \sum_{i=1}^{C_{k-1}^{n-1}} Z_i$ and $f^{\mathrm{FW}(k)}(\mathcal{Y}) = -\sum_{i=1}^{C_{k-1}^{n-1}} \log(\det(Z_i))$.*

Proof. Let K_k^n be the complete k-uniform hypergraph on n vertices. We can identify each hyperedge $\{i_1, i_2, \ldots, i_k\} \subset [n]$ in K_k^n with exactly one element $Y_J \in \mathcal{Y}$, namely the one where $\{i_1, i_2, \ldots, i_k\} = J$. Let $\{S_1, \ldots, S_{C_{k-1}^{n-1}}\}$ be a hyperedge coloring of K_k^n. Define $\mathcal{Y}_i := \{Y_J : J \in S_i\}$ and set $Z_i := \Psi(\mathcal{Y}_i)$. Then $X = \sum_{i=1}^{C_{k-1}^{n-1}} Z_i$ since $S_i \cap S_j = \emptyset$ for $i \neq j$ and $\cup_i S_i = \mathcal{J}$. Clearly, $Z_i \succ 0$. Moreover, since each S_i induces a perfect matching, there exists a permutation matrix P_i for every $i = 1, \ldots, C_{k-1}^{n-1}$

such that $P_i Z_i P_i^\top$ is a block-diagonal matrix with blocks Y_J on the diagonal for $J \in S_i$. From this we find

$$\log(\det(Z_i)) = \log(\det(P_i Z_i P_i^\top)) = \sum_{J \in S_i} \log(\det(Y_J)).$$

Hence,

$$\begin{aligned}\sum_{i=1}^{C_{k-1}^{n-1}} \log(\det(Z_i)) &= \sum_{i=1}^{C_{k-1}^{n-1}} \sum_{J \in S_i} \log(\det(Y_J)) \\ &= \sum_{J \in \mathcal{J}} \log(\det(Y_J)) = -f^{\mathrm{FW}(k)}(\mathcal{Y}),\end{aligned}$$

completing the proof. □

We continue to prove Lemma 2.7.

Proof. (Lemma 2.7) The self-concordance of $f^{\mathrm{FW}(k)}$ on $\operatorname{int}\left(\mathbb{S}_+^{(n,k)}\right)$ follows immediately from the self-concordance of $-\log\det(X)$ on $\operatorname{int}(\mathbb{S}^n)$. By assumption $X = \Psi(\mathcal{Y}) = \sum_{i=1}^{C_{k-1}^{n-1}} Z_i \in \mathrm{FW}_n(k)$. Therefore,

$$\begin{aligned}-\log(\det(X)) &= -\log\left(\det\left(\frac{1}{C_{k-1}^{n-1}} \sum_{i=1}^{C_{k-1}^{n-1}} C_{k-1}^{n-1} Z_i\right)\right) \\ &\leq -\sum_{i=1}^{C_{k-1}^{n-1}} \frac{1}{C_{k-1}^{n-1}} \log\left(\det\left(C_{k-1}^{n-1} Z_i\right)\right) \\ &= -\sum_{i=1}^{C_{k-1}^{n-1}} \frac{1}{C_{k-1}^{n-1}} \log\left({C_{k-1}^{n-1}}^n \det(Z_i)\right),\end{aligned}$$

where the inequality follows by convexity. Hence, we find

$$\begin{aligned}-C_{k-1}^{n-1} \log(\det(X)) &\leq -\sum_{i=1}^{C_{k-1}^{n-1}} \left(n \log\left(C_{k-1}^{n-1}\right) + \log(\det(Z_i))\right) \\ &= -\sum_{i=1}^{C_{k-1}^{n-1}} \log(\det(Z_i)) - C_{k-1}^{n-1} n \log\left(C_{k-1}^{n-1}\right),\end{aligned}$$

and the claim follows. □

The following corollary is analogous to Corollary 4.5 from [RSS22].

Corollary 2.9. *If*

$$\mathcal{Y}_0 = \{Y_J = 1/C_{k-1}^{n-1} I_{k\times k} : J \subset [n], |J| = k\}$$

then $X = \Psi(\mathcal{Y}_0) = I$ *and*

$$\begin{aligned} f^{\mathrm{FW}(k)}(\mathcal{Y}_0) &= C_{k-1}^{n-1} f^{\mathrm{SDP}}(X) + nC_{k-1}^{n-1} \log\left(C_{k-1}^{n-1}\right) \\ &= nC_{k-1}^{n-1} \log\left(C_{k-1}^{n-1}\right). \end{aligned}$$

Proof. The first statement follows when noting that each $i \in [n]$ lies in exactly $\binom{n-1}{k-1}$ subsets of $[n]$ of size k. The reason is that, when fixing i, there are $n-1$ elements left out of which we want to choose $k-1$ more elements to make a set of size k. For the second statement note that

$$\log\left(\det\left(\frac{1}{C_{k-1}^{n-1}} I_{k\times k}\right)\right) = \log\left(\left(C_{k-1}^{n-1}\right)^{-k}\right) = -k\log\left(C_{k-1}^{n-1}\right).$$

The result follows when noting that $k\binom{n}{k} = nC_{k-1}^{n-1}$. □

2.3 Relations of the barrier functions

To prove convergence of our algorithm we need two essential ingredients. First, we need to prove that the predictor step reduces the current objective value sufficiently, and secondly, we must prove that the corrector step converges to a point close to the central path. Moreover, we have to show that our criterion to decide which subroutine to call is valid. The issue here is that we compute the Newton decrement of $f^{\mathrm{FW}(k)}$ at $\mathcal{Y}_0$, but we need to be able to assert that the Newton decrement of f^{SDP} at X_ℓ is small enough.

The next result we present will allow us to lower bound the progress made by the corrector step. For this we need to be able to compare the barrier functions for $\mathbb{S}_+^n$ and $\mathbb{S}_+^{(n,k)}$. We assume we have a given feasible solution X_ℓ such that $\langle A_0^{(\ell)}, I\rangle = \nu$. Define the vector $b(\nu) := (\nu, b_1, \ldots, b_m)^\top$. For further reference, consider

$$\min\left\{f^{\mathrm{SDP}}(X) : \langle A_i^{(\ell)}, X\rangle = b(\nu)_i \ \forall i = 0, 1, \ldots, m, X \in \mathbb{S}_+^n\right\}, \tag{2.12}$$

which we would like to compare to

$$\min\left\{f^{\mathrm{FW}(k)}(\mathcal{Y}) : \mathcal{Y} \in L_\ell^\Psi(\nu) \cap \mathbb{S}_+^{(n,k)}\right\}. \tag{2.13}$$

Suppose $\mathcal{Y}^*$ is an approximate solution to (2.13). Defining

$$X_{\ell+1} = X_\ell^{1/2}\Psi(\mathcal{Y}^*)X_\ell^{1/2},$$

we find that $X_\ell \in \mathcal{F}_{\text{SDP}}$ for all ℓ. In other words, the points X_ℓ we obtain via this procedure are all feasible for the original SDP (2.2). The following lemma allows us to lower bound the decrease achieved by one corrector step in terms of an element in $\mathbb{S}_+^{(n,k)}$.

Lemma 2.10. *Let $\mathcal{Y}^*$ be a feasible solution to* (2.13) *and $\mathcal{Y}_0$ as in* (2.9). *Further, let $X_{\ell+1} = X_\ell^{1/2}\Psi(\mathcal{Y}^*)X_\ell^{1/2}$ for X_ℓ a feasible solution to* (2.13). *Then*

$$C_{k-1}^{n-1}\left(f^{\text{SDP}}(X_\ell) - f^{\text{SDP}}(X_{\ell+1})\right) \geq f^{\text{FW}(k)}(\mathcal{Y}_0) - f^{\text{FW}(k)}(\mathcal{Y}^*).$$

Proof. The proof follows immediately when noting that

$$\begin{aligned} C_{k-1}^{n-1}\left(f^{\text{SDP}}(X_\ell) - f^{\text{SDP}}(X_{\ell+1})\right) &= C_{k-1}^{n-1}\left(f^{\text{SDP}}(X_\ell) - f^{\text{SDP}}(X_\ell^{1/2}\Psi(\mathcal{Y}^*)X_\ell^{1/2}\right) \\ &= \underbrace{nC_{k-1}^{n-1}\log(C_{k-1}^{n-1})}_{=f^{\text{FW}(k)}(\mathcal{Y}_0) \text{ by Cor. 2.9}} \underbrace{-f^{\text{SDP}}(\Psi(\mathcal{Y}^*)) - nC_{k-1}^{n-1}\log C_{k-1}^{n-1}}_{\geq -f^{\text{FW}(k)}(\mathcal{Y}^*) \text{ by Lemma 2.7}}. \end{aligned}$$

□

2.3.1 Relation of the Newton decrements

If $f_{|L}$ is a self-concordant function restricted to a (translated) linear subspace L we have

$$\Delta\left(f_{|L}, x\right) \geq \frac{\langle d, n(x)\rangle_x}{||d||_x} \text{ for all } d \in L \setminus \{0\}, \tag{2.14}$$

see Appendix A.2 for details. We will continue to prove that we can upper bound the Newton decrement of f^{SDP} at the identity in terms of the Newton decrement of $f^{\text{FW}(k)}$ at $\mathcal{Y}_0$. To this end, define the following operator

$$\Psi^\dagger : \mathbb{S}^n \to \mathbb{S}^{(n,k)}$$

via

$$\left(\Psi^\dagger(X)\right)_J = \left(\frac{1}{C_{k-1}^{n-1}}I + \frac{1}{C_{k-2}^{n-2}}(ee^\top - I)\right) \circ X_{J,J} \quad \text{for } J \subset [n], |J| = k,$$

where $\circ$ denotes the Hadamard product.

This operator satisfies

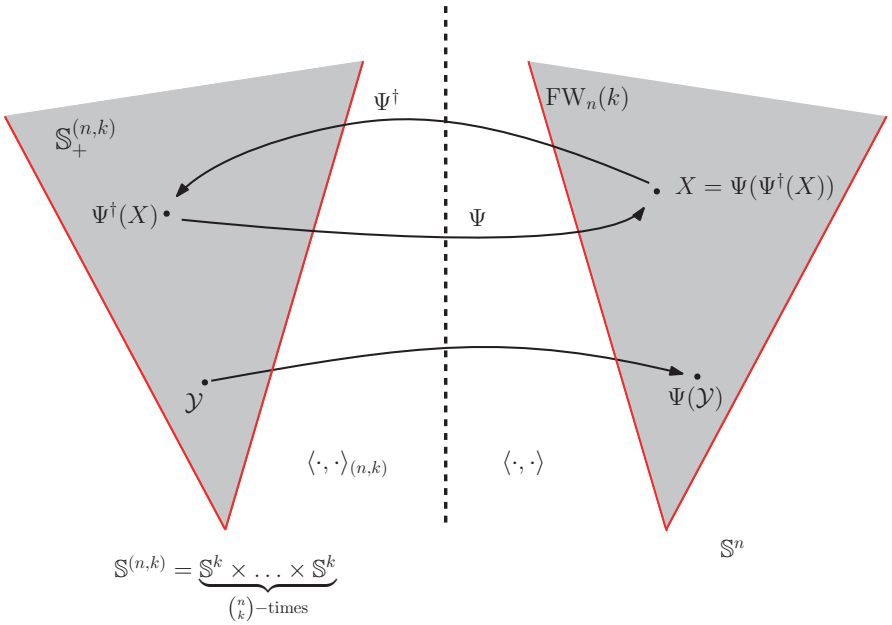


Figure 2.1: Visualization of the surjection from $\mathbb{S}_+^{(n,k)}$ to $\mathrm{FW}_n(k)$

$$\Psi(\Psi^\dagger(X)) = X \text{ for all } X \in \mathbb{S}^n.$$

In other words, $\Psi^\dagger$ is a right-inverse of the linear operator Ψ. See Figure 2.1 for a visualization of the surjection from $\mathbb{S}_+^{(n,k)}$ to $\mathrm{FW}_n(k)$.
An inner product on $\mathbb{S}^{(n,k)}$ is given by

$$\langle \mathcal{X}, \mathcal{Y} \rangle_{(n,k)} := \sum_{|J|=k} \langle X_J, Y_J \rangle,$$

and it is well-defined for $\mathcal{X} = \{X_J \in \mathbb{S}^k : |J| = k\}, \mathcal{Y} = \{Y_J \in \mathbb{S}^k : |J| = k\}$.

Lemma 2.11. *For any $X \in \mathbb{S}^n$ we have*

$$||\Psi^\dagger(X)||_{(n,k)} \leq ||X||.$$

Proof. Let $X \in \mathbb{S}^n$. For a matrix $Y \in \mathbb{S}^n$ we define $\mathrm{diag}(Y) \in \mathbb{R}^n$ to be the vector consisting of the diagonal entries of Y, and we further define a diagonal matrix $\mathrm{Diag}(Y) \in \mathbb{S}^n$ as

$$\mathrm{Diag}(Y)_{i,j} = \begin{cases} Y_{i,i} & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases}$$

Then,

$$\begin{aligned}
||\Psi^\dagger(X)||^2_{(n,k)} &= \sum_{|J|=k} \langle \left(\frac{1}{C^{n-1}_{k-1}} I + \frac{1}{C^{n-2}_{k-2}} (ee^\top - I) \right) \circ X_J, \\
&\qquad \left(\frac{1}{C^{n-1}_{k-1}} I + \frac{1}{C^{n-2}_{k-2}} (ee^\top - I) \right) \circ X_J \rangle \\
&= \sum_{|J|=k} \langle \left(\left(\frac{1}{C^{n-1}_{k-1}} - \frac{1}{C^{n-2}_{k-2}} \right) I + \frac{ee^\top}{C^{n-2}_{k-2}} \right) \circ X_J, \\
&\qquad \left(\left(\frac{1}{C^{n-1}_{k-1}} - \frac{1}{C^{n-2}_{k-2}} \right) I + \frac{ee^\top}{C^{n-2}_{k-2}} \right) \circ X_J \rangle \\
&= \sum_{|J|=k} \frac{1}{(C^{n-2}_{k-2})^2} \langle X_J, X_J \rangle + \left(\frac{1}{C^{n-1}_{k-1}} - \frac{1}{C^{n-2}_{k-2}} \right) \langle \mathrm{Diag}(X_J), \mathrm{Diag}(X_J) \rangle + \\
&\qquad 2 \left(\frac{1}{C^{n-1}_{k-1}} - \frac{1}{C^{n-2}_{k-2}} \right) \frac{1}{C^{n-2}_{k-2}} \langle X_J, \mathrm{Diag}(X_J) \rangle \\
&= \sum_{|J|=k} \frac{1}{(C^{n-2}_{k-2})^2} \langle X_J, X_J \rangle + \\
&\qquad \left(\frac{1}{C^{n-1}_{k-1}} - \frac{1}{C^{n-2}_{k-2}} \right) \left(\frac{1}{C^{n-1}_{k-1}} + \frac{1}{C^{n-2}_{k-2}} \right) \mathrm{diag}(X_J)^\top \mathrm{diag}(X_J) \\
&\leq \sum_{|J|=k} \frac{1}{C^{n-2}_{k-2}} \langle X_J, X_J \rangle + \left(\frac{1}{C^{n-1}_{k-1}} - \frac{1}{C^{n-2}_{k-2}} \right) \mathrm{diag}(X_J)^\top \mathrm{diag}(X_J) \\
&= \langle X, X \rangle = ||X||^2,
\end{aligned}$$

where the inequality follows from

$$\frac{1}{(C^{n-2}_{k-2})^2} \leq \frac{1}{C^{n-2}_{k-2}} \text{ and } \left(\frac{1}{C^{n-1}_{k-1}} + \frac{1}{C^{n-2}_{k-2}} \right) \leq 1. \qquad \square$$

Suppose now X_ℓ is a feasible solution to (2.5) such that $\langle A_0, X_\ell \rangle = \nu$. We define the vector $b(\nu) := (\nu, b_1, \ldots, b_m)^\top$ as well as the two subspaces

$$L^\Psi_\ell = \{ \mathcal{Y} \in \mathbb{S}^{(n,k)} : (\mathcal{A}^{(\ell)} \circ \Psi)(\mathcal{Y}) = b \}$$

and

$$L_\ell = \{ X \in \mathbb{S}^n : \mathcal{A}^{(\ell)}(X) = b \}.$$

Note that we may also want to keep track of the objective, in which case we will refer to the following operator

$$\mathcal{A}_0^{(\ell)}(X) = (\langle A_0^{(\ell)}, X\rangle, \langle A_1^{(\ell)}, X\rangle, \ldots, \langle A_m^{(\ell)}, X\rangle) \in \mathbb{R}^{m+1}.$$

The respective subspaces will be denoted as follows

$$L_\ell^\Psi(v) = \{\mathcal{Y} \in \mathbb{S}^{(n,k)} : (\mathcal{A}_0^{(\ell)} \circ \Psi)(\mathcal{Y}) = b(v)\} \tag{2.15}$$

and

$$L_\ell(v) = \{X \in \mathbb{S}^n : \mathcal{A}_0^{(\ell)}(X) = b(v)\}. \tag{2.16}$$

When we consider the subspaces defined via the operator with respect to the initial data matrices, we omit the subscript ℓ, e.g.,

$$L^\Psi = \{\mathcal{Y} \in \mathbb{S}^{(n,k)} : \langle A_i, \Psi(\mathcal{Y})\rangle = b_i\,,\ \forall i \in [m]\}.$$

The following lemma corresponds to Lemma A.2 in [RSS22], and allows us to bound the Newton decrement of $f_{|L}^{\mathrm{SDP}}$ in terms of $f_{|L}^{\mathrm{FW}(k)}$.

Lemma 2.12. *Assume $\mathcal{Y}_0 \in L^\Psi$ and $I \in L$. At $\mathcal{Y}_0$ one has*

$$\Delta\left(f_{|_{L^\Psi}}^{\mathrm{FW}(k)}, \mathcal{Y}_0\right) \geq \frac{\Delta\left(C_{k-1}^{n-1} f_{|_L}^{\mathrm{SDP}}, I\right)}{\sqrt{C_{k-1}^{n-1}}} = \sqrt{C_{k-1}^{n-1}}\,\Delta\left(f_{|_L}^{\mathrm{SDP}}, I\right).$$

Proof. Following (2.14) we have

$$\Delta\left(f_{|_{L^\Psi}}^{\mathrm{FW}(k)}, \mathcal{Y}\right) \geq \frac{\langle d, n^{\mathrm{FW}}(\mathcal{Y})\rangle_{(n,k),\mathcal{Y}}}{||d||_{(n,k),\mathcal{Y}}} \text{ for all } d \in L \setminus \{0\}.$$

Choosing $d = \Psi^\dagger(n_L^{\mathrm{SDP}}(X)) \in L$ leads to

$$\Delta\left(f_{|_{L^\Psi}}^{\mathrm{FW}(k)}, \mathcal{Y}\right) \geq \frac{\langle \Psi^\dagger(n_L^{\mathrm{SDP}}(X)), n^{\mathrm{FW}}(\mathcal{Y})\rangle_{(n,k),\mathcal{Y}}}{||\Psi^\dagger(n_L^{\mathrm{SDP}}(X))||_{(n,k),\mathcal{Y}}},$$

and evaluating the expression at $\mathcal{Y}_0$ we find

$$\begin{aligned}
\Delta\left(f_{|_{L^\Psi}}^{\mathrm{FW}(k)}, \mathcal{Y}_0\right) &\geq \frac{\langle \Psi^\dagger(n_L^{\mathrm{SDP}}(X)), n^{\mathrm{FW}}(\mathcal{Y}_0)\rangle_{(n,k),\mathcal{Y}}}{||\Psi^\dagger(n_L^{\mathrm{SDP}}(X))||_{(n,k),\mathcal{Y}}} \\
&= \frac{\langle \Psi^\dagger(n_L^{\mathrm{SDP}}(X)), -g^{\mathrm{FW}}(\mathcal{Y}_0)\rangle_{(n,k)}}{C_{k-1}^{n-1}||\Psi^\dagger(n_L^{\mathrm{SDP}}(X))||_{(n,k)}} \\
&\geq \frac{\langle \Psi^\dagger(n_L^{\mathrm{SDP}}(X)), (I, I, \ldots, I)\rangle_{(n,k)}}{||n_L^{\mathrm{SDP}}(X)||} \\
&= \frac{\mathrm{Tr}(n_L^{\mathrm{SDP}}(X))}{||n_L^{\mathrm{SDP}}(X)||},
\end{aligned}$$

where the second inequality follows from Lemma 2.11. Setting $X = I$ and noting

$$\begin{aligned}\mathrm{Tr}(n_L^{\mathrm{SDP}}(I)) = \langle I, n_L^{\mathrm{SDP}}(I)\rangle &= \frac{1}{C_{k-1}^{n-1}}\langle g^{\mathrm{SDP}}(I), -n_L^{\mathrm{SDP}}(I)\rangle \\ &= \frac{1}{C_{k-1}^{n-1}}\left(\Delta\left(C_{k-1}^{n-1} f^{\mathrm{SDP}}_{|_L}, I\right)\right)^2\end{aligned}$$

we conclude

$$\Delta\left(f^{\mathrm{FW}(k)}_{|_{L^\Psi}}, \mathcal{Y}_0\right) \geq \frac{1}{C_{k-1}^{n-1}} \frac{\Delta\left(C_{k-1}^{n-1} f^{\mathrm{SDP}}_{|_L}, I\right)^2}{||n_L^{\mathrm{SDP}}(I)||} = \frac{\Delta\left(C_{k-1}^{n-1} f^{\mathrm{SDP}}_{|_L}, I\right)}{\sqrt{C_{k-1}^{n-1}}},$$

because

$$||n_L^{\mathrm{SDP}}(I)|| = \frac{\Delta\left(C_{k-1}^{n-1} f^{\mathrm{SDP}}_{|L}, I\right)}{\sqrt{C_{k-1}^{n-1}}} = \sqrt{C_{k-1}^{n-1}}\,\Delta\left(f^{\mathrm{SDP}}_{|L}, I\right). \qquad \square$$

2.4 Complexity analysis

We begin the complexity analysis with the following lemma, which helps us to check whether the current point is close enough to the central path of the SDP.

Lemma 2.13. *Let X_ℓ be a feasible iterate for the SDP* (2.12) *and let the objective value at X_ℓ be ν. Define the two subspaces $L_\ell^\Psi(\nu)$, L_ℓ as in* (2.15), (2.10) *respectively. Then, if*

$$\Delta\left(f^{\mathrm{FW}(k)}_{|L_\ell^\Psi(\nu)}, \mathcal{Y}_0\right) \leq \frac{1}{14},$$

one has

$$\Delta\left(f^{\mathrm{SDP}}_{\eta_\nu|_{L_\ell}}, I\right) \leq \frac{1}{9},$$

where

$$f^{\mathrm{SDP}}_{\eta_\nu}(X) = \eta_\nu\langle A_0, X\rangle - \log\det(X),$$

and η_ν is such that

$$\nu = \min_{X\in L_\ell} f^{\mathrm{SDP}}_{\eta_\nu}(X).$$

Proof. By Lemma 2.12 we know that

$$\frac{1}{14} \geq \Delta\left(f^{\mathrm{FW}(k)}_{|L_\ell^\Psi(\nu)}, \mathcal{Y}_0\right) \geq \Delta\left(f^{\mathrm{SDP}}_{|L_\ell(\nu)}, I\right).$$

Let now $z(v)$ be the point on the central path of the rotated SDP with objective value v and let the corresponding parameter be η_v. By Theorem 2.2.5 from [Ren01] we have

$$||z(v)-I||_I \leq \Delta\left(f^{\text{SDP}}_{|L_\ell(v)}, I\right) + \frac{3\Delta\left(f^{\text{SDP}}_{|L_\ell(v)}, I\right)^2}{\left(1-\Delta\left(f^{\text{SDP}}_{|L_\ell(v)}, I\right)\right)^3} \leq \frac{1}{11}. \tag{2.17}$$

Let X_+ be the point returned by taking a Newton step at $X = I$ with respect to the function $f^{\text{SDP}}_{\eta_v}$ restricted to L_ℓ. By Theorem 2.2.3 in [Ren01] we have

$$\frac{||z(v)-I||_I^2}{1-||z(v)-I||_I} \geq ||X_+ - z(v)||_I$$

and hence

$$\begin{aligned}\Delta\left(f^{\text{SDP}}_{\eta_v|_{L_\ell}}, I\right) = ||X_+ - I||_I &\leq ||X_+ - z(v)||_I + ||z(v)-I||_I \\ &\leq \frac{||z(v)-I||_I^2}{1-||z(v)-I||_I} + ||z(v)-I||_I \leq \frac{1}{9}.\end{aligned}$$

□

The Newton decrement of the rotated SDP being smaller than 1/9 means we can safely perform the next predictor step. If the current point is too far away from the central path and one were to perform the predictor step the direction may not be approximately tangential to the central path. Hence, once the Newton decrement of the factor width program is small enough, so is the one of the SDP, and we can perform the next predictor step, knowing the direction will be approximately tangential to the central path. After each predictor step we may have to take several corrector steps, to get back close to the central path.

Corrector step

We will now find an upper bound on the number of corrector steps needed to get close to the central path. We know from Lemma 2.10 that a decrease in the barrier for the factor width cone will lead to a decrease in the barrier function for our original SDP, meaning we made progress towards its central path. The following lemma asserts that if we are too far away from the central path we can attain at least a constant reduction in the barrier of the factor width cone. Thereby a constant reduction in the SDP barrier is obtained as well.

Lemma 2.14. *Let X_ℓ be a feasible iterate for the SDP* (2.12) *and let the objective value at X_ℓ be ν. Define the subspace $L_\ell^\Psi(\nu)$ as in* (2.15)*. If*

$$\Delta\left(f^{\mathrm{FW}(k)}_{|L_\ell^\Psi(\nu)}, \mathcal{Y}_0\right) > \frac{1}{14}$$

then

$$f^{\mathrm{FW}(k)}_{|L_\ell^\Psi(\nu)}(\mathcal{Y}_0) - f^{\mathrm{FW}(k)}_{|L_\ell^\Psi(\nu)}(\mathcal{Y}^*) \geq \frac{1}{2688}.$$

Proof. If $\Delta\left(f^{\mathrm{FW}(k)}_{|L_\ell^\Psi(\nu)}, \mathcal{Y}_0\right) > \frac{1}{14}$ the corrector step will employ a line search to find $\mathcal{Y}^*$, i.e., the point in $L_\ell^\Psi(\nu)$ that minimizes $f^{\mathrm{FW}(k)}$. Let $n_{L_\ell^\Psi(\nu)}(\mathcal{Y}_0)$ be the Newton step taken from $\mathcal{Y}_0$ and let $t = \frac{1}{8\|n_{L_\ell^\Psi(\nu)}(\mathcal{Y}_0)\|_{(n,k),\mathcal{Y}_0}}$, where the norm in the denominator is the local norm at $\mathcal{Y}_0$ induced by $\langle\cdot,\cdot\rangle_{(n.k)}$. Then, for

$$\tilde{\mathcal{Y}} = \mathcal{Y}_0 + t\, n_{L_\ell^\Psi(\nu)}(\mathcal{Y}_0)$$

we find by Theorem 2.2.2 in [Ren01] (see also (A.11) in Appendix A)

$$\begin{aligned} f^{\mathrm{FW}(k)}(\tilde{\mathcal{Y}}) &\leq f^{\mathrm{FW}(k)}(\mathcal{Y}_0) - \frac{1}{14}\frac{1}{8} + \frac{1}{2}\left(\frac{1}{8}\right)^2 + \frac{(1/8)^3}{3(1-1/8)} \\ &\leq f^{\mathrm{FW}(k)}(\mathcal{Y}_0) - \frac{1}{2688}. \end{aligned}$$

□

Note that this implies together with Lemma 2.10 that

$$\begin{aligned} \frac{1}{2688} \leq f^{\mathrm{FW}(k)}(\mathcal{Y}_0) - f^{\mathrm{FW}(k)}(\tilde{\mathcal{Y}}) &\leq f^{\mathrm{FW}(k)}(\mathcal{Y}_0) - f^{\mathrm{FW}(k)}(\mathcal{Y}^*) \\ &\leq C_{k-1}^{n-1}\left(f^{\mathrm{SDP}}(X_\ell) - f^{\mathrm{SDP}}(X_{\ell+1})\right). \end{aligned} \tag{2.18}$$

With each line search we obtain a point, which reduces the function value of f^{SDP} by a constant amount compared to its predecessor. This fact will allow us to bound the number of line searches needed to find a point for which $\Delta\left(f^{\mathrm{FW}(k)}_{|L_\ell^\Psi(\nu)}, \mathcal{Y}_0\right) \leq 1/14$. Another ingredient needed for this is an upper bound on the difference of function value of f^{SDP} at the result of the predictor step and the corresponding point on the central path with the same objective value.

Lemma 2.15. *Let X_1 be close to a point $z(\nu_1)$ on the central path of the SDP in the sense that $\Delta\left(f^{\mathrm{SDP}}_{L_\ell(\nu_1)}, X_1\right) \leq \frac{1}{9}$. Further, let X_2 be the result of the predictor step and $z(\nu_2)$ be the point on the central path with the same objective value as X_2. Then*

$$f^{\mathrm{SDP}}(X_2) - f^{\mathrm{SDP}}(z(\nu_2)) \leq n\left(\log\frac{1}{1-\sigma}\right) + \frac{1}{154}.$$

Proof. A proof of this statement for generic self-concordant barriers may be found on page 54 of [Ren01]. We have used that the barrier parameter as introduced in (A.1) for the barrier of the psd cone is given by $\vartheta_{f^{\text{SDP}}} = n$. □

Lemma 2.16. *Let v_2 be the objective value of the result X_2 of the predictor step. The maximum number K of line searches needed to find a point X_{K+2} which is close enough to $z(v_2)$ in the sense that $\Delta\left(f^{\text{SDP}}_{|L_\ell(v_2)}, X_{K+2}\right) \leq \frac{1}{9}$ is*

$$K = \left\lceil 2688 C_{k-1}^{n-1} \left(n \log\left(\frac{1}{1-\sigma}\right) + \frac{1}{154} \right) \right\rceil,$$

where $z(v_2)$ is the point on the central path with objective value v_2.

Proof. We know that the distance between the result of the predictor phase and the targeted point on the central path is at most $n\left(\log \frac{1}{1-\sigma}\right) + \frac{1}{154}$ by Lemma 2.15. Moreover, using Lemma 2.14 we find that in each corrector step we reduce this distance by at least $\frac{1}{2688 C_{k-1}^{n-1}}$, unless the SDP Newton decrement at I is already small enough to perform the next predictor step. If after rescaling the Newton decrement of the factor width program satisfies

$$\Delta\left(f^{\text{FW}(k)}_{|L^{\Psi}_{\ell}(v)}, \mathcal{Y}_0\right) > \frac{1}{14},$$

thereby implying by Lemma 2.13 that I is not close to the central path of the SDP we can perform another corrector step yielding at least a constant decrease of $\frac{1}{2688 C_{k-1}^{n-1}}$ of the distance to the central path, and rescale again. This process can be continued until we do not get such a constant decrease anymore at which point we know we must be close enough to the central path, in the sense of Lemma 2.13. Because if the decrease is not greater than $\frac{1}{2688 C_{k-1}^{n-1}}$ we know that the Newton decrement cannot satisfy

$$\Delta\left(f^{\text{FW}(k)}_{|L^{\Psi}_{\ell}(v)}, \mathcal{Y}_0\right) > \frac{1}{14},$$

from which follows by Lemma 2.13 that

$$\Delta\left(f^{\text{SDP}}_{L_\ell(v)}, I\right) \leq \frac{1}{9}.$$

This implies we are close enough to the central path to perform the next predictor step. Hence, after at most

$$K = \left\lceil 2688 C_{k-1}^{n-1} \left(n \log\left(\frac{1}{1-\sigma}\right) + \frac{1}{154} \right) \right\rceil$$

corrector steps we are close enough to the central path so that we can perform the next predictor step. □

Predictor step

We will make use of the analysis of the short step interior point method discussed in Section 2.4.2 in [Ren01]; see also Appendix A.3. We will show that each predictor step reduces the objective value by an amount at least as large as the objective decrease by the short-step interior point method. This will allow us to conclude the maximum number of predictor steps needed to obtain an ε-optimal solution of the given SDP. Note that the decrease in objective value obtained by our predictor method is as follows. Let X be the point from where the predictor method starts and $-(A_0)_X := -H(X)A_0$ be the direction. Then for $\sigma \geq \frac{1}{4}$ we find

$$\begin{aligned}\langle A_0, X - s^*\sigma\,(A_0)_X\rangle &= \langle (A_0)_X, X\rangle - s^*\sigma\langle A_0, (A_0)_X\rangle \\ &\leq \langle A_0, X\rangle - \frac{1}{4}\|(A_0)_X\|_X.\end{aligned}$$

This implies the decrease is at least as large the one obtained in one iteration of the short-step method, as discussed in [Ren01, § 2.4.2]; see also Appendix A.3. Renegar's analysis shows that short-step method leads to an ε-optimal solution in at most

$$K = 10\sqrt{\vartheta_f}\log(\vartheta_f/(\varepsilon\,\eta_0))$$

steps, where η_0 is such that our starting point X_0 is close to z_{η_0}, in the sense that $\Delta(f^{\text{SDP}}, X_0) < 1/9$ and ϑ_f is the barrier parameter as introduced in (A.1).

Predictor and corrector steps combined

Combining the complexity analysis of predictor and corrector steps we arrive at the following theorem.

Theorem 2.17. *Let X_0 be a feasible solution of the SDP* (2.2) *and assume it is close to some point z_{η_0} on the corresponding central path in the sense that $\Delta\left(f^{\text{SDP}}_{|_{L(\nu)}}, X_0\right) < 1/14$, where L is as in* (2.16) *for $\nu = \langle A_0, X_0\rangle$. Algorithm 2.1 converges to an ε-optimal solution in at most*

$$\begin{aligned}K &= \left\lceil 2688 C_{k-1}^{n-1}\left(n\log\left(\frac{1}{1-\sigma}\right) + \frac{1}{154}\right)\right\rceil 10\sqrt{n}\log(n/(\varepsilon\,\eta_0)) \\ &= O\left(\binom{n-1}{k-1} n^{3/2}\log\left(\frac{1}{1-\sigma}\right)\log\left(\frac{n}{\epsilon\eta_0}\right)\right)\end{aligned}$$

steps.

The assumption of a starting point "close to the central path" may be satisfied by the self-dual embedding strategy [dKRT97]. Alternatively, one may first solve an auxiliary SDP problem, as in [Ren01, § 2.4.2], by using the algorithm we have presented. The solution of this auxiliary problem then yields a point close to the central path of the original SDP problem.

2.5 Discussion and future prospects

We finish with a brief discussion on the prospects of efficient implementation of Algorithm 2.1.

Parallelization

Essentially, the contribution of the present chapter lies in providing an algorithm for solving SDPs which is much more suitable for parallelization than the ordinary interior point method working over $\mathbb{S}^n_+$. Given common memory access, the computation of the necessary data for the respective cone factors $\mathbb{S}^k_+$ is local, meaning these tasks can be distributed among processor cores leading to a runtime decrease since each corrector step involves $\binom{n}{k}$ parallel computations of $O(k^3m+k^2m^2+m^3)$ flops. This offers the potential to perform the centering steps much more quickly than for SDP interior point methods through parallel computation. For more information about parallel implementation of IPMs for semidefinite optimization we refer the reader to [Iva08].

Replacing the predictor step

In their paper [RSS22], the authors propose to perform a fixed number of decrease steps, where a decrease step consists of solving (2.7) and rescaling with respect to the optimal solution. In our algorithm we considered a different method to decrease the objective value, i.e., the predictor method, where we use the traditional SDP affine scaling direction.

Tractability of factor width cones

The entire approach described in this chapter relies on the premise that one may optimize more efficiently over $\mathrm{FW}_n(k)$ than over $\mathbb{S}^n_+$. In practice this has not yet been demonstrated convincingly for $k > 2$, although the consensus is that it should be possible. Some recent ideas that could be useful in this regard are:

- the idea to optimize over the dual cone of $\mathrm{FW}_n(k)$ by utilizing clique trees [ZL21]
- a variation on the factor width cone involving fewer blocks [ZSP22].

In addition, it would be very helpful to know a computable self-concordant barrier functional for the cone $\mathrm{FW}_n(k)$, as well as its complexity parameter.

3

Convergence rates of RLT and Lasserre-type hierarchies

3.1 Introduction

In this chapter we consider GMP's of the following form:

$$\begin{aligned} \text{val} := \inf_{\mu \in \mathcal{M}(\mathbf{K})_+} & \int_{\mathbf{K}} f_0(\mathbf{x}) \mathrm{d}\mu(\mathbf{x}) \\ \text{s.t.} & \int_{\mathbf{K}} f_i(\mathbf{x}) \mathrm{d}\mu(\mathbf{x}) = b_i \ \forall i \in [m] \\ & \int_{\mathbf{K}} \mathrm{d}\mu(\mathbf{x}) \leq 1, \end{aligned} \tag{3.1}$$

where $m \in \mathbb{N}$, $b_i \in \mathbb{R}$ for all $i \in [m]$. Recall that $\mathcal{M}(\mathbf{K})_+$ is the convex cone of positive finite Borel measures supported on $\mathbf{K} \subset \mathbb{R}^n$, and $f_0, f_1, \ldots, f_m$ are continuous on $\mathbf{K}$. We will always assume the GMP (3.1) has a feasible solution, which implies that it has an optimal solution as well (see Theorem 3.1).

The constraint $\int_{\mathbf{K}} \mathrm{d}\mu(\mathbf{x}) \leq 1$ essentially means that we know an upper bound on the measure of $\mathbf{K}$ for the optimal solution, since in this case we may scale the functions f_i a priori to satisfy this condition.

We will consider the case where $\mathbf{K}$ is the standard (probability) simplex

$$\Delta_{n-1} = \left\{\mathbf{x} \in \mathbb{R}^n_+ : x_1 + \cdots + x_n = 1\right\},$$

or the Euclidean sphere

$$\mathcal{S}^{n-1} = \left\{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_2^2 = x_1^2 + \cdots + x_n^2 = 1\right\}.$$

Our main result is to establish a rate of convergence for the Lasserre hierarchy [Las08] for the GMP with polynomial data on the sphere, and for a related, RLT (reformulation-linearization technique)-type linear programming hierarchy for the GMP with polynomial data on the simplex. This RLT hierarchy is in fact a generalization of LP hierarchies for polynomial optimization on the simplex, as introduced by Bomze and de Klerk [BdK01], and de Klerk, Laurent and Parrilo [dKLP06], and is closely related to the original work on RLT hierarchies by Sherali and Adams [AS10].

Outline of the chapter

In Section 3.1 we review the duality theory of the GMP (3.1). For $\mathbf{K}$ the simplex we introduce a linear relaxation hierarchy in this setting in Section 3.2 and prove a convergence rate of $O(1/r)$. Section 3.3 contains the new convergence analysis of the Lasserre [Las09] SDP hierarchies of the GMP on the sphere. In Section 3.4 we take a mathematical view of how the optimal measure is obtained in the limit as the level of the hierarchies approaches infinity. In Section 3.5 we explain how our LP hierarchy is a generalization of an approximation hierarchy for the problem of minimizing a form of degree d over the simplex introduced by de Klerk, Laurent and Parrilo [dKLP06] based on earlier results obtained by Bomze and de Klerk [BdK01].

Duality of the GMP (3.1)

The dual of (3.1) is given by

$$\begin{aligned} \text{val}' = \sup_{(y,t)\in\mathbb{R}^m\times\mathbb{R}_+} & \sum_{i=1}^m y_i b_i - t \\ \text{s.t. } & f_0(\mathbf{x}) - \sum_{i=1}^m y_i f_i(\mathbf{x}) + t \geq 0 \quad \forall\, \mathbf{x} \in \mathbf{K}. \end{aligned} \tag{3.2}$$

Note that the dual problem (3.2) is always strictly feasible, due to the constraint $\int_{\mathbf{K}} \mathrm{d}\mu \leq 1$ in the primal GMP (3.1).

Weak duality holds for this pair of problems, meaning $\mathrm{val}' \leq \mathrm{val}$. In fact, the duality gap is always zero, as the next theorem shows. Note that a zero duality gap does *not* imply the existence of a dual optimal solution.

Theorem 3.1. *(see, e.g., [Las09, Theorem 1.3]) Assume problem* (3.1) *is feasible. Then it has an optimal solution (the* inf *is attained), and* $\mathrm{val} = \mathrm{val}'$.

Theorem 3.1 is a direct consequence of Theorem 1.5. We continue by recalling a sufficient condition for a dual optimal solution to exist.

Theorem 3.2. *(see, e.g., [Sha01, Proposition 2.8]) Suppose problem* (3.1) *is feasible. If*

$$b \in \mathrm{int}((\langle f_1, \mu\rangle, \ldots, \langle f_m, \mu\rangle) : \mu \in \mathcal{M}(\mathbf{K})_+) \tag{3.3}$$

then the set of optimal solutions of (3.2) is nonempty and bounded.

As discussed in Lasserre [Las08], it is customary in the literature to assume that condition (3.3) holds, but in practice it may be a non-trivial task to check whether it does. We do stress, however, that condition (3.3) does hold for the applications discussed in Section 1.1.3 for $\mathbf{K}$ the simplex or the sphere.

Another result worth mentioning is that if the GMP (3.1) has an optimal solution, it has one which is finite atomic.

Theorem 3.3. *(Tchakaloff's theorem, cf. [Tch57]) If the GMP (3.1) has an optimal solution, then it has one which is finite atomic with at most m atoms, i.e., of the form* $\mu^* = \sum_{\ell=1}^{m} \omega_\ell \delta_{\mathbf{x}^{(\ell)}}$ *where* $\omega_\ell \geq 0, \mathbf{x}^{(\ell)} \in \mathbf{K}$ *and* $\delta_{\mathbf{x}^{(\ell)}}$ *denotes the Dirac measure supported at* $\mathbf{x}^{(\ell)}$ *for* $\ell \in [m]$.

3.2 A linear relaxation hierarchy over the simplex

In the remainder of the chapter we will only deal with the GMP (3.1) with polynomial data, i.e., we assume in what follows that all f_i's are polynomials ($i \in \{0, \ldots, m\}$).

The moment sequence $(y_\alpha)_{\alpha \in \mathbb{N}^n} \subset \mathbb{R}$ of a measure $\mu \in \mathcal{M}(\mathbf{K})$ is the infinite sequence given by

$$y_\alpha = \int_{\mathbf{K}} \mathbf{x}^\alpha \mathrm{d}\mu(\mathbf{x}) \quad \forall \alpha \in \mathbb{N}^n.$$

Let $L : \mathbb{R}[\mathbf{x}] \to \mathbb{R}$ be a linear operator

$$p(\mathbf{x}) = \sum_{\alpha \in \mathbb{N}^n} p_\alpha \mathbf{x}^\alpha \mapsto L(p) = \sum_{\alpha \in \mathbb{N}^n} p_\alpha y_\alpha$$

that maps monomials to their respective moments. Thus, to an optimal solution μ^* of a GMP of the form (3.1) there is an associated linear functional L^* such that $L^*(f_0) = \text{val}$ and $L^*(f_i) = b_i$ for all $i \in [m]$ as well as $L^*(1) \leq 1$. The idea of the relaxation we are about to introduce is to approximate the optimal solution by a sequence (hierarchy) of linear functionals $L^{(r)}$ that depend on $r = 1, 2, \ldots$. Let $\mathbf{K} = \Delta_{n-1}$. For $i = 0, 1, \ldots, m$ let w.l.o.g. f_i be a real homogeneous polynomial of degree d and let $r \geq d$. Let $L^{(r)}$ be the optimal solution of the following RLT-type relaxation of (3.1):

$$\begin{aligned} \underline{f}_{\text{LP}}^{(r)} = \min_{\substack{L:\mathbb{R}[\mathbf{x}]_r \to \mathbb{R} \\ L \text{ linear}}} \quad & L(f_0) \\ \text{s.t.} \quad & L(f_i) = b_i \quad \forall\, i \in [m] \\ & L(1) \leq 1 \\ & L(\mathbf{x}^\alpha) \geq 0 \quad \forall\, |\alpha| \leq r \\ & L(\mathbf{x}^\alpha) = L\left(\mathbf{x}^\alpha \sum_{i=1}^n x_i\right) \quad \forall\, |\alpha| \leq r-1. \end{aligned} \tag{3.4}$$

Note that when considering the case where $\mathbf{K} = \Delta_{n-1}$, we may, without loss of generality, assume the f_i to be homogeneous of the same degree for all $i = 0, 1, \ldots, m$. Indeed, let $f(\mathbf{x}) = \sum_{j=0}^d f_j(\mathbf{x})$, where $\deg(f_j) = j$. Then, $g(\mathbf{x}) := \sum_{j=0}^d f_j(\mathbf{x}) \left(\sum_{i=1}^n x_i\right)^{d-j}$ is homogeneous of degree d and $f(\mathbf{x}) = g(\mathbf{x})$ for all $\mathbf{x} \in \Delta_{n-1}$. Every feasible solution μ' to (3.1) provides an upper bound for (3.4) by setting $L(\mathbf{x}^\alpha) = \langle \mathbf{x}^\alpha, \mu' \rangle$. Hence, $\underline{f}_{\text{LP}}^{(r)} \leq \text{val}$. The third constraint is reflecting the necessary condition for a positive measure μ over the simplex:

$$\langle \mathbf{x}^\alpha, \mu \rangle = \int_{\Delta_{n-1}} \mathbf{x}^\alpha \mathrm{d}\mu \geq 0 \quad \forall \alpha \in \mathbb{N}^n.$$

The last constraint in (3.4) arises from the fact that

$$L(p) = L(q) \text{ if } p(\mathbf{x}) = q(\mathbf{x}) \quad \forall \mathbf{x} \in \Delta_{n-1}.$$

Equivalently, defining the ideal $\mathcal{I} = \{p\left(1 - \sum_{i=1}^n\right) : p \in \mathbb{R}[\mathbf{x}]\}$ we require

$$L(p) = L(q) \Leftrightarrow p = q \mod \mathcal{I},$$

where $p = q \mod \mathcal{I}$ means $p(\mathbf{x}) = q(\mathbf{x}) + (1 - \sum_{i=1}^{n} x_i)h(\mathbf{x})$ for some $h \in \mathbb{R}[\mathbf{x}]$.

Formulation (3.4) is closely related to the RLT approach by Sherali and Adams [AS10], that was originally introduced for 0-1 mixed integer linear programming problems and subsequently extended for more general problems. In fact, for the special case of polynomial optimization, problem (3.4) is essentially a Sherali-Adams RLT approach. To see this, note that our linearization operator L corresponds to the approximation $L(\mathbf{x}^\alpha) \approx \langle \mathbf{x}^\alpha, \mu^* \rangle$, where μ^* again denotes an optimal solution to the GMP (3.1). For the special case of polynomial optimization, we may assume that μ^* is a Dirac delta centered at an optimal solution, say $\mathbf{x}^*$. In this case, $L(\mathbf{x}^\alpha) \approx \langle \mathbf{x}^\alpha, \mu^* \rangle = \mathbf{x}^{*\alpha}$, i.e., L corresponds to the type of linearization operator introduced by Sherali and Adams [AS10].

We now state two lemmas that will come in handy in our later analysis.

Lemma 3.4. *Let $r, k \in \mathbb{N}$ with $k \leq r$ and let L be a feasible solution to the linear relaxation (3.4) for some $f_0, f_1, \ldots, f_m$. Then for all $\mathbf{x}^\gamma$ with $\gamma \in \mathbb{N}^n$ and $|\gamma| \leq r - k$ we have*

$$L(\mathbf{x}^\gamma) = L\left(\mathbf{x}^\gamma \left(\sum_{i=1}^{n} x_i\right)^k\right).$$

Proof. The last equality constraint in the relaxation forces

$$L(\mathbf{x}^\alpha) = L\left(\mathbf{x}^\alpha \sum_{i=1}^{n} x_i\right) \quad \forall \ |\alpha| \leq r - 1.$$

Therefore, noting that $\mathbf{x}^{e_j} = x_j$ we have

$$\begin{aligned}
L(\mathbf{x}^\beta \mathbf{x}^{e_j}) = L\left(\mathbf{x}^\beta \mathbf{x}^{e_j} \sum_{i=1}^{n} x_i\right) \quad \forall \ |\beta| \leq r - 2 \\
\Rightarrow \sum_{j=1}^{n} L(\mathbf{x}^\beta \mathbf{x}^{e_j}) = \sum_{j=1}^{n} L\left(\mathbf{x}^\beta \mathbf{x}^{e_j} \sum_{i=1}^{n} x_i\right) \quad \forall \ |\beta| \leq r - 2 \\
\Leftrightarrow L\left(\mathbf{x}^\beta \sum_{j=1}^{n} \mathbf{x}^{e_j}\right) = L\left(\mathbf{x}^\beta \sum_{j=1}^{n} \mathbf{x}^{e_j} \sum_{i=1}^{n} x_i\right) \quad \forall \ |\beta| \leq r - 2 \\
= L\left(\mathbf{x}^\beta \left(\sum_{i=1}^{n} x_i\right)^2\right) \quad \forall \ |\beta| \leq r - 2.
\end{aligned}$$

Hence,

$$L(\mathbf{x}^\beta) = L\left(\mathbf{x}^\beta \sum_{i=1}^{n} x_i\right) = L\left(\mathbf{x}^\beta \left(\sum_{i=1}^{n} x_i\right)^2\right) \quad \forall \quad |\beta| \leq r - 2.$$

Repeating this procedure leads us to the desired outcome. □

Lemma 3.5. *Consider the GMP given in (3.1) and let* $(y,t) \in \mathbb{R}^m \times \mathbb{R}_+$ *be feasible solution for the dual* (3.2)*. Then the pair* (y,t) *is dual optimal only if*

$$0 = \min_{\mathbf{x}\in\mathbf{K}} \left(f_0(\mathbf{x}) - \sum_{i=1}^{m} y_i f_i(\mathbf{x}) + t \right).$$

Proof. The minimization problem

$$\min_{\mathbf{x}\in\mathbf{K}} \left(f_0(\mathbf{x}) - \sum_{i=1}^{m} y_i f_i(\mathbf{x}) + t \right)$$

is equivalent to

$$\inf_{\mu\in\mathcal{M}(\mathbf{K})_+} \left\{ \int_{\mathbf{K}} \left(f_0(\mathbf{x}) - \sum_{i=1}^{m} y_i f_i(\mathbf{x}) + t \right) \mathrm{d}\mu(\mathbf{x}) : \int_{\mathbf{K}} \mathrm{d}\mu = 1 \right\}. \tag{3.5}$$

By Theorem 3.1 there is no duality gap and there exists a primal optimal solution μ^* to the GMP (3.1). Set $\nu = \mu^*/\mu^*(\mathbf{K})$. Hence, ν is a probability measure and therefore a feasible solution to (3.5). We deduce

$$\begin{aligned}
0 &\le \min_{\mathbf{x}\in\mathbf{K}} f_0(\mathbf{x}) - \sum_{i=1}^{m} y_i f_i(\mathbf{x}) + t \\
&\le \int_{\mathbf{K}} f_0(\mathbf{x}) - \sum_{i=1}^{m} y_i f_i(\mathbf{x}) + t \,\mathrm{d}\nu(\mathbf{x}) \\
&= \frac{1}{\mu^*(\mathbf{K})} \left(\int_{\mathbf{K}} f_0(\mathbf{x}) \mathrm{d}\mu^*(\mathbf{x}) - \sum_{i=1}^{m} y_i \int_{\mathbf{K}} f_i(\mathbf{x}) \mathrm{d}\mu^*(\mathbf{x}) + t \int_{\mathbf{K}} \mathrm{d}\mu^*(\mathbf{x}) \right) \\
&= \frac{1}{\mu^*(\mathbf{K})} \left(\mathrm{val} - y^\top b + t\mu^*(\mathbf{K}) \right) \\
&\le \frac{1}{\mu^*(\mathbf{K})} \left(\mathrm{val} - y^\top b + t \right) = \frac{1}{\mu^*(\mathbf{K})} \left(\mathrm{val} - \mathrm{val}' \right) = 0,
\end{aligned}$$

where the first inequality follows from the definition of the dual (3.2) of the GMP, the third inequality from the fact that $\mu^*(\mathbf{K}) \le 1$ and the last equality from strong duality. □

3.2.1 Convergence analysis

The following theorem is a refinement of a result by Powers and Reznick [PR01], obtained by de Klerk, Laurent and Parrilo [dKLP06, Theorem 1.1]. It is a quantitative version of Pólya's Positivstellensatz (see, e.g., [Rez96] for a survey), and it will be crucial in our analysis of the simplex case.

