Sums of Nonnegative Circuit Polynomials

– Geometry and Optimization –

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften

vorgelegt beim Fachbereich 12
Informatik und Mathematik
der Goethe-Universität
in Frankfurt am Main

GOETHE
UNIVERSITÄT
FRANKFURT AM MAIN

von
Mareike Dressler
aus Hanau

Frankfurt 2018
(D 30)

vom Fachbereich 12, Informatik und Mathematik, der Goethe-Universität
als Dissertation angenommen.

DEKAN
Prof. Dr. Andreas Bernig

GUTACHTER
Prof. Dr. Thorsten Theobald
Goethe-Universität Frankfurt am Main

Priv.-Doz. Dr. Gennadiy Averkov
Otto-von-Guericke-Universität Magdeburg

DATUM DER DISPUTATION: 09.05.2018

To my family.

Acknowledgements

At the beginning of this thesis I want to convey my deepest gratefulness to many people for their help and support during my research and throughout this work.

First and foremost, I would like to thank my advisor, *Thorsten Theobald*, for giving me the opportunity to explore the beautiful areas of real algebraic geometry and optimization. During the past four years he has consistently offered advice, motivation, support, and encouragement. I appreciate the fruitful impulses he gave me and all the free space to develop my own ideas. I am also very grateful for all the conferences and workshops he enabled me to attend. They were a great inspiration and had a significant benefit on my research.

Moreover, I would like to thank *Gennadiy Averkov* for his kind willingness to act as a second reviewer for this thesis. In addition, I really appreciate that *Andreas Bernig*, *Hans Crauel*, *Raman Sanyal*, and *Jürgen Wolfart* agreed to participate in my doctoral committee.

I am deeply grateful to my collaborators *Sadik Iliman* and *Timo de Wolff* for many insightful discussions, sometimes carrying on late into the night, and shared struggles. I learned a lot from our work together and enjoyed every minute of it.

I would like to thank all former and current members of my research group for their friendship, support, and pleasant coffee breaks: *Tomáš Bajbar*, *Sadik Iliman*, *Thorsten Jörgens*, *Lukas Katthän*, *Kai Kellner*, *Martina Juhnke-Kubitzke*, and *Christian Trabandt*. My appreciation also extends to other colleagues from Frankfurt, in particular *Sven Jarohs*, *Ute Lenz*, *Sebastian Manecke*, and *Raman Sanyal*. Thank you all for the wonderful time.

I have been very fortunate to have *Samuel Hetterich*, *Janina Hüttel*, *Ralf Lehnert*, and *Felicia Raßmann* as my fellow students. We have become deep friends and I am happy you are in my life.

I owe great thanks to the people I have met at conferences or otherwise along the way for their many interesting conversations. My special thanks go to *Bernd Sturmfels* for being himself. Furthermore, I thank *Markus Schweighofer* for pointing out the implication of Krivine's special case in Remark 5.3.6.

I am very grateful to my colleagues and friends for proofreading parts of this thesis and giving valuable feedback.

Finally, I thank my family and friends for their endless patience and motivation. My deepest thanks go to my family, *Mom*, *Dad*, and *Melanie* & *Björn*, for always encouraging and believing in me. Their love and support has been constant and invaluable. Thank you.

Contents

Chapter 1

Introduction

The results of this thesis lie at the intersection of real algebraic geometry, convex geometry, and optimization, thus in the area of convex algebraic geometry.

Deciding *nonnegativity* of real polynomials is a key problem in real algebraic geometry and polynomial optimization. The question is:

Given a polynomial $f \in \mathbb{R}[\mathbf{x}] = \mathbb{R}[x_1, \ldots, x_n]$ does it hold $f(\mathbf{x}) \geqslant 0$ for all $\mathbf{x} \in \mathbb{R}^n$?

The objective of global polynomial optimization is to minimize a real multivariate polynomial f over $\mathbb{R}^n$, i.e., to find the optimal value $f^* = \inf\{f(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^n\}$. It is easy to see that searching for a global lower bound of the polynomial f is equivalent to finding the largest real number γ such that $f - \gamma$ is nonnegative. This equivalence suggests considering the alternative optimization problem

$$f^* = \sup\{\gamma \in \mathbb{R} : f(\mathbf{x}) - \gamma \geqslant 0 \text{ for all } \mathbf{x} \in \mathbb{R}^n\}.$$

Thereby, a polynomial optimization problem can be reduced to the question of deciding nonnegativity of a polynomial.

Both the decision and the optimization version of this problem have countless applications for example in dynamical systems, robotics, control theory, computer vision, signal processing, and economics, for an overview see, e.g., [BPT13] and [Las10].

Since deciding nonnegativity of a polynomial f is co-NP-hard if f is multivariate and of degree at least four [MK87], one is interested in finding sufficient conditions to *certify* nonnegativity of polynomials, which are easier to check. Such a certificate is given by *sums of squares* (SOS), which are obviously nonnegative. Thus, if we can write a polynomial f as a sum of squares of polynomials, then it is apparent from this representation that f is nonnegative. The relation between nonnegative polynomials and sums of squares is a classical question in real algebraic geometry and has its origin in Hilbert's work at the end of the nineteenth century. He intensively studied $P_{n,2d}$, the cone of nonnegative polynomials in n variables of degree at most $2d$ and $\Sigma_{n,2d}$, the cone of sums of squares respectively. This study led to his seminal result [Hil88], stating that the two cones coincide in the univariate case, in the quadratic case, and for binary quartics. In all other cases Hilbert showed the existence of nonnegative polynomials which are not sums of squares. His proof was nonconstructive and the first explicit example for such a polynomial was given only seventy years later by Motzkin [Mot67]. His observations led Hilbert to his famous question, known as *Hilbert's 17th problem*, whether every nonnegative polynomial has a representation as a sum of squares of rational functions. This was solved in the affirmative, in 1927, by Emil Artin [Art27]. See [Rez00] for a historical overview.

The benefit of using SOS certificates is apparent from the practical viewpoint, because checking if a polynomial is a sum of squares can be formulated as a *semidefinite programming problem (SDP)*, a specific subclass of convex optimization problems [BV04, VB96], which can be seen as a generalization of linear programming. There exist good numerical algorithms for solving SDPs (to any arbitrary precision) in polynomial time, see [BPT13, page 41]. Thus, one can relax the nonnegativity condition in polynomial optimization problems both in the unconstrained case and in the constrained case to an SOS condition, which can be computed efficiently by semidefinite programming. This SOS/SDP approach for polynomial optimization problems goes back to Shor [Sho87b] in 1987 and was further developed by Nesterov [Nes00], Parrilo [Par00, Par03], and Lasserre [Las01]. Starting with these works a variety of relaxation methods have been proposed in the literature, which are studied intensively by means of aspects like exactness and quality of the relaxations [dKL10, Nie13a, Nie13b, Nie14], the speed of the computations [Las10, PS03], and geometrical aspects of the underlying structures [Ble06, Ble12]. A great majority of these results are based on *Lasserre's relaxation* [Las01], which relies on the SOS/SDP method and yields a hierarchy of lower bounds converging to the optimal value of the constrained optimization problem; see, e.g., [Las10], [Las15].

A well known issue of the SOS/SDP approach is that the size of the corresponding semidefinite programs grows rapidly with the number of variables or degree of the polynomials, which makes them challenging to compute. Hence, for many applications, the problems are too large or numerical issues are too severe to find a (proper) solution. Furthermore, Blekherman [Ble06] proved that for fixed degree $2d \geqslant 4$, there are significantly more nonnegative polynomials than sums of squares as the number of variables tends to infinity.

Attacking these issues is an active area of research. The aim of this thesis is to contribute to this area both from the theoretical and practical point of view.

To address the above mentioned issues current research creates better solvers and exploits additional structure like symmetry, see, e.g., [dKS10, GP04, RTAL13, Val09], or sparsity, e.g., [Las06a], [WKKM06]. Lasserre et al. [LTY17] propose to use bounded degree SOS hierarchies and recently, Papp and Yıldız [PY17] study an approach to polynomial optimization problems which circumvents the usage of semidefinite programs. Moreover, Ahmadi and Majumdar [AM17] introduce two subcones of the SOS cone that one can optimize over using linear and second order cone programming.

In contrast, in this thesis we pursue a different approach, namely to use other nonnegativity certificates independent of SOS certificates.

Recently, Iliman and de Wolff [IdW16a] established a new certificate for nonnegativity of real polynomials via *sums of nonnegative circuit polynomials (SONC)*. These polynomials are sums of certain sparse polynomials having a special structure in terms of their support. More precisely, the Newton polytope of a circuit polynomial f forms a simplex with even vertices, the coefficients of the terms of f corresponding to the vertices of this simplex are strictly positive, and there is one additional point in the support of f which is located in the interior of the Newton polytope. For every circuit polynomial we can define the corresponding *circuit number* as a specific product which can be derived by the initial circuit polynomial immediately. The crucial fact is that nonnegativity of circuit polynomials can be decided easily by means of its circuit number alone. This naturally leads to defining the set of *sums of nonnegative circuit polynomials*, which is denoted by $C_{n,2d}$ for n-variate polynomials of degree at most $2d$. Furthermore, $C_{n,2d}$ forms a convex cone which intersects with the cone of sums of squares $\Sigma_{n,2d}$, but they do not contain each other. Hence, sums of nonnegative circuit polynomials are indeed a new *nonnegativity certificate* for real polynomials, which is *independent* of sums of squares.

In this thesis we focus on sums of nonnegative circuit polynomials and their related cone $C_{n,2d}$ and study these geometrically as well as in applications to polynomial optimization, which leads to new results in the area of both pure and applied real and convex algebraic geometry. The thesis can be divided into two parts, namely the theoretical analysis of the SONC cone and the practical in application to polynomial optimization. In the subsequent paragraphs, we outline the investigated problems and provide an overview of the results and contributions of this thesis.

The SONC Cone revisited. As sums of nonnegative circuit polynomials are a rather new concept to certify nonnegativity of real polynomials, these polynomials and their related cone are to a large extent unexplored but their study entails high potential for further research. From the theoretical point of view, being a convex cone comprised of polynomials with a certain structure, the SONC cone itself is interesting and offers many open questions to address. But even more, since $C_{n,2d}$ approximates the nonnegativity cone $P_{n,2d}$, gaining a deeper understanding of the SONC cone is highly desirable from the perspective of both pure and applied real and convex algebraic geometry. Its analysis fits therefore naturally in the long and rich theory of nonnegative polynomials and sums of squares. Hence, exploring the structure and (convex) properties of $C_{n,2d}$ as well as its relation to $P_{n,2d}$ and $\Sigma_{n,2d}$ is important.

With this in mind we first study some convex geometric aspects of the SONC cone. We show in Proposition 3.1.1 that $C_{n,2d}$ is a proper cone. In [IdW16a] the authors characterized the cases $(n, 2d)$, where the two cones $C_{n,2d}$ and $\Sigma_{n,2d}$ contain respectively not contain each other, see Theorem 2.4.8. Two cases are not covered in Theorem 2.4.8 (3), namely $(n, 2)$ for all $n \geqslant 2$ and the case $(n, 4)$ for all n. We close this gap in Theorem 3.1.2 and provide a proof for the missing cases.

So far, we limited our analysis to polynomials. A related construct are homogeneous polynomials, also called forms, which are ubiquitous in mathematics. Homogeneous polynomials are polynomials whose non-zero terms all have the same degree. In algebraic geometry forms are a fundamental object of study and often results in conjunction to nonnegative polynomials and sums of squares are stated homogeneously. Thus we also want to consider SONC forms. As a first result in this context we prove the fundamental fact that the property to be SONC is preserved under homogenization.

An interesting research subject for polynomials and forms is the study of their real zeros. There exist a large number of works studying the real zeros of nonnegative polynomials and sums of squares, which is often used to explore the difference between both cones and to get an insight into the facial structure of $P_{n,2d}$ and $\Sigma_{n,2d}$, see, e.g., [BHO+12, Ble12, CL77, CLR80, KS18, Rez78, Rez00]. Motivated by these ideas, we investigate the real zeros of SONC polynomials and forms. Our main contribution to this topic is a *complete* and *explicit characterization* of the real zeros of SONC polynomials and forms in Section 3.2. These results yield interesting further observations. For instance we show that the analog of Hilbert's 17th problem for SONC polynomials cannot hold in the general case. Based upon the study of real zeros we provide a first approach to the exposed faces of the SONC cone. Particularly, we derive estimates for the dimensions of the exposed faces of $C_{n,2d}$ and study the univariate and bivariate case in more detail together with some explicit examples.

A basic property of SOS is that the set of sums of squares is multiplicatively closed. This property is essential in the application of SOS to polynomial optimization, in particular for certain Positivstellensätze, see Section 2.3.4. We show in Lemma 3.4.1 that the set of SONC polynomials is *not closed under multiplication*, which therefore stands in strong contrast to the set of sums of squares. Another main contribution to the analysis of the SONC cone is the result that $C_{n,2d}$ is *full-dimensional* in the convex cone of nonnegative polynomials $P_{n,2d}$. This result is a necessary condition to establish SONC polynomials as a certificate useful in practice. Hence, both observations likewise have a direct impact on the application of SONC polynomials to polynomial optimization problems. This applied perspective will be discussed in the next paragraphs.

An Approach to Polynomial Optimization via SONC and GP. As already mentioned, getting a broader insight into the SONC cone and the SONC polynomials is also of crucial importance from the practical point of view. Since SONC polynomials serve as a certificate of nonnegativity they can be applied to polynomial optimization problems. The second part of this thesis is devoted to these applications.

Besides the SDP-based approach to polynomial optimization problems Ghasemi and Marshall [GM12, GM13] recently proposed using geometric programming to find lower bounds for polynomials both in the unconstrained and in the constrained case. *Geometric programs (GP)* are a special type of convex optimization problems that can be solved in polynomial time (up to an ε-error) via interior point methods [NN94]; see also [BKVH07, page 118]. Experimental results show that compared to semidefinite programs in practice the corresponding geometric programs can be solved *significantly* faster, see, e.g., [BKVH07, GM12, GM13, GLM14]. A disadvantage of the method of Ghasemi and Marshall is that the lower bounds obtained by geometric programming are, however, by construction not as good as the lower bounds obtained via semidefinite programming, and that it is restricted to very special cases.

The idea of using a GP-based approach for unconstrained optimization can be further generalized via SONC certificates for certain polynomials, as shown by Iliman and de Wolff [IdW16b]. To be more precise, deciding whether an *ST-polynomial* has a SONC decomposition can be checked efficiently with GP. ST-polynomials are polynomials having a Newton polytope that is a simplex and satisfying further conditions; see Definition 4.1.1. Thus, the connection of SONC and GP is in direct analogy to the relation between SOS and SDP. One crucial difference to Ghasemi and Marshall's approach is that there exist various classes of polynomials for which the SONC/GP-based approach is not only *faster* but, it also yields *better* bounds than the SOS/SDP approach; see [IdW16b, Corollary 3.6]. The reason is that all certificates used by Ghasemi and Marshall are always SOS, while SONC polynomials are not SOS in general; see Theorem 2.4.8.

Motivated by these recent developments we focus in the second part of this thesis on tackling constrained optimization problems with SONC polynomials, i.e., problems of the form $f_K^* = \inf\{f(\mathbf{x}) : \mathbf{x} \in K\} = \sup\{\gamma \in \mathbb{R} : f(\mathbf{x}) - \gamma \geqslant 0 \text{ for all } \mathbf{x} \in K\}$, with feasible set $K \subseteq \mathbb{R}^n$ given as the basic closed semialgebraic set K defined by polynomials $g_1, \ldots, g_s \in \mathbb{R}[\mathbf{x}]$. Essentially, we follow this aim via two different approaches. In a first step, as a *generalization* of the above mentioned SONC/GP approach, this means deriving a lower bound for the optimal value f_K^* of the constrained optimization problem by using a *single* convex optimization program, which is a GP under certain extra assumptions. In a second step, an extended approach is analyzed yielding a *hierarchy* of lower bounds which converge to f_K^*. The hierarchical approach will be discussed in the subsequent paragraph.

The first contribution of this part in this context is an *extension* of the results in [IdW16b] to constrained polynomial optimization problems for the class of ST-polynomials. The starting point is a general optimization problem from [IdW16b, Section 5], see program (4.1.5), which provides a lower bound for the constrained problem but which is not a GP. Using results from [GM13], we relax the program (4.1.5) into a geometric optimization problem; see program (4.2.2) and Theorem 4.2.1. Additionally, we show in Theorem 4.2.4 that (4.1.5) can always at least be transformed into a signomial program; see Section 4.1.2 for background information. Moreover, we prove that the new, relaxed geometric program (4.2.2) yields bounds as good as the initial program (4.1.5) for certain special cases, see Theorem 4.2.5.

In Section 4.3 we provide examples comparing our new program (4.2.2) with semidefinite programming in practice. In all these examples our program is *much faster* than semidefinite programming. Particularly, we demonstrate that, in sharp contrast to SDPs, increasing the degree of a given problem has almost *no* effect on the runtime of our program (4.2.2). Hence the GP-based approach is particularly useful for high-degree problems, where SDPs have serious issues.
Furthermore, a bound obtained by Ghasemi and Marshall in [GM13] can never be better than the bound given by the d-th Lasserre relaxation for some specific d determined by the degrees of the involved polynomials. Section 4.3 contains exam-

ples showing that our program (4.2.2) can provide bounds which are *better* than the particular d-th Lasserre relaxation.

The second contribution is to apply polynomial optimization methods based on SONC polynomials and GPs efficiently *beyond* the class of ST-polynomials. In Section 4.4, we develop an initial approach based on triangulations of the support sets of the involved polynomials. It yields bounds for nonnegativity based on SONC/GP for *arbitrary* polynomials both in the unconstrained and in the constrained case. We provide several examples and compare the new bounds to the ones obtained by SDP-based methods. In all examples, particularly those with high degree, our GP-based method is (significantly) faster than SDP.

We point out that in both approaches we make *no* assumptions about the feasible set K. Especially, it is not required to be compact as it is in the classical setting with Lasserre's relaxation.

Hierarchical Approach to Constrained Optimization Problems via SONC and REP. Directly picking up on the idea to tackle constrained optimization problems with SONC polynomials we want to extend the studied SONC/GP-based approach. Since this method only yields a single lower bound we provide a new approach leading to a hierarchy of lower bounds converging to the optimal value f_K^* of the constrained optimization problem. The main difference between the two approaches is that the latter is based on a Positivstellensatz. Positivstellensätze play a key role in the development of constrained polynomial optimization problems and have an even longer theoretical history. Roughly speaking, a Positivstellensatz guarantees that a polynomial, which is strictly positive on a semialgebraic set, can be represented algebraically in a specific way. There exist various Positivstellensätze typically relying on sums of squares, see Section 2.3.3. For instance, Lasserre's relaxation is based on *Putinar's Positivstellensatz* [Put93].

Recently, Chandrasekaran and Shah [CS16] introduced *sums of nonnegative arithmetic geometric exponentials (SAGE)* as nonnegativity certificate for signomials, which are weighted sums of exponentials. Hence this concept addresses the problem of deciding nonnegativity of polynomials on the positive orthant. Checking whether an AM/GM-exponential is nonnegative can be done by *relative entropy programming (REP)*. An REP is a convex optimization program, which is more general than a geometric program, but still efficiently solvable via interior point methods; see [CS17, NN94].

The foundation of the hierarchical approach is provided by a *Positivstellensatz for*

SONC polynomials, see Theorem 5.3.5, which is basically a consequence of Krivine's Positivstellensatz [Kri64a, Kri64b]. It roughly states that a polynomial f being strictly positive on a *compact* set K can be represented via the constrained polynomials weighted by SONC polynomials. Due to the Positivstellensatz we can define the parameter $f_{\text{sonc}}^{(d,q)}$ as the largest real number γ such that $f(\mathbf{x}) - \gamma$ has a SONC representation as given in the Positivstellensatz. Clearly, this parameter is a lower bound for f_K^*, which is based on the maximal allowed degree of the representing polynomials in the Positivstellensatz. Moreover the lower bound grows monotonically in d and q, see Lemma 5.4.1, and thus yields a hierarchy of lower bounds for f_K^*. The main contribution to the area of polynomial optimization is the key result that on the one hand the provided hierarchy is *complete*, this means the lower bounds $f_{\text{sonc}}^{(d,q)}$ converge to f_K^* for $d, q \to \infty$, see Theorem 5.4.2 and that on the other hand the bounds $f_{\text{sonc}}^{(d,q)}$ are *efficiently computable*. More precisely, we provide in (5.4.3) an optimization program for the computation of $f_{\text{sonc}}^{(d,q)}$ and prove in Theorem 5.4.3 that this program (5.4.3) is a relative entropy program. This connection was inspired by the above mentioned new concept of the SAGE cone, which is related to the SONC cone. Therefore, we additionally provide a first comparison of these two cones, see Section 5.2.

In Section 5.4.3 we illustrate the new method with an example.

1.1 Structure of the Thesis

In Chapter 2 we provide a broad overview of the theory and results needed for this thesis. After introducing notation and recalling basic concepts from the theory of convexity and about polynomials we focus in Section 2.2 on the cones of nonnegative polynomials and sums of squares. We discuss the connection of sums of squares to semidefinite programming, the quantitative relationship between the two cones, and provide some facts about their dual cones as well as their boundaries and facial structure. Hereafter we study the background of polynomial optimization problems and real algebraic geometry like SOS relaxations, Positivstellensätze, and the famous Lasserre relaxation for constrained optimization problems, see Section 2.3. In Section 2.4 we finally present the key object of study of this thesis, sums of nonnegative circuit polynomials, and state the theory required for our further study.

Chapter 3 is dedicated to the convex geometric study of the SONC cone. First, we present some properties and general results concerning the structure of the SONC cone and its relation to the SOS cone. In Section 3.2 we focus on the real zeros of

SONC polynomials and forms resulting in a complete and explicit characterization of these zeros. Afterwards, we discuss some interesting consequences of the previous observations on the real zeros of SONCs. Based on this new knowledge we provide a first approach to the exposed faces of the SONC cone, see Section 3.3. We give a deeper analysis of the univariate and bivariate case and establish estimates for the dimensions of the exposed faces of the SONC cone. In contrast to SOS we show in Section 3.4 that the set of SONC polynomials is not closed under multiplication. Furthermore, we present the important result that the SONC cone is full-dimensional in the cone of nonnegative polynomials.

The next two chapters are devoted to the practical study of the SONC cone in application to constrained polynomial optimization problems. In Chapter 4 we investigate this problem by deriving a single lower bound for the optimal value computable by a geometric program. First, we introduce ST-polynomials, the polynomials considered in the next sections, and geometric programs. Then we review the SONC/GP-based approach for the unconstrained case and an initial approach to the constrained case, which is based on the idea of tracing back the constrained problem to the unconstrained one. Unfortunately, this approach yields a lower bound for the optimal value, which is not given by a GP. In Section 4.2 we extend the result for the constrained case and provide relaxations which are computable via geometric programming. In addition, we discuss some examples comparing our new approach with SDP in practice, see Section 4.3. Finally, we generalize the SONC/GP approach in Section 4.4 both in the unconstrained and in the constrained case to non-ST-polynomials.

Chapter 5 studies an extended approach to constrained polynomial optimization problems yielding a hierarchy of lower bounds which converge to the optimal value. We begin by introducing the cone of sums of nonnegative AM/GM-exponentials and relative entropy programs. After a comparison of the SONC and the SAGE cone in Section 5.2 we state the Positivstellensatz for SONC polynomials which provides the basis for the following approach, see Section 5.3. Based on the representation given in the Positivstellensatz we establish in Section 5.4 a hierarchy of lower bounds for the optimal value of the constrained polynomial optimization problem on a basic closed semialgebraic set, and we formulate an optimization problem for the computation of these bounds. We derive the important result that for a compact constrained set the provided hierarchy is complete and efficiently solvable via relative entropy programming. Conclusively, we consider an example which provides a decomposition of a given polynomial in the form described in the Positivstellensatz for SONC polynomials.

We conclude this thesis in Chapter 6 with final remarks and a discussion of open problems.

1.2 Published Contents in Advance

Parts of this thesis are already published or submitted for publication and are based on works with co-authors. The content of Chapter 4 is based on joint work with Sadik Iliman and Timo de Wolff and is contained in [DIdW18]. Section 3.4 and Chapter 5 is based on [DIdW17], which is also a joint work with Sadik Iliman and Timo de Wolff. In the main parts of these chapters, as well as in some parts of this introduction and Chapter 6, the phrasing is a verbatim adoption from the mentioned papers with minor changes throughout for consistency with other chapters and additional comments.

Chapter 2

Preliminaries

In this chapter our aim is to discuss the motivation, the background, the main problem, and the key object of study of this thesis. We try to give a preferably broad outline of the theory required for the understanding of the following chapters and also mention suitable references to the addressed topics for a deeper study. Most proofs of the statements are omitted in favor of a more holistic exposure.

In the first section we fix terminology and recall basic concepts from convex geometry and about polynomials.

Motivated by our main problem of deciding polynomial nonnegativity, we introduce in the second section nonnegative polynomials and an important nonnegativity certificate, namely sums of squares. After providing the most important bases and a short historical classification the reader will be familiarized with an important subclass of convex optimization problems, a semidefinite programming problem. The reason for considering these problems is that checking if a polynomial is a sum of squares can be formulated as a semidefinite feasibility problem. This connection will be discussed afterwards. Then we report on the quantitative relationship between the cones of nonnegative polynomials and sums of squares as well as some convexity properties of these cones, like the dual cones, the boundaries, and their facial structure.

In the third section we study polynomial optimization, one of the most important application of nonnegative polynomials and sums of squares. We start by discussing the special case of global optimization and recall the basic idea of sums of squares relaxations for this problem. We then give a short overview of the related field of moments and derive the important result that the moment sequences yield the dual viewpoint of the study of nonnegative polynomials and sums of squares. Next we present the relationship between classical algebraic geometry and real algebraic geometry. Along the way, we state important results concerning the representation of a polynomial and the required algebraic geometric background to understand their importance. Subsequently, we discuss the constrained optimization problem and establish a common approach via sums of squares for tackling those problems. We conclude this section by giving a motivation to study the key objects of the thesis, sums of nonnegative circuit polynomials.

The last section is devoted to the introduction of the rather new concept of circuit polynomials, which provide the theoretical basis to our subsequent research. After presenting these polynomials and stating some theory, including the main outcome that these polynomials are a new nonnegativity certificate independent of sums of squares, we develop their relation to sums of squares. Finally, we give a short overview of further results about nonnegative circuit polynomials.

2.1 Notation, Convexity, and Polynomials

In this section we introduce some basic notation and preliminaries about the theory of convexity and about polynomials. For more details see, e.g., [Bar02], [BCR98], and [Zie95].

We always denote by $\mathbb{N} = \{0, 1, 2, \ldots\}$ the set of nonnegative integers, by $\mathbb{Z}$ the ring of integers, by $\mathbb{R}$ the field of real numbers, and by $\mathbb{C}$ the algebraically closed field of complex numbers. $\mathbb{R}_{\geqslant 0}$ and $\mathbb{R}_{>0}$ indicate the nonnegative and positive elements of $\mathbb{R}$, respectively. We also introduce the notation $\mathbb{N}^* = \mathbb{N}\backslash\{0\}$, and similarly $\mathbb{R}^* = \mathbb{R}\backslash\{0\}$.

Throughout the thesis bold letters denote n-dimensional vectors unless noted otherwise. Let δ_{ij} be the *ij-Kronecker symbol* and $\mathbf{e}_i = (\delta_{i1}, \ldots, \delta_{in})$ be the *i-th standard vector*. For a finite set $A \subset \mathbb{N}^n$ we denote by $\operatorname{conv}(A)$ the *convex hull* of A, and by $V(A)$ the set of all the *vertices* of $\operatorname{conv}(A)$. Analogously, we identify by $V(P)$ the vertex set of any given polytope P.

We call a lattice point $\boldsymbol{\alpha} \in \mathbb{Z}^n$ resp. $\mathbb{N}^n$ *even* if every entry α_i is even, i.e., $\boldsymbol{\alpha} \in (2\mathbb{Z})^n$ resp. $(2\mathbb{N})^n$. Furthermore, we denote by $\Delta_{n,2d}$ the *standard simplex* in n variables of edge length $2d$, i.e., the simplex satisfying $V(\Delta_{n,2d}) = \{\mathbf{0}, 2d \cdot \mathbf{e}_1, \ldots, 2d \cdot \mathbf{e}_n\}$ and we define $\mathcal{L}_{n,2d} = \Delta_{n,2d} \cap \mathbb{Z}^n$ as the *set of all integer points* in $\Delta_{n,2d}$.

Given a convex set $S \subset \mathbb{R}^n$, a *face* of S is a subset $F \subseteq S$ such that for any point $p \in F$, whenever p can be written as a convex combination of elements in S, these elements must belong to F. A face F such that $\emptyset \subsetneq F \subsetneq S$ is called *proper*. The *dimension* of a face F is defined as the dimension of its affine hull, i.e., $\dim(F) := \dim(\operatorname{aff}(F))$. An

element $p \in S$ is called *extremal* in S if $S \backslash \{p\}$ is still convex. It is fairly obvious that the extremal points are exactly the zero-dimensional faces of S. The set of all extremal elements of S is given by $\mathcal{E}S$. Moreover, we say that a face F is an *exposed face* of S if there exists a nontrivial supporting hyperplane H to S such that $F = S \cap H$.
We denote by $\text{int}(S)$ the *interior* and by ∂S the *boundary* of S.
Furthermore, we let $\text{Vol}\, S$ be the *volume* of a convex body $S \subset \mathbb{R}^n$. Observe that, if we expand S by a constant factor α, the volume satisfies $\text{Vol}(\alpha S) = \alpha^n \text{Vol}\, S$.
An important object in convexity theory is the *convex cone* $C \subset \mathbb{R}^n$. A convex cone C is a convex set such that for any $p \in C$ and $\lambda \in \mathbb{R}_{\geqslant 0}$ it holds $\lambda p \in C$. We say that p is an *extreme ray* of C if the following holds:

$$\text{If } \ p = p_1 + p_2, \ p_1, p_2 \in C, \ \text{ then } \ p_i = \lambda_i p, \ i \in \{1, 2\}, \ \text{ for some } \lambda_i \geqslant 0.$$

The set of all extreme rays of C is also denoted by $\mathcal{E}C$. $\mathcal{E}C$ plays a major role in determining the structure and the behavior of the cone C, since $C = \text{conv}(\mathcal{E}C)$. For every cone $C \subset V$ in a finite-dimensional vector space V over an ordered field K we can define its *dual cone* by $C^\vee = \{l \in V^\vee : l(x) \geqslant 0 \text{ for all } x \in C\}$, with the dual space $V^\vee = \text{Hom}(V, K)$ of all linear functionals on V. For closed convex cones the biduality theorem states that $(C^\vee)^\vee = C$, see [Bar02].
Finally we recall Carathéodory's Theorem, which provides an upper bound on the number of points in a set S needed to express a point in the convex hull of S.

Theorem 2.1.1. *Let $S \subset \mathbb{R}^n$. Then any point $p \in \text{conv}(S)$ can be written as a convex combination of at most $n+1$ points in S.*

Let $\mathbb{R}[\mathbf{x}] = \mathbb{R}[x_1, \ldots, x_n]$ be the ring of real n-variate polynomials. We usually consider *polynomials* $f \in \mathbb{R}[\mathbf{x}]$ supported on a finite set $A \subset \mathbb{N}^n$. Thus, f is of the form $f(\mathbf{x}) = \sum_{\boldsymbol{\alpha} \in A} f_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}}$ with $f_{\boldsymbol{\alpha}} \in \mathbb{R}$ and the *monomial* $\mathbf{x}^{\boldsymbol{\alpha}} = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ whose *degree* is $|\boldsymbol{\alpha}| = \sum_{i=1}^n \alpha_i$. The degree of the polynomial f is given by the maximum degree over all appearing monomials, i.e., $\deg(f) = \max\{|\boldsymbol{\alpha}| : f_{\boldsymbol{\alpha}} \neq 0\}$. We call a polynomial a *sum of monomial squares* if all terms $f_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}}$ satisfy $f_{\boldsymbol{\alpha}} > 0$ and $\boldsymbol{\alpha}$ is even.
The set of all polynomials of degree less than or equal to $2d$ is denoted by $\mathbb{R}[\mathbf{x}]_{2d}$ and if we want to emphasize the number n of variables we refer to it by $\mathbb{R}[\mathbf{x}]_{n,2d}$. Often polynomials in $\mathbb{R}[\mathbf{x}]_{n,2d}$ are termed "*n-ary $2d$-ics*". By identifying a polynomial with its $N(n, 2d) := \binom{n+2d}{2d}$ coefficients it follows that $\dim(\mathbb{R}[\mathbf{x}]_{n,2d}) = \binom{n+2d}{2d} = N(n, 2d)$. Therefore the real vector space $\mathbb{R}[\mathbf{x}]_{n,2d}$ is finite-dimensional, and in fact is isomorphic to $\mathbb{R}^{N(n,2d)}$, i.e., $\mathbb{R}[\mathbf{x}]_{n,2d} \sim \mathbb{R}^{N(n,2d)}$.

For better understanding the behavior of polynomials there is a useful tool which translates a polynomial into a geometric object, the Newton polytope. The *Newton polytope* of a polynomial f is defined as $\mathrm{New}(f) = \mathrm{conv}\{\boldsymbol{\alpha} \in A : f_{\boldsymbol{\alpha}} \neq 0\}$.

A polynomial in which all terms are of the same degree is called a *homogeneous polynomial* or a *form*. If $f \in \mathbb{R}[\mathbf{x}]_{n,2d}$ is any polynomial, then

$$\overline{f}(x_0, \dots, x_n) \;=\; x_0^{2d} f\left(\frac{x_1}{x_0}, \dots, \frac{x_n}{x_0}\right)$$

is the *homogenization* of f, which is a form of degree $2d$ in the $n+1$ variables $x_0, x_1, \dots, x_n$. Given a form $\overline{f}$ we can *dehomogenize* it by setting $x_0 = 1$. In this thesis we will mostly work with polynomials, except when analyzing the SONC cone in Chapter 3 where we use both viewpoints. To distinguish we fix the above notation and always write polynomials as $f \in \mathbb{R}[\mathbf{x}]_{n,2d}$ and forms as $\overline{f} \in \mathbb{R}[x_0, \mathbf{x}]_{n+1,2d}$.
Observe that by homogeneity we have

$$\overline{f}(\lambda x_0, \dots, \lambda x_n) \;=\; \lambda^{2d}\, \overline{f}(x_0, \dots, x_n), \;\; \lambda \in \mathbb{R},$$

for any form $\overline{f} \in \mathbb{R}[x_0, \mathbf{x}]_{n+1,2d}$. In particular it follows

$$\overline{f}(\lambda x_0, \dots, \lambda x_n) = 0 \iff \overline{f}(x_0, \dots, x_n) = 0,$$

i.e., we can consider $\overline{f}$ as a function on the real projective n-space $\mathbb{P}^n$. Finally, we define the *zero-set* of a polynomial f respectively of a form $\overline{f}$ by

$$\begin{aligned} \mathcal{V}(f) &:= \{(a_1, \dots, a_n) \in \mathbb{R}^n : f(a_1, \dots, a_n) = 0\}, \\ \mathcal{V}(\overline{f}) &:= \left\{[a_0 : \cdots : a_n] \in \mathbb{P}^n : \overline{f}(a_0, \dots, a_n) = 0\right\}. \end{aligned}$$

In the algebraic context this set is often referred to as the *real (affine) variety* resp. *real projective variety*. We denote by $|\mathcal{V}(\cdot)|$ the number of distinct elements in the zero-set. The zero-set of a form may be viewed as the set

$$\mathcal{V}(\overline{f}) \;=\; \{(a_0, \dots, a_n) \in \mathbb{R}^{n+1} \backslash \{0, \mathbf{0}\} : \overline{f}(a_0, \dots, a_n) = 0\},$$

where $|\mathcal{V}(\overline{f})|$ will be interpreted as the number of lines in $\mathcal{V}(\overline{f})$ and we only count one representative of each line.
In a natural way, there may occur *zeros of f at infinity* via homogenization. This is the case if $a_0 = 0$ for $(a_0, \boldsymbol{a}) \in \mathcal{V}(\overline{f})$. If $a_0 \neq 0$, then $(a_0, \boldsymbol{a})$ corresponds to a unique zero of f.

2.2 The Cone of Nonnegative Polynomials and Sums of Squares

In this section we establish the basis of the underlying main problem considered in this thesis, namely deciding and certifying nonnegativity of polynomials. This problem is a key challenge in real algebraic geometry and polynomial optimization. Here we give the theoretical background and examine the relationship between nonnegative polynomials and sums of squares, and in Section 2.3 we turn towards the optimization viewpoint of the problem. We refer to Reznick [Rez00] for a historical overview of this topic, and to [BPT13, Lau09, Mar08, Par00, PD01, Sch09] for a detailed discussion and background.

In real algebraic geometry nonnegative polynomials are a central object of study.

Definition 2.2.1. A multivariate real polynomial $f \in \mathbb{R}[\mathbf{x}]$ is *nonnegative* if it takes only nonnegative values, i.e.,

$$f(\mathbf{x}) \geqslant 0 \text{ for all } \mathbf{x} \in \mathbb{R}^n.$$

If the inequality is strict we call f *strictly positive* or simply *positive*. ○

Immediate questions that arise concern the decision of nonnegativity for arbitrary polynomials and the certification, i.e., the possibility to certify nonnegativity efficiently. Such problems have countless applications for example in polynomial optimization, dynamical systems, control theory, robotics, computer vision, signal processing, and economics, e.g., [BPT13], [Las10].