Theorem 3.6. *Suppose* $f \in \mathbb{R}[\mathbf{x}]$ *is a homogeneous polynomial of degree* d *of the form* $f(\mathbf{x}) = \sum_{|\alpha|=d} f_\alpha \mathbf{x}^\alpha$. *Assume* $\varepsilon = \min_{\Delta_{n-1}} f(\mathbf{x}) > 0$ *and define*

$$B(f) = \max_{|\alpha|=d} \frac{\alpha_1! \dots \alpha_n!}{d!} f_\alpha. \tag{3.6}$$

Then the polynomial $(x_1 + \cdots + x_n)^k f(\mathbf{x})$ *has only positive coefficients if*

$$k > \frac{d(d-1)}{2} \frac{B(f)}{\varepsilon} - d. \tag{3.7}$$

We continue by stating and proving one of the main results of this chapter.

Theorem 3.7. *Let* val *be the optimal value of the GMP (3.1) for input data* $\mathbf{K} = \Delta_{n-1}, f_0, f_1, \dots, f_m \in \mathbb{R}[\mathbf{x}]$ *homogeneous of degree* d *and* $b_1, \dots, b_m \in \mathbb{R}$. *Assume there exists a dual optimal solution* $(\bar{y}, t)$ *for* (3.2). *Then, setting* $y_0 = 1$ *and* $y_i = -\bar{y}_i$ *for* $i \in [m]$ *we have*

$$0 \leq \text{val} - \underline{f}_{\text{LP}}^{(r)} \leq \frac{\left(\sum_{i=0}^m B(y_i f_i) + t\right) d(d-1)}{2(r-1) - d(d-1)}, \tag{3.8}$$

for $B(\cdot)$ *as in (3.6) and* $r > d(d-1)/2 + 1$.

Proof. By Theorem 3.1 there is no duality gap. Let $r > d(d-1)/2 + 1$ and let $L^{(r)}$ be an optimal solution to (3.4) whose existence is ensured by Theorem 1.5. Fix some $\varepsilon > 0$. Then,

$$
\begin{aligned}
0 \le \operatorname{val} - \underline{f}_{\mathrm{LP}}^{(r)} &= \operatorname{val} - L^{(r)}\left(\sum_{i=1}^{m} \bar{y}_i f_i - t + f_0 - \sum_{i=1}^{m} \bar{y}_i f_i + t\right) \\
&= \operatorname{val} - \sum_{i=1}^{m} \bar{y}_i L^{(r)}(f_i) + t L^{(r)}(1) - L^{(r)}\left(f_0 - \sum_{i=1}^{m} \bar{y}_i f_i + t\right) \\
&\le \operatorname{val} - \sum_{i=1}^{m} \bar{y}_i b_i + t - L^{(r)}\left(f_0 - \sum_{i=1}^{m} \bar{y}_i f_i + t\right) \\
&= -L^{(r)}\left(f_0 - \sum_{i=1}^{m} \bar{y}_i f_i + t\right) \\
&= -L^{(r)}\left(f_0 - \sum_{i=1}^{m} \bar{y}_i f_i + t + \varepsilon\right) + \varepsilon L^{(r)}(1) \\
&\le -L^{(r)}\left(f_0 - \sum_{i=1}^{m} \bar{y}_i f_i + t + \varepsilon\right) + \varepsilon,
\end{aligned}
$$

where both inequalities follow from the fact that $L^{(r)}(1) \le 1$. Set $f_{m+1} = \left(\sum_{i=1}^{n} x_i\right)^d$ and $\bar{y}_{m+1} = -t$. By Lemma 3.5 we have $\min_{\mathbf{x}\in\Delta_{n-1}} f_0(\mathbf{x}) - \sum_{i=1}^{m+1} \bar{y}_i f_i(\mathbf{x}) + \varepsilon = \varepsilon$. The goal is now to show that $-L^{(r)}\left(f_0 - \sum_{i=1}^{m+1} \bar{y}_i f_i + \varepsilon\right) \le 0$ if

$$
r \ge \left\lceil \frac{d(d-1)}{2} \frac{\sum_{i=0}^{m+1} B(y_i f_i) + \varepsilon}{\varepsilon} \right\rceil.
$$

Define

$$
f := f_0 - \sum_{i=1}^{m+1} \bar{y}_i f_i + \varepsilon \left(\sum_{i=1}^{n} x_i\right)^d,
$$

which is homogeneous and its minimum over the simplex is ε. We continue to show that $L^{(r)}(f) \ge 0$ for the appropriate choice of r and then bound r in terms of ε. By Theorem 3.6 for k as in (3.7) we have

$$
f(\mathbf{x}) \left(\sum_{i=1}^{n} x_i\right)^k = \sum_{\beta \in \mathbb{N}^n_{d+k}} c_\beta x^\beta \tag{3.9}
$$

with $c_\beta > 0$ for all $\beta \in \mathbb{N}^n_{d+k}$. To determine the smallest integer k for which the theorem holds we will first bound $B(f)$. For this, set $y_0 = 1$ and $y_i = -\bar{y}_i$. We

may rewrite f as

$$\begin{aligned} f &= \sum_{i=0}^{m+1} y_i f_i + \varepsilon \left(\sum_{i=1}^{n} x_i \right)^d \\ &= \sum_{i=0}^{m+1} y_i f_i + \varepsilon \left(\sum_{|\alpha|=d} \binom{d}{\alpha_1 \dots \alpha_n} x^\alpha \right) \\ &= \sum_{|\alpha|=d} \left(\sum_{i=0}^{m+1} y_i f_{i,\alpha} + \varepsilon \binom{d}{\alpha_1 \dots \alpha_n} \right) x^\alpha. \end{aligned}$$

Then,

$$\begin{aligned} B(f) &= \max_\alpha \left[\left(\sum_{i=0}^{m+1} y_i f_{i,\alpha} + \frac{d!}{\alpha_1! \dots \alpha_n!} \varepsilon \right) \frac{\alpha_1! \dots \alpha_n!}{d!} \right] \\ &= \left(\max_\alpha \left(\sum_{i=0}^{m+1} y_i f_{i,\alpha} \right) \frac{\alpha_1! \dots \alpha_n!}{d!} \right) + \varepsilon \\ &\le \sum_{i=0}^{m+1} \left(\max_\alpha y_i f_{i,\alpha} \frac{\alpha_1! \dots \alpha_n!}{d!} \right) + \varepsilon \\ &= \sum_{i=0}^{m+1} B(y_i f_i) + \varepsilon. \end{aligned}$$

With this bound on $B(f)$ we find that if r is large enough, i.e.,

$$r \ge \left\lceil \frac{d(d-1)}{2} \frac{\sum_{i=0}^{m+1} B(y_i f_i) + \varepsilon}{\varepsilon} \right\rceil \ge \left\lceil \frac{d(d-1)}{2} \frac{B(f)}{\varepsilon} \right\rceil,$$

it follows from Lemma 3.4 that

$$\begin{aligned} -L^{(r)} \left(f_0 - \sum_{i=1}^{m+1} \bar{y}_i f_i + \varepsilon \right) + \varepsilon &= \varepsilon - L^{(r)}(f) \\ &= \varepsilon - L^{(r)} \left(f \left(\sum_{i=1}^{n} x_i \right)^k \right) \\ &= \varepsilon - L^{(r)} \left(\sum_{\beta \in \mathbb{N}^n_{k+d}} c_\beta x^\beta \right) \le \varepsilon, \end{aligned}$$

where the last equality follows from (3.9) and the inequality follows from the fact that $L^{(r)}(\mathbf{x}^\alpha) \geq 0$ for all $|\alpha| \leq r$ and $c_\beta \geq 0$ for all $|\beta| \leq k+d$. To find a bound on r in terms of ε we set

$$r = \left\lceil \frac{d(d-1)}{2} \frac{\sum_{i=0}^{m+1} B(y_i f_i) + \varepsilon}{\varepsilon} \right\rceil.$$

Then, one may bound r as follows

$$\begin{aligned} r - 1 &\leq \frac{d(d-1)}{2} \left(\frac{\sum_{i=0}^{m+1} B(y_i f_i)}{\varepsilon} + 1 \right) \\ \Leftrightarrow \varepsilon &\leq \frac{\sum_{i=0}^{m+1} B(y_i f_i) d(d-1)}{2(r-1) - d(d-1)}, \end{aligned}$$

concluding the proof. □

Remark 3.8. *The bound we give in Theorem 3.7 depends on the dual optimal solution $(\bar{y}, t)$. We cannot bound the dual variables in terms of the problem data a priori in general, as they may become arbitrarily large. There are, however, cases in which one can bound the variables in terms of the problem data. An example of this case can be found in Section 3.5.*

3.3 Moment-SOS hierarchy over the sphere

We now consider the GMP (3.1) over the sphere, i.e., the case $\mathbf{K} = \mathcal{S}^{n-1}$. Additionally, we assume the $f_0, f_1, \ldots, f_m$ in (3.1) are homogeneous polynomials of even degree $2d$.

The moment-SOS hierarchy [Las01] of semidefinite relaxations of the GMP (3.1) over the sphere is given by

$$\begin{aligned} \underline{f}_{\text{SDP}}^{(2r)} = \min_{\substack{L:\mathbb{R}[\mathbf{x}]_{2r} \to \mathbb{R} \\ L \text{ linear}}} \; & L(f_0) \\ \text{s.t.} \quad & L(f_i) = b_i \quad \forall i \in [m] \\ & L(1) \leq 1 \\ & L\left([\mathbf{x}]_r [\mathbf{x}]_r^\top\right) \succeq 0 \\ & L(\mathbf{x}^\alpha) = L\left(\mathbf{x}^\alpha \|\mathbf{x}\|_2^2\right) \quad \forall \quad |\alpha| \leq 2r-2, \end{aligned} \tag{3.10}$$

where the L operator is now applied entry-wise to matrix-valued functions, where needed, and the optimal solution is denoted by $L^{(2r)}$.

The following lemma enables us to use a quantitative Positivstellensatz by Fang and Fawzi [FF21] for positive polynomials on the sphere, to obtain a rate of convergence of the Lasserre hierarchy. It is a basic but crucial result and certainly known to be true, however we did not find a suitable reference. Hence, we give a short proof for completeness.

Lemma 3.9. *Let $L : \mathbb{R}[\mathbf{x}]_{2k} \to \mathbb{R}$ be a linear operator and suppose $L\left([\mathbf{x}]_k[\mathbf{x}]_k^\top\right) \succeq 0$, where the operator is applied entrywise to the matrix $[\mathbf{x}]_k[\mathbf{x}]_k^\top$. Then, $L(\sigma) \geq 0$ for all $\sigma \in \Sigma[\mathbf{x}]_k$.*

Proof. Let $\sigma \in \Sigma[\mathbf{x}]_k$ be a sum of squares of degree $2k$. Then there exists $A \succeq 0$ such that $\sigma = [\mathbf{x}]_k^\top A[\mathbf{x}]_k$. Let $\langle \cdot, \cdot \rangle$ denote the trace inner product. We have

$$L(\sigma) = L\left([\mathbf{x}]_k^\top A[\mathbf{x}]_k\right) = \sum_{i,j} A_{i,j} L\left(([\mathbf{x}]_k)_i([\mathbf{x}]_k)_j\right) = \langle A, L\left([\mathbf{x}]_k[\mathbf{x}]_k^\top\right)\rangle \geq 0,$$

since both A and $L\left([\mathbf{x}]_k[\mathbf{x}]_k^\top\right)$ are psd. □

The quantitative Positivstellensatz by Fang and Fawzi [FF21] is as follows.

Theorem 3.10. *[FF21, Theorem 3.8] Assume f is a homogeneous polynomial of degree $2d$ such that $0 \leq f(\mathbf{x}) \leq 1$ for all $\mathbf{x} \in \mathcal{S}^{n-1}$ and $d \leq n$. There are constants C_d, C_d' that depend only on d such that if $r \geq C_d n$ then*

$$f + C_d'(d/r)^2 = \sigma(\mathbf{x}) + (1 - \|\mathbf{x}\|_2^2)h(\mathbf{x})$$

for $\sigma(\mathbf{x}) \in \Sigma[\mathbf{x}]_r$ and $h \in \mathbb{R}[\mathbf{x}]_{2r-2}$.

We may now use the theorem by Fang and Fawzi [FF21] and Lemma 3.9 to derive a rate of convergence for the Lasserre hierarchy [Las09] of the GMP on the sphere as follows.

Theorem 3.11. *Let* val *be the optimal value of the GMP (3.1) for input data $\mathbf{K} = \mathcal{S}^{n-1}, f_0, f_1, \ldots, f_m \in \mathbb{R}[\mathbf{x}]$ homogeneous of even degree $2d$, $b_1, \ldots, b_m \in \mathbb{R}$ and $d \leq n$. Let $(\bar{y}, t)$ be a dual optimal solution and let $f_{m+1}(\mathbf{x}) := 1$ for every $\mathbf{x} \in \mathcal{S}^{n-1}$, set $\bar{y}_{m+1} = -t$ and set $y_0 = 1$ and $y = -\bar{y}$. Further, let $f_{\max}^{i,y_i} = \max_{\mathbf{x} \in \mathcal{S}^{n-1}} y_i f_i(\mathbf{x})$. There exist constants C_d, C_d', only dependent on d, such that if $r \geq C_d n$ we have*

$$0 \leq \text{val} - \underline{f}_{\text{SDP}}^{(2r)} \leq \frac{C_d' d^2 \sum_{i=0}^{m+1} f_{\max}^{i,y_i}}{r^2}.$$

Proof. As in the proof of Theorem 3.7, Theorem 3.1 gives us strong duality. Let $r \geq C_d n$ and let $L^{(2r)}$ be an optimal solution to (3.10). Then by the same reasoning as in Theorem 3.7,

$$0 \leq \text{val} - \underline{f}_{\text{SDP}}^{(2r)} \leq -L^{(2r)}\left(f_0 - \sum_{i=1}^{m+1} \bar{y}_i f_i\right).$$

Set $f := f_0 - \sum_{i=1}^{m+1} y_i f_i$ and $f_{\max} = \max_{\mathbf{x}\in\mathcal{S}^{n-1}} f(\mathbf{x})$. Then $\tilde{f} = f/f_{\max}$ satisfies $\tilde{f}_{\max} = 1$ and by Lemma 3.5 we have $\tilde{f}_{\min} = 0$. We find for any $\delta \geq 0$

$$\begin{aligned} -L^{(2r)}\left(f_0 - \sum_{i=1}^{m+1} \bar{y}_i f_i\right) &= -f_{\max} L^{(2r)}\left(\tilde{f}\right) \\ &\leq -f_{\max} L^{(2r)}\left(\tilde{f} + \delta\right) + \delta f_{\max}. \end{aligned}$$

Choosing $\delta = C_d' \frac{d^2}{r^2}$ and applying Theorem 3.10 we see that $\tilde{f} + \delta = \sigma + (1 - \|\mathbf{x}\|_2^2)h$ for $\sigma \in \Sigma[\mathbf{x}]_r$ and $h \in \mathbb{R}[\mathbf{x}]_{2r-2}$.

Thus, since $L^{(2r)}(\mathbf{x}^\alpha) = L^{(2r)}(\mathbf{x}^\alpha \|\mathbf{x}\|_2^2)$ we have

$$\begin{aligned} -f_{\max} L^{(2r)}\left(\tilde{f} + \frac{C_d' d^2}{r^2}\right) + \frac{C_d' d^2}{r^2} f_{\max} &= -f_{\max} L^{(2r)}\left(\sigma + (1 - \|\mathbf{x}\|_2^2)h\right) + \frac{C_d' d^2}{r^2} f_{\max} \\ &= -f_{\max} L^{(2r)}(\sigma) + \frac{C_d' d^2}{r^2} f_{\max} \\ &\leq \frac{C_d' d^2}{r^2} f_{\max}, \end{aligned}$$

where the last inequality follows from Lemma 3.9. Noting that

$$f_{\max} = \max_{\mathbf{x}\in\mathcal{S}^{n-1}} \left(f_0(\mathbf{x}) - \sum_{i=1}^{m+1} \bar{y}_i f_i(\mathbf{x})\right) \leq \sum_{i=0}^{m+1} \max_{\mathbf{x}\in\mathcal{S}^{n-1}} y_i f_i(\mathbf{x}) = \sum_{i=0}^{m+1} f_{\max}^{i,y_i}$$

we arrive at the result. □

3.4 Limiting behavior of the hierarchies of linear operators

The purpose of this section is to show that the limit functionals of the introduced hierarchies correspond to measures, in the sense that they are the Riesz functional of an optimal solution of the corresponding GMP. In the following we will define

the limit of the optimal solutions $L^{(r)}$ of the introduced hierarchies in a meaningful way and prove that the corresponding moment sequences have a representing measure.

3.4.1 The simplex case

Consider the case when $\mathbf{K} = \Delta_{n-1}$. When looking at the linear operators in the relaxation hierarchies (3.4) one would expect that in the limit, i.e., for $r \to \infty$, the operators $L^{(r)}(\cdot)$ behave like $\langle \cdot, \mu \rangle$ for some positive measure μ. In the rest of this section we prove that this is in fact the case, and we will define the limit in a meaningful way. Consider again the ideal $\mathcal{I} = \{p\left(1 - \sum_{i=1}^n x_i\right) : p \in \mathbb{R}[\mathbf{x}]\}$ and let $\bar{L} : \mathbb{R}[\mathbf{x}]/\mathcal{I} \to \mathbb{R}$ be a linear operator such that

1. $\bar{L}(\mathbf{x}^\alpha) \geq 0$ for all $\alpha \in \mathbb{N}^n$
2. $\bar{L}(1) \leq 1$

and let

$$\mathcal{L} = \{\bar{L} : \mathbb{R}[\mathbf{x}]/\mathcal{I} \to \mathbb{R} : \bar{L} \text{ fulfills conditions 1. and 2.}\}$$

be the class of all linear operators that satisfy the conditions above. Note that for every $\bar{L} \in \mathcal{L}$ the relation

$$\bar{L}\left(\left(1 - \sum_{i=1}^n x_i\right)\mathbf{x}^\alpha\right) = 0 \text{ for all } \alpha \in \mathbb{N}^n$$

trivially holds. Defining $\|f\| = \|f\|_{\infty,\Delta_{n-1}} := \sup_{\mathbf{x}\in\Delta_{n-1}} |f(\mathbf{x})|$ we obtain a normed vector space given by $(\mathbb{R}[\mathbf{x}]/\mathcal{I}, \|\cdot\|)$.

Theorem 3.12. *(see, e.g. [Meg98, Theorem 1.4.2]) Suppose $F : X \to Y$ is a linear operator between two normed vector spaces $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$, then the following are equivalent*

1. *F is continuous*
2. *$\|Fx\|_Y \leq M\|x\|_X$ for some $M \in \mathbb{R}$.*

Using Theorem 3.12 we can prove that the operators we consider are continuous in the limit.

Lemma 3.13. *Every $\bar{L} \in \mathcal{L}$ is continuous.*

Proof. Invoking Theorem 3.12 for $M = 1$ it suffices to show that every $\bar{L} \in \mathcal{L}$ satisfies

$$|\bar{L}(f)| \leq \|f\| = \sup_{\mathbf{x}\in\Delta_{n-1}} |f(\mathbf{x})|$$

for all $f \in \mathbb{R}[\mathbf{x}]/\mathcal{I}$. We begin by showing that for any $f \in \mathbb{R}[\mathbf{x}]/\mathcal{I}$ and $\bar{L} \in \mathcal{L}$ we have $\bar{L}(f) \geq -\|f\|$. Consider the GMP given by

$$\begin{aligned} \text{val} = \inf_{\mu\in\mathcal{M}(\Delta_{n-1})_+} & \int_{\Delta_{n-1}} f(\mathbf{x})\mathrm{d}\mu(\mathbf{x}) \\ \text{s.t.} & \int_{\Delta_{n-1}} \mathrm{d}\mu(\mathbf{x}) \leq 1. \end{aligned} \tag{3.11}$$

Set

$$f_{\min} = \min_{x\in\Delta_{n-1}} f(\mathbf{x}) \geq -\|f\| \text{ and } f_{\max} = \max_{x\in\Delta_{n-1}} f(\mathbf{x}) \leq \|f\|.$$

Let μ^* be an optimal solution to (3.11) whose existence is ensured by Theorem 1.5. Note that if $f_{\min} \geq 0$ then

$$\text{val} = \int_{\Delta_{n-1}} f(\mathbf{x})\mathrm{d}\mu^*(\mathbf{x}) \geq f_{\min} \int_{\Delta_{n-1}} \mathrm{d}\mu^*(\mathbf{x}) \geq 0 \geq -\|f\|.$$

If $f_{\min} < 0$ then

$$\text{val} = \int_{\Delta_{n-1}} f(\mathbf{x})\mathrm{d}\mu^*(\mathbf{x}) \geq |f_{\min}| \left(- \int_{\Delta_{n-1}} \mathrm{d}\mu^*(\mathbf{x}) \right) \geq f_{\min} \geq -\|f\|,$$

where we used $-\int_{\Delta_{n-1}} \mathrm{d}\mu^*(\mathbf{x}) \geq -1$. It follows that $\text{val} \geq -\|f\|$. Every $\bar{L} \in \mathcal{L}$ is feasible for the LP relaxation (3.4) of (3.11). Hence, $\bar{L}(f) \geq \underline{f}_{\text{LP}}^{(r)}$ for all $r \geq \deg(f)$. Thus, $\bar{L}(f) \geq \lim_{r\to\infty} \underline{f}_{\text{LP}}^{(r)} = \text{val} \geq -\|f\|$ by Theorem 3.7. To show that $\bar{L}(f) \leq \|f\|$ consider the following GMP

$$\begin{aligned} \text{val} = \sup_{\mu\in\mathcal{M}(\Delta_{n-1})_+} & \int_{\Delta_{n-1}} f(\mathbf{x})\mathrm{d}\mu(\mathbf{x}) \\ \text{s.t.} & \int_{\Delta_{n-1}} \mathrm{d}\mu(\mathbf{x}) \leq 1. \end{aligned} \tag{3.12}$$

Let again μ^* be an optimal solution to (3.12). Then, if $f_{\max} \geq 0$ we find

$$\text{val} = \int_{\Delta_{n-1}} f(\mathbf{x})\mathrm{d}\mu^*(\mathbf{x}) \leq f_{\max} \int_{\Delta_{n-1}} \mathrm{d}\mu^*(\mathbf{x}) \leq f_{\max} \leq \|f\|.$$

If $f_{\max} < 0$ then

$$\text{val} = \int_{\Delta_{n-1}} f(\mathbf{x}) \mathrm{d}\mu^*(\mathbf{x}) \leq f_{\max} \int_{\Delta_{n-1}} \mathrm{d}\mu^*(\mathbf{x}) \leq 0 \leq \|f\|.$$

Replacing the min operator in (3.4) by the max operator leads to a relaxation of (3.12) whose optimal values $\overline{f}_{\text{LP}}^{(r)}$ form a non-increasing sequence lower bounded by val and $\lim_{r\to\infty} \overline{f}_{\text{LP}}^{(r)} = \text{val}$. Moreover, any $\bar{L} \in \mathcal{L}$ is feasible for that relaxation, hence $\bar{L}(f) \leq \overline{f}_{\text{LP}}^{(r)}$ for all $r \geq \deg(f)$. Thus, $\bar{L}(f) \leq \lim_{r\to\infty} \overline{f}_{\text{LP}}^{(r)} = \text{val} \leq \|f\|$. Therefore,

$$|\bar{L}(f)| \leq \|f\|,$$

for all $f \in \mathbb{R}[\mathbf{x}]/\mathcal{I}, \bar{L} \in \mathcal{L}$. □

The set $\mathbb{R}[\mathbf{x}]/\mathcal{I}$ is dense in $\mathcal{C}(\Delta_{n-1})$. This means we can employ the following theorem in the next step.

Theorem 3.14. *(see, e.g. [Meg98, Theorem 1.9.1]) Suppose that M is a dense subspace of a normed space X, that Y is a Banach space, and that $T_0 : M \to Y$ is a bounded linear operator. Then there is a unique continuous function $T : X \to Y$ that agrees with T_0 on M. This function T, called a continuous linear extension of T_0, is a bounded linear operator and $\|T\| = \|T_0\|$.*

Now let

$$\mathcal{T} = \left\{T : \mathcal{C}(\Delta_{n-1}) \to \mathbb{R} : T \text{ is the continuous linear extension of some } \bar{L} \in \mathcal{L}\right\}.$$

Proposition 3.15. *Let $T \in \mathcal{T}$ and $f \in \mathcal{C}(\Delta_{n-1})$. Then*

$$T(f) = \int_{\Delta_{n-1}} f(\mathbf{x}) \mathrm{d}\mu(\mathbf{x})$$

for some positive measure μ supported on Δ_{n-1}, satisfying $\mu(\Delta_{n-1}) \leq 1$.

Proof. It is sufficient to show $T(f) \geq 0$ for all $f \in \mathcal{C}(\Delta_{n-1})_+ = \{f \in \mathcal{C}(\Delta_{n-1}) : f(\mathbf{x}) \geq 0 \ \forall \mathbf{x} \in \Delta_{n-1}\}$. To see this, note that the space $\mathcal{C}(\Delta_{n-1})$ can be ordered by the convex cone $\mathcal{C}(\Delta_{n-1})_+$. Now $T(f) \geq 0$ for all $f \in \mathcal{C}(\Delta_{n-1})_+$ implies that $T \in (\mathcal{C}(\Delta_{n-1})_+)^*$, i.e., the dual cone of $\mathcal{C}(\Delta_{n-1})_+$ which is known to be the set of finite Borel measures on Δ_{n-1}. Let f be a homogeneous continuous function that

is non-negative on the simplex and consider its Bernstein approximation of order r given by

$$\mathcal{B}_f^r(\mathbf{x}) = \sum_{\substack{\alpha\in\mathbb{N}_r^n \\ |\alpha|=r}} f\left(\frac{\alpha}{r}\right)\binom{r}{\alpha}\mathbf{x}^\alpha.$$

The approximation converges uniformly to f as $r\to\infty$ since f is continuous. Using Lemma 3.13 we see

$$\begin{aligned} T(f) &= T\left(\lim_{r\to\infty}\mathcal{B}_f^r\right) \\ &\overset{T \text{ cont.}}{=} \lim_{r\to\infty} T\left(\mathcal{B}_f^r\right) \\ &= \lim_{r\to\infty}\sum_{\substack{\alpha\in\mathbb{N}_r^n \\ |\alpha|=r}} \underbrace{f\left(\frac{\alpha_1}{r},\ldots,\frac{\alpha_n}{r}\right)}_{\geq 0}\underbrace{\binom{r}{\alpha}}_{\geq 0}\underbrace{T(\mathbf{x}^\alpha)}_{\geq 0} \geq 0. \end{aligned}$$

Hence, it follows that $T(f) = \langle f,\mu\rangle$ for some positive measure μ, such that $\mu(\Delta_{n-1}) \leq 1$. □

Remark 3.16. *By the proof given above, it becomes clear that the continuous linear extension can in fact be defined in terms of the limit of the Bernstein approximation, i.e., define* $T(f) := \lim_{r\to\infty}\bar{L}(\mathcal{B}_f^r)$ *for* $f\in\mathcal{C}(\Delta_{n-1})$ *and* $\bar{L}\in\mathcal{L}$.

3.4.2 The sphere case

For the sphere case, i.e., $\mathbf{K} = \mathcal{S}^{n-1}$ consider the following theorem.

Theorem 3.17. *(see, e.g. [Las09, Theorem 3.8] or [Sch91] for the original reference) Let* $\mathbf{y} = (y_\alpha)_{\alpha\in\mathbb{N}^n}\subset\mathbb{R}^\infty$ *be a given infinite real sequence,* $\bar{L}:\mathbb{R}[\mathbf{x}]\to\mathbb{R}$ *be the linear operator defined by*

$$p(\mathbf{x}) = \sum_{\alpha\in\mathbb{N}^n} p_\alpha\mathbf{x}^\alpha \mapsto \bar{L}(p) = \sum_{\alpha\in\mathbb{N}^n} p_\alpha y_\alpha,$$

and let $\mathbf{K} = \{\mathbf{x}\in\mathbb{R}^n : g_1(\mathbf{x})\geq 0,\ldots,g_m(\mathbf{x})\geq 0\}$ *be compact. The sequence* $\mathbf{y}$ *has a finite Borel representing measure with support contained in* $\mathbf{K}$ *if and only if*

$$\bar{L}(f^2 g_J)\geq 0\ \forall J\subseteq\{1,\ldots,m\} \text{ and } f\in\mathbb{R}[\mathbf{x}],$$

where $g_J(\mathbf{x}) = \prod_{j\in J} g_j(\mathbf{x})$.

Now, let $\bar{L}$ be a linear operator such that

1. $\bar{L}(1) \leq 1$
2. $\bar{L}([\mathbf{x}]_t[\mathbf{x}]_t^\top) \succeq 0 \ \forall t \in \mathbb{N}$
3. $\bar{L}(\mathbf{x}^\alpha) = \bar{L}(\mathbf{x}^\alpha \|\mathbf{x}\|_2^2) \ \forall \alpha \in \mathbb{N}^n$

and let $\mathcal{L}' = \{\bar{L} : \mathbb{R}[\mathbf{x}] \to \mathbb{R} : \bar{L} \text{ satisfies 1. — 3.}\}$. Recall that as a semialgebraic set the sphere can be written as $\mathcal{S}^{n-1} = \{\mathbf{x} \in \mathbb{R}^n : g_1(\mathbf{x}) := 1 - \|\mathbf{x}\|_2^2 \geq 0, g_2(\mathbf{x}) := \|\mathbf{x}\|_2^2 - 1 \geq 0\}$. Then for $\mathbf{K} = \mathcal{S}^{n-1}$ every $\bar{L} \in \mathcal{L}'$ satisfies all conditions of Theorem 3.17. To see this, note that the only possibilities for J are $\{\emptyset, \{1\}, \{2\}, \{1,2\}\}$. Because of condition 3 we have that $\bar{L}(\pm(1-\|\mathbf{x}\|_2^2)p) = 0$ for all $p \in \mathbb{R}[\mathbf{x}]$ covering all cases except $J = \emptyset$. For $J = \emptyset$ the condition reduces to $\bar{L}(p^2) \geq 0$ which holds for all $p \in \mathbb{R}[\mathbf{x}]$ because of Lemma 3.9. Hence, every $\bar{L} \in \mathcal{L}'$ has a representing measure whose support is contained in $\mathcal{S}^{n-1}$.

3.5 Concluding remarks

In this last section we conclude by outlining the connection of our results to previous work. We show that — in the special case of polynomial optimization on the simplex — our RLT hierarchy reduces to one studied earlier by Bomze and de Klerk [BdK01], and de Klerk, Laurent and Parrilo [dKLP06].

De Klerk, Laurent and Parrilo [dKLP06] introduced the following hierarchy for minimizing a homogeneous polynomial $p \in \mathbb{R}[\mathbf{x}]$ of degree d over the simplex.

$$p^{(r)} = \max \lambda \text{ s.t. the polynomial } \left(\sum_{i=1}^n x_i\right)^r \left(p(\mathbf{x}) - \lambda \left(\sum_{i=1}^n x_i\right)^d\right) \text{ has only non-neg. coefficients.} \tag{3.13}$$

It was proved that $\lim_{r\to\infty} p^{(r)} = p_{\min} = \min_{\mathbf{x}\in\Delta_{n-1}} p(\mathbf{x})$. The LP hierarchy introduced in Section 3.2 of this chapter is a generalization of the hierarchy (3.13), in the sense made precise in the following theorem.

Theorem 3.18. *For some homogeneous polynomial $p \in \mathbb{R}[\mathbf{x}]$ of degree d let $\underline{f}_{\text{LP}}^{(r+d)}$ be the solution to the LP relaxation of the problem*

$$\min_{\mathbf{x}\in\Delta_{n-1}} p(\mathbf{x}) = \text{val} = \inf_{\mu \in \mathcal{M}(\Delta_{n-1})_+} \left\{ \int_{\Delta_{n-1}} p(\mathbf{x}) \mathrm{d}\mu(\mathbf{x}) : \int_{\Delta_{n-1}} \mathrm{d}\mu(\mathbf{x}) = 1 \right\}$$

for some $r \in \mathbb{N}$. Then,

$$p^{(r)} = \underline{f}_{\text{LP}}^{(r+d)}.$$

Proof. " $\leq$ " : Let $\lambda^* = p^{(r)}$ be optimal for (3.13). Then the polynomial given by $\left(\sum_{i=1}^n x_i\right)^r \left(p(\mathbf{x}) - \lambda \left(\sum_{i=1}^n x_i\right)^d\right)$ has only negative coefficients, and we find

$$\begin{aligned} 0 &\leq L^{(r+d)}\left(\left(\sum_{i=1}^n x_i\right)^r \left(p(\mathbf{x}) - \lambda^* \left(\sum_{i=1}^n x_i\right)^d\right)\right) \\ &= L^{(r+d)}\left(\left(\sum_{i=1}^n x_i\right)^r p(\mathbf{x})\right) - \lambda^* L^{(r+d)}\left(\left(\sum_{i=1}^n x_i\right)^{r+d}\right) \\ &= \underline{f}_{\mathrm{LP}}^{(r+d)} - \lambda^* \end{aligned}$$

for $L^{(r+d)}$ being the optimal solution to the LP relaxation.
" $\geq$ " : For the multinomial coefficient

$$\binom{k}{\alpha} = \binom{k}{\alpha_1, \ldots, \alpha_n} = \frac{k!}{\alpha_1! \ldots \alpha_n!}$$

we define $\binom{k}{\alpha} = 0$ if $\alpha_i < 0$ for some $i \in [n]$.

Consider the expansion

$$\begin{aligned} \left(\sum_{i=1}^n x_i\right)^r \left(p(\mathbf{x}) - \lambda \left(\sum_{i=1}^n x_i\right)^d\right) &= \sum_{|\beta|=r} \binom{r}{\beta} \mathbf{x}^\beta \sum_{|\alpha|=d} p_\alpha \mathbf{x}^\alpha - \lambda \sum_{|\beta|=r+d} \binom{r+d}{\beta} \mathbf{x}^\beta \\ &= \sum_{|\beta|=r+d} \left(\sum_{|\alpha|=d} \binom{r}{\beta-\alpha} p_\alpha - \lambda \binom{r+d}{\beta}\right) \mathbf{x}^\beta. \end{aligned}$$

Thus, the LP formulation of (3.13) reads

$$\begin{aligned} p^{(r)} = \max\ & \lambda \\ \text{s.t. } & \binom{r+d}{\beta} \lambda \leq \sum_{|\alpha|=d} \binom{r}{\beta-\alpha} p_\alpha \quad \forall |\beta| = r+d \end{aligned}$$

with its dual

$$\begin{aligned} p^{(r)} = \min\ & \sum_{|\beta|=r+d} \sum_{|\alpha|=d} y_\beta \binom{r}{\beta-\alpha} p_\alpha \\ \text{s.t. } & y_\beta \geq 0 \quad \forall |\beta| = r+d \\ & \sum_{|\beta|=d+r} \binom{r+d}{\beta} y_\beta = 1. \end{aligned}$$

Let y be an optimal solution for the dual and define

$$L^{(r+d)}(\mathbf{x}^\beta) = y_\beta \quad \forall |\beta| = r+d.$$

Then for $|\alpha| = r+d-1$ we let

$$L^{(r+d)}(\mathbf{x}^\alpha) = \sum_{i=1}^n y_{\alpha+e_i}$$

and proceed in this manner for all $|\gamma| \leq r+d-2$. The last constraint of the dual then implies

$$1 = \sum_{|\beta|=d+r} \binom{r+d}{\beta} y_\beta = \sum_{|\beta|=d+r} \binom{r+d}{\beta} L^{(r+d)}(\mathbf{x}^\beta) = L^{(r+d)}\left(\left(\sum_{i=1}^n x_i\right)^{r+d}\right).$$

By construction, we have

1. $L^{(r+d)}(\mathbf{x}^\alpha) \geq 0$ for all $|\alpha| \leq r+d$
2. $L^{(r+d)}(\mathbf{x}^\alpha) = L^{(r+d)}\left(\mathbf{x}^\alpha \sum_{i=1}^n x_i\right)$ for all $|\alpha| \leq r+d-1$
3. $1 = L^{(r+d)}\left(\left(\sum_{i=1}^n x_i\right)^{r+d}\right) \overset{2.}{=} L^{(r+d)}(1)$.

Hence, the constructed solution for the LP relaxation is feasible. Further,

$$\begin{aligned}
p^{(r)} &= \sum_{|\beta|=r+d} \sum_{|\alpha|=d} y_\beta \binom{r}{\beta-\alpha} p_\alpha \\
&= \sum_{|\beta|=r+d} \sum_{|\alpha|=d} L^{(r+d)}(\mathbf{x}^\beta) \binom{r}{\beta-\alpha} p_\alpha \\
&= L^{(r+d)}\left(\sum_{|\beta|=r+d} \sum_{|\alpha|=d} \binom{r}{\beta-\alpha} p_\alpha \mathbf{x}^\beta\right) \\
&= L^{(r+d)}\left(\left(\sum_{i=1}^n x_i\right)^r p\right) \\
&= L^{(r+d)}(p) \geq \underset{\text{LP}}{f}^{(r+d)}.
\end{aligned}$$

□

As has been noted before, the estimate of Theorem (3.7) depends on the dual variables. While it is in general not possible to get rid of these variables in the estimate, there are cases in which we can. In the following we present an example of such a case.

Example 3.19. *Consider the case of polynomial optimization over the simplex. Let* $f \in \mathbb{R}[\boldsymbol{x}]$ *be of degree* d *and set*

$$f_{\min} = \min_{\boldsymbol{x} \in \Delta_{n-1}} f(\boldsymbol{x}),$$

and analogously define $f_{\max}$*. We can cast this as a GMP of type* (3.1)

$$f_{\min} = \inf_{\mu \in \mathcal{M}(\Delta_{n-1})} \left\{ \int_{\Delta_{n-1}} f(\boldsymbol{x}) \mathrm{d}\mu : \int_{\Delta_{n-1}} \mathrm{d}\mu = 1, \int_{\Delta_{n-1}} \mathrm{d}\mu \leq 1 \right\}.$$

A dual optimal solution is in this case given by $(y^*, t^*) = (f_{\min}, 0)$*. Noting that in the estimate we set* $y_0 = 1$*, our estimate* (3.8) *becomes*

$$f_{\min} - \underset{\sim}{f}_{\mathrm{LP}}^{(r+d)} \leq \frac{d(d-1)}{2(r+d-1) - d(d-1)} (B(f) - f_{\min})$$

and applying the inequality

$$B(f) - f_{\min} \leq \binom{2d-1}{d} d^d \, (f_{\max} - f_{\min}),$$

shown in [dKLP06, Theorem 2.2], we find

$$f_{\min} - \underset{\sim}{f}_{\mathrm{LP}}^{(r+d)} \leq \frac{d(d-1)}{2(r+d-1) - d(d-1)} \binom{2d-1}{d} d^d \, (f_{\max} - f_{\min}).$$

This is essentially the same result as was obtained in [dKLP06, Theorem 1.3]. The presented example highlights the fact that results for convergence rate of the GMP may not be as clean as for simpler problems like polynomial optimization, even though the tools that are used to obtain these results are the same. This, of course, is due to the fact that the GMP is much more complicated in general.

Moreover, we would like to emphasize that the conceptual tools of this chapter are not limited to the cases that were treated. In fact, given a quantitative version of a Positivstellensatz, it is possible to perform a convergence analysis of the kinds we proposed in this chapter as long as the nature of the relaxation hierarchy, i.e., linear or semidefinite, is coherent with the positivity certificate given by the Positivstellensatz. For example, for more general sets $\mathbf{K}$ there is a (much weaker) quantitative Positivstellensatz available found by Nie and Schweighofer [NS07] as well as a recent improvement by Baldi and Mourrain [BM22]. This result can be used to bound the rate of convergence of the GMP for more general sets. We chose to discuss the simplex and the sphere as there are strong Positivstellensätze available in these cases and to expose the fact that the relaxation must be in line with the certificate. For the sphere case, one could also make use of the following Positivstellensatz by Reznick.

Theorem 3.20 (cf. Theorem 3.12 in [Rez95])**.** *Assume f is a homogeneous polynomial of degree $2d$ such that $0 \leq f(x) \leq 1$ for all $x \in \mathcal{S}^{n-1}$. Then one has*

$$f(\mathbf{x}) + \frac{d(d-1)n}{r \log 2} = \sigma(\mathbf{x}) + (1 - \|\mathbf{x}\|^2)\|h(\mathbf{x})$$

for some $\sigma \in \Sigma[\mathbf{x}]_{r+d}$ and $h \in \mathbb{R}[\mathbf{x}]_{2(r+d)-2}$.

By using this theorem instead of Theorem 3.10, one obtains a convergence result with fewer assumptions than the one presented in Theorem 3.11, but at the cost of a worse convergence rate. In particular, one may avoid the assumption $n \leq d$ in Theorem 3.10 by using the result by Reznick, leading to a convergence rate of $O(1/r)$ on the sphere (as opposed to the $O(1/r^2)$ in Theorem 3.11). Finally, we would like to mentioned that in a recent article [Slo22b] provided a stronger quantitative Positivstellensatz for the simplex than Theorem 3.6 using similar techniques as [FF21]. Using the results from [Slo22b] a convergence rate of $O(1/r^2)$ can be obtained.

4

Computing bounds for option prices

4.1 Introduction

Derivative securities have become an integral part in financial economics and constitute attractive instruments for a wide variety of parties. Such products may be used to hedge portfolios, ensure financial planning security in supply chains and for investment purposes. The value of a derivative security relies on the value of one or multiple assets, called *underlyings*, like stocks, currencies or commodities. The most commonly used derivative securities are futures, forwards, swaps or options. A central question of financial economics is at what price to sell such products. Important in this respect is to ensure the price put on the security creates no possibility of arbitrage, i.e., there must not be a risk-free possibility to make money. Two main approaches to finding bounds on such prices are used throughout the literature. The first one assumes the prices of the underlying assets follow a stochastic differential equation (SDE) and tools from the theory of SDEs are used to solve the problem of finding a price. The most famous model in this regard is the Black and Scholes model, which provides closed formula solutions to many problems. However, this has the drawback that the assumed model is highly susceptible to model misspecifications and to parameter estimation errors. The other approach, which is the one we will follow, has no underlying model or assump-

tions on the price dynamics, but solely assumes the non-existence of arbitrage. It is based on the idea of using observable data like prices of other options on the same asset or prices of correlation-based derivatives and then using semidefinite optimization techniques to obtain solutions.

In this chapter we will focus on the problem of deriving bounds on the price of *European call options*. A European call option is a contract that gives the owner the right, but no obligation, to buy an underlying asset at fixed price, referred to as *strike* (or *strike price*) at a predetermined date in the future, called *maturity*. Since the owner is not obliged to exercise the option, it has nonnegative value. For example, consider a European call option with strike K on a stock, whose price at time t is given by S_t. If at maturity T the price S_T of the stock is greater than the strike price K, a rational owner will exercise the option and make a profit of $S_T - K$. If, however, the price of the asset is less than the strike, the owner will not exercise the option (since they could buy the stock cheaper at the stock market) and therefore not make a profit. Thus, the *payoff function* of the option is given by $\max(S_T - K, 0)$. There are many different types of options and we will introduce the ones that will be relevant in this chapter. A *rainbow option* is an option on multiple underlyings $S_T^{(1)}, \dots, S_T^{(n)}$ that pays on the level of one option, for example a call on max with payoff function $\max(0, \max(S_t^{(1)}, \dots, S_t^{(n)}) - K)$. This is equivalent to a *lookback option* on one asset if $S_t^{(i)}$ is the price of the same asset at n points in time. A *basket option* also depends on multiple assets and pays on the level of more than one. For example, it could be a weighted linear combination of the prices of the assets at maturity with payoff function $\max(0, \sum_{i=1}^{n} \alpha_i S_t^{(i)} - K)$, where $\alpha_i \geq 0$. Examples for options of this type are index options or currency basket options. Because markets are incomplete in general, it is not possible to compute exact prices of options. However, one can compute bounds, such that, if the price of the option lies within the given range, it is consistent with the given information and does not create the possibility of arbitrage.

4.1.1 Prior work

The problem of computing bounds on option prices without assuming a specific price dynamic of the underlying assets has been studied since the 1970s beginning with the poineering work of Merton [Mer73]. Similar problems have already been studied in the late 1950s, see [Sca57]. Cox and Ross [CR76] and Harrison and Kreps [HK79] show that the assumption of no arbitrage possibilities is equivalent

to the existence of a probability measure under which the option prices become Martingales. Boyle and Lin [BL97] extended prior contributions of Lo [Lo87] considering the problem of deriving upper bounds on basket options on multiple assets given the means and the covariance matrix of the underlying assets by constructing a semidefinite program. In [BP02], Bertsimas and Popescu considered a more general setting assuming observable options prices as well as moment information of the underlying distribution of the assets like means and variances are available. Using semidefinite programming techniques they solve the univariate case and give closed form solutions to some cases. For the multivariate case, i.e., options depending on multiple assets they prove that the problem is NP-hard in general and present a relaxation to the problem by enlarging the set of possible values the assets can attain. They follow up by identifying the cases in which their relaxation can be solved efficiently, which is the case if the objective and the constraint functions are quadratic or linear over d disjoint polyhedra $D_1, \ldots, D_d$ which form a partition of $\mathbb{R}^n$, where n is the number of assets considered. Davis and Hobson [DH07] study the structure of the underlying problem and give sufficient and necessary conditions for the existence of measures specified in [CR76], [HK79]. In a series of papers (see [LW05], [HLW05b], [HLW05a]) Hobson, Laurence, and Wang consider the case of multivariate basket options and give sharp upper and lower bounds when the constraints consist of observable vanilla options prices. They do not employ semidefinite programming techniques, but approach the problem by constructing primal and dual solutions with a zero duality gap. Primbs [Pri06] constructs dynamic replicating portfolios using semidefinite programming to get upper and lower bound on option prices, using knowledge of piecewise polynomial data. In his dissertation [d'A04], d'Aspremont computes bounds for basket options by constructing static replicating portfolios assuming knowledge on prices of different basket options with the same maturity. Li et al. [LSS05] extend the work of Bertsimas and Popescu using SOS relaxations to obtain a hierarchy of bounds on option prices. Another approach was taken by Peña and Zuluaga [ZP05]. They used tools from conic programming to reformulate the considered problem and prove strong duality in many cases. To give approximate solutions to the problem they propose to use increasingly tight outer approximations of the cone of interest. For certain sets K they provide explicit outer approximation sequences for the cone of measures supported on K, and use these to compute upper bounds for option prices. We also point out the paper by Peña, Vera and Zuluaga [PVZ12] where the authors consider the problem of computing upper bound on European basket options under the no-arbitrage as-

sumption taking into account the presence of bid-ask spreads. The problem is recast as an LP linear in size with respect to the input data.

4.1.2 Contribution of this chapter

The work presented in this chapter builds on the work of Bertsimas and Popescu [BP02]. We analyze and computationally explore cases which they determined to be NP-hard. We consider a model similar to the one treated by Li et al. [LSS05], which in itself is a generalization of the problem Boyle and Lin [BL97] considered. While the authors in [LSS05] focus on a dual approach using inner (i.e., sum of squares) approximations of the cone of positive polynomials, our main interest lies in a primal method relying on an outer approximation of the moment cone. In contrast to Li et al. we give a rigorous argument as to why we consider compact underlying sets whenever we do so. To complement our primal method of outer approximation we analyze an inner approximation of the moment cone as well. Our inner approximation does not rely on any compactness assumption. In special cases we give explicit bounds on the support of the optimal solution of the treated problem. Our method of outer approximations takes the same approach as Peña and Zuluaga in [ZP05]. Our analysis contributes additional insights into when optimal solutions exist and the proposed hierarchies converge. Several numerical examples are provided to illustrate the effectiveness of our methods.

4.1.3 Outline of the chapter

In Section 4.2 we present the problem we intend to study in this chapter, which is finding bounds on the prices of options depending on multiple assets without assuming any underlying stochastic processes of the assets prices. This can be modeled as a generalized moment problem over a non-compact set. We consider a general model assuming the information given is in the form of observable prices of vanilla options on the given assets, as well as priced moment information like mean, variance and covariance of the underlying assets. Higher order moment information like skewness and kurtosis can also be used in this model. Also in Section 4.2 we prove the existence of an optimal solution of the problem formulation we proposed. Equipped with this knowledge we continue in Section 4.3 to demonstrate how to obtain a bound on the support of the optimal solution in the univariate case. The reason for this is that this allows to consider a new problem over a compact set with the same solution. This ensures that the moment-SOS hi-

erarchy converges to the optimal solution in this case. Section 4.4 contains a few examples of numerical computations for problems with real world data as well as some explanation of the implementation techniques. In Section 4.5 we apply a relaxation technique for the non-compact generalized moment problem to our setting and conclude the section with a numerical example to show its effectiveness. This procedure provides upper (resp. lower) bounds to the minimization (resp. maximization) problem, and we refer to these bounds as the *inner range*. The last section serves as a conclusion and gives direction for future work.