Obviously, a necessary condition for a polynomial to be nonnegative is that its degree is even. Moreover, we can formulate nonnegativity conditions regarding the support of a polynomial. More precisely, a polynomial is nonnegative on the entire $\mathbb{R}^n$ only if the following necessary conditions are satisfied; see, e.g., [Rez78].

Proposition 2.2.2. *Let $A \subset \mathbb{N}^n$ be a finite set and $f \in \mathbb{R}[\mathbf{x}]$ be supported on A such that* $\text{New}(f) = \text{conv}(A)$. *Then f is nonnegative on $\mathbb{R}^n$ only if the following hold:*

(1) *All elements of $V(A)$ are even.*

(2) *If $\boldsymbol{\alpha} \in V(A)$, then the corresponding coefficient $f_{\boldsymbol{\alpha}}$ is strictly positive.*

In other words, if $\boldsymbol{\alpha} \in V(A)$, then the term $f_{\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}}$ has to be a monomial square.

The statement remains true for *real Laurent polynomials* $g \in \mathbb{R}[\mathbf{x}^{\pm 1}] = \mathbb{R}[x_1^{\pm 1}, \dots, x_n^{\pm 1}]$, since we can consider g as a polynomial f divided by a monomial square $\mathbf{x}^{\boldsymbol{\alpha}}$ for an even $\boldsymbol{\alpha}$; this is of relevance in Section 4.4.

Furthermore in some simple cases we can give direct nonnegativity characterizations and for univariate polynomials there exist several explicit algorithms to tackle the nonnegativity decision question, like via *Sturm sequences* or the *Hermite form* method; see, e.g., [BCR98]. But unfortunately, in the general multivariate case deciding polynomial nonnegativity is co-NP-hard whenever the degree is greater than or equal to four [BCSS98, MK87]. Therefore, one is interested in finding sufficient conditions to *certify* nonnegativity of polynomials, which can be checked efficiently. Clearly, a nonnegativity certificate is given by sums of squares, i.e., if we can write a polynomial as a sum of squares of polynomials, then it is apparent from this representation that it is nonnegative.

Definition 2.2.3. A polynomial $f \in \mathbb{R}[\mathbf{x}]_{2d}$ is a *sum of squares (SOS)* if there exist polynomials $f_1, \dots, f_k \in \mathbb{R}[\mathbf{x}]_d$ such that

$$f(\mathbf{x}) \;=\; \sum_{i=1}^{k} f_i^2(\mathbf{x}).$$

○

The property to be nonnegative respectively SOS is preserved under homogenization and dehomogenization. Note that this does not hold for the property of being positive, because it is possible for a strictly positive f to have zeros at infinity. Consider for example the strictly positive polynomial $f(x_1, x_2) = x_1^2 + (x_1x_2 - 1)^2$. Then $[0:1:0]$ and $[0:0:1]$ are zeros of the homogenization $\overline{f}(x_0, x_1, x_2) = x_1^2x_0^2 + (x_1x_2 - x_0^2)^2$.

Definition 2.2.4. We define the set of n-variate nonnegative polynomials and sums of squares with degree at most $2d$ as follows:

$$\begin{aligned} P_{n,2d} &= \{f \in \mathbb{R}[\mathbf{x}]_{n,2d} : f(\mathbf{x}) \geqslant 0 \text{ for all } \mathbf{x} \in \mathbb{R}^n\}, \\ \Sigma_{n,2d} &= \left\{ f \in P_{n,2d} : f(\mathbf{x}) = \sum_i f_i^2(\mathbf{x}) \text{ for some } f_i \in \mathbb{R}[\mathbf{x}]_{n,d} \right\}. \end{aligned}$$

○

We omit the index $2d$ if there is no bound on the degree, i.e., $\Sigma_{n,2d} = \Sigma_n \cap \mathbb{R}[\mathbf{x}]_{2d}$, similarly $P_{n,2d} = P_n \cap \mathbb{R}[\mathbf{x}]_{2d}$.
One can show, see, e.g., [Rez92]:

Proposition 2.2.5. *$P_{n,2d}$ and $\Sigma_{n,2d}$ are proper cones (i.e., closed, convex, pointed, and solid) in $\mathbb{R}[\mathbf{x}]_{n,2d} \sim \mathbb{R}^{N(n,2d)}$.*

The question of the relationship between the two cones $P_{n,2d}$ and $\Sigma_{n,2d}$ arises in a natural way and goes back to work of Hilbert at the end of the 19th century. Motivated by Minkowski's claim that there exist nonnegative polynomials that are not sums of squares, Hilbert intensively studied $P_{n,2d}$ and $\Sigma_{n,2d}$. In his seminal paper [Hil88] he finally classified all cases in which the two cones coincide:

Theorem 2.2.6 (Hilbert, 1888)**.** *Let $P_{n,2d}$ and $\Sigma_{n,2d}$ be as explained, then*
$P_{n,2d} = \Sigma_{n,2d}$ *if and only if* $n = 1$ *or* $d = 2$ *or* $(n, 2d) = (2, 4)$.

Proof. We only give a proof outline here. The first case, that every univariate nonnegative polynomial is a sum of squares, follows from the factorization theory. In fact, thereby it can be shown that every univariate nonnegative polynomial is a sum of two squares. For quadratic polynomials the argument follows easily from the diagonalization theorem. The third statement $P_{2,4} = \Sigma_{2,4}$ is non-trivial. Hilbert originally proved this statement for the homogeneous case of ternary quartics. Moreover, he showed that every nonnegative ternary quartic is a sum of three squares. His proof is based on the theory of algebraic curves.
For the "only if" part Hilbert described (homogeneously) a construction of forms which are nonnegative and not SOS for the two smallest cases where the two cones differ, namely for $(n+1, 2d) = (3, 6)$ and $(n+1, 2d) = (4, 4)$. From these two crucial cases all remaining cases can be easily deduced. For the construction, he used the fact that forms of degree d satisfy linear relations, known as the Cayley-Bacharach relations, which are not satisfied by forms of full degree $2d$. □

Hilbert's proof was nonconstructive and the first explicit example to verify Minkowski's claim was given by Motzkin in 1967 [Mot67] :

$$m(x_1, x_2) \;=\; 1 + x_1^4 x_2^2 + x_1^2 x_2^4 - 3x_1^2 x_2^2. \tag{2.2.1}$$

This binary sextic is nonnegative, which follows from the arithmetic-geometric mean inequality applied to the monomials $(1, x_1^4x_2^2, x_1^2x_2^4)$, but not a sum of squares. The non-existence of an SOS decomposition can be shown by assuming $m(\mathbf{x}) = \sum_i f_i^2(\mathbf{x})$, $\deg(f_i) \leqslant 3$. Now by inspecting monomials and comparing coefficients, we reach a contradiction. Thus, the Motzkin polynomial shows that $\Sigma_{2,6} \subsetneq P_{2,6}$. With exactly the same argument one can even show that $\lambda + m(\mathbf{x})$, for any real constant λ, is not a sum of squares in $\mathbb{R}[x_1, x_2]$.
After that, many other examples have been considered. For instance, for the case

$(n, 2d) = (3, 4)$ Choi and Lam provided the following polynomial:

$$q(x_1, x_2, x_3) \;=\; 1 + x_1^2 x_2^2 + x_2^2 x_3^2 + x_1^2 x_3^2 - 4x_1 x_2 x_3 \;\in P_{3,4} \backslash \Sigma_{3,4}.$$

Theorem 2.2.6 led Hilbert to asking the famous question, whether every nonnegative polynomial is a sum of squares of rational functions. This question is known as *Hilbert's 17th problem* and can be equivalently stated as whether there always exists a suitable multiplier for a nonnegative polynomial to be a finite sum of squares. Hilbert himself gave in 1893 [Hil93] an affirmative answer for the special case $n = 2$, and Artin provided in 1927 a solution to this problem in the general case, see [Art27].

Theorem 2.2.7 (Artin, 1927)**.** *Let $f \in P_{n,2d}$. Then there is a sum of squares multiplier $h \in \Sigma_{n,2d'}$, $h \neq 0$, such that $h \cdot f$ is a sum of squares.*

Example 2.2.8. For instance, multiplying the Motzkin polynomial with the square factor $h(x_1, x_2) = (x_1^2 + x_2^2)$ yields the following SOS decomposition

$$(x_1^2 + x_2^2) \cdot m(x_1, x_2) \;=\; x_2^2(1 - x_1^2)^2 + x_1^2(1 - x_2^2)^2 + x_1^2 x_2^2 (x_1^2 + x_2^2 - 2)^2.$$

$\bigcirc$

Artin's proof used the Artin-Schreier theory of ordered fields and was again nonconstructive. However, a constructive method for strictly positive polynomials was given by Habicht [Hab40] and is based on the following theorem by Pólya about forms:

Theorem 2.2.9. *Let $\overline{f}(x_0, \ldots, x_n)$ be a strictly positive form on $\mathbb{R}^{n+1}_{\geqslant 0} \backslash \{0, \mathbf{0}\}$, then $\overline{f}$ can be represented as $\overline{h} \cdot \overline{f} = \overline{g}$, where $\overline{h}$ and $\overline{g}$ are forms with positive coefficients. In particular, we can choose*

$$\overline{h} \;=\; (x_0 + x_1 + \cdots + x_n)^N,$$

for a suitable $N \in \mathbb{N}$.

Note that by homogeneity, the condition of strict positivity on $\mathbb{R}^{n+1}_{\geqslant 0} \backslash \{0, \mathbf{0}\}$ is equivalent to strict positivity on the unit simplex $\Delta^1_{n+1} = \{(x_0, \ldots, x_n) : x_i \geqslant 0, \sum_{i=0}^n x_i = 1\}$. Reznick [Rez95] generalized this statement by showing that for any strictly positive form $\overline{f}$ there exists a uniform denominator $\overline{h} = (x_0^2 + x_1^2 + \cdots + x_n^2)^N$ such that $\overline{h} \cdot \overline{f}$ is a sum of squares for N large enough. A lower estimate for N is provided by Powers and Reznick in [PR01].

2.2.1 Semidefinite Programming and Detecting Sums of Squares

In this section we address the main reason why sums of squares are not only an important nonnegativity certificate from theoretical point of view but also in practical applications. Namely, membership in $\Sigma_{n,2d}$ can be checked efficiently by semidefinite programming. Therefore, we first provide an overview of semidefinite programming, where we also introduce positive semidefinite matrices, and then we study the relation between sums of squares and semidefinite programming. For more details the reader may consult [Las10], [LV12], and [VB96].

We denote by $\mathcal{S}^n$ the set of real symmetric $n \times n$ matrices,

$$\mathcal{S}^n = \left\{ A \in \mathbb{R}^{n \times n} : A = A^T \right\},$$

which is a vector space with dimension $n(n+1)/2$. A real symmetric matrix A is *positive semidefinite (psd)*, if the quadratic form $\mathbf{x}^T A \mathbf{x} \geqslant 0$ for all $\mathbf{x} \in \mathbb{R}^n$ and similarly, A is *positive definite (pd)* if $\mathbf{x}^T A \mathbf{x} > 0$ for all $\mathbf{x} \in \mathbb{R}^n \backslash \{\mathbf{0}\}$. We use the shorthands $A \succeq 0$ resp. $A \succ 0$. The set of positive semidefinite matrices is denoted as $\mathcal{S}^n_+$, and its interior, the set of positive definite matrices, as $\mathcal{S}^n_{++}$.

The psd property has some equivalent conditions:

Proposition 2.2.10. *For $A \in \mathcal{S}^n$, the following statements are equivalent:*

(1) *The matrix A is positive semidefinite, i.e., $A \succeq 0$.*

(2) *Each eigenvalue of A is nonnegative.*

(3) *All $2^n - 1$ principal minors of A are nonnegative.*

(4) *There exists a factorization $A = LL^T$, where $L \in \mathbb{R}^{n \times r}$ and $r = \operatorname{rank}(A)$ (Cholesky decomposition).*

For the pd property there are similar characterizations:

Proposition 2.2.11. *For $A \in \mathcal{S}^n$, the following statements are equivalent:*

(1) *The matrix A is positive definite, i.e., $A \succ 0$.*

(2) *Each eigenvalue of A is strictly positive.*

(3) *All n leading principal minors of A are strictly positive.*

(4) *There exists a factorization $A = LL^T$, with $L \in \mathbb{R}^{n \times n}$ nonsingular.*

The set $\mathcal{S}_+^n$ forms a convex cone in $\mathcal{S}^n$, with $\text{int}(\mathcal{S}_+^n) = \mathcal{S}_{++}^n$. In fact, one can show that $\mathcal{S}_+^n$ is a proper cone and it is full-dimensional in $\mathcal{S}^n$.
The *standard scalar product* on the algebra of all $n \times n$ matrices $\mathbb{R}^{n \times n}$ is defined by

$$\langle A, B \rangle \ := \ \text{Tr}(A^T B) \ = \ \sum_{i,j=1}^{n} A_{ij} B_{ij},$$

where Tr(A) denotes the *trace* of A. Obviously, if $A, B \in \mathcal{S}^n$, then $\langle A, B \rangle = \text{Tr}(AB)$. Under this inner product, the cone $\mathcal{S}_+^n$ is *self-dual*, i.e., $(\mathcal{S}_+^n)^\vee = \mathcal{S}_+^n$:

Proposition 2.2.12. *A matrix $A \in \mathcal{S}^n$ is positive semidefinite if and only if $\langle A, B \rangle \geqslant 0$ holds for all $B \in \mathcal{S}_+^n$.*

A *semidefinite program (SDP)* is the problem of maximizing a linear function over the intersection of the cone of positive semidefinite matrices with an affine space. When restricting $\mathcal{S}_+^n$ to diagonal matrices in $\mathcal{S}^n$ we get $\mathbb{R}_{\geqslant 0}^n$. Thus, semidefinite programming generalizes linear programming, which is the problem of maximizing a linear function over an affine slice of the nonnegative orthant.
The (standard) primal form of a semidefinite program is

$$(2.2.2) \qquad p^* \ = \ \sup_{X \in \mathcal{S}^n} \{\langle C, X \rangle : \ \langle A_i, X \rangle = b_i \ \text{ for } \ i = 1, \ldots, m, \ X \succeq 0\},$$

where $C, A_i \in \mathcal{S}^n$ and $\mathbf{b} \in \mathbb{R}^m$ are monomial functions. The feasible set of an SDP is called a *spectrahedron* and is always a convex set. Hence, SDPs are convex optimization programs. In the special case where $C = 0$, the problem reduces to a *feasibility problem.* Note that the optimal value p^* might not be attained in the program (2.2.2). In general, $p^* \in \mathbb{R} \cup \{\pm\infty\}$, with $p^* = -\infty$ if problem (2.2.2) is infeasible and $p^* = +\infty$ might occur in which case we say the problem is unbounded. A very important feature of SDP problems is the associated duality theory. The dual semidefinite program reads:

$$(2.2.3) \qquad d^* \ = \ \inf_{\mathbf{y} \in \mathbb{R}^m} \left\{ \mathbf{b}^T \mathbf{y} : \ \sum_{i=1}^{m} y_i A_i - C \succeq 0 \right\},$$

where the positive semidefinite constraint $-C + y_1 A_1 + \cdots + y_m A_m \succeq 0$ is also named a *linear matrix inequality (LMI).* The spectrahedron can be parametrized by the LMI to

$$S \ = \ \{\mathbf{y} \in \mathbb{R}^m : -C + y_1 A_1 + \cdots + y_m A_m \succeq 0\},$$

with $C, A_i \in \mathcal{S}^n$.

Obviously, for a primal/dual pair of feasible solutions $(X, \mathbf{y})$ it holds $\langle C, X\rangle \leqslant \mathbf{b}^T\mathbf{y}$, therefore $p^* \leqslant d^*$, which is known as *weak duality.* The quantity $d^* - p^*$ is called the *duality gap*, and in contrast to linear programming there might be a positive duality gap. One crucial issue in duality theory is to identify sufficient conditions that ensure $p^* = d^*$, i.e., a zero duality gap, in which case one speaks of *strong duality.* Under specific constraint qualifications, SDP problems have strong duality, and thus zero duality gap. We say that the program (2.2.2) is *strictly feasible* if there exists a feasible $X \in \mathcal{S}^n$ with $X \succ 0$, analogously is (2.2.3) strictly feasible, if a feasible $\mathbf{y} \in \mathbb{R}^m$ fulfills $\sum_{i=1}^m y_i A_i - C \succ 0$.

Theorem 2.2.13 (Strong duality)**.** *If the primal program* (2.2.2) *is strictly feasible and its dual* (2.2.3) *is feasible, then* $p^* = d^*$ *and* (2.2.3) *attains its infimum.*
Analogously, if (2.2.3) *is strictly feasible and* (2.2.2) *is feasible, then* $p^* = d^*$ *and* (2.2.2) *attains its supremum.*

Even though the duality results for semidefinite programming are weaker than for linear programming, the key strength of SDP relies on the fact that one can also use interior point methods to find an approximate solution (to any given precision) in polynomially many iterations and their running time is efficient in practice for medium sized problems; see, e.g., [dK02]. There are many good software packages for semidefinite programming, see [BPT13] for an overview and the references therein for an in-depth treatment. Moreover, semidefinite programs provide a powerful tool for constructing convex relaxations for problems coming from combinatorial or polynomial optimization. Well known examples for applications in combinatorial optimization are the SDP approximation of the max-cut of a graph given by Goemans and Williamson [GW95] and Lovász's SDP relaxation on the Shannon capacity of a graph, see [Lov79]. For more general surveys we recommend [Lau08] and [Lov03]. SDP relaxations for polynomial optimization will be discussed in Section 2.3.

Now, we address the remarkable connection between semidefinite programs and sums of squares. This is given by the fact, that the problem of deciding whether a polynomial is a sum of squares can be reduced to a semidefinite feasibility problem. To be more specific, consider $f(\mathbf{x}) = \sum_{\boldsymbol{\alpha} \in A} f_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}} \in \mathbb{R}[\mathbf{x}]_{n,2d}$ and let $\mathbf{z}_d$ denote the vector of all monomials x_i with degree at most d, i.e., $\mathbf{z}_d = (1, x_1, x_2, \ldots, x_n, x_1x_2, \ldots, x_n^d)$. Those monomials form the canonical basis of $\mathbb{R}[\mathbf{x}]_{n,d}$. Notice that the length of the vector $\mathbf{z}_d$ is $\binom{n+d}{d} = N(n,d)$. Then we have the following relationship:

Theorem 2.2.14. *Let $f \in \mathbb{R}[\mathbf{x}]_{n,2d}$. Then f is a sum of squares if and only if there exists a symmetric matrix $Q \in \mathcal{S}^{N(n,d)}$ such that*

$$f(\mathbf{x}) \;=\; \mathbf{z}_d^T Q \mathbf{z}_d,\; Q \succeq 0. \tag{2.2.4}$$

Proof. Suppose f is a sum of squares $f(\mathbf{x}) \;=\; \sum_{i=1}^k f_i^2(\mathbf{x})$. Denoting the vector of coefficients of f_i by $\mathbf{f}_i$ yields

$$f(\mathbf{x}) \;=\; \sum_{i=1}^k \mathbf{z}_d^T \mathbf{f}_i \mathbf{f}_i^T \mathbf{z}_d \;=\; \mathbf{z}_d^T \left(\sum_{i=1}^k \mathbf{f}_i \mathbf{f}_i^T \right) \mathbf{z}_d \;=\; \mathbf{z}_d^T L L^T \mathbf{z}_d,$$

where L is the matrix with i-th column containing the coefficients $\mathbf{f}_i$. Having a Cholesky decomposition, the matrix $Q = LL^T$ is positive semidefinite and is of size $\binom{n+d}{d}$. Conversely, if (2.2.4) holds, then we can factorize the matrix $Q = LL^T$ with L of size $\binom{n+d}{d} \times \operatorname{rank}(Q)$ and obtain an SOS decomposition as given above. □

The matrix Q is often called the *Gram matrix*. By comparing the coefficients of the equation $f(\mathbf{x}) \;=\; \mathbf{z}_d^T Q \mathbf{z}_d$, we obtain

$$f_{\boldsymbol{\alpha}} = \sum_{\boldsymbol{\beta}+\boldsymbol{\gamma}=\boldsymbol{\alpha}} Q_{\boldsymbol{\beta},\boldsymbol{\gamma}}, \tag{2.2.5}$$

where $|\boldsymbol{\beta}|, |\boldsymbol{\gamma}| \leqslant d$. Obviously, this is a system of $\binom{n+2d}{2d}$ linear equations. Thus, the feasible set of (2.2.4) is the intersection of an affine subspace, given by the linear constraints, with the cone of psd matrices, which is a SDP problem. Hence, membership in $\Sigma_{n,2d}$ can be decided with semidefinite programming. Notice that (2.2.5) also yields:

Corollary 2.2.15. *The SOS cone $\Sigma_{n,2d}$ is a* projected spectrahedron *of dimension $\binom{n+2d}{2d}$.*

The Gram matrix Q is of size $\binom{n+d}{d} \times \binom{n+d}{d}$, which grows rapidly as the number of variables and the degree grow. However, the size for fixed d is polynomial in n and for fixed n polynomial in d. Sometimes, one can restrict to a smaller sized Gram matrix Q or monomial vector $\mathbf{z}$ depending on certain structures of f, like sparsity or symmetry. For example one can use a result by Reznick [Rez78, Theorem 1], which relates Newton polytopes to sums of squares:

Theorem 2.2.16. *If $f(\mathbf{x}) = \sum_i f_i^2(\mathbf{x})$, then* $\operatorname{New}(f_i) \subseteq \frac{1}{2}\operatorname{New}(f)$.

Hence, it suffices in (2.2.5) to consider monomials in $\frac{1}{2}\operatorname{New}(f)$, which reduces the size of the Gram matrix Q.

2.2.2 Quantitative Relationship between Σ and P

Having seen in Hilbert's Theorem 2.2.6 all cases where $P_{n,2d}$ and $\Sigma_{n,2d}$ coincide, entails the question about the size of the gap between the two cones, i.e., the set theoretic difference $P_{n,2d}\backslash\Sigma_{n,2d}$. Actually, the answer depends on whether we fix the number of variables and the degree or not. In what follows, we provide results and their conclusions for different assumptions.

We first give a result of Blekherman [Ble06] which is rather negative as it shows that if the degree is fixed and the number of variables grows, then the gap between nonnegative polynomials and sums of squares is unbounded. Namely, let $\overline{P}_{n+1,2d}$ resp. $\overline{\Sigma}_{n+1,2d}$ denote the cone of nonnegative forms resp. the cone of sums of squares of forms in $\mathbb{R}[x_0, \mathbf{x}]_{n+1,2d}$. In order to compare both cones, Blekherman's idea is to define subsets of finite volume by intersecting the cones with the following hyperplane H consisting of all forms with integral average one on the unit sphere $\mathbb{S}^n$ in $\mathbb{R}^{n+1}$:

$$H \;:=\; \left\{\overline{f} \in \mathbb{R}[x_0, \mathbf{x}]_{n+1,2d} : \int_{\mathbb{S}^n} \overline{f}\, d\sigma = 1\right\},$$

where σ is the rotation invariant probability measure on $\mathbb{S}^n$. Now define the compact sections of $\overline{P}_{n+1,2d}$ and $\overline{\Sigma}_{n+1,2d}$ with H as

$$\hat{P}_{n+1,2d} \;:=\; \overline{P}_{n+1,2d} \cap H \quad\text{and}\quad \hat{\Sigma}_{n+1,2d} \;:=\; \overline{\Sigma}_{n+1,2d} \cap H.$$

The dimension of the ambient space of these sections is $\binom{n+2d}{2d} - 1 =: D$.

Theorem 2.2.17 ([Ble06])**.** *There exist constants $c_1(d)$ and $c_2(d)$ both depending only on d such that for $n+1$ large enough*

$$c_1(d)(n+1)^{(d-1)/2} \leqslant \left(\frac{\operatorname{Vol}\hat{P}_{n+1,2d}}{\operatorname{Vol}\hat{\Sigma}_{n+1,2d}}\right)^{\frac{1}{D}} \leqslant c_2(d)(n+1)^{(d-1)/2}.$$

Thus, if the degree is fixed and at least 4 and the number of variables grows, then there are significantly more nonnegative polynomials than sums of squares. We point out that for a small number of variables the distinction between the two cones is quite delicate, and it is not known at what point $P_{n,2d}$ becomes much larger than $\Sigma_{n,2d}$.

However, there are also positive results showing that we can perturb every nonnegative polynomial f to make it a sum of squares. First, Berg [Ber87] showed existentially that for a fixed number of variables and variable degrees the cone $\Sigma_{n,2d}$ is dense in $P_{n,2d}$ on $[-1,1]^n$ with respect to the l_1-norm $||f||_1 = \sum_{\boldsymbol{\alpha}\in\mathbb{N}^n} |f_{\boldsymbol{\alpha}}|$. We state the result by Lasserre and Netzer [LN07] which provides an explicit SOS approximation:

Theorem 2.2.18. *Let $f \in \mathbb{R}[\mathbf{x}]$ be a polynomial nonnegative on $[-1,1]^n$. For any $\varepsilon > 0$, there exists a nonnegative integer $t_0 \in \mathbb{N}$ such that the polynomial*

$$f + \varepsilon \left(1 + \sum_{i=1}^{n} x_i^{2t}\right),$$

is a sum of squares for all $t \geqslant t_0$.

Previously, Lasserre [Las06b] had given an analogous result for polynomials nonnegative on the whole space $\mathbb{R}^n$.

Theorem 2.2.19. *Let $f \in \mathbb{R}[\mathbf{x}]$ be a polynomial nonnegative on $\mathbb{R}^n$.*

(1) *There exist some $r_0 \in \mathbb{N}, \lambda_0 \geqslant 0$ such that for all $r \geqslant r_0$ and $\lambda \geqslant \lambda_0$ the polynomial*

$$f + \lambda \sum_{k=0}^{r} \sum_{i=1}^{n} \frac{x_i^{2k}}{k!}$$

is a sum of squares.

(2) *For every $\varepsilon > 0$, there exists some $r_\varepsilon \in \mathbb{N}$ such that*

$$f_\varepsilon := f + \varepsilon \sum_{k=0}^{r_\varepsilon} \sum_{i=1}^{n} \frac{x_i^{2k}}{k!}$$

is a sum of squares. Hence, $||f - f_\varepsilon||_1 \to 0$ as $\varepsilon \downarrow 0$.

Notice that statement (2) of the above Theorem 2.2.19 provides an explicit converging sequence $\{f_\varepsilon\}$. However, finding explicit bounds for r_0, λ_0, and r_ε is still an open problem.

2.2.3 Dual Cones, Boundary, and Facial Structure of Σ and P

Studying convex geometric structures such as the boundary, the facial structure, and the dual cones of $P_{n,2d}$ and $\Sigma_{n,2d}$ is an active area of research in convex algebraic geometry with many pending issues. In this section we give a brief overview of basic results of those properties, see [BPT13] and the references stated in this section. As most of the results are given for forms, we will consider the cones $\overline{P}_{n+1,2d}$ and $\overline{\Sigma}_{n+1,2d}$ in this section.

We begin with describing the conceptually simple *dual cone* $\overline{P}^{\vee}_{n+1,2d}$. Consider the dual space $\mathbb{R}[x_0,\mathbf{x}]^{\vee}_{n+1,2d}$ of linear functionals on $\mathbb{R}[x_0,\mathbf{x}]_{n+1,2d}$. For $(v_0,\mathbf{v}) \in \mathbb{R}^{n+1}$, let $l_{(v_0,\mathbf{v})} \in \mathbb{R}[x_0,\mathbf{x}]^{\vee}_{n+1,2d}$ be the linear functional given by evaluation at $(v_0,\mathbf{v})$, i.e., $l_{(v_0,\mathbf{v})}(\overline{f}) = \overline{f}(v_0,\mathbf{v})$, for $\overline{f} \in \mathbb{R}[x_0,\mathbf{x}]_{n+1,2d}$. Note that a form is globally nonnegative on $\mathbb{R}^{n+1}$ if and only if it is nonnegative on the unit sphere $\mathbb{S}^n$. Thus, it follows easily:

Proposition 2.2.20. *The dual cone $\overline{P}^{\vee}_{n+1,2d}$ is the conical hull of functionals $l_{(v_0,\mathbf{v})}$ with $(v_0,\mathbf{v}) \in \mathbb{S}^n$:*

$$\overline{P}^{\vee}_{n+1,2d} \;=\; \operatorname{cone}\left(l_{(v_0,\mathbf{v})} : (v_0,\mathbf{v}) \in \mathbb{S}^n\right).$$

In fact, the cone $\overline{P}^{\vee}_{n+1,2d}$ is the conical hull of the $2d$*-th Veronese variety*. Moreover, one can show that the functional $l_{(v_0,\mathbf{v})}$ spans an *extreme ray* of $\overline{P}^{\vee}_{n+1,2d}$ for all $(v_0,\mathbf{v}) \in \mathbb{S}^n$, and those functionals even form the complete set of extreme rays of $\overline{P}^{\vee}_{n+1,2d}$.
In contrast to $\overline{P}^{\vee}_{n+1,2d}$, describing the extreme rays of $\overline{\Sigma}^{\vee}_{n+1,2d}$ is significantly more complicated. Therefore we restrict ourselves to giving only the description of the dual cone of $\overline{\Sigma}_{n+1,2d}$ here. We associate to every linear functional $l \in \mathbb{R}[x_0,\mathbf{x}]^{\vee}_{2d}$ a quadratic form Q_l defined on $\mathbb{R}[x_0,\mathbf{x}]_d$ by $Q_l(\overline{f}) = l(\overline{f}^2)$ for all $\overline{f} \in \mathbb{R}[x_0,\mathbf{x}]_d$. Then we can view $\mathbb{R}[x_0,\mathbf{x}]^{\vee}_{2d}$ as a subspace of $S^{n+1,d}$, the vector space of real quadratic forms on $\mathbb{R}[x_0,\mathbf{x}]_d$, by identifying l with Q_l. Hence, the cone of positive semidefinite forms in $S^{n+1,d}$ is defined as

$$S^{n+1,d}_{+} \;:=\; \left\{Q \in S^{n+1,d} : Q(\overline{f}) \geqslant 0 \text{ for all } \overline{f} \in \mathbb{R}[x_0,\mathbf{x}]_d\right\}.$$

Proposition 2.2.21. *The dual cone $\overline{\Sigma}^{\vee}_{n+1,2d}$ is given by the intersection of the cone of psd matrices $S^{n+1,d}_{+}$ with the subspace $\mathbb{R}[x_0,\mathbf{x}]^{\vee}_{2d}$, i.e.,*

$$\overline{\Sigma}^{\vee}_{n+1,2d} \;=\; S^{n+1,d}_{+} \cap \mathbb{R}[x_0,\mathbf{x}]^{\vee}_{2d}.$$

Observe that on the basis of the above description the cone $\overline{\Sigma}^{\vee}_{n+1,2d}$ is a spectrahedron. Recall that $\overline{\Sigma}_{n+1,2d}$ is only a projected spectrahedron. By explicitly choosing the monomial basis of $\mathbb{R}[x_0,\mathbf{x}]_{n+1,2d}$, we get the viewpoint of Section 2.2.1. An important interpretation of the dual cones in terms of moment sequences of probability measures will be discussed in Section 2.3.2.

The *boundaries* of $\overline{P}_{n+1,2d}$ and $\overline{\Sigma}_{n+1,2d}$ are hypersurfaces in $\mathbb{R}[x_0, \mathbf{x}]_{2d}$. First of all note that the interior of $\overline{P}_{n+1,2d}$ consists of all strictly positive forms and the boundary of $\overline{P}_{n+1,2d}$ consists of forms with a nontrivial zero. Now, we are interested in the *algebraic boundary* of the cones $\overline{P}_{n+1,2d}$ and $\overline{\Sigma}_{n+1,2d}$, i.e., the *Zariski closure* of the boundary hypersurfaces. The algebraic boundary of $\overline{P}_{n+1,2d}$ is extensively studied in [Nie12]. Nie shows therein that the algebraic boundary of $\overline{P}_{n+1,2d}$ lies on the irreducible hypersurface defined by the discriminant. Again the algebraic boundary of $\overline{\Sigma}_{n+1,2d}$ is significantly more complicated. So far, there exist only partial results. For instance, it is known that the algebraic boundary of $\overline{P}_{n+1,2d}$ is contained in the algebraic boundary of $\overline{\Sigma}_{n+1,2d}$, as the discriminant is a component of the algebraic boundary of $\overline{\Sigma}_{n+1,2d}$. Furthermore, in [BHO+12] it is shown that the algebraic boundary of $\overline{\Sigma}_{3,6}$ respectively $\overline{\Sigma}_{4,4}$ has a unique non-discriminant component. It has degree 83200 resp. 38475 and it is the Zariski closure of the sextics that are sums of three squares of cubics, resp. the quartics that are sums of four squares of quadrics.

Another open problem is to analyze the *facial structures* of $\overline{P}_{n+1,2d}$ and $\overline{\Sigma}_{n+1,2d}$ as well as the possible *dimensions* of their faces. In the cases $n+1 = 2$, $2d = 2$, and, to some extent, the case $(n+1, 2d) = (3, 4)$, thus, the Hilbert cases where $\overline{P}_{n+1,2d} = \overline{\Sigma}_{n+1,2d}$, this problem is relatively well understood, see [Bar02]. But generally, only partial results are known. In particular, the facial analysis in the case where the two cones differ is important for a better understanding of the gap between $\overline{P}_{n+1,2d}$ and $\overline{\Sigma}_{n+1,2d}$. Recently, Blekherman [Ble12] provided a geometric construction for the faces of the SOS cone that are not faces of $\overline{P}_{n+1,2d}$, and in [BIK15] especially the dimensions of the exposed faces are investigated, this will be further discussed in Section 3.3. Since every form in $\overline{P}_{n+1,2d}$ is a finite sum of extremal forms, the cone $\overline{P}_{n+1,2d}$ is completely determined when all its extremal forms are known. Therefore, the study of $\mathcal{E}\overline{P}_{n+1,2d}$ and $\mathcal{E}\overline{\Sigma}_{n+1,2d}$ is extremely important and a wealth of research is dedicated to it, see, e.g., [CL78], [CKLR82], and [Rez78]. For $\overline{\Sigma}_{n+1,2d}$ there exists the necessary condition that all extremal forms are perfect squares. But not every perfect square is extremal: $(x_1^2 + x_2^2)^2 = (x_1^2 - x_2^2)^2 + (2x_1x_2)^2$. We point out that the product of two extremal forms need not be extremal, see [CL77]. In [CKLR82] the authors give a complete classification of the cases when the inclusion $\mathcal{E}\overline{\Sigma}_{n+1,2d} \subset \mathcal{E}\overline{P}_{n+1,2d}$ holds:

Theorem 2.2.22. *The inclusion $\mathcal{E}\overline{\Sigma}_{n+1,2d} \subset \mathcal{E}\overline{P}_{n+1,2d}$ holds precisely in the following cases:*

(1) $n + 1 = 2$,

(2) $2d \leqslant 6$,

(3) $(n+1, 2d) = (3, 8)$,

(4) $(n+1, 2d) = (3, 10)$.

Exposed extreme rays of $\overline{P}_{n+1,2d}$ are conceptually simple, since a nonnegative form $\overline{f} \in \overline{P}_{n+1,2d}$ is an exposed extreme ray if and only if the variety given by $\overline{f}$ is maximal among all varieties defined by nonnegative forms:

Proposition 2.2.23. *A form $\overline{f} \in \overline{P}_{n+1,2d}$ is an exposed extreme ray of $\overline{P}_{n+1,2d}$ if and only if for all $\overline{p} \in \overline{P}_{n+1,2d}$ with $\mathcal{V}(\overline{f}) \subseteq \mathcal{V}(\overline{p})$ it follows that $\overline{p} = \lambda \overline{f}$ for some $\lambda \in \mathbb{R}$.*

In light of Proposition 2.2.23 and the fact that the boundary $\overline{P}_{n+1,2d}$ consists of forms with a nontrivial zero, the study of the facial structure is closely related to the study of the real zeros of the cones. See also the following result by Harris [Har99] in this context:

Lemma 2.2.24. *If $\overline{f} \in \overline{P}_{n+1,2d}$ is extremal, then $\mathcal{V}(\overline{f}) \neq \varnothing$.*

Hence, the investigation of the real zeros of $\overline{P}_{n+1,2d}$ and $\overline{\Sigma}_{n+1,2d}$ enqueues the active research in convex algebraic geometry, see, e.g., [CL77, CKLR82, Qua15, Rez00, XY17], and also Section 3.2. In [CLR80] the authors investigate the real zeros of nonnegative forms and provide various consequential results. The main result concerns the set theoretic difference between $\overline{P}_{n+1,2d}$ and $\overline{\Sigma}_{n+1,2d}$ in the minimal cases:

Theorem 2.2.25.

(1) *If $\overline{f} \in \overline{P}_{3,6}$ and $|\mathcal{V}(\overline{f})| > 10$, then $\overline{f} \in \overline{\Sigma}_{3,6}$. In fact, $\overline{f}$ is a sum of three squares of cubics.*

(2) *If $\overline{f} \in \overline{P}_{4,4}$ and $|\mathcal{V}(\overline{f})| > 11$, then $\overline{f} \in \overline{\Sigma}_{4,4}$. In fact, $\overline{f}$ is a sum of six squares of quadratics.*

Both cases require $|\mathcal{V}(\overline{f})| = \infty$.

Statement (2) has been slightly improved in [BHO+12] to forms $\overline{f}$ with $|\mathcal{V}(\overline{f})| > 10$. We close this section by defining certain numbers given by the maximal number of real zeros of forms in the SOS and the nonnegativity cone, which were as well studied by Choi, Lam, and Reznick in [CKLR82]. Let

$$B_{n+1,2d} := \sup_{\substack{\overline{f} \in \overline{P}_{n+1,2d} \\ |\mathcal{V}(\overline{f})| < \infty}} |\mathcal{V}(\overline{f})| \quad \text{and} \quad B'_{n+1,2d} := \sup_{\substack{\overline{f} \in \overline{\Sigma}_{n+1,2d} \\ |\mathcal{V}(\overline{f})| < \infty}} |\mathcal{V}(\overline{f})|.$$

To determine these numbers exactly is a rather difficult task. In what follows, we list some known results in few special cases.

Theorem 2.2.26 ([BHO+12, CLR80, Sha77])**.** *Let $B_{n+1,2d}$ and $B'_{n+1,2d}$ be as defined above, then:*

(1) $B_{2,2d} = B'_{2,2d} = d$ *and* $B_{n+1,2} = B'_{n+1,2} = 1$.

(2) $B_{3,4} = 4$ *and* $B_{3,6} = B_{4,4} = 10$.

(3) $d^2 \leqslant B_{3,2d} \leqslant \frac{3d(d-1)}{2} + 1$ *for* $2d \geqslant 6$, *and*

(4) $B_{3,6k} \geqslant 10k^2$, $B_{3,6k+2} \geqslant 10k^2 + 1$, $B_{3,6k+4} \geqslant 10k^2 + 4$.

(5) *Let* $\beta(2d) = \frac{B_{3,2d}}{4d^2}$. *Then* $\beta = \lim_{2d \to \infty} \beta(2d)$ *exists. Moreover,* $\beta(2d) \leqslant \beta$ *for all* $\frac{5}{18} \leqslant \beta \leqslant \frac{1}{2}$.

(6) $B'_{n+1,2d} \geqslant d^n$.

(7) $B'_{3,2d} = \frac{(2d)^2}{4} = d^2$.

Theorem 2.2.26 shows that $B_{3,2d}$ is always finite. But already for $B_{n+1,4}$ with $n+1 \geqslant 5$, we do not know if the number needs to be finite in general.

2.3 Polynomial Optimization and Real Algebraic Geometry

Building upon Section 2.2, we now discuss one of the most important applications of nonnegative polynomials and sums of squares, namely its application to polynomial optimization. The objective of polynomial optimization is to minimize a real polynomial f over some set K,

$$f_K^* \quad = \quad \inf\{f(\mathbf{x}) : \mathbf{x} \in K\}.$$

In the case of global respectively unconstrained optimization, we have $K = \mathbb{R}^n$, and for constrained optimization K is the semialgebraic set defined by a finite number of real polynomials. The special case that all involved polynomials in this problem have degree one boils down to a linear program. Obviously, the general problem is equivalent to finding the largest real number γ such that $f - \gamma$ is nonnegative on K:

$$f_K^* \quad = \quad \sup\{\gamma \in \mathbb{R} : f(\mathbf{x}) - \gamma \geqslant 0 \text{ for all } \mathbf{x} \in K\}.$$

By means of this observation we may see the relation between nonnegative polynomials and polynomial optimization. Thus, a polynomial optimization problem can be reduced to our main problem of deciding nonnegativity of polynomials. Since this is co-NP-hard in general, see Section 2.2, a natural idea is to replace the hard nonnegativity condition with a more tractable condition, i.e., to relax the problem. This section is dedicated to giving relaxation methods for this problem based on relaxing nonnegativity over K by sums of squares decompositions, and the dual theory of moments. Since SOS decompositions can be formulated as a semidefinite programming problem, the approach via SOS relaxations leads to efficiently computable approximations for the initial problem. There is a vast literature devoted to this topic, see [Las01], [Las08], [Par00], [Par03], [PS03], and [Sho87a]. In what follows, we first discuss the relaxation idea in the particular case of global optimization. Then we study nonnegative polynomials from the dual perspective, namely with the theory of moments, which allows us also to formulate a dual polynomial optimization problem, the moment problem. As a first step towards constrained optimization, we provide several results from real algebraic geometry concerning representation theorems of polynomials which are nonnegative respectively positive on a given set. Finally, we analyze the general constrained optimization problem and its common approach via Lasserre's relaxation.

2.3.1 Global Optimization and Sums of Squares Relaxations

Given a real polynomial $f \in \mathbb{R}[\mathbf{x}]$, the *global polynomial optimization problem (POP)* for f is the problem of minimizing f over the full space $\mathbb{R}^n$:

$$f^* \quad = \quad \inf\{f(\mathbf{x}) \ : \ \mathbf{x} \in \mathbb{R}^n\}.$$

The *optimal value* of this program will be denoted by f^*. Clearly, this problem is equivalent to determining the real number

$$f^* \quad = \quad \sup\{\gamma \in \mathbb{R} \ : \ f(\mathbf{x}) - \gamma \geqslant 0 \text{ for all } \mathbf{x} \in \mathbb{R}^n\}.$$

Since this problem is hard in general, one can find a lower bound for f^* by relaxing the nonnegativity condition in the above problem to finding the real number

$$(2.3.1) \qquad f_{\mathrm{sos}} \quad = \quad \sup\left\{\gamma \in \mathbb{R} \ : \ f - \gamma = \sum_{i=1}^{k} q_i^2 \ \text{ for some } \ q_i \in \mathbb{R}[\mathbf{x}]\right\}.$$

The bound f_{sos} for the optimal SOS decomposition of f can be determined by semidefinite programming, see Section 2.2.1. Since every sum of squares is nonnegative, the SOS relaxation yields a lower bound for the optimal value of (POP), i.e., $f_{\text{sos}} \leqslant f^*$. The subsequent question is when the SOS relaxation provides the exact number, in which this case the relaxation is said to be *exact.* This can be answered as follows, see [Las01]:

Theorem 2.3.1. *For $f \in \mathbb{R}[\mathbf{x}]$ it holds $f_{\text{sos}} = f^*$ if and only if the polynomial $f(\mathbf{x}) - f^*$ is a sum of squares.*

Remember the Motzkin polynomial $m(\mathbf{x}) = 1 + x_1^4 x_2^2 + x_1^2 x_2^4 - 3x_1^2 x_2^2$, cf. (2.2.1). It holds $m^* = 0$ with zeros attained at $(\pm 1, \pm 1)$. But we noted that $m - \lambda$ for any real λ is not a sum of squares, leading to $m_{\text{sos}} = -\infty$.

2.3.2 Duality and the Moment Problem

A field closely interlinked with the theory of nonnegative polynomials and sums of squares decompositions is the problem of moments. In this section we provide a brief overview of the theory of moments, the moment problem, the relation to polynomial optimization, and the connection to the duals of $P_{n,2d}$ and $\Sigma_{n,2d}$.

Let μ be a measure on $\mathbb{R}^n$. For $\boldsymbol{\alpha} \in \mathbb{N}^n$, the quantity $y_{\boldsymbol{\alpha}} := \int \mathbf{x}^{\boldsymbol{\alpha}} d\mu$ is called the *moment* of order $\boldsymbol{\alpha}$ of the measure μ. Then, the sequence $y = (y_{\boldsymbol{\alpha}})_{\boldsymbol{\alpha} \in \mathbb{N}^n}$ is the *sequence of moments*, and for a given $t \in \mathbb{N}$, the truncated sequence $(y_{\boldsymbol{\alpha}})_{\boldsymbol{\alpha} \in \mathbb{N}_t^n}$ is the sequence of moments up to order t, where $\mathbb{N}_t^n := \{\boldsymbol{\alpha} \in \mathbb{N}^n : |\boldsymbol{\alpha}| \leqslant t\}$. When y is the sequence of moments of a measure, we refer to μ as a *representing measure* for the sequence y.
A basic problem in the theory of moments concerns the characterization of (infinite or truncated) moment sequences, i.e., the characterization of those sequences $y = (y_{\boldsymbol{\alpha}})_{\boldsymbol{\alpha}}$ that are the sequences of moments of some measure μ. Given a sequence y, the *K-moment problem* asks for the existence of a representing (Borel) measure μ supported on $K \subseteq \mathbb{R}^n$, i.e., a measure μ with $y_{\boldsymbol{\alpha}} := \int_K \mathbf{x}^{\boldsymbol{\alpha}} d\mu$. The case $K = \mathbb{R}^n$ yields the basic *moment problem.* Solving this problem is in general a hard task and is related to polynomial optimization.
Given a sequence $(y_{\boldsymbol{\alpha}})_{\boldsymbol{\alpha} \in \mathbb{N}^n}$, its *moment matrix* is defined by the (infinite) matrix $M(y) := (y_{\boldsymbol{\alpha}+\boldsymbol{\beta}})_{\boldsymbol{\alpha},\boldsymbol{\beta}}$, for $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{N}^n$. Similarly, for $t \geqslant 1$ and $(y_{\boldsymbol{\alpha}})_{\boldsymbol{\alpha} \in \mathbb{N}_{2t}^n}$ the *truncated moment matrix* is given by $M_t(y) := (y_{\boldsymbol{\alpha}+\boldsymbol{\beta}})_{\boldsymbol{\alpha},\boldsymbol{\beta}}$, for $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{N}_t^n$. Moreover, shifting a sequence $(y_{\boldsymbol{\alpha}})_{\boldsymbol{\alpha} \in \mathbb{N}^n}$ by a polynomial $g \in \mathbb{R}[\mathbf{x}]$ leads to a new sequence defined as $g * y := M(y)g$, with $\boldsymbol{\alpha}$-th entry $(g * y)_{\boldsymbol{\alpha}} = \sum_{\boldsymbol{\beta}} g_{\boldsymbol{\beta}} y_{\boldsymbol{\alpha}+\boldsymbol{\beta}}$ for all $\boldsymbol{\alpha} \in \mathbb{N}^n$. The moment

matrices of $(g * y)$ are often called *localizing matrices.*
To a sequence y corresponds a *linear functional* $l \in \mathbb{R}[\mathbf{x}]^\vee$, defined as

$$\begin{aligned} l\colon \mathbb{R}[\mathbf{x}] &\to \mathbb{R} \\ \mathbf{x}^{\boldsymbol{\alpha}} &\mapsto y_{\boldsymbol{\alpha}} \\ f = \sum_{\boldsymbol{\alpha}} f_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}} &\mapsto l(f) = \sum_{\boldsymbol{\alpha}} f_{\boldsymbol{\alpha}} y_{\boldsymbol{\alpha}}. \end{aligned}$$

Observe that in the case of y being a sequence of vectors $\mathbf{v} \in \mathbb{R}^n$, the linear functional l is the evaluation at $\mathbf{v}$ (denoted by $l_{\mathbf{v}}$), i.e., $l_{\mathbf{v}}(f) = f(\mathbf{v})$, see Section 2.2.3.
Before proceeding, we give a fundamental result from Haviland [Hav36], which shows the relation of the moment problem and our main problem of deciding nonnegativity.

Theorem 2.3.2 (Haviland)**.** *Let $y = (y_{\boldsymbol{\alpha}})_{\boldsymbol{\alpha} \in \mathbb{N}^n}$ be a sequence and $l \in \mathbb{R}[\mathbf{x}]^\vee$ be the corresponding linear functional. Then the following statements are equivalent:*

(1) *There exists a Borel measure μ on $\mathbb{R}^n$ such that $y_{\boldsymbol{\alpha}} = \int \mathbf{x}^{\boldsymbol{\alpha}} d\mu$ for all $\boldsymbol{\alpha} \in \mathbb{N}^n$.*

(2) *$l(f) \geqslant 0$ for all nonnegative polynomials f on $\mathbb{R}^n$.*

Next, given a linear form l to a sequence y we consider its associate *bilinear form* $\mathcal{L}$ by

$$\begin{aligned} \mathcal{L}\colon \mathbb{R}[\mathbf{x}] \times \mathbb{R}[\mathbf{x}] &\to \mathbb{R} \\ (f, g) &\mapsto l(f \cdot g) = \mathbf{f}^T M(y) \mathbf{g}, \end{aligned}$$

where $\mathbf{f}$ denotes the coefficient vector of f, similarly $\mathbf{g}$. Obviously, the moment matrix $M(y)$ is the representation matrix of $\mathcal{L}$. Observe that the related quadratic form $\mathcal{L}(f, f) = l(f^2) = \mathbf{f}^T M(y) \mathbf{f}$ is psd for all $f \in \mathbb{R}[\mathbf{x}]$, i.e., $l(f^2) \geqslant 0$ if and only if the moment matrix $M(y)$ is psd, i.e., $M(y) \succeq 0$. Hence, we can deduce an easy necessary condition for moment sequences, namely if y is a truncated sequence of moments up to order $2t$ of a measure μ, then $M_t(y) \succeq 0$. This observation results from the following calculation for $f \in \mathbb{R}[\mathbf{x}]$: $\mathcal{L}(f, f) = l(f^2) = \int f(\mathbf{x})^2 d\mu \geqslant 0$, since every $f(\mathbf{x})^2 \geqslant 0$ for all $\mathbf{x} \in \mathbb{R}^n$.

Based on this fundamental knowledge, we study the moment relations for polynomial optimization problems. Here, we restrict to the unconstrained case, but note that the provided approach can be generalized to the constrained case, as we will see for

Lasserre's relaxation in Section 2.3.4. For (POP) it is

$$f^* = \inf_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}) = \inf_{\mu} \int_{\mathbb{R}^n} f(\mathbf{x})d\mu,$$

where the second infimum is taken over all probability measures μ on $\mathbb{R}^n$ supported on $\mathbb{R}^n$. Recall that for a given polynomial $f = \sum_{\boldsymbol{\alpha}} f_{\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}} \in \mathbb{R}[\mathbf{x}]_{n,2d}$ the SOS relaxation is

$$f_{\mathrm{sos}} = \sup\left\{\gamma \in \mathbb{R} \ : \ f - \gamma \in \Sigma_{n,2d}\right\}.$$

The *moment relaxation* is given by

$$\begin{aligned} f_{\mathrm{mom}} &= \inf_{y} \left\{ \sum_{\boldsymbol{\alpha}} f_{\boldsymbol{\alpha}} y_{\boldsymbol{\alpha}} \ : \ y_{\mathbf{0}} = 1, M_d(y) \succeq 0 \right\} \\ &= \inf_{l} \left\{ l(f) \ : \ l(1) = 1,\, l(g^2) \geqslant 0 \text{ for all } g \in \mathbb{R}[\mathbf{x}]_{n,d} \right\}. \end{aligned}$$

Using the truncated moment matrix in this relaxation has its reason in the computation. Involving only the infinite moment matrix, it would be unclear how to compute f_{mom}. However, due to the degree bound we obtain a (finite-dimensional) semidefinite program. The SDPs for f_{sos} and f_{mom} can be viewed as dual programs, and actually there is no duality gap:

Theorem 2.3.3. *Given $f \in \mathbb{R}[\mathbf{x}]_{2d}$, we have $f_{\mathrm{sos}} = f_{\mathrm{mom}}$. Additionally, if $f_{\mathrm{mom}} > -\infty$ then the SOS relaxation f_{sos} has an optimal solution.*

Moreover, if $f - f^*$ is a sum of squares, then the subsequent theorem explains how to extract a minimizer for (POP).

Theorem 2.3.4. *Let $f \in \mathbb{R}[\mathbf{x}]_{2d}$ with global minimum f^*. If the nonnegative polynomial $f - f^*$ is SOS, then $f^* = f_{\mathrm{sos}} = f_{\mathrm{mom}}$, and if x^* is a minimizer for f on $\mathbb{R}^n$, then the moment vector $y^* = (x^*_{\boldsymbol{\alpha}})_{\boldsymbol{\alpha}\in\mathbb{N}^n_{2d}}$ is a minimizer of the moment relaxation.*

We now return to the question of exactness of a relaxation already mentioned in Section 2.3.1. Using duality theory there are some important results which tackle this question. We present one based on the so called *flat extension* of moment matrices:

Theorem 2.3.5. *Let $f \in \mathbb{R}[\mathbf{x}]_{2d}$, and suppose that the optimal value f_{mom} of the moment relaxation is attained at some optimal solution y^*. If it holds that* $\mathrm{rank}(M_{d-1}(y^*)) = \mathrm{rank}(M_d(y^*))$*, then $f^* = f_{\mathrm{mom}}$ and there are at least* $\mathrm{rank}(M_d(y^*))$ *global minimizers.*

We conclude this section by sketching the duality between the cone P of nonnegative polynomials respectively the cone Σ of sums of squares and moment sequences. For this, consider the cones $\mathcal{M} := \{y = (y_{\boldsymbol{\alpha}})_{\boldsymbol{\alpha}\in\mathbb{N}^n} : y \text{ has a representing measure}\}$ and $\mathcal{M}_+ := \{y = (y_{\boldsymbol{\alpha}})_{\boldsymbol{\alpha}\in\mathbb{N}^n} : M(y) \succeq 0\}$. It holds:

Theorem 2.3.6. *The cones P and $\mathcal{M}$ respectively Σ and $\mathcal{M}_+$ are duals of each other.*

2.3.3 Positivstellensätze

We begin this section by establishing the relationship between classical algebraic geometry and real algebraic geometry as well as certain basic results from the latter. Then we introduce several Positivstellensätze, which provide the essential tool to attack constrained optimization problems. We refer to [Las07], [Lau09], and [Mar08] for an overview including different representation theorems.

While classical algebraic geometry deals with zeros of polynomial equations as subsets of algebraically closed fields, like $\mathbb{C}^n$, real algebraic geometry deals with subsets of $\mathbb{R}^n$ defined by polynomial (in-)equalities. These subsets are called *semialgebraic sets*, which are finite unions of basic sets. Given a semialgebraic set, a fundamental problem in this context is to decide whether it is empty or not. In real algebraic geometry one often considers a *basic closed semialgebraic set* $K \subseteq \mathbb{R}^n$, which can be described as the set of solutions of the form

$$K = \{\mathbf{x} \in \mathbb{R}^n : g_1(\mathbf{x}) \geqslant 0, \ldots, g_s(\mathbf{x}) \geqslant 0\}, \tag{2.3.2}$$

where $g_1, \ldots, g_s \in \mathbb{R}[\mathbf{x}]$.

In both classical and real algebraic geometry one deals with "Stellensätzen". In the classical setting, these are the so called *Nullstellensätze*. The first one goes back to Hilbert:

Theorem 2.3.7 (Hilbert's Nullstellensatz)**.** *Given polynomials $f, g_1, \ldots, g_s \in \mathbb{R}[\mathbf{x}]$. Denote by $I = \langle g_1, \ldots, g_s\rangle$ the ideal generated by g_i, then we have that*

$$f = 0 \text{ on } \mathcal{V}_{\mathbb{C}}(I) \quad \text{if and only if} \quad f^N = \sum_{i=1}^{s} p_i g_i,$$

for some $p_i \in \mathbb{R}[\mathbf{x}]$, $N \in \mathbb{N}^*$, and $\mathcal{V}_{\mathbb{C}}(I) = \{\mathbf{x} \in \mathbb{C}^n : g_1(\mathbf{x}) = 0, \ldots, g_s(\mathbf{x}) = 0\}$.

This theorem is often referred to as the strong Nullstellensatz and has a useful corollary, which is called Hilbert's weak Nullstellensatz.

Corollary 2.3.8. *Let* $I = \langle g_1, \ldots, g_s \rangle$*, with* $g_1, \ldots, g_s \in \mathbb{R}[\mathbf{x}]$*. Then we have the following equivalence:*

(1) $\mathcal{V}_{\mathbb{C}}(I) = \varnothing$.

(2) $1 \in I$*, i.e.,* $1 = \sum_{i=1}^{s} p_i g_i$ *for some* $p_i \in \mathbb{R}[\mathbf{x}]$.

Thus, the weak Nullstellensatz gives an explicit *algebraic certificate* of the emptiness of any algebraic set.

The analog of a Nullstellensatz in real algebraic geometry is a *Positivstellensatz*. A Positivstellensatz is a representation theorem for a polynomial f which is strictly positive on a semialgebraic set K. The representation yields an algebraic evidence of the positivity of f on K. In the literature there is a huge amount of such representation theorems, dealing with different sets and having different focus and versions.
In fact, we already came across a Positivstellensatz that provides a certificate of positivity for a homogeneous polynomial on the unit simplex, namely Pólya's Thereom 2.2.9. The first who characterized positive (and nonnegative) polynomials on basic closed semialgebraic sets was Krivine [Kri64a] in 1964. His result was rediscovered by Stengle [Ste74] in 1973. But with a view to optimization this result does not help us to tackle constrained polynomial optimization problems over a semialgebraic set K, because the reformulation using the Positivstellensatz by Krivine/Stengle cannot be computed efficiently. Therefore, we have to use other Positivstellensätze which add further assumptions on K.
Before stating two prominent representatives of such results, we introduce the necessary terminology.

Definition 2.3.9.

(i) A subset $T \subseteq \mathbb{R}[\mathbf{x}]$ is called a *preorder* if it contains all squares in $\mathbb{R}[\mathbf{x}]$ and is closed under addition and multiplication, i.e.,

$$f^2 \in T \text{ for all } f \in \mathbb{R}[\mathbf{x}],\ T + T \subseteq T, \text{ and } T \cdot T \subseteq T.$$

(ii) A subset $M \subseteq \mathbb{R}[\mathbf{x}]$ is called a *quadratic module* if it contains 1, is closed under addition and under multiplication by squares, i.e.,

$$1 \in M,\ M + M \subseteq M, \text{ and } f^2 M \subseteq M \text{ for all } f \in \mathbb{R}[\mathbf{x}].$$ ○

Observe that every preorder is a quadratic module. The preorder respectively the quadratic module generated by polynomials $g_1, \ldots, g_s \in \mathbb{R}[\mathbf{x}]$ is given by

$$T(g_1, \ldots, g_s) = \left\{ \sum_{J \subseteq \{1,\ldots,s\}} \sigma_J \prod_{j \in J} g_j \, : \, \sigma_J \in \Sigma_n \right\} = \left\{ \sum_{\mathbf{e} \in \{0,1\}^s} \sigma_{\mathbf{e}} g_1^{e_1} \cdots g_s^{e_s} \, : \, \sigma_{\mathbf{e}} \in \Sigma_n \right\},$$
$$M(g_1, \ldots, g_s) = \left\{ \sigma_0 + \sum_{i=1}^{s} \sigma_i g_i \, : \, \sigma_0, \sigma_i \in \Sigma_n \right\}.$$

Obviously, every polynomial f contained in $T(g_1, \ldots, g_s)$ or $M(g_1, \ldots, g_s)$ has a sum of square decomposition and thus, is nonnegative. We now discuss two theorems showing that the converse holds under certain extra assumptions. In what follows, we consider basic closed semialgebraic sets K given by the polynomials $g_1, \ldots, g_s$, as defined in (2.3.2). In 1991, Schmüdgen [Sch91] proved the subsequent Positivstellensatz for the additional assumption that K is compact:

Theorem 2.3.10 (Schmüdgen's Positivstellensatz). *Assume the set K is compact. Given a polynomial $f \in \mathbb{R}[\mathbf{x}]$ which is positive on K, i.e., $f(\mathbf{x}) > 0$ for all $\mathbf{x} \in K$, then $f \in T(g_1, \ldots, g_s)$.*

Although this result naturally leads to a hierarchy of semidefinite relaxations for the constrained polynomial optimization problem, one drawback is that the representation $\sum_J \sigma_J g_J \in T(g_1, \ldots, g_s)$ involves 2^s sums of squares σ_J. Hence, the representation is exponential in the number s of the constraints defining K. Next we will present a Positivstellensatz first proven in 1993 by Putinar [Put93], which only involves a linear number of terms for the representation, and thus is more suitable for optimization purposes.
Roughly speaking, Putinar showed the analog of Schmüdgens Positivstellensatz, where the preorder is replaced by the quadratic module, but with the additional assumption that the quadratic module has to be Archimedean. To clarify, a quadratic module $M(g_1, \ldots, g_s)$ is called *Archimedean*, if for all $f \in \mathbb{R}[\mathbf{x}]$ there exists an integer $N \geqslant 1$ such that $N - f \in M(g_1, \ldots, g_s)$. This definition has several equivalent conditions:

(1) There exists a polynomial $p(\mathbf{x}) \in M(g_1, \ldots, g_s)$ such that the level set $\{\mathbf{x} \in \mathbb{R}^n : p(\mathbf{x}) \geqslant 0\}$ is compact.

(2) $N - \sum_{i=1}^n x_i^2 \in M(g_1, \ldots, g_s)$ for some integer $N \geqslant 1$.

(3) $N \pm x_i \in M(g_1, \ldots, g_s)$ for $i = 1, \ldots, n$ and some integer $N \geqslant 1$.

Clearly, condition (1) implies that the set K is compact. On the other hand, if K is compact, then it is contained in a ball of radius R for some $R \in \mathbb{N}$. By definition of boundedness, one can always force $M(g_1, \ldots, g_s)$ to be Archimedean simply by adding the (redundant) inequality $R^2 - \sum_{i=1}^n x_i^2 \geqslant 0$ to the description of K. However, in the general situation, the radius is not known a priori. An explicit example for a compact set K without $M(g_1, \ldots, g_s)$ being Archimedean is the Jacobi-Prestel counterexample, see [PD01]. With this in mind, we state

Theorem 2.3.11 (Putinar's Positivstellensatz)**.** *Assume that the quadratic module $M(g_1, \ldots, g_s)$ is Archimedean. For $f \in \mathbb{R}[\mathbf{x}]$, if $f(\mathbf{x}) > 0$ for all $\mathbf{x} \in K$, then $f \in M(g_1, \ldots, g_s)$.*

Note that in both given Positivstellensätzen the strict positivity of f is necessary. Putinar's Positivstellensatz provides the basis for the SOS relaxation established by Lasserre, this relaxation is subject of the next section.

2.3.4 Constrained Optimization and Lasserre's Relaxation

In this section the different strands offered above will be woven together. We use the knowledge of the last sections to tackle constrained optimization problems. In this context we establish the well-known Lasserre relaxation which yields a hierarchy of lower bounds converging to the optimal value.

We begin by stating the underlying problem. We consider the *constrained polynomial optimization problem (CPOP)*

$$f_K^* \;=\; \inf\{f(\mathbf{x}) : \mathbf{x} \in K\} \;=\; \sup\{\gamma \in \mathbb{R} : f(\mathbf{x}) - \gamma \geqslant 0 \text{ for all } \mathbf{x} \in K\},$$

with feasible set $K \subseteq \mathbb{R}^n$ given as the basic closed semialgebraic set K defined by polynomials $g_1, \ldots, g_s$ as in (2.3.2). The optimal value of (CPOP) is denoted by f_K^*.

By means of Putinar's Positivstellensatz we consider following SOS relaxation for (CPOP):

$$\begin{aligned} f_{\text{sos}} &= \sup\{\gamma \in \mathbb{R} \;:\; f - \gamma \in QM(g_1, \ldots, g_s)\} \\ &= \sup\left\{\gamma \;:\; f - \gamma = \sigma_0 + \sum_{i=1}^{s} \sigma_i g_i,\ \sigma_0, \sigma_i \in \Sigma_n \text{ for all } i = 1, \ldots, s\right\}. \end{aligned}$$

As already stated, this relaxation can be reformulated using positive semidefinite matrices with infinite dimension. In order to get an SDP involving finite-dimensional matrices, the key idea is to add degree bounds. Following Lasserre [Las01] we now construct such a relaxation. Let $2d \geqslant \max\{\deg(f), \deg(g_1), \ldots, \deg(g_s)\}$ then the *d-th Lasserre relaxation* is given by the parameter

$$(2.3.3) \qquad f_{\text{sos}}^{(d)} = \sup\left\{\gamma : f - \gamma = \sigma_0 + \sum_{i=1}^{s} \sigma_i g_i, \quad \begin{matrix} \sigma_0, \sigma_i \in \Sigma_n \text{ for all } i = 1, \ldots, s, \\ \text{with} \deg(\sigma_0), \deg(\sigma_i g_i) \leqslant 2d \end{matrix}\right\}.$$

The relaxation based on moments is given by

$$(2.3.4) \quad f_{\text{mom}}^{(d)} = \inf_y \left\{\sum_{\boldsymbol{\alpha}} f_{\boldsymbol{\alpha}} y_{\boldsymbol{\alpha}} : y_{\mathbf{0}} = 1, \, M_d(y) \succeq 0, \, M_{d-t_{g_i}}(g_i * y) \succeq 0 \, (i = 1, \ldots, s)\right\},$$

with sequence $y = (y_{\boldsymbol{\alpha}})_{\boldsymbol{\alpha} \in \mathbb{N}^n_{2d}}$, $t_{g_i} := \lceil \deg(g_i)/2 \rceil$, and the truncated localizing matrices $M_{d-t_{g_i}}(g_i * y)$. Both programs are semidefinite programs involving matrices of size $\binom{n+d}{d}$. Thus, for fixed d both parameters $f_{\text{sos}}^{(d)}$ and $f_{\text{mom}}^{(d)}$ can be computed in polynomial time (up to an ε-error).
Actually, the programs are dual semidefinite programs, i.e., it holds $f_{\text{sos}}^{(d)} \leqslant f_{\text{mom}}^{(d)} \leqslant f_K^*$. Under some conditions on K one can show that the bounds coincide, see [Las01] and [Sch05]. Again, an important question concerns the exactness of the bounds. Analogously to the moment relaxation in the unconstrained case, a flat extension condition provides a sufficient condition such that the bound $f_{\text{mom}}^{(d)}$ is exact. Moreover, Lasserre [Las01] showed that by Putinar's Positivstellensatz the bounds converge asymptotically if the quadratic module is Archimedean:

Theorem 2.3.12. *If the quadratic module $M(g_1, \ldots, g_s)$ is Archimedean, then it holds $\lim_{d\to\infty} f_{\text{sos}}^{(d)} = \lim_{d\to\infty} f_{\text{mom}}^{(d)} = f_K^*$.*

Hence, the Lasserre relaxation yields a hierarchy of lower bounds which converge monotonously to the optimal value. A crucial subsequent question concerns the *finite convergence* of the hierarchy, i.e., the convergence after finitely many steps. In general, only in some situations finite convergence can be guaranteed. For example, if the description of K additionally involves equality constraints $h_j(\mathbf{x}) = 0, j = 1, \ldots, m$, Laurent [Lau09, Theorem 6.15] showed:

Theorem 2.3.13. *Consider the optimization problem of minimizing* $f \in \mathbb{R}[\mathbf{x}]$ *over the set* $K = \{\mathbf{x} \in \mathbb{R}^n : h_1(\mathbf{x}) = 0, \ldots, h_m(\mathbf{x}) = 0, g_1(\mathbf{x}) \geqslant 0, \ldots, g_s(\mathbf{x}) \geqslant 0\}$. *Define the ideal* $I = \langle h_1, \ldots, h_m \rangle$.

(1) $|\mathcal{V}_{\mathbb{C}}(I)| < \infty$, *then* $f_{\mathrm{sos}}^{(d)} = f_{\mathrm{mom}}^{(d)} = f_K^*$, *for* d *large enough.*

(2) $|\mathcal{V}_{\mathbb{R}}(I)| < \infty$, *then* $f_{\mathrm{mom}}^{(d)} = f_K^*$, *for* d *large enough.*

Recently, Nie [Nie13a] proposed a new condition called *flat truncation* for the purpose of certifying exactness and finiteness.

For the problems of finding an SOS decomposition as well as the computation of the SOS respectively the moment relaxation for both (POP) and (CPOP), there exist several software packages, see [Lau09] for a short overview.

We conclude this section by mentioning some well known issues of the relaxations based on sums of squares. There are minor issues like there being cases with non-finite convergence or that no efficient bounds for finite convergence are known in the generic case. But the most crucial one is due to the usage of SDP to compute these relaxations, because these have long running times for a large number of variables or high-degree polynomials, which makes this approach challenging to use for problems arising in practice. Attacking these issues is an active area of research, and motivates the search for other nonnegativity certificates independent of sums of squares. In light of this, we discover a new cone, which approximates the nonnegativity cone independent of the SOS cone. Such a cone will be introduced in the next section.

2.4 The Cone of Sums of Nonnegative Circuit Polynomials

In this section we introduce sums of nonnegative circuit polynomials, the key object of this thesis. Circuit polynomials are certain sparse polynomials having a special structure in terms of their Newton polytopes and supports. The key property that makes the study of these polynomials attractive is that nonnegativity of circuit polynomials can be checked easily by a certain number which can be derived from the initial circuit polynomial immediately. Thereby, we have a new nonnegativity certificate, which even appears to be independent of SOS certificates. In what follows, we introduce these polynomials and discuss some important results, see also [IdW16a] and [dW15] for an overview.

We start by defining the class of *circuit polynomials*:

Definition 2.4.1. Let $f \in \mathbb{R}[\mathbf{x}]$ be supported on $A \subset \mathbb{N}^n$ such that all elements of $V(A)$ are even. Then f is called a *circuit polynomial* if it is of the form

$$f(\mathbf{x}) \;=\; \sum_{j=0}^{r} f_{\boldsymbol{\alpha}(j)} \mathbf{x}^{\boldsymbol{\alpha}(j)} + f_{\boldsymbol{\beta}} \mathbf{x}^{\boldsymbol{\beta}}, \tag{2.4.1}$$

with $r \leqslant n$, exponents $\boldsymbol{\alpha}(j)$, $\boldsymbol{\beta} \in A$, and coefficients $f_{\boldsymbol{\alpha}(j)} \in \mathbb{R}_{>0}$, $f_{\boldsymbol{\beta}} \in \mathbb{R}$, such that the following conditions hold:

(C1) The points $\boldsymbol{\alpha}(0), \boldsymbol{\alpha}(1), \ldots, \boldsymbol{\alpha}(r)$ are affinely independent and equal $V(A)$.

(C2) The exponent $\boldsymbol{\beta}$ can be written uniquely as

$$\boldsymbol{\beta} \;=\; \sum_{j=0}^{r} \lambda_j \boldsymbol{\alpha}(j) \;\text{ with }\; \lambda_j \;>\; 0 \;\text{ and }\; \sum_{j=0}^{r} \lambda_j \;=\; 1$$

in *barycentric coordinates* λ_j relative to the vertices $\boldsymbol{\alpha}(j)$ with $j = 0, \ldots, r$.

We call the terms $f_{\boldsymbol{\alpha}(0)} \mathbf{x}^{\boldsymbol{\alpha}(0)}, \ldots, f_{\boldsymbol{\alpha}(r)} \mathbf{x}^{\boldsymbol{\alpha}(r)}$ the *outer terms* and $f_{\boldsymbol{\beta}} \mathbf{x}^{\boldsymbol{\beta}}$ the *inner term* of f. For their corresponding exponents we refer to *outer exponents* respectively *inner exponent*. We denote the set of all circuit polynomials with support A by Circ_A.

For every circuit polynomial we define the corresponding *circuit number* as

$$\Theta_f \;=\; \prod_{j=0}^{r} \left(\frac{f_{\boldsymbol{\alpha}(j)}}{\lambda_j} \right)^{\lambda_j}. \tag{2.4.2}$$

⬡

Observe that by condition (C1) the set $V(A) = \{\boldsymbol{\alpha}(0), \ldots, \boldsymbol{\alpha}(r)\}$ is the vertex set of an r-dimensional *even* simplex, i.e., each $\boldsymbol{\alpha}(j) \in (2\mathbb{N})^n$, which coincides with $\mathrm{New}(f) = \mathrm{conv}(A)$. In such a case we say that $\mathrm{New}(f)$ is a *simplex Newton polytope*. Moreover, since the barycentric coordinates are strictly positive, we assume that $\boldsymbol{\beta} \in \mathrm{int}(\mathrm{New}(f))$, which is justified by Lemma 2.4.3. To clarify, if $\boldsymbol{\beta}$ was located on the boundary of $\mathrm{New}(f)$, then f could be written as a sum of two circuit polynomials, namely the face containing $\boldsymbol{\beta}$ and a monomial square, see Example 2.4.2 (ii).

The terms "circuit polynomial" and "circuit number" are chosen since the support $A = \{\boldsymbol{\beta}, \boldsymbol{\alpha}(0), \ldots, \boldsymbol{\alpha}(r)\}$ forms a *circuit*; this is a minimally affine dependent set. This terminology comes from matroid theory; see, e.g., [Oxl11, page 9].

There is a well-known representative in the class of circuit polynomials:

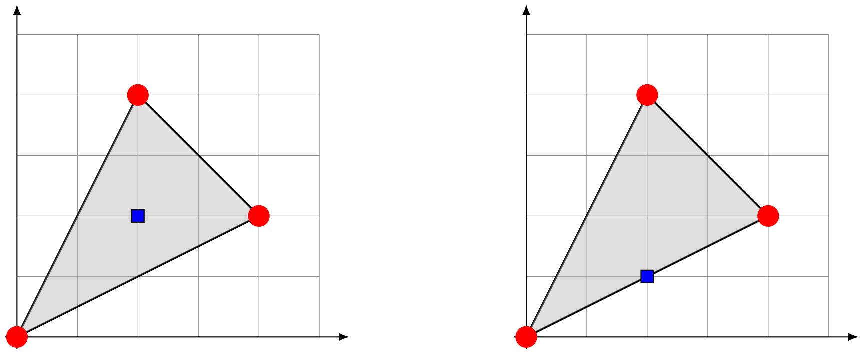

Figure 2.1: The support sets of the Motzkin polynomial $m(x_1, x_2)$ and the polynomial $p(x_1, x_2)$ in Example 2.4.2.