4.2 Bounds on options via the GMP formulation

In this section we will cast the problem of computing bounds on the price of European call options as a particular instance of the GMP. The option will be dependent on n assets $S_1, \ldots, S_n$. We will denote the (piecewise polynomial) payoff function by $\varphi : \mathbb{R}_+^n \to \mathbb{R}_+$, which may depend on the prices of the n different assets. We assume the payoff is nonnegative, since we consider options, meaning there is no obligation of the owner to exercise it, in which case the payoff is zero. The range of possible prices for asset S_i will be the nonnegative reals, i.e., $x_i \in \mathbb{R}_+$. Note that the payoff function is what defines the type of the option. As mentioned in Section 4.1.1 the no-arbitrage assumption is equivalent to the existence of a probability measure μ such that asset prices become martingales under μ. This measure is referred to as the *equivalent martingale measure* or the *risk-neutral measure*. The price of the option is then given by the expectation of the payoff function with respect to this measure. Here and throughout this chapter we assume for simplicity an interest rate of 0.

4.2.1 Problem statement

Let for a finite index set $\mathcal{I}$ information pairs (f_i, q_i) be given where $i \in \mathcal{I}, f_i : \mathbb{R}_+^n \to \mathbb{R}$ and $q_i \in \mathbb{R}$. These pairs might consist of (piecewise polynomial) payoff functions f_i of options on the assets $S_1, \ldots, S_n$ with the observable prices q_i at which these options are traded, or prices of derivatives on moments of underlying asset, such as mean, variance or correlation. In order to find bounds for the option at hand we will look for a probability measure that is consistent with this given information. In other words, the feasible set of measures μ will consist of measures such that

$$\int_{\mathbb{R}^n_+} f_i(\mathbf{x})\mathrm{d}\mu(\mathbf{x}) \leqq q_i\,, \text{ for all } i \in \mathcal{I},$$

where " $\leqq$ " means either " $\leq$ " or " $=$ ". We will also assume that the d-th order moments of the corresponding distributions are finite for some $d \in \mathbb{N}$. To fix ideas we will consider the following problem adapted from [BP02]. Given n assets $S_1, \ldots, S_n$ whose prices are given by $x_1, \ldots, x_n$, we want to find a lower bound on a European call option whose payoff may depend on the assets S_i for $i \in [n]$. The available information may come in the form of observable option prices in which case the corresponding f_i is a piecewise polynomial which we assume to be continuous and q_i is the respective price of the option or in the form of observable moments in which case $f_i \in \mathbb{R}[\mathbf{x}]$. For example, if γ_i is the observed mean of asset i and the observed covariance of assets i and j is $\sigma_{i,j}$, one can add the constraint

$$\int_{\mathbb{R}^n_+} (x_i - \gamma_i)(x_j - \gamma_j)\mathrm{d}\mu(\mathbf{x}) = \sigma_{i,j}\,.$$

Define the degree of a piecewise polynomial as the largest degree of its constituents. We assume the d-th order moments under a risk-neutral pricing measure are finite, where

$$d = \max_{i \in \mathcal{I}} \{\deg(\varphi), \deg(f_i)\} + 1.$$

What we mean by this is that

$$\int_{\mathbb{R}^n_+} \|\mathbf{x}\|_2^d \mathrm{d}\mu(\mathbf{x}) \leq M$$

for some $M \in \mathbb{R}_+$, where $\|\mathbf{x}\|_2 = \sqrt{x_1^2 + \cdots + x_n^2}$ is the standard Euclidean ℓ_2-norm. A risk-neutral pricing measure is a measure such that the asset prices are equal to the expectation under this measure discounted by the risk-free interest rate. For convenience, we assume that d is even, otherwise we set $d \leftarrow d+1$. This way we make sure that we are dealing with a GMP with (piecewise) polynomial data. The optimal value of the optimization problem below will serve as bound for the given option that is consistent with the available information.

$$\begin{aligned} \sup_{\mu\in\mathcal{M}(\mathbb{R}^n_+)_+} \Big/ \inf_{\mu\in\mathcal{M}(\mathbb{R}^n_+)_+} & \int_{\mathbb{R}^n_+} \varphi(\mathbf{x})\mathrm{d}\mu(\mathbf{x}) \\ \text{s.t.} & \int_{\mathbb{R}^n_+} f_i(\mathbf{x})\mathrm{d}\mu(\mathbf{x}) \leqq q_i\,, \text{ for } i\in\mathcal{I} \\ & \int_{\mathbb{R}^n_+} \mathrm{d}\mu(\mathbf{x}) = 1 \\ & \int_{\mathbb{R}^n_+} \|\mathbf{x}\|_2^d \mathrm{d}\mu(\mathbf{x}) \leq M\,. \end{aligned} \tag{4.1}$$

To obtain upper bounds we maximize and for lower bounds we minimize. In a nutshell, one is looking for the probability distribution of the asset price, that is consistent with the known information and minimizes (respectively maximizes) the objective.

4.2.2 Existence of an optimal solution

Now we prove that the infimum in (4.1) is attained. In order to do so, we will use the Prokhorov theorem [Pro56] asserting a weak sequential compactness of a family of tight measures, as well as the monotone convergence theorem.

Definition 4.1 (Tightness)**.** A sequence of measures $(\mu_k)_{k=1}^{\infty}$ defined on $\mathbb{R}^n$ is called tight if for every $\epsilon > 0$ there exists a compact set K such that $\mu_k(K^c) < \epsilon$ for all $k \in \mathbb{N}$.

Theorem 4.2 (Prokhorov)**.** *Let $(\mu_k)_{k=1}^{\infty}$ be a tight sequence of Borel probability measures on $\mathbb{R}^n$. Then there exists a Borel probability measure μ and a subsequence $(\mu_{k_i})_{i=1}^{\infty}$ converging weakly to μ, i.e.,*

$$\lim_{i\to\infty} \int g\,\mathrm{d}\mu_{k_i} = \int g\,\mathrm{d}\mu \tag{4.2}$$

for all bounded continuous functions g on $\mathbb{R}^n$.

Theorem 4.3 (Monotone Convergence Theorem)**.** *Let X be a measure space with a positive measure μ and let $\{g_\ell : X \to [0,\infty)\}_{\ell=1}^{\infty}$ be a sequence of pointwise non-decreasing μ-measurable functions, i.e., $g_1 \leq g_2 \leq \ldots$. Further, let g be the pointwise limit of $(g_\ell)_{\ell=1}^{\infty}$, i.e.,*

$$g(\boldsymbol{x}) = \lim_{\ell\to\infty} g_\ell(\boldsymbol{x}).$$

Then, g is μ measurable and

$$\lim_{\ell\to\infty}\int g_\ell \,\mathrm{d}\mu = \int g \,\mathrm{d}\mu\,.$$

Theorem 4.4. *If Problem (4.1) is feasible, then its supremum/infimum is attained.*

Proof. We begin by observing that if (4.1) is feasible, then the infimum in (4.1) is finite since the objective function is nonnegative. Also, by the choice of d the supremum is finite because of the last constraint $\int_{\mathbb{R}^n_+} \|\mathbf{x}\|_2^d \mathrm{d}\mu(\mathbf{x}) \le M$. Let $(\mu_k)_{k=1}^\infty$ be a sequence of feasible solutions for (4.1) converging to an optimal solution. Denote by ϕ_k the measures defined by

$$\mathrm{d}\phi_k = (1+\|\mathbf{x}\|_2^{d-1})\mathrm{d}\mu_k\,.$$

Moving on, we show that the sequence $(\phi_k)_{k=1}^\infty$ is tight. Let $\epsilon > 0$ be given and let K be the closed ball of radius a. Then we have

$$\begin{aligned}\phi_k(K^c) &= \int_{\mathbb{R}^n} \mathbb{1}_{\{\|\mathbf{x}\|_2\ge a\}}(1+\|\mathbf{x}\|_2^{d-1})\mathrm{d}\mu_k \\ &\le \frac{1}{a}\int_{\mathbb{R}^n} \|\mathbf{x}\|_2(1+\|\mathbf{x}\|_2^{d-1})\mathrm{d}\mu_k \le \frac{M^{1/d}+M}{a}\,,\end{aligned}$$

where we used Jensen's inequality [Jen06] in the last step. By picking a sufficiently large, we make $\phi_k(K^c) < \epsilon$, hence establishing tightness. By Theorem 4.2, there exists a weakly convergent subsequence (that we do not relabel) that converges weakly to a measure ϕ. We set

$$\mathrm{d}\mu := \frac{\mathrm{d}\phi}{1+\|\mathbf{x}\|_2^{d-1}}$$

to be the candidate optimizer for (4.1). We first show that the equality constraints for (4.1) are satisfied by μ. For a given pair (f_i, q_i) corresponding to an equality constraint we have

$$\begin{aligned}q_i &= \lim_{k\to\infty}\int f_i \,\mathrm{d}\mu_k = \lim_{k\to\infty}\int \frac{f_i}{1+\|\mathbf{x}\|_2^{d-1}}\,\mathrm{d}\phi_k \\ &= \int \frac{f_i}{1+\|\mathbf{x}\|_2^{d-1}}\,\mathrm{d}\phi = \int f_i \,\mathrm{d}\mu\,,\end{aligned} \tag{4.3}$$

where in the third equality we used the fact that the function $f_i/(1+\|\mathbf{x}\|_2^{d-1})$ is continuous and bounded since $\deg(f_i) \le d+1$. The same argument applies if

(f_i, q_i) corresponds to an inequality constraint only the first equality in (4.3) is replaced by a "$\geq$". Moreover, (4.3) also shows that μ leads to the same objective value as the optimizing sequence $(\mu_k)_{k=1}^{\infty}$. Finally, we establish that $\int \|\mathbf{x}\|_2^d \, \mathrm{d}\mu < M$. We define $g_\ell(\mathbf{x}) := \min(\|\mathbf{x}\|_2^d, \ell)$. Note that g_ℓ is a pointwise non-decreasing function and its pointwise limit is $\|\mathbf{x}\|_2^d$. Therefore, Theorem 4.3 applies and we have

$$\begin{aligned}\int \|\mathbf{x}\|_2^d \, \mathrm{d}\mu &\overset{(i)}{=} \lim_{\ell\to\infty} \int g_\ell \, \mathrm{d}\mu = \lim_{\ell\to\infty} \int \frac{g_\ell}{1+\|\mathbf{x}\|_2^{d-1}} \, \mathrm{d}\phi \\ &\overset{(ii)}{=} \lim_{\ell\to\infty} \lim_{k\to\infty} \int \frac{g_\ell}{1+\|\mathbf{x}\|_2^{d-1}} \, \mathrm{d}\phi_k = \lim_{\ell\to\infty} \lim_{k\to\infty} \int g_\ell \, \mathrm{d}\mu_k \\ &\overset{(iii)}{\leq} \lim_{\ell\to\infty} \lim_{k\to\infty} \int \|\mathbf{x}\|_2^d \, \mathrm{d}\mu_k \leq M \,.\end{aligned}$$

In (i) we used Theorem 4.3, the weak convergence of ϕ_k to ϕ in (ii) and the fact that $g_\ell \leq \|\mathbf{x}\|_2^d$ in (iii). □

Combining this result with the Richter theorem (see [Ric57, Satz 4] for an original reference or [dS18, Theorem 19] for a modern statement and historical remarks), we get the following immediate corollary.

Corollary 4.5. *If Problem (4.1) is feasible, then the optimal value of (4.1) is attained by an atomic measure with finitely many atoms.*

We finish this section by showing that finite d-th order moments are necessary for the existence of an optimal solution.

Proposition 4.6. *The last constraint in* (4.1) *cannot be omitted in Theorem 4.4.*

Proof. Consider the following problem

$$\begin{aligned} p^* = \inf & \int_0^\infty \max(0, x-k_1)\mathrm{d}\mu \\ \text{s.t.} & \int_0^\infty \max(0, x-k_2)\mathrm{d}\mu = a \\ & \int_0^\infty \mathrm{d}\mu = 1 \,, \end{aligned} \tag{4.4}$$

where we assume $k_1 < k_2$ and $a \neq 0$. Note that this implies that for the optimal value we have $p^* \geq a$. We will show that there exists no measure for which the

optimal value is attained. The following is a minimizing sequence for (4.4)

$$\mu_n = \left(1 - \frac{1}{n}\right)\delta_{k_1} + \frac{1}{n}\delta_{k_2+na}\,.$$

For every $n \in \mathbb{N}$ we see that μ_n is a probability measure as it is a convex combination of atomic measures and

$$\int_0^\infty \max(0, x-k_2)\mathrm{d}\mu_n = \frac{1}{n}(k_2 + na - k_2) = a\,.$$

So the sequence is indeed feasible. For the objective value we get

$$\begin{aligned}\int_0^\infty \max(0, x-k_1)\mathrm{d}\mu_n &= \left(1-\frac{1}{n}\right)(k_1-k_1) + \frac{1}{n}(k_2+na-k_1)\\ &= a + \frac{1}{n}(k_2-k_1)\,.\end{aligned}$$

So we have that μ_n is a minimizing sequence as it is feasible and converges to $a \leq p^*$. The limit $\lim_{n\to\infty}\mu_n = \delta_{k_1}$, however, is not feasible. We now show that there exists no probability measure $\mu \in \mathcal{M}(\mathbb{R}_+)_+$ that is optimal for (4.4). For this we assume that μ is an optimizer of (4.4). Then we have

$$\int_{\mathbb{R}_+} \max(0, x-k_1)\mathrm{d}\mu(x) = a = \int_{\mathbb{R}_+} \max(0, x-k_2)\mathrm{d}\mu(x)\,.$$

Thus,

$$\begin{aligned}0 &= \int_{\mathbb{R}_+} \max(0, x-k_1)\mathrm{d}\mu(x) - \int_{\mathbb{R}_+} \max(0, x-k_2)\mathrm{d}\mu(x)\\ &= \int_{k_1}^{k_2} \underbrace{(x-k_1)}_{\geq 0}\mathrm{d}\mu(x) + \int_{k_2}^{\infty} \underbrace{(k_2-k_1)}_{>0}\mathrm{d}\mu(x)\,.\end{aligned}$$

The latter integral must be zero which implies that $\mathrm{supp}(\mu) \cap [k_2, \infty) = \emptyset$. But if that is the case we have

$$a = \int_{\mathbb{R}_+} \max(0, x-k_2)\mathrm{d}\mu(x) = \int_{k_2}^{\infty} (x-k_2)\mathrm{d}\mu(x) = 0.$$

Therefore, μ cannot be feasible. □

Example 4.4 illustrates that the support of a minimizing sequence may be unbounded.

4.3 Bounding the support

By Corollary 4.5 the optimal solution to (4.1) is a measure with finitely many atoms. This section is devoted to the question whether it is possible to bound the support of the optimal solution in terms of the problem data of (4.1). If this were possible, i.e., if we knew the optimal solution is attained in a box $[0,B]^n$ for some $B \in \mathbb{R}_+$, we could consider a compact version of (4.1), where $\mathbb{R}^n_+$ is replaced by $[0,B]^n$. This has the advantage that we know that the moment-SOS hierarchy converges if the underlying sets are compact (recall Archimedian assumption in connection with [Las09, Theorem 4.10]).

4.3.1 Atomic representation

We now present an approach to the problem of bounding the support. We consider the univariate case for problem (4.1) and assume for all $i \in \mathcal{I}$ that $f_i(x) = \max(0, x - k_i)$ is the payoff function of an observable European call option with associated price q_i. By Corollary 4.5 there exists an atomic solution of the form $\sum_{j=1}^{N} \alpha_j \delta_{x_j}$. Let $0 \le x_1 \le x_2 \le \cdots \le x_N$. The following lemma shows that we may assume w.l.o.g. that $x_{N-1} \le k_m \le x_N$ if $\varphi(x) = \max(0, x - k)$ with $k \le \max_{i\in\mathcal{I}}\{k_i\}$.

Lemma 4.7. *Consider* (4.1) *for* $n = 1$, *let* $\varphi(x) = \max(0, x - k_j)$ *and* $f_i(x) = \max(0, x - k_i), q_i > 0$ *for all* $i \in \mathcal{I} = [m] \setminus \{j\}$ *be the payoff functions of European call options with strike* k_i *such that* $k_1 \le \cdots \le k_{j-1} \le k_j \le k_{j+1} \cdots \le k_m$ *with associated prices* q_i. *If there exists an optimal solution, then there exists one such that exactly one atom* x_i *lies in* (k_m, ∞). *Moreover, there exists a solution such that in each of the intervals*

$$[0, k_1), [k_1, k_2), \ldots, [k_{m-1}, k_m), [k_m, \infty)$$

there is at most one atom.

Proof. For the first claim asserting that there exists an optimal measure μ^* such that exactly one atom lies in $[k_m, \infty)$, let us assume that all atoms lie in $[0, k_m)$. Then

$$q_m = \int_{\mathbb{R}_+} \max(0, x - k_m) \mathrm{d}\mu^* = \int_0^{k_m} \max(0, x - k_m) \mathrm{d}\mu^* = 0\,,$$

which is a contradiction since $q_i > 0$. Thus, at least one atom lies in $[k_m, \infty)$. Suppose two atoms lie in $[k_m, \infty)$ and let the associated weighted Dirac measures be $\alpha\delta_{x_1}$ and $\beta\delta_{x_2}$ with $\alpha, \beta > 0$ and $x_1, x_2 > k_m$. Now, since $x_1, x_2 > k_m$ these two Dirac measures influence every constraint of (4.1) as well as the objective because all input functions are strictly positive at x_1, x_2. Their influence on the i-th constraint is exactly

$$\alpha(x_1 - k_i) + \beta(x_2 - k_i) = (\alpha+\beta)\left(\frac{\alpha}{\alpha+\beta}x_1 + \frac{\beta}{\alpha+\beta}x_2 - k_i\right).$$

It follows that these two Dirac measures can be combined to a single one with weight $\omega = \alpha + \beta > 0$ and support $x = \frac{\alpha}{\alpha+\beta}x_1 + \frac{\beta}{\alpha+\beta}x_2 > k_m$ and the corresponding measure remains feasible. Also, because $\|\cdot\|^d$ is convex, the inequality constraint is also satisfied. By similar reasoning one can prove the second claim of the lemma. □

Lemma 4.8. *In the setting of Lemma 4.7 the support of an optimal measure is contained in* $[0, B]$ *for*

$$B = \frac{M + \sqrt{M(M - 4q_m k_m)}}{2q_m}.$$

Proof. By Lemma 4.7 we have for an optimal solution of the form $\sum_{j=1}^{N} \alpha_j \delta_{x_j}$ that $\alpha_N(x_N - k_N) = q_m$. We also know $\alpha_N x_N^2 \leq M \iff \alpha_N \leq M/x_N^2$, from which follows that $q_m \leq (M/x_N^2)(x_N - k_m)$. Hence,

$$x_N \leq \frac{M + \sqrt{M(M - 4q_m k_m)}}{2q_m} =: B\,.$$

Hence, the support of an optimal solution lies in $[0, B]$. □

4.4 Examples for outer range

We will now present some examples of numerical computations of bounds on option prices in the framework specified in the previous sections. The moment-SOS hierarchy provides a lower bound to the minimization problem and an upper bound to the maximization problem, which is why we call these outer bounds. For the implementation was coded in Julia, and we used the MOSEK solver [MOS19] version 9.1.9. The code is available online [1] and relies partly on the Julia package MomentOpt.jl [WLC+19].

[1] `https://github.com/FelixKirschner/boundingOptionPricesCode`

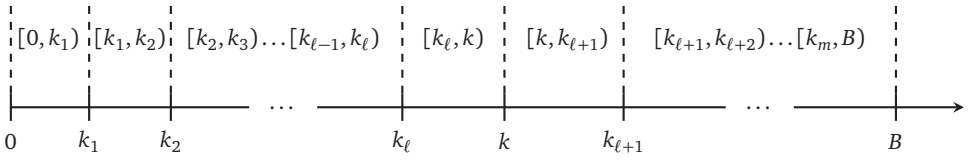


Figure 4.1: Visualization of the segmentation of the interval $[0,B]$

4.4.1 Univariate case

Let us describe our implementation strategy for the univariate case. Assume we want to find bounds on the price of an option with strike k given strikes k_i and prices q_i of other options on the same asset such that $0 < k_1 < k_2 < \cdots < k_m$. We assume there exists an $\ell \in [m-1]$ such that $k_\ell < k < k_{\ell+1}$.

$$\begin{aligned} \sup_{\mu\in\mathcal{M}(\mathbb{R}_+)_+} \Big/ \inf_{\mu\in\mathcal{M}(\mathbb{R}_+)_+} & \int_{\mathbb{R}_+} \max(0,x-k)\mathrm{d}\mu(x) \\ \text{s.t.} & \int_{\mathbb{R}_+} \max(0,x-k_i)\mathrm{d}\mu(x) = q_i\,, \text{ for } i\in[m] \\ & \int_{\mathbb{R}_+} \mathrm{d}\mu(x) = 1 \\ & \int_{\mathbb{R}_+} x^2\mathrm{d}\mu(x) \le M\,. \end{aligned} \tag{4.5}$$

Since we know from Theorem 4.4 that feasibility implies the existence of an optimal solution, we will assume the optimal solution will be attained in a box $[0,B]$ for some $B \in \mathbb{R}$. A suitable B can be obtained via the procedure described in Section 4.3.

To circumvent the problem of dealing with piecewise affine functions we split the interval $[0,B]$ into subintervals and define measures supported on each of the subintervals. For this let ℓ be the index such that $k_\ell < k < k_{\ell+1}$. We define intervals $[0,k_1],[k_i,k_{i+1}]$ for $i = 1,\ldots,\ell-1$, as well as $[k_\ell,k),[k,k_{\ell+1})$ and $[k_j,k_{j+1})$ for $j = \ell+1,\ldots,m-1$ and finally $[k_m,B)$. The situation is visualized in Figure 4.1 Let S be the collection of these subsets. Elements in S are pairwise disjoint and the union of all sets in S is $[0,B]$. The collection S contains $m+2$ intervals and to each one a measure μ_i is assigned for $i = 1,\ldots,m+2$. This way we can formulate a problem equivalent to (4.5).

$$
\begin{aligned}
\sup / \inf \ & \sum_{i=\ell+2}^{m+2} \int (x-k) \mathrm{d}\mu_i(x) \\
\text{s.t.} \ & \sum_{i=j+2}^{m+2} \int (x-k_j) \mathrm{d}\mu_i(x) = q_j \,, \text{ for } j = \ell+1, \ldots, m \\
& \sum_{i=j+1}^{m+2} \int (x-k_j) \mathrm{d}\mu_i(x) = q_j \,, \text{ for } j \in [\ell] \\
& \sum_{i=1}^{m+2} \int x^2 \mathrm{d}\mu_i(x) \leq M \\
& \sum_{i=1}^{m+2} \int \mathrm{d}\mu_i(x) = 1 \,. \\
& \operatorname{supp}(\mu_i) \subseteq s_i,
\end{aligned}
\tag{4.6}
$$

where $s_i = [s_{i_1}, s_{i_2})$ for $i = 1, \ldots, m+2$ are the elements of S. Introduce a linear operator $L_i^r : \mathbb{R}[x]_{2r} \to \mathbb{R}$ for every μ_i. The level r relaxation is then given by

$$
\begin{aligned}
\sup / \inf \ & \sum_{i=\ell+2}^{m+2} L_i^r(x-k) \\
\text{s.t.} \ & \sum_{i=j+2}^{m+2} L_i^r(x-k_j) = q_j \,, \text{ for } j = \ell+1, \ldots, m \\
& \sum_{i=j+1}^{m+2} L_i^r(x-k_j) = q_j \,, \text{ for } j \in [\ell] \\
& \sum_{i=1}^{m+2} L_i^r(x^2) \leq M \\
& \sum_{i=1}^{m+2} L_i^r(1) = 1 \\
& L_i^r([x]_r [x]_r^T) \in \mathcal{DNN} \,, \text{ for } i \in [m+2] \\
& L_i^r((s_{i_2} - x)(x - s_{i_1})[x]_{r-1} [x]_{r-1}^T) \in \mathcal{DNN} \,, \text{ for } i \in [m+2],
\end{aligned}
\tag{4.7}
$$

where $\mathcal{DNN}$ is the doubly nonnegative cone, i.e., $\mathbb{S}_+^n \cap \mathbb{R}_+^{n \times n}$ and the operator L_i^r is applied entry-wise to the matrices $[x]_r [x]_r^T$. The decision variables here are the linear operators L_i^r. By introducing a variable $y_j^{(i)} = L_i^r(x^j)$, for $i \in [m+2], j = 0, 1, \ldots, 2r + d_{\max}$ problem (4.7) becomes a regular semidefinite program. For

i	1	2	3	4	5
k_i	95	100	110	115	120
q_i	12.875	8.375	1.875	0.625	0.25

Table 4.1: Prices of European call options on the Microsoft stock from July '98 with strikes k_i

the actual calculation it is expedient to normalize everything, that is dividing the given data by B. Consider problem (4.5) with the data displayed in Table 4.1 and with $m = 5, k = 105$ and $M = 200\,000$. Using the relaxation given in (4.7) we can approximate the optimal solution and we find the first level is tight, meaning we obtained the optimal bounds proposed by Bertsimas and Popescu in [BP02]. For the considered case we get a lower bound of 3.875 and an upper bound of 5.125 and the computation took 0.01 seconds.

The domain in this problem is partitioned into 7 parts. For each part we define a measure for each of which we introduce moment variables $y_j^{(i)}$ for $j = 0, 1, \ldots, 2r + d_{\max}$, where

$$d_{\max} = \max_{i\in[n], j\in[N_i], \ell\in[m]} \{\deg(\varphi), \deg(f_{i,j}), \deg(f_\ell)\}.$$

Thus, for this particular problem, we introduced $7\times5 = 35$ variables. The number of involved matrices was $7 \times 2 = 14$, each of size 2×2. In total we had 14 linear matrix inequality (LMI) constraints, 6 equality constraints as well as $1 + 35 = 36$ inequality constraints, one to ensure finite d-th order moments and one for each variable to ensure $y_j^{(i)} \geq 0$.

4.4.2 Explicit examples with two assets

Consider the following artificial example where we want to compute bounds on the price of a basket option on a basket with two assets whose prices are given by x_1 and x_2, respectively. As a payoff function we choose $\max(0, 1/2x_1+1/2x_2-k)$. We assume we can observe the prices of two single call options on each asset. The corresponding optimization program is given in (4.8) below.

$$\begin{aligned}\sup_{\mu\in\mathcal{M}(\mathbb{R}^2_+)_+} \Big/ \inf_{\mu\in\mathcal{M}(\mathbb{R}^2_+)_+} & \int_{\mathbb{R}^2_+} \max\left(0, \frac{1}{2}x_1 + \frac{1}{2}x_2 - k\right) \mathrm{d}\mu(\mathbf{x}) \\ \text{s.t.} & \int_{\mathbb{R}^2_+} \max(0, x_i - k_{x_{i,j}}) \mathrm{d}\mu(\mathbf{x}) = a_{x_{i,j}}\,, \text{ for } i,j = 1,2 \\ & \int_{\mathbb{R}^2_+} \|\mathbf{x}\|_2^2 \mathrm{d}\mu(\mathbf{x}) \le M \\ & \int_{\mathbb{R}^2_+} \mathrm{d}\mu(\mathbf{x}) = 1.\end{aligned} \tag{4.8}$$

To solve this numerically we slice up the domain into an irregular grid along the kinks of the max-functions, under the assumption that the support of the optimal solution is contained in $[0,B]^2$ for some $B \in \mathbb{R}$. The domain then may look as depicted in Figure 4.2, where the dotted lines indicate where the objective ascends from 0, i.e., where $0.5x_1 + 0.5x_2 - k = 0$. We index the tiles from bottom to top, left to right. For each tile i in the grid we introduce a new measure μ_i. For example for tile 12 in Figure 4.2 we get

$$\begin{aligned}\operatorname{supp}(\mu_{12}) = \{\mathbf{x} \in \mathbb{R}^2 : (B - x_1)(x_1 - k_{x_{1,2}}) \ge 0\,,\ x_2(k_{x_{2,1}} - x_2) \ge 0\,, \\ 1/2x_1 + 1/2x_2 - k \ge 0\}\,.\end{aligned}$$

Consider the following (strike, price) pairs

- x_1: (100, 12), (110, 3)
- x_2: (102, 10), (107, 6)

and let $M = 200\,000$, $B = 400$ and $k = 105$.

Applying the above described procedure to problem (4.8) with the data given above results in problem (4.9). Note that (4.8) and (4.9) are equivalent. With respect to Figure 4.2 the index sets J_i for $i = 0, 1, \ldots, 4$ correspond to the sets on which the functions which define problem (4.8) are not identically zero, i.e.,

- $\max(0, \frac{1}{2}x_1 + \frac{1}{2}x_2 - k) = \frac{1}{2}x_1 + \frac{1}{2}x_2 - k$ on $J_0 = \{4, 6, 8, 10, 12, 13, 14\}$
- $\max(0, x_1 - k_{x_1,1}) = x_1 - k_{x_1,1}$ on $J_1 = \{5, 6, \ldots, 14\}$
- $\max(0, x_1 - k_{x_1,2}) = x_1 - k_{x_1,2}$ on $J_2 = \{11, \ldots, 14\}$

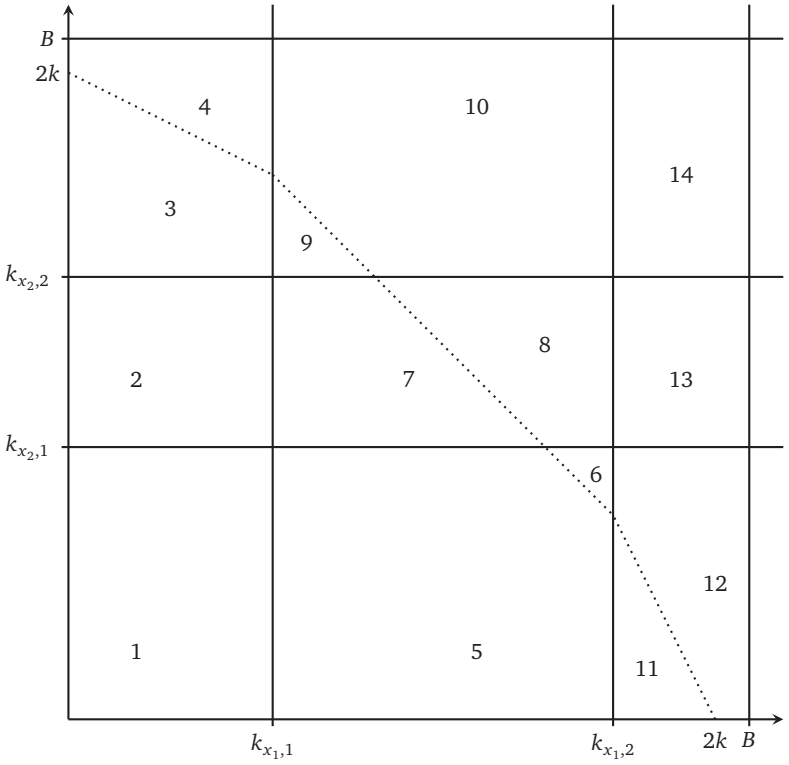


Figure 4.2: Example of how the support might be split

- $\max(0, x_2 - k_{x_2,1}) = x_2 - k_{x_2,1}$ on $J_3 = \{2,3,4,7,8,9,10,13,14\}$
- $\max(0, x_2 - k_{x_2,2}) = x_2 - k_{x_2,2}$ on $J_4 = \{3,4,9,10,14\}$.

Thus, we obtain the following problem:

$$\begin{aligned}
\sup_{\mu_i} / \inf_{\mu_i} & \sum_{i \in J_0} \int \left(\frac{1}{2} x_1 + \frac{1}{2} x_2 - k \right) \mathrm{d}\mu_i(\mathbf{x}) \\
\text{s.t.} & \sum_{j \in J_1} \int (x_1 - k_{x_1,1}) \, \mathrm{d}\mu_j(\mathbf{x}) = q_{x,1} \\
& \sum_{j \in J_2} \int (x_1 - k_{x_1,2}) \, \mathrm{d}\mu_j(\mathbf{x}) = q_{x_1,2} \\
& \sum_{j \in J_3} \int (x_2 - k_{x_2,1}) \, \mathrm{d}\mu_j(\mathbf{x}) = q_{x_2,1} \\
& \sum_{j \in J_4} \int (x_2 - k_{x_2,2}) \, \mathrm{d}\mu_j(\mathbf{x}) = q_{x_2,2} \\
& \sum_{i=1}^{14} \int \mathrm{d}\mu_i(\mathbf{x}) = 1 \\
& \sum_{i=1}^{14} \int x_1^2 + x_2^2 \, \mathrm{d}\mu_i(\mathbf{x}) \leq M \, .
\end{aligned} \tag{4.9}$$

Applying the moment-SOS hierarchy to this problem and solving the first level results in an upper bound of 7.4 and a lower bound of 2.387, which are in the optimal values of (4.8). The SDP consisted of $14 \times 15 = 210$ variables, 80 LMIs involving matrices of size 3×3, 4 equality constraints and 211 inequality constraints.

Varying strikes

We are now going to give an example to see how changing the strike price affects the optimal values of the optimization problems. Consider the data presented in Table 4.2 and let the objective function be $\max(0, 1/2x_1 + 1/2x_2 - k)$, $B = 400$ and $M = 200\,000$. The optimal values are given in Table 4.3. All values stem from the first level of the moment-SOS hierarchy and increasing the level up to level 10 did not change the objective values. For each of the strike prices specified in Table 4.2 the resulting programm for the first level of the hierarchy consisted

i	1	2	3	4	5
$k_{x_1,i}$	90	95	100	110	120
$q_{x_1,i}$	20	15.5	12	5.5	1
$k_{x_2,i}$	90	96	102	107	115
$q_{x_2,i}$	20.5	15	10	6	0.75

Table 4.2: Strikes and corresponding prices for European call options

k	90	95	100	105	110	115
lower bound on price	16.875	12.792	8.708	4.625	1.675	0.0
computation time [s]	0.12	0.14	0.13	0.15	0.14	0.13
upper bound on price	20.25	15.7	11.55	8.016	4.75	2
computation time [s]	0.12	0.14	0.16	0.15	0.15	0.14

Table 4.3: Optimal lower and upper bounds w.r.t. the data given in Table 4.2

of $47 \times 15 = 705$ variables, 257 LMIs, 11 equality constraints and 706 inequality constraints. All moment and localizing matrices are of size 3×3.

Currency Basket

A currency basket is simply a way to determine the value of a national currency by calculating the weighted average of exchange rates of selected foreign currencies. These objects became popular in 1971 after the abolition of the gold standard. Options on currency baskets are attractive tools for multinational corporations to manage exposure to multiple currencies. Consider the following currency basket option on Euro and British Pounds in US Dollars. For both EUR/USD and GBP/USD two options are observable in the form (strike, price):

- EUR/USD: $\{(135.5, 2.77), (138.5, 1.17)\}$
- GBP/USD: $\{(116, 2.21), (119, 0.67)\}$

We choose the weights $(2/3, 1/3)$ for the objective function, i.e. $\varphi(\mathbf{x}) = \max(0, 2/3x_1 + 1/3x_2 - k)$ and we compute bounds for different values of k. We obtain an optimization problem similar to (4.8). The optimal values for the first level of the hierarchy are shown in Table 4.4. In the corresponding optimization

k	100	105	110	115	120
lower bound on price	1.4933	1.2599	1.0266	0.7933	0.56
computation time [s]	0.22	0.23	0.22	0.20	0.16
upper bound on price	31.5834	26.5833	21.5833	16.5833	11.5833
computation time [s]	0.15	0.16	0.18	0.18	0.16

Table 4.4: Optimal lower and upper bounds for a currency basket option with different strikes for level $r = 1$

problem the domain is partitioned into 14 sets, for each of which 15 moment variables are introduced. In total there are $14 \times 15 = 210$ variables, 80 LMIs each involving a matrix of size 3×3, 211 inequality constraints and 5 equality constraints. For this particular example, it is clear that the bounds are not very useful in practice. This is, however, not due to our approach but to the number of data points given. If more data is available, the bounds improve.

Example from Boyle and Lin [BL97]

In this example we compute bounds for a different type of option. We assume we only have data like mean, variance and correlation of the assets under the risk-neutral pricing measure is available, instead of observable option prices with different strikes. The type of option is specified through the payoff function, which will be given by $\max(0, \max(x_1, \ldots, x_n) - k)$ in this case. This type of option is called *call on max*. It is based on n assets $S_1, \ldots, S_n$, and gives the owner the right to buy the asset which at maturity is the most valuable for the predetermined strike K.

The data in the following example is taken from Boyle and Lin [BL97], where they introduced a different method to compute upper bounds. Consider three assets with means $(44.21, 44.21, 44.21)$ and the covariance matrix given by

$$C = \begin{bmatrix} 184.04 & 164.88 & 164.88 \\ 164.88 & 184.04 & 164.88 \\ 164.88 & 164.88 & 184.04 \end{bmatrix}.$$

Then, in our setting, the smallest upper bound on the price on the call on max option on these three assets is the optimal value of the following optimization problem:

$$\begin{aligned}
\sup_{\mu\in\mathcal{M}(\mathbb{R}^3_+)_+} & \int_{\mathbb{R}^3_+} \max(0,\max(x_1,x_2,x_3)-k)\,\mathrm{d}\mu(\mathbf{x}) \\
\text{s.t.} & \int_{\mathbb{R}^3_+} x_i\,\mathrm{d}\mu(\mathbf{x}) = 44.21\,, \text{for } i=1,2,3 \\
& \int_{\mathbb{R}^3_+} (x_i-44.21)(x_j-44.21)\mathrm{d}\mu(\mathbf{x}) = C_{i,j}\,, \text{for } i,j=1,2,3 \\
& \int_{\mathbb{R}^3_+} \|\mathbf{x}\|_2^2\mathrm{d}\mu(\mathbf{x}) \le M \\
& \int_{\mathbb{R}^3_+} \mathrm{d}\mu(\mathbf{x}) = 1.
\end{aligned} \tag{4.10}$$

The upper and lower bounds we obtain for different strikes

$$k \in \{30, 35, 40, 45, 50\}$$

are given in Table 4.5 as well as the bounds obtained by Boyle and Lin. Since all constraint functions are polynomial the only function contributing to the partition is the objective $\max\{0, \max(x_1, x_2, x_3) - k\}$. The resulting partition consists of 4 sets. To solve the first level of the hierarchy we introduce $4 \times 35 = 140$ moment variables. The final problem has 13 equality constraints, 61 inequality constraints and 22 LMIs, each involving a matrix of size 4×4. As in the previous example, the weakness of the bound is due to the fact that not enough information is available and is not inherent to the approach. Note that in their paper, Boyle and Lin only give a procedure for upper bounds. Also, in the original reference Boyle and Lin include a discount factor of $\exp(-0.1)$ to account for an assumed risk free interest rate. This has no effect on the optimization problem, they simply multiply their solution by the discount factor in the end.

Basket option on tech stocks

As a last example we consider four different tech stocks, namely Apple Inc. (AAPL), Meta Platforms, Inc. (FB), Nvidia Corporation (NVDA), Qualcomm Incorporated (QCOM). Suppose one wants to price a basket option on these given the data provided in Table 4.6 with payoff function $\max(0, \frac{1}{4}(x_1 + \cdots + x_4) - k)$, where the x_i are the prices of the stocks of the given companies. The bounds obtained

k	30	35	40	45	50
Boyle & Lin [BL97]	21.51	17.17	13.2	9.84	7.3
upper bound on price	21.51	17.17	13.2	9.84	7.3
computation time [s]	0.02	0.01	0.01	0.02	0.02
lower bound on price	14.21	9.21	4.21	0	0
computation time [s]	0.02	0.02	0.01	0.01	0.01

Table 4.5: Revisiting an example from Boyle and Lin, computing bounds on prices of a basket options given means and covariance of the underlying assets for different strikes.

Company	AAPL	FB	NVDA	QCOM
Option 1	(120, 45.2)	(155, 52.7)	(175, 57.9)	(130, 35.35)
Option 2	(130, 35.7)	(170, 38.5)	(180, 53.2)	(145, 20.5)
Option 3	(145, 21.75)	(180, 29.85)	(190, 43.85)	(157.5, 8.8)
Option 4	(160, 9.1)	(190, 22)	(195, 39.35)	(167.5, 2.32)
Option 5	(170, 3.35)	(200, 14.75)	(227.5, 10.75)	(175, 0.47)

Table 4.6: Strikes and corresponding prices for European call options observed on March 1st 2022, all prices in USD.

by solving the first level of the hierarchy for different strike prices are shown in Table 4.7. We set $B = 400$ and $M = 200\,000$. For this problem with $k = 140$, the partition consisted of 1938 sets, for each of which we introduce 70 variables, making 135 660 variables, 135 661 inequality constraints, 21 equality constriants and 18 726 LMIs, each involving a 5×5 matrix. It is clear that the size of the partition necessary to compute these bound grows exponentially in the number of assets considered. Even though for low levels of the hierarchy the involved matrices are very small, size of the partition is the limiting factor in the computations, since for every subset we need to introduce moment variables and at least n LMIs. Note that changing k may slightly change the number of partitions.

4.5 Relaxations of the non-compact case

We will consider another approach to approximate the solutions to problems like (4.1), but without assuming the underlying set is compact.

k	140	150	160	170	180	190	200
upper bound on price	52.79	42.89	33.48	24.53	15.68	8.51	6.99
computation time [s]	32.38	43.10	40.92	42.89	44.11	46.46	39.44
lower bound on price	46.26	36.26	26.27	16.28	6.28	0.0	0.0
computation time [s]	47.16	50.48	53.89	52.67	57.01	45.74	37.70

Table 4.7: Bounds for basket options on tech firms subject to observable data of Table 4.6

4.5.1 Lasserre hierarchy of inner range

The one considered in this section, known as the Lasserre measure-based hierarchy of inner bounds introduced by Lasserre [Las11], consists of fixing a reference measure ν on $\mathbb{R}^n_+$ such that $\nu(\mathbb{R}^n_+) < \infty$ and then approximating the density function of the optimal measure μ for (4.1) by SOS polynomials $h_r(\mathbf{x}) \in \Sigma[\mathbf{x}]_r$, such that $\mathrm{d}\mu(\mathbf{x}) = h_r(\mathbf{x})\mathrm{d}\,\nu(\mathbf{x})$. This has the advantage that instead of searching for an optimal measure in the infinite dimensional cone $\mathcal{M}(\mathbb{R}^n_+)_+$ we optimize over the set of sums of squares of fixed degree, which can be done with SDP techniques. For recent result on the convergence behavior of these measure-based hierarchies in the compact case, see, e.g. [LS22b] and [LS21]. Opposed to before, the cone of measures $\mathcal{M}(\mathbb{R}^n_+)_+$ is here approximated from inside, while before, we used an outer approximation. A possible choice for the reference measure is

$$\mathrm{d}\,\nu(\mathbf{x}) = \exp\left(-\sum_{i=1}^{n} x_i\right)\mathrm{d}\mathbf{x}\,.$$

An important assumption on the reference measure is that its moments must be available in closed form or efficiently computable. In the case above the moments are given by $\int_{\mathbb{R}^n_+} \mathbf{x}^\alpha d\,\nu(\mathbf{x}) = \alpha!$. The level r relaxation of problem (4.1) can be

formulated as follows

$$\begin{aligned}
\inf_{h_r \in \Sigma[\mathbf{x}]_r} & \int_{\mathbb{R}^n_+} \varphi(\mathbf{x}) h_r(\mathbf{x}) \mathrm{d}\nu(\mathbf{x}) \\
\text{s.t.} & \int_{\mathbb{R}^n_+} f_{i,j}(\mathbf{x}) h_r(\mathbf{x}) \mathrm{d}\nu(\mathbf{x}) = q_{i,j}\,, \text{ for } i \in [n], j \in [N_i] \\
& \int_{\mathbb{R}^n_+} f_\ell(\mathbf{x}) h_r(\mathbf{x}) \mathrm{d}\nu(\mathbf{x}) = p_\ell\,, \text{ for } \ell \in [m] \\
& \int_{\mathbb{R}^n_+} \|\mathbf{x}\|_2^2 h_r(\mathbf{x}) \mathrm{d}\nu(\mathbf{x}) \leq M\,.
\end{aligned} \tag{4.11}$$

This problem can be cast as an SDP. It should be noted that the above SDP might be infeasible even if the GMP has an optimal solution. As a simple example consider the following constraint for some $\alpha \in \mathbb{N}^n$

$$\int_{\mathbb{R}^n_+} \mathbf{x}^\alpha \mathrm{d}\mu(\mathbf{x}) = 0\,.$$

While the atomic Dirac delta measure δ_0 at 0 certainly satisfies this equation, there is no $r \in \mathbb{N}$ such that there is a degree r SOS polynomial density function that does. One can, however, relax the constraints slightly, by searching for an h_r such that one lands in (increasingly) close proximity of the right-hand side. Consider the following generalized moment problem

$$b_0 = \inf_{\nu \in \mathcal{P}(K)_+} \left\{ \int_{K_0} f_0(\mathbf{x}) \mathrm{d}\nu(\mathbf{x}) : \int_{K_i} f_i(\mathbf{x}) \mathrm{d}\nu(\mathbf{x}) = b_i\,, i \in [m] \right\}, \tag{4.12}$$

where $\mathcal{P}(K)_+$ is the set of probability measures on $K \subset \mathbb{R}^n$, $\mathrm{int}K \neq \emptyset$ and $K_i \subset K$ is closed for every $i = 0, 1, \ldots, m$. De Klerk et al. proved the following result in [dKKP20].

Theorem 4.9. *Let μ be a reference measure with known (or efficiently computable) moments such that the moments are finite and $\int_K x_i^{2k} \mathrm{d}\mu(\boldsymbol{x}) \leq (2k)! M$ for some $M > 0$ and all $i \in [n], k \in \mathbb{N}$. If all f_i for $i = 0, 1, \ldots, m$ are polynomials, then, as $r \to \infty$ we have*

$$\varepsilon(r) := \inf_{h \in \Sigma[\boldsymbol{x}]_r} \max_{i=0,1,\ldots,m} \left| \int_{K_i} f_i(\boldsymbol{x}) h(\boldsymbol{x}) \mathrm{d}\mu(\boldsymbol{x}) - b_i \right|$$

tends to zero ($\varepsilon(r) = o(1)$).

This means that if we fix an $\varepsilon > 0$ and relax the equality constraints to an ε neighborhood of the RHS, then we will eventually (for r large enough) find a feasible solution for the relaxation such that the optimal value is at most ε away from the true optimum. Theorem 4.9 promises convergence, but we cannot say anything about the rate at which ε goes to zero. It shall be mentioned that adding the $\varepsilon(r)$ in the relaxation does not necessarily result in the inner range of the bounds of the sought option prices, since this is basically an outer approximation of the inner range. Another way to think of it is first relaxing the equality constraints of problem (4.11) resulting in an increase of the possible range and then applying the inner approximation to the obtained optimization problem. When adding the ε_r-relaxation it is clear that we cannot expect monotonicity of the bounds, which will become apparent in the numerical results of Section 4.5.2.

4.5.2 Univariate example

Consider the following example with data taken from [BP02]:

$$\begin{aligned} \sup_{h_r \in \Sigma[x]_r} / \inf_{h_r \in \Sigma[x]_r} & \int_{\mathbb{R}_+} \max(0, x-105) h_r(x) \mathrm{d}\nu(x) \\ \text{s.t.} & \int_{\mathbb{R}_+} \max(0, x-100) h_r(x) \mathrm{d}\nu(x) = 8.375 \\ & \int_{\mathbb{R}_+} \max(0, x-110) h_r(x) \mathrm{d}\nu(x) = 1.875 \\ & \int_{\mathbb{R}_+} h_r(x) \mathrm{d}\nu(x) = 1. \end{aligned} \tag{4.13}$$

We know that the optimal lower and upper bounds for this data set are 3.375 and 5.125, respectively. To improve the numerical stability of SDP (4.13), one can use a basis which is orthogonal on $\mathbb{R}_+$ w.r.t. the measure $\mathrm{d}\nu(x) = \exp(-x)\mathrm{d}x$, namely the Laguerre basis defined by

$$L_n(x) = \sum_{i=0}^{n} \binom{n}{i} \frac{(-1)^i}{i!} x^i .$$

These polynomials form an orthogonal system for the Hilbert space given by $L^2(\mathbb{R}_+, w(x)\mathrm{d}x)$ with $w(x) = \exp(-x)$, i.e.,

$$\int_0^\infty L_n(x) L_m(x) \exp(-x) \mathrm{d}x = \begin{cases} 1, & \text{if } m = n \\ 0, & \text{otherwise.} \end{cases}$$

To implement the program we used the fact that

$$\int_k^\infty x^n \mathrm{d}\nu(x) = \exp(-k)\left(\sum_{\ell=0}^{n} \frac{n!}{\ell!} k^\ell\right) \tag{4.14}$$

and relaxed it to

$$\begin{aligned} \sup_{h_r \in \Sigma[x]_r} / \inf_{h_r \in \Sigma[x]_r} & \int_{\mathbb{R}_+} \max\left(0, x - \frac{105}{110}\right) h_r(x) \mathrm{d}\nu(x) \\ \text{s.t.} & \left| \int_{\mathbb{R}_+} \max\left(0, x - \frac{100}{110}\right) h_r(x) \mathrm{d}\nu(x) - \frac{8.375}{110} \right| \leq \varepsilon_r \\ & \left| \int_{\mathbb{R}_+} \max\left(0, x - \frac{110}{110}\right) h_r(x) \mathrm{d}\nu(x) - \frac{1.875}{110} \right| \leq \varepsilon_r \\ & \left| \int_{\mathbb{R}_+} h_r(x) \mathrm{d}\nu(x) - \frac{1}{110} \right| \leq \varepsilon_r \,. \end{aligned} \tag{4.15}$$

As a normalization step, we divided the data by 110. We indicate in Table 4.8 how the optimal values change if for level r we choose ε_r to be the smallest value such that the corresponding relaxation still has a feasible solution. In other words, decreasing ε_r in this case results in infeasibility. Observe that no monotonicity appears, which is expected because the equality constraint is relaxed. We mention that in Table 4.8 for $r \in \{6, 7\}$ MOSEK could not solve the maximization problem. However, the upper bound approximations were already reasonably accurate at the previous levels. It seems that the approach considered in Section 4.2 is superior to the one presented in this section, since there we get the optimal values of 5.125 and 3.375 for the first level of the hierarchy already. Especially, when considering the fact that increasing r quickly results in numerical problems and the problem is highly susceptible to small changes in ε_r. Additionally, it is difficult to estimate how much the ε relaxation perturbs the optimal value of the optimization problem.