Example 2.4.2.

(i) Consider the Motzkin polynomial $m(x_1, x_2) = 1 + x_1^4 x_2^2 + x_1^2 x_2^4 - 3x_1^2 x_2^2$. See Figure 2.1 for the support of m. Obviously, m is a circuit polynomial since the Newton polytope $\text{New}(m) = \text{conv}\{(\mathbf{0}, \mathbf{0}), (\mathbf{4}, \mathbf{2}), (\mathbf{2}, \mathbf{4})\}$ is a simplex and m has one additional support point coming from the inner term $-3x_1^2 x_2^2$, which is located strictly in the interior of this simplex, namely $(\mathbf{2}, \mathbf{2}) \in \text{int}(\text{New}(m))$.

(ii) Now consider $p(x_1, x_2) = 1 + x_1^4 x_2^2 + x_1^2 x_2^4 - 3x_1^2 x_2$. This polynomial arises from the Motzkin polynomial by shifting the inner exponent $(2, 2)$ to the boundary of $\text{New}(m) = \text{New}(p)$ onto $(2, 1)$, see also Figure 2.1. We can write p as the sum $p = p_1 + p_2 = (1 + x_1^4 x_2^2 - 3x_1^2 x_2) + (x_1^2 x_2^4)$, with a circuit polynomial p_1 in the sense of Definition 2.4.1 and p_2 being a monomial square, which can be interpreted as a degenerated circuit polynomial. ◇

In what follows, we characterize nonnegativity for circuit polynomials. But at first, we provide the observation allowing us to assume that $\boldsymbol{\beta} \in \text{int}(\text{New}(f))$, see [IdW16a, Lemma 3.7].

Lemma 2.4.3. *Let* $f(\mathbf{x}) = \sum_{j=0}^{r} f_{\boldsymbol{\alpha}(j)} \mathbf{x}^{\boldsymbol{\alpha}(j)} + f_{\boldsymbol{\beta}} \mathbf{x}^{\boldsymbol{\beta}}$ *be such that* $\text{New}(f)$ *is a simplex and* $\boldsymbol{\beta} \in \partial\,\text{New}(f)$. *Furthermore, let* F *be the face of* $\text{New}(f)$ *containing* $\boldsymbol{\beta}$. *Then,* f *is nonnegative if and only if the restriction to the face* F *is nonnegative.*

We now state the fundamental fact that nonnegativity of a circuit polynomial f can be decided easily via its circuit number Θ_f alone:

Theorem 2.4.4 ([IdW16a], Theorem 3.8)**.** *Let f be a circuit polynomial with inner term $f_{\boldsymbol{\beta}}\mathbf{x}^{\boldsymbol{\beta}}$ and let Θ_f be the corresponding circuit number, as defined in* (2.4.2)*. Then the following statements are equivalent:*

(1) *f is nonnegative.*

(2) $|f_{\boldsymbol{\beta}}| \leqslant \Theta_f$ *and* $\boldsymbol{\beta} \notin (2\mathbb{N})^n$ *or* $f_{\boldsymbol{\beta}} \geqslant -\Theta_f$ *and* $\boldsymbol{\beta} \in (2\mathbb{N})^n$*.*

Note that statement (2) is equivalent to: $|f_{\boldsymbol{\beta}}| \leqslant \Theta_f$ or f is a sum of monomial squares. Thus, Theorem 2.4.4 shows that we only have to solve a system of linear equations to check nonnegativity of a circuit polynomial. The proof of the above theorem is based on the norm relaxation method. We denote by $\langle \cdot, \cdot \rangle$ the standard inner product. A key observation for circuit polynomials $f(\mathbf{x})$ is that nonnegativity of $f(\mathbf{x})$ is equivalent to nonnegativity of $f(e^{\mathbf{x}}) = \sum_{j=0}^{r} f_{\boldsymbol{\alpha}(j)} e^{\langle \mathbf{x}, \boldsymbol{\alpha}(j) \rangle} + f_{\boldsymbol{\beta}} e^{\langle \mathbf{x}, \boldsymbol{\beta} \rangle}$ for all $\mathbf{x} \in \mathbb{R}^n$, where without loss of generality $f_{\boldsymbol{\beta}}$ is assumed to be strictly negative after a possible transformation of variables $x_j \mapsto -x_j$. Then one can show that $f(e^{\mathbf{x}})$ with $f_{\boldsymbol{\beta}} = -\Theta_f$ has a unique global minimizer and one can even compute this root in special cases, see Section 3.2. The knowledge about the minimum finally leads to the statements of Theorem 2.4.4.
We remark that for special instances the result of Theorem 2.4.4 was known before. In [Rez89] Reznick showed this for f being an *agiform*, that is a special case of a circuit polynomial when choosing $f_{\boldsymbol{\alpha}(j)} = \lambda_j$ and $f_{\boldsymbol{\beta}} = -1$. The term agiform is implied by the fact that its nonnegativity follows by the arithmetic-geometric mean inequality. And Fidalgo and Kovacec proved the result for circuit polynomials with standard simplex Newton polytope, i.e., $\text{New}(f) = \Delta_{n,2d}$, see [FK11].

An immediate consequence of the above theorem is an upper bound for the number of zeros of circuit polynomials, but we will postpone this discussion to Section 3.2. Furthermore, a direct corollary can be drawn from the observations above:

Corollary 2.4.5. *A circuit polynomial f with constant term is located on the boundary of the cone of nonnegative polynomials, i.e., $f \in \partial P_{n,2d}$, if and only if $f_{\boldsymbol{\beta}} \in \{\pm\Theta_f\}$ and $\boldsymbol{\beta} \notin (2\mathbb{N})^n$ or $f_{\boldsymbol{\beta}} = -\Theta_f$ and $\boldsymbol{\beta} \in (2\mathbb{N})^n$.*

Another equivalent condition for $f \in \partial P_{n,2d}$ is that the *A-discriminant* vanishes at f, for more information we refer to [IdW16a, Section 4.3].

The example hereinafter demonstrates the result of Theorem 2.4.4.

Example 2.4.6. Again we consider the Motzkin polynomial $m = 1+x_1^4x_2^2+x_1^2x_2^4-3x_1^2x_2^2$. To test nonnegativity of m, we have to compare its circuit number Θ_m with the coefficient $m_{\boldsymbol{\beta}} = -3$. Since $\boldsymbol{\beta} = (2,2) \in (2\mathbb{N})^2$, m is nonnegative if and only if $m_{\boldsymbol{\beta}} = -3 \geqslant -\Theta_m$. The barycentric coordinates of the inner exponent $(2,2)$ are given by $\lambda_j = \frac{1}{3}$ for $j \in \{0,1,2\}$. This yields the circuit number

$$\Theta_m = \prod_{j=0}^{2}\left(\frac{m_{\boldsymbol{\alpha}(j)}}{\lambda_j}\right)^{\lambda_j} = \left(\frac{1}{1/3}\right)^{1/3}\cdot\left(\frac{1}{1/3}\right)^{1/3}\cdot\left(\frac{1}{1/3}\right)^{1/3} = 3.$$

Thus, $m_{\boldsymbol{\beta}} = -3 = -\Theta_m$ and we can conclude that the Motzkin polynomial is nonnegative, hence it is a nonnegative circuit polynomial, and by Corollary 2.4.5 it is contained in the boundary of the cone of nonnegative polynomials. ○

On the basis of the provided results, we deduce that writing a polynomial as a sum of nonnegative circuit polynomials is a certificate of nonnegativity. More formally we define the set of such polynomials as follows:

Definition 2.4.7. We define for every $n, d \in \mathbb{N}^*$ the set of *sums of nonnegative circuit polynomials (SONC)* in n variables of degree $2d$ as

$$C_{n,2d} = \left\{p \in \mathbb{R}[\mathbf{x}]_{n,2d} \ : \ p = \sum_{i=1}^{k}\mu_i f_i, \mu_i \geqslant 0, f_i \in \mathrm{Circ}_A \cap P_{n,2d}, A \subseteq \mathcal{L}_{n,2d}, k \in \mathbb{N}^*\right\}.$$

○

We denote by *SONC* both the class of polynomials that are *sums of nonnegative circuit polynomials* and the property of a polynomial to be in this class. Indeed, SONC polynomials form a convex cone independent of the SOS cone.

Theorem 2.4.8 ([IdW16a], Proposition 7.2)**.** *$C_{n,2d}$ is a convex cone satisfying:*

(1) *$C_{n,2d} \subseteq P_{n,2d}$ for all $n, d \in \mathbb{N}^*$,*

(2) *$C_{n,2d} \subseteq \Sigma_{n,2d}$ if and only if $(n, 2d) \in \{(1, 2d), (n, 2), (2, 4)\}$,*

(3) *$\Sigma_{n,2d} \nsubseteq C_{n,2d}$ for all $(n, 2d)$ with $2d \geqslant 6$.*

Obviously, $C_{n,2d}$ is a convex cone as for $\alpha, \beta \in \mathbb{R}_{>0}$ and $p, q \in C_{n,2d}$ it holds that $\alpha p + \beta q \in C_{n,2d}$. The first statement of Theorem 2.4.8 is trivial. The "if" part of the second statement follows from Hilbert's Theorem 2.2.6 and for the "only if" part one can explicitly obtain polynomials in $C_{n,2d} \backslash \Sigma_{n,2d}$. Note in this context, that the Motzkin polynomial is a SONC polynomial, which is not SOS. The last statement is a consequence of a zero argument. The only case where this argument fails is the case $(n, 2d) = (n, 4)$. One contribution of this thesis is to complete the cone containment statement, see Theorem 3.1.2. In sum, Theorem 2.4.8 yields that sums of nonnegative circuit polynomials are a new type of *nonnegativity certificates*, which are *independent* of sums of squares.

The obvious follow-up question is how to detect a SONC decomposition of a given polynomial, or in other words how to check membership in the cone $C_{n,2d}$. It turns out that in many cases this can be done by *geometric programming (GP)*. A GP is a special type of convex optimization problem, which can be solved in polynomial time using interior point methods; see [DPZ67] and [NN94]. To be more precise, a polynomial with simplex Newton polytope is SONC if and only if a GP is feasible. This is in direct analogy to the relation between SOS and SDP. A detailed discussion in context of polynomial optimization is given in Chapter 4.

Note that just like the SOS condition, the condition to be a sum of nonnegative circuit polynomials is only sufficient for nonnegativity of a polynomial. However, for special cases it is indeed necessary, see [IdW16a, Corollary 7.5]:

Proposition 2.4.9. *Let* $f = \sum_{j=0}^{r} f_{\boldsymbol{\alpha}(j)} \mathbf{x}^{\boldsymbol{\alpha}(j)} + \sum_{i=1}^{k} f_{\boldsymbol{\beta}(i)} \mathbf{x}^{\boldsymbol{\beta}(i)}$ *be nonnegative with* $f_{\boldsymbol{\alpha}(j)} \in \mathbb{R}_{>0}$ *and* $f_{\boldsymbol{\beta}(i)} \in \mathbb{R}^*$ *such that* $\operatorname{New}(f) = \operatorname{conv}\{\boldsymbol{\alpha}(0), \boldsymbol{\alpha}(1), \ldots, \boldsymbol{\alpha}(r)\}$ *is an even simplex and all* $\boldsymbol{\beta}(i) \in (\operatorname{int}(\operatorname{New}(f)) \cap \mathbb{N}^n)$*. If there exists a vector* $\mathbf{y} \in (\mathbb{R}^*)^n$ *such that* $f_{\boldsymbol{\beta}(i)} \mathbf{y}^{\boldsymbol{\beta}(i)} < 0$ *for all* $i = 1, \ldots, k$*, then* f *is SONC.*

2.4.1 Sums of Squares supported on a Circuit

In this section we provide a short overview of the results concerning the question in which cases a circuit polynomial f is a sum of squares. Remarkably, this question depends only on the lattice point configuration of the Newton polytope of f and the location of the interior point, see Theorem 2.4.14. The given results generalize results by Reznick [Rez89] for agiforms. Thus, we first provide the needed background and mention some of Reznick's results on agiforms, and then we consider general circuit polynomials.

Recall that a circuit polynomial with $f_{\boldsymbol{\alpha}(j)} = \lambda_j$ and $f_{\boldsymbol{\beta}} = -1$ with simplex Newton polytope $\mathrm{New}(f) =: \Delta$ is called an agiform:

$$f(\Delta, \boldsymbol{\lambda}, \boldsymbol{\beta})(\mathbf{x}) = \sum_{j=0}^{r} \lambda_j \mathbf{x}^{\boldsymbol{\alpha}(j)} - \mathbf{x}^{\boldsymbol{\beta}}.$$

Definition 2.4.10. Let $\hat{\Delta} := \{\boldsymbol{\alpha}(0), \boldsymbol{\alpha}(1), \ldots, \boldsymbol{\alpha}(r)\} \subset (2\mathbb{N})^n$ be such that $\mathrm{conv}(\hat{\Delta})$ is a simplex and let $\mathcal{B} \subseteq \mathrm{conv}(\hat{\Delta}) \cap \mathbb{Z}^n$.

(i) Define by $A(\mathcal{B}) := \left\{\frac{1}{2}(\boldsymbol{s}+\boldsymbol{t}) \in \mathbb{Z}^n : \boldsymbol{s}, \boldsymbol{t} \in \mathcal{B} \cap (2\mathbb{Z})^n\right\}$ and $\overline{A}(\mathcal{B}) := \{\frac{1}{2}(\boldsymbol{s}+\boldsymbol{t}) \in \mathbb{Z}^n : \boldsymbol{s} \neq \boldsymbol{t}, \boldsymbol{s}, \boldsymbol{t} \in \mathcal{B} \cap (2\mathbb{Z})^n\}$ the set of averages of even resp. distinct even points in $\mathcal{B}$.

(ii) We say that $\mathcal{B}$ is $\hat{\Delta}$-*mediated*, if

$$\hat{\Delta} \subseteq \mathcal{B} \subseteq (\overline{A}(\mathcal{B}) \cup \hat{\Delta}),$$

i.e., every $\mathbf{b} \in \mathcal{B} \backslash \hat{\Delta}$ is an average of two distinct even points in $\mathcal{B}$. ○

The following result shows the existence of a maximal mediated set, and the proof actually gives a brute force algorithm for constructing such a set, see [Rez89, Theorem 2.2].

Theorem 2.4.11. *There is a $\hat{\Delta}$-mediated set Δ^* satisfying $A(\hat{\Delta}) \subseteq \Delta^* \subseteq (\Delta \cap \mathbb{Z}^n)$, which contains every $\hat{\Delta}$-mediated set.*

If $A(\hat{\Delta}) = \Delta^*$ resp. $\Delta^* = (\Delta \cap \mathbb{Z}^n)$, we call Δ an *M-simplex* resp. *H-simplex*. These terms are chosen due to the following well-known representatives:

Example 2.4.12.

(i) The standard simplex $\Delta_{n,2d} = \mathrm{conv}\{\mathbf{0}, 2d \cdot \mathbf{e}_1, \ldots, 2d \cdot \mathbf{e}_n\} \subset \mathbb{R}^n$ for $d \in \mathbb{N}^*$ is an H-simplex. Sometimes $\Delta_{n,2d}$ is called the Hurwitz simplex, since by Hurwitz' Theorem, see [Hur91], it follows that for $\boldsymbol{\beta} \in A(\hat{\Delta}_{n,2d}) = \{\boldsymbol{\beta} : 0 \leqslant \beta_i \in \mathbb{Z} \text{ and } |\boldsymbol{\beta}| = 2d\}$ the agiform $2d \cdot f(\Delta_{n,2d}, \frac{1}{2d}\boldsymbol{\beta}, \boldsymbol{\beta})$ is a sum of squares.

(ii) The Newton polytope $\mathrm{conv}\{(0,0),(2,4),(4,2)\} \subset \mathbb{R}^2$ of the Motzkin polynomial $m(\mathbf{x}) = 1 + x_1^4 x_2^2 + x_1^2 x_2^4 - 3x_1^2 x_2^2$ is an M-simplex. ○

Now we are able to state Reznick's result [Rez89, Corollary 4.9] concerning the question when agiforms are sums of squares:

Proposition 2.4.13. *Let $f(\Delta, \boldsymbol{\lambda}, \boldsymbol{\beta})(\mathbf{x})$ be an agiform. Then it holds that $f \in \Sigma_{n,2d}$ if and only if $\boldsymbol{\beta} \in \Delta^*$.*

This generalizes as follows to arbitrary circuit polynomials, see [IdW16a, Theorem 5.2]:

Theorem 2.4.14. *Let $f \in C_{n,2d}$ be a nonnegative circuit polynomial with* New$(f) = \Delta$. *Then,*

$$f \in \Sigma_{n,2d} \quad \textit{if and only if} \quad \boldsymbol{\beta} \in \Delta^* \quad \textit{or} \quad f_{\boldsymbol{\beta}} > 0 \textit{ and } \boldsymbol{\beta} \in (2\mathbb{N})^n.$$

Note again that $f_{\boldsymbol{\beta}} > 0$ and $\boldsymbol{\beta} \in (2\mathbb{N})^n$ holds if and only if f is a sum of monomial squares, and hence the statement obviously holds.

We conclude this section by providing two immediate corollaries of Theorem 2.4.14.

Corollary 2.4.15. *Let f be a circuit polynomial such that* New(f) *is an H-simplex. Then, $f \in C_{n,2d}$ if and only if $f \in \Sigma_{n,2d}$.*

The second result concerns the polynomial optimization for circuit polynomials, namely

Corollary 2.4.16. *Let $f \in C_{n,2d}$ be a circuit polynomial with* New$(f) = \Delta$. *Then, $f_{\text{sos}} = f^*$ if and only if $\boldsymbol{\beta} \in \Delta^*$.*

Most of the results above can be extended to the case of polynomials having a simplex Newton polytope with various interior points, see again [IdW16a] for a careful treatment.

2.4.2 Further Results

In this section we briefly outline further results from [IdW16a, Section 4] for SONC polynomials.

A question directly following by the discussions in Section 2.4.1 is whether the results can be extended to polynomials with *non-simplex* Newton polytopes P. In this case it can be shown that the lattice point criterion given by the maximal mediated set from the simplex case does not suffice to characterize sums of squares. But on the positive side there exists a necessary and sufficient criterion based on the set of possible triangulations of P combined with maximal mediated sets.

Another interesting class of polynomials is the set of convex polynomials. A polynomial $f \in \mathbb{R}[\mathbf{x}]$ is *convex* if its Hessian H_f is positive semidefinite for all $\mathbf{x} \in \mathbb{R}^n$. We denote the cone of all n-variate degree $2d$ convex polynomials by $K_{n,2d}$. As it is the case for nonnegativity, deciding convexity of a polynomial is co-NP-hard in general, see [AOPT13]. A fundamental fact is that the convexity property is not preserved under homogenization, which stands in contrast to nonnegative polynomials, sums of squares, and also SONC polynomials as we will see in Proposition 3.1.5. Interestingly, every convex form is nonnegative, thus this cone relates to the study of the considered three cones ($P_{n,2d}$, $\Sigma_{n,2d}$, and $C_{n,2d}$). Despite the fact that the relation between convex forms and sums of squares are not well understood except that their cones are not contained in each other, convex forms supported on a circuit are easier to understand. In fact, one can completely characterize convex polynomials supported on a circuit and get the following interesting result:

Theorem 2.4.17. *Let $f \in C_{n,2d}$ be a nonnegative circuit polynomial and $n \geqslant 2$. Then f is not convex, i.e., $f \notin K_{n,2d}$.*

Hence, $C_{n,2d} \cap K_{n,2d} = \varnothing$ for $n \geqslant 2$.

Finally, circuit polynomials have a nice connection to the theory of amoebas.

Chapter 3

The SONC Cone revisited

In this chapter we take a deeper look at the SONC cone $C_{n,2d}$ introduced in Section 2.4. To understand this cone in more detail is motivated in several ways. On the one hand $C_{n,2d}$ is a new cone approximating the nonnegativity cone $P_{n,2d}$. Therefore, the study of the SONC cone is naturally embedded in the rich theory of nonnegative polynomials and sums of squares. Thus it is important to explore the structure and properties of this cone as well as to make general observations concerning $C_{n,2d}$. On the other hand a deeper knowledge of the cone $C_{n,2d}$ is also desirable from the practical viewpoint. Since SONC polynomials serve as a certificate of nonnegativity, they can be used in applications to polynomial optimization problems, as we will see in Chapters 4 and 5. In consequence, a better understanding of this cone hopefully yields an improvement of these methods.

Hence, in this chapter we dive into the theoretical study of the SONC cone. We start with presenting some general results concerning the structure and properties of $C_{n,2d}$ and dedicate the second section to the analysis of the real zeros of SONC polynomials and forms. Here, the main contribution is a complete and explicit characterization of the real zeros of SONC polynomials and forms. We add some interesting observations as a result of the new knowledge about the real zeros. In the third section we provide, based on the preceding sections, an approach to the exposed faces of $C_{n,2d}$. Finally, in the fourth section, we show that the set of SONC polynomials is not closed under multiplication, which stands in strong contrast to the set of sums of squares. Moreover, we derive the important result, that the SONC cone is always full-dimensional in the cone of nonnegative polynomials. Both results of Section 3.4 can be seen as contextual transition to the aim of the subsequent chapters, namely the application of SONC polynomials to polynomial optimization problems.

3.1 Deeper Analysis of the SONC Cone

In what follows, we present some important, interesting properties and observations of the SONC cone itself. First we prove that this cone is proper, then we give the missing piece of the statement about the (non-)containment of the SONC cone and the SOS cone. After that we look at the realizability of nonnegative circuit polynomials with a certain degree. We conclude this section by taking a first step towards the analysis of SONC forms.

Proposition 3.1.1. *The SONC cone $C_{n,2d}$ is a proper cone in $\mathbb{R}[\mathbf{x}]_{n,2d} \sim \mathbb{R}^{N(n,2d)}$.*

Proof. For $C_{n,2d}$ to be a proper cone, it has to be convex, solid, closed, and pointed. Clearly, the cone is convex.
To prove that $C_{n,2d}$ is solid, we have to show that the interior of the cone is nonempty. This follows immediately from the fact that the SONC cone is full-dimensional in $\mathbb{R}[\mathbf{x}]_{n,2d}$, which will be proven in Theorem 3.4.3.
To evidence the closedness of $C_{n,2d}$ suppose $p_j(\mathbf{x}) = \sum_{i=1}^{k_j} f_i^{(j)}(\mathbf{x}) \in C_{n,2d}$ and $p_j(\mathbf{x}) \to p(\mathbf{x})$, for $j \to \infty$, with respect to some norm. As $p_j \in C_{n,2d}$ and each $f_i^{(j)}$ is a nonnegative circuit polynomial, it holds $\deg(f_i^{(j)}) \leqslant 2d$ for all $i = 1, \ldots, k_j$. Because p_j is a convex combination of polynomials of a finite-dimensional vector space, Carathéodory's Theorem, see Theorem 2.1.1, implies that the number k_j of summands of p_j is bounded by $\binom{n+2d}{2d} + 1$, thus, independent of j. Since the norm of each $f_i^{(j)}$ is bounded by p_j, each $f_i^{(j)}$ lies in a compact set, due to the compactness of any bounded set in a finite-dimensional vector space. Consequently, the coefficients of the polynomials $f_i^{(j)}$ are uniformly bounded and we can choose a convergent subsequence $f_i^{(j_l)}(\mathbf{x}) \to f_i(\mathbf{x})$, for $l \to \infty$. Hence, there is a convergent subsequence $p_j^{(l)}(\mathbf{x}) \to s(\mathbf{x}) \in C_{n,2d}$, for $l \to \infty$, thus we can conclude $p(\mathbf{x}) \in C_{n,2d}$.
It remains to show that $C_{n,2d}$ is pointed, i.e., the cone contains no line. For closed, solid cones this geometric property is equivalent to $C_{n,2d} \cap -C_{n,2d} = \{0\}$. This condition is apparent as $C_{n,2d} \subseteq P_{n,2d}$ and $-C_{n,2d} \subseteq -P_{n,2d}$, and clearly $P_{n,2d} \cap -P_{n,2d} = \{0\}$. □

We now give the missing cases of Theorem 2.4.8 statement (3), where the noncontainment of the SOS cone in the SONC cone was only shown for degree $2d \geqslant 6$. Thus, we get the following result, whereby we actually prove the full statement rather than only the missing cases.

Theorem 3.1.2. $\Sigma_{1,2} = C_{1,2}$, $\Sigma_{n,2} \nsubseteq C_{n,2}$ *for all* $n \geqslant 2$, *and* $\Sigma_{n,2d} \nsubseteq C_{n,2d}$ *for all* $(n, 2d)$ *with* $2d \geqslant 4$.

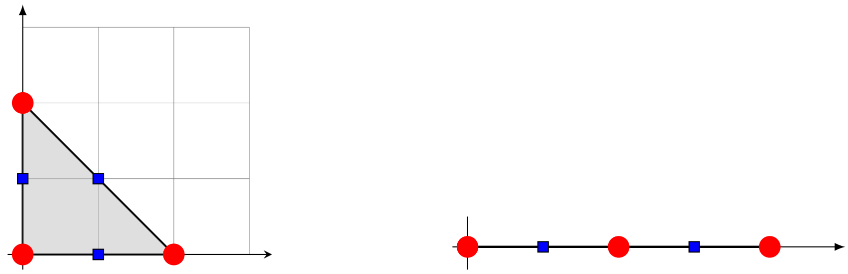

Figure 3.1: The support sets of $q(x_1, x_2)$ and of $r(x)$. The even points are the red ones.

Proof. The first statement was already observed in [IdW16a], but not proven as it is rather obvious. For the sake of completeness we give a short argument here. Consider an arbitrary polynomial $p \in \mathbb{R}[x]_{1,2}$, i.e., $p(x) = ax^2 + bx + c$, with $a, b, c \in \mathbb{R}$. Note that the support of p forms a circuit. Obviously, p is nonnegative if and only if p is a nonnegative circuit polynomial. Hence, $C_{1,2} = P_{1,2} = \Sigma_{1,2}$, where the last equality follows by Hilbert's Theorem 2.2.6.

For proving the second assertion, we explicitly construct a polynomial which is SOS but not SONC. First we explain the bivariate case in detail, then we generalize this idea to an arbitrary number of variables. Consider the following bivariate polynomial of degree 2

$$q(x_1, x_2) = 1 + x_1^2 + x_2^2 + 2x_1x_2 + 2x_1 + 2x_2 = (1 + x_1 + x_2)^2,$$

which is clearly SOS. The Newton polytope of q is the standard simplex $\Delta_{2,2}$. Additionally to the even vertices we have odd support points on every edge of $\Delta_{2,2}$, see Figure 3.1. Therefore, there is only one possibility to write q as a sum of circuit polynomials:

$$q = f_1 + f_2 + f_3 = \left(\frac{1}{2} + \frac{1}{2}x_1^2 + 2x_1\right) + \left(\frac{1}{2} + \frac{1}{2}x_2^2 + 2x_2\right) + \left(\frac{1}{2}x_1^2 + \frac{1}{2}x_2^2 + 2x_1x_2\right).$$

Clearly, all circuit polynomials f_i are not nonnegative, because $2 > \Theta_{f_i} = 1$ for all $i \in \{1, 2, 3\}$. Thus, q is not a SONC polynomial. For $n \geqslant 2$ we generalize this idea by constructing a polynomial whose support consists of the vertices of the standard simplex $\Delta_{n,2}$ and, in addition, the midpoints of each of the $\binom{n}{2}$ edges:

$$\hat{q}(\mathbf{x}) = (1 + x_1 + x_2 + \cdots + x_n)^2.$$

Showing that this polynomial is not a SONC polynomial is analogous to the bivariate case.

To proof the third statement, we again explicitly construct polynomials with the desired property. We start with $n = 1$, i.e., the univariate case and with the following polynomial of degree 4:

$$r(x) = 1 + 4x + 6x^2 + 4x^3 + x^4 = (1 + 2x + x^2)^2.$$

Figure 3.1 shows the support of r. The only meaningful option to split the polynomial into circuit polynomials which potentially are nonnegative is to divide it into the two symmetric circuit polynomials:

$$r = f_1 + f_2 = (1 + 4x + 3x^2) + (3x^2 + 4x^3 + x^4).$$

By computing the circuit number $\Theta_{f_i} = (12)^{\frac{1}{2}} \approx 3.4641 < 4$, we can easily see that both f_i are not nonnegative. To get a univariate polynomial with an arbitrary degree $2d$ which is SOS but not SONC, we only need to shift the above polynomial:

$$r'(x) = x^{2d-4} + 4x^{2d-3} + 6x^{2d-2} + 4x^{2d-1} + x^{2d} = (x^{d-2} + 2x^{d-1} + x^d)^2,$$

with $d \geqslant 2$, which means $2d \geqslant 4$.
By transferring this construction to further variables, we can prove the general multivariate case for $2d \geqslant 4$:

$$\hat{r}(\mathbf{x}) = (x_1^{d-2} + 2x_1^{d-1} + x_1^d)^2 + \cdots + (x_n^{d-2} + 2x_n^{d-1} + x_n^d)^2 \in \Sigma_{n,2d} \backslash C_{n,2d}. \tag{3.1.1}$$

□

Remark 3.1.3. Choosing $2d = 4$ in (3.1.1) yields a polynomial in $\Sigma_{n,4} \backslash C_{n,4}$, which explicitly witnesses the missing case of Theorem 2.4.8 (3).

Note that the degree of each variable part in the sum of the above polynomial (3.1.1) need not to be the same, i.e., a more general polynomial in $\Sigma_{n,2d} \backslash C_{n,2d}$ is

$$\hat{r}(\mathbf{x}) = (x_1^{d_1-2} + 2x_1^{d_1-1} + x_1^{d_1})^2 + \cdots + (x_n^{d_n-2} + 2x_n^{d_n-1} + x_n^{d_n})^2,$$

where $d := \max_{i=1,\ldots,n}\{d_1, \ldots, d_n\}$ and $d_i \geqslant 2$ for all $i = 1, \ldots, n$.

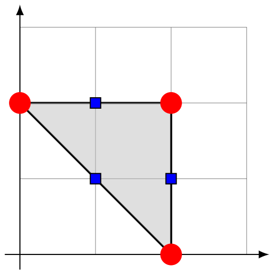

Figure 3.2: The support set of $s(x_1, x_2)$.

Another reasonable approach to construct a polynomial showing $\Sigma_{n,2d} \nsubseteq C_{n,2d}$ for all $(n, 2d)$ with $n \geqslant 2$ and $2d \geqslant 4$ is by making use of the idea in the proof of the second assertion of Theorem 3.1.2. Namely, to construct a polynomial whose support contains points on the boundary of the Newton polytope. More precisely, we can reflect the standard simplex $\Delta_{2,2}$ with additional boundary points on each edge with respect to the hypotenuse, see Figure 3.2, to get the following bivariate polynomial of degree 4:

$$s(x_1, x_2) = (x_1 + x_2 + x_1x_2)^2 = x_1^2x_2^2 + x_1^2 + x_2^2 + 2x_1x_2 + 2x_1^2x_2 + 2x_1x_2^2.$$

This polynomial is SOS but not SONC. By adding additional variables of certain degree, we get a polynomial in $\Sigma_{n,2d} \backslash C_{n,2d}$:

$$\hat{s}(\mathbf{x}) = (x_1 + x_2 + x_1x_2)^2 + \sum_{i=3}^{n} x_i^{2d_i},$$

with $d := \max_{i=1,\ldots,n}\{d_1, \ldots, d_n\}$.

A natural question arising due to the special structure of a circuit polynomial is from which degree on is it possible for an n-variate polynomial f to be a proper circuit polynomial. By *proper* we mean that f is not a sum of monomial squares. Indeed, the answer to that question depends on the exponent of the inner term.

Lemma 3.1.4. *Let f be an n-variate circuit polynomial of degree $2d$ with inner term $f_{\boldsymbol{\beta}}\mathbf{x}^{\boldsymbol{\beta}}$. Then f can be a proper circuit polynomial if and only if*

(1) $2d \geqslant n + 1$, *for* $\boldsymbol{\beta} \notin (2\mathbb{N})^n$.

(2) $d \geqslant n + 1$, *for* $\boldsymbol{\beta} \in (2\mathbb{N})^n$.

Proof. Consider a circuit polynomial f as in (2.4.1)

$$f(x_1, \ldots, x_n) = \sum_{j=0}^{r} f_{\boldsymbol{\alpha}(j)} \mathbf{x}^{\boldsymbol{\alpha}(j)} + f_{\boldsymbol{\beta}} \mathbf{x}^{\boldsymbol{\beta}}.$$

Recall that we assume $\boldsymbol{\beta} \in \text{int}(\text{New}(f))$. Thus, it follows that the smallest possible inner exponent $\boldsymbol{\beta}$ such that f may be a circuit polynomial is $\beta_i = 1$, $i = 1, \ldots, n$ for an odd inner exponent and respectively $\beta_i = 2$, $i = 1, \ldots, n$ if the inner exponent is even. Hence, it is easy to see that if the number of variables is strictly smaller than the degree, i.e., $n < 2d$, a circuit polynomial with odd inner exponent is realizable. And a circuit polynomial with $\boldsymbol{\beta} \in (2\mathbb{N})^n$ is only realizable if $2n < 2d$. The result follows directly from these considerations. □

The observation in Lemma 3.1.4 holds of course as well for nonnegative circuit polynomials. The above statement is in particular useful for the study of the real zeros of SONCs.

So far we only studied SONC polynomials. In the literature often results on nonnegative polynomials and SOS are stated homogeneously, i.e., for forms. For better comparability, in what follows we also consider SONC forms and investigate their behavior. A first important observation is to show that the property to be SONC is inherited under homogenization and, conversely, is preserved when a form is dehomogenized into a polynomial.

Proposition 3.1.5. *If a polynomial $p \in \mathbb{R}[x_1, \ldots, x_n]$ is a SONC polynomial of degree $2d$, then its homogenization $\overline{p}(x_0, \ldots, x_n) = x_0^{2d} p\left(\frac{x_1}{x_0}, \ldots, \frac{x_n}{x_0}\right)$ is also SONC; and vice versa.*

Proof. Since a SONC polynomial p is sum of nonnegative circuit polynomials f_i, i.e., $p = \sum_{i=1}^{k} \mu_i f_i, \mu_i \geqslant 0$, it suffices to prove the statement for circuit polynomials. So let f be a circuit polynomial and $\overline{f}$ its homogenization:

$$f(\mathbf{x}) = f(x_1, \ldots, x_n) = \sum_{j=0}^{r} f_{\boldsymbol{\alpha}(j)} \mathbf{x}^{\boldsymbol{\alpha}(j)} + f_{\boldsymbol{\beta}} \mathbf{x}^{\boldsymbol{\beta}},$$

$$\overline{f}(x_0, \mathbf{x}) = x_0^{2d} f\left(\frac{x_1}{x_0}, \ldots, \frac{x_n}{x_0}\right) = \sum_{j=0}^{r} f_{\boldsymbol{\alpha}(j)} \mathbf{x}^{\boldsymbol{\alpha}(j)} x_0^{2d - \sum_{i=1}^{n} \alpha_i(j)} + f_{\boldsymbol{\beta}} \mathbf{x}^{\boldsymbol{\beta}} x_0^{2d - \sum_{i=1}^{n} \beta_i},$$

where $2d = \deg(f)$. The nonnegativity of f and $\overline{f}$ depends only on the corresponding circuit numbers Θ_f and $\Theta_{\overline{f}}$, see Theorem 2.4.4. Due to equality of the coefficients of the outer terms of f and $\overline{f}$, the circuit numbers may only differ by the barycentric coordinates λ_j and $\overline{\lambda}_j$, see (2.4.2). Recall that the barycentric coordinates are given by the convex combination of the interior point in terms of the vertices, thus by a system of n resp. $n+1$ linear equations in r unknowns:

$$\boldsymbol{\beta} = (\beta_1, \ldots, \beta_n) = \sum_{j=0}^{r} \lambda_j \boldsymbol{\alpha}(j) = \lambda_0(\alpha_1(0), \ldots, \alpha_n(0)) + \cdots + \lambda_r(\alpha_1(r), \ldots, \alpha_n(r)) ,$$

$$\left(2d - \sum_{i=1}^{n} \beta_i, \boldsymbol{\beta}\right) = \overline{\lambda}_0 \left(2d - \sum_{i=1}^{n} \alpha_i(0), \alpha_1(0), \ldots, \alpha_n(0)\right) + \cdots + \overline{\lambda}_r \left(2d - \sum_{i=1}^{n} \alpha_i(r), \alpha_1(r), \ldots, \alpha_n(r)\right).$$

Obviously, the row given by the homogenization is linearly dependent on the other rows, thereby we get $\lambda_j = \overline{\lambda}_j$ for all $j = 0, \ldots, r$, and on account of this $\Theta_f = \Theta_{\overline{f}}$. So we conclude that f is a nonnegative circuit polynomial if and only if $\overline{f}$ is a nonnegative circuit form. □

3.2 Real Zeros of SONCs

In this section we present a complete classification of the real zeros of SONC polynomials as well as SONC forms. To study real zeros of polynomials is an interesting research subject itself with a long and rich history and is especially helpful for polynomials with certain properties like nonnegativity. Recall that the real zeros of nonnegative polynomials and SOS are in particular used to study the set theoretic difference $P_{n,2d} \backslash \Sigma_{n,2d}$ and to construct explicit examples of nonnegative non-SOS forms. This started with Hilbert's seminal paper, see [Hil88], which influenced many works on this subject, e.g., [CL77, Rez00, KS18, XY17]. In [CKLR82] relations between elements of $\mathcal{E}\Sigma_{n,2d}$ and $\mathcal{E}P_{n,2d}$ are studied and by means of the presence of certain zeros (so called "transversal zeros") all pairs $(n, 2d)$ are determined such that the former set is contained in the latter, see Section 2.2.3. Real zeros of nonnegative biquadratic forms are analyzed in particular in [Qua15, BŠ16]. Finally, we want to mention [CLR80], where Choi, Lam, and Reznick discuss nonnegative forms and properties of their real zeros, including their number. Motivated by this, we take a closer look on the number of real zeros of SONC polynomials and forms.

A first result for an upper bound of affine real zeros for a nonnegative circuit polynomial having a constant term is given by Iliman and de Wolff [IdW16a, Corollary 3.9] as a corollary of Theorem 2.4.4. In this thesis, we study a more general case of circuit polynomials, which do not need to have a constant term. If we want to emphasize that a circuit polynomial does not have a constant term, we refer to a *non-constant term* circuit polynomial. For such a polynomial certainly there appears one more zero, namely the origin, and in some cases, i.e., if every outer term of the polynomial contains the variable x_i, there are infinitely many zeros in addition. But those additional zeros are zeros on the coordinate hyperplanes, being special in the sense that they are not invariant with respect to shifting the polynomial. Loosely speaking, we can shift every Newton polytope such that one vertex is located at the origin and thus a general circuit polynomial can be transformed into a circuit polynomial with constant term by shifting. To be more specific, we consider Laurent polynomials. Then we can write a non-constant term circuit polynomial f as an irreducible product

$$f \;=\; \mathbf{x}^{-\boldsymbol{\alpha}(j)} \cdot f^c,$$

with some exponent $\boldsymbol{\alpha}(j)$ and f^c being a constant term circuit polynomial. Hence, the original upper bound is valid on the real locus of the one nontrivial irreducible component, if we omit the obvious redundant part of the circuit polynomial. Thus, we often reduce ourselves to observe the zeros on $(\mathbb{R}^*)^n$, if we are only interested in a finite zero set. For convenience we define

$$\mathcal{V}^*(f) \;=\; \{(a_1, \ldots, a_n) \in (\mathbb{R}^*)^n : \; f(a_1, \ldots, a_n) = 0\}.$$

Now we state a slightly modified version of the upper bound of affine real zeros for a nonnegative circuit polynomial due to the above consideration.

Corollary 3.2.1. *A nonnegative circuit polynomial $f \in \partial C_{n,2d}$ has at most 2^n affine real zeros $\mathbf{v}$ in $(\mathbb{R}^*)^n$ all of which only differ in the signs of their entries. Therefore, every entry of $\mathbf{v}$ has the same norm, i.e., the zeros are of the form $(\pm v_1, \pm v_2, \ldots, \pm v_n)$.*

Moreover the boundary condition of Corollary 2.4.5 can be generalized to:

Corollary 3.2.2. *A proper circuit polynomial f is located on the boundary of the cone of nonnegative polynomials, i.e., $f \in \partial P_{n,2d}$, if and only if $f_{\boldsymbol{\beta}} \in \{\pm\Theta_f\}$ and $\boldsymbol{\beta} \notin (2\mathbb{N})^n$ or $f_{\boldsymbol{\beta}} = -\Theta_f$ and $\boldsymbol{\beta} \in (2\mathbb{N})^n$.*

In fact, by combinatorial arguments we can refine the observation of an upper bound for the number of zeros and state the exact number of zeros in dependence of the exponent of the inner term. Recall that we always count distinct zeros.

Theorem 3.2.3. *The number of affine real zeros* $\mathbf{v}$ *in* $(\mathbb{R}^*)^n$ *of a proper nonnegative circuit polynomial* $f \in \partial C_{n,2d}$ *is* 2^n *if* $\boldsymbol{\beta} \in (2\mathbb{N})^n$ *and* 2^{n-1} *if* $\boldsymbol{\beta} \notin (2\mathbb{N})^n$.

Proof. First note that $f \in \partial C_{n,2d}$ if and only if the coefficient of the inner term equals the circuit number, see Corollary 3.2.2. More specifically, if and only if $|f_{\boldsymbol{\beta}}| = \Theta_f$ for $\boldsymbol{\beta} \notin (2\mathbb{N})^n$ and $f_{\boldsymbol{\beta}} = -\Theta_f$ for $\boldsymbol{\beta} \in (2\mathbb{N})^n$. For $\boldsymbol{\beta} \notin (2\mathbb{N})^n$ we can assume without loss of generality that $f_{\boldsymbol{\beta}} = -\Theta_f$ after a possible transformation of variables $x_i \mapsto -x_i$. Thus, we consider a nonnegative circuit polynomial f of the subsequent form:

$$f(\mathbf{x}) \;=\; \sum_{j=0}^{r} f_{\boldsymbol{\alpha}(j)}\mathbf{x}^{\boldsymbol{\alpha}(j)} - \Theta_f \mathbf{x}^{\boldsymbol{\beta}}.$$

Observe that the special structure of the circuit polynomial gives rise to zeros whose entries are of the same norm, see also Corollary 3.2.1. Therefore the number of zeros only depends on the exponent of the inner term. In what follows, we denote the zeros of f by $\mathbf{v} = (v_1, \ldots, v_n)$. Clearly, if $\boldsymbol{\beta} \in (2\mathbb{N})^n$, then all components v_i of zeros of f can have both negative and positive sign. Hence, f has the maximal number of affine real zeros, namely 2^n. It remains to investigate the case $\boldsymbol{\beta} \notin (2\mathbb{N})^n$. In this case a necessary condition for f to have zeros at all is $\operatorname{sgn}(\mathbf{x}^{\boldsymbol{\beta}}) = 1$. Thus, only an even number of entries β_i may be odd. Since every $x_i^{\beta_i}$ corresponds to an entry v_i, this yields that the number of negative components of $\mathbf{v}$ has to be even. As a first step we assume that all β_i are odd, i.e., $\beta_i \notin 2\mathbb{N}$ for all $i = 1, \ldots, n$. Hence the number of affine real zeros is given by the number of possibilities such that for the zero $\mathbf{v}$ an even number of entries v_i is negative. Let k be the number of negative v_i, then the number of affine real zeros is given by the following basic calculation:

$$\sum_{\substack{k=0 \\ k \text{ even}}}^{n} \binom{n}{k} \;=\; 2^{n-1}.$$

As a second step we suppose that exactly one β_i is odd, without loss of generality let β_1 be odd and β_i be even for $i = 2, \ldots, n$. Obviously, v_1 has to be positive and all other entries v_i may be both positive and negative. Thus, this case is equivalent to the case of $n-1$ variables with $\boldsymbol{\beta} \in (2\mathbb{N})^{n-1}$. As seen above, in this case we have 2^{n-1} zeros.
Now we look at the general case of $\boldsymbol{\beta} \notin (2\mathbb{N})^n$. We may assume after a possible renumbering of variables that $\beta_1, \ldots, \beta_s \notin 2\mathbb{N}$ and $\beta_{s+1}, \ldots, \beta_n \in 2\mathbb{N}$. As seen in the second step the signs of $v_{s+1}, \ldots, v_n$ can be chosen positive as well as negative. Consequently, there are 2^{n-s} possibilities to choose their sign. However, for the entries $v_1, \ldots, v_s$

only an odd number of entries is allowed to be negative. According to the first step, there are 2^{s-1} such possibilities. Taking together these two numbers we get in total $2^{s-1} \cdot 2^{n-s} = 2^{n-1}$ numbers of affine real zeros for the case $\boldsymbol{\beta} \notin (2\mathbb{N})^n$. $\square$

This specific number of zeros for nonnegative circuit polynomials yields the following number of zeros for SONC polynomials.

Corollary 3.2.4. *Let $p \in \partial C_{n,2d} \cap \partial P_{n,2d}$ and $p = \sum_{i=1}^{k} f_i$, where f_i are proper nonnegative circuit polynomials for all i with corresponding inner exponent $\boldsymbol{\beta}^{(i)}$. Then $|\mathcal{V}^*(p)| = 2^n$ if $\boldsymbol{\beta}^{(i)} \in (2\mathbb{N})^n$ for all $i = 1, \ldots, k$ and $1 \leqslant |\mathcal{V}^*(p)| \leqslant 2^{n-1}$, with $|\mathcal{V}^*(p)| \mid 2^{n-1}$ otherwise. In particular, if every j-th entry, $j = 1, \ldots, n$, of each $\boldsymbol{\beta}^{(i)}$ coincides in whether $\beta_j^{(i)}$ is even or odd, then $|\mathcal{V}^*(p)| = 2^{n-1}$.*

Proof. The statement follows by the fact that a SONC polynomial p is a sum of nonnegative circuit polynomials f_i. Hence, p is zero if and only if every summand f_i is zero. Therefore, the assertion follows by counting the common zeros of the f_i. Due to the special structure of the zeros of a circuit polynomial it holds that if two nonnegative circuit polynomials f_i and f_l, $i \neq l$, have more than one zero in common, they have an even number of zeros in common. If both circuit polynomials f_i and f_l have an even inner exponent and one zero in common, then they have all their 2^n zeros in common. Analogously, if both circuit polynomials have an odd inner exponent where each entry $\beta_j^{(i)}$ and $\beta_j^{(l)}$ coincide whether it is odd or even, then the zero set of f_i and f_l is identical, i.e., both polynomials have all their 2^{n-1} zeros in common. $\square$

The subsequent example illustrates the last case of the above Corollary 3.2.4.

Example 3.2.5. Consider the bivariate SONC polynomial $p = f_1 + f_2$ with the following two nonnegative circuit polynomials

$$
\begin{aligned}
f_1 &= 3/8 + 3/8 \cdot x_1^4 + 1/4 \cdot x_1^2 x_2^4 - x_1^2 x_2, \\
f_2 &= 1/8 + 1/2 \cdot x_1^4 + 3/8 \cdot x_2^8 - x_1^2 x_2^3.
\end{aligned}
$$

Obviously, both $\beta_1^{(1)} = 2$ and $\beta_1^{(2)} = 2$ are even and both $\beta_2^{(1)} = 1$ and $\beta_2^{(2)} = 3$ are odd. Therefore, we have $\mathcal{V}(f_1) = \mathcal{V}(f_2) = \mathcal{V}(p) = \{(1,1), (-1,1)\}$ and hence the number of zeros is $|\mathcal{V}(p)| = 2^{2-1} = 2$. $\bigcirc$

If we take also infinitely many zeros and sums of monomial squares into account we get the following result, which is a direct conclusion of the observations at the beginning of this section and Corollary 3.2.4.

Corollary 3.2.6. *Let $p \in \partial C_{n,2d} \cap \partial P_{n,2d}$ be given as in Corollary* 3.2.4. *Generally, it is possible to have $|\mathcal{V}(p)| = \infty$. If $\mathcal{V}(p)$ is finite, then $|\mathcal{V}(p)| = 2^n$ or $|\mathcal{V}(p)| = 2^n + 1$ if $\boldsymbol{\beta}^{(i)} \in (2\mathbb{N})^n$ for all $i = 1, \ldots, k$, and otherwise $1 \leqslant |\mathcal{V}(p)| \leqslant 2^{n-1} + 1$.*

At this point we insert a brief discussion on the determination of the zeros.
In [IdW16a] it is shown that the zeros $\mathbf{v} \in \mathbb{R}^n$ of the nonnegative circuit polynomial $f \in C_{n,2d}$ with $\boldsymbol{\alpha}(0) = 0$ and $f_0 = \lambda_0$, i.e., $f = \lambda_0 + \sum_{j=1}^n f_j \mathbf{x}^{\boldsymbol{\alpha}(j)} - \Theta_f \mathbf{x}^{\boldsymbol{\beta}}$, satisfy $|v_j| = e^{s_j^*}$ for all $j = 1, \ldots, n$, where $\mathbf{s}^* \in \mathbb{R}^n$ is the unique vector satisfying $e^{\langle \mathbf{s}^*, \boldsymbol{\alpha}(j) \rangle} = \frac{\lambda_j}{f_j}$ for all j. Thus, $\mathbf{s}^*$ is given by a linear system of equations.
In the specific case that $\mathrm{New}(f) = \Delta_{n,2d}$, it is even possible to exactly specify the zeros of f. To be more precise, consider the nonnegative circuit polynomial $f \in C_{n,2d}$ with $f = \lambda_0 + \sum_{j=1}^n f_j x_j^{2d} - \Theta_f \mathbf{x}^{\boldsymbol{\beta}}$. Then every entry v_j of every zero $\mathbf{v} \in \mathbb{R}^n$ of this polynomial is given by $|v_j| = (\lambda_j / f_j)^{1/(2d)}$, see also Lemma 3.4.1.
An interesting question is, whether it is possible to determine the zeros of nonnegative circuit polynomials in the general case.

We now continue the study of the numbers of zeros and analyze the homogeneous case. Recall that depending on the polynomial p, the zero set $\mathcal{V}(\overline{p})$ may have no, finitely many, or infinitely many additional zeros at infinity. Therefore, we now address ourselves to the task of determining the number of real zeros additionally appearing due to homogenizing SONC polynomials.
In the affine case we were mainly interested in finitely many zeros, i.e., zeros in $(\mathbb{R}^*)^n$, which corresponds to the investigation of constant term circuit polynomials. In what follows, we distinguish between SONC forms which arise out of homogenization of a constant term SONC polynomial and those arising from homogenizing a non-constant term SONC polynomial.
To begin with, we consider nonnegative circuit polynomials and their homogenizations. The following result on the additional zeros in the homogeneous case will be explained in great detail and clearly described including specific representatives in which the specific number of zeros arises. Note that by properness, see Lemma 3.1.4, not all possibilities of numbers of zeros are realizable for every degree, see also the comments in the proof below. Let $\overline{C}_{n+1,2d}$ denote the cone of SONC forms in $\mathbb{R}[x_0, \mathbf{x}]_{n+1,2d}$.

Theorem 3.2.7. *Let $f \in \partial C_{n,2d}$ be a nonnegative circuit polynomial and $\overline{f} \in \partial \overline{C}_{n+1,2d}$ be its homogenization.*

(1) *For $n = 2$, we have:*

(i) *If f has a constant term, then f has at most 4 zeros in $\mathbb{R}^2$ and $\overline{f}$ has at most 6 zeros in $\mathbb{P}^2$.*

(ii) *If f does not have a constant term, then f has at most 5 zeros in $\mathbb{R}^2$ or infinitely many and $\overline{f}$ has at most 7 zeros or infinitely many in $\mathbb{P}^2$.*

Consequently, if $\mathcal{V}(f)$ is finite, then in both cases $\overline{f}$ has at most 2 zeros at infinity in addition.

(2) *In the general case $n > 2$ by homogenizing a constant term resp. a non-constant term circuit polynomial $f \in \partial C_{n,2d}$, $\overline{f}$ has either up to 3 resp. 2 additional zeros at infinity, or else infinitely many. Whereby we stress the fact that in contrast to the bivariate case, infinitely many additional zeros are also possible in the constant term case. That means if $\mathcal{V}(f)$ is finite, the homogenization $\overline{f}$ may have infinitely many zeros.*

Proof. Without loss of generality we assume $\mathrm{New}(f)$ is n-dimensional for $f \in \partial C_{n,2d}$. We point out that circuit polynomials with one-dimensional Newton polytope are always homogeneous.

(1) (i) The fact that f has at most 4 zeros follows directly from Corollary 3.2.1 and the assumption that we consider constant term circuit polynomials. Next we show that $\overline{f}$ has at most 6 zeros. For this, consider first a nonnegative circuit polynomial $f \in \partial C_{2,2d}$ with $\boldsymbol{\beta} \in (2\mathbb{N})^2$ and its homogenization. Note that hence f has exactly $2^2 = 4$ zeros. Therefore, f and $\overline{f}$ have the following shape:

$$\begin{aligned} f(x_1,x_2) &= f_0 + f_{\boldsymbol{\alpha}(1)} x_1^{\alpha_1(1)} x_2^{\alpha_2(1)} + f_{\boldsymbol{\alpha}(2)} x_1^{\alpha_1(2)} x_2^{\alpha_2(2)} - \Theta_f x_1^{\beta_1} x_2^{\beta_2}, \\ \overline{f}(x_0,\mathbf{x}) &= f_0 x_0^{2d} + f_{\boldsymbol{\alpha}(1)} \mathbf{x}^{\boldsymbol{\alpha}(1)} x_0^{2d-|\boldsymbol{\alpha}(1)|} + f_{\boldsymbol{\alpha}(2)} \mathbf{x}^{\boldsymbol{\alpha}(2)} x_0^{2d-|\boldsymbol{\alpha}(2)|} - \Theta_{\overline{f}} \mathbf{x}^{\boldsymbol{\beta}} x_0^{2d-|\boldsymbol{\beta}|} \,. \end{aligned}$$

Recall that we assume $\boldsymbol{\beta} \in \mathrm{int}(\mathrm{New}(f))$, where we know that both β_1 and β_2 are non-zero as well as $2d - |\boldsymbol{\beta}| \neq 0$. The zeros of f are of the form $(\pm v_1, \pm v_2)$, see Corollary 3.2.1. Thus, in the considered case $\overline{f}$ has at least 4 zeros corresponding to the zeros of the affine circuit polynomial f, which precisely are $[1 : \pm v_1 : \pm v_2]$. The number of additional zeros of $\overline{f}$ depends on the exponents of the monomials of f corresponding to the vertices of $\mathrm{New}(f)$, i.e., the outer exponents of f. In what follows, we analyze this exponent structure in detail and point out which cases may occur. Observe

that the origin is always a vertex of $\text{New}(f)$ corresponding to the constant term. Consequently, it suffices to study the exponent structure of the other two vertices $\boldsymbol{\alpha}(1)$ and $\boldsymbol{\alpha}(2)$ of $\text{New}(f)$. Also note that at least one of the vertices $\boldsymbol{\alpha}(1)$ and $\boldsymbol{\alpha}(2)$ has to be of full degree, i.e., $|\boldsymbol{\alpha}(i)| = 2d$ for at least one $i \in \{1, 2\}$ and that there are exactly three cases.
Case 1: $|\mathcal{V}(\overline{f})| = 4$, i.e., $\overline{f}$ has no additional zeros. This occurs only in the case that one vertex of $\text{New}(f)$ has full degree of the variable x_1 and the other full degree of x_2. Without loss of generality we assume $\alpha_1(1) = 2d$ and $\alpha_2(2) = 2d$. Then:

$$\begin{aligned} f &= f_0 + f_{\boldsymbol{\alpha}(1)} x_1^{2d} + f_{\boldsymbol{\alpha}(2)} x_2^{2d} - \Theta_f x_1^{\beta_1} x_2^{\beta_2}, \\ \overline{f} &= f_0 x_0^{2d} + f_{\boldsymbol{\alpha}(1)} x_1^{2d} + f_{\boldsymbol{\alpha}(2)} x_2^{2d} - \Theta_{\overline{f}} x_1^{\beta_1} x_2^{\beta_2} x_0^{2d-|\boldsymbol{\beta}|} \ . \end{aligned}$$

Obviously, $\overline{f}$ only has the zeros coming from f.
Case 2: $|\mathcal{V}(\overline{f})| = 5$. This may happen in three different exponent structures of the outer terms of f, which we discuss subsequently. If one of the vertices is of full degree in exactly one variable, we may assume $\alpha_1(1) = 2d$, and the second vertex is of one of the following forms:

(a) only the other variable x_2 occurs, but not of full degree, i.e., $\alpha_1(2) = 0$ and $\alpha_2(2) < 2d$,

(b) both variables occur but the vertex is not of full degree, hence $\alpha_1(2) \neq 0$, $\alpha_2(2) \neq 0$, and $|\boldsymbol{\alpha}(2)| < 2d$, or

(c) both variables occur and the vertex is of full degree, thus $\alpha_1(2) \neq 0$, $\alpha_2(2) \neq 0$, and $|\boldsymbol{\alpha}(2)| = 2d$.

For case (a) observe that

$$\begin{aligned} f &= f_0 + f_{\boldsymbol{\alpha}(1)} x_1^{2d} + f_{\boldsymbol{\alpha}(2)} x_2^{\alpha_2(2)} - \Theta_f x_1^{\beta_1} x_2^{\beta_2}, \\ \overline{f} &= f_0 x_0^{2d} + f_{\boldsymbol{\alpha}(1)} x_1^{2d} + f_{\boldsymbol{\alpha}(2)} x_2^{\alpha_2(2)} x_0^{2d-\alpha_2(2)} - \Theta_{\overline{f}} x_1^{\beta_1} x_2^{\beta_2} x_0^{2d-|\boldsymbol{\beta}|} \ , \end{aligned}$$

with $\alpha_2(2) < 2d$. It is easy to see that the additional zero is $[0 : 0 : 1]$. The other cases follow analogously.
Case 3: $|\mathcal{V}(\overline{f})| = 6$, i.e., $\overline{f}$ has two additional zeros. This case arises from the exponent structure that one of the vertices is full-dimensional and of full degree, without loss of generality $|\boldsymbol{\alpha}(1)| = 2d$, and the other vertex exhibits one of the following structures:

(a) only one variable x_1 or x_2 occurs, but not of full degree, hence for example $\alpha_1(2) = 0$ and $\alpha_2(2) < 2d$,

(b) both variables occur but the vertex is not of full degree, thus $\alpha_1(2) \neq 0$, $\alpha_2(2) \neq 0$, and $|\boldsymbol{\alpha}(2)| < 2d$, or

(c) both variables occur and this vertex is also of full degree, i.e., $\alpha_1(2) \neq 0$, $\alpha_2(2) \neq 0$, and $|\boldsymbol{\alpha}(2)| = 2d$.

Similarly as in case 2, we consider only case (c), the cases (a) and (b) follow analogously. For (c) the circuit polynomial and corresponding circuit form have the following shape:

$$\begin{aligned} f &= f_0 + f_{\boldsymbol{\alpha}(1)}\mathbf{x}^{\boldsymbol{\alpha}(1)} + f_{\boldsymbol{\alpha}(2)}\mathbf{x}^{\boldsymbol{\alpha}(2)} - \Theta_f \mathbf{x}^{\boldsymbol{\beta}}, \\ \overline{f} &= f_0 x_0^{2d} + f_{\boldsymbol{\alpha}(1)}\mathbf{x}^{\boldsymbol{\alpha}(1)} + f_{\boldsymbol{\alpha}(2)}\mathbf{x}^{\boldsymbol{\alpha}(2)} - \Theta_{\overline{f}} \mathbf{x}^{\boldsymbol{\beta}} x_0^{2d-|\boldsymbol{\beta}|} . \end{aligned}$$

It is easily traceable that the additional zeros are $[0 : 0 : 1]$ and $[0 : 1 : 0]$. To finish the proof of (1)(i), it remains to consider a nonnegative circuit polynomial $f \in \partial C_{2,2d}$ with $\boldsymbol{\beta} \notin (2\mathbb{N})^2$ and to show that $\overline{f}$ has at most 4 zeros. However, in this case the reasoning is as for $\boldsymbol{\beta} \in (2\mathbb{N})^2$ except that now the number of zeros starts with $|\mathcal{V}(\overline{f})| = 2$, as $|\mathcal{V}(f)| = 2$ in this case. We point out one small difference for degree 4. Despite the general existence of a proper bivariate circuit polynomial of degree 4 with an odd inner exponent, there does not occur one with the exponent structure of case 3, i.e., for two additional zeros. Because such a circuit polynomial would be only a sum of monomial squares, since there is no inner term. Whereas there is no proper bivariate circuit polynomial with even inner exponent for degree 4. Thus, by homogenizing a circuit polynomial $f \in C_{2,4}$ we get at most one zero at infinity.

(ii) We now have to take also the third vertex of $\text{New}(f)$ into account and again there are exactly three cases.

Case 1: $|\mathcal{V}(\overline{f})| = |\mathcal{V}(f)|$. Analogous to (i) this case only appears if two vertices are of full degree in exactly one variable each. The last vertex must be of degree $< 2d$ but it is irrelevant if it is comprised of both or only one variable.

For both other cases we have to consider that not two vertices of $\text{New}(f)$ can be bivariate, as otherwise f would have infinitely many zeros, since the third vertex is non-constant.

Case 2: $|\mathcal{V}(\overline{f})| = |\mathcal{V}(f)|+1$, i.e., $\overline{f}$ has one zero at infinity. Here the situation is relatively similar to the constant term case. One vertex has to be of full degree in one variable and we assume $\alpha_1(0) = 2d$. The remaining two vertices have various possibilities. If in the second vertex only the other variable x_2 occurs not of full degree, than the third vertex can have all compositions also possible in (i) for case 2, namely:

(a) only the variable x_2 occurs, but not of full degree, i.e., $\alpha_1(2) = 0$ and $\alpha_2(2) < 2d$,

(b) both variables occur but the vertex is not of full degree, hence $\alpha_1(2) \neq 0$, $\alpha_2(2) \neq 0$, and $|\boldsymbol{\alpha}(2)| < 2d$, or

(c) both variables occur and the vertex is of full degree, thus $\alpha_1(2) \neq 0$, $\alpha_2(2) \neq 0$, and $|\boldsymbol{\alpha}(2)| = 2d$.

If the second vertex also consists of the variable x_1 with degree $< 2d$, then the third vertex has to be comprised of x_2 with degree $< 2d$. In those cases we get the additional zero $[0 : 0 : 1]$.

Case 3: $|\mathcal{V}(\overline{f})| = |\mathcal{V}(f)| + 2$. Finally, this case is only possible if f has the following vertex constellation:

$$f = f_{\boldsymbol{\alpha}(0)} x_1^{\alpha_1(0)} x_2^{2d-\alpha_1(0)} + f_{\boldsymbol{\alpha}(1)} x_1^{<2d} + f_{\boldsymbol{\alpha}(2)} x_2^{<2d} - \Theta_f \mathbf{x}^{\boldsymbol{\beta}},$$

with $0 < \alpha_1(0) < 2d$ and where the monomial $x_i^{<2d}$ abbreviates $x_i^{\alpha_i(i)}$ with $\alpha_i(i) < 2d$.

Like in the constant term case (i), there is one exception for degree 4 due to the existence of a proper circuit polynomial with odd inner exponent. The only proper bivariate circuit polynomial with degree 4 after possible renumbering of variables is $f = f_{\boldsymbol{\alpha}(0)} x_1^2 + f_{\boldsymbol{\alpha}(1)} x_1^4 + f_{\boldsymbol{\alpha}(2)} x_2^4 - \Theta_f x_1^2 x_2$. It is easy to see that $|\mathcal{V}(\overline{f})| = |\mathcal{V}(f)| = 3$.

Independent of whether f has a constant term or not, it is evident that if $\mathcal{V}(f)$ is finite there are not more than two zeros at infinity. Namely, the additional zeros are of the form $[0 : a_1 : a_2]$, with $a_1, a_2 \in \mathbb{R}$, not both zero. However, we can also exclude the case $a_1, a_2 > 0$. Indeed, as already mentioned, one vertex of $\text{New}(f)$ has to be of full degree, whereby the corresponding outer term need not be homogenized. As a necessary condition for the circuit form $\overline{f}$ to vanish at a point $[0 : a_1 : a_2]$, all outer terms have to vanish, because of the homogenization the inner term is zero at this point. This, however, cannot happen if a_1 and a_2

both are strictly greater than zero. Hence, in the projective space there are only two opportunities for zeros if one of the a_i has to be zero, namely $[0:0:1]$ and $[0:1:0]$.

(2) First of all we show that even a homogenized constant term nonnegative circuit polynomial $\overline{f} \in \partial\overline{C}_{n+1,2d}$ with $n \geqslant 3$ possibly has infinitely many zeros at infinity. Strictly speaking this ensues from the last considerations of part (1), namely that in the bivariate case it is not possible that $\overline{f}$ may have more than two additional zeros. To clarify, for some $\overline{f} \in \partial\overline{C}_{n+1,2d}$ the additional zeros are of the form $[0:0:a_2:a_3:\cdots:a_n]$, where we assumed without loss of generality that $a_1 = 0$. The existence of such a case can easily be seen by the following examples. Let $f \in \partial C_{n,2d}$ be of the form

$$f = f_0 + f_{\boldsymbol{\alpha}(1)}x_1^{2d} + f_{\boldsymbol{\alpha}(2)}x_1^{\alpha_1(2)}x_2^{2d-\alpha_1(2)} + \cdots + f_{\boldsymbol{\alpha}(n)}x_1^{\alpha_1(n)}x_n^{2d-\alpha_1(n)} - \Theta_f\mathbf{x}^{\boldsymbol{\beta}},$$

with $\alpha_1(i) \in 2\mathbb{N}^*$ and $\alpha_1(i) < 2d$ for $i = 2,\ldots,n$. Or

$$f = f_0 + f_{\boldsymbol{\alpha}(1)}x_1^{2d} + f_{\boldsymbol{\alpha}(2)}x_2^{<2d} + \cdots + f_{\boldsymbol{\alpha}(n)}x_n^{<2d} - \Theta_f\mathbf{x}^{\boldsymbol{\beta}}.$$

Obviously, in both examples the zeros of $\overline{f}$ are $[0:0:a_2:a_3:\cdots:a_n]$, with $(a_2,\ldots,a_n) \in \mathbb{R}^{n-1}\backslash\{\mathbf{0}\}$, and consequently $|\mathcal{V}(\overline{f})| = \infty$.

We now discuss once again the additional zeros of $\overline{f} \in \partial\overline{C}_{n+1,2d}$ compared to the zeros of $f \in \partial C_{n,2d}$. In what follows, we limit the detailed discussion to the constant term case. We only show that by homogenizing a non-constant term circuit polynomial f, the form $\overline{f}$ cannot have three zeros at infinity, the other cases follow analogously to the constant term case. We proceed as in (1) by studying the structures of the exponents of f, or rather the vertex constellations of $\mathrm{New}(f)$, where we exclude the origin from consideration by assumption. Here, there are four cases.

Case 1: $|\mathcal{V}(\overline{f})| = |\mathcal{V}(f)|$. Like in the bivariate case we have no zeros at infinity only for $f \in \partial C_{n,2d}$, where all vertices of $\mathrm{New}(f)$ are of full degree in one variable each. Hence,

$$f = f_0 + f_{\boldsymbol{\alpha}(1)}x_1^{2d} + f_{\boldsymbol{\alpha}(2)}x_2^{2d} + \cdots + f_{\boldsymbol{\alpha}(n)}x_n^{2d} - \Theta_f\mathbf{x}^{\boldsymbol{\beta}}.$$

In what follows, we confine ourselves to presenting one possible representative for each occurring case.

Case 2: $|\mathcal{V}(\overline{f})| = |\mathcal{V}(f)| + 1$. One additional zero at infinity appears if all but one vertex of $\mathrm{New}(f)$ is of full degree of one different variable and the last vertex consists either of the not yet appearing variable but not of full degree or of various variables optional if of full degree or not. For instance, consider $f \in \partial C_{n,2d}$ with

$$f = f_0 + f_{\boldsymbol{\alpha}(1)} x_1^{2d} + f_{\boldsymbol{\alpha}(2)} x_2^{2d} + \cdots + f_{\boldsymbol{\alpha}(n)} x_n^{<2d} - \Theta_f \mathbf{x}^{\boldsymbol{\beta}}.$$

By homogenizing f, we get the additional zero $[0 : 0 : \cdots : 0 : a_n]$, $a_n \in \mathbb{R}^*$.
Case 3: $|\mathcal{V}(\overline{f})| = |\mathcal{V}(f)| + 2$. For two zeros at infinity, $n-1$ of the non-origin vertices of $\mathrm{New}(f)$ have to be of full degree in one variable, whereas for the remaining two vertices there are different exponent structures possible. Representatively, we look at the following possibility

$$f = f_0 + f_{\boldsymbol{\alpha}(1)} x_1^{2d} + \cdots + f_{\boldsymbol{\alpha}(n-2)} x_{n-2}^{2d} + f_{\boldsymbol{\alpha}(n-1)} x_{n-1}^{<2d} + f_{\boldsymbol{\alpha}(n)} x_{n-1}^{\alpha} x_n^{2d-\alpha} - \Theta_f \mathbf{x}^{\boldsymbol{\beta}},$$

where $\alpha \in 2\mathbb{N}^*$ and $\alpha < 2d$. The homogenization $\overline{f}$ has the two supplementary zeros $[0 : 0 : \cdots : 0 : a_{n-1} : 0]$ and $[0 : 0 : \cdots : 0 : a_n]$ with $a_{n-1}, a_n \in \mathbb{R}^*$.
Case 4: $|\mathcal{V}(\overline{f})| = |\mathcal{V}(f)| + 3$. This case arises if three vertices of $\mathrm{New}(f)$ are of full degree of two variables, which use three variables circulant. All other vertices are of full degree of one different variable each:

$$\begin{aligned} f &= f_0 + f_{\boldsymbol{\alpha}(1)} x_1^{2d} + \cdots + f_{\boldsymbol{\alpha}(n-3)} x_{n-3}^{2d} \\ &+ f_{\boldsymbol{\alpha}(n-2)} x_{n-2}^{\alpha} x_{n-1}^{2d-\alpha} + f_{\boldsymbol{\alpha}(n-1)} x_{n-1}^{\delta} x_n^{2d-\delta} + f_{\boldsymbol{\alpha}(n)} x_n^{\gamma} x_{n-2}^{2d-\gamma} - \Theta_f \mathbf{x}^{\boldsymbol{\beta}}, \end{aligned}$$

where $\alpha, \delta, \gamma \in 2\mathbb{N}^*$ and each is strictly smaller than $2d$. The three additional zeros of $\overline{f}$ are $[0 : 0 : \cdots : 0 : a_{n-2} : 0 : 0]$, $[0 : 0 : \cdots : 0 : a_{n-1} : 0]$, and $[0 : 0 : \cdots : 0 : a_n]$, with $a_{n-2}, a_{n-1}, a_n \in \mathbb{R}^*$.
From the constellation just given, it is apparent that case 4 cannot occur if f does not have a constant term, since in this case f already has infinitely many (affine) zeros.
It remains to show that, if $\mathcal{V}(\overline{f})$ is finite, the number of additional zeros is bounded by 3, i.e., $|\mathcal{V}(\overline{f})| \geqslant |\mathcal{V}(f)| + 4$ is not possible. We proceed by contradiction to prove that we may exclude the case of 4 additional zeros. All other cases follow analogously. First note that 4 additional zeros are only possible from (affine) dimension 4 on. Thus, to show the general n-variate case, it suffices to prove that the homogenization $\overline{f}$ never gets 4 additional zeros for $f \in C_{4,2d}$. We suppose there

exists a nonnegative circuit polynomial $f \in C_{4,2d}$ such that $|\mathcal{V}(\overline{f})| = |\mathcal{V}(f)| + 4$. The additional zeros have to be $[0:1:0:0:0]$, $[0:0:1:0:0]$, $[0:0:0:1:0]$, and $[0:0:0:0:1]$. In order for such a zero set to exist, f has to be a sum of monomials, such that every term consists of a product of two variables each $x_i x_j$, $i \neq j$ and all pairings have to appear. Each summand has to be of full degree. To receive a circuit polynomial, the mentioned terms of variable pairings together with the origin form the outer terms and $x_1^{\beta_1} \cdots x_4^{\beta_4}$ forms the inner term. Hence, there are $\binom{4}{2} + 1 = 7$ outer terms, which is a contradiction to f being a circuit polynomial, because f may only have up to 5 outer terms. This completes the proof.

□

Remark 3.2.8. The considerations of the final step of the last proof are in line of the possibility to receive 3 additional zeros in the constant term case. Following the train of thought beginning with $f \in C_{3,2d}$, we get $\binom{3}{2} + 1 = 4$ outer terms. This is consonant to the number of vertices of a simplex.

A direct corollary for SONC forms can be drawn from the above arguments.

Corollary 3.2.9. *Let $p \in \partial C_{n,2d} \cap \partial P_{n,2d}$ be a SONC polynomial with $p = \sum_{i=1}^{k} f_i$, where f_i are proper nonnegative circuit polynomials for all i with corresponding inner exponent $\boldsymbol{\beta}^{(i)}$. Consider the homogenization $\overline{p} \in \partial \overline{C}_{n+1,2d} \cap \partial \overline{P}_{n+1,2d}$.*

(1) *For $n + 1 = 3$:*

(i) *If p has a constant term, then $|\mathcal{V}(p)| \leqslant |\mathcal{V}(\overline{p})| \leqslant |\mathcal{V}(p)| + 2$. More precisely, if $\boldsymbol{\beta}^{(i)} \in (2\mathbb{N})^2$ for all $i = 1, \ldots, k$, then $4 \leqslant |\mathcal{V}(\overline{p})| \leqslant 6$, if every j-th entry of each $\boldsymbol{\beta}^{(i)}$ coincides in whether $\beta_j^{(i)}$ is even or odd, then $2 \leqslant |\mathcal{V}(\overline{p})| \leqslant 4$, and otherwise $1 \leqslant |\mathcal{V}(\overline{p})| \leqslant 4$. In all three cases the given bounds are sharp and the intermediate case occurs, with the exception of degree $2d = 4$, where the upper bound for the second and the third case is 3.*

(ii) *If p does not have a constant term, then each f_i is a non-constant term circuit polynomial and it holds that either $|\mathcal{V}(p)| \leqslant |\mathcal{V}(\overline{p})| \leqslant |\mathcal{V}(p)| + 2$ or $|\mathcal{V}(\overline{p})| = \infty$. More precisely, if $\mathcal{V}(p)$ is finite then we have the following three cases: If $\boldsymbol{\beta}^{(i)} \in (2\mathbb{N})^2$ for all $i = 1, \ldots, k$, then $5 \leqslant |\mathcal{V}(\overline{p})| \leqslant 7$, if every j-th entry of each $\boldsymbol{\beta}^{(i)}$ coincides in whether $\beta_j^{(i)}$ is even or odd, then $3 \leqslant |\mathcal{V}(\overline{p})| \leqslant 5$, and otherwise $1 \leqslant |\mathcal{V}(\overline{p})| \leqslant 5$.*
Here again, in all three cases the bounds are sharp and the intermediate case occurs. The only exception is for degree 4, where the number of zeros is either 2 or 3.