4.6 Conclusion and further research

In this section we reflect on our results. The model we considered has the advantage that it combines different possibilities of using observable data. Option prices with different strikes as well as moment information like mean, (co-)variance etc.

r	2	3	4	5	6	7	∞
ε_r	0.0273	0.02525	0.022125	0.01755	0.0161	0.0161	0
upper bound	5.1279	5.1366	5.1288	5.1264	-	-	5.125
time in s	0.01	0.01	0.01	0.01	-	-	-
lower bound	5.122	5.1136	5.1221	5.1251	4.224	3.3522	3.375
time in s	0.01	0.01	0.01	0.01	0.02	0.03	-

Table 4.8: Optimal solutions for the level-r relaxation of the measure-based Lasserre hierarchy applied to the ε_r relaxation given in (4.15) for Laguerre basis with varying ε_r for $r = 2, \ldots, 7$. The ε_r are the smallest possible such that the resulting SDP still has a feasible solution.

can be taken into account, which is very useful in practice. The moment-SOS hierarchy, which was used to obtain the outer range delivers good approximations for low hierarchy levels. The method for the inner range quickly fails but in the considered cases still gave reasonable bounds. However, comparing the two, the outer range clearly outperformed the inner range.

Regarding the compactness argument it should be noted that in practice it might be prohibitive to carry out the core variety procedure in a setting with many assets and constraints.

5

Construction of approximation kernels via semidefinite programming

A classical problem in approximation theory is the approximation of a function by orthogonal polynomials. Orthogonality of polynomials may be defined in the following way. Let μ be a positive finite Borel measure supported on a compact semi-algebraic set

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

for $g_j \in \mathbb{R}[\mathbf{x}]$. We say two functions $f, g \in \mathcal{C}(\mathbf{K}) \setminus \{h \in \mathcal{C}(\mathbf{K}) : h \equiv 0 \text{ on } \mathbf{K}\}$ are orthogonal (with respect to μ), whenever

$$\langle f, g\rangle_\mu := \int_{\mathbf{K}} f(\mathbf{x})g(\mathbf{x})\mathrm{d}\mu(\mathbf{x}) = 0.$$

It is possible to define systems of pairwise orthogonal polynomials with respect to the inner product $\langle\cdot,\cdot\rangle_\mu$ so that they form a basis of $\mathcal{C}(\mathbf{K})$.

Let $\{p_\alpha\}_{\alpha\in\mathbb{N}^n}$ be a system of orthogonal polynomials with respect to a measure μ with orthogonality relation

$$\langle p_\alpha, p_\beta\rangle_\mu = c_\alpha \delta_{\alpha,\beta}, \tag{5.1}$$

for some positive constants c_α. Then, any μ-integrable function $f = \sum_{\alpha\in\mathbb{N}^n} f_\alpha \mathbf{x}^\alpha$ can be approximated within that system in the following way

$$f(\mathbf{x}) \approx \sum_{\alpha\in\mathbb{N}^n} a_\alpha p_\alpha(\mathbf{x}), \tag{5.2}$$

where

$$a_\alpha = \frac{\langle p_\alpha, f\rangle_\mu}{c_\alpha}. \tag{5.3}$$

To obtain a degree r approximation of a function f one could simply truncate the above sum leading to an approximation f_r defined by

$$f_r(x) = \sum_{\alpha\in\mathbb{N}^n_r} a_\alpha p_\alpha(\mathbf{x}), \tag{5.4}$$

with a_α as in (5.3). For non-differentiable functions these approximations usually do not behave well in practice as reviewed in [WWAF06], because of a phenomenon called *Gibbs oscillations*, see also Figure 5.1. This phenomenon occurs in the vicinity of discontinuities of the function to be approximated, near the boundary of the approximation domain, as well as close to points of non-differentiability [Car25, HH79, GS97].

Our goal in this chapter is to find a polynomial of degree r to approximate f while mitigating this effect. Another way to frame this question is as follows. We are looking for a mapping that sends f to an approximation f_r, i.e.,

$$f \mapsto \sum_{\alpha\in\mathbb{N}^n_r} b_\alpha p_\alpha =: f_r,$$

such that an error which measures the quality of the approximation is minimized. Such a mapping can be constructed in the following way. Consider a kernel $K_r(\mathbf{x},\mathbf{y}) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ given by

$$K_r(\mathbf{x},\mathbf{y}) := \sum_{\alpha\in\mathbb{N}^n_r} g_\alpha p_\alpha(\mathbf{x}) p_\alpha(\mathbf{y}), \tag{5.5}$$

for some constants $g_\alpha, \alpha \in \mathbb{N}^n_r$. Then the integral (convolution) operator defined as

$$\mathcal{K}^{(r)}(f)(\mathbf{x}) := \int_{\mathbf{K}} f(\mathbf{y}) K_r(\mathbf{x},\mathbf{y}) \mathrm{d}\mu(\mathbf{y})$$

maps any function $f = \sum_{\alpha\in\mathbb{N}^n} f_\alpha \mathbf{x}^\alpha$ to a polynomial of degree at most r. More accurately,

$$\mathcal{K}^{(r)}(f)(\mathbf{x}) = \sum_{\alpha,\beta\in\mathbb{N}^n_r} a_\alpha g_\beta \langle p_\alpha, p_\beta\rangle_\mu \, p_\beta(\mathbf{x}) = \sum_{\alpha\in\mathbb{N}^n_r} b_\alpha p_\alpha(\mathbf{x}),$$

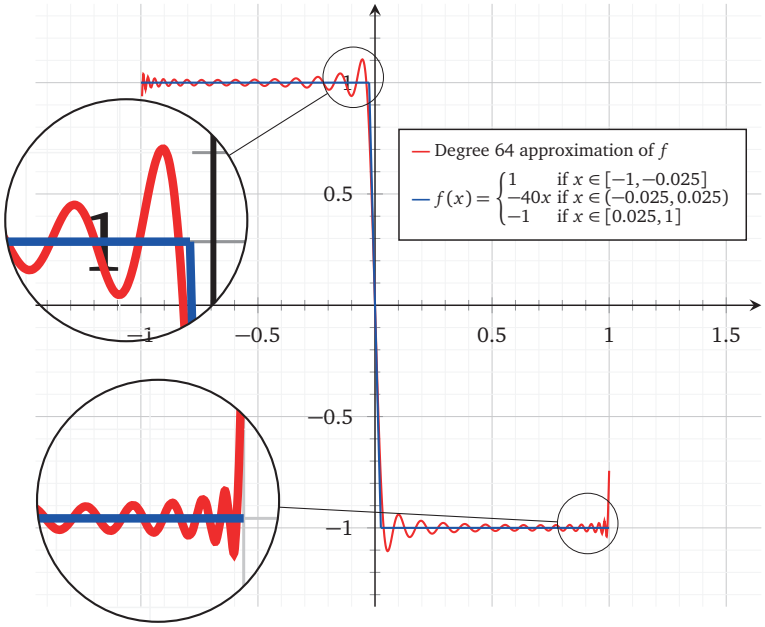


Figure 5.1: Plot of a continuous function f and its approximation via the truncation of the series (5.2) at degree $r = 64$ visualizing Gibbs oscillations.

where

$$b_\alpha = \langle p_\alpha, f \rangle_\mu \, g_\alpha.$$

The coefficients g_α of the kernel K_r determine the approximation. For example, setting $g_\alpha = 1/c_\alpha$ results in the truncation of the expansion (5.2). In the univariate case, where the orthogonal system is given by the Chebyshev polynomials of the first kind this kernel is known as the *Dirichlet kernel* [Dir29], i.e., the approximation operator used in Figure 5.1. In the multivariate case it is known as the Christoffel-Darboux kernel (named after [Chr58, Dar78]). This approximation works well for analytic functions, as reviewed in [Tre17]. Lasserre [Las21] draws an interesting connection between the celebrated moment-SOS hierarchy [Las01] and the Christoffel-Darboux kernel. In [MPW+21] the authors develop a method for approximating possibly discontinuous functions using the Christoffel-Darboux kernel, where the Gibbs phenomenon does not occur. Other approaches to get rid of unwanted oscillations is to make use of non-negative kernels as has been done in [WWAF06]. For this reason, positive approximation kernels are popular in physics for the approximation of non-smooth functions in various settings. The aim of our work is to generalize these kernels to several variables in a natural way,

thus providing computational alternatives to using products of univariate kernels.

Outline and contributions

The aim of this chapter is to present a computational procedure, based on semidefinite programming (SDP) (cf. [Tod01, BV96]), to construct non-negative polynomial kernels on $\mathbf{K} = [-1,1]^n$ suitable for approximation. We show that these kernels generalize a kernel which is called the Jackson kernel in [WWAF06], but this is different from the original kernels introduced by Jackson in [Jac12]; we give more details on this in Section 5.2.2. As the orthogonal basis we will use products of univariate Chebyshev polynomials, as reviewed in Section 5.1, and the fixed measure μ will be the corresponding product of measures so that the Chebyshev polynomials are orthogonal. The resulting *kernel polynomial method* is reviewed in Section 5.2 for the univariate case ($n = 1$), and extended to the multivariate case in Section 5.2.4. In Section 5.3 we discuss how to form the SDP problems that yield the optimal kernels, in the sense that their resolution is minimal. In Section 5.4 we show how to exploit algebraic symmetry by using techniques from [RTAL13, Val09] to reduce the size of these SDP problems. We show in Section 5.5 that our constructions are superior to simply multiplying optimal univariate kernels in a well-defined sense. Finally, in Section 5.6 we give further details of our numerical computations and show they are useful in practice to approximate non-differentiable functions and related applications in physics.

5.1 Chebyshev polynomials

In this section we review properties of univariate Chebyshev polynomials for later use. Our exposition closely follows the survey [WWAF06].

Let $\mathbf{K} = [-1,1]$ and fix the measure μ on $\mathbf{K}$ defined by

$$\mathrm{d}\mu(x) = \frac{1}{\pi\sqrt{1-x^2}}\mathrm{d}x, \quad x \in \mathbf{K}.$$

The Chebyshev polynomials of the first kind form a system of orthogonal polynomials. We will refer to the k-th Chebyshev polynomial of first kind as $T_k(x)$. We have for $k \in \mathbb{Z}$

$$T_k(x) = \cos(k \arccos(x)).$$

The following recurrence relation is well known for $m \in \mathbb{N}$

$$\begin{aligned} T_0(x) &= 1, \\ T_{-1}(x) &= T_1(x) = x, \\ T_{m+1}(x) &= 2xT_m(x) - T_{m-1}(x) \end{aligned}$$

Define for $f, g : [-1,1] \to \mathbb{R}$

$$\langle f, g \rangle_\mu = \int_{-1}^{1} \frac{f(x)g(x)}{\pi\sqrt{1-x^2}} \mathrm{d}x,$$

to obtain the following orthogonality relations for the Chebyshev polynomials of the first kind

$$\langle T_k, T_m \rangle_\mu = \frac{1+\delta_{k,0}}{2} \delta_{k,m}.$$

Chebyshev polynomials exhibit nice stability and convergence properties in practice which is why they are the first choice in many applications. It is straightforward to generalize the Chebyshev polynomials to the multivariate case. Let $\mathbf{K} = [-1,1]^n$ and define

$$\mathrm{d}\mu(\mathbf{x}) := \prod_{i=1}^{n} \frac{1}{\pi\sqrt{1-x_i^2}} \mathrm{d}\mathbf{x}.$$

Then, for $\alpha \in \mathbb{N}^n$ the corresponding multivariate Chebyshev polynomial of the first kind is defined as

$$T_\alpha(\mathbf{x}) = \prod_{i=1}^{n} T_{\alpha_i}(x_i).$$

The orthogonality relations extend in the following way

$$\begin{aligned} \langle T_\alpha, T_\beta \rangle_\mu = \int_{\mathbf{K}} T_\alpha(\mathbf{x}) T_\beta(\mathbf{x}) \mathrm{d}\mu(\mathbf{x}) &= \prod_{i=1}^{n} \int_{-1}^{1} \frac{T_{\alpha_i}(x_i) T_{\beta_i}(x_i)}{\pi\sqrt{1-x_i^2}} \mathrm{d}x_i \\ &= \prod_{i=1}^{n} \frac{1+\delta_{\alpha_i,0}}{2} \delta_{\alpha_i,\beta_i} = c_\alpha \delta_{\alpha,\beta}, \end{aligned}$$

with $c_\alpha = \left(\frac{1}{2}\right)^{H(\alpha)}$, where $H(\alpha)$ is the Hamming weight of α, i.e., the number of non-zero entries.

Convergence result

Our goal is to approximate a given continuous f defined on $\mathbf{K} = [-1,1]^n$, by a sequence of polynomials of increasing degree, such that the sequence converges to f, uniformly on $\mathbf{K}$. We further introduce a quantity

$$\sigma_r := \left(\int_{\mathbf{K}\times\mathbf{K}} \|\mathbf{x}-\mathbf{y}\|^2 K_r(\mathbf{x},\mathbf{y}) \mathrm{d}\mu(\mathbf{x}) \mathrm{d}\mu(\mathbf{y}) \right)^{1/2},$$

called the *resolution* of the kernel $K_r(\mathbf{x},\mathbf{y})$ in [WWAF06]. The resolution may be interpreted as a measure of how much mass of the kernel is concentrated away from the line where $\mathbf{x} = \mathbf{y}$. If all $g_\alpha = 1/c_\alpha$ in expression (5.5) then $\mathcal{K}^{(r)}$ is the identity operator on the space of polynomials of degree at most r, and the associated resolution is zero. This kernel will have all of its mass concentrated at $\mathbf{x} = \mathbf{y}$. To ensure uniform convergence we want a kernel that has as much mass as possible at the line $\mathbf{x} = \mathbf{y}$ for every $r \in \mathbb{N}$, while fulfilling some other properties stated below:

P1. $K_r(\mathbf{x},\mathbf{y}) = \sum_{\alpha\in\mathbb{N}^n_r} g_\alpha T_\alpha(\mathbf{x}) T_\alpha(\mathbf{y})$;

P2. $K_r(\mathbf{x},\mathbf{y}) \geq 0$ for all $(\mathbf{x},\mathbf{y}) \in \mathbf{K}\times\mathbf{K}$ and all r;

P3. $\int_{\mathbf{K}} K_r(\mathbf{x},\mathbf{y}) \mathrm{d}\mu(\mathbf{y}) = 1$ for all $\mathbf{x} \in \mathbf{K}$ for all r;

P4. $\lim_{r\to\infty} \sigma_r = 0$.

In the statement of the result, recall that the modulus of continuity of $f \in C(\mathbf{K})$ is defined as

$$\omega_f(\delta) := \max_{\substack{\mathbf{x},\mathbf{y}\in\mathbf{K} \\ \|\mathbf{x}-\mathbf{y}\|\leq\delta}} |f(\mathbf{x}) - f(\mathbf{y})|.$$

Proposition 5.1. *Let* $\mathbf{K} = [-1,1]^n$ *and* $f : \mathbf{K} \to \mathbb{R}$ *be continuous on* $\mathbf{K}$ *with modulus of continuity* ω_f. *Under the above conditions* P1-P4 *on* $K_r(\boldsymbol{x},\boldsymbol{y})$, *one has* $\mathcal{K}^{(r)}(f) \to f$ *as* $r \to \infty$, *uniformly on* $\mathbf{K}$. *Moreover,*

$$\|\mathcal{K}^{(r)}(f) - f\|_{\infty,\mathbf{K}} \leq 2\left(1 + \frac{\pi}{\sqrt{2}}\right) \omega_f(\sigma_r). \tag{5.6}$$

Our main result is the construction of kernels whose resolutions satisfy $\sigma_r = O(1/r)$ using semidefinite programming techniques (see Proposition 5.9). This proves that our kernels yield the best possible rate of convergence for continuous

f that are not differentiable, due to Bernstein's theorem (see [Ber12]). The proof of Proposition 5.1 will be postponed to later, until we have given all necessary definitions and auxiliary results. Let us mention at this point that Proposition 5.1 is a known result in approximation theory. Indeed, the argument is essentially as given in the PhD thesis of Jackson [Jac11]. We give a proof in our specific setting for completeness, since we could not find a statement of Proposition 5.1 in a suitable form in the literature.

5.2 The kernel polynomial method

We begin by considering kernels to approximate univariate functions. Let K_r be a kernel of the following form

$$K_r(x,y) = g_0 + 2\sum_{k=1}^{r} g_k T_k(x) T_k(y). \tag{5.7}$$

Kernels of this kind clearly satisfy property P1. If we set $g_0 = 1$, the resulting kernel also satisfies P3. In the following, we will explore how to find kernels that of this form that additionally satisfy P2 and are therefore suitable for approximation. For this we will first introduce trigonometric polynomials.

5.2.1 Trigonometric polynomials

A trigonometric polynomial $p(t)$ of degree r is defined as

$$p(t) = p_0 + \sum_{k=1}^{r} p_k \cos(kt) + p_{-k}\sin(kt),$$

for $p_k \in \mathbb{R}$ for $k = -r, -r+1, \ldots, r-1, r$. The following lemma is proved in [LP04], but it is a classical result.

Lemma 5.2. *If $p(t)$ is a non-negative trigonometric polynomial of degree r, then there exists a positive semidefinite matrix $Q \in \mathbb{S}_+^{r+1}$ such that $p(t) = v^\top Q v$ where*

$$v^\top = [1, \cos(t), \sin(t) \ldots, \cos(kt), \sin(kt)]$$

if $r = 2k$ is even and

$$v^\top = [\cos(t/2), \sin(t/2), \cos(t + t/2), \ldots, \cos(kt + t/2), \sin(kt + t/2)]$$

if $r = 2k+1$ is odd.

Remark 5.3. *Let us mention that there are stronger results of the kind of Lemma 5.2. For example Corollary 2 in [FP16]. We state this weaker result for the ease of exposition.*

Note that every trigonometric polynomial of the form

$$p(t) = g_0 + 2\sum_{k=1}^{r} g_k \cos(kt) \tag{5.8}$$

gives rise to a kernel of the form

$$K_r(x,y) = g_0 + 2\sum_{k=1}^{r} g_k T_k(x) T_k(y).$$

To see this, consider the following substitution

$$\begin{aligned}
&\frac{1}{2}[p(\arccos(x)+\arccos(y)) + p(\arccos(x)-\arccos(y))] \\
&\qquad = g_0 + 2\sum_{k=1}^{r} g_k \frac{1}{2}[\cos(k(\arccos(x)+\arccos(y))) \\
&\qquad\qquad + \cos(k(\arccos(x)-\arccos(y)))] \\
&\qquad = g_0 + 2\sum_{k=1}^{r} g_k \cos(k\arccos(x))\cos(k\arccos(y)) \\
&\qquad = g_0 + 2\sum_{k=1}^{r} g_k T_k(x) T_k(y).
\end{aligned} \tag{5.9}$$

If $p(t)$ is non-negative on $[-\pi,\pi]$, then $K_r(x,y)$ is non-negative on $[-1,1]^2$.

Theorem 5.4. *(Fejér (1915)) Every non-negative trigonometric polynomial of degree r of the form*

$$p(t) = \lambda_0 + \lambda_1\cos(t) + \mu_1\sin(t) + \cdots + \lambda_r\cos(rt) + \mu_r\sin(rt)$$

can be written as

$$p(t) = \left|\sum_{m=0}^{r} c_m \mathrm{e}^{\imath m t}\right|^2$$

for $c_m \in \mathbb{C}$.

In other words, there is a one-to-one correspondence between trigonometric polynomials of the form

$$t \mapsto \lambda_0 + \lambda_1 \cos(t) + \mu_1 \sin(t) + \cdots + \lambda_r \cos(rt) + \mu_r \sin(rt)$$

that are non-negative for every t and functions of the form

$$t \mapsto \left| \sum_{m=0}^{r} c_m e^{\imath m t} \right|^2 .$$

This correspondence may be leveraged to obtain kernels with minimum resolution, which is done in the next section.

5.2.2 Constructing optimal kernels

In this subsection we will revisit the approach described in [WWAF06], showing the kernel they obtain has minimum resolution among all non-negative kernels on $[-1,1]^2$. To avoid ambiguity, note the following. The authors in [WWAF06] refer to their kernel as the *Jackson kernel*, even though in the literature there is another object which is referred to in that name. Therefore, we will refer to the kernel from [WWAF06] as the *minimum resolution kernel*, reserving the term "Jackson kernel" for the object Jackson used in [Jac12] to prove his theorems. We explain the different notions of the term Jackson kernel in Appendix C.1. We are interested in non-negative trigonometric polynomials with cosine terms only, as these are the ones giving rise to kernels of the form that we want as we have seen in (5.8), (5.9). It is easy to see that if all sine-terms are zero, then the c_m terms are real. Thus, this gives us a way to characterize kernels of the form (5.7) that are non-negative. Consider a function of the form

$$p(t) = \left| \sum_{m=0}^{r} a_m e^{\imath m t} \right|^2 ,$$

for $a_m \in \mathbb{R}$. Rewriting this expression we find

$$\begin{aligned} p(t) &= \sum_{m,\ell=0}^{r} a_m a_\ell \cos([\ell - m]t) \\ &= \sum_{m=0}^{r} a_m^2 + 2 \sum_{k=1}^{r} \sum_{m=0}^{r-k} a_m a_{m+k} \cos(kt) \\ &= g_0 + 2 \sum_{k=1}^{r} g_k \cos(kt), \end{aligned}$$

for

$$g_k = \sum_{m=0}^{r-k} a_m a_{m+k}. \tag{5.10}$$

Therefore, every set of real numbers $a_0, a_1, \ldots, a_r$ satisfying $\sum_{k=0}^{r} a_k^2 = 1$ gives rise to a kernel of the form (5.7) satisfying P1, P2, P3 when the g_k are set as in (5.10). A first idea to construct such a kernel would be to set all $a_k = \frac{1}{\sqrt{r+1}}$, to ensure that $g_0 = 1$. Then we find $g_k^F = 1 - \frac{k}{r+1}$. The resulting object is known as the Fejér kernel [Fej04]. However, this kernel is not optimal in the sense that is does not have minimum resolution. We next take a look at kernels satisfying P1-P3, with minimal resolution σ_r. A simple calculation using the orthogonality of the polynomials defining K_r shows that

$$\sigma_r^2 = \int_{\mathbf{K}} (x-y)^2 K_r(x,y) \mathrm{d}\mu(x) \mathrm{d}\mu(y) = g_0 - g_1.$$

We can formulate an optimization problem to minimize resolution with respect to a_k.

$$\begin{array}{lll} \min g_0 - g_1 & \Leftrightarrow & \min \sum_{k=0}^{r} a_k^2 - \sum_{k=0}^{r-1} a_k a_{k+1} \\ \text{s.t. } g_0 = 1 & & \text{s.t. } \sum_{k=0}^{r} a_k^2 = 1. \end{array} \tag{5.11}$$

Solving this problem results in the minimum resolution kernel mentioned earlier which is called the Jackson kernel in [WWAF06], whose coefficients are given by

$$g_{k,r}^{\mathrm{KPM}} = \frac{(r-k+2)\cos\left(\frac{\pi k}{r+2}\right) + \sin\left(\frac{\pi k}{r+2}\right)\cot\left(\frac{\pi}{r+2}\right)}{r+2}. \tag{5.12}$$

Thus, we see that $\sigma_r^2 = g_{0,r}^{\mathrm{KPM}} - g_{1,r}^{\mathrm{KPM}} = 1 - \cos\left(\frac{\pi}{r+2}\right) = O(1/r^2)$. We refer to Figure 5.2 as an illustration of the absence of Gibbs' phenomenon when approximating a function using the minimum resolution kernel.

5.2.3 Uniform convergence in terms of σ_r in the multivariate case

In this subsection we prove we may also bound the rate of convergence in terms of σ_r in the multivariate case. Note that a simple calculation using the orthogonality of the chosen polynomial basis shows that σ_r can be expressed in terms of the coefficients g_α:

$$\sigma_r^2 = \sum_{i=1}^{n} (1 - g_{e_i}). \tag{5.13}$$

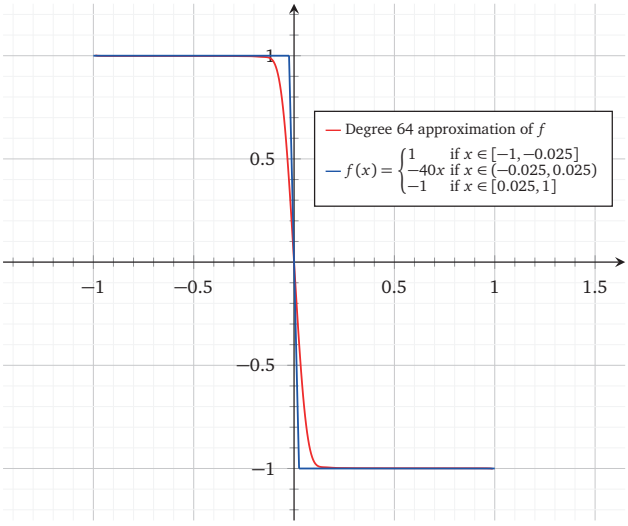


Figure 5.2: Plot of a continuous function f and its approximation via the minimum resolution kernel defined via (5.12) for degree $r = 64$ visualizing absence of Gibbs oscillations.

We first prove the result for uniformly continuous periodic functions on $[-\pi, \pi]^n$ and then extend the results for the case of uniformly continuous functions on $[-1, 1]^n$. Recall that the modulus of continuity for a uniformly continuous function f is defined as

$$\omega_f(\delta) := \max_{\substack{\mathbf{x},\mathbf{y}\in\mathbf{K} \\ \|\mathbf{x}-\mathbf{y}\|\leq\delta}} |f(\mathbf{x}) - f(\mathbf{y})|.$$

Further, note the following properties of the modulus of continuity

1. For $\lambda, \delta > 0 : \omega_f(\lambda\delta) \leq (1+\lambda)\omega_f(\delta)$ (Lemma 1.3 in [Riv69])
2. For continuous f and $\delta > 0$ one has $|f(\mathbf{x}-\mathbf{y}) - f(\mathbf{x})| \leq \left(1 + \frac{1}{\delta^2}||\mathbf{y}||^2\right)\omega_f(\delta)$ (by proof of Proposition 5.1.5 in [AC11]).

Let f be uniformly continuous and periodic on $\mathbf{B} := [-\pi, \pi]^n$ and let the degree r trigonometric kernel

$$L_r(\mathbf{x}) = 1 + \sum_{\alpha\in\mathbb{N}_r^n\setminus\{0\}} 2^{H(\alpha)} g_\alpha \prod_{i=1}^{n} \cos(\alpha_i x_i) \tag{5.14}$$

be non-negative on $\mathbf{B}$ for some set of coefficients $g_\alpha, \alpha \in \mathbb{N}_r^n \setminus \{0\}$. Define

$$\mathcal{L}^{(r)}(f)(\mathbf{x}) = \frac{1}{(2\pi)^n} \int_{\mathbf{B}} f(\mathbf{x}-\mathbf{y}) L_r(\mathbf{y}) d\mathbf{y}.$$

Further, note that

$$
\begin{aligned}
|\mathcal{L}^{(r)}(f)(\mathbf{x})-f(\mathbf{x})| &\le \frac{1}{(2\pi)^n}\int_{\mathbf{B}} |f(\mathbf{x}-\mathbf{y})-f(\mathbf{x})|L_r(\mathbf{y})\mathrm{d}\mathbf{y} \\
&\le \frac{1}{(2\pi)^n}\int_{\mathbf{B}} \left(1+\frac{1}{\delta^2}\|\mathbf{y}\|^2\right)\omega_f(\delta)L_r(\mathbf{y})\mathrm{d}\mathbf{y} \qquad (5.15)\\
&= \omega_f(\delta)+\frac{\omega_f(\delta)}{\delta^2}\frac{1}{(2\pi)^n}\int_{\mathbf{B}}\|\mathbf{y}\|^2 L_r(\mathbf{y})\mathrm{d}\mathbf{y},
\end{aligned}
$$

where the second inequality follows from the second property of the modulus of continuity. Note that since $-\pi \le y_i \le \pi$ one can check that

$$
||\mathbf{y}||^2 = \sum_{i=1}^n y_i^2 \le \frac{\pi^2}{2}\sum_{i=1}^n (1-\cos(y_i))\,.
$$

A simple calculation then shows

$$
\begin{aligned}
\int_{\mathbf{B}}\|\mathbf{y}\|^2 Lr(\mathbf{y})\mathrm{d}\mathbf{y} &\le \frac{\pi^2}{2}\left(\sum_{i=1}^n (2\pi)^n - \sum_{\alpha\in\mathbb{N}_r^n} 2^{H(\alpha)} g_\alpha \int_{\mathbf{B}} \cos(y_i)\prod_{i=1}^n \cos(\alpha_i y_i)\mathrm{d}\mathbf{y}\right)\\
&= (2\pi)^n \frac{\pi^2}{2}\sum_{i=1}^n \left(1-g_{e_i}\right),
\end{aligned}
$$

where we used the fact that

$$
\cos(x)\cos(kx) = \frac{1}{2}\left(\cos((k-1)x)+\cos((k+1)x)\right).
$$

Recalling identity (5.13) and choosing

$$
\delta = \sqrt{\frac{\pi^2}{2}\sum_{i=1}^n (1-g_{e_i})} = \pi/\sqrt{2}\sigma_r
$$

we find by (5.15) and the first property of the modulus of continuity:

$$
|\mathcal{L}^{(r)}(f)(\mathbf{x})-f(\mathbf{x})| \le 2\omega_f(\pi/\sqrt{2}\sigma_r) \le 2(1+\pi/\sqrt{2})\omega_f(\sigma_r).
$$

We next deal with the case where f is a continuous function on $[-1,1]^n$. Define $g(\theta) = f(\cos(\theta)) = f(\cos(\theta_1),\ldots,\cos(\theta_n))$ for $\theta \in [0,\pi]^n$. Further, let $g(\theta) = g(\theta_1,\ldots,-\theta_i,\ldots,\theta_n)$ for $\theta_i \in [-\pi,0]$. Similarly, we may define $g(\theta)$ for all $\theta \in [-\pi,\pi]^n$. We see $g(\theta)$ is even and periodic on $[-\pi,\pi]^n$. Since it is even, the

convolution with a trigonometric kernel of the form (5.14) will have only cosine terms. The argument is as follows:

$$\begin{aligned}(2\pi)^n \mathcal{L}^{(r)}(g)(\theta) &= \int_{\mathbf{B}} g(\theta-\varphi)L_r(\varphi)\mathrm{d}\varphi \\ &= \int_{\mathbf{B}} g(\varphi)L_r(\theta-\varphi)\mathrm{d}\varphi \\ &= \int_{\mathbf{B}} g(\varphi)\left(1+\sum_{\alpha\in\mathbb{N}_r^n\setminus\{0\}} 2^{H(\alpha)} g_\alpha \prod_{i=1}^n \cos\left(\alpha_i(\theta_i-\varphi_i)\right)\right)\mathrm{d}\varphi. \\ &= \int_{\mathbf{B}} g(\varphi)\mathrm{d}\varphi + \\ &\sum_{\alpha\in\mathbb{N}_r^n\setminus\{0\}} 2^{H(\alpha)} g_\alpha \int_{\mathbf{B}} g(\varphi)\prod_{i=1}^n \left(\cos(\alpha_i\theta_i)\cos(\alpha_i\varphi_i)+\sin(\alpha_i\theta_i)\sin(\alpha_i\varphi_i)\right)\mathrm{d}\varphi.\end{aligned}$$

The integrand in the last integral is the function g times the sum of products of sine and cosine functions. The domain $\mathbf{B} = [-\pi, \pi]^n$ is symmetric and g is even by construction. Hence, every integral containing a sine function will evaluate to zero, since the sine function is odd. We may therefore assume the approximation will take the following form

$$q_r(\theta) = \mathcal{L}^{(r)}(g)(\theta) = a_0 + \sum_{\alpha\in\mathbb{N}_r^n\setminus\{0\}} 2^{H(\alpha)} g_\alpha a_\alpha \prod_{i=1}^n \cos\left(\alpha_i\theta_i\right). \tag{5.16}$$

Substituting $\theta_i = \arccos(x_i)$ results in a polynomial

$$p_r(\mathbf{x}) = a_0 + \sum_{\alpha\in\,\mathbb{N}_r^n} 2^{H(\alpha)} g_\alpha a_\alpha T_\alpha(\mathbf{x}). \tag{5.17}$$

This polynomial will serve as an approximation for f, and we will bound the absolute error in terms of σ_r. For this we will need the following lemma.

Lemma 5.5. *For f, g as defined above we have*

$$\omega_g(\delta) \le \omega_f(\delta).$$

Proof. Note that

$$\begin{aligned}\omega_g(\delta) &= \sup_{\substack{\theta_1,\theta_2\in[-\pi,\pi]^n\\ \|\theta_1-\theta_2\|\le\delta}} |g(\theta_1)-g(\theta_2)|\\ &= \sup_{\substack{\theta_1,\theta_2\in[0,\pi]^n\\ \|\theta_1-\theta_2\|\le\delta}} |f(\cos\theta_1)-f(\cos\theta_2)|\\ &= \sup_{\substack{\mathbf{x},\mathbf{y}\in[-1,1]^n\\ \|\arccos\mathbf{x}-\arccos\mathbf{y}\|\le\delta}} |f(\mathbf{x})-f(\mathbf{y})|\\ &\le \sup_{\substack{\mathbf{x},\mathbf{y}\in[-1,1]^n\\ \|\mathbf{x}-\mathbf{y}\|\le\delta}} |f(\mathbf{x})-f(\mathbf{y})| = \omega_f(\delta),\end{aligned}$$

where we set $\arccos\mathbf{x} = (\arccos(x_1),\dots,\arccos(x_n))$ to shorten the notation. Note that we used $|\cos(x)-\cos(y)| \le |x-y|$. □

We have gathered everything required to prove Proposition 5.1.

Proof of Proposition 5.1. First, note that

$$\begin{aligned}\sup_{\mathbf{x}\in[-1,1]^n} |f(\mathbf{x})-p_r(\mathbf{x})| &\le \sup_{\theta\in[-\pi,\pi]^n} |g(\theta)-q_r(\theta)|\\ &\le 2(1+\pi/\sqrt{2})\omega_g(\sigma_r)\\ &\le 2(1+\pi/\sqrt{2})\omega_f(\sigma_r),\end{aligned}$$

where the last inequality follows by Lemma 5.5. Above we have obtained a polynomial p_r of degree less than r which approximates $f \in \mathcal{C}(\mathbf{K})$. We did so by using the convolution with a trigonometric kernel of the form (5.14). Using the substitution given in (5.9) we may transform the trigonometric kernel into a positive polynomial kernel of the form

$$K_r(\mathbf{x},\mathbf{y}) = 1 + \sum_{\alpha\in\mathbb{N}^n_r} 2^{H(\alpha)} g_\alpha T_\alpha(\mathbf{x}) T_\alpha(\mathbf{y}). \tag{5.18}$$

It is left to show that both approaches are equivalent, i.e., lead to the same approximation of the function f. For this note the following. The polynomial we obtain via the approximation process defined as per (5.18) is

$$\begin{aligned}\mathcal{K}^{(r)}(f)(\mathbf{x}) &= \int_{\mathbf{K}} f(\mathbf{y})\mathrm{d}\mu(\mathbf{y}) + \sum_{\alpha\in\mathbb{N}^n_r} 2^{H(\alpha)} g_\alpha \left(\int_{\mathbf{K}} f(\mathbf{y}) T_\alpha(\mathbf{y})\mathrm{d}\mu(\mathbf{y})\right) T_\alpha(\mathbf{x})\\ &= c_0 + \sum_{\alpha\in\mathbb{N}^n_r} 2^{H(\alpha)} g_\alpha c_\alpha T_\alpha(\mathbf{x}),\end{aligned}$$

where

$$c_\alpha = \langle f, T_\alpha \rangle_\mu = \int_{\mathbf{K}} f(\mathbf{y}) T_\alpha(\mathbf{y}) \mathrm{d}\mu(\mathbf{y}).$$

We need to check whether we get the same coefficients from both approaches. Recall the approximation from (5.17):

$$p_r(\mathbf{x}) = a_0 + \sum_{\alpha \in \mathbb{N}^n_r} 2^{H(\alpha)} g_\alpha a_\alpha T_\alpha(\mathbf{x}),$$

where

$$a_\alpha = \frac{1}{(2\pi)^n} \int_{\mathbf{B}} g(\varphi) \prod_{i=1}^n \cos(\alpha_i \varphi_i) \mathrm{d}\varphi.$$

We would like to show that $a_\alpha = c_\alpha$, i.e.,

$$\int_{\mathbf{K}} f(\mathbf{y}) \prod_{i=1}^n \frac{\cos(\alpha_i \arccos(y_i))}{\sqrt{1-y_i^2}} \mathrm{d}\mathbf{y} = \frac{1}{2^n} \int_{\mathbf{B}} g(\varphi) \prod_{i=1}^n \cos(\alpha_i \varphi_i) \mathrm{d}\varphi.$$

For this note that

$$a_\alpha = \frac{1}{2^n} \int_{\mathbf{B}} g(\varphi) \prod_{i=1}^n \cos(\alpha_i \varphi_i) \mathrm{d}\varphi = \int_{[0,\pi]^n} f(\cos(\varphi)) \prod_{i=1}^n \cos(\alpha_i \varphi_i) \mathrm{d}\varphi. \quad (5.19)$$

Next we can make use of the following substitution

$$\varphi_i = \arccos(x_i) \Rightarrow \mathrm{d}\varphi_i = -\frac{1}{\sqrt{1-x_i^2}} \mathrm{d}x_i.$$

Finally, using (5.19) we find

$$a_\alpha = \int_{[0,\pi]^n} f(\cos(\varphi)) \prod_{i=1}^n \cos(\alpha_i \varphi_i) \mathrm{d}\varphi = \int_{\mathbf{K}} f(\mathbf{x}) \prod_{i=1}^n \frac{\cos(\alpha_i \arccos(x_i))}{\sqrt{1-x_i^2}} \mathrm{d}\mathbf{x} = c_\alpha.$$

This proves we indeed have that both approaches are equivalent, i.e., they lead to the same approximation. This completes the proof of Proposition 5.1. □

5.2.4 Positivstellensatz for the multivariate case

In this subsection we present a classical Positivstellensatz for multivariate trigonometric polynomials. This result allows for the construction of semidefinite programs whose solutions provide optimal kernels with respect to σ_r. To this end, recall the identities: if $p_k(\phi) = \cos(k\phi)$, then, for $x, y \in [-1, 1]$,

$$\begin{aligned} p_k(\arccos(x)) &= T_k(x) \\ \frac{1}{2}\left(p_k(\arccos(x) + \arccos(y)) + p_k(\arccos(x) - \arccos(y))\right) &= T_k(x) T_k(y). \end{aligned}$$

As a consequence, if we start with a non-negative multivariate trigonometric polynomial of the form

$$(\phi_1, \ldots, \phi_n) \mapsto 1 + \sum_{\alpha \in \mathbb{N}_r^n \setminus \{0\}} 2^{H(\alpha)} g_\alpha \prod_{i \in [n]} \cos(\alpha_i \phi_i),$$

then replacing each $p_{\alpha_i}(\phi_i) := \cos(\alpha_i \phi_i)$ by

$$\frac{1}{2}\left(p_{\alpha_i}(\arccos(x_i) + \arccos(y_i)) + p_{\alpha_i}(\arccos(x_i) - \arccos(y_i))\right)$$

as above (this operation preserves non-negativity), one obtains the non-negative kernel

$$K(\mathbf{x}, \mathbf{y}) = 1 + \sum_{\alpha \in \mathbb{N}_r^n \setminus \{0\}} 2^{H(\alpha)} g_\alpha \prod_{i \in [n]} T_{\alpha_i}(x_i) T_{\alpha_i}(y_i) \quad \mathbf{x}, \mathbf{y} \in [-1, 1]^n.$$

In contrast to the polynomial case, each multivariate, positive, trigonometric polynomial is a sum of squares of trigonometric polynomials. (The degrees appearing in the sums-of-squares may be arbitrarily large, though.)

Theorem 5.6 (e.g. Theorem 3.5 in [Dum07])**.** *If p is a positive trigonometric polynomial, then there exists an $k \in \mathbb{N}$ and a hermitian p.s.d. matrix M of order $\binom{n+k}{k}$ such that*

$$p(\phi) = \left[\exp(\imath \alpha^\top \phi)\right]^*_{\alpha \in \mathbb{N}_k^n} M \left[\exp(\imath \alpha^\top \phi)\right]_{\alpha \in \mathbb{N}_k^n}.$$

Again, the value k in the theorem may be arbitrarily large, but for fixed $k \geq r$, one may consider the kernel given by solving the SDP:

$$\sigma_r^2 = \min \sum_{i=1}^n 1 - g_{e_i} \tag{5.20}$$

subject to

$$\begin{aligned} 1 + \sum_{\alpha \in \mathbb{N}_r^n \setminus \{0\}} 2^{H(\alpha)} g_\alpha \prod_{i \in [n]} \cos \alpha_i \phi_i &= \left[\exp(\imath \alpha^\top \phi)\right]^*_{\alpha \in \mathbb{N}_k^n} M \left[\exp(\imath \alpha^\top \phi)\right]_{\alpha \in \mathbb{N}_k^n} \\ M &\succeq 0. \end{aligned}$$

In what follows we will focus on the first level of the hierarchy, that is, $r = k$. We discuss the case where $k > r$ in Section 5.6.1.

5.3 Reformulations of the SDP

In this section we present two ways to find numerical solutions to problem (5.20). But first we show that the optimal solution to the problem maybe assumed w.l.o.g. to be real.

5.3.1 Existence of a real solution

Let $M = P + \imath Q$, for $P \in \mathbb{S}^{s(n,r)}$ and Q being skew-symmetric of the same size. Then,

$$\left[\exp(\imath\alpha^\top\phi)\right]^*_{\alpha\in\mathbb{N}^n_r}(P+\imath Q)\left[\exp(\imath\alpha^\top\phi)\right]_{\alpha\in\mathbb{N}^n_r} = \sum_{\alpha,\beta\in\mathbb{N}^n_r}\cos\left[(\alpha-\beta)^\top\phi\right]P_{\alpha,\beta} - \sum_{\alpha,\beta\in\mathbb{N}^n_r}\sin\left[(\alpha-\beta)^\top\phi\right]Q_{\alpha,\beta},$$

because of the identities

$$\cos(\alpha^\top\phi)\cos(\beta^\top\phi)+\sin(\alpha^\top\phi)\sin(\beta^\top\phi)=\cos\left[(\alpha-\beta)^\top\phi\right]$$

and

$$\cos(\alpha^\top\phi)\sin(\beta^\top\phi)-\sin(\alpha^\top\phi)\cos(\beta^\top\phi)=\sin\left[(\alpha-\beta)^\top\phi\right].$$

We continue to show we may without loss of generality assume there exists an optimal solution which is real symmetric, i.e., $Q = 0$. Let $P + \imath Q$ be an optimal solution to (5.20) and define

$$\begin{aligned} G(\phi_1,\ldots,\phi_n) &= \sum_{\gamma\in\mathbb{N}^n_r} 2^{H(\gamma)} g_\gamma \prod_{i\in[n]}\cos(\gamma_i\phi_i) \\ &= \sum_{\alpha,\beta\in\mathbb{N}^n_r}\cos\left[(\alpha-\beta)^\top\phi\right]P_{\alpha,\beta} - \sum_{\alpha,\beta\in\mathbb{N}^n_r}\sin\left[(\alpha-\beta)^\top\phi\right]Q_{\alpha,\beta}. \end{aligned}$$

Now, since the cosine is even we find $G(\phi_1,\ldots,\phi_n) = G(-\phi_1,\ldots,-\phi_n)$. So

$$\begin{aligned} G(\phi_1,\ldots,\phi_n) &= \frac{1}{2}(G(\phi_1,\ldots,\phi_n)+G(-\phi_1,\ldots,-\phi_n)) \\ &= \frac{1}{2}\Big(\sum_{\alpha,\beta\in\mathbb{N}^n_r}\cos\left[(\alpha-\beta)^\top\phi\right]P_{\alpha,\beta} - \sum_{\alpha,\beta\in\mathbb{N}^n_r}\sin\left[(\alpha-\beta)^\top\phi\right]Q_{\alpha,\beta} \\ &\quad + \sum_{\alpha,\beta\in\mathbb{N}^n_r}\cos\left[(\beta-\alpha)^\top\phi\right]P_{\alpha,\beta} - \sum_{\alpha,\beta\in\mathbb{N}^n_r}\sin\left[(\beta-\alpha)^\top\phi\right]Q_{\alpha,\beta}\Big) \\ &= \sum_{\alpha,\beta\in\mathbb{N}^n_r}\cos\left[(\alpha-\beta)^\top\phi\right]P_{\alpha,\beta}. \end{aligned}$$

Therefore, we may subsequently assume that in (5.20) the matrix M is real symmetric.

5.3.2 Sampling based formulation

The first strategy for the implementation is based on sampling, a concept introduced by Löfberg and Parrilo in [LP04]. Instead of equating coefficients, the idea is to equate evaluations of two representations of a (trigonometric) polynomial on a finite number of discrete sampling points. In our problem at hand we are looking for a real symmetric psd matrix M such that

$$1+\sum_{\alpha\in\mathbb{N}_r^n\setminus\{0\}} 2^{H(\alpha)} g_\alpha \prod_{i\in[n]} \cos(\alpha_i\phi_i) = \sum_{\alpha,\beta\in\mathbb{N}_r^n} \cos\left[(\alpha-\beta)^\top \phi\right] M_{\alpha,\beta}, \qquad (5.21)$$

for all $\phi\in[0,2\pi]^n$. We discretize the set of values for $\phi\in[0,2\pi]^n$ and enforce equality after evaluating the cosine terms for these sample points. Consider the set

$$\Phi=\left\{\varphi\in[0,2\pi]^n : \varphi_i=\frac{k_i\pi}{(r+2)} \text{ for } (k_1,\ldots,k_n)\in\{0,1,\ldots,2(r+2)\}^n\right\},$$

which has $(2(r+2)+1)^n$ elements. This set may be seen as uniformly distributed points on the unit sphere of dimension $n-1$. Then we can formulate an SDP as follows

$$\sigma_r^2 = \min \sum_{i=1}^n 1-g_{e_i} \qquad (5.22)$$

subject to

$$\begin{aligned}
\sum_{\gamma\in\mathbb{N}_r^n} 2^{H(\gamma)} g_\gamma \prod_{i\in[n]} \cos(\gamma_i\varphi_i) &= \sum_{\alpha,\beta\in\mathbb{N}_r^n} \cos\left[(\alpha-\beta)^\top\varphi\right] M_{\alpha,\beta} \quad \forall\varphi\in\Phi\\
g_\gamma &= 1 \text{ for } \gamma=(0,\ldots,0)\\
M &\succeq 0.
\end{aligned}$$

Note that the number of sample points is large enough, so that the optimal value of the above program is equal to σ_r^2. In practice this approach works well for small instances, i.e., for $n=2, r\le 8$. For larger r the solver we used (MOSEK [MOS19]) terminated because of slow progress. It did, however, return solutions that were close to being optimal. Using other solvers resulted in a significant increase in running time.

5.3.3 Equating coefficients via trigonometric identities

The other approach is using the standard idea of SOS-optimization, i.e., equating coefficients. The following lemma is essential for this step.

Lemma 5.7. *Let for $I \subseteq [n]$ the function $\omega_I : \mathbb{R}^n \to \mathbb{R}^n$ be defined as follows*

$$\omega_I(\boldsymbol{x})_i = \begin{cases} -x_i, \text{ if } i \in I \\ x_i, \text{ otherwise.} \end{cases}$$

The following trigonometric identity holds for all $n \in \mathbb{N}$

$$2^{H(\boldsymbol{x})} \prod_{i=1}^{n} \cos(x_i) = \sum_{I \subseteq [n]} \cos\left(\sum_{i=1}^{n} \omega_I(\boldsymbol{x})_i \right). \tag{5.23}$$

Proof. Note that in other words, ω_I flips the sign of x_i for all $i \in I$. We will prove the statement by induction on n. Let $\mathbf{x} \in \mathbb{R}^n$ and assume without loss of generality that the support of $\mathbf{x}$ is n. For $n = 0, 1$ the identity is obvious. For the induction step consider the identity for $n \leftarrow n+1$. Let $\tilde{\mathbf{x}} = (\mathbf{x}, x_{n+1}) \in \mathbb{R}^{n+1}$ with full support. Then,

$$\begin{aligned} 2^{n+1} \prod_{i=1}^{n+1} \cos(x_i) &= 2\cos(x_{n+1}) \left(2^n \prod_{i=1}^{n} \cos(x_i) \right) \\ &\overset{(i)}{=} \sum_{I \subset [n]} \cos\left(\sum_{i=1}^{n} \omega_I(\mathbf{x})_i \right) (\cos(x_{n+1}) + \cos(-x_{n+1})) \\ &\overset{(ii)}{=} \sum_{I \subset [n]} \cos\left(\sum_{i=1}^{n} \omega_I(\mathbf{x})_i + x_{n+1} \right) + \sum_{I \subset [n]} \cos\left(\sum_{i=1}^{n} \omega_I(\mathbf{x})_i - x_{n+1} \right) \\ &= \sum_{I \subset [n+1]} \cos\left(\sum_{i=1}^{n+1} \omega_I(\tilde{\mathbf{x}})_i \right), \end{aligned}$$

where in (i) we used the induction assumption and in (ii) we used the well-known identity

$$\cos(x+y) = \cos(x)\cos(y) - \sin(x)\sin(y),$$

and the fact that $\sin(-x) = -\sin(x)$. □

For convenience, we restate the optimization problem below. Recall that it is enough to consider real symmetric matrices.

$$\sigma_r^2 = \min \sum_{i=1}^{n} 1 - g_{e_i} \tag{5.24}$$

subject to

$$1 + \sum_{\gamma \in \mathbb{N}_r^n \setminus \{0\}} 2^{H(\gamma)} g_\gamma \prod_{i \in [n]} \cos(\gamma_i \phi_i) = \sum_{\alpha,\beta \in \mathbb{N}_r^n} \cos[(\alpha - \beta)^\top \phi] M_{\alpha,\beta} \tag{5.25}$$

$$M \succeq 0.$$

The identity (5.23) will allow us to compare coefficients of trigonometric polynomials in (5.24). On the right-hand side in (5.25) we will find all $\zeta \in \mathbb{Z}^n$ for which there exist $\alpha, \beta \in \mathbb{N}_r^n$, such that $\zeta = \alpha - \beta$. The identity (5.23) now tells us that we have to make sure that for a given $\gamma \in \mathbb{N}_r^n$ the following holds

$$\sum_{\substack{\alpha,\beta \in \mathbb{N}_r^n \\ \alpha - \beta = \gamma}} M_{\alpha,\beta} = \sum_{\substack{\alpha,\beta \in \mathbb{N}_r^n \\ \alpha - \beta = \omega_I(\gamma)}} M_{\alpha,\beta} \quad \text{for all } I \subset [n],$$

since then we can factor out the same sum for each $\omega_I(\gamma)$ and apply identity (5.23). Noting that $\alpha - \beta = -(\beta - \alpha)$ we can construct symmetric constraint matrices. For each $\gamma \in \mathbb{N}_r^n$ let $C^{(\gamma,I)} \in \{0,1\}^{s(n,r) \times s(n,r)}$

$$C_{\alpha,\beta}^{(\gamma,I)} = \begin{cases} 1, \text{ if } \alpha - \beta = \omega_I(\gamma) \vee \omega_{I^c}(\gamma) \\ 0, \text{ otherwise.} \end{cases}$$

These matrices will always be symmetric since if $\alpha - \beta = \omega_I(\gamma)$ then $\beta - \alpha = \omega_{I^c}(\gamma)$. We define $\mathcal{I}$ as a set of subsets of $[n]$ such that no complement I^c of a set $I \in \mathcal{I}$ lies in $\mathcal{I}$ and $\cup_{I \in \mathcal{I}} \{I, I^c\} = \{I : I \subseteq [n]\}$. With this we can formulate the first set of constraints, i.e.,

$$\langle M, C^{(\gamma,\emptyset)} \rangle = \langle M, C^{(\gamma,I)} \rangle \ \forall I \in \mathcal{I}, \forall \gamma \in \mathbb{N}_r^n.$$

Then $g_\gamma = \frac{1}{2} \langle M, C^{(\gamma,I)} \rangle$ for any $I \in \mathcal{I}$. Additionally, we need the following. Let

$$\Gamma_{(n,r)} = \left\{ \zeta \in \mathbb{Z}^n : \exists \alpha, \beta \in \mathbb{N}_r^n, \alpha - \beta = \zeta \wedge \sum_{i=1}^{n} |\zeta_i| > r \right\},$$

which leads us to the next set of constraints

$$\langle M, C^{(\zeta,I)} \rangle = 0 \ \forall I \in \mathcal{I}, \forall \zeta \in \Gamma_{(n,r)}.$$

This way we ensure that there will not appear any unwanted terms in the resulting polynomial. Any $\zeta \in \Gamma_{(n,r)}$ will not find the necessary pairs to use identity (5.23). Therefore, we force all such terms to be zero. We can equivalently formulate the SDP now as

$$\sigma_r^2 = \min \sum_{i=1}^{n} 1 - \frac{1}{2}\langle M, C^{(e_i,\emptyset)}\rangle \tag{5.26}$$

subject to

$$\begin{aligned} \langle M, C^{(\alpha,\emptyset)} - C^{(\alpha,I)}\rangle &= 0 \quad \forall I \in \mathcal{I}, \forall \alpha \in \mathbb{N}_r^n \setminus \{0\} \\ \langle M, C^{(\zeta,I)}\rangle &= 0 \quad \forall I \in \mathcal{I}, \forall \zeta \in \Gamma_{(n,r)} \\ \mathrm{Tr}(M) &= 1 \\ M &\succeq 0. \end{aligned}$$

Note that $\mathrm{Tr}(M) = 1$ ensures $g_\alpha = 1$ for $\alpha = (0,\dots,0)$. This implementation works better in practice than the previous one (5.22), especially for larger values of r.

5.4 Symmetry reduction

In this section we will present an approach that exploits existing symmetries in semidefinite programs in order to improve numerical tractability. There has been done research on the exploitation of symmetries in semidefinite programming, see, e.g., [Val09]. There are also results available focusing on symmetry exploitation for semidefinite relaxations of polynomial optimization problems for which we refer the reader to [GP04], [RTAL13]. The name readily implies that we will use some *symmetry* to *reduce* the size of the SDP. The goal is to set up an equivalent SDP, where we can impose a block diagonal structure on the matrix variable. This is helpful because we then only have to enforce the positive semidefiniteness for the individual blocks instead of the whole matrix. We present the necessary background information in Appendix B.

5.4.1 Symmetry adapted basis

Let $\mathcal{S}_n$ be the symmetric group acting on the variables x_i for $i \in [n]$ by permuting the elements, i.e.,

$$\sigma(x_i) = x_{\sigma(i)} \text{ for } i \in [n], \sigma \in \mathcal{S}_n.$$

The action of $\mathcal{S}_n$ may be defined on functions as well in the following way. Let $f : \mathbb{R}^n \to \mathbb{R}$, then

$$\sigma(f) = f(\sigma(\mathbf{x})),$$

where, if an element $\sigma \in \mathcal{S}_n$ is applied to an n-tupel we define for $\mathbf{x} \in \mathbb{R}^n$

$$\sigma(\mathbf{x}) = (\sigma(x_1), \sigma(x_2) \ldots, \sigma(x_n)) = (x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(n)}),$$

i.e., elementwise application. We will call a function f invariant under $\mathcal{S}_n$ if $\sigma(f) = f(\mathbf{x})$ for all $\sigma \in \mathcal{S}_n$. Note that our kernel is defined over the set $[-\pi, \pi]^n$, which is invariant under the action of $\mathcal{S}_n$. We can also assume without loss of generality that the optimal kernel K_r will be invariant under the action of $\mathcal{S}_n$, meaning that for all coefficients we will have

$$g_\alpha = g_{\sigma(\alpha)} \text{ for all } \sigma \in \mathcal{S}_n, \alpha \in \mathbb{N}_r^n,$$

To see that the optimal kernel will be invariant under $\mathcal{S}_n$, note that all constraints in problem (5.24) are invariant under $\mathcal{S}_n$. Thus, any optimal solution to (5.24) can be "symmetrized" using the Reynolds-operator, which is defined as

$$\mathcal{R}^{\mathcal{S}_n}(f) := \frac{1}{|\mathcal{S}_n|} \sum_{\sigma \in \mathcal{S}_n} \sigma(f).$$

Let $K_r = \sum_{\alpha \in \mathbb{N}_r^n} \tilde{g}_\alpha \prod_{i \in [n]} \cos \alpha_i \varphi_i$ be a feasible solution to (5.24), then

$$\mathcal{R}^{\mathcal{S}_n}(K_r) = \frac{1}{|\mathcal{S}_n|} \sum_{\sigma \in \mathcal{S}_n} \sum_{\alpha \in \mathbb{N}_r^n} \tilde{g}_\alpha \sigma \left(\prod_{i \in [n]} \cos \alpha_i \varphi_i \right)$$

is also feasible and will lead to the same objective value. Let $\mathbb{T}[\varphi]_r = \mathbb{T}[\varphi_1, \ldots, \varphi_n]_r$ be the set of trigonometric polynomials of degree less that r. We will define $\mathbb{T}[\varphi]_r^{\mathcal{S}_n}$ to be the set of trigonometric polynomials of degree at most r which are invariant under the action of $\mathcal{S}_n$. A basis for $\mathbb{T}[\varphi]_r$ is given by $\{\exp(\imath \alpha^\top \phi)\}_{\alpha \in \mathbb{N}_r^n}$. To exploit the symmetry we will construct a new basis $\mathcal{B}$, which we call the *symmetry adapted basis*, see Appendix B. The set $\mathcal{B}$ may be seen as a collection of $k(n,r) \in \mathbb{N}$ sub-bases $\mathcal{B}_i = \{b_{i_j} \in \mathbb{T}[\varphi]_r : \text{ for } j \in [k_i]\}$, for some $k_i \in \mathbb{N}$ in the sense that $\mathcal{B} = \{\mathcal{B}_i : i \in [k(n,r)]\}$. In general there are no closed form expressions for $k(n,r)$ and k_i available as functions of n and r, but to give some impression of these numbers we provide Table C.2 in Appendix C.2. We call $\mathcal{B}$ a basis because

$$\operatorname{span}\left\{\mathcal{R}^{\mathcal{S}_n}(b_{i_l} b_{i_m}^*),\ i \in [k(n,r)],\ l, m \in [k_i]\right\} = \mathbb{T}[\varphi]_r^{\mathcal{S}^n}.$$

The basis $\mathcal{B}$ has the property that its elements are pairwise orthogonal in the sense that for $b_{i_l} \in \mathcal{B}_i, b_{j_m} \in \mathcal{B}_j$ with $i \neq j$ the symmetrized product is zero, i.e.,

$$\mathcal{R}^{\mathcal{S}_n}(b_{i_l} b_{j_m}^*) = 0.$$

Before, we were interested in suitable kernels that could be written as

$$K_r = \left[\exp(\imath \alpha^\top \phi)\right]_{\alpha \in \mathbb{N}_r^n}^* M \left[\exp(\imath \alpha^\top \phi)\right]_{\alpha \in \mathbb{N}_r^n},$$

where $M \in \mathbb{S}_{\succeq 0}^{s(n,r)}$. Knowing that the optimal K_r is invariant under $\mathcal{S}_n$, we can write

$$K_r = \sum_{i=1}^{k(n,r)} [b_{i_j}^*]_{j \in [k_i]} M^{(i)} [b_{i_j}]_{j \in [k_i]}$$

for $M^{(i)} \succeq 0$ for all $i \in [k(n,r)]$. The pairwise orthogonality of $\mathcal{B}$ means that we can consider a block diagonal matrix

$$\begin{bmatrix} M^{(1)} & 0 & 0 & \dots & 0 \\ 0 & M^{(2)} & 0 & \dots & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & 0 & M^{(k(n,r))} \end{bmatrix}$$

with $M^{(i)} \in \mathbb{S}_{\succeq 0}^{k_i}$ in our SDP. The computational advantage is that we only have to ensure the positive semidefiniteness of the individual blocks, instead of the much larger matrix M.

Example 5.8. *Consider the following example for a symmetry adapted basis with $n = 2, r = 2$. In this case $k(n,r) = k(2,2) = 2$ and $k_1 = 4, k_2 = 2$. For a corresponding SDP we would have to consider two psd blocks of sizes 4 and 2 instead of one psd matrix of size 6×6. For $\mathcal{B}_1, \mathcal{B}_2$ we find*

$$\begin{aligned} \mathcal{B}_1 = \big\{ & \exp(\imath\,(0\varphi_1 + 0\varphi_2)) = 1, \\ & \exp(\imath\,(\varphi_1 + \varphi_2)), \\ & \exp(\imath\,\varphi_1) + \exp(\imath\,\varphi_2), \\ & \exp(\imath\,2\varphi_1) + \exp(\imath\,2\varphi_2)\big\} \end{aligned} \qquad \begin{aligned} \mathcal{B}_2 = \big\{ & \exp(\imath\,\varphi_1) - \exp(\imath\,\varphi_2), \\ & \exp(\imath\,2\varphi_1) - \exp(\imath\,2\varphi_2)\big\}. \end{aligned}$$

5.4.2 Construction of the symmetry adapted SDP

For $\alpha \in \mathbb{N}_r^n$ we define the corresponding *orbit* as $\mathcal{O}_\alpha = \{\sigma(\alpha) \ : \sigma \in \mathcal{S}_n\}$. For each orbit we choose a representative α which is sorted, i.e., $\alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_n$ and index the orbit by that α. We define $S(\mathbb{N}_r^n) = \{\alpha \in \mathbb{N}_r^n : \alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_n\}$ to be the set of representatives for the set of orbits. The set $\mathbb{N}_r^n$ can be written as the union of orbits, i.e.,

$$\mathbb{N}_r^n = \bigcup_{\alpha \in S(\mathbb{N}_r^n)} \mathcal{O}_\alpha.$$

The invariance of K_r means that $g_\alpha = g_\beta$ for every $\beta \in \mathcal{O}_\alpha$. We are now equipped to reformulate (5.24) as an equivalent optimization problem which is easier to solve. We first note that for the invariant kernel we have

$$\sum_{\alpha \in \mathbb{N}_r^n} 2^{H(\alpha)} g_\alpha \prod_{i=1}^n \cos \alpha_i \varphi_i = \sum_{\alpha \in S(\mathbb{N}_r^n)} 2^{H(\alpha)} g_\alpha \left(\sum_{\beta \in \mathcal{O}_\alpha} \prod_{i=1}^n \cos \beta_i \varphi_i \right).$$

For every $\mathcal{B}_i \in \mathcal{B}$ we define a matrix $M^{(i)}$ of size $k_i \times k_i$ with $k_i = |\mathcal{B}_i|$. Then the program may be written as follows.

$$\sigma_r^2 = \min \ n\left(1 - g_{e_n}\right) \tag{5.27}$$

subject to

$$\begin{aligned} \sum_{\alpha \in S(\mathbb{N}_r^n)} 2^{H(\alpha)} g_\alpha \left(\sum_{\beta \in \mathcal{O}_\alpha} \prod_{i=1}^n \cos \beta_i \phi_i \right) &= \sum_{i=1}^{k(n,r)} \sum_{j,\ell=1}^{k_i} \mathcal{R}^{\mathcal{S}_n}(b_{i_j} b_{i_\ell}^*) M_{j,\ell}^{(i)} \\ M^{(i)} &\succeq 0, \ i \in [k]. \end{aligned}$$

Let us take a closer look at the terms $\mathcal{R}^{\mathcal{S}_n}(b_{i_l} b_{i_m}^*)$. Assume that the elements $b_{i_j} \in \mathcal{B}_i$ are given in the following form

$$b_{i_j} = \sum_{m=1}^{k_{i_j}} b_{i_j}^{(m)} \exp\left(\imath \left(\alpha^{(m)}\right)^\top \phi\right).$$

Then,

$$\begin{aligned}\mathcal{R}^{\mathcal{S}_n}(b_{i_j}b_{i_\ell}^*) &= \frac{1}{|\mathcal{S}_n|}\sum_{\sigma\in\mathcal{S}_n}\sigma(b_{i_j}b_{i_\ell}^*)\\ &= \frac{1}{|\mathcal{S}_n|}\sum_{\sigma\in\mathcal{S}_n}\sum_{m=1}^{k_{i_j}}\sum_{p=1}^{k_{i_\ell}} b_{i_j}^{(m)}b_{i_\ell}^{(p)}\exp\left(\imath\left(\alpha^{(m)}\right)^\top\sigma(\phi)\right)\\ &\qquad\qquad \exp\left(\imath\left(\beta^{(p)}\right)^\top\sigma(\phi)\right)^*\\ &= \frac{1}{|\mathcal{S}_n|}\sum_{\sigma\in\mathcal{S}_n}\sum_{m=1}^{k_{i_j}}\sum_{p=1}^{k_{i_\ell}} b_{i_j}^{(m)}b_{i_\ell}^{(p)}\cos[\sigma(\alpha^{(m)}-\beta^{(p)})^\top\phi].\end{aligned} \tag{5.28}$$

Recalling the trigonometric identity (5.23) from before, we can now construct the constraint matrices. Let $\gamma\in S(\mathbb{N}_r^n)$. For each $i\in[k(n,r)], I\in\mathcal{I}$ for $\mathcal{I}$ as in the previous subsection we define

$$\left(C_i^{(\gamma,I)}\right)_{j,\ell} = \begin{cases} c(\gamma,I,i,j,\ell), \text{ if } \omega_I(\gamma) \text{ or } \omega_{I^c}(\gamma) \text{ occurs in } \mathcal{R}^{\mathcal{S}_n}(b_{i_j}b_{i_\ell}^*)\\ 0, \text{ otherwise,}\end{cases}$$

where

$$c(\gamma,I,i,j,\ell) = \frac{1}{|\mathcal{S}_n|}\sum_{\substack{m,p\in[k_i]\\ \alpha^{(m)}-\beta^{(p)}=\pm\omega_I(\gamma)}} b_{i_j}^{(m)}b_{i_\ell}^{(p)}.$$

As before, we must ensure that for all $I\in\mathcal{I}$ we have that the corresponding coefficients in our resulting polynomial are equal. Therefore, we arrive at the set of constraints

$$\sum_{i=1}^{k(n,r)}\langle M^{(i)}, C_i^{(\gamma,\emptyset)} - C_i^{(\gamma,I)}\rangle = 0\,, \text{ for every } \gamma\in S(\mathbb{N}_r^n), I\in\mathcal{I}, I\neq\emptyset.$$

We will now define the set $S\Gamma_{(n,r)} = \Gamma_{(n,r)}/\mathcal{S}_n$, which is the "symmetry adapted" version of $\Gamma_{(n,r)}$ of the previous subsection, where we factored out all permutations $\sigma\in\mathcal{S}_n$ of a reference element γ except the identity. This leads to the constraint set

$$\sum_{i=1}^{k(n,r)}\langle M^{(i)}, C_i^{(\gamma,I)}\rangle = 0\,, \text{ for every } \gamma\in S\Gamma_{(n,r)}, I\in\mathcal{I}.$$

The resulting SDP reads as follows.

$$\sigma_r^2 = \min\ n\left(1-\frac{1}{2}\sum_{i=1}^{k(n,r)}\langle M^{(i)}, C_i^{(e_n,\emptyset)}\rangle\right) \tag{5.29}$$

subject to

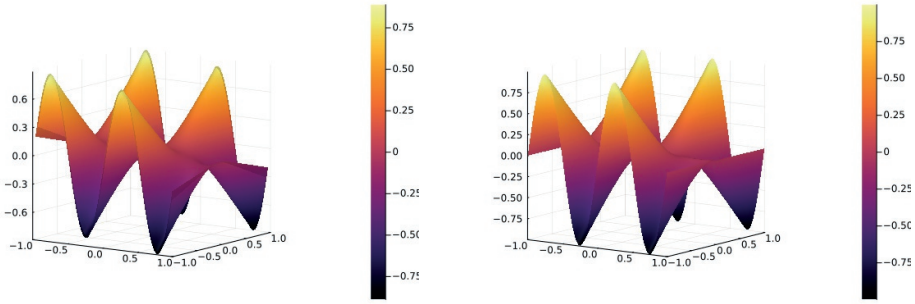
$$\begin{aligned}
\sum_{i=1}^{k(n,r)} \langle M^{(i)}, C_i^{(\gamma,\emptyset)} - C_i^{(\gamma,I)} \rangle &= 0\,, \text{ for every } \gamma \in S(\mathbb{N}_r^n)/\{(0,\ldots,0)\}, I \in \mathcal{I}, I \neq \emptyset \\
\sum_{i=1}^{k(n,r)} \langle M^{(i)}, C_i^{(\gamma,I)} \rangle &= 0\,, \text{ for every } \gamma \in S\Gamma_{(n,r)}, I \in \mathcal{I} \\
\sum_{i=1}^{k(n,r)} \langle M^{(i)}, C_i^{((0,\ldots,0),\emptyset)} \rangle &= 1 \\
M^{(i)} &\succeq 0 \text{ for all } i \in [k(n,r)].
\end{aligned}$$

The efficiency of using this symmetry reduction of course increases when n grows, since the underlying group is $\mathcal{S}_n$. Fixing n and increasing the degree r the size of the underlying program still grows exponentially. It is also possible to use software for the symmetry reduction, such as the Julia package SDPSymmetryReduction.jl[1] which is based on the paper [BdK22] by Brosch and de Klerk. This software takes as input a semidefinite program and numerically performs a symmetry reduction without any need to specify the underlying group. The advantage that comes with this is that there is no need for the construction of a specific symmetry adapted basis. But even for small n, if r becomes too large the resulting optimization problem becomes numerically unstable. It is still worthwhile to compare the two approaches. The block sizes which are returned by the software are the same as the ones we obtained by our approach presented in this section. This suggests that the symmetry is fully exhausted by the symmetric group $\mathcal{S}_n$. We present a plot of a non-differentiable function as well its degree $r = 50$ approximation obtained via the convolution of the minimum resolution kernel in Figure 5.3.

5.5 Comparison to products of univariate minimum resolution kernels

In the following we will have a look at what kernels we get when we take the shortcut and multiply univariate kernels instead of solving the corresponding SDP. The clear advantage is that some optimal univariate kernels are available in closed

[1] see https://github.com/DanielBrosch/SDPSymmetryReduction.jl



(a) Approximation level $r = 50$ (b) Plot of function

Figure 5.3: Comparison of plot of the function $(x, y) \mapsto \sin(2\pi x)|y|$ and its degree $r = 50$ approximation via the minimum resolution kernel.

form. Generating kernels as products means solving the corresponding SDP is unnecessary. Recall that in the univariate case kernels of the form

$$K_r^{\mathrm{KPM}}(x, y) = 1 + 2\sum_{k=1}^{r} g_{k,r}^{\mathrm{KPM}} T_k(x) T_k(y), \tag{5.30}$$

for $g_{k,r}^{\mathrm{KPM}}$ as in (5.12), have minimum resolution σ_r. The product of n univariate degree r kernels of the form (5.30) results in an n-variate kernel of degree nr that is feasible for the optimization problem (5.24). A natural question is to ask how these kernels compare to the ones obtained by solving the SDP. Consider the product of n degree r kernels

$$\prod_{i=1}^{n}\left(1 + 2\sum_{k=1}^{r} g_{k,r}^{\mathrm{KPM}} T_k(x_i) T_k(y_i)\right) = \sum_{\substack{\alpha \in \mathbb{N}_{nr}^n \\ \alpha_i \le r, i \in [n]}} 2^{H(\alpha)} \tilde{g}_\alpha^{\mathrm{KPM}} T_\alpha(\mathbf{x}) T_\alpha(\mathbf{y}),$$

where

$$\tilde{g}_\alpha^{\mathrm{KPM}} = \prod_{i=1}^{n} g_{\alpha_i,r}^{\mathrm{KPM}}$$

We know the resolution is

$$\sigma_r^2 = \sum_{i=1}^{n}(1 - g_{e_i}) = n(1 - g_{e_1}).$$

Thus, we can generate a feasible n-variate kernel with a degree nr multiplying n univariate degree r kernels with minimum resolution and the corresponding resolution is

$$\sigma^2_{nr,\mathrm{KPM}} = n\left(1 - g^{\mathrm{KPM}}_{1,r}\right) \approx \frac{n\pi^2}{2(r+2)^2}.$$

We would expect these to have a worse resolution than the kernels we obtain via solving the SDP where σ_r^2 is minimized. The reason for this is that the product kernels would be feasible to the same SDP with the additional set of constraints

$$g_\alpha = 0 \text{ for all } \alpha \in \mathbb{N}^n_{nr} \text{ with } \alpha_i > r \text{ for some } i \in [n].$$

In particular, we have the following result.

Proposition 5.9. *Fix $n \in \mathbb{N}$. For $r \geq n$ we have*

$$\sigma_r^2 \leq n\left(1 - \cos\left(\frac{n\pi}{r+n}\right)\right) \sim \frac{n^3\pi^2}{2(r+n)^2} \text{ if } r \gg 0.$$

Proof. Clearly, $\sigma_r^2 \leq \sigma^2_{r-1}$ for any $r \geq 1$. Let now $k \in \mathbb{N}$ be such that $kn \leq r \leq (k+1)n$. Then we find

$$\sigma_r^2 \leq \sigma^2_{nk} \leq \sigma^2_{nk,\mathrm{KPM}} = n\left(1 - \cos\left(\frac{\pi}{k+2}\right)\right) \leq n\left(1 - \cos\left(\frac{\pi}{\frac{r}{n}+1}\right)\right) \sim \frac{n^3\pi^2}{2(r+n)^2},$$

for $r \gg 0$. □

Looking at Table 5.1 we find that the values $\sigma_{r,\mathrm{KPM}}$ are larger than σ_r. Therefore, our generalization of the minimum resolution kernels to the multivariate case leads to better approximations than simply multiplying univariate kernels together. Also, for large n, i.e., the case for which our symmetry reduction is efficient, multiplying identical univariate kernels together is not always a feasible approach as the degree is always a multiple of n. In Figure 5.4 and Figure 5.5 we compare the errors of the approximation via the products of Jackson kernels with the approximation via minimum resolution for two different functions.

5.6 Numerical computations

In this section we discuss the numerical computations that were conducted. All code was written in the Julia programming language and is available on GitHub[2].

[2] see https://github.com/FelixKirschner/Approximation-Kernels

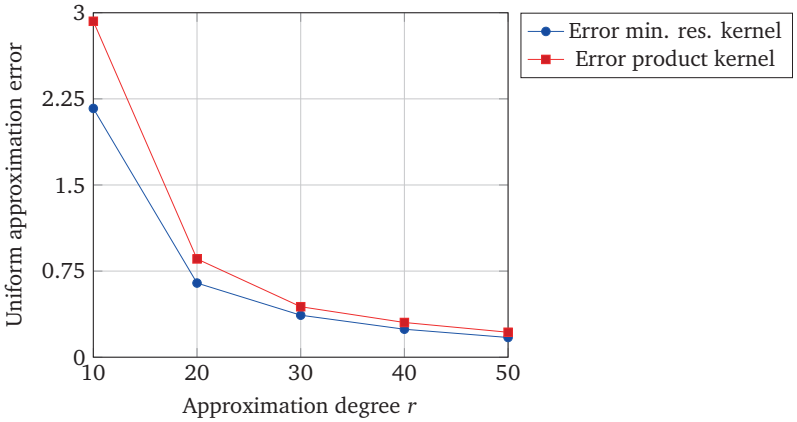


Figure 5.4: Comparison of uniform approximation errors of several approximations of the function $q(\mathbf{x}) := x_2 \sin(2\pi x_1)$. We plotted the errors for the kernel with minimal resolution σ_r and for the product of two univariate degree $r/2$ kernels, i.e., $K^{\mathrm{KPM}}_{r/2}$ as in (5.30).

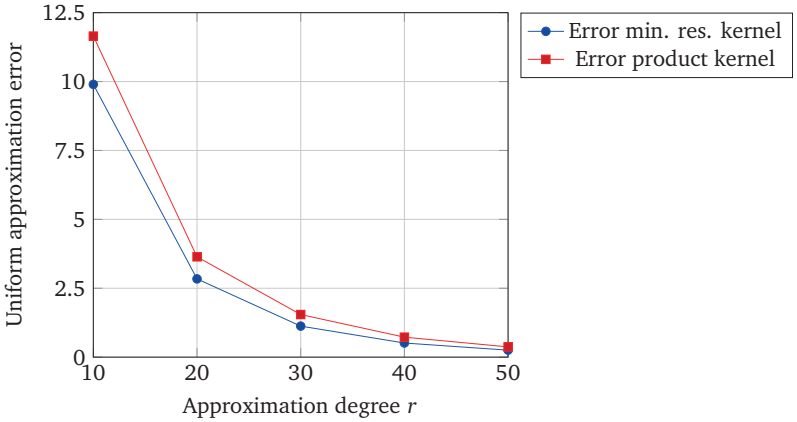


Figure 5.5: Comparison of uniform approximation errors of several approximations of the peaks function $p(\mathbf{x}) := 3(1-x_1)^2 \exp(-x_1^2-(x_2+1)^2) - 10(x_1/5 - x_1^3 - x_2^5)\exp(-x_1^2-x_2^2) - (1/3)\exp(-(x_1+1)^2-x_2^2)$. We plotted the errors for the kernel with minimal resolution σ_r and for the product of two univariate degree $r/2$ kernels, i.e., $K^{\mathrm{KPM}}_{r/2}$ as in (5.30).

At the same website we also list the coefficients of the minimum resolution kernels for various values of n and r. We present some values of σ_r^2 for different values of n and r in Table 5.1. We also compare them to the resolution of the product of identical univariate minimum resolution kernels with the same degree. The results show our method is superior to simple multiplication of identical univariate minimum resolution kernels. In Figure 5.8 we plotted some values of σ_r for different values of n.

For $n = 2$ we were able to compute the coefficients for up to $r = 50$ in a reasonable amount of time (375.5 seconds for σ_{50}^2 on an Apple M1 Pro with 32GB of RAM). After the symmetry reduction, the corresponding program contains two semidefinite matrix variables of order 676 and 650 and has 1277 constraints. We used the CSDP solver version 6.2.0[3] (see [Bor99]) to compute these values. Without the symmetry reduction the program would have one matrix variable of order $\binom{52}{2} = 1326$. We computed the values of σ_r for up to $r = 22$ for $n = 3$ and $r = 13$ for $n = 4$. In the latter case, i.e., $n = 4, r = 13$, without the symmetry reduction the program contains one matrix of size 2380. Using the symmetry reduction we can reduce the size to five matrices of the orders $194, 370, 192, 218$ and 38. For values of $n > 2$ the limiting factor was time.

5.6.1 Decoupling the degrees

Taking a look at problem (5.20) it is clear the value of r on the right-hand-side could be increased to obtain a kernel with potentially smaller resolution. Consider the following problem for fixed r and r' such that $r' \geq r$.

$$\sigma_{r,r'}^2 = \min_{g_\alpha \,:\, \alpha \in \mathbb{N}_r^n} \sum_{i=1}^{n} \left(1 - g_{e_i}\right) \tag{5.31}$$

subject to

$$\begin{aligned} 1 + \sum_{\alpha \in \mathbb{N}_r^n \setminus \{0\}} 2^{H(\alpha)} g_\alpha \prod_{i \in [n]} \cos(\alpha_i \phi_i) &= \left[\exp(\imath \alpha^\top \phi)\right]_{\alpha \in \mathbb{N}_{r'}^n}^* M \left[\exp(\imath \alpha^\top \phi)\right]_{\alpha \in \mathbb{N}_{r'}^n} \\ M &\succeq 0. \end{aligned}$$

For $n = 2$ (resp. $n = 3$) we show how the resolution evolves for $r = 3, \ldots, 10$ (resp. $r = 2, \ldots, 10$) and $r' = r, r+1, \ldots, 20$ in Figure 5.6 (resp. Figure 5.7). We note that the optimal values in the case $n = 2$ seem to stabilize for $r' \geq r + \lfloor \frac{r-1}{2} \rfloor$, whereas such a stabilization pattern may not be observed for $n \geq 3$. We leave further investigation in this direction for future research.

[3] available at https://github.com/coin-or/Csdp

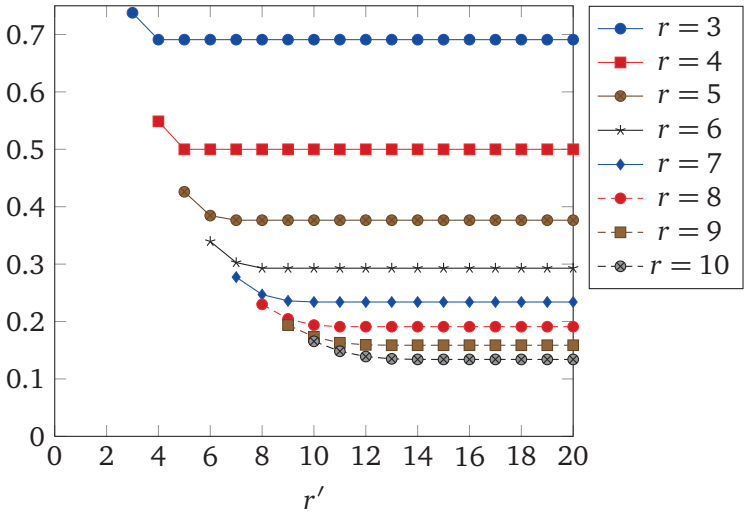


Figure 5.6: Plot $\sigma^2_{r,r'}$ vs r' for $r = 3, \ldots, 10$ and $r' = r, r+1, \ldots, 20$ and $n = 2$.

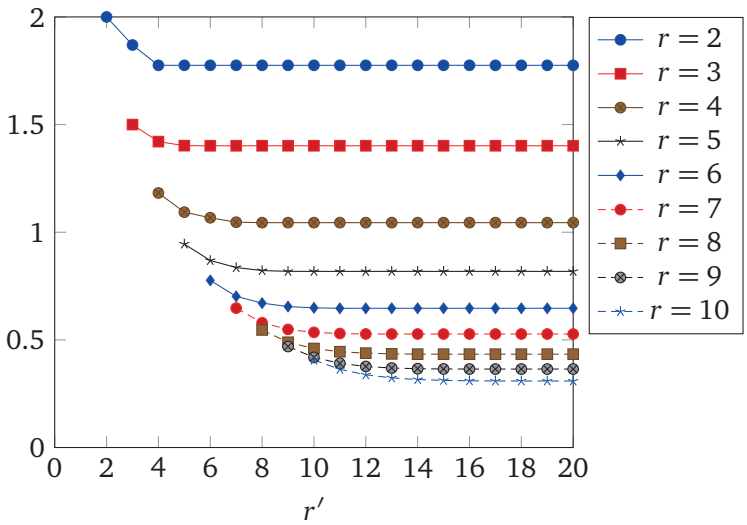


Figure 5.7: Plot $\sigma^2_{r,r'}$ vs r' for $r = 2, \ldots, 10$ and $r' = r, r+1, \ldots, 20$ and $n = 3$

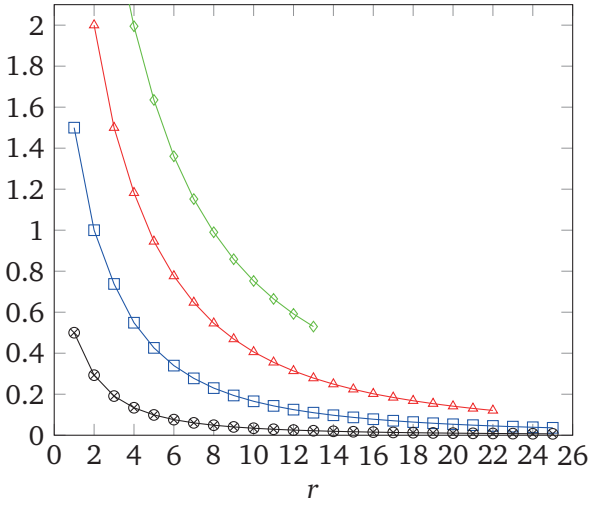


Figure 5.8: Plot σ_r^2 vs r for $n = 1, 2, 3, 4$ ($\otimes, \square, \triangle, \diamond$ resp.)

5.7 Concluding remarks

We have shown how to construct polynomial approximation kernels with minimal resolution on the hypercube. A major open question is if one may find closed form solutions of the semidefinite programs that yield these kernels.

These type of results are also of independent interest in the study of SDP hierarchies for polynomial optimization on the hypercube, as shown recently by Laurent and Slot [LS22a]. In particular, our kernels may be useful to study hierarchies of the Lasserre-type [Las01] on the hypercube (see also [dKHL17, dK10]).

The advantage of our approach over the multiplication of univariate minimum resolution kernels is that it is more efficient (fewer coefficients needed for the same quality approximation), while the clear disadvantage is that we have no closed form solution for the coefficients. Having said that, the tables of coefficients only have to computed once using SDP, and we provide a partial list online[4], as well as a smaller list in Appendix C. Moreover, our approach should become more viable in practice as SDP solvers continue to improve, allowing to compute the coefficients of the kernels in higher dimensions and for larger values of r.

[4]Available at https://github.com/FelixKirschner/Approximation-Kernels/tree/master/SigmaKernels

	$n=2$		$n=3$		$n=4$	
r	σ_r^2	$\sigma_{r,\mathrm{KPM}}^2$	σ_r^2	$\sigma_{r,\mathrm{KPM}}^2$	σ_r^2	$\sigma_{r,\mathrm{KPM}}^2$
1	1.5	-	2.5	-	3.5	-
2	1	1	2	-	2.9310	-
3	0.7378	-	1.5	1.5	2.4561	-
4	0.5487	0.5858	1.1823	-	1.9948	2
5	0.4260	-	0.9451	-	1.6354	-
6	0.3395	0.3820	0.7764	0.8787	1.3605	-
7	0.2774	-	0.6474	-	1.1518	-
8	0.2299	0.2679	0.5461	-	0.9901	1.1716
9	0.1939	-	0.4692	0.5729	0.8584	-
10	0.1655	0.1981	0.4062	-	0.7524	-
11	0.1431	-	0.3556	-	0.6648	-
12	0.1248	0.1522	0.3136	0.4019	0.5917	0.7639
13	0.1099	-	0.2787	-	0.5299	-
14	0.0975	0.1206	0.2493	-	-	-
15	0.0871	-	0.2243	0.2971	-	-
16	0.0782	0.0979	0.2028	-	-	0.5359
17	0.0706	-	0.1843	-	-	-
18	0.0641	0.0810	0.1682	0.2284	-	-
19	0.0585	-	0.1541	-	-	-
20	0.0535	0.0681	0.1417	-	-	0.3961
21	0.0492	-	0.1307	0.1809	-	-
22	0.0453	0.0581	0.1209	-	-	-
23	0.0419	-	-	-	-	-
24	0.0389	0.0501	-	0.1468	-	0.3045
25	0.0362	-	-	-	-	-
30	0.0261	0.0341	-	0.1022	-	-
35	0.0197	-	-	-	-	-
40	0.0154	0.0204	-	-	-	0.1363
45	0.0123	-	-	0.0511	-	-
50	0.0101	0.01352	-	-	-	-

Table 5.1: Computational results for σ_r^2 and $\sigma_{r,\mathrm{KPM}}^2$ for different values of n and r obtained by solving (5.29).

6

Conclusion and outlook

To end this thesis let us give a brief conclusion to reflect on the contributions and present an outlook for possible research directions. The overarching topic of the thesis is conic programming, which is a general framework that allows us to formulate a large variety of optimization problems using only a handful of cones. It is emphasized that the generalized moment problem (GMP) is too difficult to solve in full generality. Nevertheless, we are interested in finding approximate solutions for it. The moment-SOS hierarchy for the GMP is defined over the cone of positive semidefinite matrices, see (1.16). Conceivably, one would like to solve instances of the moment-SOS hierarchies corresponding to levels as high as possible. However, the size of both the primal (1.16) and the dual (1.15) semidefinite programming (SDP) relaxation grows rapidly with the level and the limit cases of solvable problems involve $n \times n$ matrices of size $n \approx 1,000$ and a few tens of thousands of constraints. The first contribution of this thesis is to address this particular problem by constructing an algorithm more suitable for large scale instances of SDP in Chapter 2. The main advantage of the presented algorithm is its suitability for parallelization. Since the average number of cores in modern processors is constantly rising, we believe that this feature will prove advantageous in the future. An efficient implementation of the algorithm and numerical experiments are still needed to prove its superiority to existing software. The underlying idea

of performing subroutines of the interior point method (IPM) in more tractable cones and rescaling the problem is surely not bound the factor width cone alone. Possible future research directions include the analysis of other tractable cones which allow a similar scheme, like the cone of diagonally dominant matrices, see [RSS22].

In Chapter 3 we consider two approximation hierarchies for the GMP over the simplex Δ_{n-1} and the sphere $\mathcal{S}^{n-1}$. The contribution is of a theoretical nature by providing a convergence rate analysis of the hierarchies. Convergence hierarchies of this kind and their convergence analysis remain an active field of research, as many questions regarding theoretical guarantees remain open. For general underlying sets the quantitative Positivstellensatz from Baldi and Mourrain [BM22] can be used to derive a bound on the rate of convergence similarly to the procedure discussed in Chapter 3. However, the resulting convergence rate guarantees are very weak. To obtain stronger results more information about the underlying sets must be taken into account. The moment-SOS hierarchies of lower bounds often exhibit so-called *finite convergence*, meaning the optimal value is achieved for a finite level. In fact, Nie showed in [Nie14] that the moment-SOS hierarchy for polynomial optimization problems exhibits finite convergence generically, that is under some mild assumptions on the global minimizers of the problem. This implies that in practice these hierarchies behave better than some of the theoretical results might suggest. Getting a better understanding of the cases and the orders at which the finite convergence takes place is a promising future research direction.

Chapter 4 deals with the problem of pricing options when information about observable data is available. It is known that this problem can be modeled as a GMP over the non-negative orthant, i.e., $\mathbb{R}^n_+$ with (piecewise) polynomial data functions. In practice option pricing problems are usually solved using numerical techniques like Monte-Carlo simulations [SC23, JL11]. Our main goal is to provide a framework based on semidefinite programming which provably converges to the optimal bounds. For this we prove that, under a mild assumption, the optimal solution of our problem is attained, which allows us to consider an equivalent GMP with compact underlying set. The moment-SOS hierarchy is known to converge when the underlying set satisfies an assumption slightly stronger than compactness. Our intention is to spark the interest in using hierarchies like the one we proposed in practice. We believe there are cases in which the solution obtained via solving moderately sized SDPs can be superior to using common techniques such as Monte Carlo simulation based methods. This, however, is yet

to be demonstrated. The relaxation of the non-compact case, which is treated in Section 4.5 was rather numerically unstable. The relaxation is based on inner approximations of the truncated moment cone. Convergence analyses for such relaxations with several underlying sets were studied by Laurent and Slot, see Lucas Slot's dissertation [Slo22a] for a collection of results. Even though convergence guarantees for hierarchies based on inner approximations of the moment cone are often better than for outer approximations, the convergence behavior and numerical stability in practice is much worse, which is certainly a topic we would like to understand better in the future.

Finally, in Chapter 5 we use SDP techniques to compute multivariate polynomial approximation kernels which lead to uniform approximation of continuous but non-differentiable functions on the hypercube. To avoid unwanted oscillations we consider non-negative kernel functions of minimum resolution. In the univariate case there is a closed form solution of the optimal kernels. These objects are frequently used by physicists in practice. For the multivariate case usually products of univariate kernels are used. We provided a framework which allows for the computation of multivariate minimum resolution kernels via SDP, which we show to be superior to the product of univariate kernels. We hope to encourage physicists who tend to use products of univariate kernels in practice to make use of the multivariate minimum resolution kernels we provide online, as they are provable better than product kernels. It is not clear if there exists a closed form solution of the kernel coefficients in the multivariate case. Here again the scalability of SDP is the key to numerically obtain optimal kernels of higher orders and more variables.

A

Interior Point Methods

In this appendix we will review a class of algorithms called *interior point methods* (IPMs), which can be used to solve (finite dimensional) conic optimization problems [NN94]. These algorithms gained widespread use in the past few decades for multiple reasons. For one, they can solve large-scale linear optimization problems efficiently with good accuracy, where other algorithms like the Simplex method may fail to do so. Moreover, the increasing computational power of modern computers made them suitable for solving medium-sized semidefinite programs, which are known to provide good bounds on NP-hard problems [GW95].

Our exposition follows [Ren01] closely and may be skipped by readers familiar with the topic. Throughout this section $f : D_f \subseteq \mathbb{R}^n \to \mathbb{R}$ will denote a twice continuously differentiable function whose gradient and Hessian we denote by $g(x)$ and $H(x)$, respectively. The domain D_f of f is assumed to be an open, convex set and f is such that $H(x) \succ 0$, i.e., the Hessian is positive definite for all $x \in D_f$. Note that this implies that f is a convex function. We associate a *local inner product* $\langle u, v \rangle_x := \langle u, H(x)v \rangle$ with f, where $\langle \cdot, \cdot \rangle$ is some reference inner product on $\mathbb{R}^n$. Even though $\langle \cdot, \cdot \rangle_x$ depends on the Hessian of f at x, we indicate the local inner product solely by x since f will remain fixed. Note that since we will assume that $H(x)$ is positive definite for all $x \in D_f$ the local inner product is well-defined. In fact, all inner products in $\mathbb{R}^n$ arise this way. With respect to the

local inner product $\langle \cdot, \cdot \rangle_x$, the gradient of f at y becomes $H(x)^{-1}g(y)$ and the Hessian is given by $H(x)^{-1}H(y)$. The local inner product induces a local norm, which we denote by $||\cdot||_x$. With respect to this norm we define

$$B_x(y,r) := \{z \in \mathbb{R}^n : ||z-y||_x < r\},$$

i.e., all points in $\mathbb{R}^n$ whose distance to y measured by the local norm at x is less than $r > 0$.

Definition A.1. A functional f is called *(strongly non-degenerate) self-concordant* if for all $x \in D_f$ we have $B_x(x,1) \subseteq D_f$, and if whenever $y \in B_x(x,1)$ we have

$$1 - ||y-x||_x \leq \frac{||v||_y}{||v||_x} \leq \frac{1}{1-||y-x||_x} \text{ for all } v \neq 0.$$

We will denote the class of self-concordant functions by $\mathcal{SC}$. The class of functions defined this way play a crucial role in optimization as they exhibit many useful properties. For instance, adding a linear functional to a self-concordant function f does not affect its self-concordance, i.e., if $f \in \mathcal{SC}$ then the functional $x \mapsto \langle c, x \rangle + f(x) \in \mathcal{SC}$ for any $c \in \mathbb{R}^n$. Similarly, restricting f to an affine subspace L, which we denote by $f_{|L}$ leads to a self-concordant function, i.e., $f_{|L} \in \mathcal{SC}$. Both these claims follow from the fact that the mentioned transformations leave the Hessian of f unchanged. A functional f is called a self-concordant *barrier* functional if $f \in \mathcal{SC}$ and

$$\vartheta_f := \sup_{x \in D_f} ||H(x)^{-1}g(x)||_x^2 < \infty. \tag{A.1}$$

We refer to ϑ_f as the complexity value of f. Let $\mathcal{SCB}$ denote the set of self-concordant barrier functionals. Nesterov and Nemirovskii [NN94] proved that $\vartheta_f \geq 1$ for all $f \in \mathcal{SCB}$. A crucial property of the functions in $\mathcal{SCB}$ is stated in the next theorem.

Theorem A.2 (cf. Theorem 2.3.3. in [Ren01])**.** *Let $f \in \mathcal{SCB}$ and $x, y \in D_f$. Then,*

$$\langle g(x), y-x \rangle < \vartheta_f.$$

Self-concordant barrier functions can serve as penalty functions for convex cones. If D_f is the interior of a convex cone $\mathcal{K}$ and $f \in \mathcal{SCB}$ then for any $x \in D_f$ we find that $f(x) < \infty$ and for $x \in \partial\mathcal{K}$, i.e., on the boundary of the cone we have $f(x) = \infty$. This property can be exploited to avoid leaving the cone when approaching its boundary, where we know every optimal solution to a conic optimization problem will be attained.

A.1 The central path

Let $f \in \mathcal{SCB}$ be a barrier function for a convex cone $\mathcal{K}$ and say our aim is to solve a conic optimization problem of the following form

$$\begin{aligned} \text{val} = \inf \langle c, x \rangle \\ \text{s.t.} \langle a_i, x \rangle = b_i \,, i \in [m] \\ x \in \mathcal{K}. \end{aligned} \tag{A.2}$$

Let $L = \{x \in \mathbb{R}^n \,: \langle a_i, x \rangle = b_i \,, i \in [m]\}$. Consider now the following sequence of optimization problems for $\eta \in \mathbb{R}_+$

$$z_\eta = \text{argmin}_{x \in L \cap \bar{D}_f} f_\eta(x) := \eta \langle c, x \rangle + f(x). \tag{A.3}$$

The collection of minimizers for all $\eta > 0$ is called the central path and is an analytic curve in $L \cap \bar{D}_f$. Every point on the central path is feasible for (A.2) and provides an upper bound on the optimal value val. At the end of this section we will provide a proof that z_η converges to a minimizer of (A.2) for $\eta \to \infty$. Given some $\varepsilon > 0$ interior point methods approximate a sequence $z_{\eta_0}, z_{\eta_1}, \ldots$ of points on the central path with $\eta_0 < \eta_1 < \ldots$ until a feasible point x' is found such that $\langle c, x' \rangle \leq \text{val} + \varepsilon$. There are many ways to achieve this, and we will present two in the following. The first is conceptually easy to grasp, while the second one works better in practice and builds the foundation to the algorithm we develop in the Chapter 2. We close this section with a proof that z_η converges to a minimizer of (A.2) for $\eta \to \infty$. Evaluating the gradient g_L of $f_{|L}$, i.e., f restricted to L at a point z_η on the central path we find for P_L being the orthogonal projection operator onto L with respect to the reference inner product $\langle \cdot, \cdot \rangle$ that

$$g_L(z_\eta) = -\eta P_L c,$$

since z_η is the minimizer of (A.3). Hence, for any $y \in \bar{D}_f \cap L$ it follows by Theorem A.2 that

$$\begin{aligned} \langle c, z_\eta \rangle - \langle c, y \rangle &= \frac{1}{\eta} \langle P_L g(z_\eta), y - z_\eta \rangle \\ &< \frac{1}{\eta} \vartheta_{f_{|L}} \leq \frac{1}{\eta} \vartheta_f, \end{aligned}$$

where we used the fact that we can write $c = P_L c + c'$, with $c' \in L^\perp$. Finally, we deduce

$$\text{val} \leq \langle c, z_\eta \rangle \leq \text{val} + \frac{1}{\eta} \vartheta_f.$$

Since ϑ_f is finite if $f \in \mathcal{SCB}$, the claim follows.

A.2 Newton's method

In this section we introduce Newton's method for minimizing functionals. Let f be the function to be minimized and $x^{(0)} \in D_f$ be given. Newton's method works by iteratively minimizing the quadratic approximation of f at $x^{(i)}$ for $i = 0, 1 \ldots$ and defining $x^{(i+1)}$ as the minimizer. Let us define the quadratic approximation q_x of f at x as

$$q_x(y) := f(x) + \langle g(x), y - x \rangle + \frac{1}{2} \langle y - x, H(x)(y - x) \rangle.$$

Given $x^{(0)}$ Newton's method asks for the minimizer of $q_{x^{(0)}}$, i.e.,

$$x^{(1)} := \operatorname{argmin}_{y \in D_f} q_{x^{(0)}}(y)$$

and proceeds by using this point as the starting point for the following iteration. Reiterating this process one obtains a sequence of points $\{x^{(i)}\}_{i \in \mathbb{N}}$. It is well-known that Newton's method does not converge to the true optimum in general. It can be shown, however, that if the given starting point is sufficiently close to a local minimum, then the procedure converges to this minimum. Note the following proposition about the gradient and Hessian of the quadratic approximation of f.

Proposition A.3 (cf. Proposition 1.6.1 in [Ren01]). *The gradient of q_x at y is given by $g(x) + H(x)(y - x)$ and its Hessian is $H(x)$, with respect to the reference inner product.*

Since we are looking for an extreme point of q_x, and we have a formula for its gradient, we can enforce the necessary condition at the minimizer, call it x^+. We have

$$g(x^+) = g(x) + H(x)(x^+ - x) = 0.$$

Suppose f is such that its Hessian is positive definite for all $x \in D_f$. It follows that

$$x^+ := x - H(x)^{-1} g(x)$$

(this expression is independent of the reference inner product). For further reference we define the *Newton step at x* as

$$n(x) = x^+ - x = -H(x)^{-1} g(x).$$

For generic optimization problems the gradient and Hessian are usually not available in closed form or too difficult to compute. In special cases however, such

as, e.g., LP and SDP, there exist self-concordant barrier functions for the underlying cones whose gradients and Hessians can be computed efficiently. This implies that interior point methods are suitable to tackle conic optimization problems of this kind.

Let us continue by analyzing Newton's method. For this note the following theorem.

Theorem A.4 (cf. Theorem 1.6.2 in [Ren01]). *If z minimizes f and $H(x)$ is invertible, then*

$$||z - x^+|| \leq ||x - z||\, ||H(x)^{-1}|| \int_0^1 ||H(x + t(z-x)) - H(x)||\, \mathrm{d}t.$$

Essentially, this theorem implies that if a point x is *close enough* to the minimizer z of f then the result x^+ of an iteration of Newton's methods lies closer to z than x. The following results will be crucial to our analysis.