(2) *If* $n+1 \geqslant 4$ *either*

(a) $|\mathcal{V}(p)| \leqslant |\mathcal{V}(\overline{p})| \leqslant |\mathcal{V}(p)| + 3$.
In particular: $2^n \leqslant |\mathcal{V}(\overline{p})| \leqslant 2^n + 3$ *if* $\boldsymbol{\beta}^{(i)} \in (2\mathbb{N})^n$ *for all* $i = 1, \ldots, k$ *and* $2^{n-1} \leqslant |\mathcal{V}(\overline{p})| \leqslant 2^{n-1} + 3$ *if every* j*-th entry of each* $\boldsymbol{\beta}^{(i)}$ *coincides in whether* $\beta_j^{(i)}$ *is even or odd, or*

(b) $|\mathcal{V}(\overline{p})| = \infty$.

The bounds of (*a*) *are sharp as well and all intermediate cases can occur.*

Remark 3.2.10. In particular, we point out that in contrast to a SONC polynomial with a constant term its homogenization may have infinitely many zeros (at infinity) even in the case $n+1 = 3$. Furthermore, the properness condition in Corollary 3.2.9 is necessary since we also take lower bounds on the zeros into account.

Before we proceed, we provide some explicit examples demonstrating the considered cases in the proof of Theorem 3.2.7.

Example 3.2.11.

(i) First we give an example for $|\mathcal{V}(\overline{f})| = |\mathcal{V}(f)|$. Let $f \in C_{2,4}$ be the following nonnegative circuit polynomial

$$f = \frac{1}{2} + x_1^4 + x_2^4 - 2x_1x_2.$$

The zeros of f are $\mathbf{v}_1 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ and $\mathbf{v}_2 = \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$. Homogenizing f yields $\overline{f} = \frac{1}{2}x_0^4 + x_1^4 + x_2^4 - 2x_1x_2x_0^2$ and $\mathcal{V}(\overline{f}) = \left\{\left[1 : \frac{1}{\sqrt{2}} : \frac{1}{\sqrt{2}}\right], \left[1 : -\frac{1}{\sqrt{2}} : -\frac{1}{\sqrt{2}}\right]\right\}$.

(ii) For $|\mathcal{V}(\overline{f})| = |\mathcal{V}(f)| + 1$ consider

$$f = \frac{1}{3} + \frac{1}{6}x_1^6 + \frac{1}{2}x_1^2x_2^4 - x_1^2x_2^2,$$

its 4 zeros are $(\pm 1, \pm 1)$. Then $\overline{f} = \frac{1}{3}x_0^6 + \frac{1}{6}x_1^6 + \frac{1}{2}x_1^2x_2^4 - x_1^2x_2^2x_0^2$, which has the additional zero at infinity $[0 : 0 : 1]$.

(iii) An example for the case $|\mathcal{V}(\overline{f})| = |\mathcal{V}(f)| + 2$ is the Motzkin polynomial

$$f = 1 + x_1^4x_2^2 + x_1^2x_2^4 - 3x_1^2x_2^2,$$

with zeros $(\pm 1, \pm 1)$. The Motzkin form $\overline{f} = x_0^6 + x_1^4 x_2^2 + x_1^2 x_2^4 - 3x_1^2 x_2^2 x_0^2$ additionally has the zeros $[0:1:0]$ and $[0:0:1]$.

(iv) The subsequent polynomial $f \in C_{3,8}$ serves as an example for $|\mathcal{V}(\overline{f})| = |\mathcal{V}(f)| + 3$:

$$f = 5 + x_1^4 x_2^4 + x_2^4 x_3^4 + x_1^4 x_3^4 - 8x_1 x_2 x_3.$$

The zero set is $\mathcal{V}(f) = \{(1,1,1), (1,-1,-1), (-1,1,-1), (-1,-1,1)\}$. The homogenization $\overline{f} = 5x_0^8 + x_1^4 x_2^4 + x_2^4 x_3^4 + x_1^4 x_3^4 - 8x_1 x_2 x_3 x_0^5$ has the following zeros in addition $[0:1:0:0]$, $[0:0:1:0]$, and $[0:0:0:1]$.

(v) Finally, we give an example for $|\mathcal{V}(\overline{f})| = \infty$ in the case that $\mathcal{V}(f)$ is finite. Via homogenizing

$$f = 1 + x_1^4 x_2^2 x_3^2 + x_1^2 x_2^4 x_3^2 + x_1^2 x_2^2 x_3^4 - 4x_1^2 x_2^2 x_3^2,$$

the form $\overline{f}$ has the zeros hereinafter in addition $[0:0:a_2:a_3]$, $[0:a_1:0:a_3]$, and $[0:a_1:a_2:0]$ with $(a_i, a_j) \in \mathbb{R}^2 \backslash \{\mathbf{0}\}, i \neq j$. Thus, $\overline{f}$ has infinitely many zeros.

○

3.2.1 Consequences of the Zero Statements

In this section we discuss resulting properties which can be deduced from the preceding section on real zeros of SONCs.

An important question for SONC polynomials (resp. forms) is whether the analog to Hilbert's 17th problem is true, see Section 2.2. That is, if every nonnegative polynomial is representable as a finite sum of nonnegative circuit polynomials of rational functions, i.e., a sum of quotients of nonnegative circuit polynomials. By means of a zero argument we are able to answer this analog question of Hilbert for SONC polynomials in the negative.

Corollary 3.2.12. *Let $f \in P_{n,2d}$. Then, in general, there are no nonnegative circuit polynomials $g_1, \ldots, g_r, h_1, \ldots, h_r \in C_{n,2d}$, and $h_j \neq 0$ for $j = 1, \ldots, r$, such that*

$$f = \sum_{j=1}^{r} \left(\frac{g_j}{h_j} \right).$$

Proof. We prove the assertion by inspecting the zero set of the left hand side and the right hand side of the equation. By Corollary 3.2.4 it holds $|\mathcal{V}^*(\sum_j g_j)| \leqslant 2^n$, thus $|\mathcal{V}^*(\sum_j (\frac{g_j}{h_j}))| \leqslant 2^n$.
For all n and d there exist polynomials $f \in P_{n,2d}$ such that $|\mathcal{V}^*(f)| \geqslant d^n$, see [CLR80, Proposition 4.1]. Hence, we have $|\mathcal{V}^*(f)| \geqslant d^n > 2^n$, for $d \geqslant 3$. □

Note that by Corollary 3.2.9 and the subsequent Theorem 3.2.14, the analog of Hilbert's 17th problem cannot be true in general in the homogeneous case as well.

As we will see in Lemma 3.4.1 the set of SONC polynomials is not closed under multiplication. Therefore, the question, if there always exists a suitable multiplier for a nonnegative polynomial to be a sum of nonnegative circuit polynomials is not equivalent to Corollary 3.2.12. However, with similar arguments as above one can show, that the multiplier question can also be answered in the negative.

In [CLR80] Choi, Lam, and Reznick considered the numbers $B_{n+1,2d}$ and $B'_{n+1,2d}$, where $B_{n+1,2d}$ (resp. $B'_{n+1,2d}$) is defined as $\sup|\mathcal{V}(\overline{p})|$, where $\overline{p}$ ranges over all forms in $\overline{P}_{n+1,2d}$ (resp. in $\overline{\Sigma}_{n+1,2d}$) with $\sup|\mathcal{V}(\overline{p})| < \infty$, see also Section 2.2.3. They noticed that the determination of these numbers is quite challenging and presented some partial results. Moreover, they observed that for general n and d it is unclear if $B_{n+1,2d}$ always needs to be finite. See also Theorem 2.2.26 for results in special cases.
Inspired by this, we define an analog number for SONC forms.

Definition 3.2.13.

$$B''_{n+1,2d} := \sup_{\substack{\overline{p} \in \overline{C}_{n+1,2d} \\ |\mathcal{V}(\overline{p})| < \infty}} |\mathcal{V}(\overline{p})|.$$

⬡

A crucial difference to the numbers $B_{n+1,2d}$ and $B'_{n+1,2d}$ is that in our case such a number $B''_{n+1,2d}$ is *always finite* and actually can be given *explicitly.*

Theorem 3.2.14. *Let $B''_{n+1,2d}$ be defined as above, then:*

(1) *Special case $d = 1$: $B''_{2,2} = 1$.*

(2) *$B''_{2,4} = 2$, $B''_{2,6} = 3$, and if $2d \geqslant 8$ we have $B''_{2,2d} = 4$.*

(3) *$B''_{3,4} = 3$ and $B''_{3,2d} = 7$ for $2d \geqslant 6$.*

(4) *For all $n+1 \geqslant 4$: $B''_{n+1,2d} = 2^{2d-2} + 3$ for $2d < n+1$, $B''_{n+1,2d} = 2^{n-1} + 3$ if $n+1 \leqslant 2d < 2(n+1)$, and $B''_{n+1,2d} = 2^n + 3$ for $2(n+1) \leqslant 2d$.*

Proof.

(1) For $d = 1$ we have a special case, since the only possibility for a proper SONC form of degree 2 is the circuit form $\overline{f} = f_{\boldsymbol{\alpha}(0)}x_0^2 + f_{\boldsymbol{\alpha}(1)}x_1^2 - \Theta_f x_0 x_1$, which has only one zero $[1:1]$. Even if we consider a sum of monomial squares, the only zero in the case of degree 2 would be $[0:0] \notin \mathbb{P}^2$.

(2) First, note that the maximum number of zeros of a sum of monomial squares in the case $n+1 = 2$ is 2, namely the single monomial square $x_0^2 x_1^2$ has the zeros $[0:1]$ and $[1:0]$. If we consider proper SONC forms, then the number of zeros depends on the degree, because certain vertex constellations are only possible from a certain degree on. For $2d = 4$ we have, up to renumbering of the variables, only two possible circuit forms $\overline{f}_1 = f_{\boldsymbol{\alpha}(0)}x_0^4 + f_{\boldsymbol{\alpha}(1)}x_1^4 - \Theta_{f_1}x_0^2x_1^2$ with zeros $[1:1], [1:-1]$ and $\overline{f}_2 = f_{\boldsymbol{\alpha}(0)}x_0^2x_1^2 + f_{\boldsymbol{\alpha}(1)}x_0^4 - \Theta_{f_2}x_0^3x_1$ with zeros $[1:1], [0:1]$. Therefore, the first assertion in (2) holds. The second follows by the observation that for $2d = 6$ there exists a circuit form the first time, for which one outer term consists of both variables and the inner term has an even exponent: $\overline{f} = f_{\boldsymbol{\alpha}(0)}x_0^2x_1^4 + f_{\boldsymbol{\alpha}(1)}x_0^6 - \Theta_f x_0^4x_1^2$. This gives the zeros $[1:1], [1:-1]$, and $[0:1]$. Lastly, if the degree is greater or equal than 8, a circuit form with even inner exponent exists, for which both outer terms consist of both variables: $\overline{f} = f_{\boldsymbol{\alpha}(0)}x_0^2x_1^6 + f_{\boldsymbol{\alpha}(1)}x_0^6x_1^2 - \Theta_f x_0^4x_1^4$. It has the four zeros $[1:1], [1:-1], [0:1]$, and $[1:0]$. Obviously, a bivariate SONC form cannot have more than 4 zeros.

(3) Observe that in the case of $n+1 = 3$ the maximum number of zeros by a sum of monomial squares is 3. More precisely consider the following sum of monomial squares without loss of generality in degree 4: $\overline{m} = x_0^2x_1^2 + x_1^2x_2^2 + x_0^2x_2^2$. Obviously, $[1:0:0], [0:1:0]$, and $[0:0:1]$ are the zeros of $\overline{m}$.
For reasons of realizability, see Lemma 3.1.4, a proper circuit form of degree 4 has an odd inner exponent, and for $2d \geqslant 6$, also a proper circuit form with even inner exponent is possible. Thus, the statements follow immediately by Corollary 3.2.9 (1) and the preliminary consideration.

(4) The last two assertions are a direct result of Lemma 3.1.4 and Corollary 3.2.9 (2). Note that as for $n+1 = 3$, the maximum number of zeros by a sum of monomial squares is 3. In the case $2d < n+1$ there exists no proper circuit form. Though there are SONC forms $\overline{p}$, which consist of a sum of a proper circuit form $\overline{f}$ and a sum of monomial squares. We know that a proper circuit form with odd inner exponent exists if the number of variables is equal to $2d$. Hence, if $2d < n+1$ we

have the following SONC form:

$$\overline{p} = \overline{f} + x_{2d+1}^{2d} + \cdots + x_n^{2d},$$

where $\overline{f}$ is a $2d$-variate proper circuit form. Thus, $|\mathcal{V}(\overline{p})| = |\mathcal{V}(\overline{f})|$, which leads to the equality $B''_{n+1,2d} = B''_{2d,2d}$. By Corollary 3.2.9 (2), with $n+1 = 2d$, it follows $B''_{2d,2d} = 2^{(2d-1)-1} + 3$. □

The following example serves to illustrate the considerations of the case (4) for $2d < n+1$ in the proof above.

Example 3.2.15. We want to verify the calculation $B''_{7,4} = 2^{4-2} + 3 = 7$. Let $\overline{p} \in \partial\overline{P}_{7,4}$ be a SONC polynomial. Obviously $4 < 7$, therefore we search a proper 4-variate circuit form. Consider for instance

$$\overline{f} = \frac{1}{4}x_0^2x_1^2 + \frac{1}{4}x_1^2x_2^2 + \frac{1}{4}x_2^2x_0^2 + \frac{1}{4}x_3^4 - x_0x_1x_2x_3.$$

The zero set of this form is

$$\mathcal{V}(\overline{f}) = \{[1:1:1:1], [1:1:-1:-1], [1:-1:1:-1], [-1:1:1:-1], \\ [1:0:0:0], [0:1:0:0], [0:0:1:0]\}.$$

Hence, $|\mathcal{V}(\overline{f})| = 7$. Thus, the SONC form $\overline{p}$,

$$\overline{p} = \frac{1}{4}x_0^2x_1^2 + \frac{1}{4}x_1^2x_2^2 + \frac{1}{4}x_2^2x_0^2 + \frac{1}{4}x_3^4 - x_0x_1x_2x_3 + x_4^4 + x_5^4 + x_6^4,$$

has the same number of zeros as $\overline{f}$, namely 7. ○

We conclude this section with a short comparison of the "B-numbers" of the different cones $\overline{P}_{n+1,2d}$, $\overline{\Sigma}_{n+1,2d}$, and $\overline{C}_{n+1,2d}$.

Remark 3.2.16.

(i) First note that $B_{2,2} = B'_{2,2} = B''_{2,2} = 1$, which is in line with the fact, that for $(n+1, 2d) = (2,2)$ the three cones coincide, see Theorem 3.1.2.

(ii) In the bivariate case one has $B_{2,2d} = B'_{2,2d} = d$, which equals $B''_{2,2d}$ for $2d \leqslant 8$. Therefore, we have a first difference in the number of real zeros for degree 10.

(iii) In the case $n + 1 = 3$ we have the following observations, $B_{3,4} = B'_{3,4} = B''_{3,4} = 4$. But from degree 6 on, there are differences in the numbers of zeros: $B_{3,6} = 10$, $B'_{3,6} = 9$, and $B''_{3,6} = 7$.

(iv) Finally, we take a look at $n+1 = 4$. Here, we already have differences for quartics: $B_{4,4} = 10$, $B'_{4,4} = 8$, and $B''_{4,4} = 7$.

3.3 Exposed Faces of the SONC Cone in Small Dimension and Dimension Bounds

The aim of this section is to provide a first approach to the study of the exposed faces of $C_{n,2d}$. Understanding the facial structure of the cone as well as the relationship between $P_{n,2d}$ and $C_{n,2d}$ is interesting from many perspectives in both pure and applied real algebraic geometry. Unfortunately, even for the cones $P_{n,2d}$ and $\Sigma_{n,2d}$ this is still an active area of research, which is not yet well understood, see Section 2.2.3. Building upon the results of the real zeros of Section 3.2 we analyze the dimensions of the exposed faces of the SONC cone and compare those with the exposed faces of the nonnegativity cone.

First we provide a brief theoretical overview of exposed faces, where we also recall results for the exposed faces of $P_{n,2d}$ and $\Sigma_{n,2d}$. Afterwards, we derive estimates for the dimensions of the exposed faces of $C_{n,2d}$ and study some first special cases in small dimension, leading to interesting directions for further research.

In what follows, we restrict ourselves to the affine case again.
Recall from Section 2.1 that given a convex set $S \subset \mathbb{R}^n$, a face F of S is exposed if there exists a nontrivial supporting hyperplane H with $F = S \cap H$. Let Γ be a finite set of points in $\mathbb{R}^n$. The polynomials in $C_{n,2d}$ vanishing at all points of Γ form an *exposed face* of $C_{n,2d}$, which we define as $C_{n,2d}(\Gamma)$:

$$C_{n,2d}(\Gamma) \;=\; \{p \in C_{n,2d} : \; p(\mathbf{s}) = 0 \text{ for all } \mathbf{s} \in \Gamma\}.$$

Analogously, let $P_{n,2d}(\Gamma)$ and $\Sigma_{n,2d}(\Gamma)$ denote the exposed faces of $P_{n,2d}$ and $\Sigma_{n,2d}$ respectively, i.e., $P_{n,2d}(\Gamma)$ and $\Sigma_{n,2d}(\Gamma)$ are the sets of all polynomials in $P_{n,2d}$ and $\Sigma_{n,2d}$, resp., that vanish at all points of Γ. In fact, any exposed face of $P_{n,2d}$ has this description, see [BPT13]. We start with collecting some observations for $P_{n,2d}(\Gamma)$ and $\Sigma_{n,2d}(\Gamma)$. For this, following [Rez07], we denote by $I(\Gamma)_{r,2d}$ the vector space of those polynomials

$p \in \mathbb{R}[\mathbf{x}]_{n,2d}$, which have an r-th order zero at each $\mathbf{s} \in \Gamma$. Then,

$$\begin{aligned} I(\Gamma)_{1,d} &= \{p \in \mathbb{R}[\mathbf{x}]_{n,d} : \ p(\mathbf{s}) = 0 \text{ for all } \mathbf{s} \in \Gamma\}, \\ I(\Gamma)_{2,2d} &= \{p \in \mathbb{R}[\mathbf{x}]_{n,2d} : \ \nabla p(\mathbf{s}) = 0 \text{ for all } \mathbf{s} \in \Gamma\}. \end{aligned}$$

Without degree bounds $I(\Gamma)_1$ is the vanishing ideal of Γ, and $I(\Gamma)_2$ the second symbolic power of $I(\Gamma)_1$.

Clearly, $P_{n,2d}(\Gamma) \subset I(\Gamma)_{2,2d}$, since for nonnegative polynomials p zeros are local minima, which implies that the gradient of p at the zeros must vanish as well. Whereas for the set of exposed faces of the SOS cone we have $\Sigma_{n,2d}(\Gamma) \subset I(\Gamma)^2_{1,d}$, where $I(\Gamma)^2_{1,d} = \{\sum_i \alpha_i f_i g_i : f_i, g_i \in I(\Gamma)_{1,d}, \alpha_i \in \mathbb{R}\} = \{\sum_i \alpha_i h_i^2 : h_i \in I(\Gamma)_{1,d}, \alpha_i \in \mathbb{R}\}$. Actually, one can show that this inclusion is full-dimensional, i.e., $\dim(\Sigma_{n,2d}(\Gamma)) = \dim(I(\Gamma)^2_{1,d})$. Obviously it holds $I(\Gamma)^2_{1,d} \subseteq I(\Gamma)_{2,2d}$. Therefore subsequent questions concern the full-dimensionality of $P_{n,2d}(\Gamma)$ in $I(\Gamma)_{2,2d}$ and then, the equality of $I(\Gamma)_{2,2d}$ and $I(\Gamma)^2_{1,d}$. These questions were discussed in [BIK15], where the authors showed that $\dim(P_{n,2d}(\Gamma)) = \dim(I(\Gamma)_{2,2d})$ under some assumptions on the set Γ, namely, if Γ is "d-independent". Moreover, they provided an answer for the second question again under some assumptions on the set Γ and characterized those cases, where $\dim(I(\Gamma)_{2,2d})$ is strictly greater than $\dim(I(\Gamma)^2_{1,d})$.

In what follows, we use the subsequent observation for the computation of $\dim(P_{n,2d}(\Gamma))$. Since in n variables a second order zero imposes $n+1$ linear conditions which not necessarily are all independent, it holds

$$\dim(I(\Gamma)_{2,2d}) \ \geqslant \ \dim(\mathbb{R}[\mathbf{x}]_{n,2d}) - |\Gamma| \cdot (n+1) \ = \ \binom{n+2d}{2d} - |\Gamma| \cdot (n+1).$$

By the Alexander-Hirschowitz-Theorem [Mir99] it follows that generically, with exception of $2d = 2$, we have equality in the above inequality.

We now analyze $C_{n,2d}(\Gamma)$. Clearly, we have $C_{n,2d}(\Gamma) \subseteq P_{n,2d}(\Gamma)$. Immediate subsequent questions are: What is the dimension of $C_{n,2d}(\Gamma)$? Are there cases where $C_{n,2d}(\Gamma)$ is full-dimensional in $P_{n,2d}(\Gamma)$, i.e., $\dim(C_{n,2d}(\Gamma)) = \dim(P_{n,2d}(\Gamma))$?

In what follows, we consider $|\Gamma| = 2^n$ and $|\Gamma| = 2^{n-1}$.

To begin with, we state an important observation regarding the dimension of $C_{n,2d}(\Gamma)$. Obviously, $\dim(C_{n,2d}(\Gamma))$ equals the number of linear independent SONC polynomials in $C_{n,2d}$ vanishing at all points of Γ. Hence, it suffices to study nonnegative circuit polynomials or, in fact, agiforms f, where all entries of all zeros $\mathbf{v}$ of f have norm one, i.e., $|v_i| = 1$ for all $i = 1, \dots, n$. We therefore limit our subsequent analysis to agiforms.

As a first result in this context we give an upper bound on the dimension of $C_{n,2d}(\Gamma)$.

Proposition 3.3.1.

(1) *Let* $|\Gamma| = 2^n$. *Then* $\dim(C_{n,2d}(\Gamma)) \leqslant \binom{n+d}{d}$.

(2) *Let* $|\Gamma| = 2^{n-1}$. *Then* $\dim(C_{n,2d}(\Gamma)) \leqslant \binom{n+2d}{2d}$.

Proof.

(1) Let $|\Gamma| = 2^n$ and let f be an agiform in $C_{n,2d}(\Gamma)$. By Theorem 3.2.3 we know that f must have an even inner exponent, so the complete support of f is even. As already noted above, the dimension of $C_{n,2d}(\Gamma)$ is equal to the number of linear independent agiforms in $C_{n,2d}$ vanishing at all points of Γ. Thus, $\dim(C_{n,2d}(\Gamma))$ is equivalent to the rank of the matrix $A \in \mathbb{R}^{m \times N(n,d)}$, where m is the number of all agiforms in $C_{n,2d}$ vanishing at all points of Γ and $N(n,d) = \binom{n+d}{d}$ is the number of all monomials $\mathbf{x}^{\boldsymbol{\alpha}} \in \mathbb{R}[\mathbf{x}]_{2d}$ with even exponents $\boldsymbol{\alpha}$. Since $\operatorname{rank}(A) \leqslant \min\{m, N(n,d)\}$ and $m \geqslant N(n,d)$ for $2d \geqslant 6$ we can conclude that the rank of A is at most $\binom{n+d}{d}$.

(2) Now let $|\Gamma| = 2^{n-1}$. Observe that the inner exponent of an agiform f vanishing at all points of Γ may be both even and odd. Therefore we have to take all monomials up to degree $2d$ into account, whereby the matrix A above is in $\mathbb{R}^{m \times N(n,2d)}$. The result now follows by similar arguments as for statement (1). □

Note that the dimension bound for $|\Gamma| = 2^{n-1}$ is very naive and by comparison with $P_{n,2d}(\Gamma)$ also not likely to be sharp at all. At the end of this section we give an improvement of this bound.

3.3.1 The univariate Case

In this section we study the univariate case. Since we actually may count the dimension by hand in this case, we initially compute $\dim(C_{1,2d}(\Gamma))$ for some small degree. Then, we fully determine the dimension of $C_{n,2d}(\Gamma)$ in the univariate case.

First, let $\Gamma = \{1, -1\}$, hence $|\Gamma| = 2$. There is no agiform of degree 2 vanishing at both $s \in \Gamma$, since the inner exponent cannot be even. In what follows, we determine the number of univariate linear independent agiforms vanishing at 1 and -1 for the degree $4 \leqslant 2d \leqslant 14$. We also compute $\dim(P_{1,2d}(\Gamma))$ for these cases:

2d	4	6	8	10	12	14
$\dim(C_{1,2d}(\Gamma))$	1	2	3	4	5	6
$\dim(P_{1,2d}(\Gamma))$		2	5	7	9	11

For the degree $2d = 4$ the dimension of $P_{1,4}(\Gamma)$ is omitted since it is unclear if $\dim(P_{1,4}(\Gamma)) = \dim(I(\Gamma)_{2,4})$ holds in this case. Furthermore, it can be seen that the first difference between the dimensions of the exposed faces of $C_{1,2d}$ and $P_{1,2d}$ is in degree 8.

We now look at $|\Gamma| = 1$, i.e., $\Gamma = \{1\}$. Recall that we reduce our study to the case of agiforms, for which -1 is not a zero. This leads to the following dimensions:

2d	2	4	6	8	10	12
$\dim(C_{1,2d}(\Gamma))$	1	3	5	7	9	11
$\dim(P_{1,2d}(\Gamma))$		3	5	7	9	11

Obviously, there is no dimensional gap between the dimensions of the exposed faces of $C_{1,2d}$ and $P_{1,2d}$ for the considered degrees. Recall that $2d = 2$ is one exception in the Alexander-Hirschowitz Theorem.

With these calculations in mind, we provide the result for determining the dimension of $C_{1,2d}(\Gamma)$ in the general case of degree $2d$:

Lemma 3.3.2.

(1) *For* $|\Gamma| = 2$, *it holds* $\dim(C_{1,2d}(\Gamma)) = d - 1$, *if* $d \geqslant 2$.

(2) *For* $|\Gamma| = 1$, *we have* $\dim(C_{1,2d}(\Gamma)) = 2d - 1$.

Proof.

(1) Let $|\Gamma| = 2$. It is easy to see that in this case the number of linear independent agiforms with even inner exponent equals the number of even lattice points in $\Delta_{1,2d}$ without the vertices, which is equivalent to the number of all monomials x^α in $\mathbb{R}[x]_{2d}$ with even degree $0 < \alpha < 2d$. Hence, $\dim(C_{1,2d}(\Gamma)) = \binom{n+d}{d} - 2 = d - 1$. In this case we have $d \geqslant 2$, because, as already noted, in dimension $2d = 2$ there exists no agiform with even inner exponent.

(2) Now let $|\Gamma| = 1$. Analogously to case (1) the sought number equals the number of all monomials x^α in $\mathbb{R}[x]_{2d}$ with $0 < \alpha < 2d$. Thus, it follows immediately $\dim(C_{1,2d}(\Gamma)) = \binom{n+2d}{2d} - 2 = 2d + 1 - 2 = 2d - 1$.

□

The analysis of $\dim(P_{1,2d}(\Gamma))$ for both finite sets Γ yields:

(1) For $|\Gamma| = 2$, it holds $\dim(P_{1,2d}(\Gamma)) = 2d - 3$, if $d \geqslant 3$.

(2) For $|\Gamma| = 1$, we have $\dim(P_{1,2d}(\Gamma)) = 2d - 1$, if $d \geqslant 2$.

Observe that, if $|\Gamma| = 1$ we have $\dim(P_{1,2d}(\Gamma)) = \dim(C_{1,2d}(\Gamma))$, and on the contrary for $|\Gamma| = 2$ the dimension of the exposed face $P_{1,2d}(\Gamma)$ is nearly twice as large as $\dim(C_{1,2d}(\Gamma))$.

3.3.2 The bivariate Case

Now we turn towards the bivariate case. Again we calculate the dimensions of the exposed faces $C_{2,2d}(\Gamma)$ for some degrees. Already in this case we have to limit the determination of explicit cases to the two smallest degrees because of the amount of agiforms in $C_{2,2d}$ even for low degrees.
For $|\Gamma| = 4$, namely $\Gamma = \{(1,1), (-1,-1), (1,-1), (-1,1)\}$, we have:

2d	4	6
$\dim(C_{2,2d}(\Gamma))$	3	8
$\dim(P_{2,2d}(\Gamma))$		16

Already in degree $2d = 6$ there is a noticeable dimensional difference between the exposed face of $C_{2,6}$ and $P_{2,6}$. Moreover, observe that the exposed face $C_{2,6}(\Gamma)$ contains the agiform

$$f(x_1, x_2) = \frac{1}{3} + \frac{1}{3}x_1^4x_2^2 + \frac{1}{3}x_1^2x_2^4 - x_1^2x_2^2.$$

This polynomial can easily be detected to be one third of the Motzkin polynomial m, i.e., $\frac{1}{3} \cdot m = f$. Thus, we can conclude $C_{2,6}(\Gamma) \nsubseteq \Sigma_{2,6}(\Gamma)$ which is in line with our knowledge regarding the cone containment of $C_{n,2d}$ and $\Sigma_{n,2d}$.

In the case of $|\Gamma| = 2$, i.e., $\Gamma = \{(1,1), (-1,-1)\}$, we compute the following dimensions:

2d	2	4
$\dim(C_{2,2d}(\Gamma))$	1	6
$\dim(P_{2,2d}(\Gamma))$		9

Also for $|\Gamma| = 2$ we may already detect in the first calculable case dimensional differences of the exposed faces of the analyzed cones.
During counting the agiforms f which vanish on Γ in this case, we observe that not all $\binom{2+2d}{2d}$ possible monomials $\mathbf{x}^{\boldsymbol{\alpha}} \in \mathbb{R}[\mathbf{x}]_{2d}$ appear in the support of f. Since f has to vanish on both $\mathbf{s} \in \Gamma$ the inner exponent must have a special structure. To be more precise, the degree of the inner monomial $\mathbf{x}^{\boldsymbol{\beta}}$ of an agiform vanishing on $\Gamma = \{(1,1),(-1,-1)\}$ has to be even, i.e., $|\boldsymbol{\beta}| \in 2\mathbb{N}$, but not necessarily each component of $\boldsymbol{\beta}$ has to be even, that is, it may hold that $\boldsymbol{\beta} \notin (2\mathbb{N})^n$.

Due to this observation, we can give a refined dimension bound of the exposed face $C_{2,2d}(\Gamma)$ in the case $|\Gamma| = 2$ compared to Proposition 3.3.1 (2).

Lemma 3.3.3. *For* $|\Gamma| = 2$ *it holds* $\dim(C_{2,2d}(\Gamma)) \leqslant d^2 + 2d + 1$.

Proof. This bound follows by counting the involved monomials in the support of the agiforms vanishing on $\Gamma = \{(1,1),(-1,-1)\}$, which are all monomials $\mathbf{x}^{\boldsymbol{\alpha}} \in \mathbb{R}[\mathbf{x}]_{2d}$ with $|\boldsymbol{\alpha}|$ even. □

3.3.3 Improved Dimension Bound

In case of n even we can be even more precise, which yields the following improved bound for $\dim(C_{n,2d}(\Gamma))$ if $|\Gamma| = 2^{n-1}$:

Proposition 3.3.4. *Let the number of variables n be even and* $|\Gamma| = 2^{n-1}$. *Then it holds* $\dim(C_{n,2d}(\Gamma)) \leqslant \sum_{i=0}^{d} \binom{n+2i-1}{2i}$.

Proof. The dimension of $C_{n,2d}(\Gamma)$ is bounded by the number of all monomials of degree $\leqslant 2d$ with even degree. The number of monomials $\mathbf{x}^{\boldsymbol{\alpha}} \in \mathbb{R}[\mathbf{x}]_{2d}$ which have exactly even degree $2i$, i.e., $|\boldsymbol{\alpha}| = 2i$, is given by $\binom{n+2i-1}{2i}$. Hence, the summation over all $i = 0, \ldots, d$ leads to the right number. □

This argumentation does not hold for n odd, since the inner monomial $\mathbf{x}^{\boldsymbol{\beta}}$ of an agiform vanishing on all $\mathbf{s} \in \Gamma$ with $|\Gamma| = 2^{n-1}$ may also have an odd degree. For instance, consider the 3-variate case and $\Gamma = \{(1,1,1),(1,-1,-1),(-1,1,-1),(-1,-1,1)\}$. The agiform $f = \frac{1}{4} + \frac{1}{4}x_1^4 + \frac{1}{4}x_2^4 + \frac{1}{4}x_3^4 - x_1x_2x_3$ vanishes on every $\mathbf{s} \in \Gamma$, but $|\boldsymbol{\beta}| = 3$.

Furthermore, observe that for $n \geqslant 4$ the dimension bound of Proposition 3.3.4 can be further refined. In the above sum we count monomials $\mathbf{x}^{\boldsymbol{\alpha}}$ with $\alpha_i = 0$ for some i as well. For example in the 4-variate case we also take the monomials $x_1^3 x_2$, $x_2 x_3$, or $x_2 x_3 x_4^2$ into account. But clearly, these monomials cannot be inner monomials of an agiform vanishing at all $\mathbf{s} \in \Gamma$, with $|\Gamma| = 8$.
It would be an interesting task to further improve the dimension bounds (also in the case of an odd number of variables) or actually to exactly determine the dimension of the exposed faces of $C_{n,2d}$. Moreover, the gaps between the dimensions of the exposed faces of $C_{n,2d}$ and $P_{n,2d}$ need to be explored in more detail.

3.4 Multiplicative Closedness and Full-Dimensionality

In this section we establish results on the SONC cone, which are not only of importance for themselves, but also in the context of polynomial optimization, the content of the next two chapters.

First we analyze if the set of SONC polynomials is closed under multiplication. Multiplicative closedness is a basic property of sums of squares, which especially takes effect in the application to polynomial optimization, namely for certain Positivstellensätze, see Section 2.3.4. Then, we prove that $C_{n,2d}$ is full-dimensional in the nonnegativity cone $P_{n,2d}$ for every n and d; see Theorem 3.4.3. This is a necessary condition to establish SONC polynomials as a certificate, which is useful in practice and hence of crucial importance for the subsequent chapters.

The following property of SONC polynomials stands in strong contrast to SOS polynomials.

Lemma 3.4.1. *For every $n, d \in \mathbb{N}^*$ there exist $f, g \in C_{n,2d}$ such that $f \cdot g \notin C_{n,4d}$.*

Proof. A circuit polynomial in $C_{n,2d}$ has at most 2^n affine real zeros in $(\mathbb{R}^*)^n$, which is a sharp bound for every $d \in \mathbb{N}^*$; see Corollary 3.2.1 and Theorem 3.2.3. Thus, the same holds for a SONC polynomial since it is a sum of nonnegative circuit polynomials. More precisely, if we choose a circuit polynomial $f(\mathbf{x}) = \lambda_0 + \sum_{j=1}^n f_j x_j^{2d} + f_{\boldsymbol{\beta}} \mathbf{x}^{\boldsymbol{\beta}} \in \partial C_{n,2d}$ such that $\mathrm{New}(f) = \Delta_{n,2d}$, then every entry v_j of every zero $\mathbf{v} \in \mathbb{R}^n$ of f satisfies $|v_j| = (\lambda_j / f_j)^{1/(2d)}$. Then $f(\mathbf{x})$ is nonnegative and has exactly 2^n affine zeros in $(\mathbb{R}^*)^n$ if $f_{\boldsymbol{\beta}} = -\Theta_f$ and $\boldsymbol{\beta} \in (2\mathbb{N})^n$. Therefore, for such a given $f(\mathbf{x})$ we can construct a new nonnegative circuit polynomial $g(\mathbf{x})$ with 2^n different affine zeros in $(\mathbb{R}^*)^n$ by changing every f_j by a small $\varepsilon_j \in \mathbb{R}$ and adjusting $f_{\boldsymbol{\beta}}$ to the new circuit number $-\Theta_g$; see (2.4.2).

The product $f(\mathbf{x}) \cdot g(\mathbf{x})$, a product of two SONC polynomials, is a polynomial with $2^n + 2^n = 2^{n+1}$ affine real zeros in $(\mathbb{R}^*)^n$ and of degree at most $4d$. Consequently, this product cannot be a SONC polynomial in $C_{n,4d}$. □

An immediate consequence of the proof of this lemma is the following statement:

Corollary 3.4.2. *Not every square of a polynomial is a SONC polynomial.*

These results will be of relevance in Chapter 5, where we study the application of SONC polynomials to constrained optimization problems based on a Positivstellensatz involving SONC polynomials. Namely, the observations above imply that SONC polynomials form neither a *preorder* nor a *quadratic module*; see Section 2.3.3 for the formal definitions. Hence, we cannot expect to exploit several of the classical techniques from real algebraic geometry to derive a *Putinar-like* Positivstellensatz, since these techniques rely heavily on the fact that sums of squares form both a preorder and a quadratic module. However, this does not contradict the possibility of deriving a similar result or even the exact equivalent of Putinar's Positivstellensatz for SONC polynomials. We address this topic again in Chapter 6.