Theorem A.5 (cf. Theorem 2.2.2 in [Ren01]). *If $f \in \mathcal{SC}$, $x \in D_f$ and $y \in B_x(x, 1)$, then*

$$|f(y) - q_x(y)| \leq \frac{||y - x||_x^3}{3(1 - ||y - x||_x)}.$$

Theorem A.6 (cf. Theorem 2.2.3 in [Ren01]). *Let $f \in \mathcal{SC}$ and $x \in D_f$. If z minimizes f and $z \in B_x(x, 1)$, then*

$$||x^+ - z||_x \leq \frac{||x - z||_x^2}{1 - ||x - z||_x}.$$

The following theorem will allow us to decide whether a given point x is *close* to the minimizer of f in terms of the Newton step at x. The quantity $||n(x)||_x$ is also known as the *Newton decrement* at x; see Definition A.8 below.

Theorem A.7 (cf. Theorem 2.2.5 in [Ren01]). *Assume $f \in \mathcal{SC}$. If $||n(x)||_x \leq \frac{1}{4}$ for some $x \in D_f$, then f has a minimizer z and*

$$||z - x_+||_x \leq \frac{3||n(x)||_x^2}{(1 - ||n(x)||_x)^3}.$$

Thus,

$$||z - x||_x \leq ||n(x)||_x + \frac{3||n(x)||_x^2}{(1 - ||n(x)||_x)^3}.$$

Newton's method on subspaces

Suppose we want to use Newton's method to minimize a function $f_{|L}$, i.e., a functional f restricted to a subspace $L \subset \mathbb{R}^n$. Let P_L be the orthogonal projection of $\mathbb{R}^n$ onto L with respect to the reference inner product $\langle \cdot, \cdot \rangle$. The gradient and the Hessian of $f_{|L}$ are given by $P_L g(x)$ and $P_L H(x)$, respectively. Therefore, the Newton step at $x \in L \cap D_f$ is given by $n_L(x) \in L$ such that

$$n_L(x) = -(P_L H(x))^{-1} P_L g(x) \quad \Leftrightarrow \quad P_L \left(H(x) n_L(x) + g(x)\right) = 0,$$

i.e., by a vector $n_L \in L$ such that $H(x)n_L(x)+g(x)$ is orthogonal to L. Suppose L is the nullspace of a linear operator $A : \mathbb{R}^n \to \mathbb{R}^m$, then we cast these two restrictions as a linear system:

$$\begin{aligned} H(x) n_L(x) + g(x) &= A^* y \\ A n_L(x) &= 0. \end{aligned}$$

Newton decrements for functions restricted to subspaces

The convergence rate of interior point methods depends on the so-called Newton decrement.

Definition A.8. If $f : \mathbb{R}^n \to \mathbb{R}$ has a gradient $g(x)$ and positive definite Hessian $H(x) \succ 0$ at a point x in its domain, then the Newton decrement of f at x is defined as

$$\Delta(f, x) = \sqrt{\langle g(x), g_x(x) \rangle} = ||g_x(x)||_x = ||n(x)||_x,$$

where we define $g_x(x) = H^{-1}(x)g(x)$, i.e., $g_x(x)$ is the gradient of f at x with respect to the $\langle \cdot, \cdot \rangle_x$ inner product.

For self-concordant functions f, a sufficiently small value of $\Delta(f, x)$ implies that x is close to the minimizer of f see Theorem A.7, where the estimates are given in terms of the Newton decrement.

If a self-concordant function f is restricted to a (translated) linear subspace L, and denoted by $f_{|L}$, then the Newton decrement at x becomes

$$\Delta\left(f_{|L}, x\right) = ||P_{L,x} H^{-1}(x) g(x)||_x,$$

where $\| \cdot \|_x$ is the norm induced by the inner product $\langle u, v \rangle_x = \langle u, H(x)v \rangle$, and $P_{L,x}$ is the orthogonal projection onto L for the $\| \cdot \|_x$ norm; see [Ren01, § 1.6].

Note that we have

$$\begin{aligned}\Delta(f,x) &= \langle g(x), H^{-1}(x)g(x)\rangle^{1/2} = \langle g(x), -n(x)\rangle^{1/2} \\ &= \langle n(x), n(x)\rangle_x^{1/2} = ||n(x)||_x = \sup_{||d||_x=1} \langle d, n(x)\rangle_x,\end{aligned}$$

where $n(x)$ is the Newton step at x, i.e., $n(x) = -H(x)^{-1}g(x)$. Hence, restricting the function f to a subspace L we find

$$\begin{aligned}\Delta\left(f_{|_L}, x\right) &= \sup_{||d||_x=1} \langle d, P_{L,x}n(x)\rangle_x = \sup_{\substack{||d||_x=1 \\ d\in L}} \langle d, n(x)\rangle_x \\ &= \sup_{0\neq d\in L} \frac{\langle d, n(x)\rangle_x}{||d||_x} \geq \frac{\langle d, n(x)\rangle_x}{||d||_x} \text{ for all } d \in L\setminus\{0\}.\end{aligned} \tag{A.4}$$

A.3 Short-step method

Let now $D_f = \{x \in \mathcal{K} : \langle a_i, x\rangle = b_i, i \in [m]\}$. In the following we assume a point $x^{(0)}$ is given, which is close to a point z_{η_0} on the central path in the sense that Newton's method started at $x^{(0)}$ converges to z_{η_0}. The method produces a sequence of points which all lie close to the central path, by taking *short* steps, hence the name. Essentially, the algorithm works by increasing the parameter η by a specified *safe* amount and using one iteration of Newton's method to obtain a new point near the central path corresponding to the increased η value. This point will serve as a starting point for the next iteration, where η is increased again by the safe amount. This procedure is repeated until a satisfactory solution is found. In the following we will assume that only a single iteration of Newton's method is applied, i.e., starting with $x^{(i)}$ close to z_{η_i} we set

$$x^{(i+1)} = x^{(i)} + n^{\eta_{i+1}}(x^{(i)}),$$

where the Newton step with respect to η is defined as

$$n^{\eta}(x) := -H(x)^{-1}\left(\eta c + g(x)\right).$$

Some caution is required when increasing η_0 to η_1. If we choose η_1 too large with respect to η_0, then $x^{(1)}$ may fail to approximate z_{η_1}. To avoid this situation, we will derive an amount κ which is so that when $\eta_1 = \kappa\eta_0$, then $x^{(1)}$ is close to z_{η_1}. Conceivably, the length of the Newton steps $n^{\eta}(x^{(i)})$ will decrease the closer $x^{(i)}$ is to z_{η_i}. This relation provides a tool of measuring distance to the central path.

We say a point x is close to a point z_η on the central path if the Newton decrement $\Delta(f_\eta, x) = ||n^\eta(x)||_x$ is small. For the time being, we will consider the *concept* of smallness and defer quantification until later. Let $x^{(0)}$ be close to z_{η_0} in the sense that $\Delta(f_{\eta_0}, x^{(0)})$ is given and small. Now increase η_0 to $\kappa\eta_0 =: \eta_1$. Our aim is to come up with a suitable value for κ, such that $\Delta(f_{\eta_1}, x^{(1)})$ is still small. The Newton step taken from $x^{(0)}$ with respect to the increased η value will be given by $n^{\eta_1}(x^{(0)})$. We can relate the Newton steps for η_0 and η_1 in the following way

$$n^{\eta_1}(x) = \frac{\eta_1}{\eta_0} n^{\eta_0}(x) + \left(\frac{\eta_1}{\eta_0} - 1\right) g_x(x), \tag{A.5}$$

where we set $g_x(x) := H(x)^{-1} g(x)$. This equality is easily verified. Using the triangle inequality we obtain

$$\Delta(f_{\eta_1}, x) = ||n^{\eta_1}(x)||_x \leq \frac{\eta_1}{\eta_0} \Delta\left(f_{\eta_0}, x\right) + \left|\frac{\eta_1}{\eta_0} - 1\right| \sqrt{\vartheta_f}. \tag{A.6}$$

We want to find a value of κ such that if $||n^{\kappa\eta_0}(x^{(0)})||_{x^{(0)}}$ is small, then $||n^{\kappa\eta_0}(x^{(1)})||_{x^{(1)}}$ is small. The following theorem lets us upper bound the latter expression in terms of the former.

Theorem A.9. *[see, e.g., Thm. 2.2.4 in [Ren01]] Let $f \in \mathcal{SC}$. If $\Delta(f, x) = ||n(x)||_x < 1$, then*

$$||n(x^+)||_{x^+} \leq \left(\frac{||n(x)||_x}{1 - ||n(x)||_x}\right)^2.$$

If $||n^{\eta_1}(x^{(0)})||_{x^{(0)}} < 1$

$$||n^{\eta_1}(x^{(1)})||_{x^{(1)}} \leq \left(\frac{||n^{\eta_1}(x^{(0)})||_{x^{(0)}}}{1 - ||n^{\eta_1}(x^{(0)})||_{x^{(0)}}}\right)^2.$$

With these tools we can upper bound $\Delta(f_{\eta_1}, x^{(1)})$ in terms of $\Delta(f_{\eta_0}, x^{(0)})$. Let $\Delta(f_{\eta_0}, x^{(0)}) < \alpha$, for α to be determined. Then, by (A.6) we have

$$\Delta(f_{\eta_1}, x^{(0)}) \leq \beta := \kappa\alpha + (\kappa - 1)\sqrt{\vartheta_f}.$$

Suppose we find values for $\alpha > 0$ and $\kappa > 1$ such that $\beta < 1$ and

$$\left(\frac{\beta}{1 - \beta}\right)^2 \leq \alpha.$$

Then, using Theorem A.9 we find

$$\Delta(f_{\eta_1}, x^{(1)}) \leq \left(\frac{\Delta(f_{\eta_1}, x^{(0)})}{1-\Delta(f_{\eta_1}, x^{(0)})}\right)^2 \leq \left(\frac{\beta}{1-\beta}\right)^2 \leq \alpha.$$

Choosing

$$\kappa := 1 + \frac{1}{8\max\{1, \sqrt{\vartheta_f}\}} = 1 + \frac{1}{8\sqrt{\vartheta_f}}$$

and $\alpha := 1/9$ we find these relations satisfied.

Complexity

Recall from Section A.1 that $\langle c, z_\eta \rangle \leq \text{val} + \frac{1}{\eta}\vartheta_f$. Hence, to achieve a solution x approximating z_η such that $\langle c, z_\eta \rangle \leq \text{val} + \varepsilon$ one must increase η from η_0 to $\eta \geq \vartheta_f/\varepsilon$. If at each step κ is chosen as exactly $1 + \frac{1}{8\sqrt{\vartheta_f}}$ we find that the number K of steps required to increase η_0 to some value $\eta > \eta_0$ is

$$K = \frac{\log(\eta/\eta_0)}{\log(1 + 1/8\sqrt{\vartheta_f})} \leq 10\sqrt{\vartheta_f}\log(\eta/\eta_0) = O(\sqrt{\vartheta_f}\log(\eta/\eta_0)).$$

Thus, the number of increases needed to approximate an ε-optimal solution is

$$O\left(\sqrt{\vartheta_f}\log\left(\frac{\vartheta_f}{\varepsilon\eta_0}\right)\right).$$

The question remains what is the quality of a solution x approximating z_η. Let x now be an arbitrary point in $D_f \cap L$. By the definition of self-concordance we have that $B_x(x, 1) \subseteq D_f$, and thus $x - tc_x \in D_f$ for all $t \in [0, ||c_x||_x^{-1})$, where $c_x := H(x)^{-1}c$. Plugging $P_L(x - tc_x)$ into the objective function we find

$$\text{val} \leq \langle c, P_L(x - tc_x)\rangle \Rightarrow ||P_L c_x||_x \leq \langle c, x\rangle - \text{val}. \tag{A.7}$$

It follows for every $y \in L$ that

$$\frac{\langle c, y\rangle - \text{val}}{\langle c, x\rangle - \text{val}} = 1 + \frac{\langle P_L c_x, y - x\rangle_x}{\langle c, x\rangle - \text{val}} \leq 1 + \frac{||P_L c_x||_x \, ||y - x||_x}{\langle c, x\rangle - \text{val}} \leq 1 + ||y - x||_x.$$

Therefore, we find that

$$\langle c, y\rangle \leq \text{val} + \frac{1}{\eta}\vartheta_f\left(1 + ||y - z_\eta||_{z_\eta}\right),$$

meaning the objective value for any feasible point y is bounded in terms of its distance to the central path. Now, suppose that x is returned by the algorithm for given η. Then by the choice of $\alpha = 1/9$ we know that $||n^\eta(x)||_x \leq 1/9$. Applying the second part of Theorem A.7 we get $||x - z_\eta||_x \leq 1/6$. Recalling the definition of self-concordance (and setting $v = x - z_\eta$) note that $z_\eta \in B_x(x, 1)$ and thus

$$\frac{||x - z_\eta||_{z_\eta}}{||x - z_\eta||_x} \leq \frac{1}{1 - ||x - z_\eta||_x} \Leftrightarrow ||x - z_\eta||_{z_\eta} \leq 1/5.$$

We will close this section with the following remark about the maximal decrease in objective value that can be achieved in on iteration of the presented method. Let $x \in D_f$ be the current point and $n^\eta(x)$ the corresponding Newton step. The decrease is then given by

$$\langle c, x\rangle - \langle c, x + n^\eta(x)\rangle = \langle c_x, n^\eta(x)\rangle_x \leq ||c_x||_x \, ||n^\eta(x)||_x \leq \frac{1}{4}||c_x||_x, \tag{A.8}$$

where we used (A.5) to bound

$$||n^\eta(x)||_x = \left|\left|\left(1 + \frac{1}{8\sqrt{\vartheta_f}}\right) n^{\eta/\kappa}(x) + \frac{1}{8\sqrt{\vartheta_f}} g_x(x)\right|\right|_x \leq \frac{9}{8}\frac{1}{9} + \frac{1}{8} = \frac{1}{4},$$

because $||n^{\eta/\kappa}(x)||_x \leq 1/9$ by construction and $||g_x(x)||_x \leq \sqrt{\vartheta_f}$ by definition.

A.4 Predictor-Corrector method

In this section we present the predictor-corrector method, which will serve as the foundation of the algorithm we introduce in Chapter 2. The method consists of two phases. In the first phase, called the predictor phase, a feasible point is produced which reduces the current objective value. This point is not necessarily close to the central path. In the second phase, called the corrector phase, a sequence of points with the same reduced objective value is generated such that the distance to the central path of each point in the sequence is less that its predecessor's. Once a point is found which is *close enough* to the central path, the algorithm terminates if it is an ε optimal solution, or switches to the predictor phase to reduce the objective value again, see Figure A.1.

Predictor step

Taking a look at the Newton step at x, i.e., $n^\eta(x) = -\eta c_x - g_x(x)$ we find that it consists of two components. The first, i.e., $-\eta c_x$ reduces the objective value and

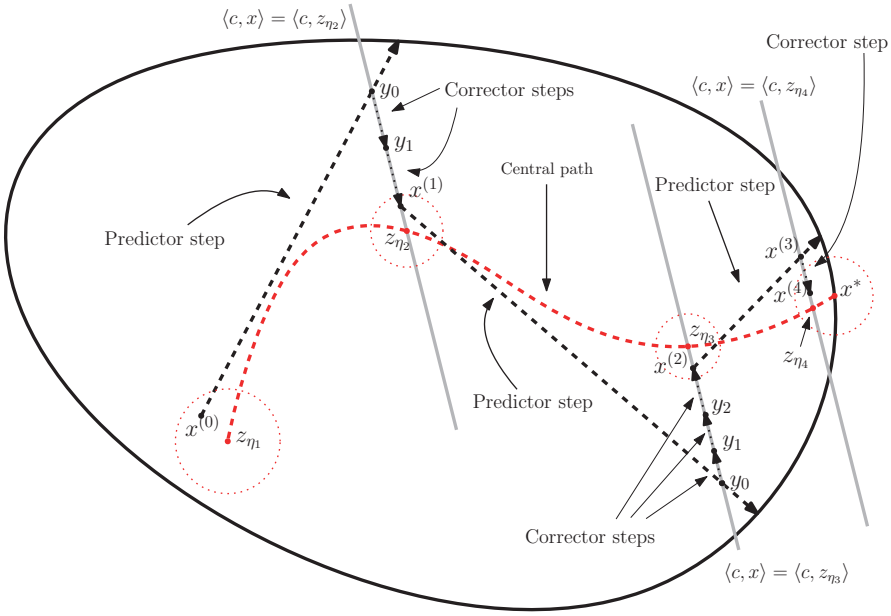


Figure A.1: Visualization of predictor-corrector method. Initial feasible solution close to central path (red) is given by $x^{(0)}$. Algorithm performs predictor step returning y_0. Corrector steps are taken until point close enough to central path ($x^{(1)}$) is found. Next predictor step returns new point y_0. Corrector steps are taken until $x^{(2)}$ is found, which is close enough to central path to perform next predictor step. After one corrector step the final point $x^{(4)}$ is ε-close to x^*.

is approximately tangential to the central path if x is close enough to the central path. The second quantity $g_x(x)$ contains information about the curvature of the central path at x. Given a point $x^{(0)}$ close to a point z_{η_0} on the central path for some $\eta_0 > 0$ the predictor step moves a fixed fraction of the way towards the boundary of the feasible region in the affine scaling direction, which is given by $c_{x^{(0)}}$. Let $\sigma \in (0,1)$ be fixed. To perform the predictor step one computes

$$s^* = \sup_{s \in \mathbb{R}} \{s : x^{(0)} - s c_{x^{(0)}} \in D_f\},$$

and returns $y_0 := x^{(0)} - s^* \sigma c_{x^{(0)}}$. Let $v_0 = \langle c, x^{(0)} \rangle$. Then we find

$$v_1 = \langle c, y_0 \rangle = \langle c, x^{(0)} - s^* \sigma c_{x^{(0)}} \rangle = v_0 - s^* \sigma ||c_{x^{(0)}}||^2_{x^{(0)}}.$$

By definition of self-concordance, we have that $s^* \geq ||c_{x^{(0)}}||^{-1}_{x^{(0)}}$. Hence, choosing $\sigma \geq 1/4$ the decrease in the objective value for the predictor step is at least as large as the maximal decrease (A.8) that can be achieved by an iteration of the short step method. Usually, σ is chosen to be much larger, e.g., $\sigma = .99$.

Corrector step

The corrector step takes as input a feasible point y_0 which is not close to the central path and returns a point $x^{(1)}$ with the same objective value as y_0 which is close to the central path. Let $v_1 = \langle c, y_0 \rangle$ and define $L(v_1) = \{x \in D_f : \langle c, x \rangle = v_1, \langle a_i, x \rangle = b_i, i \in [m]\}$. The corrector step seeks to find the minimum of the barrier $f_{|L(v_1)}$, which clearly is the point z_{η_1} on the central path such that $\langle c, z_{\eta_1} \rangle = v_1$. This can be done by iteratively computing y_k for $k \geq 1$ as the minimizer of the univariate functional

$$t \mapsto f\left(y_{k-1} + t n_{L(v_1)}(y_{k-1})\right). \tag{A.9}$$

Here, $n_{L(v_1)}(y_k)$ is the Newton step of the restricted functional $f_{|L(v_1)}$ at y_k and the minimizer can be computed using (exact) line search. This is repeated until a y_k is found for which $||n_{L(v_1)}(y_k)||_{y_k}$ is small enough, implying y_k is close to the central path and one sets $x^{(1)} = y_k$. To prove this method converges we will first provide an upper bound on the difference $f(y_0) - f(z_{\eta_1})$ and then show that each y_k for $k \geq 1$ reduces this distance by at least a constant amount. For the termination criterion we claim that if $||n_{L(v_1)}(y_k)||_{y_k} \leq 1/14$ we can set $x_1 = y_k$. To see this, assume $||n_{L(v_1)}(y_k)||_{y_k} \leq 1/14$ and note that z_{η_1} is the minimizer of $f_{|L(v_1)}$. Hence, applying Theorem A.7 we find $||z_{\eta_1} - y_k||_{y_k} \leq 1/11$. Theorem A.6

applied to the functional f_{η_1} implies

$$||n^{\eta_1}(y_k)||_{y_k} \leq ||z_{\eta_1} - y_k||_{y_k} + \frac{||z_{\eta_1} - y_k||^2_{y_k}}{1 - ||z_{\eta_1} - y_k||_{y_k}} \leq 1/9,$$

which is exactly the threshold we obtained for the short step method.

Let us now prove an upper bound for the difference $f(y_0) - f(z_{\eta_1})$. For this consider the following identity

$$f(y_0) - f(z_{\eta_1}) = \underbrace{\big(f(y_0) - f(x^{(0)})\big)}_{=:\rho_1} + \underbrace{\big(f(x^{(0)}) - f(z_{\eta_0})\big)}_{=:\rho_2} + \underbrace{\big(f(z_{\eta_0}) - f(z_{\eta_1})\big)}_{\rho_3}.$$

We will bound these one at a time. For ρ_1 we can use the following theorem.

Theorem A.10 (cf. Theorem 2.3.8 in [Ren01]). *Assume $f \in \mathcal{SCB}$ and $x \in D_f$. If $y \in \bar{D}_f$, then for all $0 < t \leq 1$,*

$$f(y + t(x - y)) \leq f(x) - \vartheta_f \log(t).$$

Note that

$$y_0 = x^{(1)} - \sigma s^* c_x = \big(x^{(0)} - s^* c_x\big) + (1 - \sigma)\big(x^{(0)} - (x^{(0)} - s^* c_x)\big),$$

which after applying Theorem A.10 leads to $f(y_1) \leq f(x^{(0)}) - \vartheta_f \log(1 - \sigma)$. Hence, it follows that $\rho_1 \leq -\vartheta_f \log(1 - \sigma)$. To bound ρ_2 we will use the fact that for a convex functional f and $x, y \in D_f$ we have

$$f(x) - f(y) \leq \langle g(x), x - y \rangle. \tag{A.10}$$

Clearly,

$$\begin{aligned} \rho_2 = f(x^{(0)}) - f(z_{\eta_1}) &= f_{|L(v_0)}(x^{(0)}) - f_{|L(v_0)}(z_{\eta_0}) \\ &\leq ||n_{L(v_0)}(x^{(0)})||_{x^{(0)}} \; ||z_{\eta_0} - x^{(0)}||_{x^{(0)}} \leq \frac{1}{14}\frac{1}{11} = \frac{1}{154}. \end{aligned}$$

Finally, for ρ_3 we use (A.10) again to find

$$\rho_3 = f(z_{\eta_0}) - f(z_{\eta_1}) \leq \langle -\eta_0 c, z_{\eta_0} - z_{\eta_1} \rangle = -\eta_0(v_0 - v_1) \leq 0,$$

since $g(z_{\eta_0}) = -\eta_0 c$. Putting it all together we obtain

$$f_{\eta_1}(y_0) - f_{\eta_1}(z_{\eta_1}) \leq \vartheta_f \log\left(\frac{1}{1 - \sigma}\right) + \frac{1}{154}.$$

The remainder of the proof consists of providing a lower bound on the differences $f(y_k)-f(y_{k+1})$, i.e., we will show that while $||n_{|L(v_1)}(y_k)||_{y_k} > 1/14$ the decrease of the function value stepping from y_k to y_{k+1} is at least a constant amount, which is to be specified. Recall that by Theorem A.5 we have

$$|f(y)-q_x(y)| \leq \frac{||y-x||_x^3}{3(1-||y-x||_x)}.$$

Let $y = x + tn(x)$ for some $x \in D_f$. Then,

$$|f(x+tn(x))-q_x(x+tn(x))| \leq \frac{t^3||n(x)||_x^3}{3(1-t||n(x)||_x)}.$$

Setting $t = \frac{1}{8||n(x)||_x}$ the RHS becomes

$$\frac{t^3||n(x)||_x^3}{3(1-t||n(x)||_x)} = \frac{(1/8)^3}{3(1-1/8)} = \frac{1}{1344}.$$

Moreover, we find that

$$\begin{aligned} f(y) = f(x+tn(x)) &\leq q_x(x+tn(x)) + \frac{1}{1344} \\ &= f(x) + \frac{1}{8||n(x)||_x}\langle g(x), n(x)\rangle \\ &\quad + \frac{1}{2}\left(\frac{1}{8||n(x)||_x}\right)^2 \langle n(x), H(x)n(x)\rangle + \frac{1}{1344} \\ &= f(x) - \frac{1}{8||n(x)||_x}\langle -H(x)^{-1}g(x), H(x)n(x)\rangle \\ &\quad + \frac{1}{2}\left(\frac{1}{8}\right)^2 + \frac{1}{1344} \\ &= f(x) - \frac{1}{8}||n(x)||_x + \frac{1}{2}\left(\frac{1}{8}\right)^2 + \frac{1}{1344} \\ &\leq f(x) - \frac{1}{2688}, \end{aligned} \tag{A.11}$$

where the last inequality follows from the fact that by assumption $||n(x)||_x > 1/14$. The line search returns a value of t which minimizes the functional (A.9). Hence,

$$f(y_{k+1}) - f(y_k) \geq \tau := \frac{1}{2688}.$$

We conclude that the number K of line searches necessary to move from $x^{(0)}$ to $x^{(1)}$ satisfies

$$K = O\left(\vartheta_f \log\left(\frac{1}{1-\sigma}\right)\right).$$

B

Symmetry reduction

In this chapter we present the theoretical foundation of a tool which can be used to reduce the size of convex optimization problems which exhibit symmetries.

B.1 Groups and their representations

We begin by reviewing some concepts from representation theory. Recall that a group $(\mathcal{G}, *)$ is a set $\mathcal{G}$ together with a binary operation $* : \mathcal{G} \times \mathcal{G} \to \mathcal{G}$ such that

- there exists one unit element id such that $\mathrm{id} * g = g = g * \mathrm{id}$ for all $g \in \mathcal{G}$;
- for every $g \in \mathcal{G}$ there exists an inverse element $g^{-1} \in \mathcal{G}$, i.e., $g * g^{-1} = g^{-1} * g = \mathrm{id}$;
- associativity holds, i.e., for $g, h, k \in \mathcal{G}$ we have

$$(g * h) * k = g * (h * k).$$

We will be particularly interested in the so-called *symmetric group* $\mathcal{S}_n$, which consists of all bijections from the set $\{1, \ldots, n\}$ to itself. For a comprehensive study of the symmetric group we refer to [Sag01]. The elements $\pi \in \mathcal{S}_n$ are called

permutations and the binary group action is composition. As the elements of $\mathcal{S}_n$ permute elements of the set $\{1,\dots,n\}$, a group may be interpreted as acting on a set. Define an *action map* $a : \mathcal{G} \times X \to X$ such that the map is compatible with the group law, i.e., for $g,h \in \mathcal{G}, x \in X$

$$a(h,a(g,x)) = a(hg,x)$$

and $a(\mathrm{id},x) = x$. For an element $\pi \in \mathcal{S}_n$, which acts on the set $\{1,\dots,n\}$, we may simply write $a(\pi,i) = \pi(i) = j$ for the action of π on the element i.

B.1.1 Representation theory

Representation theory concerns itself with representing groups in a way that facilitates their study. Let GL_d be the set of all invertible matrices in $\mathbb{C}^{d\times d}$. Recall that for two groups $(\mathcal{G},*),(\mathcal{H},\cdot)$ a *group homomorphism* $h : \mathcal{G} \to \mathcal{H}$ is a function such that for all $g_1,g_2 \in \mathcal{G}$ we have

$$h(g_1 * g_2) = h(g_1) \cdot h(g_2).$$

A matrix representation of a group $\mathcal{G}$ is a group homomorphism $X : \mathcal{G} \to GL_d$. Equivalently, for each element $g \in \mathcal{G}$ there is an invertible matrix $X(g) \in GL_d$ such that

- $X(g_1g_2) = X(g_1)X(g_2)$ for all $g_1,g_2 \in \mathcal{G}$.

Clearly, this property implies $X(\mathrm{id}) = I$.

Example B.1 (cf. Example 1.2.4 in [Sag01])**.** *The following representation of $\mathcal{S}_n$ is called the* defining representation *for $\mathcal{S}_n$. Let $\pi \in \mathcal{S}_n$ and define the matrix $X(\pi) \in \{0,1\}^{n\times n}$ by*

$$X(\pi)_{i,j} = \begin{cases} 1 & \text{if } \pi(i) = j, \\ 0 & \text{otherwise.} \end{cases}$$

It is easy to see that this is a matrix representation.

If a group $\mathcal{G}$ acts on a vector space V, then the action of $g \in \mathcal{G}$ on an element $v \in V$ can be represented by a matrix $X(g)$, where X is a matrix representation of $\mathcal{G}$, i.e.,

$$g(v) = X(g)v, \text{ where } \forall g \in \mathcal{G}, v \in V.$$

We continue by defining $\mathcal{G}$-*modules*. For this let V be a finite dimensional vector space over the complex numbers $\mathbb{C}$. The general linear group of V, denoted

by $GL(V)$, consists of all invertible linear transformations from V to itself. Let $\dim(V) = d$. Choosing a basis for V, every linear transformation can be written as a matrix and hence $GL(V)$ and GL_d are isomorphic as groups.

Definition B.2. Let V be a vector space and $\mathcal{G}$ be a group. If there exists a group homomorphism

$$\rho : \mathcal{G} \to GL(V),$$

then we call (V, ρ) a $\mathcal{G}$-module.

Further, we call two $\mathcal{G}$-modules $(V_1, \rho_1), (V_2, \rho_2)$ *equivalent*, if there exists a linear homomorphism $\phi : V_1 \to V_2$ such that $\rho_1(\pi) = \phi \circ \rho_2(\pi) \circ \phi^{-1}$ for all $\pi \in \mathcal{G}$. We now aim to study how $\mathcal{G}$-modules can be decomposed into simpler objects.

Definition B.3. Let V be a $\mathcal{G}$-module. We call a subspace $W \subset V$ a proper $\mathcal{G}$-submodule, if W is a proper subspace of V and W is closed under the action of $\mathcal{G}$, i.e., for all $w \in W$ one has $\pi(w) \in W$ for all $\pi \in \mathcal{G}$. We may also say that W is *$\mathcal{G}$-invariant*.

A $\mathcal{G}$-module V is called *reducible*, if there exists a proper submodule W of V. If V has as its only $\mathcal{G}$-invariant subspaces the spaces V and $\{0\}$, then V is called *irreducible*.

Maschke's theorem states that any non-zero (finite dimensional) $\mathcal{G}$-module can be decomposed into a finite number of irreducible submodules.

Theorem B.4. *(Maschke's Theorem) Let $\mathcal{G}$ be a finite group. If V is a non-zero $\mathcal{G}$-module, then*

$$V \simeq \bigoplus_{i=1}^{k} W^{(i)}, \tag{B.1}$$

where each $W^{(i)}$ is an irreducible submodule of V.

Having a decomposition as above, Schur's Lemma gives insight about the relationship of its constituents.

Lemma B.5. *(Schur's Lemma) Let V, W be two irreducible $\mathcal{G}$-modules and let $\theta : V \to W$ be a $\mathcal{G}$-homomorphism. Then either*

- *θ is a multiple of the identity, or*
- *θ is the zero map.*

What this lemma tells us is that the components of the Maschke decomposition (B.1) are either isomorphic or inequivalent. Grouping the pairwise isomorphic components together into $V^{(i)}$ we find that for a $\mathcal{G}$-module there is a decomposition

$$V = \bigoplus_{i=1}^{\ell} V^{(i)} = \bigoplus_{i=1}^{\ell} \bigoplus_{j=1}^{m_i} V_{i,j},$$

with finite m_i. For fixed i, the $V_{i,j}$ are pairwise isomorphic irreducible $\mathcal{G}$-submodules, and if $i_1 \neq i_2$, then $V_{i_1,j}$ and $V_{i_2,k}$ are inequivalent. The decomposition into the *isotypic components* $V^{(i)}$ is unique, while the decomposition into the $V_{i,j}$ is not unique in general.

B.2 Symmetric polynomials

Our interest lies in the case where the set on which the symmetric group acts is the ring of polynomials in n variables, i.e., $\mathbb{R}[\mathbf{x}]$. We refer to [GP04] for more information of symmetry reduction in polynomial optimization. An element $\pi \in \mathcal{S}_n$ acts on a polynomial $p \in \mathbb{R}[\mathbf{x}]$ by permuting the (indices of the) variables, i.e., $\pi(p) = p(\pi(\mathbf{x})) = p(x_{\pi(1)}, \ldots, x_{\pi(n)})$. We call a polynomial invariant under $\mathcal{S}_n$ if $\pi(p) = p$ for all $\pi \in \mathcal{S}_n$. Furthermore, we define the set of symmetric polynomials as $\mathbb{R}[\mathbf{x}]^{\mathcal{S}_n}$. Given a polynomial $p \in \mathbb{R}[\mathbf{x}]$ we can symmetrize it with respect to a group $\mathcal{G}$ using the *Reynolds operator*

$$\mathcal{R}^{\mathcal{G}}(p) := \frac{1}{|\mathcal{G}|} \sum_{\pi \in \mathcal{G}} \pi(p).$$

If the group is clear from the context we may omit the superscript and simply write $\mathcal{R}(p)$. The symmetry reduction for trigonometric polynomials, i.e., the case we considered in Chapter 5 works the same as the one carried out in the following. Essentially, our goal is to block-diagonalize a matrix $M \succeq 0$ whose rows and columns are indexed by $\alpha \in \mathbb{N}_r^n$ and is invariant under $\mathcal{S}_n$, i.e., $M_{\alpha,\beta} = M_{\pi(\alpha),\pi(\beta)}$ for all $\pi \in \mathcal{S}_n$. Whether this matrix gives rise to a sum of squares in the ordinary or the trigonometric framework makes no difference.

B.2.1 Block-diagonalization

Our aim from now on is to construct a basis which allows us to block-diagonalize the SDP formulation of a symmetric polynomial optimization problem of the fol-

lowing form

$$\begin{aligned} &\min \langle C, M \rangle \\ &\text{s.t. } [\mathbf{x}]_r^\top M [\mathbf{x}]_r = p(\mathbf{x}) \\ &\quad M \succeq 0, \end{aligned} \tag{B.2}$$

for a polynomial $p \in \mathbb{R}[\mathbf{x}]^{\mathcal{S}_n}$ and given $C \in \mathbb{S}^n$. The idea is roughly as follows. Sums of squares of degree less than $2r$ can be written as $[\mathbf{x}]_r^\top A [\mathbf{x}]_r = \langle A, [\mathbf{x}]_r [\mathbf{x}]_r^\top \rangle$ for a positive semidefinite matrix A, where $[\mathbf{x}]_r$ is a monomial basis vector, e.g., as defined in (1.12). The idea of the block diagonalization is to find a change-of-basis matrix B such that we can write symmetric sums of squares as $\mathcal{R}\left(B^\top [\mathbf{x}]_r^\top A B [\mathbf{x}]_r\right) = \langle A, \mathcal{R}\left(B [\mathbf{x}]_r [\mathbf{x}]_r^\top B^\top\right) \rangle$ for $A \succeq 0$ where $\mathcal{R}\left(B [\mathbf{x}]_r [\mathbf{x}]_r^\top B^\top\right)$ is a block-diagonal matrix and the Reynolds operator is applied entry-wise. Let $B_1, \ldots, B_\ell$ be the blocks of $\mathcal{R}\left(B [\mathbf{x}]_r [\mathbf{x}]_r^\top B^\top\right)$. Then we can write a symmetric sum of squares $p(\mathbf{x})$ as

$$p(\mathbf{x}) = \langle A, \mathcal{R}\left(B [\mathbf{x}]_r [\mathbf{x}]_r^\top B^\top\right) \rangle = \sum_{i=1}^{\ell} \langle A_i, B_i \rangle,$$

for $A \succeq 0$ and the A_i are the blocks of A corresponding to B_i. In the final decomposition we may find multiple copies of the same blocks B_i. By convexity, we can assume that the corresponding A_i are identical as well, allowing us to delete all but one of the copies. It is therefore enough to enforce the psd constraint on the blocks A_i instead of the whole matrix A. This can result in a significant decrease of the size of the underlying SDP. We will interpret $\mathbb{R}[\mathbf{x}]_r$ as an $\mathcal{S}_n$- module. Then we can use a structured decomposition as in (B.1) to obtain a so-called *symmetry adapted basis*, which leads to a block-diagonal A. Let V be a $\mathcal{G}$-module with associated orthogonal representation $\rho : \mathcal{G} \to GL(V)$. We know there exists a decomposition of V as follows:

$$\begin{aligned} V &= m_1 V_1 \oplus \cdots \oplus m_\ell V_\ell \\ &= \left(V_{1,1} \oplus \cdots \oplus V_{1,m_1}\right) \oplus \cdots \oplus \left(V_{\ell,1} \oplus \cdots \oplus V_{\ell,m_\ell}\right), \end{aligned}$$

where for fixed $i \in [\ell]$ the irreducible n_i-dimensional modules $V_{i,j}$ are equivalent for all $j \in [m_i]$ and V_{i,k_1}, V_{j,k_2} are inequivalent for $i \neq j$. Let a basis for $V_{i,k}$ be given by

$$V_{i,k} = \operatorname{span}\{s_{j,k}^i : j = 1, \ldots, n_i\}.$$

This basis can be chosen to be orthonormal with respect to the standard inner product $(s_{j,1}^i)^\top s_{\ell,1}^i = \delta_{j,\ell}$. For fixed i let $\varphi_i : \mathcal{G} \to O(n_i)$ be an orthogonal representation of the irreducible module V_i.

We call a basis of V given by

$$\mathcal{B} = \{s^i_{j,k} : i \in [\ell], j \in [n_i], k \in [m_i]\}$$

a *symmetry adapted basis* if the bases of the pairwise equivalent copies $V_{i,k}$ transform according to the same orthogonal representation φ_i we are about to define.

In other words, $\mathcal{B}$ is a symmetry adapted basis, if for fixed $i \in [\ell]$ and all $k \in [m_i]$ the vector of basis elements $(s^i_{1,k}, \ldots, s^i_{n_i,k})$ satisfies for all $g \in \mathcal{G}$

$$\begin{pmatrix} \rho(g)s^i_{1,k} \\ \rho(g)s^i_{2,k} \\ \vdots \\ \rho(g)s^i_{n_i,k} \end{pmatrix} = \varphi_i(g) \begin{pmatrix} s^i_{1,k} \\ s^i_{2,k} \\ \vdots \\ s^i_{n_i,k} \end{pmatrix}.$$

On the left-hand-side we have a vector of images of elements of V under ρ. On the right-hand-side we can interpret φ_i as a basis transformation, where the output is a vector of linear combinations of basis elements of the submodule $V_{i,k}$. As ρ is orthogonal by assumption, we can assume that the given basis $\mathcal{B}$ is orthonormal, i.e.,

$$\langle s^{i_1}_{j_1,k_1}, s^{i_2}_{j_2,k_2} \rangle = \left(s^{i_1}_{j_1,k_1}\right)^\top s^{i_2}_{j_2,k_2} = \delta_{\{i_1,j_1,k_1\},\{i_2,j_2,k_2\}}$$

When dealing with polynomials up to a fixed degree $r \in \mathbb{N}$ that are symmetric with respect to a group $\mathcal{G}$, i.e., $\mathbb{R}[\mathbf{x}]_r$, the situation is essentially the same. We can view $\mathbb{R}[\mathbf{x}]_r$ as a vector space over $\mathbb{R}$ of dimension $s(n,r)$, where we associate to each polynomial its unique coefficient vector. Then $\mathbb{R}[\mathbf{x}]_r$ becomes a $\mathcal{G}$-module for suitable $\rho : \mathcal{S}_n \to GL(\mathbb{R}^{s(n,r)})$. The following theorem states that rewriting a sum of squares problem in a symmetry adapted basis block-diagonalizes the problem into $\sum_{i=1}^{\ell} n_i$ blocks, i.e, for each $i \in [\ell]$ there are n_i blocks of size $m_i \times m_i$. Moreover, there will be n_i copies of the same block for every $i \in [\ell]$. It is therefore enough to consider ℓ blocks of size $m_i \times m_i$.

Theorem B.6. *(see [Bro22, Theorem 7.8], cf. [GP04, §5]) Let $f \in \mathbb{R}[\boldsymbol{x}]^{\mathcal{G}}$ be a $\mathcal{G}$ invariant polynomial that can be written as a sum of squares, where each term has degree at most r. Let $\{s^i_{j,k} \in \mathbb{R}[\boldsymbol{x}] : i \in [\ell], j \in [n_i], k \in [m_i]\}$ be an orthonormal symmetry adapted basis of $\mathbb{R}[\boldsymbol{x}]_r$ (as a $\mathcal{G}$-module). Then,*

$$f = \sum_{i=1}^{\ell} \sum_{t=1}^{h_i} \mathcal{R}\left(\left(\sum_{k=1}^{m_i} \alpha_{i,t,k} s^i_{1,k}\right)^2\right),$$

for $h_i \in \mathbb{N}$. Equivalently, we can write

$$f = \sum_{i=1}^{\ell} \langle A_i, S_i \rangle ,$$

for $A_i \in \mathbb{S}_+^{m_i}$ and

$$S_i = \left(\mathcal{R}(s^i_{1,k} s^i_{1,l}) \right)_{k,l=1,\dots,m_i} \in \mathbb{R}[\boldsymbol{x}]^{m_i \times m_i},$$

where we define $\mathcal{R}(s^i_{1,k} s^i_{1,l}) := \mathcal{R}\left(\left(s^i_{1,k} \right)^\top [\boldsymbol{x}]_r [\boldsymbol{x}]_r^\top s^i_{1,\ell} \right)$.

For a detailed proof of this statement we refer to [Bro22].

B.3 Specht modules

The ring of polynomials $\mathbb{R}[\mathbf{x}]_r$ under the action of the symmetric group becomes an $\mathcal{S}_n$-module. Our goal is to decompose this module into its irreducible components and construct a symmetry adapted basis to block-diagonalize a given symmetric optimization problem. The motivation is that irreducible modules of a group $\mathcal{G}$ correspond to the conjugacy classes of $\mathcal{G}$. As we will see, the conjugacy classes of $\mathcal{S}_n$ are given by the so-called *cycle-types*, which correspond to *partitions* of n.

B.3.1 Partitions and cycle types

When characterizing elements of $\mathcal{S}_n$, we will find that it is useful to associate an element $\pi \in \mathcal{S}_n$ with its cycle type. Given $\pi \in \mathcal{S}_n$ we can represent its action by fixing one element $i \in [n]$ and considering the cycle

$$\pi(i), \pi^2(i), \dots, \pi^{k-1}(i),$$

where k is such that $\pi^k(i) = i$. Taking next an element $j \in \{1, \dots, n\} \setminus \{\pi^m(i) : m = 1, \dots, k-1\}$ we can reiterate this process until we have captured all elements as part of exactly one cycle.

Example B.7. *Consider the permutation $\pi \in \mathcal{S}_5$ defined by*

$$\pi(1) = 2, \pi(2) = 1, \pi(3) = 5, \pi(4) = 4, \pi(5) = 3.$$

The corresponding cycles are given by

$$(1,2)(3,5)(4).$$

We define the *cycle type* of a permutation as

$$(1^{m_1}, 2^{m_2}, \ldots, n^{m_n})$$

where m_i denotes the number of cycles of length i for the given permutation. For the above example this would be $(1^1, 2^2, 3^0, 4^0, 5^0)$. We continue by defining a *partition* of n.

Definition B.8. A *partition* $\lambda = (\lambda_1, \ldots, \lambda_k)$ of an integer $n \in \mathbb{N}$ is defined as a weakly decreasing sequence consisting of positive integers such that

$$\sum_{i=1}^{k} \lambda_i = n.$$

We write $\lambda \vdash n$ if λ is a partition of n.

Cycle types correspond to partitions when they are sorted by the size of the cycles. In the aforementioned example the corresponding partition would be $\lambda = (2, 2, 1)$. We call two elements $a, b \in \mathcal{G}$ of a group $(\mathcal{G}, *)$ *conjugate* if there exists $g \in \mathcal{G}$ such that $a = g * b * g^{-1}$. Conjugacy is an equivalence relation. The corresponding equivalence class is called *conjugacy class*. Two elements $\pi, \sigma \in \mathcal{S}_n$ are in the same conjugacy class if and only if their cycle types are equal. This means the conjugacy classes of $\mathcal{S}_n$ may be identified with the set of partitions of n.

For a permutation π we can define its *sign*, denoted by $\operatorname{sgn}(\pi)$ in the following way. Consider the conjugacy class of transpositions, i.e., elements of $\mathcal{S}_n$ of cycle type $(1^0, 2^1, 3^0, \ldots, n^0)$. These are all permutations swapping exactly two elements in $\{1, \ldots, n\}$. Every permutation $\pi \in \mathcal{S}_n$ can be written as the composition of transpositions. Let $\pi = \tau_1 \ldots \tau_k$, where τ_i is a transposition for $i \in [k]$. Then

$$\operatorname{sgn}(\pi) = (-1)^k.$$

It can be shown that this is well-defined. We will make use of these concepts later when we show how to block-diagonalize a symmetric optimization problem of the form (B.2).

The whole point now is to decompose the $\mathcal{S}_n$-module $\mathbb{R}[\mathbf{x}]$ in its irreducible submodules to achieve the block-diagonalization presented above in Theorem B.6. For this we need to define *Young tableaux*.

Definition B.9. Given a partition $\lambda \vdash n$ with $\lambda = (\lambda_1, \ldots, \lambda_k)$ a *Young tableau* consists of k rows aligned to the left containing λ_i elements each. All numbers $\{1, \ldots, n\}$ occur exactly once in a Young tableau. A Young tableau is called a *standard Young tableau* if all rows and columns are increasing.

Example B.10. *Let $\lambda = (4, 2)$ with $n = 6$. Then a Young tableau is given by*

$$t = \begin{array}{|c|c|c|c|}\hline 5 & 1 & 6 & 4 \\ \hline 2 & 3 \\ \cline{1-2}\end{array}.$$

For the same λ a standard Young tableau is given by

$$t = \begin{array}{|c|c|c|c|}\hline 1 & 2 & 3 & 4 \\ \hline 5 & 6 \\ \cline{1-2}\end{array}.$$

A permutation $\pi \in \mathcal{S}_n$ acts on Young tableaux by replacing the elements in the cells by their images under π. We call two Young tableaux row equivalent, if the elements in the respective rows are the same. For a given Young tableau t the class of row equivalent Young tableaux is called a tabloid and denoted by $\{t\}$. Tabloids are display with horizontal lines only. For example, consider the tableau

$$t = \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 & 5 \\ \hline 4 \\ \cline{1-1}\end{array},$$

then we write

$$\{t\} = \left\{ \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 & 5 \\ \hline 4 \\ \cline{1-1}\end{array}, \begin{array}{|c|c|}\hline 3 & 1 \\ \hline 2 & 5 \\ \hline 4 \\ \cline{1-1}\end{array}, \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 5 & 2 \\ \hline 4 \\ \cline{1-1}\end{array}, \begin{array}{|c|c|}\hline 3 & 1 \\ \hline 5 & 2 \\ \hline 4 \\ \cline{1-1}\end{array} \right\} = \begin{array}{c}\overline{1\ 3} \\ \overline{2\ 5} \\ \overline{4}\end{array},$$

to indicate the order of the elements within each row does not matter. Note that for any $\pi \in \mathcal{S}_n$ we have that $\pi\{t\} = \{\pi t\}$.

With these object we can define permutation modules.

Definition B.11. Let $\lambda \vdash n$. The permutation module M^λ corresponding to λ is the $\mathcal{S}_n$-module defined by

$$M^\lambda := \operatorname{span}\{\{t_1\}, \ldots, \{t_k\}\},$$

where $\{t_1\}, \ldots, \{t_k\}$ is a complete list of λ-tabloids.

Since the $\{t_1\}, \ldots, \{t_k\}$ form a complete list of λ-tabloids, it is clear that the resulting vector space M^λ is an $\mathcal{S}_n$-module. Note that a permutation module is not irreducible in general. It can be further decomposed into so-called *Specht modules*.

These objects are pairwise inequivalent and constitute a full set of irreducible submodules of the symmetric group. The next step is show how to decompose permutation modules into Specht modules.

For $\lambda \vdash n$ consider a Young tableau t and let $\mathcal{C}_1, \ldots, \mathcal{C}_k$ denote the columns of t. Then the group $\mathcal{C}_t = S_{\mathcal{C}_1} \times \cdots \times S_{\mathcal{C}_k}$ is called the column stabilizer of t, where $S_{\mathcal{C}_i}$ is the symmetric group over the content of column i. The column stabilizer allows us to define polytabloids.

Definition B.12. Let $\lambda \vdash n$ and t a Young tableau of shape λ. Then we define the *polytabloid* e_t as

$$e_t := \sum_{\pi \in \mathcal{C}_t} \operatorname{sgn}(\pi)\{\pi t\}.$$

It is clear that for a given t with corresponding λ that $e_t \in M^\lambda$.

Definition B.13. Let $\lambda \vdash n$. The *Specht module* corresponding to λ is

$$S^\lambda := \operatorname{span}\{e_t : t \text{ is a tableau of shape } \lambda\}.$$

Clearly, for a partition λ the Specht module S^λ is contained in the permutation module M^λ. To describe all Specht modules that are contained in a permutation module M^λ we need to define the *domination ordering* of partitions.

Definition B.14. Let $\lambda, \mu \vdash n$. We say μ is dominating λ, writing $\mu \trianglerighteq \lambda$, if

$$\mu_1 + \cdots + \mu_i \geq \lambda_1 + \cdots + \lambda_i \text{ for all } i \geq 1,$$

where we set $\mu_i, \lambda_i = 0$ if i is not contained in their index set.

Consider now a *generalized* tableau t, i.e., a tableau where we allow for repetition of entries. Such a tableau is called *semistandard* if all entries along the columns are strictly increasing and along the rows are weakly increasing. For a given generalized tableau t we define the content of t to be the array μ where μ_i is the number of occurrences of i in t. Take for example the generalized semistandard tableau

1	1	3	4	4
3	4	4		
5	5	6		
6				

.

The content is $\mu = (2, 0, 2, 4, 2, 2)$. Important for us are the semistandard generalized Young-tableaux for which the content is a partition itself. Given a

generalized semistandard tableau t of shape λ it is not difficult to see that if the content μ of t is a partition, then $\lambda \trianglerighteq \mu$. The following theorem states how permutation modules decompose into Specht modules.

Theorem B.15. *The permutation module M^λ decomposes into the (irreducible) Specht modules*

$$M^\lambda = \bigoplus_{\mu \trianglerighteq \lambda} K_{\mu\lambda} S^\mu,$$

where the multiplicities are given by the Kostka-numbers $K_{\mu\lambda}$, which is the number of semistandard generalized Young-tableaux with shape μ and content λ.