We now show, that the convex cone of SONC polynomials is always full-dimensional in the convex cone of nonnegative polynomials.

Theorem 3.4.3. *Let $n, d \in \mathbb{N}^*$. Then the SONC cone $C_{n,2d}$ is full-dimensional in the cone of nonnegative polynomials $P_{n,2d}$.*

Proof. To prove the theorem it is sufficient to provide a single polynomial $f \in C_{n,2d}$ such that for every $g \in P_{n,2d}$ there exists a sufficiently small $\varepsilon > 0$ such that we have $f + \varepsilon g \in C_{n,2d}$. We choose f as follows: Let $\mathrm{New}(f) = \Delta_{n,2d}$ be the standard simplex with edge length $2d$, i.e., $V(\mathrm{New}(f)) = \{\mathbf{0}, 2d \cdot \mathbf{e}_1, \ldots, 2d \cdot \mathbf{e}_n\}$. Moreover, assume that f has full support, i.e., $\mathrm{supp}(f) = \mathcal{L}_{n,2d}$. Since f is a SONC polynomial, we can write f as a sum of nonnegative circuit polynomials $f_1, \ldots, f_s$ such that for every $j = 1, \ldots, s$ it holds that

$$f_j(\mathbf{x}) \quad = \quad f_{j,\mathbf{0}} + \sum_{i=1}^{r_j} f_{j,i} x_i^{2d} - f_{\boldsymbol{\beta}(j)} \mathbf{x}^{\boldsymbol{\beta}(j)},$$

$r_j \leqslant n$. Furthermore, we assume that every f_j is in the interior of $C_{n,2d}$, that is, $|f_{\boldsymbol{\beta}(j)}| < \Theta_{f_j}$. Thus, f is in the interior of $C_{n,2d}$, too. Let

$$\delta \quad = \quad \min_{1 \leqslant j \leqslant s} \left\{ \Theta_{f_j} - |f_{\boldsymbol{\beta}(j)}| \right\} \; > \; 0. \tag{3.4.1}$$

Let $g(\mathbf{x}) = \sum_{\boldsymbol{\alpha}\in\mathcal{L}_{n,2d}} g_{\boldsymbol{\alpha}}\mathbf{x}^{\boldsymbol{\alpha}} \in P_{n,2d}$ be arbitrary. By Proposition 2.2.2 we have $g_{\mathbf{0}} \geqslant 0$ and $g_{2d\cdot\mathbf{e}_i} \geqslant 0$ for $i = 1,\ldots,n$. For a given δ we choose

$$\varepsilon \;=\; \min_{\substack{\boldsymbol{\alpha}\in\mathcal{L}_{n,2d}\setminus V(\mathrm{New}(f)),\\ g_{\boldsymbol{\alpha}}\neq 0}} \left\{\frac{\delta}{2\cdot|g_{\boldsymbol{\alpha}}|}\right\} \;>\; 0. \tag{3.4.2}$$

Since f has full support and every nonnegative circuit polynomial f_j has exactly one inner term and satisfies $V(\mathrm{New}(f_j)) \subseteq V(\mathrm{New}(f)) = V(\Delta_{n,2d})$, the exponent $\boldsymbol{\alpha} \in \mathcal{L}_{n,2d}\setminus\{\mathbf{0}, 2d\cdot\mathbf{e}_1,\ldots,2d\cdot\mathbf{e}_n\}$ of a term in g equals the exponent $\boldsymbol{\beta}(j)$ of an inner term of exactly one nonnegative circuit polynomial f_j. Therefore, it holds that

$$f(\mathbf{x}) + \varepsilon\cdot g(\mathbf{x}) \;=\; \sum_{j=1}^{s}\left(f_j(\mathbf{x}) + \varepsilon\cdot g_{\boldsymbol{\beta}(j)}\mathbf{x}^{\boldsymbol{\beta}(j)}\right) + \varepsilon\cdot\left(g_{\mathbf{0}} + \sum_{i=1}^{n} g_{2d\cdot\mathbf{e}_i}\cdot x_i^{2d}\right) \tag{3.4.3}$$

for a suitable matching of the $g_{\boldsymbol{\alpha}}$'s of $g(\mathbf{x})$ and the $g_{\boldsymbol{\beta}(j)}$'s. For every $j = 1,\ldots,s$ we have

$$\begin{aligned} & f_j(\mathbf{x}) + \varepsilon\cdot g_{\boldsymbol{\beta}(j)}\mathbf{x}^{\boldsymbol{\beta}(j)} + \frac{\varepsilon}{s}\cdot\left(g_{\mathbf{0}} + \sum_{i=1}^{n} g_{2d\cdot\mathbf{e}_i}\cdot x_i^{2d}\right) \\ = \;& f_{j,\mathbf{0}} + \frac{\varepsilon}{s}\cdot g_{\mathbf{0}} + \sum_{i=1}^{r_j} f_{j,i}x_i^{2d} + \sum_{i=1}^{n}\frac{\varepsilon}{s}\cdot g_{2d\cdot\mathbf{e}_i}x_i^{2d} - (f_{\boldsymbol{\beta}(j)} - \varepsilon\cdot g_{\boldsymbol{\beta}(j)})\mathbf{x}^{\boldsymbol{\beta}(j)} \\ \geqslant \;& f_{j,\mathbf{0}} + \sum_{i=1}^{r_j} f_{j,i}x_i^{2d} - (f_{\boldsymbol{\beta}(j)} - \varepsilon\cdot g_{\boldsymbol{\beta}(j)})\mathbf{x}^{\boldsymbol{\beta}(j)}. \end{aligned}$$

Every polynomial $f_{j,\mathbf{0}} + \sum_{i=1}^{r_j} f_{j,i}x_i^{2d} - (f_{\boldsymbol{\beta}(j)} - \varepsilon\cdot g_{\boldsymbol{\beta}(j)})\mathbf{x}^{\boldsymbol{\beta}(j)}$ is a circuit polynomial. Hence, we can conclude that it is nonnegative if we show that the norm of the coefficient of its inner term is bounded by the corresponding circuit number. This is the case since

$$\begin{aligned} & |f_{\boldsymbol{\beta}(j)} - \varepsilon\cdot g_{\boldsymbol{\beta}(j)}| \overset{(3.4.2)}{\leqslant} |f_{\boldsymbol{\beta}(j)}| + \min_{\substack{\boldsymbol{\alpha}\in\mathcal{L}_{n,2d}\setminus V(\mathrm{New}(f)),\\ g_{\boldsymbol{\alpha}}\neq 0}} \left\{\frac{\delta}{2\cdot|g_{\boldsymbol{\alpha}}|}\right\}\cdot|g_{\boldsymbol{\beta}(j)}| \\ \leqslant \;& |f_{\boldsymbol{\beta}(j)}| + \frac{\delta}{2} \overset{(3.4.1)}{<} \Theta_{f_j}. \end{aligned}$$

Thus, for every $j = 1,\ldots,s$ we conclude that $f_j(\mathbf{x}) + \varepsilon g_{\boldsymbol{\beta}(j)}\mathbf{x}^{\boldsymbol{\beta}(j)} + \frac{\varepsilon}{s}\cdot\left(g_{\mathbf{0}} + \sum_{i=1}^{n} g_{2d\cdot\mathbf{e}_i}x_i^{2d}\right)$ is a nonnegative circuit polynomial. Hence, by (3.4.3), it follows that $f + \varepsilon\cdot g \in C_{n,2d}$.

□

3.5 Conclusion

In this chapter we have given new theoretical insights about the SONC cone. Besides showing that $C_{n,2d}$ is a proper cone, we provided the missing piece about the cone containment of the SONC and the SOS cone, and we proved the important fact that the property to be SONC is preserved under homogenization. Moreover, we gave a complete classification of the real zeros of SONC forms and polynomials yielding some interesting additional results. Using the observations of the real zeros, we took a first step towards the analysis of the exposed faces $C_{n,2d}(\Gamma)$. Finally, we observed that the set of SONC polynomials is not closed under multiplication and we provided the important result that the cone of SONC polynomials is full-dimensional in the cone of nonnegative polynomials.

The study of theoretical aspects of $C_{n,2d}$ and the relationship between the SONC cone and the cones $P_{n,2d}$ and $\Sigma_{n,2d}$ is far from being complete. In addition to the mentioned questions in the chapter, there are many future questions and open problems concerning this topic. We address these further in Chapter 6.

Chapter 4

An Approach to Polynomial Optimization via SONC and GP

In this and the next chapter we apply SONC polynomials to constrained polynomial optimization problems. A common approach to tackle these problems is Lasserre's relaxation which is based on sums of squares using semidefinite programming. In spite of the fact, that SDPs can be solved in polynomial time (up to an ε-error), e.g., [BPT13, page 41] and references therein, these programs quickly get very large in size, which often is an issue for problems with high degrees or many variables. Hence, this approach is challenging to use in practice.

Recently, Ghasemi and Marshall suggested a promising alternative approach both for (POPs) and (CPOPs) based on *geometric programming (GP)* [GM12, GM13]. GPs can be solved in polynomial time (up to an ε-error) as well [NN94]; see also [BKVH07, page 118], but, by experimental results, e.g., [BKVH07, GM12, GM13, GLM14], in practice the corresponding geometric programs can be solved *significantly* faster than their counterparts in semidefinite programming. The lower bounds obtained by Ghasemi and Marshall are, however, by construction worse than lower bounds obtained via semidefinite programming, and they can only be applied in very special cases.

In [IdW16b] Iliman and de Wolff showed that the GP-based approach for unconstrained optimization by Ghasemi and Marshall can be generalized crucially via SONC certificates. In consequence, the presented geometric programs are linked to sums of nonnegative circuit polynomials similarly as semidefinite programming relaxations are linked to sums of squares. Particularly, there exist various classes of polynomials for which the SONC/GP-based approach is not only *faster* but, it also yields *better* bounds than the SOS/SDP approach. The reason is that all certificates used by Ghasemi and

Marshall are always SOS, while SONCs are not SOS in general; see Theorem 2.4.8.

Motivated by these recent developments, we provide an extension for the SONC/GP method which yields a new approach to solve a huge class of (CPOPs), particularly for high-degree polynomials. Experimentally, the resulting method is *significantly* faster than semidefinite programming as we demonstrate in various examples.

First we introduce ST-polynomials, the considered polynomials in the predominant part of this chapter, and geometric programs, the underlying optimization problems. Afterwards we recall the SONC/GP approach for unconstrained polynomial optimization problems as well as an initial approach to the constrained case yielding a program which leads to a lower bound for the optimal value but which is not computable via geometric programming. The aim of the third section is to provide relaxations for this program such that we can use geometric programming. Moreover, we show that for certain cases the new, relaxed GP yields bounds as good as the initial non-GP program. In the fourth section we provide several examples comparing our new approach to Lasserre's relaxation. Concerning the speed of the computation we demonstrate that in all examples our program is significantly faster than semidefinite programming. The main observation is, that in sharp antagonism to SDPs, increasing the degree of a given problem does not entail significant amendments in the runtime of our program, which fits into the previously mentioned narrative that a GP-based approach is especially useful for high-degree problems, where SDP methods break down. Furthermore, the examples demonstrate that in contrast to the bounds obtained by Ghasemi and Marshall, our program can provide bounds which are better than the bounds given by the d-th Lasserre relaxation for some specific d determined by the degrees of the involved polynomials. Finally, we generalize the SONC/GP approach in the fifth section to non-ST-polynomials both in the unconstrained and in the constrained case. Again we provide some examples including a comparison of the new bounds to the ones obtained by semidefinite programming methods. Repeatedly we observe that in all examples our GP-based approach is much faster.

4.1 Preliminaries

In this section we introduce the considered ST-polynomials and geometric programming, a special type of convex optimization problems. Then we review key results on the SONC/GP-based approach to the unconstrained case.

4.1.1 ST-Polynomials

Let $f(\mathbf{x}) = \sum_{\boldsymbol{\alpha} \in A} f_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}}$ be a polynomial as defined in Section 2.1. For a given set $A \subset \mathbb{N}^n$ we define $\Delta(A) = A \backslash V(A)$ and we denote by $\Delta(f)$ the elements of $\Delta(A)$ which appear as exponents of non-zero terms, that are no monomial squares. I.e., we have

$$\Delta(f) \;=\; \{\boldsymbol{\alpha} \in \Delta(A) \;:\; |f_{\boldsymbol{\alpha}}| \neq 0 \text{ and } (f_{\boldsymbol{\alpha}} < 0 \text{ or } \boldsymbol{\alpha} \notin (2\mathbb{N})^n)\}.$$

For the remainder of this chapter, we assume that the necessary conditions in Proposition 2.2.2 are satisfied including $\mathrm{New}(f) = \mathrm{conv}(A)$. For simplicity, we denote this assumption by the symbol ($\clubsuit$) from now on.

In what follows, we consider the class of *ST-polynomials*, which can be seen as a certain generalization of circuit polynomials to polynomials supported on a simplex Newton polytope with various interior points, for further details see [IdW16b]. This class generalizes a class of polynomials considered by Fidalgo and Kovacec in [FK11] and by Ghasemi and Marshall in [GM12, GM13].

Definition 4.1.1. Let $f \in \mathbb{R}[\mathbf{x}]$ be supported on $A \subset \mathbb{N}^n$ such that ($\clubsuit$) holds. Then f is called an *ST-polynomial* if it is of the form

$$f(\mathbf{x}) \;=\; \sum_{j=0}^{r} f_{\boldsymbol{\alpha}(j)} \mathbf{x}^{\boldsymbol{\alpha}(j)} + \sum_{\boldsymbol{\beta} \in \Delta(A)} f_{\boldsymbol{\beta}} \mathbf{x}^{\boldsymbol{\beta}}, \tag{4.1.1}$$

with $r \leqslant n$, exponents $\boldsymbol{\alpha}(j)$ and $\boldsymbol{\beta}$, and coefficients $f_{\boldsymbol{\alpha}(j)}, f_{\boldsymbol{\beta}}$, for which the following conditions hold:

(ST1) $V(A) = \{\boldsymbol{\alpha}(0), \ldots, \boldsymbol{\alpha}(r)\}$ is the vertex set of an r-dimensional even simplex, coinciding with $\mathrm{New}(f) = \mathrm{conv}(A)$.

(ST2) Every exponent $\boldsymbol{\beta} \in \Delta(A)$ can be written uniquely as

$$\boldsymbol{\beta} \;=\; \sum_{j=0}^{r} \lambda_j^{(\boldsymbol{\beta})} \boldsymbol{\alpha}(j) \;\text{ with }\; \lambda_j^{(\boldsymbol{\beta})} \;\geqslant\; 0 \;\text{ and }\; \sum_{j=0}^{r} \lambda_j^{(\boldsymbol{\beta})} \;=\; 1,$$

where the $\lambda_j^{(\boldsymbol{\beta})}$ denote the barycentric coordinates of $\boldsymbol{\beta}$ relative to the vertices $\boldsymbol{\alpha}(j)$ with $j = 0, \ldots, r$.

○

The "ST" in "ST-polynomial" stands for "simplex tail". The tail part is given by the sum $\sum_{\boldsymbol{\beta} \in \Delta(A)} f_{\boldsymbol{\beta}} \mathbf{x}^{\boldsymbol{\beta}}$, while the other terms define the simplex part. Note that here we allow the barycentric coordinates $\lambda_j^{(\boldsymbol{\beta})}$ to be nonnegative instead of strictly positive as in the circuit polynomial definition, see Definition 2.4.1. Thus in this chapter, the non vertex points $\boldsymbol{\beta} \in \Delta(A)$ may be located on the boundary of the simplex. Despite this difference, we call an ST-polynomial, which has a tail part consisting of at most one term, a circuit polynomial.

However, nonnegativity of ST-polynomials is as well closely related to the circuit number. This invariant is adapted in the following way in this chapter:

Definition 4.1.2. Let f be an ST-polynomial with support set A. For every $\boldsymbol{\beta} \in \Delta(A)$ we define the corresponding *circuit number* as

$$\Theta_f(\boldsymbol{\beta}) \;=\; \prod_{j \in \mathrm{nz}(\boldsymbol{\beta})} \left(\frac{f_{\boldsymbol{\alpha}(j)}}{\lambda_j^{(\boldsymbol{\beta})}} \right)^{\lambda_j^{(\boldsymbol{\beta})}}$$

with $\mathrm{nz}(\boldsymbol{\beta}) := \{j \in \{0, \ldots, r\} : \lambda_j^{(\boldsymbol{\beta})} \neq 0\}$, $f_{\boldsymbol{\alpha}(j)}$, and $\lambda_j^{(\boldsymbol{\beta})}$ as before. ○

4.1.2 Geometric Programming

Geometric programming was introduced in [DPZ67]. It is a convex optimization problem and has applications for example in nonlinear network flow problems, optimal control, optimal location problems, chemical equilibrium problems and particularly in circuit design problems.

Definition 4.1.3. A function $p\colon \mathbb{R}^n_{>0} \to \mathbb{R}$ of the form $p(\mathbf{z}) = p(z_1, \ldots, z_n) = c z_1^{\alpha_1} \cdots z_n^{\alpha_n}$ with $c > 0$ and $\alpha_i \in \mathbb{R}$ is called a *monomial (function)*. A sum $\sum_{i=0}^{k} c_i z_1^{\alpha_1(i)} \cdots z_n^{\alpha_n(i)}$ of monomials with $c_i > 0$ is called a *posynomial (function)*.

A *geometric program* (GP) has the following form:

$$\text{(4.1.2)} \qquad \begin{cases} \text{minimize} & p_0(\mathbf{z}), \\ \text{subject to:} & \begin{array}{ll} (1) & p_i(\mathbf{z}) \leqslant 1 \;\text{ for all }\; i = 1, \ldots, m, \\ (2) & q_j(\mathbf{z}) = 1 \;\text{ for all }\; j = 1, \ldots, l, \end{array} \end{cases}$$

where $p_0, \ldots, p_m$ are posynomials and $q_1, \ldots, q_l$ are monomial functions. ○

Geometric programs can be solved with interior point methods. In [NN94], the authors prove worst-case polynomial time complexity of this method; see also [BKVH07, page 118]. A *signomial program* is given like a geometric program except that the coefficients c_i of the involved posynomials can be arbitrary real numbers.

For an introduction to geometric programming, signomial programming, and an overview about applications see [BKVH07, BV04].

4.1.3 SONC Certificates via Geometric Programming in the Unconstrained Case

In this section we recall the main results from [IdW16b] about SONC certificates obtained via geometric programming for unconstrained polynomial optimization problems. These results always require that the polynomial in the optimization problem is an ST-polynomial in the sense of Section 4.1.1.

Theorem 4.1.4. ([IdW16b, Theorems 3.4 and 3.5]) *Assume that f is an ST-polynomial as in (4.1.1) and let $k \in \mathbb{R}$. Suppose that for every $(\boldsymbol{\beta}, j) \in \Delta(f) \times \{1, \ldots, r\}$ there exists an $a_{\boldsymbol{\beta},j} \geqslant 0$, such that:*

(1) $a_{\boldsymbol{\beta},j} > 0$ *if and only if* $\lambda_j^{(\boldsymbol{\beta})} > 0$,

(2) $|f_{\boldsymbol{\beta}}| \leqslant \prod_{j \in \mathrm{nz}(\boldsymbol{\beta})} \left(\frac{a_{\boldsymbol{\beta},j}}{\lambda_j^{(\boldsymbol{\beta})}} \right)^{\lambda_j^{(\boldsymbol{\beta})}}$ *for every* $\boldsymbol{\beta} \in \Delta(f)$ *with* $\lambda_0^{(\boldsymbol{\beta})} = 0$,

(3) $f_{\boldsymbol{\alpha}(j)} \geqslant \sum_{\boldsymbol{\beta} \in \Delta(f)} a_{\boldsymbol{\beta},j}$ *for all* $j = 1, \ldots, r$,

(4) $(f_{\boldsymbol{\alpha}(0)} - k)\mathbf{x}^{\boldsymbol{\alpha}(0)} \geqslant \sum_{\substack{\boldsymbol{\beta} \in \Delta(f) \\ \lambda_0^{(\boldsymbol{\beta})} \neq 0}} \lambda_0^{(\boldsymbol{\beta})} |f_{\boldsymbol{\beta}}|^{1/\lambda_0^{(\boldsymbol{\beta})}} \prod_{\substack{j \in \mathrm{nz}(\boldsymbol{\beta}) \\ j \geqslant 1}} \left(\frac{\lambda_j^{(\boldsymbol{\beta})}}{a_{\boldsymbol{\beta},j}} \right)^{\lambda_j^{(\boldsymbol{\beta})}/\lambda_0^{(\boldsymbol{\beta})}}$.

Then $f - k\mathbf{x}^{\boldsymbol{\alpha}(0)}$ is a sum of nonnegative circuit polynomials $g_1, \ldots, g_s$ such that $s := |\Delta(f)|$, and for every g_i the Newton polytope $\mathrm{New}(g_i)$ *is a face of* $\mathrm{New}(f)$.

Let f_{sonc} be the supremum of all $k \in \mathbb{R}$ such that for every $\boldsymbol{\beta} \in \Delta(f)$ there exist nonnegative reals $a_{\boldsymbol{\beta},1}, \ldots, a_{\boldsymbol{\beta},r}$ such that the conditions (1) to (4) are satisfied. Then f_{sonc} coincides with the supremum of all $k \in \mathbb{R}$ such that there exist nonnegative circuit polynomials $g_1, g_2, \ldots, g_s$ whose Newton polytopes are faces of $\mathrm{New}(f)$ *and which satisfy $f - k\mathbf{x}^{\boldsymbol{\alpha}(0)} = \sum_{i=1}^{s} g_i$.*

For the special case of scaled standard simplices the theorem was shown earlier by Ghasemi and Marshall [GM12, Theorem 3.1]. In this special case every sum of nonnegative circuit polynomials is also a sum of binomial squares which is not true in general. For example, the Motzkin polynomial is an ST-polynomial with one interior term, which is not even a SOS.

Theorem 4.1.4 states

$$f_{\text{sonc}} \;=\; \sup\{k \in \mathbb{R} \;:\; f - k\mathbf{x}^{\boldsymbol{\alpha}(0)} \text{ is a SONC }\}.$$

The bound f_{sonc} is given by a geometric program [IdW16b, Corollary 4.2]:

Corollary 4.1.5. *Let $f \in \mathbb{R}[\mathbf{x}]$ be an ST-polynomial. Let R be the subset of an $r|\Delta(f)|$-dimensional real space given by*

$$R \;=\; \{(a_{\boldsymbol{\beta},j}) \;:\; a_{\boldsymbol{\beta},j} \in \mathbb{R}_{>0} \textit{ for every } \boldsymbol{\beta} \in \Delta(f) \textit{ and } j \in \operatorname{nz}(\boldsymbol{\beta})\}.$$

Then $f_{\text{sonc}} = f_{\boldsymbol{\alpha}(0)} - m^*$, where m^* is given as the output of the following geometric program:

$$\begin{cases} \text{minimize} & \displaystyle\sum_{\substack{\boldsymbol{\beta}\in\Delta(f)\\ \lambda_0^{(\boldsymbol{\beta})}\neq 0}} \lambda_0^{(\boldsymbol{\beta})} |f_{\boldsymbol{\beta}}|^{1/\lambda_0^{(\boldsymbol{\beta})}} \prod_{\substack{j\in\operatorname{nz}(\boldsymbol{\beta})\\ j\geqslant 1}} \left(\frac{\lambda_j^{(\boldsymbol{\beta})}}{a_{\boldsymbol{\beta},j}}\right)^{\lambda_j^{(\boldsymbol{\beta})}/\lambda_0^{(\boldsymbol{\beta})}} \quad \text{over the subset } R' \text{ of } R \\ & \\ \text{defined by:} & \begin{array}{l} (1) \quad \displaystyle\sum_{\boldsymbol{\beta}\in\Delta(f)} (a_{\boldsymbol{\beta},j}/f_{\boldsymbol{\alpha}(j)}) \leqslant 1 \;\text{ for every } j = 1,\ldots,r, \\ (2) \quad |f_{\boldsymbol{\beta}}| \displaystyle\prod_{j\in\operatorname{nz}(\boldsymbol{\beta})} \left(\frac{\lambda_j^{(\boldsymbol{\beta})}}{a_{\boldsymbol{\beta},j}}\right)^{\lambda_j^{(\boldsymbol{\beta})}} \leqslant 1 \;\text{ for every } \boldsymbol{\beta} \in \Delta(f) \text{ with } \lambda_0^{(\boldsymbol{\beta})} = 0. \end{array} \end{cases}$$

Hence, the optimal bound to find a SONC decomposition of an ST-polynomial is provided by geometric programming. Since a polynomial with a SONC decomposition is nonnegative, geometric programming can be used to find certificates of nonnegativity.

A key observation is that the bounds obtained by the new SONC approach can be better than the ones obtained by SOS as the following result shows; see [IdW16b, Corollary 3.6]. For this, remember from Section 2.3.1 that $f^* = \inf\{f(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^n\}$ and $f_{\text{sos}} = \sup\{\gamma \in \mathbb{R} : f - \gamma \text{ is SOS}\}$.

Corollary 4.1.6. *Let f be an ST-polynomial with $\Delta(A) = \Delta(f)$ such that $\Delta(f)$ is contained in the interior of* New(f). *Let $\boldsymbol{\alpha}(0)$ be the origin and suppose that there exists a vector $\mathbf{v} \in (\mathbb{R}^*)^n$ such that $f_{\boldsymbol{\alpha}} \cdot \mathbf{v}^{\boldsymbol{\alpha}} < 0$ for all $\boldsymbol{\alpha} \in \Delta(f)$. Then*

$$f_{\text{sonc}} \;=\; f^* \;\geqslant\; f_{\text{sos}}.$$

4.1.4 SONC Certificates for the Constrained Case

Based on the results of the previous subsection, Iliman and de Wolff [IdW16b, Section 5] derive some initial applications for constrained polynomial optimization problems, which were restated in a more careful way in [DIdW18]. Before deepening these results and describing our contributions in Section 4.2, we recall the refined version of these initial results.

Let $f, g_1, \ldots, g_s$ be elements of the polynomial ring $\mathbb{R}[\mathbf{x}]$ and let

$$K \;=\; \{\mathbf{x} \in \mathbb{R}^n \;:\; g_i(\mathbf{x}) \geqslant 0 \text{ for all } i = 1, \ldots, s\}$$

be the basic closed semialgebraic set defined by $g_1, \ldots, g_s$. We consider the constrained polynomial optimization problem (CPOP)

$$f_K^* \;=\; \inf_{\mathbf{x} \in K} f(\mathbf{x}) \;=\; \sup\{\gamma \in \mathbb{R} \;:\; f(\mathbf{x}) - \gamma \geqslant 0 \text{ for all } \mathbf{x} \in K\},$$

as defined in Section 2.3.4. If $s = 0$, then we have no polynomial constraints g_i and therefore $K = \mathbb{R}^n$, which leads to the global optimization problem explained in Section 4.1.3.

To obtain a general lower bound for f on K which is computable by geometric programming we replace the considered polynomials by a new function. Let

$$G(\boldsymbol{\mu})(\mathbf{x}) \;=\; f(\mathbf{x}) - \sum_{i=1}^{s} \mu_i g_i(\mathbf{x}) \;=\; -\sum_{i=0}^{s} \mu_i g_i(\mathbf{x}) \tag{4.1.3}$$

for $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_s) \in \mathbb{R}^s_{\geqslant 0}$, $g_0 = -f$ and $\mu_0 = 1$. For every fixed $\boldsymbol{\mu}^* \in \mathbb{R}^s_{\geqslant 0}$ the function $G(\mathbf{x}) = G(\boldsymbol{\mu}^*)(\mathbf{x})$ is a polynomial in $\mathbb{R}[\mathbf{x}]$. Following an argument in [GM13] we can assume that all monomial squares of $-g_i$ are vertices of $\mathrm{New}(G(\boldsymbol{\mu}))$: One can reduce to this case by neglecting all monomial squares not corresponding to such a vertex. That is, for all $i = 0, \ldots, s$ one can replace g_i by $\tilde{g}_i$, which resemble g_i without monomial squares of $-g_i$ in the interior of $\mathrm{New}(G(\boldsymbol{\mu}))$. Then $-\tilde{g}_i \leqslant -g_i$ on $\mathbb{R}^n$ for $i = 0, \ldots, s$, thus, $K \subseteq \tilde{K}$, where $\tilde{K} = \{\mathbf{x} \in \mathbb{R}^n : \tilde{g}_i(\mathbf{x}) \geqslant 0,\ i = 1, \ldots, s\}$, as well as $f_{\tilde{K}}^* \leqslant f_K^*$.

Let $A_i \subset \mathbb{N}^n$ be the support of the polynomial g_i for $i = 0, \ldots, s$ and let $A = \bigcup_{i=0}^{s} A_i$ be the union of all supports of polynomials g_i. We remark that while considering a fixed support, the Newton polytope of $G(\boldsymbol{\mu})$ is not invariant in general since certain μ_i might vanish or term cancellation might occur. If for some $\boldsymbol{\mu} \in \mathbb{R}^s_{\geqslant 0}$ the polynomial

$G(\boldsymbol{\mu})$ is an ST-polynomial, then we assume that $\mathrm{New}(G(\boldsymbol{\mu})) = \mathrm{conv}(A)$ and $V(A) = \{\boldsymbol{\alpha}(0), \ldots, \boldsymbol{\alpha}(r)\} \subset (2\mathbb{N})^n$ and we denote $G(\boldsymbol{\mu})_{\mathrm{sonc}}$ as the optimal value of the GP from Corollary 4.1.5. Theorem 4.1.4 implies that $G(\boldsymbol{\mu}) - G(\boldsymbol{\mu})_{\mathrm{sonc}} \mathbf{x}^{\boldsymbol{\alpha}(0)} \geqslant 0$ and $G(\boldsymbol{\mu})_{\mathrm{sonc}} \in \mathbb{R}$ is the maximal possible choice for nonnegativity. Hence, we obtain a bound for the coefficient of the term $\mathbf{x}^{\boldsymbol{\alpha}(0)}$ depending on the other coefficients of $G(\boldsymbol{\mu})$ certifying nonnegativity of $G(\boldsymbol{\mu})$. If $G(\boldsymbol{\mu})$ is not an ST-polynomial for some $\boldsymbol{\mu} \in \mathbb{R}^s_{\geqslant 0}$, then we set $G(\boldsymbol{\mu})_{\mathrm{sonc}} = -\infty$, since the corresponding geometric program is infeasible. Thus, by (4.1.3), if $\boldsymbol{\mu}$ is fixed, then $G(\boldsymbol{\mu})_{\mathrm{sonc}}$ is a lower bound for f on the semialgebraic set K regarding the coefficient of $\mathbf{x}^{\boldsymbol{\alpha}(0)}$. Let $\mathbf{g} = (g_1, \ldots, g_s)$. We define

$$s(f, \mathbf{g}) \;=\; \sup\{G(\boldsymbol{\mu})_{\mathrm{sonc}} : \boldsymbol{\mu} \in \mathbb{R}^s_{\geqslant 0}\}.$$

Thus, we have particularly for $\boldsymbol{\alpha}(0) = \mathbf{0}$:

$$s(f, \mathbf{g}) \;\leqslant\; f^*_K. \tag{4.1.4}$$

For every fixed $\boldsymbol{\mu}$ the bound $G(\boldsymbol{\mu})_{\mathrm{sonc}}$ is computable by a geometric program. Unfortunately, this does not imply that the supremum is computable by a geometric program as well. However, following ideas by Ghasemi and Marshall [GM12] we present a general optimization program for a lower bound of $s(f, \mathbf{g})$, which is a geometric program under special conditions. Therefore, we need to fix some notation.

In the sense of Section 4.1.1 let $\Delta(A)$ be the set of exponents of the tail terms of $G(\boldsymbol{\mu})$ and $\Delta(G(\boldsymbol{\mu})) \subseteq \Delta(A)$ be the set of exponents which have a non-zero coefficient and are not a monomial square. Moreover, we define $\Delta(G) = \bigcup_{\boldsymbol{\mu} \in \mathbb{R}^s_{\geqslant 0}} \Delta(G(\boldsymbol{\mu}))$. Note that $\Delta(G(\boldsymbol{\mu})) \subseteq \Delta(G) \subseteq \Delta(A)$ for all $\boldsymbol{\mu}$. We have by Section 4.1.1, Definition 4.1.1

$$G(\boldsymbol{\mu})(\mathbf{x}) \;=\; -\sum_{i=0}^{s} \mu_i g_i(\mathbf{x}) \;=\; \sum_{j=0}^{r} G(\boldsymbol{\mu})_{\boldsymbol{\alpha}(j)} \mathbf{x}^{\boldsymbol{\alpha}(j)} + \sum_{\boldsymbol{\beta} \in \Delta(G)} G(\boldsymbol{\mu})_{\boldsymbol{\beta}} \mathbf{x}^{\boldsymbol{\beta}}$$

with coefficients $G(\boldsymbol{\mu})_{\boldsymbol{\alpha}(j)}, G(\boldsymbol{\mu})_{\boldsymbol{\beta}} \in \mathbb{R}$ depending on $\boldsymbol{\mu}$. Here we set the coefficients $G(\boldsymbol{\mu})_{\boldsymbol{\beta}} = 0$ for all $\boldsymbol{\beta} \in \Delta(G) \backslash \Delta(G(\boldsymbol{\mu}))$.

As before, we denote by $\{\lambda_0^{(\boldsymbol{\beta})}, \ldots, \lambda_r^{(\boldsymbol{\beta})}\}$ the barycentric coordinates of the lattice point $\boldsymbol{\beta} \in \Delta(A)$ with respect to the vertices of the simplex $\mathrm{New}(G(\boldsymbol{\mu})) = \mathrm{conv}(A)$. We define for every $\boldsymbol{\beta} \in \Delta(G)$ a set

$$R_{\boldsymbol{\beta}} \quad = \quad \{\mathbf{a}_{\boldsymbol{\beta}} \ : \ \mathbf{a}_{\boldsymbol{\beta}} = (a_{\boldsymbol{\beta},1}, \ldots, a_{\boldsymbol{\beta},r}) \in \mathbb{R}^r_{>0}\}.$$

Furthermore, we construct the nonnegative real set R as

$$R \quad = \quad [0, \infty)^s \times \bigtimes_{\boldsymbol{\beta} \in \Delta(G)} (R_{\boldsymbol{\beta}} \times \mathbb{R}_{\geqslant 0}).$$

Hence, R is the Cartesian product of $[0, \infty)^s$ and $|\Delta(G)|$ many copies $\mathbb{R}^r_{>0} \times \mathbb{R}_{\geqslant 0}$; each given by one $R_{\boldsymbol{\beta}}$ with $\boldsymbol{\beta} \in \Delta(G)$ and one $\mathbb{R}_{\geqslant 0}$. We define the function p from R to $\mathbb{R}_{\geqslant 0}$ as

$$p(\boldsymbol{\mu}, \{(\mathbf{a}_{\boldsymbol{\beta}}, b_{\boldsymbol{\beta}}) \ : \ \boldsymbol{\beta} \in \Delta(G)\}) \ = \ \sum_{i=1}^{s} \mu_i g_{i,\boldsymbol{\alpha}(0)} + \sum_{\substack{\boldsymbol{\beta} \in \Delta(G) \\ \lambda_0^{(\boldsymbol{\beta})} \neq 0}} \lambda_0^{(\boldsymbol{\beta})} \cdot b_{\boldsymbol{\beta}}^{\frac{1}{\lambda_0^{(\boldsymbol{\beta})}}} \cdot \prod_{\substack{j \in \mathrm{nz}(\boldsymbol{\beta}) \\ j \geqslant 1}} \left(\frac{\lambda_j^{(\boldsymbol{\beta})}}{a_{\boldsymbol{\beta},j}}\right)^{\frac{\lambda_j^{(\boldsymbol{\beta})}}{\lambda_0^{(\boldsymbol{\beta})}}}$$

where, as before, $\boldsymbol{\alpha}(0)$ is a vertex of $\mathrm{New}(G(\boldsymbol{\mu}))$ and $g_{i,\boldsymbol{\alpha}(0)}$ is the coefficient of the monomial $\mathbf{x}^{\boldsymbol{\alpha}(0)}$ in the polynomial g_i.