Note that above the S^μ are pairwise inequivalent, and for $m \in \mathbb{N}$ we recall

$$mS^\mu = \underbrace{S^\mu \oplus \cdots \oplus S^\mu}_{m \text{ times}}.$$

We will now show how to construct a symmetry adapted basis when $V = M^\lambda$. Let T be a generalized semistandard Young tableau of shape μ and with content λ. This object defines an isomorphism $\vartheta_T : M^\mu \to M^\lambda$ in the following way. Let t be a Young tableau of shape μ which we intend to send to a tableau of shape λ. Define $t(i,j)$ to be the entry in row i column j of t. Mapping $t(i,j)$ of t to the row given by the entry $T(i,j)$ will result in a tabloid of shape λ. Moreover, summing these up over all elements in $\{T\}$ we get a function that is constant over all elements of $\{t\}$. Given a Young tableau t of shape μ and T a generalized Young tableau of shape μ and content μ we define $t[T]$[1] as the tabloid of shape λ where we find each entry $t(i,j)$ in row $T(i,j)$. We define

$$\begin{aligned} \vartheta_T : M^\mu &\to M^\lambda \\ \{t\} &\mapsto \sum_{T' \in \{T\}} \{t[T']\}. \end{aligned}$$

Example B.16. *Let $\mu = (3,2,1) \trianglerighteq (2,2,2) = \lambda$. Then there are two semistandard generalized Young tableaux of shape μ and content λ, namely*

$$T_1 = \begin{array}{|c|c|c|}\hline 1&1&2\\\hline 2&3\\\cline{1-2} 3\\\cline{1-1}\end{array} \text{ and } T_2 = \begin{array}{|c|c|c|}\hline 1&1&3\\\hline 2&2\\\cline{1-2} 3\\\cline{1-1}\end{array}.$$

[1] This notation was introduced in [Bro22].

Let us demonstrate what ϑ_{T_1} does. For this, let

$$t = \begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline 4 & 5 \\ \cline{1-2} 6 \\ \cline{1} \end{array}\,.$$

Then,

$$\begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline 4 & 5 \\ \cline{1-2} 6 \\ \cline{1} \end{array} \overset{\vartheta_{T_1}}{\rightarrow} \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & 4 \\ \cline{1-2} 5 & 6 \\ \cline{1-2} \end{array} + \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 & 4 \\ \cline{1-2} 5 & 6 \\ \cline{1-2} \end{array} + \begin{array}{|c|c|}\hline 2 & 3 \\ \hline 1 & 4 \\ \cline{1-2} 5 & 6 \\ \cline{1-2} \end{array} + \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & 5 \\ \cline{1-2} 4 & 6 \\ \cline{1-2} \end{array} + \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 & 5 \\ \cline{1-2} 4 & 6 \\ \cline{1-2} \end{array} + \begin{array}{|c|c|}\hline 2 & 3 \\ \hline 1 & 5 \\ \cline{1-2} 4 & 6 \\ \cline{1-2} \end{array}\,.$$

For a given μ the image of the Specht module $S^\mu \subseteq M^\mu$ under the homomorphisms defined by generalized semistandard tableaux of shape μ and content λ decompose the permutation module M^λ fully into irreducible submodules. Note that there will be many copies of the same Specht module in the final decomposition.

"Recipe":

Let us construct a symmetry adapted basis with respect to $\mathcal{S}_n$ for $\mathbb{R}[\mathbf{x}]_r$.

Step 1: Decompose a given basis of $\mathbb{R}[\mathbf{x}]_r$ into orbits of $\mathcal{S}_n$ and identify which partitions they correspond to. For a monomial $\mathbf{x}^\alpha$ the corresponding orbit is defined as $\mathcal{O}(\mathbf{x}^\alpha) = \{\mathbf{x}^{\sigma(\alpha)} : \sigma \in \mathcal{S}_n\}$. For each orbit we can choose a representative monomial $\mathbf{x}^\alpha$ such that α is sorted, i.e., $\alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_n$. We will denote the corresponding orbit by $\mathcal{O}_\alpha$ or $\mathcal{O}(\mathbf{x}^\alpha)$. Clearly, we have that $\operatorname{span}\{\mathbf{x}^\beta \in \mathcal{O}_\alpha\}$ is an $\mathcal{S}_n$-module. The next step is identify orbits with permutation modules. Let $\mathcal{O}_\alpha$ be an orbit. Now define the set $\mathcal{I} := \{j : \alpha_i = j \text{ for some } i \in n\}$, i.e., the set of all appearing exponents in the monomial $\mathbf{x}^\alpha$. Further, let $b_j = |\{k : \alpha_k = i_j\}|$ for $i_j \in \mathcal{I}$. The vector b indicates how often an exponent $i_j \in \mathcal{I}$ appears in the monomial $\mathbf{x}^\alpha$. Assume without loss of generality that b is sorted, i.e., $b_1 \geq b_2 \geq \cdots \geq b_{|\mathcal{I}|}$. By construction b is a partition of n. Moreover, b is the partition of the permutation

module corresponding to the given orbit. For example

$$\begin{aligned}
\operatorname{span}(\mathcal{O}(1)) &= \operatorname{span}(1) \simeq M^{(n)} \\
\operatorname{span}(\mathcal{O}(x_1)) &= \operatorname{span}(x_1, \dots, x_n) \simeq M^{(n-1,1)} \\
\operatorname{span}\left(\mathcal{O}(x_1^2)\right) &= \operatorname{span}\left(x_1^2, \dots, x_n^2\right) \simeq M^{(n-1,1)} \\
\operatorname{span}(\mathcal{O}(x_1x_2)) &= \operatorname{span}\left(x_ix_j : i < j\right) \simeq M^{(n-2,2)} \\
\operatorname{span}\left(\mathcal{O}(x_1^3)\right) &= \operatorname{span}\left(x_1^3, \dots, x_n^3\right) \simeq M^{(n-1,1)} \\
\operatorname{span}\left(\mathcal{O}(x_1^2x_2)\right) &= \operatorname{span}\left(x_i^2x_j : i < j\right) \simeq M^{(n-2,1,1)} \\
\operatorname{span}(\mathcal{O}(x_1x_2x_3)) &= \operatorname{span}\left(x_ix_jx_k : i < j < k\right) \simeq M^{(n-3,3)} \\
&\vdots \\
\operatorname{span}\left(\mathcal{O}(x_1^r)\right) &= \operatorname{span}\left(x_1^r, \dots, x_n^r\right) \simeq M^{(n-1,1)}.
\end{aligned}$$

Example B.17. *Consider the monomial $x_1^3x_3^2x_4^3x_5^2x_6^3x_7$ where we assume $n = 7$. Sorting the exponent vector we find this is in the same orbit as $x_1^3x_2^3x_3^3x_4^2x_5^2x_6$. Therefore, the corresponding partition will be $(3,2,1,1)$, and we can associate it to a tabloid:*

$$x_1^3x_2^3x_3^3x_4^2x_5^2x_6 \longleftrightarrow \begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline 4 & 5 \\ \cline{1-2} 6 \\ \cline{1-1} 7 \\ \cline{1-1}\end{array} \in M^{(3,2,1,1)}.$$

Step 2: After identifying the appearing monomials with tabloids we can decompose the permutation modules above into Specht modules. For every orbit $\mathcal{O}_\alpha$ we do the following. Let M^λ be the permutation module corresponding to $\mathcal{O}_\alpha$ and let $\{\mu^{(1)}, \dots, \mu^{(k)}\}$ be a complete list of partitions $\mu^{(i)} \vdash n$ such that $\mu^{(i)} \trianglerighteq \lambda$. Each $\mu^{(i)}$ defines a Specht module

$$S^{\mu^{(i)}} = \operatorname{span}\{e_t : t \text{ is a Young tableau of shape } \mu^{(i)}\}.$$

By Theorem B.6 it is enough to consider a single element in the spanning set of $S^{\mu^{(i)}}$. Hence, we will for any partition $\mu^{(i)}$ only consider the tableau $t^{\mu^{(i)}}$ of shape $\mu^{(i)}$ with strictly increasing rows and columns where for each entry we have $t^{\mu^{(i)}}(\ell, m+1) = t^{\mu^{(i)}}(\ell, m) + 1$ if $m + 1 \leq \mu_\ell^{(i)}$. For every $\mu^{(i)}$ let $\{T_1, \dots, T_{k_i}\}$ be a complete list of generalized semistandard tableaux with shape $\mu^{(i)}$ and content λ. Each T_j defines a homomorphism $\vartheta_{T_j} : S^{\mu^{(i)}} \to M^\lambda$. The corresponding elements of the symmetry adapted basis are given by $\vartheta_{T_j}(t^{\mu^{(i)}})$ interpreted in terms of the orbit $\mathcal{O}_\alpha$.

Example B.18. *Let $n = 3$ and say we want to decompose the permutation module $M^{(2,1)}$ corresponding to the orbit $\mathcal{O}(x_1)$. We will interpret the images of the Specht modules, which will be linear combinations of tabloids, in terms of this orbit. We collect all λ such that $\lambda \trianglerighteq (2,1)$, i.e., $\lambda^{(1)} = (3), \lambda^{(2)} = (2,1)$. Take $\lambda^{(1)} = (3)$. How many generalized semistandard tableaux T with shape (3) and content $(2,1)$ do exist? Only one, namely*

$$T_1 = \begin{array}{|c|c|c|}\hline 1 & 1 & 2 \\ \hline \end{array}.$$

How many generalized semistandard tableaux T with shape $(2,1)$ and content $(2,1)$ exist? Also, only one:

$$T_2 = \begin{array}{|c|c|}\hline 1 & 1 \\ \hline 2 \\ \cline{1-1} \end{array}.$$

From this follows that the permutation module $M^{(2,1)}$ decomposes as

$$M^{(2,1)} = S^{(3)} \oplus S^{(2,1)}.$$

We have yet to determine the corresponding map from the Specht module to elements in $\mathbb{R}[\boldsymbol{x}]$. Let us compute the image of the Specht module $S^{(3)}$ under ϑ_{T_1}. Note that

$$S^{(3)} = \operatorname{span}\{e_t : t \text{ is a Young tableau of shape } (3)\},$$

i.e.,

$$\begin{aligned}\vartheta_T\left(\begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline \end{array}\right) &= \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1} \end{array} + \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1} \end{array} + \begin{array}{|c|c|}\hline 2 & 3 \\ \hline 1 \\ \cline{1-1} \end{array} \\ &\simeq x_3 + x_2 + x_1.\end{aligned}$$

This polynomial $x_1 + x_2 + x_3$ will be one basis element of the subbasis belonging to the appearing $S^{(3)}$ modules. Consider the next Specht module $S^{(2,1)}$. We let

$$t = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1} \end{array}$$

and T_2 is the only generalized semistandard tableau with shape $(2,1)$ and content $(2,1)$. We see that

$$e_t = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1} \end{array} - \begin{array}{|c|c|}\hline 2 & 3 \\ \hline 1 \\ \cline{1-1} \end{array},$$

and so

$$\vartheta_{T_2}(e_t) = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1} \end{array} - \begin{array}{|c|c|}\hline 2 & 3 \\ \hline 1 \\ \cline{1-1} \end{array} \simeq x_3 - x_1.$$

This polynomial will be an element of the basis belonging to $S^{(2,1)}$. Similarly, we can go about decomposing the other permutation modules to obtain a complete symmetry adapted basis.

C

Additional notes on approximation kernels

C.1 A note on Jackson kernels

Our exposition here follows closely Section 5.4 of [AC11]. Our goal is to expose the relations and differences between the Jackson kernel from [Jac11] and the one from [WWAF06].

Dirichlet and Fejér kernels

The story begins with the approximation of *trigonometric polynomials*. We define the space of functions f that are Lebesgue integrable over $[-\pi, \pi]$ and satisfy $f(x) = f(x + 2\pi)$ almost everywhere on $\mathbb{R}$ as $L^1_{2\pi}$. We further define the ℓ_1-norm on $L^1_{2\pi}$:

$$\|f\| := \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)| \mathrm{d}x. \tag{C.1}$$

Let us define periodic kernels and approximate identities. A sequence of functions $\{\chi_r\}_{r\in\mathbb{N}}$ in $L^1_{2\pi}$ is called a *(periodic) kernel* if

$$\lim_{r\to\infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} \chi_r(x) \mathrm{d}x = 1.$$

The kernels that we consider are *normalized* for any $r \in \mathbb{N}$, meaning $\int_{-\pi}^{\pi} \chi_r(x)\mathrm{d}x = 2\pi$ for all $r \in \mathbb{N}$. We say a kernel is non-negative if $\chi_r(x) \geq 0$ for all $x \in [-\pi, \pi]$ and $r \in \mathbb{N}$. We further refer to a periodic kernel as an *approximate identity* if there exists $N \in \mathbb{N}$ such that the sequence $\{\chi_r\}_{r \geq N}$ is bounded in $L^1_{2\pi}$ with respect to the norm defined in (C.1) and moreover,

$$\lim_{r\to\infty} \left(\int_{-\pi}^{-\delta} |\chi_r(x)|\mathrm{d}x + \int_{\delta}^{\pi} |\chi_r(x)|\mathrm{d}x \right) = 0,$$

for every $\delta \in \mathbb{R}$, $0 < \delta < \pi$. For non-negative kernels, the approximate identity property therefore implies that the χ_r approximate the Dirac-δ at zero as $r \to \infty$. More precisely, $\{\chi_r/2\pi\}_{r\in\mathbb{N}}$ may be viewed as a series of probability density functions, where the mass is increasingly concentrated around zero, and that converge weakly to the Dirac-δ at zero. For a fixed kernel $\{\chi_r\}_{r\in\mathbb{N}}$ the convolution with a function $f \in L^1_{2\pi}$, i.e.,

$$f * \chi_r(x) := \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x-t)\chi_r(t)\mathrm{d}t = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t)\chi_r(x-t)\mathrm{d}t, \quad r \in \mathbb{N},$$

defines a sequence of linear operators $L_r : L^1_{2\pi} \to L^1_{2\pi}$ for $r \in \mathbb{N}$ via $L_r(f) := f * \chi_r$. If $\{\chi_r\}_{r\in\mathbb{N}}$ is non-negative and an approximate identity, then $\lim_{r\to\infty} L_r(f) = f$, uniformly in $L^1_{2\pi}$ for every $f \in L^1_{2\pi}$, (cf. [AC11, Theorem 5.4.1]). Let $\mathcal{C}_{2\pi}$ be the set of all 2π-periodic real functions that are continuous. A real trigonometric polynomial of degree r is defined as

$$u_r(x) = a_0 + 2\sum_{k=1}^{r} (a_k \cos kx + b_k \sin kx), \; x \in \mathbb{R},$$

for $a_0, a_1, \ldots, a_r$ and $b_1, \ldots, b_r$ are real numbers. A sequence of trigonometric polynomials can only uniformly converge to a function in $\mathcal{C}_{2\pi}$. To approximate a function $f \in \mathcal{C}_{2\pi}$ by a trigonometric polynomial of degree r one can use the *Fourier coefficients*. These are defined as

$$a_k = a_k(f) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t)\cos kt\mathrm{d}t, \quad k \geq 0,$$
$$b_k = b_k(f) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t)\sin kt\mathrm{d}t, \quad k \geq 1.$$

With these coefficients we can define a linear operator $S^*_r : \mathcal{C}_{2\pi} \to \mathcal{C}_{2\pi}$, which returns the r-th partial sum of the Fourier series of f:

$$S^*_r(f)(x) := a_0(f) + 2\sum_{k=1}^{r} (a_k(f)\cos kx + b_k(f)\sin kx),$$

which serves as an approximation of f. This approximation may also be interpreted as the orthogonal projection of f onto the subspace of trigonometric polynomials spanned by $(\cos kx, \sin kx)$ for $k = 0, \dots, r$, since this basis is orthogonal w.r.t. the integral inner product. We can also write S_r^* as

$$S_r^*(f)(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x-t) D_r(t) \mathrm{d}t, \quad r \in \mathbb{N},$$

i.e., $S^*(f) = f * D_r$, where D_r is known as the *Dirichlet kernel* (cf. [Dir29]). More precisely, for $r \in \mathbb{N}$ define

$$D_r(t) := 1 + 2 \sum_{k=1}^{r} \cos kt.$$

To see this note that

$$\begin{aligned}
&\frac{1}{2\pi} \int_{-\pi}^{\pi} f(x-t) D_r(t) \mathrm{d}t \\
&\qquad = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) D_r(x-t) \mathrm{d}t \\
&\qquad = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \left(1 + 2 \sum_{i=1}^{r} \cos k(x-t) \right) \mathrm{d}t \\
&\qquad = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \left(1 + 2 \sum_{i=1}^{r} \cos kx \cos kt + \sin kx \sin kt \right) \mathrm{d}t \\
&\qquad = a_0(f) + 2 \sum_{k=1}^{r} (a_k(f) \cos kx + b_k(f) \sin kx) = S_r^*(f)(x),
\end{aligned}$$

where we used the well-known identity

$$\cos(x-y) = \cos x \cos y + \sin x \sin y.$$

In fact, there is a simpler way to write the Dirichlet kernel, namely

$$D_r(t) = \begin{cases} \frac{\sin((2r+1)\frac{t}{2})}{\sin \frac{t}{2}}, & \text{if } t \text{ is not a multiple of } 2\pi, \\ 2r+1 & \text{otherwise.} \end{cases} \tag{C.2}$$

This kernel is not an *approximate identity*, as it is not a bounded sequence in $L^1_{2\pi}$. In fact,

$$\|D_r\| = \frac{4}{\pi^2} \log(r) + O(1), \text{ for } r \gg 0 \text{ (asymptotically)},$$

which was shown in [BN71, Proposition 1.2.3]. From this follows that the Fourier series does not converge uniformly for all functions $f \in \mathcal{C}_{2\pi}$ (cf. [AC11, §5.4]).

An improvement is given by the following kernel. Consider for $r \in \mathbb{N}$ and $x \in \mathbb{R}$ the arithmetic means of the partial sum of the Fourier series. We call the following object the *Fejér kernel*:

$$\varphi_r(t) := \frac{1}{r+1} \sum_{k=0}^{r} D_k(t),$$

with $D_0 := 1$. A simple calculation shows that

$$\varphi_r(t) = 1 + 2 \sum_{k=1}^{r} \left(1 - \frac{k}{r+1}\right) \cos kt\,,$$

and this kernel can be rewritten as

$$\varphi_r(t) = \begin{cases} \frac{\sin^2((r+1)t/2)}{(r+1)\sin^2 t/2}, & \text{if } t \text{ is not a multiple of } 2\pi \\ r+1, & \text{otherwise.} \end{cases}$$

As has been pointed out, one drawback of the Dirichlet kernel is that it is not non-negative. Taking its square and norming it by dividing by

$$\int_{-\pi}^{\pi} \frac{\sin^2((2r+1)\frac{t}{2})}{\sin^2 \frac{t}{2}} \mathrm{d}t = 2r+1,$$

leads to an approximate identity. Consider the normalized non-negative kernel $\tilde{\varphi}_k$ defined as

$$\tilde{\varphi}_k(t) = \frac{\sin^2((2k+1)t/2)}{(2k+1)\sin^2 t/2},$$

which is the Fejér kernel for even $r = 2k$. Clearly, the Fejér kernel is a *generalization* of the normed square of a Dirichlet kernel. Note that for this kernel we have for any $r \in \mathbb{N}$ that

$$\sup_{\delta \le |x| \le \pi} |\varphi_r(x)| \le \frac{1}{(r+1)\sin^2 \delta/2},$$

which implies, in connection with the fact that the kernel is normed and positive, that $(\varphi_r)_{r\in\mathbb{N}}$ is an approximate identity.

C.1.1 Jackson kernels

Let us consider now two kernels A and B. Let the kernel A be a sequence in $L^1_{2\pi}$ defined as

$$A_r(x) := a_r \varphi_r(x)^2,$$

where a_r is chosen such that

$$\frac{1}{a_r} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \varphi_r(t)^2 \mathrm{d}t.$$

In that way it is ensured that the resulting function A_r is a normalized periodic kernel. Now, in these terms the kernel A_r is fully determined. We can find a closed form expression for a_r

$$\frac{1}{a_r} = \frac{1}{(r+1)^2}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\sin^4((r+1)t/2)}{\sin^4 t/2}\mathrm{d}t\right) = \frac{(2(r+1)^2+1)}{3(r+1)}.$$

This means we can write out the kernel A_r explicitly as

$$A_r(t) = \frac{3}{(r+1)(2(r+1)^2)+1}\frac{\sin^4((r+1)t/2)}{\sin^4 t/2}.$$

This object is sometimes referred to as the *Jackson kernel* or *Jackson Convolution operator* (see, e.g., [AC11]). However, it is not the same object which is called the Jackson kernel in [WWAF06]. We will denote the kernel from [WWAF06] as

$$B_r(t) = 1 + 2\sum_{k=1}^{r} g_{k,r}^{\mathrm{KPM}} \cos kt, \text{ where } g_{k,r}^{\mathrm{KPM}} \text{ as in (5.12).}$$

These objects are not the same. To see this consider again the Dirichlet kernel (C.2) for $r = 1$. Taking the fourth power leads to

$$\begin{aligned}\left(\frac{\sin 3x/2}{\sin x/2}\right)^4 &= (1+2\cos x)^4 \\ &= 19 + 32\cos x + 20\cos 2x + 8\cos 3x + 2\cos 4x,\end{aligned}$$

which after normalizing the coefficients by $19 = 1/(2\pi)\int_{-\pi}^{\pi}(1+2\cos x)^4 \mathrm{d}x$ results in a kernel of the form

$$A_4(t) = 1 + 2\sum_{k=1}^{4} g_k \cos kt.$$

We compare the coefficients to those of kernel B_r in Table C.1.

So the kernels A_r and B_r are not the same. They are, however, strongly related, which we will investigate further in the next section.

k	0	1	2	3	4
g_k	1	0.842	0.526	0.211	0.105
$g_{k,4}^{\mathrm{KPM}}$	1	0.866	0.583	0.289	0.083

Table C.1: Comparison of coefficients for $A_4(t)$ and $B_4(t)$.

C.1.2 Jackson's original work

In the following we will look more closely at what Jackson did in his PhD thesis [Jac11] and a publication [Jac12] in the following year. Let f be a real 2π periodic function which is Lipschitz continuous, i.e.,

$$|f(x_2)-f(x_1)| \leq \lambda |x_2-x_1|,$$

for some constant $\lambda \in \mathbb{R}$. Jackson defined the *approximating function*

$$I_r(x) = h_r \int_{-\pi/2}^{\pi/2} f(x+2u)\left(\frac{\sin ru}{r\sin u}\right)^4 \mathrm{d}u,$$

with

$$\frac{1}{h_r} = \int_{-\pi/2}^{\pi/2} \left(\frac{\sin ru}{r\sin u}\right)^4 \mathrm{d}u.$$

He showed that the integrand in the former integral is a trigonometric polynomial of order no more than $2(r-1)$. From there it is not difficult to find that

$$I_r(x)-f(x) = h_r \int_{-\pi/2}^{\pi/2} (f(x+2u)-f(x))\left(\frac{\sin ru}{r\sin u}\right)^4 \mathrm{d}u.$$

Using the fact that the function f is Lipschitz continuous we see that

$$|I_r(x)-f(x)| \leq 2\lambda h_r \int_{-\pi/2}^{\pi/2} |u|\left(\frac{\sin ru}{r\sin u}\right)^4 \mathrm{d}u. \tag{C.3}$$

From here, there are two possible ways to proceed. Jackson continued by arguing

$$|I_r(x)-f(x)| \leq 2\lambda \frac{\int_0^{\pi/2} u\left(\frac{\sin ru}{r\sin u}\right)^4 \mathrm{d}u}{\int_0^{\pi/2}\left(\frac{\sin ru}{r\sin u}\right)^4 \mathrm{d}u},$$

and then finding estimates for denominator and numerator achieving

$$|I_r(x)-f(x)| \leq 2\lambda \frac{1}{r}\left(\frac{\pi}{2}\right)^4 \frac{\int_0^{\infty} \frac{\sin^4 u}{u^3}\mathrm{d}u}{\int_0^{\pi/2}\left(\frac{\sin u}{u}\right)^4 \mathrm{d}u}.$$

The fraction containing integrals on the right-hand-side is a finite constant:

$$\frac{\int_0^\infty \frac{\sin^4 u}{u^3}\mathrm{d}u}{\int_0^{\pi/2}\left(\frac{\sin u}{u}\right)^4 \mathrm{d}u} \leq 2\log(2).$$

With this he proved that his kernel achieved the best possible convergence rate of $O(1/r)$, which had been proved earlier by Bernstein [Ber12].

Another possibility

Going back to the fork (C.3) one can also go about this in the following way[1]

$$\begin{aligned}
|I_r(x) - f(x)| &\leq 2\lambda h_r \int_{-\pi/2}^{\pi/2} |u| \left(\frac{\sin ru}{r\sin u}\right)^4 \mathrm{d}u \\
&= 2\lambda h_r \int_{-\pi/2}^{\pi/2} |u| \left(\left(\frac{\sin ru}{r\sin u}\right)^2\right)^2 \mathrm{d}u \\
&\leq 2\lambda h_r \left(\int_{-\pi/2}^{\pi/2} u^2 \left(\frac{\sin ru}{r\sin u}\right)^4 \mathrm{d}u\right)^{\frac{1}{2}} \left(\int_{-\pi/2}^{\pi/2} \left(\frac{\sin ru}{r\sin u}\right)^4 \mathrm{d}u\right)^{\frac{1}{2}} \\
&= 2\lambda h_r^{\frac{1}{2}} \left(\int_{-\pi/2}^{\pi/2} u^2 \left(\frac{\sin ru}{r\sin u}\right)^4 \mathrm{d}u\right)^{\frac{1}{2}},
\end{aligned}$$

where we used Schwarz's inequality for integrals and the definition of h_r. Since $-\pi/2 \leq u \leq \pi/2$ we can use $u^2 \leq \frac{\pi^2}{2}(1-\cos u)$. Continuing, we find

$$2\lambda h_r^{\frac{1}{2}} \left(\int_{-\pi/2}^{\pi/2} u^2 \left(\frac{\sin ru}{r\sin u}\right)^4 \mathrm{d}u\right)^{\frac{1}{2}} \leq \sqrt{2}\pi\lambda h_r^{\frac{1}{2}} \left(\int_{-\pi}^{\pi} (1-\cos u) \left(\frac{\sin ru}{r\sin u}\right)^4 \mathrm{d}u\right)^{\frac{1}{2}}.$$

Here we changed the domain of integration, but since the integrand is positive the inequality works the right way. Now, it can be shown that

$$\left(\frac{\sin ru}{r\sin u}\right)^4 = c_0 + 2\sum_{k=1}^{2(r-1)} c_k \cos ku,$$

for some coefficients $c_k \in \mathbb{R}$. Using the identity

$$\cos u \cos ku = \frac{1}{2}\left(\cos(k-1)u + \cos(k+1)u\right)$$

[1] Idea from [Riv69] Chapter 1

we see the only term remaining in the integral is c_1, leading to

$$|I_r(x) - f(x)| \leq \sqrt{2}\lambda\pi\sqrt{1-\bar{c}_1},$$

where $\bar{c}_1$ is the coefficient of the normed kernel. This shows that the convergence rate may be bounded in terms of $\sqrt{1-\bar{c}_1}$. This is where the "other" Jackson kernel comes from. One can use a Jackson-type kernel of order r

$$1 + 2\sum_{k=1}^{r} c_k \cos kt$$

with coefficients c_k that are yet to be determined. This kernel is normalized by default. Then, looking for coefficients enforcing non-negativity while maximizing c_1 we find the kernel B_r, which was called Jackson kernel in [WWAF06]. This second construction uses a kernel whose structure is close to the kernel Jackson used in his original work. However, Jackson did not explicitly construct a kernel with coefficients $g_{k,r}^{\mathrm{KPM}}$. Our suggestion is to not refer to the kernel with coefficients $g_{k,r}^{\mathrm{KPM}}$ as the Jackson kernel. In the literature, especially in the approximation community, the Jackson kernel is defined as the kernel Jackson used in his PhD thesis [Jac11]:

$$A_r(t) = \frac{3}{(r+1)(2(r+1)^2)+1}\frac{\sin^4(r+1)t/2}{\sin^4 t/2}.$$

The kernel which was recovered in [WWAF06] should be given a name which reflects that it has minimal resolution among all normed non-negative kernels, e.g., *minimum resolution kernel*. In fact, the only references we were able to find who referred to the minimum resolution kernel as the Jackson kernel were citing [WWAF06].

C.2 Plots and tables

In this section we present some data relevant to Chapter 5. In Table C.2 we compare the sizes of the SDPs (5.26) and (5.29) for different values of n and r. Tables C.3 — C.12 contain the values of the (orbit) coefficients of the minimum resolution kernels, i.e., the solutions of (5.29) for a few values of n and r. In Figure C.1 we compare the approximations resulting from a minimal resolution kernel with the function $(x, y) \mapsto \sin(2\pi x)y$ for different values of r. In Figure C.2 we compare the approximations resulting from a minimal resolution kernel with the peaks

function $p(x,y) := 3(1-x)^2\exp(-x^2-(y+1)^2) - 10(x/5-x^3-y^5)\exp(-x^2-y^2) - (1/3)\exp(-(x+1)^2-y^2)$ for different values of r. In Figure C.3 we show the result of the convolution of the minimum resolution kernels with the Dirac-δ measure at the origin, i.e., $\delta_{(0,0)}$.

n	r	$k(n,r)$	$k_1,\ldots,k_{k(n,r)}$	$s(n,r)$
2	1	2	2, 1	3
2	2	2	4, 2	6
2	3	2	6, 4	10
2	4	2	9, 6	15
2	5	2	12, 9	21
2	6	2	16, 12	28
2	7	2	20, 16	36
2	8	2	25, 20	45
2	9	2	30, 25	55
2	10	2	36, 30	66
3	1	2	2, 1	4
3	2	2	4, 3	10
3	3	3	7, 6, 1	20
3	4	3	11, 11, 2	35
3	5	3	16, 18, 4	56
3	6	3	23, 27, 7	84
3	7	3	31, 39, 11	120
3	8	3	41, 54, 16	165
3	9	3	53, 72, 23	220
3	10	3	67, 94, 31	286
4	1	2	2, 1	5
4	2	3	4, 3, 1	15
4	3	4	7, 7, 2, 1	35
4	4	4	12, 13, 5, 3	70
4	5	4	18, 23, 9, 7	126
4	6	5	27, 37, 16, 13, 1	210
4	7	5	38, 57, 25, 23, 2	330
4	8	5	53, 83, 39, 37, 4	495
4	9	5	71, 118, 56, 57, 7	715
4	10	5	94, 162, 80, 83, 12	1001
5	1	2	2, 1,	6
5	2	3	4, 3, 1	21
5	3	4	7, 7, 3, 1	56
5	4	5	12, 14, 7, 3, 1	126
5	5	5	19, 25, 14, 8, 3	252
5	6	6	29, 42, 26, 16, 7, 1	462
5	7	6	42, 67, 44, 30, 14, 3	792
5	8	6	60, 102, 71, 51, 26, 7	1287
5	9	6	83, 150, 109, 83, 44, 14	2002
5	10	7	113, 214, 162, 128, 71, 25, 1	3003

Table C.2: Comparison for size of SDP (5.29) and (5.26) for different values of n and r.

α	g_α	α	g_α	α	g_α	α	g_α
$(0,0)$	1.0	$(0,1)$	0.78699	$(1,1)$	0.62968	$(0,2)$	0.40199
$(1,2)$	0.31816	$(2,2)$	0.14946	$(0,3)$	0.11188	$(1,3)$	0.08431
$(2,3)$	0.02954	$(0,4)$	0.00538	$(1,4)$	0.00384	$(0,5)$	0.0

Table C.3: Coefficients of minimum resolution kernel for $r = 5$, $n = 2$.

α	g_α	α	g_α	α	g_α	α	g_α
$(0,0)$	1.0	$(0,1)$	0.91725	$(1,1)$	0.8452	$(0,2)$	0.72312
$(1,2)$	0.66675	$(2,2)$	0.52709	$(0,3)$	0.48749	$(1,3)$	0.44866
$(2,3)$	0.35198	$(3,3)$	0.23167	$(0,4)$	0.27302	$(1,4)$	0.25031
$(2,4)$	0.19314	$(3,4)$	0.12244	$(4,4)$	0.06021	$(0,5)$	0.11905
$(1,5)$	0.10857	$(2,5)$	0.08168	$(3,5)$	0.04826	$(4,5)$	0.02028
$(5,5)$	0.0047	$(0,6)$	0.03512	$(1,6)$	0.03208	$(2,6)$	0.02324
$(3,6)$	0.01204	$(4,6)$	0.00337	$(0,7)$	0.00532	$(1,7)$	0.00496
$(2,7)$	0.00352	$(3,7)$	0.0013	$(0,8)$	0.00024	$(1,8)$	0.00026
$(2,8)$	0.00013	$(0,9)$	7.0e-5	$(1,9)$	5.0e-5	$(0,10)$	0.0

Table C.4: Coefficients of minimum resolution kernel for $r = 10$, $n = 2$.

α	g_α	α	g_α	α	g_α	α	g_α
$(0,0)$	1.0	$(0,1)$	0.95647	$(1,1)$	0.91655	$(0,2)$	0.84498
$(1,2)$	0.81035	$(2,2)$	0.71818	$(0,3)$	0.6908	$(1,3)$	0.66249
$(2,3)$	0.58704	$(3,3)$	0.47975	$(0,4)$	0.52059	$(1,4)$	0.49892
$(2,4)$	0.44107	$(3,4)$	0.3589	$(4,4)$	0.26675	$(0,5)$	0.35808
$(1,5)$	0.34275	$(2,5)$	0.3017	$(3,5)$	0.24347	$(4,5)$	0.17847
$(5,5)$	0.11689	$(0,6)$	0.22081	$(1,6)$	0.21103	$(2,6)$	0.18465
$(3,6)$	0.14712	$(4,6)$	0.10543	$(5,6)$	0.06653	$(6,6)$	0.03569
$(0,7)$	0.1185	$(1,7)$	0.11309	$(2,7)$	0.09822	$(3,7)$	0.07689
$(4,7)$	0.05322	$(5,7)$	0.03158	$(6,7)$	0.01525	$(7,7)$	0.0054
$(0,8)$	0.05275	$(1,8)$	0.0503	$(2,8)$	0.04337	$(3,8)$	0.03317
$(4,8)$	0.02178	$(5,8)$	0.01164	$(6,8)$	0.00461	$(7,8)$	0.00103
$(0,9)$	0.01804	$(1,9)$	0.01726	$(2,9)$	0.01486	$(3,9)$	0.01105
$(4,9)$	0.00666	$(5,9)$	0.00296	$(6,9)$	0.00073	$(0,10)$	0.00416
$(1,10)$	0.00404	$(2,10)$	0.00353	$(3,10)$	0.00253	$(4,10)$	0.00132
$(5,10)$	0.00037	$(0,11)$	0.00052	$(1,11)$	0.00052	$(2,11)$	0.00047
$(3,11)$	0.00032	$(4,11)$	0.00011	$(0,12)$	5.0e-5	$(1,12)$	5.0e-5
$(2,12)$	4.0e-5	$(3,12)$	1.0e-5	$(0,13)$	3.0e-5	$(1,13)$	2.0e-5
$(2,13)$	1.0e-5	$(0,14)$	1.0e-5	$(1,14)$	0.0	$(0,15)$	0.0

Table C.5: Coefficients of minimum resolution kernel for $r = 15$, $n = 2$

α	g_α	α	g_α	α	g_α	α	g_α
$(0,0)$	1.0	$(0,1)$	0.97325	$(1,1)$	0.94811	$(0,2)$	0.90176
$(1,2)$	0.87892	$(2,2)$	0.81609	$(0,3)$	0.79699	$(1,3)$	0.77697
$(2,3)$	0.72182	$(3,3)$	0.63911	$(0,4)$	0.67169	$(1,4)$	0.65475
$(2,4)$	0.60811	$(3,4)$	0.53814	$(4,4)$	0.45283	$(0,5)$	0.53849
$(1,5)$	0.52475	$(2,5)$	0.48685	$(3,5)$	0.43002	$(4,5)$	0.36082
$(5,5)$	0.28643	$(0,6)$	0.40878	$(1,6)$	0.39813	$(2,6)$	0.36875
$(3,6)$	0.32468	$(4,6)$	0.27109	$(5,6)$	0.21366	$(6,6)$	0.15784
$(0,7)$	0.29168	$(1,7)$	0.2839	$(2,7)$	0.26235	$(3,7)$	0.22999
$(4,7)$	0.19066	$(5,7)$	0.14866	$(6,7)$	0.10812	$(7,7)$	0.07248
$(0,8)$	0.1935	$(1,8)$	0.18821	$(2,8)$	0.17347	$(3,8)$	0.15124
$(4,8)$	0.12418	$(5,8)$	0.09536	$(6,8)$	0.06779	$(7,8)$	0.04396
$(8,8)$	0.02543	$(0,9)$	0.11748	$(1,9)$	0.11421	$(2,9)$	0.10497
$(3,9)$	0.09093	$(4,9)$	0.07375	$(5,9)$	0.05547	$(6,9)$	0.03816
$(7,9)$	0.02356	$(8,9)$	0.01265	$(9,9)$	0.00562	$(0,10)$	0.06382
$(1,10)$	0.06203	$(2,10)$	0.05689	$(3,10)$	0.04894	$(4,10)$	0.03908
$(5,10)$	0.02857	$(6,10)$	0.01876	$(7,10)$	0.01074	$(8,10)$	0.00512
$(9,10)$	0.00184	$(10,10)$	0.00039	$(0,11)$	0.03006	$(1,11)$	0.02924
$(2,11)$	0.02681	$(3,11)$	0.02292	$(4,11)$	0.01796	$(5,11)$	0.01263
$(6,11)$	0.00772	$(7,11)$	0.00392	$(8,11)$	0.00151	$(9,11)$	0.00034
$(0,12)$	0.01176	$(1,12)$	0.01148	$(2,12)$	0.01059	$(3,12)$	0.00902
$(4,12)$	0.0069	$(5,12)$	0.00458	$(6,12)$	0.00251	$(7,12)$	0.00103
$(8,12)$	0.00025	$(0,13)$	0.00361	$(1,13)$	0.00355	$(2,13)$	0.00332
$(3,13)$	0.00284	$(4,13)$	0.0021	$(5,13)$	0.00128	$(6,13)$	0.00058
$(7,13)$	0.00015	$(0,14)$	0.00081	$(1,14)$	0.00081	$(2,14)$	0.00078
$(3,14)$	0.00066	$(4,14)$	0.00047	$(5,14)$	0.00025	$(6,14)$	7.0e-5
$(0,15)$	0.00015	$(1,15)$	0.00015	$(2,15)$	0.00014	$(3,15)$	0.00011
$(4,15)$	7.0e-5	$(5,15)$	2.0e-5	$(0,16)$	5.0e-5	$(1,16)$	5.0e-5
$(2,16)$	4.0e-5	$(3,16)$	2.0e-5	$(4,16)$	1.0e-5	$(0,17)$	3.0e-5
$(1,17)$	3.0e-5	$(2,17)$	2.0e-5	$(3,17)$	0.0	$(0,18)$	1.0e-5
$(1,18)$	1.0e-5	$(2,18)$	0.0	$(0,19)$	0.0	$(1,19)$	0.0
$(0,20)$	0.0	-	-	-	-	-	-

Table C.6: Coefficients of minimum resolution kernel for $r = 20$, $n = 2$

α	g_α	α	g_α	α	g_α	α	g_α
(0,0)	1.0	(0,1)	0.98191	(1,1)	0.96467	(0,2)	0.93234
(1,2)	0.91628	(2,2)	0.87123	(0,3)	0.85734	(1,3)	0.84273
(2,3)	0.80174	(3,3)	0.7385	(0,4)	0.76377	(1,4)	0.7508
(2,4)	0.71437	(3,4)	0.65817	(4,4)	0.5868	(0,5)	0.65878
(1,5)	0.64755	(2,5)	0.61599	(3,5)	0.5673	(4,5)	0.50548
(5,5)	0.43512	(0,6)	0.54933	(1,6)	0.53987	(2,6)	0.51327
(3,6)	0.47223	(4,6)	0.42015	(5,6)	0.36095	(6,6)	0.29868
(0,7)	0.44172	(1,7)	0.43399	(2,7)	0.41225	(3,7)	0.37871
(4,7)	0.33617	(5,7)	0.28787	(6,7)	0.23718	(7,7)	0.18732
(0,8)	0.34125	(1,8)	0.33516	(2,8)	0.31802	(3,8)	0.29155
(4,8)	0.25798	(5,8)	0.21992	(6,8)	0.18008	(7,8)	0.14105
(8,8)	0.10507	(0,9)	0.25198	(1,9)	0.24739	(2,9)	0.23443
(3,9)	0.21438	(4,9)	0.18895	(5,9)	0.16012	(6,9)	0.13001
(7,9)	0.10067	(8,9)	0.07385	(9,9)	0.05086	(0,10)	0.17659
(1,10)	0.17331	(2,10)	0.164	(3,10)	0.14956	(4,10)	0.13117
(5,10)	0.11032	(6,10)	0.0886	(7,10)	0.06755	(8,10)	0.04851
(9,10)	0.03246	(10,10)	0.01994	(0,11)	0.11635	(1,11)	0.11416
(2,11)	0.10789	(3,11)	0.09808	(4,11)	0.08555	(5,11)	0.07128
(6,11)	0.05644	(7,11)	0.04215	(8,11)	0.0294	(9,11)	0.0189
(10,11)	0.01097	(11,11)	0.00557	(0,12)	0.07119	(1,12)	0.06984
(2,12)	0.06593	(3,12)	0.05976	(4,12)	0.0518	(5,12)	0.04268
(6,12)	0.03318	(7,12)	0.0241	(8,12)	0.01615	(9,12)	0.0098
(10,12)	0.00523	(11,12)	0.00233	(12,12)	0.00079	(0,13)	0.03979
(1,13)	0.03905	(2,13)	0.03686	(3,13)	0.03334	(4,13)	0.0287
(5,13)	0.02333	(6,13)	0.01771	(7,13)	0.0124	(8,13)	0.00785
(9,13)	0.00437	(10,13)	0.00204	(11,13)	0.00073	(12,13)	0.00015
(0,14)	0.0199	(1,14)	0.01956	(2,14)	0.01849	(3,14)	0.01672
(4,14)	0.0143	(5,14)	0.01144	(6,14)	0.00842	(7,14)	0.0056
(8,14)	0.00327	(9,14)	0.00159	(10,14)	0.00059	(11,14)	0.00013
(0,15)	0.00867	(1,15)	0.00854	(2,15)	0.00812	(3,15)	0.00736
(4,15)	0.00627	(5,15)	0.00491	(6,15)	0.00347	(7,15)	0.00214
(8,15)	0.0011	(9,15)	0.00043	(10,15)	9.0e-5	(0,16)	0.00317
(1,16)	0.00314	(2,16)	0.00302	(3,16)	0.00276	(4,16)	0.00234
(5,16)	0.00179	(6,16)	0.00119	(7,16)	0.00065	(8,16)	0.00027
(9,16)	6.0e-5	(0,17)	0.00094	(1,17)	0.00094	(2,17)	0.00092
(3,17)	0.00085	(4,17)	0.00071	(5,17)	0.00053	(6,17)	0.00032
(7,17)	0.00015	(8,17)	4.0e-5	(0,18)	0.00023	(1,18)	0.00023
(2,18)	0.00023	(3,18)	0.00021	(4,18)	0.00017	(5,18)	0.00011
(6,18)	6.0e-5	(7,18)	2.0e-5	(0,19)	7.0e-5	(1,19)	7.0e-5
(2,19)	6.0e-5	(3,19)	5.0e-5	(4,19)	3.0e-5	(5,19)	2.0e-5
(6,19)	1.0e-5	(0,20)	4.0e-5	(1,20)	4.0e-5	(2,20)	3.0e-5
(3,20)	2.0e-5	(4,20)	1.0e-5	(5,20)	0.0	(0,21)	2.0e-5
(1,21)	2.0e-5	(2,21)	1.0e-5	(3,21)	1.0e-5	(4,21)	0.0
(0,22)	1.0e-5	(1,22)	1.0e-5	(2,22)	0.0	(3,22)	0.0
(0,23)	0.0	(1,23)	0.0	(2,23)	0.0	(0,24)	0.0
(1,24)	0.0	(0,25)	0.0	-	-	-	-

Table C.7: Coefficients of minimum resolution kernel for $r = 25$, $n = 2$

α	g_α	α	g_α	α	g_α	α	g_α
$(0,0,0)$	1.0	$(0,0,1)$	0.68496	$(0,1,1)$	0.48031	$(1,1,1)$	0.33143
$(0,0,2)$	0.23185	$(0,1,2)$	0.15668	$(1,1,2)$	0.10399	$(0,2,2)$	0.04188
$(1,2,2)$	0.02408	$(0,0,3)$	0.02117	$(0,1,3)$	0.01319	$(1,1,3)$	0.00757
$(0,2,3)$	0.00092	$(0,0,4)$	7.0e-5	$(0,1,4)$	−5.0e-5	$(0,0,5)$	−1.0e-5

Table C.8: Coefficients of minimum resolution kernel for $r = 5$, $n = 3$.

α	g_α	α	g_α	α	g_α	α	g_α
$(0,0,0)$	1.0	$(0,0,1)$	0.86461	$(0,1,1)$	0.75281	$(1,1,1)$	0.65401
$(0,0,2)$	0.57852	$(0,1,2)$	0.5036	$(1,1,2)$	0.43793	$(0,2,2)$	0.33561
$(1,2,2)$	0.28996	$(2,2,2)$	0.18874	$(0,0,3)$	0.29062	$(0,1,3)$	0.25169
$(1,1,3)$	0.21767	$(0,2,3)$	0.16346	$(1,2,3)$	0.14054	$(2,2,3)$	0.08819
$(0,3,3)$	0.07549	$(1,3,3)$	0.06385	$(2,3,3)$	0.03779	$(3,3,3)$	0.01388
$(0,0,4)$	0.09872	$(0,1,4)$	0.08484	$(1,1,4)$	0.07271	$(0,2,4)$	0.05252
$(1,2,4)$	0.04453	$(2,2,4)$	0.02619	$(0,3,4)$	0.02154	$(1,3,4)$	0.01782
$(2,3,4)$	0.00935	$(3,3,4)$	0.00238	$(0,4,4)$	0.00478	$(1,4,4)$	0.00372
$(2,4,4)$	0.00146	$(0,0,5)$	0.01757	$(0,1,5)$	0.0151	$(1,1,5)$	0.01284
$(0,2,5)$	0.00872	$(1,2,5)$	0.00718	$(2,2,5)$	0.00361	$(0,3,5)$	0.00277
$(1,3,5)$	0.00212	$(2,3,5)$	0.00081	$(0,4,5)$	0.00028	$(1,4,5)$	0.00016
$(0,5,5)$	2.0e-5	$(0,0,6)$	0.00108	$(0,1,6)$	0.00095	$(1,1,6)$	0.00079
$(0,2,6)$	0.00051	$(1,2,6)$	0.00039	$(2,2,6)$	0.00014	$(0,3,6)$	0.00011
$(1,3,6)$	6.0e-5	$(0,4,6)$	1.0e-5	$(0,0,7)$	8.0e-5	$(0,1,7)$	5.0e-5
$(1,1,7)$	2.0e-5	$(0,2,7)$	1.0e-5	$(1,2,7)$	0.0	$(0,3,7)$	0.0
$(0,0,8)$	1.0e-5	$(0,1,8)$	0.0	$(1,1,8)$	0.0	$(0,2,8)$	0.0
$(0,0,9)$	0.0	$(0,1,9)$	0.0	$(0,0,10)$	0.0	-	-

Table C.9: Coefficients of minimum resolution kernel for $r = 10$, $n = 3$.

α	g_α	α	g_α	α	g_α	α	g_α
(0,0,0)	1.0	(0,0,1)	0.92525	(0,1,1)	0.85868	(1,1,1)	0.79631
(0,0,2)	0.74623	(0,1,2)	0.6932	(1,1,2)	0.64369	(0,2,2)	0.56094
(1,2,2)	0.5199	(2,2,2)	0.41863	(0,0,3)	0.52376	(0,1,3)	0.48624
(1,1,3)	0.45127	(0,2,3)	0.39215	(1,2,3)	0.36346	(2,2,3)	0.29109
(0,3,3)	0.27243	(1,3,3)	0.25169	(2,3,3)	0.19998	(3,3,3)	0.13526
(0,0,4)	0.31461	(0,1,4)	0.29152	(1,1,4)	0.27006	(0,2,4)	0.23319
(1,2,4)	0.21574	(2,2,4)	0.17145	(0,3,4)	0.15935	(1,3,4)	0.14698
(2,3,4)	0.11543	(3,3,4)	0.07607	(0,4,4)	0.09079	(1,4,4)	0.08327
(2,4,4)	0.06428	(3,4,4)	0.04086	(4,4,4)	0.02055	(0,0,5)	0.15601
(0,1,5)	0.14422	(1,1,5)	0.13325	(0,2,5)	0.11392	(1,2,5)	0.10508
(2,2,5)	0.08235	(0,3,5)	0.07573	(1,3,5)	0.06961	(2,3,5)	0.0538
(3,3,5)	0.03407	(0,4,5)	0.04116	(1,4,5)	0.03757	(2,4,5)	0.0283
(3,4,5)	0.01699	(4,4,5)	0.00763	(0,5,5)	0.01731	(1,5,5)	0.01561
(2,5,5)	0.01131	(3,5,5)	0.0062	(4,5,5)	0.00229	(5,5,5)	0.00043
(0,0,6)	0.05986	(0,1,6)	0.05526	(1,1,6)	0.05095	(0,2,6)	0.04299
(1,2,6)	0.03951	(2,2,6)	0.03032	(0,3,6)	0.02746	(1,3,6)	0.02509
(2,3,6)	0.01887	(3,3,6)	0.01123	(0,4,6)	0.01382	(1,4,6)	0.01249
(2,4,6)	0.00905	(3,4,6)	0.00494	(4,4,6)	0.00181	(0,5,6)	0.00508
(1,5,6)	0.00451	(2,5,6)	0.00306	(3,5,6)	0.00142	(4,5,6)	0.00034
(0,6,6)	0.00119	(1,6,6)	0.00101	(2,6,6)	0.0006	(3,6,6)	0.00019
(0,0,7)	0.01607	(0,1,7)	0.01487	(1,1,7)	0.01372	(0,2,7)	0.01142
(1,2,7)	0.01046	(2,2,7)	0.0078	(0,3,7)	0.00693	(1,3,7)	0.00626
(2,3,7)	0.00449	(3,3,7)	0.0024	(0,4,7)	0.0031	(1,4,7)	0.00274
(2,4,7)	0.00183	(3,4,7)	0.00083	(4,4,7)	0.00019	(0,5,7)	0.00089
(1,5,7)	0.00076	(2,5,7)	0.00044	(3,5,7)	0.00013	(0,6,7)	0.00014
(1,6,7)	0.0001	(2,6,7)	4.0e-5	(0,7,7)	1.0e-5	(1,7,7)	0.0
(0,0,8)	0.00273	(0,1,8)	0.00255	(1,1,8)	0.00237	(0,2,8)	0.00195
(1,2,8)	0.00178	(2,2,8)	0.00128	(0,3,8)	0.00113	(1,3,8)	0.001
(2,3,8)	0.00067	(3,3,8)	0.0003	(0,4,8)	0.00044	(1,4,8)	0.00037
(2,4,8)	0.00021	(3,4,8)	6.0e-5	(0,5,8)	0.0001	(1,5,8)	7.0e-5
(2,5,8)	3.0e-5	(0,6,8)	1.0e-5	(1,6,8)	0.0	(0,7,8)	0.0
(0,0,9)	0.00026	(0,1,9)	0.00024	(1,1,9)	0.00022	(0,2,9)	0.00018
(1,2,9)	0.00016	(2,2,9)	0.0001	(0,3,9)	0.0001	(1,3,9)	8.0e-5
(2,3,9)	5.0e-5	(3,3,9)	1.0e-5	(0,4,9)	4.0e-5	(1,4,9)	3.0e-5
(2,4,9)	1.0e-5	(0,5,9)	1.0e-5	(1,5,9)	0.0	(0,6,9)	0.0
(0,0,10)	2.0e-5	(0,1,10)	2.0e-5	(1,1,10)	1.0e-5	(0,2,10)	1.0e-5
(1,2,10)	0.0	(2,2,10)	0.0	(0,3,10)	0.0	(1,3,10)	0.0
(2,3,10)	0.0	(0,4,10)	0.0	(1,4,10)	0.0	(0,5,10)	0.0
(0,0,11)	0.0	(0,1,11)	0.0	(1,1,11)	0.0	(0,2,11)	0.0
(1,2,11)	0.0	(2,2,11)	0.0	(0,3,11)	0.0	(1,3,11)	0.0
(0,4,11)	0.0	(0,0,12)	0.0	(0,1,12)	0.0	(1,1,12)	0.0
(0,2,12)	0.0	(1,2,12)	0.0	(0,3,12)	0.0	(0,0,13)	0.0
(0,1,13)	0.0	(1,1,13)	0.0	(0,2,13)	0.0	(0,0,14)	0.0
(0,1,14)	0.0	(0,0,15)	0.0	-	-	-	-

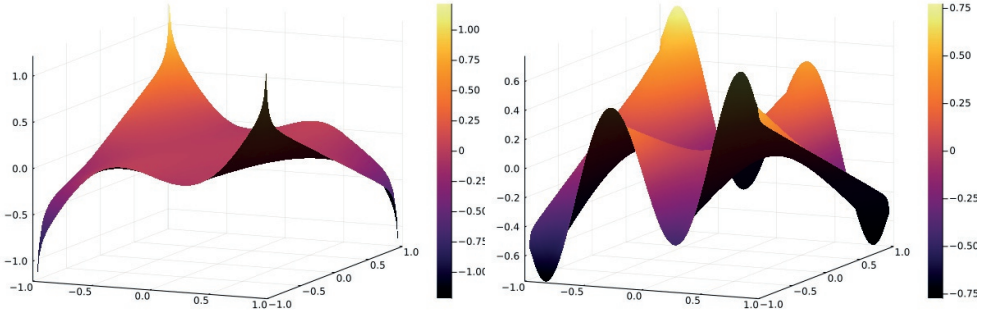
Table C.10: Coefficients of minimum resolution kernel for $r = 15$, $n = 3$.

α	g_α	α	g_α	α	g_α	α	g_α
$(0,0,0,0)$	1.0	$(0,0,0,1)$	0.59115	$(0,0,1,1)$	0.35488	$(0,1,1,1)$	0.20728
$(1,1,1,1)$	0.11764	$(0,0,0,2)$	0.12466	$(0,0,1,2)$	0.06911	$(0,1,1,2)$	0.03576
$(1,1,1,2)$	0.01705	$(0,0,2,2)$	0.00736	$(0,1,2,2)$	0.00225	$(0,0,0,3)$	0.00251
$(0,0,1,3)$	0.00147	$(0,1,1,3)$	0.00051	$(0,0,2,3)$	9.0e-5	$(0,0,0,4)$	−3.0e-5
$(0,0,1,4)$	0.0	$(0,0,0,5)$	0.0	-	-	-	-

Table C.11: Coefficients of minimum resolution kernel for $r = 5$, $n = 4$.

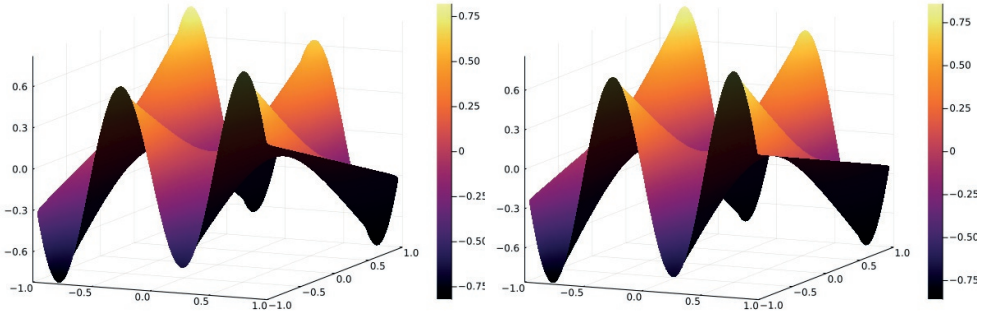
α	g_α	α	g_α	α	g_α	α	g_α
$(0,0,0,0)$	1.0	$(0,0,0,1)$	0.81189	$(0,0,1,1)$	0.66535	$(0,1,1,1)$	0.54349
$(1,1,1,1)$	0.44544	$(0,0,0,2)$	0.45316	$(0,0,1,2)$	0.37047	$(0,1,1,2)$	0.3025
$(1,1,1,2)$	0.24635	$(0,0,2,2)$	0.20357	$(0,1,2,2)$	0.16459	$(1,1,2,2)$	0.13334
$(0,2,2,2)$	0.08681	$(1,2,2,2)$	0.06968	$(2,2,2,2)$	0.03479	$(0,0,0,3)$	0.15888
$(0,0,1,3)$	0.12858	$(0,1,1,3)$	0.10384	$(1,1,1,3)$	0.08364	$(0,0,2,3)$	0.06743
$(0,1,2,3)$	0.05385	$(1,1,2,3)$	0.04282	$(0,2,2,3)$	0.02653	$(1,2,2,3)$	0.02072
$(2,2,2,3)$	0.00914	$(0,0,3,3)$	0.02034	$(0,1,3,3)$	0.01576	$(1,1,3,3)$	0.01217
$(0,2,3,3)$	0.00692	$(1,2,3,3)$	0.0051	$(2,2,3,3)$	0.00166	$(0,3,3,3)$	0.00136
$(1,3,3,3)$	0.0008	$(0,0,0,4)$	0.02515	$(0,0,1,4)$	0.02016	$(0,1,1,4)$	0.01603
$(1,1,1,4)$	0.01265	$(0,0,2,4)$	0.00981	$(0,1,2,4)$	0.00759	$(1,1,2,4)$	0.00582
$(0,2,2,4)$	0.00324	$(1,2,2,4)$	0.00237	$(2,2,2,4)$	0.00076	$(0,0,3,4)$	0.00247
$(0,1,3,4)$	0.00178	$(1,1,3,4)$	0.00125	$(0,2,3,4)$	0.00059	$(1,2,3,4)$	0.00034
$(0,3,3,4)$	2.0e-5	$(0,0,4,4)$	0.00027	$(0,1,4,4)$	0.00014	$(1,1,4,4)$	7.0e-5
$(0,2,4,4)$	0.0	$(0,0,0,5)$	4.0e-5	$(0,0,1,5)$	0.00016	$(0,1,1,5)$	0.00017
$(1,1,1,5)$	0.00017	$(0,0,2,5)$	0.00015	$(0,1,2,5)$	0.00012	$(1,1,2,5)$	9.0e-5
$(0,2,2,5)$	5.0e-5	$(1,2,2,5)$	3.0e-5	$(0,0,3,5)$	7.0e-5	$(0,1,3,5)$	3.0e-5
$(1,1,3,5)$	1.0e-5	$(0,2,3,5)$	0.0	$(0,0,4,5)$	1.0e-5	$(0,1,4,5)$	0.0
$(0,0,5,5)$	0.0	$(0,0,0,6)$	0.00024	$(0,0,1,6)$	0.00017	$(0,1,1,6)$	0.0001
$(1,1,1,6)$	−7.0e-5	$(0,0,2,6)$	−3.0e-5	$(0,1,2,6)$	−2.0e-5	$(1,1,2,6)$	−1.0e-5
$(0,2,2,6)$	0.0	$(0,0,3,6)$	1.0e-5	$(0,1,3,6)$	0.0	$(0,0,4,6)$	0.0
$(0,0,0,7)$	−4.0e-5	$(0,0,1,7)$	−1.0e-5	$(0,1,1,7)$	−1.0e-5	$(1,1,1,7)$	0.0
$(0,0,2,7)$	0.0	$(0,1,2,7)$	0.0	$(0,0,3,7)$	0.0	$(0,0,0,8)$	0.0
$(0,0,1,8)$	0.0	$(0,1,1,8)$	0.0	$(0,0,2,8)$	0.0	$(0,0,0,9)$	0.0
$(0,0,1,9)$	0.0	$(0,0,0,10)$	0.0	-	-	-	-

Table C.12: Coefficients of minimum resolution kernel for $r = 10$, $n = 4$.



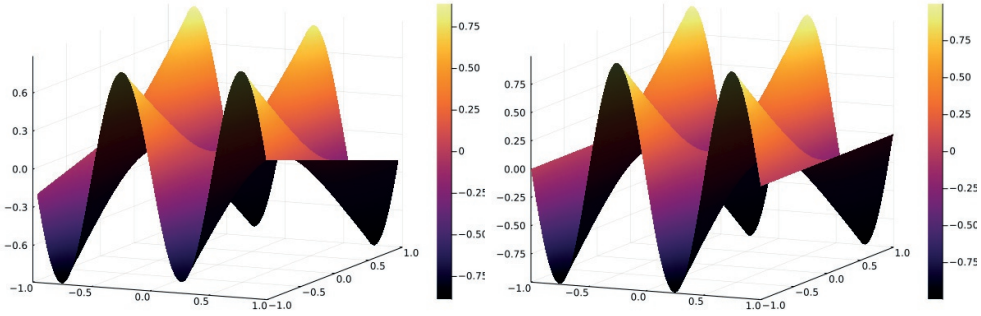
(a) Approximation level $r = 10$

(b) Approximation level $r = 20$



(c) Approximation level $r = 30$

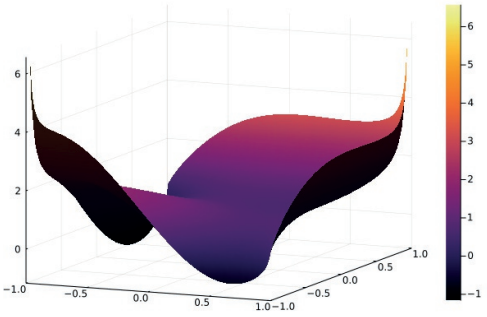
(d) Approximation level $r = 40$



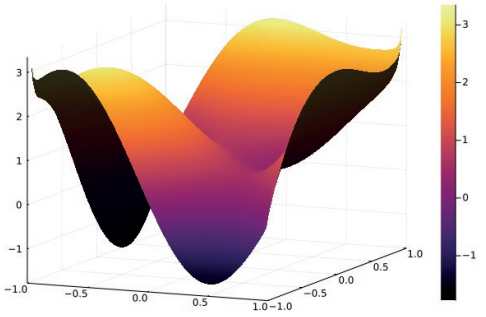
(e) Approximation level $r = 50$

(f) $\sin(2\pi x)y$

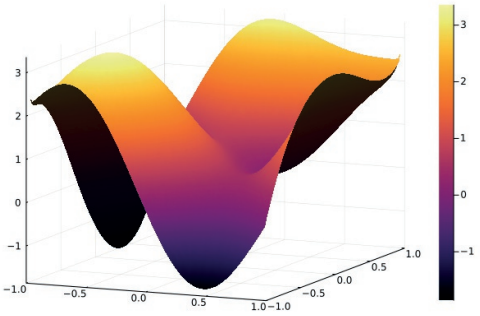
Figure C.1: Approximation of the function $(x, y) \mapsto \sin(2\pi x)y$ via the kernel with respect to σ_r for $r = 10, 20, \ldots, 50$ as well as a plot of the function itself.



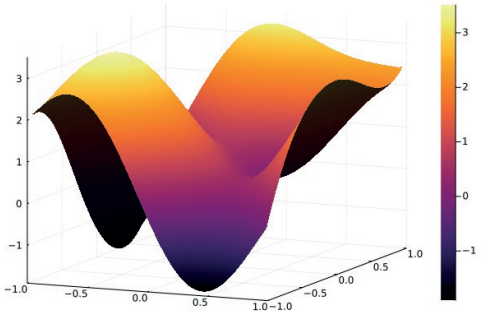
(a) Approximation level $r = 10$



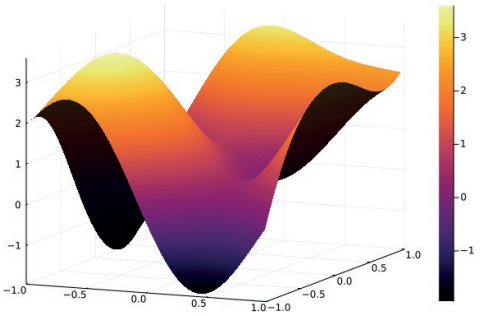
(b) Approximation level $r = 20$



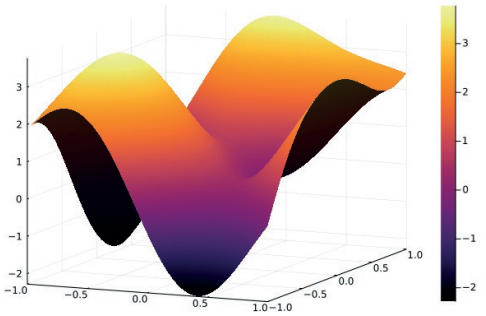
(c) Approximation level $r = 30$



(d) Approximation level $r = 40$

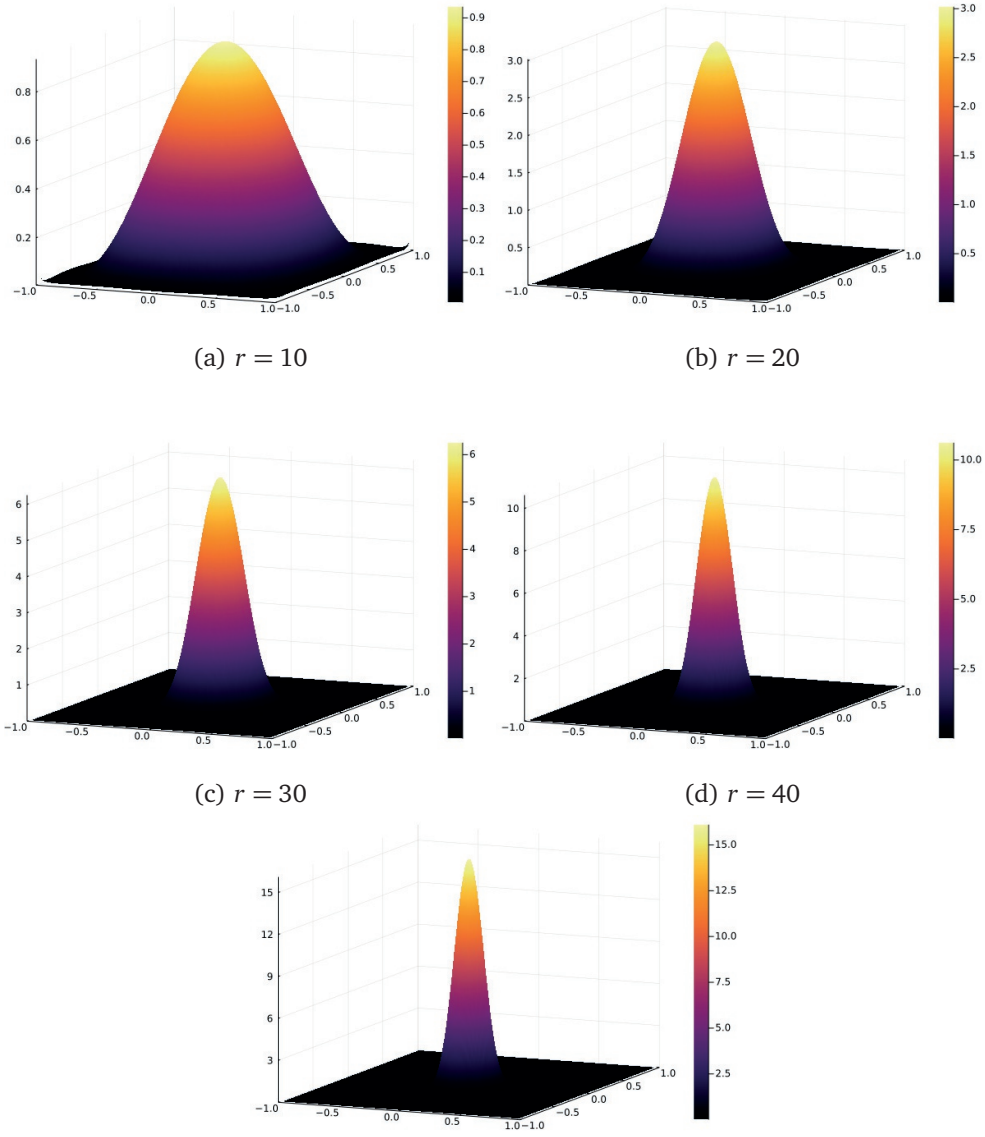


(e) Approximation level $r = 50$



(f) Peaks function $p(x, y)$

Figure C.2: Approximation of the *peaks* function $p(x,y) := 3(1-x)^2\exp(-x^2-(y+1)^2) - 10(x/5-x^3-y^5)\exp(-x^2-y^2) - (1/3)\exp(-(x+1)^2-y^2)$ via the kernel with respect to σ_r for $r = 10, 20, \ldots, 50$ as well as a plot of the function itself in graph C.2f.



(a) $r = 10$ (b) $r = 20$

(c) $r = 30$ (d) $r = 40$

(e) $r = 50$

Figure C.3: Convolution of Dirac-δ at $(0, 0)$ with optimal kernel with respect to σ_r for different values of r

Bibliography

[AC11] Francesco Altomare and Michele Campiti. *Korovkin-type Approximation Theory and its Applications*, volume 17 of *De Gruyter Studies in Mathematics*. De Gruyter, 2011.

[ADH17] Amir A. Ahmadi, Sanjeeb Dash, and Georgina Hall. Optimization over structured subsets of positive semidefinite matrices via column generation. *Discrete Optimization*, 24:129–151, 2017.

[AH17] Amir A. Ahmadi and Georgina Hall. *Sum of squares basis pursuit with linear and second order cone programming*, volume 685 of *Contemporary Mathematics*, pages 27–53. American Mathematical Society, United States, 2017.

[AHO98] Farid Alizadeh, Jean-Pierre A. Haeberly, and Michael L. Overton. Primal-dual interior-point methods for semidefinite programming: Convergence rates, stability and numerical results. *SIAM Journal on Optimization*, 8(3):746–768, 1998.

[AM14] Amir A. Ahmadi and Anirudha Majumdar. DSOS and SDSOS optimization: LP and SOCP-based alternatives to sum of squares optimization. In *Proceedings of the 48th Annual Conference on Information Sciences and Systems (CISS)*, pages 1–5, 2014.

[AM19] Amir A. Ahmadi and Anirudha Majumdar. DSOS and SDSOS optimization: More tractable alternatives to sum of squares and semidefinite optimization. *SIAM Journal on Applied Algebra and Geometry*, 3(2):193–230, 2019.

[AOPT13] Amir A. Ahmadi, Alex Olshevsky, Pablo A. Parrilo, and John N. Tsitsiklis. NP-hardness of deciding convexity of quartic polynomials

and related problems. *Mathematical Programming*, 137(1):453–476, 2013.

[AS10] Warren Adams and Hanif Sherali. *A Reformulation-Linearization Technique for Solving Discrete and Continuous Nonconvex Problems*, volume 31 of *Nonconvex Optimization and Its Applications*. Springer, 2010.

[Bar75] Zsolt Baranyai. On the factorization of the complete uniform hypergraph. *Infinite and Finite Sets*, 1:91–108, 1975. Proceedings of a Colloquium held at Keszthely, June 25-July 1, 1973. Dedicated to Paul Erdős on his 60th birthday.

[Bar02] Alexander Barvinok. *A course in convexity*. Providence, RI : American Mathematical Society, 2002.

[BCPT05] Erik G. Boman, Doron Chen, Ojas Parekh, and Sivan Toledo. On factor width and symmetric H-matrices. *Linear Algebra and its Applications*, 405:239–248, 2005.

[BdK01] Immanuel Bomze and Etienne de Klerk. Solving standard quadratic optimization problems via linear, semidefinite and copositive programming. *Journal of Global Optimization*, 24:163–185, 2001.

[BdK22] Daniel Brosch and Etienne de Klerk. Jordan symmetry reduction for conic optimization over the doubly nonnegative cone: theory and software. *Optimization Methods and Software*, 37(6):2001–2020, 2022.

[BDMS22] Grigoriy Blekherman, Santanu S. Dey, Marco Molinaro, and Shengding Sun. Sparse psd approximation of the psd cone. *Mathematical Programming*, 191(2):981–1004, 2022.

[Ber12] S. Bernstein. *Sur l'ordre de la meilleure approximation des fonctions continues par des polynomes de degré donné*. (Acad. Royale de Belgique. Classe d. sc. Mémoires. Coll. in-4"). Hayez, imprimeur des académies royales, 1912.

[Bil86] Patrick Billingsley. *Probability and Measure*. John Wiley and Sons, third edition, 1986.

[BL97] Phelim P. Boyle and X. Sheldon Lin. Bounds on contingent claims based on several assets. *Journal of Financial Economics*, 46(3):383–400, 1997.

[BM22] Lorenzo Baldi and Bernard Mourrain. On the effective Putinar's positivstellensatz and moment approximation. *Mathematical Programming*, 2022. Advance online publication.

[BN71] Paul L. Butzer and Rolf J. Nessel. *Fourier analysis and approximation*. Academic Press New York, 1971.

[Bor99] Brian Borchers. CSDP, A C library for semidefinite programming. *Optimization Methods and Software*, 11(1-4):613–623, 1999.

[BP02] Dimitris Bertsimas and Ioana Popescu. On the relation between option and stock prices: A convex optimization approach. *Operations Research*, 50(2):358–374, 2002.

[Bro22] Daniel Brosch. *Symmetry reduction in convex optimization with applications in combinatorics*. PhD thesis, Tilburg University, 2022.

[BV96] Stephen Boyd and Lieven Vandenberghe. Semidefinite programming. *SIAM Review*, 38(1):49–95, 1996.

[BV04] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[Car25] H. S. Carslaw. A historical note on Gibbs' phenomenon in Fourier's series and integrals. *Bulletin of the American Mathematical Society*, 31(8):420 – 424, 1925.

[Chr58] Elwin B. Christoffel. Über die Gaußische Quadratur und eine Verallgemeinerung derselben. *Journal für die reine und angewandte Mathematik*, 1858(55):61–82, 1858.

[CN21] Jordi Castro and Stefano Nasini. A specialized interior-point algorithm for huge minimum convex cost flows in bipartite networks. *European Journal of Operational Research*, 290(3):857–869, 2021.

[CR76] John C. Cox and Stephen A. Ross. The valuation of options for alternative stochastic processes. *Journal of Financial Economics*, 3(1):145–166, 1976.

[d'A04] Alexandre d'Aspremont. *Shape constrained optimization with application in finance and engineering*. PhD thesis, Stanford University, 2004.

[Dan51] George B. Dantzig. Maximization of a linear function of variables subject to linear inequalities. In *T. C. Koopmans, Ed., Activity Analysis of Production and Allocation, Cowles Commission Monograph, No. 13*, pages 339–347, 1951.

[Dar78] Gaston Darboux. Mémoire sur l'approximation des fonctions de très-grands nombres, et sur une classe étendue de développements en série. *Journal de Mathématiques Pures et Appliquées*, pages 5–56, 1878.

[DH07] Mark H. A. Davis and David G. Hobson. The range of traded option prices. *Mathematical Finance*, 17(1):1–14, 2007.

[Dir29] Lejeune G. Dirichlet. Sur la convergence des séries trigonométriques qui servent à représenter une fonction arbitraire entre des limites données. *Journal für die reine und angewandte Mathematik*, 1829(4):157–169, 1829.

[dK10] Etienne de Klerk. Exploiting special structure in semidefinite programming: A survey of theory and applications. *European Journal of Operational Research*, 201(1):1–10, 2010.

[dKHL17] Etienne de Klerk, Roxana Hess, and Monique 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.

[dKKP20] Etienne de Klerk, Daniel Kuhn, and Krzysztof Postek. Distributionally robust optimization with polynomial densities: Theory, models and algorithms. *Mathematical Programming*, 181:265–296, 2020.

[dKLP06] Etienne de Klerk, Monique Laurent, and Pablo A. Parrilo. A PTAS for the minimization of polynomials of fixed degree over the simplex. *Theoretical Computer Science*, 361(2-3):210–225, 2006.

[dKP02] Etienne de Klerk and Dmitrii V. Pasechnik. Approximation of the stability number of a graph via copositive programming. *SIAM Journal on Optimization*, 12(4):875–892, 2002.

[dKRT97] Etienne de Klerk, Cornelis Roos, and Támas Terlaky. Initialization in semidefinite programming via a self-dual skew-symmetric embedding. *Operations Research Letters*, 20(5):213–221, 1997.

[dKV16] Etienne de Klerk and Frank Vallentin. On the Turing model complexity of interior point methods for semidefinite programming. *SIAM Journal on Optimization*, 26(3):1944–1961, 2016.

[dS18] Philipp J. di Dio and Konrad Schmüdgen. The multidimensional truncated moment problem: Atoms, determinacy, and core variety. *Journal of Functional Analysis*, 274(11):3124–3148, 2018.

[DSdSG+22] Valentina De Simone, Daniela di Serafino, Jacek Gondzio, Spyridon Pougkakiotis, and Marco Viola. Sparse approximations with interior point methods. *SIAM Review*, 64(4):954–988, 2022.

[DT14] Robin Deits and Russ Tedrake. Computing large convex regions of obstacle-free space through semidefinite programming. In H. Levent Akin, Nancy M. Amato, Volkan Isler, and A. Frank van der Stappen, editors, *Algorithmic Foundations of Robotics XI - Selected Contributions of the Eleventh International Workshop on the Algorithmic Foundations of Robotics, WAFR 2014, 3-5 August 2014, Boğaziçi University, İstanbul, Turkey*, Springer Tracts in Advanced Robotics, pages 109–124. Springer, 2014.

[Dum07] Bogdan Dumitrescu. *Positive Trigonometric Polynomials and Signal Processing Applications*. Signals and Communication Technology. Springer, Dordrecht, 2007.

[Fej04] Leopold Fejér. Untersuchungen über Fouriersche Reihen. *Mathematische Annalen*, 58:51–69, 1904.

[FF21] Kun Fang and Hamza Fawzi. The sum-of-squares hierarchy on the sphere and applications in quantum information theory. *Mathematical Programming*, 190:331–360, 2021.

[FP16] Simon Foucart and Vladlena Powers. Basc: constrained approximation by semidefinite programming. *IMA Journal of Numerical Analysis*, 37(2):1066–1085, 2016.

[Gon12] Jacek Gondzio. Interior point methods 25 years later. *European Journal of Operational Research*, 218(3):587–601, 2012.

[GP04] Karin Gatermann and Pablo A. Parrilo. Symmetry groups, semidefinite programs, and sums of squares. *Journal of Pure and Applied Algebra*, 192(1):95–128, 2004.

[GS97] David Gottlieb and Chi-Wang Shu. On the Gibbs phenomenon and its resolution. *SIAM Review*, 39(4):644–668, 1997.

[GS19] J. Gondzio and F. N. C. Sobral. Quasi-newton approaches to interior point methods for quadratic problems. *Computational Optimization and Applications*, 74(1):93–120, 2019.

[GW95] Michel X. Goemans and David P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *J. ACM*, 42(6):1115–1145, 1995.

[HH79] Edwin Hewitt and Robert E. Hewitt. The Gibbs-Wilbraham phenomenon: An episode in Fourier analysis. *Archive for History of Exact Sciences*, 21(2):129–160, 1979.

[HK79] Michael J. Harrison and David M. Kreps. Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory*, 20(3):381–408, 1979.

[HKdK+23] Didier Henrion, Felix Kirschner, Etienne de Klerk, Milan Korda, Jean B. Lasserre, and Victor Magron. Revisiting semidefinite programming approaches to options pricing: complexity and computational perspectives. *Informs Journal on Computing*, 2023. Advance online publication.

[HLW05a] David Hobson, Peter Laurence, and Tai-Ho Wang. Static-arbitrage optimal subreplicating strategies for basket options. *Insurance: Mathematics and Economics*, 37:553–572, 2005.

[HLW05b] David Hobson, Peter Laurence, and Tai-Ho Wang. Static-arbitrage upper bounds for the prices of basket options. *Quantitative Finance*, 5(4):329–342, 2005.

[Iva08] Ivan D. Ivanov. *Parallel Implementation of Interior-Point methods for Semidefinite Optimization*. PhD thesis, Technische Universiteit Delft, 2008.

[Jac11] Dunham Jackson. *Über die Genauigkeit der Annäherung stetiger Funktionen durch ganze rationale Funktionen gegebenen Grades und trigonometrische Summen gegebener Ordnung*. PhD thesis, Universität Göttingen, 1911.

[Jac12] Dunham Jackson. On Approximation by Trigonometric Sums and Polynomials. *Transactions of the American Mathematical Society*, 13(4):491–515, 1912.

[JdK06] Dorina Jibetean and Etienne de Klerk. Global optimization of rational functions: A semidefinite programming approach. *Mathematical Programming*, 106:93–109, 2006.

[Jen06] Johan L. W. V. Jensen. Sur les fonctions convexes et les inégalités entre les valeurs moyennes. *Acta Mathematica*, 30(1):175–193, 1906.

[JL11] George M. Jabbour and Yi-Kang Liu. Option pricing and Monte Carlo simulations. *Journal of Business & Economics Research*, 3, 2011.

[JV22] Xin Jiang and Lieven Vandenberghe. Bregman primal-dual first-order method and application to sparse semidefinite programming. *Computational Optimization and Applications*, 81(1):127–159, 2022.

[Kar84] Narendra Karmarkar. A new polynomial-time algorithm for linear programming. *Combinatorica*, 4(4):373–395, 1984.

[KdK22] Felix Kirschner and Etienne de Klerk. Convergence rates of RLT and Lasserre-type hierarchies for the generalized moment problem over the simplex and the sphere. *Optimization Letters*, 16:2191–2208, 2022.

[KdK23] Felix Kirschner and Etienne de Klerk. A predictor-corrector algorithm for semidefinite programming that uses the factor width cone. arXiv preprint, arXiv:2301.06368 [math.OC], 2023.

[KdKng] Felix Kirschner and Etienne de Klerk. Construction of multivariate approximation kernels via semidefinite programming. *SIAM Journal on Optimization*, forthcoming.

[Las01] Jean B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, 11(3):796–817, 2001.

[Las08] Jean B. Lasserre. A semidefinite programming approach to the generalized problem of moments. *Mathematical Programming*, 112(1):65–92, 2008.

[Las09] Jean B. Lasserre. *Moments, Positive Polynomials and Their Applications*. Imperial College Press, 2009.

[Las11] Jean B. Lasserre. A new look at nonnegativity on closed sets and polynomial optimization. *SIAM Journal on Optimization*, 21(3):864–885, 2011.

[Las21] Jean B. Lasserre. The moment-SOS hierarchy and the Christoffel-Darboux kernel. *Optimization Letters*, 15(6):1835–1845, 2021.

[Lo87] Andrew W. Lo. Semi-parametric upper bounds for option prices and expected payoffs. *Journal of Financial Economics*, 19(2):373–387, 1987.

[LP04] Johan Löfberg and Pablo A. Parrilo. From coefficients to samples: a new approach to sos optimization. *43rd IEEE Conference on Decision and Control (CDC) (IEEE Cat. No.04CH37601)*, 3:3154–3159, 2004.

[LS21] Monique Laurent and Lucas Slot. Near-optimal analysis of Lasserre's univariate measure-based bounds for multivariate polynomial optimization. *Mathematical Programming*, 188(2):443–460, 2021.

[LS22a] Monique Laurent and Lucas Slot. An effective version of Schmüdgen's Positivstellensatz for the hypercube. *Optimization Letters*, 2022. Advance online publication.

[LS22b] Monique Laurent and Lucas Slot. Improved convergence analysis of Lasserre's measure-based upper bounds for polynomial minimization on compact sets. *Mathematical Programming*, 193(2):831–871, 2022.

[LSS05] Xun Li, Defeng Sun, and Jie Sun. Bounding option prices of multi-assets: A semidefinite programming approach. *Pacific Journal of Optimization*, 1:59–79, 2005.

[LW05] Peter Laurence and Tai-Ho Wang. Sharp upper and lower bounds for basket options. *Applied Mathematical Finance*, 12(3):253–282, 2005.

[LY15] David G. Luenberger and Yinyu Ye. *Linear and Nonlinear Programming*. Springer, 2015.

[Meg98] Robert E. Megginson. *An Introduction to Banach Space Theory*, volume 183 of *Graduate Texts in Mathematics*. Springer, 1998.

[Mer73] Robert Merton. The theory of rational option pricing. *Bell J Econ Manage Sci*, 4:141–183, 1973.

[MK87] Katta G. Murty and Santosh N. Kabadi. Some NP-complete problems in quadratic and nonlinear programming. *Mathematical Programming*, 39(2):117–129, 1987.

[MOS19] ApS MOSEK. *The MOSEK Optimization Software. Version 9.1.9*, 2019. http://docs.mosek.com/9.1/toolbox/index.html.

[MPW+21] Swann Marx, Edouard Pauwels, Tillmann Weisser, Didier Henrion, and Jean B. Lasserre. Semi-algebraic approximation using Christoffel-Darboux kernel. *Constructive Approximation*, 54(3):391–429, December 2021.

[MS65] Theodore S. Motzkin and Ernst G. Straus. Maxima for graphs and a new proof of a theorem of Turán. *Canadian Journal of Mathematics*, 17:533–540, 1965.

[Nie14] Jiawang Nie. Optimality conditions and finite convergence of Lasserre's hierarchy. *Mathematical Programming*, 146(1):97–121, 2014.

[NN94] Yurii Nesterov and Arkadii Nemirovskii. *Interior-Point Polynomial Algorithms in Convex Programming*. Society for Industrial and Applied Mathematics, 1994.

[NS07] Jiawang Nie and Markus Schweighofer. On the complexity of Putinar's Positivstellensatz. *Journal of Complexity*, 23:135–150, 2007.

[Par00] Pablo A. Parrilo. *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization*. PhD thesis, California Institute of Technology, 2000.

[PR01] Victoria Powers and Bruce Reznick. A new bound for Pólya's theorem with applications to polynomials positive on polyhedra. *Journal of Pure and Applied Algebra*, 164(1-2):221–229, 2001.

[Pri06] James Primbs. Option pricing bounds via semidefinite programming. In *Proceedings of the American Control Conference*, pages 1266–1271, 2006.

[Pro56] Yuri V. Prokhorov. Convergence of random processes and limit theorems in probability theory. *Theory of Probability & Its Applications*, 1(2):157–214, 1956.

[Put93] Mihai Putinar. Positive polynomials on compact semi-algebraic sets. *Indiana University Mathematics Journal*, 42(3):969–984, 1993.

[PVZ12] Javier Peña, Juan C. Vera, and Luis F. Zuluaga. Computing arbitrage upper bounds on basket options in the presence of bid–ask spreads. *European Journal of Operational Research*, 222(2):369–376, 2012.

[Ren01] James Renegar. *A Mathematical View of Interior-Point Methods in Convex Optimization*, volume 3 of *MPS-SIAM Series on Optimization*. Society for Industrial and Applied Mathematics, 2001.

[Rez95] Bruce Reznick. Uniform denominators in Hilbert's seventeenth problem. *Mathematische Zeitschrift*, 220(1):75–97, December 1995.

[Rez96] Bruce Reznick. Some concrete aspects of Hilbert's 17th problem. In *In Contemporary Mathematics*, pages 251–272. American Mathematical Society, 1996.

[Ric57] Hans Richter. Parameterfreie Abschätzung und Realisierung von Erwartungswerten. *Blätter der DGVFM*, 3:147–162, 1957.

[Riv69] Theodore Rivlin. *An introduction to the approximation of functions*. Blaisdell Pub. Co, 1969.

[RSS22] Biel Roig-Solvas and Mario Sznaier. A globally convergent LP and SOCP-based algorithm for semidefinite programming. arXiv preprint, arXiv:2202.12374 [math.OC], 2022.

[RTAL13] Cordian Riener, Thorsten Theobald, Lina Jansson Andrén, and Jean B. Lasserre. Exploiting symmetries in SDP-relaxations for polynomial optimization. *Mathematics of Operations Research*, 38(1):122–141, 2013.

[Sag01] Bruce E. Sagan. *The symmetric group : Representations, combinatorial algorithms and symmetric functions*. Springer New York, 2001.

[SC23] Devang Sinha and Siddhartha P. Chakrabarty. A review of efficient multilevel monte carlo algorithms for derivative pricing and risk management. *MethodsX*, 10:102078, 2023.

[Sca57] Herbert E. Scarf. *A Min-Max Solution of an Inventory Problem*. RAND Corporation, Santa Monica, CA, 1957.

[Sch91] Konrad Schmüdgen. The K-moment problem for compact semi-algebraic sets. *Mathematische Annalen*, 289(1):203–206, 1991.

[Sha01] Alexander Shapiro. *On Duality Theory of Conic Linear Problems*, volume 57 of *Nonconvex Optimization and Its Applications*, pages 135–165. Springer US, Boston, MA, 2001.

[SKL09] M. Stingl, M. Kočvara, and G. Leugering. Free material optimization with fundamental eigenfrequency constraints. *SIAM Journal on Optimization*, 20(1):524–547, 2009.

[Slo22a] Lucas Slot. *Asymptotic Analysis of Semidefinite Bounds for Polynomial Optimization and Independent Sets in Geometric Hypergraphs*. PhD thesis, Tilburg University, 2022.

[Slo22b] Lucas Slot. Sum-of-squares hierarchies for polynomial optimization and the Christoffel–Darboux kernel. *SIAM Journal on Optimization*, 32(4):2612–2635, 2022.

[Tch57] Ljubomir Tchakaloff. Formules de cubature ḿecanique cóefficients nonnégatifs. *Bulletin des Sciences Mathématiques*, 81:123–134, 1957.

[Tod01] M. J. Todd. Semidefinite optimization. *Acta Numerica*, 10:515–560, 2001.

[Tre17] LLoyd N. Trefethen. Multivariate polynomial approximation in the hypercube. *Proceedings of the American mathematical society*, 145(11):4837–4844, 2017.

[Val09] Frank Vallentin. Symmetry in semidefinite programs. *Linear Algebra and its Applications*, 430(1):360–369, 2009.

[WLC+19] Tillmann Weisser, Benoît Legat, Chris Coey, Lea Kapelevich, and Juan Pablo Vielma. *MomentOpt.jl*, 2019. Polynomial and moment optimization in Julia and JuMP, v0.2.0.

[WWAF06] Alexander Weiße, Gerhard Wellein, Andreas Alvermann, and Holger Fehske. The kernel polynomial method. *Rev. Mod. Phys.*, 78:275–306, 2006.

[ZJJ+20] Fariba Zohrizadeh, Cedric Josz, Ming Jin, Ramtin Madani, Javad Lavaei, and Somayeh Sojoudi. A survey on conic relaxations of optimal power flow problem. *European Journal of Operational Research*, 287(2):391–409, 2020.

[ZL21] Richard Y. Zhang and Javad Lavaei. Sparse semidefinite programs with guaranteed near-linear time complexity via dualized clique tree conversion. *Mathematical Programming*, 188(1):351–393, 2021.

[ZP05] Luis Zuluaga and Javier Pena. A conic programming approach to generalized Tchebycheff inequalities. *Math. Oper. Res.*, 30(2):369–388, 2005.

[ZSP22] Yang Zheng, Aivar Sootla, and Antonis Papachristodoulou. Block factor-width-two matrices and their applications to semidefinite and sum-of-squares optimization. *IEEE Transactions on Automatic Control*, 2022.

CENTER DISSERTATION SERIES

CentER for Economic Research, Tilburg University, the Netherlands

No.	Author	Title	ISBN	Published
638	Pranav Desai	Essays in Corporate Finance and Innovation	978 90 5668 639 0	January 2021
639	Kristy Jansen	Essays on Institutional Investors, Asset Allocation Decisions, and Asset Prices	978 90 5668 640 6	January 2021
640	Riley Badenbroek	Interior Point Methods and Simulated Annealing for Nonsymmetric Conic Optimization	978 90 5668 641 3	February 2021
641	Stephanie Koornneef	It's about time: Essays on temporal anchoring devices	978 90 5668 642 0	February 2021
642	Vilma Chila	Knowledge Dynamics in Employee Entrepreneurship: Implications for parents and offspring	978 90 5668 643 7	March 2021
643	Minke Remmerswaal	Essays on Financial Incentives in the Dutch Healthcare System	978 90 5668 644 4	July 2021
644	Tse-Min Wang	Voluntary Contributions to Public Goods: A multi-disciplinary examination of prosocial behavior and its antecedents	978 90 5668 645 1	March 2021
645	Manwei Liu	Interdependent individuals: how aggregation, observation, and persuasion affect economic behavior and judgment	978 90 5668 646 8	March 2021
646	Nick Bombaij	Effectiveness of Loyalty Programs	978 90 5668 647 5	April 2021
647	Xiaoyu Wang	Essays in Microeconomics Theory	978 90 5668 648 2	April 2021
648	Thijs Brouwer	Essays on Behavioral Responses to Dishonest and Anti-Social Decision-Making	978 90 5668 649 9	May 2021
649	Yadi Yang	Experiments on hold-up problem and delegation	978 90 5668 650 5	May 2021
650	Tao Han	Imperfect information in firm growth strategy: Three essays on M&A and FDI activities	978 90 5668 651 2	June 2021

No.	Author	Title	ISBN	Published
651	Johan Bonekamp	Studies on labour supply, spending and saving before and after retirement	978 90 5668 652 9	June 2021
652	Hugo van Buggenum	Banks and Financial Markets in Microfounded Models of Money	978 90 5668 653 6	August 2021
653	Arthur Beddock	Asset Pricing with Heterogeneous Agents and Non-normal Return Distributions	978 90 5668 654 3	September 2021
654	Mirron Adriana Boomsma	On the transition to a sustainable economy: Field experimental evidence on behavioral interventions	978 90 5668 655 0	September 2021
655	Roweno Heijmans	On Environmental Externalities and Global Games	978 90 5668 656 7	August 2021
656	Lenka Fiala	Essays in the economics of education	978 90 5668 657 4	September 2021
657	Yuexin Li	Pricing Art: Returns, Trust, and Crises	978 90 5668 658 1	September 2021
658	Ernst Roos	Robust Approaches for Optimization Problems with Convex Uncertainty	978 90 5668 659 8	September 2021
659	Joren Koëter	Essays on asset pricing, investor preferences and derivative markets	978 90 5668 660 4	September 2021
660	Ricardo Barahona	Investor Behavior and Financial Markets	978 90 5668 661 1	October 2021
660	Stefan ten Eikelder	Biologically-based radiation therapy planning and adjustable robust optimization	978 90 5668 662 8	October 2021
661	Maciej Husiatyński	Three essays on Individual Behavior and New Technologies	978 90 5668 663 5	October 2021
662	Hasan Apakan	Essays on Two-Dimensional Signaling Games	978 90 5668 664 2	October 2021
663	Ana Moura	Essays in Health Economics	978 90 5668 665 9	November 2021
664	Frederik Verplancke	Essays on Corporate Finance: Insights on Aspects of the General Business Environment	978 90 5668 666 6	October 2021

No.	Author	Title	ISBN	Published
665	Zhaneta Tancheva	Essays on Macro-Finance and Market Anomalies	978 90 5668 667 3	November 2021
666	Claudio Baccianti	Essays in Economic Growth and Climate Policy	978 90 5668 668 0	November 2021
667	Hongwei Zhang	Empirical Asset Pricing and Ensemble Machine Learning	978 90 5668 669 7	November 2021
668	Bart van der Burgt	Splitsing in de Wet op de vennootschapsbelasting 1969 Een evaluatie van de Nederlandse winstbelastingregels voor splitsingen ten aanzien van lichamen	978 90 5668 670 3	December 2021
669	Martin Kapons	Essays on Capital Markets Research in Accounting	978 90 5668 671 0	December 2021
670	Xolani Nghona	From one dominant growth mode to another: Switching between strategic expansion modes	978 90 5668 672 7	December 2021
671	Yang Ding	Antecedents and Implications of Legacy Divestitures	978 90 5668 673 4	December 2021
672	Joobin Ordoobody	The Interplay of Structural and Individual Characteristics	978 90 5668 674 1	February 2022
673	Lucas Avezum	Essays on Bank Regulation and Supervision	978 90 5668 675 8	March 2022
674	Oliver Wichert	Unit-Root Tests in High-Dimensional Panels	978 90 5668 676 5	April 2022
675	Martijn de Vries	Theoretical Asset Pricing under Behavioral Decision Making	978 90 5668 677 2	June 2022
676	Hanan Ahmed	Extreme Value Statistics using Related Variables	978 90 5668 678 9	June 2022
677	Jan Paulick	Financial Market Information Infrastructures: Essays on Liquidity, Participant Behavior, and Information Extraction	978 90 5668 679 6	June 2022
678	Freek van Gils	Essays on Social Media and Democracy	978 90 5668 680 2	June 2022

No.	Author	Title	ISBN	Published
679	Suzanne Bies	Examining the Effectiveness of Activation Techniques on Consumer Behavior in Temporary Loyalty Programs	978 90 5668 681 9	July 2022
680	Qinnan Ruan	Management Control Systems and Ethical Decision Making	978 90 5668 682 6	June 2022
681	Lingbo Shen	Essays on Behavioral Finance and Corporate Finance	978 90 5668 683 3	August 2022
682	Joshua Eckblad	Mind the Gales: An Attention-Based View of Startup Investment Arms	978 90 5668 684 0	August 2022
683	Rafael Greminger	Essays on Consumer Search	978 90 5668 685 7	August 2022
684	Suraj Upadhyay	Essay on policies to curb rising healthcare expenditures	978 90 5668 686 4	September 2022
685	Bert-Jan Butijn	From Legal Contracts to Smart Contracts and Back Again: An Automated Approach	978 90 5668 687 1	September 2022
686	Sytse Duiverman	Four essays on the quality of auditing: Causes and consequences	978 90 5668 688 8	October 2022
687	Lucas Slot	Asymptotic Analysis of Semidefinite Bounds for Polynomial Optimization and Independent Sets in Geometric Hypergraphs	978 90 5668 689 5	September 2022
688	Daniel Brosch	Symmetry reduction in convex optimization with applications in combinatorics	978 90 5668 690 1	October 2022
689	Emil Uduwalage	Essays on Corporate Governance in Sri Lanka	978 90 5668 691 8	October 2022
690	Mingjia Xie	Essays on Education and Health Economics	978 90 5668 692 5	October 2022
691	Peerawat Samranchit	Competition in Digital Markets	978 90 5668 693 2	October 2022
692	Jop Schouten	Cooperation, allocation and strategy in interactive decision-making	978 90 5668 694 9	December 2022
693	Pepijn Wissing	Spectral Characterizations of Complex Unit Gain Graphs	978 90 5668 695 6	November 2022

No.	Author	Title	ISBN	Published
694	Joris Berns	CEO attention, emotion, and communication in corporate financial distress	978 90 5668 696 3	November 2022
695	Tom Aben	The (long) road towards smart management and maintenance: Organising the digital transformation of critical infrastructures	978 90 5668 697 0	December 2022
696	Gülbike Mirzaoğlu	Essays in Economics of Crime Prevention and Behavior Under Uncertainty	978 90 5668 698 7	February 2023
697	Suwei An	Essays on incentive contracts, M&As, and firm risk	978 90 5668 699 4	February 2023
698	Jorgo Goossens	Non-standard Preferences in Asset Pricing and Household Finance	978 90 5668 700 7	February 2023
699	Santiago Bohorquez Correa	Risk and rewards of residential energy efficiency	978 90 5668 701 4	April 2023
700	Gleb Gertsman	Behavioral Preferences and Beliefs in Asset Pricing	978 90 5668 702 1	May 2023
701	Gabriella Massenz	On the Behavioral Effects of Tax Policy	978 90 5668 703 8	May 2023
702	Yeqiu Zheng	The Effect of Language and Temporal Focus on Cognition, Economic Behaviour, and Well-Being	978 90 5668 704 5	May 2023
703	Michela Bonani	Essays on Innovation, Cooperation, and Competition Under Standardization	978 90 5668 705 2	June 2023
704	Fabien Ize	Managing Middle Managers: How Management Control Design Choices Can Improve Managerial Decision-Making	978 90 5668 706 9	June 2023
705	Kristel de Nobrega	Cyber Defensive Capacity and Capability: A Perspective from the Financial Sector of a Small State…Shield Up!	978 90 5668 707 6	July 2023
706	Christian Peters	The Microfoundations of Audit Quality	978 90 5668 708 3	June 2023
707	Felix Kirschner	Conic Optimization with Applications in Finance and Approximation Theory	978 90 5668 709 0	July 2023

This dissertation explores conic optimization techniques with applications in the fields of finance and approximation theory. One of the most general types of conic optimization problems is the so-called generalized moment problem (GMP), which plays a fundamental part in this work. While being a powerful modeling framework, the GMP is notoriously difficult to solve. Semidefinite programming problems (SDPs) can be used to define approximation hierarchies for the GMP. The thesis includes an analysis of an interior point algorithm for SDPs, as well as a convergence analysis of an approximation hierarchy for the GMP defined over special sets. Additionally, the dissertation investigates the problem of pricing options that depend on multiple underlyings, which can be modeled as a GMP. Finally, the dissertation applies tools from conic optimization to address a classical question in approximation theory.

Felix Constantin Kirschner (Hof, Germany, 1993) received his bachelor's and master's degrees in "Wirtschaftsmathematik" (hybrid program in applied mathematics, computer science and economics) from the University of Cologne (2016 & 2019). In 2019 he joined the Department of Econometrics and Operations Research at Tilburg University as a PhD candidate as part of the Marie Skłodowska-Curie Innovative Training Network POEMA.

ISBN: 978 90 5668 709 0
DOI: 10.26116/x0f5-4k25

TILBURG SCHOOL OF ECONOMICS
AND MANAGEMENT

CentER Dissertation Series