For the coefficient $G(\boldsymbol{\mu})_{\boldsymbol{\beta}}$ corresponding to the term with exponent $\boldsymbol{\beta}$ in $G(\boldsymbol{\mu})$ we use the notation $G(\boldsymbol{\mu})_{\boldsymbol{\beta}} = -\sum_{i=0}^{s} \mu_i \cdot g_{i,\boldsymbol{\beta}}$. In other words, $G(\boldsymbol{\mu})_{\boldsymbol{\beta}}$ is a linear form in the μ_i's given by the coefficients of the polynomials g_i; analogously for $G(\boldsymbol{\mu})_{\boldsymbol{\alpha}(j)}$.
We consider the following optimization problem:

$$(4.1.5) \quad \begin{cases} \text{minimize} & p(\boldsymbol{\mu}, \{(\mathbf{a}_{\boldsymbol{\beta}}, b_{\boldsymbol{\beta}}) \ : \ \boldsymbol{\beta} \in \Delta(G)\}) \text{ over the subset of } R \\ & \\ & (1) \ \sum_{\boldsymbol{\beta} \in \Delta(G)} a_{\boldsymbol{\beta},j} \ \leqslant \ G(\boldsymbol{\mu})_{\boldsymbol{\alpha}(j)} \ \text{ for all } \ j = 1, \ldots, r, \\ \text{defined by:} & (2) \ \prod_{j \in \mathrm{nz}(\boldsymbol{\beta})} \left(\frac{a_{\boldsymbol{\beta},j}}{\lambda_j^{(\boldsymbol{\beta})}}\right)^{\lambda_j^{(\boldsymbol{\beta})}} \geqslant b_{\boldsymbol{\beta}} \ \text{ for every } \boldsymbol{\beta} \in \Delta(G) \text{ with } \lambda_0^{(\boldsymbol{\beta})} = 0, \text{ and} \\ & (3) \ |G(\boldsymbol{\mu})_{\boldsymbol{\beta}}| \leqslant b_{\boldsymbol{\beta}} \ \text{ for every } \boldsymbol{\beta} \in \Delta(G) \text{ with } \lambda_0^{(\boldsymbol{\beta})} \neq 0. \end{cases}$$

The optimal value of (4.1.5) yields a lower bound for $s(f, \mathbf{g})$. In general, this program is neither a signomial program nor a geometric program. Though, the program can be relaxed to a signomial program and with further, stronger assumptions to a geometric program, see [IdW16b, Theorems 5.1 and 5.2].

Theorem 4.1.7. *Let γ be the optimal value of the optimization problem* (4.1.5). *Then we have $f_{\boldsymbol{\alpha}(0)} - \gamma \leqslant s(f, \mathbf{g})$. The optimization problem* (4.1.5) *restricted to $\boldsymbol{\mu} \in (0, \infty)^s$ is a signomial program if for every $\boldsymbol{\beta} \in \Delta(G)$ it holds that $G(\boldsymbol{\mu})_{\boldsymbol{\beta}}$ has the same sign for every choice of $\boldsymbol{\mu}$.*

Assume additionally that every linear form $G(\boldsymbol{\mu})_{\boldsymbol{\alpha}(j)} = -\sum_{i=0}^{s} \mu_i \cdot g_{i,\boldsymbol{\alpha}(j)}$ corresponding to a vertex $\boldsymbol{\alpha}(j)$ of $\mathrm{New}(G(\boldsymbol{\mu}))$ *has only one summand and is strictly positive. Assume moreover that for all $\boldsymbol{\beta} \in \Delta(G)$ the linear form $G(\boldsymbol{\mu})_{\boldsymbol{\beta}} = -\sum_{i=0}^{s} \mu_i \cdot g_{i,\boldsymbol{\beta}}$ has only positive terms. If furthermore all $g_{i,\boldsymbol{\alpha}(0)}$ for $i = 1, \ldots, s$ are greater than or equal to zero, then* (4.1.5) *is a geometric program.*

4.2 Constrained Polynomial Optimization via Signomial and Geometric Programming

In this section, we provide relaxations of the program (4.1.5) following the ideas of Ghasemi and Marshall in [GM13]. The goal is to weaken the assumptions which are needed to obtain a geometric program or at least a signomial program. We provide such relaxations in the programs (4.2.2) and (4.2.3) and provide the desired properties in the Theorems 4.2.1 and 4.2.4. Moreover, we show that under certain extra assumptions the bound obtained by the new program (4.2.2) equals the optimal bound $s(f, \mathbf{g})$ from the previous section; see Theorem 4.2.5. Furthermore, we demonstrate in the following Sections 4.3 and 4.4 that the resulting programs can be an alternative for SDP in cases where Lasserre's relaxation has issues.

Let all notation regarding $G(\boldsymbol{\mu})$ be given as in Section 4.1.4. Assume that we have for each $i = 0, \ldots, s$

$$g_i \quad = \quad \sum_{\boldsymbol{\beta} \in A_i} g_{i,\boldsymbol{\beta}} \cdot \mathbf{x}^{\boldsymbol{\beta}}$$

with $g_{i,\boldsymbol{\beta}} \in \mathbb{R}$. We have $\Delta(A_i) \subseteq \Delta(A)$ and hence write

$$g_i \quad = \quad \sum_{j=0}^{r} g_{i,\boldsymbol{\alpha}(j)} \mathbf{x}^{\boldsymbol{\alpha}(j)} + \sum_{\boldsymbol{\beta} \in \Delta(A)} g_{i,\boldsymbol{\beta}} \mathbf{x}^{\boldsymbol{\beta}}$$

and set $g_{i,\boldsymbol{\alpha}(j)} = 0$ for all $\boldsymbol{\alpha}(j) \in V(A) \backslash A_i$ and $g_{i,\boldsymbol{\beta}} = 0$ for all $\boldsymbol{\beta} \in \Delta(A) \backslash A_i$. We remark that three cases can occur for $\boldsymbol{\beta} \in \Delta(A) \cap A_i$:

(1) $-g_{i,\boldsymbol{\beta}}\mathbf{x}^{\boldsymbol{\beta}}$ is not a monomial square. Then we have $\boldsymbol{\beta} \in \Delta(G)$.

(2) $-g_{i,\boldsymbol{\beta}}\mathbf{x}^{\boldsymbol{\beta}}$ is a monomial square, but there exists another g_l such that $-g_{l,\boldsymbol{\beta}}\mathbf{x}^{\boldsymbol{\beta}}$ is not a monomial square. Then we have $\boldsymbol{\beta} \in \Delta(G)$.

(3) $-g_{i,\boldsymbol{\beta}}\mathbf{x}^{\boldsymbol{\beta}}$ is a monomial square, and there exists no other g_l such that $-g_{l,\boldsymbol{\beta}}\mathbf{x}^{\boldsymbol{\beta}}$ is not a monomial square. Then we have $\boldsymbol{\beta} \notin \Delta(G)$.

Sums of monomial squares as described in case (3) are ignored in our program (4.1.5). Thus, we can also ignore this case here. We now investigate the other two cases in detail. As already mentioned in Section 4.1.4 we can interpret the coefficients $G(\boldsymbol{\mu})_{\boldsymbol{\alpha}(j)}$ and $G(\boldsymbol{\mu})_{\boldsymbol{\beta}}$ as linear forms in $\boldsymbol{\mu}$ since we have for all $j = 0, \ldots, r$

$$G(\boldsymbol{\mu})_{\boldsymbol{\alpha}(j)} \;=\; -\sum_{i=0}^{s} \mu_i \cdot g_{i,\boldsymbol{\alpha}(j)} \quad \text{and} \quad G(\boldsymbol{\mu})_{\boldsymbol{\beta}} \;=\; -\sum_{i=0}^{s} \mu_i \cdot g_{i,\boldsymbol{\beta}}.$$

We decompose every coefficient $G(\boldsymbol{\mu})_{\boldsymbol{\beta}}$ into a positive and a negative part such that $G(\boldsymbol{\mu})_{\boldsymbol{\beta}} = G(\boldsymbol{\mu})^+_{\boldsymbol{\beta}} - G(\boldsymbol{\mu})^-_{\boldsymbol{\beta}}$, where

$$G(\boldsymbol{\mu})^-_{\boldsymbol{\beta}} \;=\; \sum_{g_{i,\boldsymbol{\beta}}>0} \mu_i \cdot g_{i,\boldsymbol{\beta}} \quad \text{and} \quad G(\boldsymbol{\mu})^+_{\boldsymbol{\beta}} \;=\; -\sum_{g_{i,\boldsymbol{\beta}}<0} \mu_i \cdot g_{i,\boldsymbol{\beta}}. \tag{4.2.1}$$

This decomposition is *independent* of the choice of $\boldsymbol{\mu}$ in the sense that no $g_{i,\boldsymbol{\beta}}$ can be a summand of *both* $G(\boldsymbol{\mu})^+_{\boldsymbol{\beta}}$ and $G(\boldsymbol{\mu})^-_{\boldsymbol{\beta}}$ for different choices of $\boldsymbol{\mu}$ since $\boldsymbol{\mu} \in \mathbb{R}^s_{\geqslant 0}$. The key idea is to redefine the constraint $b_{\boldsymbol{\beta}} \geqslant |G(\boldsymbol{\mu})_{\boldsymbol{\beta}}|$ of the program (4.1.5) by a new constraint $b_{\boldsymbol{\beta}} \geqslant \max\{G(\boldsymbol{\mu})^+_{\boldsymbol{\beta}}, G(\boldsymbol{\mu})^-_{\boldsymbol{\beta}}\}$. Let R be defined as in Section 4.1.4 and let $g^+_{i,\boldsymbol{\alpha}(0)} = \max\{g_{i,\boldsymbol{\alpha}(0)}, 0\}$, i.e., we only consider the terms with exponents $\boldsymbol{\alpha}(0)$ which are positive in the g_i and hence negative in $G(\boldsymbol{\mu})$. We redefine p as

$$p(\boldsymbol{\mu}, \{(\mathbf{a}_{\boldsymbol{\beta}}, b_{\boldsymbol{\beta}}) \;:\; \boldsymbol{\beta} \in \Delta(G)\}) \;=\; \sum_{i=1}^{s} \mu_i g^+_{i,\boldsymbol{\alpha}(0)} + \sum_{\substack{\boldsymbol{\beta}\in\Delta(G) \\ \lambda_0^{(\boldsymbol{\beta})}\neq 0}} \lambda_0^{(\boldsymbol{\beta})} \cdot b_{\boldsymbol{\beta}}^{\frac{1}{\lambda_0^{(\boldsymbol{\beta})}}} \cdot \prod_{\substack{j\in \mathrm{nz}(\boldsymbol{\beta}) \\ j\geqslant 1}} \left(\frac{\lambda_j^{(\boldsymbol{\beta})}}{a_{\boldsymbol{\beta},j}}\right)^{\frac{\lambda_j^{(\boldsymbol{\beta})}}{\lambda_0^{(\boldsymbol{\beta})}}}.$$

We consider the following optimization problem in the variables $\mu_1, \ldots, \mu_s$ and $a_{\boldsymbol{\beta},1}, \ldots, a_{\boldsymbol{\beta},r}, b_{\boldsymbol{\beta}}$ for every $\boldsymbol{\beta} \in \Delta(G)$:

$$(4.2.2)\quad \begin{cases} \text{minimize} & p(\boldsymbol{\mu}, \{(\mathbf{a}_{\boldsymbol{\beta}}, b_{\boldsymbol{\beta}}) \ : \ \boldsymbol{\beta} \in \Delta(G)\}) \text{ over the subset of } R \\ & (1) \quad \sum_{\boldsymbol{\beta} \in \Delta(G)} a_{\boldsymbol{\beta},j} \ \leqslant \ G(\boldsymbol{\mu})_{\boldsymbol{\alpha}(j)} \ \text{ for all } \ j = 1, \ldots, r, \\ \text{defined by:} & (2) \quad \prod_{j \in \mathrm{nz}(\boldsymbol{\beta})} \left(\frac{a_{\boldsymbol{\beta},j}}{\lambda_j^{(\boldsymbol{\beta})}} \right)^{\lambda_j^{(\boldsymbol{\beta})}} \ \geqslant \ b_{\boldsymbol{\beta}}, \ \text{ for all } \ \boldsymbol{\beta} \in \Delta(G) \text{ with } \lambda_0^{(\boldsymbol{\beta})} = 0, \\ & (3) \quad G(\boldsymbol{\mu})^+_{\boldsymbol{\beta}} \leqslant b_{\boldsymbol{\beta}} \ \text{ for all } \ \boldsymbol{\beta} \in \Delta(G), \ \text{ and} \\ & (4) \quad G(\boldsymbol{\mu})^-_{\boldsymbol{\beta}} \leqslant b_{\boldsymbol{\beta}} \ \text{ for all } \ \boldsymbol{\beta} \in \Delta(G). \end{cases}$$

By condition (1), this problem is feasible only for choices of $\boldsymbol{\mu}$ such that $G(\boldsymbol{\mu})_{\boldsymbol{\alpha}(j)} > 0$ for all $\boldsymbol{\alpha}(j)$ since all $a_{\boldsymbol{\beta},j}$ are strictly positive. We set the output as $-\infty$ in all other cases. Indeed, with some additional assumptions the program (4.2.2) is a geometric program. Moreover, it is a relaxation of the program (4.1.5).

Theorem 4.2.1. *Assume that for all $j = 1, \ldots, r$ the form $G(\boldsymbol{\mu})_{\boldsymbol{\alpha}(j)} = -\sum_{i=0}^{s} \mu_i \cdot g_{i,\boldsymbol{\alpha}(j)}$ has exactly one strictly positive term, i.e., there exists exactly one strictly negative $g_{i,\boldsymbol{\alpha}(j)}$. Then the optimization problem* (4.2.2) *restricted to $\boldsymbol{\mu} \in (0,\infty)^s$ is a geometric program. Assume that γ_{sonc} denotes the optimal value of* (4.2.2) *and γ denotes the optimal value of* (4.1.5)*. Then we have*

$$f_{\boldsymbol{\alpha}(0)} - \gamma_{\text{sonc}} \ \leqslant \ f_{\boldsymbol{\alpha}(0)} - \gamma \ \leqslant \ s(f, \mathbf{g}).$$

The typical choice for $\boldsymbol{\alpha}(0)$ is the origin which yields a lower bound for f to be nonnegative on K with the inequality (4.1.4):

Corollary 4.2.2. *Let all assumptions be as in Theorem* 4.2.1*. If $\boldsymbol{\alpha}(0)$ is the origin, then we have*

$$f_{\mathbf{0}} - \gamma_{\text{sonc}} \ \leqslant \ f_{\mathbf{0}} - \gamma \ \leqslant \ s(f, \mathbf{g}) \ \leqslant \ f_K^*.$$

Proof. (Theorem 4.2.1) If we restrict ourselves to $\boldsymbol{\mu} \in (0,\infty)^s$, then all functions involved in (4.2.2) depend on variables in $\mathbb{R}_{>0}$. By assumption every $G(\boldsymbol{\mu})_{\boldsymbol{\alpha}(j)}$ has exactly one strictly positive term. Thus, we can express constraint (1) in (4.2.2) as

$$\frac{\sum_{\boldsymbol{\beta} \in \Delta(G)} a_{\boldsymbol{\beta},j} + G(\boldsymbol{\mu})^-_{\boldsymbol{\alpha}(j)}}{G(\boldsymbol{\mu})^+_{\boldsymbol{\alpha}(j)}} \ \leqslant \ 1,$$

with $G(\boldsymbol{\mu})^-_{\boldsymbol{\alpha}(j)}$ and $G(\boldsymbol{\mu})^+_{\boldsymbol{\alpha}(j)}$ defined analogously as in (4.2.1). Since $G(\boldsymbol{\mu})^+_{\boldsymbol{\alpha}(j)}$ is a monomial the left hand side is a posynomial in $\boldsymbol{\mu}$ and $\mathbf{x}$. The constraints (2) – (4) are posynomial constraints in the sense of Definition 4.1.3 of a geometric program. The function p is also a posynomial since all terms are nonnegative by construction and all exponents are rational. Moreover, every $b_{\boldsymbol{\beta}}$ in (4.2.2) has to be greater or equal than the corresponding $b_{\boldsymbol{\beta}}$ in (4.1.5) because $\max\{a,b\} \geqslant |a-b|$ for all $a,b \in \mathbb{R}\backslash\{0\}$. Furthermore, since $g^+_{i,\boldsymbol{\alpha}(0)} \geqslant g_{i,\boldsymbol{\alpha}(0)}$ holds, the inequality $\gamma_{\text{sonc}} \leqslant \gamma$ follows by the definitions of (4.2.2) and (4.1.5). The last inequality follows from Theorem 4.1.7. □

One expects the programs (4.1.5) and (4.2.2) to have a similar optimal value if, for example, $g_{i,\boldsymbol{\alpha}(0)} \geqslant 0$ for most $i = 1,\ldots,s$ and if either $G(\boldsymbol{\mu})^+_{\boldsymbol{\beta}}$ or $G(\boldsymbol{\mu})^-_{\boldsymbol{\beta}}$ is identically zero for most $\boldsymbol{\beta} \in \Delta(G)$. Note that one of $G(\boldsymbol{\mu})^+_{\boldsymbol{\beta}}, G(\boldsymbol{\mu})^-_{\boldsymbol{\beta}}$ is zero if and only if $\max\{G(\boldsymbol{\mu})^+_{\boldsymbol{\beta}}, G(\boldsymbol{\mu})^-_{\boldsymbol{\beta}}\} = |G(\boldsymbol{\mu})^+_{\boldsymbol{\beta}} - G(\boldsymbol{\mu})^-_{\boldsymbol{\beta}}| = |G(\boldsymbol{\mu})_{\boldsymbol{\beta}}|$ and the latter holds if and only if $g_{i,\boldsymbol{\beta}} \geqslant 0$ or $g_{i,\boldsymbol{\beta}} \leqslant 0$ for $i = 0,\ldots,s$.

We give an example to demonstrate how a given constrained polynomial optimization problem can be translated into the geometric program (4.2.2). In Section 4.3, we provide several further examples including actual computations of infima using the GP-solver CVX.

Example 4.2.3. Let $f = 1 + 2x^2y^4 + \frac{1}{2}x^3y^2$ and $g_1 = \frac{1}{3} - x^6y^2$. From these two polynomials we obtain a function

$$G(\mu) = \left(1 - \frac{1}{3}\mu\right) + 2x^2y^4 + \mu x^6y^2 + \frac{1}{2}x^3y^2.$$

For $G(\mu)$ to be an ST-polynomial, we have to choose $\mu \in (0,3)$. Here, the vertices of $\text{New}(G(\mu))$ are $\boldsymbol{\alpha}(0) = (0,0)$, $\boldsymbol{\alpha}(1) = (2,4)$, and $\boldsymbol{\alpha}(2) = (6,2)$, and we have $\Delta(G) = \{\boldsymbol{\beta}\} = \{(3,2)\}$. Thus, we introduce 4 variables $(a_{\boldsymbol{\beta},1}, a_{\boldsymbol{\beta},2}, b_{\boldsymbol{\beta}}, \mu)$. First, we compute the barycentric coordinates of $\boldsymbol{\beta}$ and get

$$\lambda_0^{(\boldsymbol{\beta})} = \frac{3}{10}, \quad \lambda_1^{(\boldsymbol{\beta})} = \frac{3}{10}, \quad \lambda_2^{(\boldsymbol{\beta})} = \frac{2}{5}.$$

We match the coefficients of $G(\mu)$ with the vertices $\boldsymbol{\alpha}(j)$:

- $g^+_{1,\boldsymbol{\alpha}(0)} = \max\{\frac{1}{3}, 0\} = \frac{1}{3}$,
- $G(\mu)_{\boldsymbol{\alpha}(1)} = 2$, $G(\mu)_{\boldsymbol{\alpha}(2)} = \mu$,
- $G(\mu)^+_{\boldsymbol{\beta}} = \frac{1}{2}$, $G(\mu)^-_{\boldsymbol{\beta}}$ does not exist.

Hence, program (4.2.2) is of the form:

$$\inf \left\{ \frac{1}{3}\mu + \frac{3}{10} \cdot b_{\boldsymbol{\beta}}^{\frac{10}{3}} \cdot \left(\frac{3}{10}\right)^{1} \cdot \left(\frac{2}{5}\right)^{\frac{4}{3}} \cdot (a_{\boldsymbol{\beta},1})^{-1} \cdot (a_{\boldsymbol{\beta},2})^{-\frac{4}{3}} \right\}$$

such that:

(1) $a_{\boldsymbol{\beta},1} \leqslant 2,\ a_{\boldsymbol{\beta},2} \leqslant \mu$.

(2) The second constraint does not appear, because we do not have $\lambda_0^{(\boldsymbol{\beta})} = 0$.

(3) $\frac{1}{2} \leqslant b_{\boldsymbol{\beta}}$.

(4) The fourth constraint does not appear, because we do not have a $G(\mu)_{\boldsymbol{\beta}}^{-}$.

○

In what follows, we extend Theorem 4.1.7 by reformulating the program (4.2.2) such that it is *always* applicable. On the one hand, the new program is only a signomial program instead of a geometric program in general. On the other hand, the reformulated program covers the missing cases of Theorem 4.2.1 and also yields better bounds than the corresponding geometric program (4.2.2) in general. We define

$$q(\boldsymbol{\mu}, \{(\mathbf{a}_{\boldsymbol{\beta}}, c_{\boldsymbol{\beta}})\ :\ \boldsymbol{\beta} \in \Delta(G)\}) = \sum_{i=1}^{s} \mu_i g_{i,\boldsymbol{\alpha}(0)} + \sum_{\substack{\boldsymbol{\beta} \in \Delta(G) \\ \lambda_0^{(\boldsymbol{\beta})} \neq 0}} \lambda_0^{(\boldsymbol{\beta})} \cdot c_{\boldsymbol{\beta}}^{\frac{1}{\lambda_0^{(\boldsymbol{\beta})}}} \cdot \prod_{\substack{j \in \mathrm{nz}(\boldsymbol{\beta}) \\ j \geqslant 1}} \left(\frac{\lambda_j^{(\boldsymbol{\beta})}}{a_{\boldsymbol{\beta},j}}\right)^{\frac{\lambda_j^{(\boldsymbol{\beta})}}{\lambda_0^{(\boldsymbol{\beta})}}}$$

and consider the subsequent program:

$$(4.2.3)\quad \begin{cases} \text{minimize} & q(\boldsymbol{\mu}, \{(\mathbf{a}_{\boldsymbol{\beta}}, c_{\boldsymbol{\beta}})\ :\ \boldsymbol{\beta} \in \Delta(G)\}) \text{ over the subset of } R \\ & (1)\ \sum_{\boldsymbol{\beta} \in \Delta(G)} a_{\boldsymbol{\beta},j} \leqslant G(\boldsymbol{\mu})_{\boldsymbol{\alpha}(j)} \text{ for all } j = 1, \ldots, r, \\ \text{defined by:} & (2)\ \prod_{j \in \mathrm{nz}(\boldsymbol{\beta})} \left(\frac{a_{\boldsymbol{\beta},j}}{\lambda_j^{(\boldsymbol{\beta})}}\right)^{\lambda_j^{(\boldsymbol{\beta})}} \geqslant c_{\boldsymbol{\beta}} \text{ for all } \boldsymbol{\beta} \in \Delta(G) \text{ with } \lambda_0^{(\boldsymbol{\beta})} = 0, \\ & (3)\ G(\boldsymbol{\mu})_{\boldsymbol{\beta}}^{+} - G(\boldsymbol{\mu})_{\boldsymbol{\beta}}^{-} \leqslant c_{\boldsymbol{\beta}} \text{ for all } \boldsymbol{\beta} \in \Delta(G), \text{ and} \\ & (4)\ G(\boldsymbol{\mu})_{\boldsymbol{\beta}}^{-} - G(\boldsymbol{\mu})_{\boldsymbol{\beta}}^{+} \leqslant c_{\boldsymbol{\beta}} \text{ for all } \boldsymbol{\beta} \in \Delta(G). \end{cases}$$

The key difference between this program and (4.2.2) is that

$$c_{\boldsymbol{\beta}} \geqslant \max\{G(\boldsymbol{\mu})_{\boldsymbol{\beta}}^{+} - G(\boldsymbol{\mu})_{\boldsymbol{\beta}}^{-}, G(\boldsymbol{\mu})_{\boldsymbol{\beta}}^{-} - G(\boldsymbol{\mu})_{\boldsymbol{\beta}}^{+}\} = |G(\boldsymbol{\mu})_{\boldsymbol{\beta}}|.$$

We obtain the following statement.

Theorem 4.2.4. *The optimization problem* (4.2.3) *restricted to* $\boldsymbol{\mu} \in (0,\infty)^s$ *is a signomial program. Assume that* γ_{snp} *denotes the optimal value of* (4.2.3) *and* $\gamma_{\text{sonc}}, \gamma$ *denote the optimal values of* (4.2.2) *and* (4.1.5) *as before. Then we have*

$$f_{\boldsymbol{\alpha}(0)} - \gamma_{\text{sonc}} \ \leqslant\ f_{\boldsymbol{\alpha}(0)} - \gamma_{\text{snp}} \ \leqslant\ f_{\boldsymbol{\alpha}(0)} - \gamma \ \leqslant\ s(f,\mathbf{g}).$$

Particularly, we have $\gamma_{\text{snp}} = \gamma$ *if the program* (4.1.5) *attains its optimal value for* $\boldsymbol{\mu} \in (0,\infty)^s$.

Proof. The proof is analogous to the proof of Theorem 4.2.1. The only difference is that now certain terms can have a negative sign and hence posynomials then become signomials. The statement follows with the definition of a signomial program; see Section 4.1.2. □

Finally, we show that if we strengthen the assumptions in Theorem 4.2.1, then, the output $f_{\boldsymbol{\alpha}(0)} - \gamma_{\text{sonc}}$ of the program (4.2.2) equals the output $f_{\boldsymbol{\alpha}(0)} - \gamma$ of the program (4.1.5) and particularly equals the bound $s(f,\mathbf{g})$.

Theorem 4.2.5. *Assume that for every* $1 \leqslant j \leqslant r$ *the form* $G(\boldsymbol{\mu})_{\boldsymbol{\alpha}(j)} = -\sum_{i=0}^{s} \mu_i \cdot g_{i,\boldsymbol{\alpha}(j)}$ *has exactly one strictly positive term. Furthermore, assume that* $g_{i,\boldsymbol{\alpha}(0)} \geqslant 0$ *for all* $i = 1,\ldots,s$, *and that* $\Delta(A) \cap A_i \cap A_l = \varnothing$ *for all* $0 \leqslant i < l \leqslant s$. *Let* γ *be the optimal value of the program* (4.1.5). *If the optimal value* $s(f,\mathbf{g}) = \sup\{G(\boldsymbol{\mu})_{\text{sonc}} : \boldsymbol{\mu} \in \mathbb{R}^s_{\geqslant 0}\}$ *is attained for some* $\boldsymbol{\mu} \in (0,\infty)^s$, *then* $f_{\boldsymbol{\alpha}(0)} - \gamma_{\text{sonc}} = f_{\boldsymbol{\alpha}(0)} - \gamma = s(f,\mathbf{g})$, *where, as before,* γ_{sonc} *denotes the optimal value of* (4.2.2).

Note that the condition $\Delta(A) \cap A_i \cap A_l = \varnothing$ is satisfied if the supports of g_i and g_l differ in all elements that are not vertices of $\text{New}(G(\boldsymbol{\mu}))$.

Proof. The assumption $\Delta(A) \cap A_i \cap A_l = \varnothing$ for all $0 \leqslant i < l \leqslant s$ implies for every $\boldsymbol{\beta} \in \Delta(G)$ that $G(\boldsymbol{\mu})_{\boldsymbol{\beta}} = -\sum_{i=0}^{s} \mu_i \cdot g_{i,\boldsymbol{\beta}} = -\mu_k \cdot g_{k,\boldsymbol{\beta}}$, for some $k = 0,\ldots,s$. Therefore, we have for every $\boldsymbol{\beta} \in \Delta(G)$ that

$$\max\{G(\boldsymbol{\mu})^+_{\boldsymbol{\beta}}, G(\boldsymbol{\mu})^-_{\boldsymbol{\beta}}\} \ = \ |\mu_k \cdot g_{k,\boldsymbol{\beta}}| \ = \ |G(\boldsymbol{\mu})_{\boldsymbol{\beta}}|.$$

Moreover, we have $g_{i,\boldsymbol{\alpha}(0)} \geqslant 0$ for all $i = 1,\ldots,s$ by assumption and thus we obtain $\sum_{i=1}^{s} \mu_i g_{i,\boldsymbol{\alpha}(0)} = \sum_{i=1}^{s} \mu_i g^+_{i,\boldsymbol{\alpha}(0)}$. Hence, the two programs (4.1.5) and (4.2.2) coincide.

By assumption, every $G(\boldsymbol{\mu})_{\boldsymbol{\alpha}(j)}$ consists of exactly one positive term. Therefore, (4.2.2) is a GP by Theorem 4.2.1. Considering Theorem 4.2.1 it suffices to show that the inequality $f_{\boldsymbol{\alpha}(0)} - \gamma_{\text{sonc}} \geqslant s(f, \mathbf{g})$ holds, such that $f_{\boldsymbol{\alpha}(0)} - \gamma_{\text{sonc}} = f_{\boldsymbol{\alpha}(0)} - \gamma = s(f, \mathbf{g})$ is fulfilled. Let $\boldsymbol{\mu}^* \in (0, \infty)^s$ be such that $G(\boldsymbol{\mu}^*)_{\text{sonc}} = s(f, \mathbf{g})$. By Corollary 4.1.5 $G(\boldsymbol{\mu}^*)_{\text{sonc}}$ is given by a feasible point $(a_{\boldsymbol{\beta},1}, \ldots, a_{\boldsymbol{\beta},r})$ of the program

$$\begin{cases} \text{minimize} & \sum\limits_{\substack{\boldsymbol{\beta} \in \Delta(G) \\ \lambda_0^{(\boldsymbol{\beta})} \neq 0}} \lambda_0^{(\boldsymbol{\beta})} \cdot |\mu_k^* \cdot g_{k,\boldsymbol{\beta}}|^{\frac{1}{\lambda_0^{(\boldsymbol{\beta})}}} \cdot \prod\limits_{\substack{j \in \text{nz}(\boldsymbol{\beta}) \\ j \geqslant 1}} \left(\frac{\lambda_j^{(\boldsymbol{\beta})}}{a_{\boldsymbol{\beta},j}} \right)^{\frac{\lambda_j^{(\boldsymbol{\beta})}}{\lambda_0^{(\boldsymbol{\beta})}}} \text{ over the subset } R' \text{ of } R \\ \text{defined by:} & \begin{array}{l} (1) \quad \sum\limits_{\boldsymbol{\beta} \in \Delta(G)} a_{\boldsymbol{\beta},j} \leqslant G(\boldsymbol{\mu}^*)_{\boldsymbol{\alpha}(j)} \text{ for all } j = 1, \ldots, r, \text{ and} \\ (2) \quad \prod\limits_{j \in \text{nz}(\boldsymbol{\beta})} \left(\frac{a_{\boldsymbol{\beta},j}}{\lambda_j^{(\boldsymbol{\beta})}} \right)^{\lambda_j^{(\boldsymbol{\beta})}} \geqslant |\mu_k^* \cdot g_{k,\boldsymbol{\beta}}| \text{ for all } \boldsymbol{\beta} \in \Delta(G) \text{ with } \lambda_0^{(\boldsymbol{\beta})} = 0. \end{array} \end{cases}$$

Then every $(a_{\boldsymbol{\beta},1}, \ldots, a_{\boldsymbol{\beta},r}, b_{\boldsymbol{\beta}}, \boldsymbol{\mu}^*)$ with $b_{\boldsymbol{\beta}} \geqslant |\mu_k^* \cdot g_{k,\boldsymbol{\beta}}|$ for all $\boldsymbol{\beta} \in \Delta(G)$ is a feasible point of (4.2.2). Furthermore,

$$\begin{aligned} f_{\boldsymbol{\alpha}(0)} - \sum_{i=1}^{s} \mu_i^* g_{i,\boldsymbol{\alpha}(0)}^{+} - \sum_{\substack{\boldsymbol{\beta} \in \Delta(G) \\ \lambda_0^{(\boldsymbol{\beta})} \neq 0}} \lambda_0^{(\boldsymbol{\beta})} \cdot b_{\boldsymbol{\beta}}^{\frac{1}{\lambda_0^{(\boldsymbol{\beta})}}} \cdot \prod_{\substack{j \in \text{nz}(\boldsymbol{\beta}) \\ j \geqslant 1}} \left(\frac{\lambda_j^{(\boldsymbol{\beta})}}{a_{\boldsymbol{\beta},j}} \right)^{\frac{\lambda_j^{(\boldsymbol{\beta})}}{\lambda_0^{(\boldsymbol{\beta})}}} \\ = G(\boldsymbol{\mu}^*)(0) - \sum_{\substack{\boldsymbol{\beta} \in \Delta(G) \\ \lambda_0^{(\boldsymbol{\beta})} \neq 0}} \lambda_0^{(\boldsymbol{\beta})} \cdot b_{\boldsymbol{\beta}}^{\frac{1}{\lambda_0^{(\boldsymbol{\beta})}}} \cdot \prod_{\substack{j \in \text{nz}(\boldsymbol{\beta}) \\ j \geqslant 1}} \left(\frac{\lambda_j^{(\boldsymbol{\beta})}}{a_{\boldsymbol{\beta},j}} \right)^{\frac{\lambda_j^{(\boldsymbol{\beta})}}{\lambda_0^{(\boldsymbol{\beta})}}} \\ \geqslant G(\boldsymbol{\mu}^*)(0) - \sum_{\substack{\boldsymbol{\beta} \in \Delta(G) \\ \lambda_0^{(\boldsymbol{\beta})} \neq 0}} \lambda_0^{(\boldsymbol{\beta})} \cdot |\mu_k^* \cdot g_{k,\boldsymbol{\beta}}|^{\frac{1}{\lambda_0^{(\boldsymbol{\beta})}}} \cdot \prod_{\substack{j \in \text{nz}(\boldsymbol{\beta}) \\ j \geqslant 1}} \left(\frac{\lambda_j^{(\boldsymbol{\beta})}}{a_{\boldsymbol{\beta},j}} \right)^{\frac{\lambda_j^{(\boldsymbol{\beta})}}{\lambda_0^{(\boldsymbol{\beta})}}}. \end{aligned}$$

Hence, $f_{\boldsymbol{\alpha}(0)} - \gamma_{\text{sonc}} \geqslant G(\boldsymbol{\mu}^*)_{\text{sonc}} = s(f, \mathbf{g})$.

□

4.3 Examples for Constrained Optimization via Geometric Programming and a Comparison to Lasserre Relaxations

We consider constrained polynomial optimization problems of the form

$$f_K^* \;=\; \inf_{\mathbf{x}\in K} f(\mathbf{x}),$$

where K is a basic closed semialgebraic set defined by $g_1,\ldots,g_s \geqslant 0$. One of the main results in [IdW16b] is the observation that lower bounds for global optimization problems arising from SONCs via GP can not only be computed faster, but also provide better bounds than those obtained by SOS via SDP. Here, we show that competitive bounds arising from SONC via GP can also be obtained for constrained problems. Particularly, if $2d$ is the maximal total degree of f and $g_1,\ldots,g_s$, then the bound given by the d-th Lasserre relaxation is not necessarily as good as our optimal solution, which is in contrast to the bounds obtained by Ghasemi and Marshall; see Example 4.3.5 for further details. Moreover, we provide examples demonstrating that the runtime of the GP approach is not sensitive to increasing the degree of a given problem, which is in sharp contrast to the runtime of SDPs.

Recall that the d-th Lasserre relaxation, see (2.3.3), is given by the parameter

$$f_{\mathrm{sos}}^{(d)} = \sup\left\{\gamma : f-\gamma = \sigma_0 + \sum_{i=1}^{s}\sigma_i g_i,\ \sigma_0,\sigma_i\in\Sigma_n, \text{with}\deg(\sigma_0),\deg(\sigma_i g_i)\leqslant 2d\right\},$$

where $d \geqslant \max\{\lceil\deg(f)/2\rceil, \max_{1\leqslant i\leqslant s}\{\lceil\deg(g_i)/2\rceil\}\}$. In what follows, we provide several examples comparing the Lasserre relaxation using the Matlab SDP solver Gloptipoly [HLL09] to our approach given in program (4.2.2) using the Matlab GP solver CVX [BG08, GBY06]. In every example in this section we optimize with respect to the constant term when applying program (4.2.2).

Example 4.3.1. Let $f = 1 + x^4y^2 + x^2y^4 - 3x^2y^2$ be the Motzkin polynomial and $g_1 = x^3y^2$. Then the corresponding feasible set is

$$K \;=\; \{(x,y)\in\mathbb{R}^2 \;:\; x\geqslant 0 \text{ or } y=0\}.$$

Since f is globally nonnegative and has two zeros $(1,1),(1,-1)$ on K, e.g., [Rez00], we

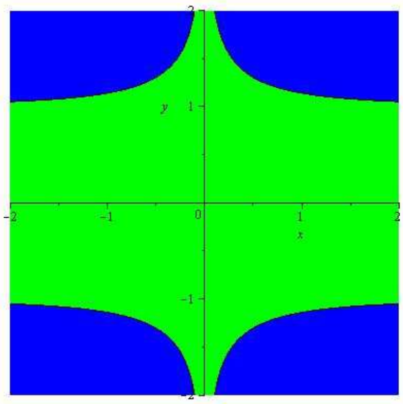

Figure 4.1: The feasible set for the constrained optimization problem in Example 4.3.2 is the unbounded green area.

have $f_K^* = 0$. We consider the third Lasserre relaxation and obtain

$$f_{\text{sos}}^{(3)} = \sup\{\gamma \ : \ f - \gamma = \sigma_0 + \sigma_1 \cdot g_1, \ \sigma_0, \sigma_1 \in \Sigma_2, \deg(\sigma_0) \leqslant 6, \deg(\sigma_1 g_1) \leqslant 6\} = -\infty,$$

since the problem is infeasible. Note that K is unbounded. Hence, it is not necessarily the case that $f_{\text{sos}}^{(d)} > -\infty$ for sufficiently high relaxation order d. Here, using GLOPTIPOLY, one can find that $f_{\text{sos}}^{(7)} = 0 = f_K^*$.

Now, we consider $s(f, g_1) = \sup\{G(\mu)_{\text{sonc}} : \mu \in \mathbb{R}_{\geqslant 0}\} \leqslant f_K^*$ where $G(\mu) = f - \mu g_1$ with $\mu \geqslant 0$. Note that $\text{New}(G(\mu))$ is a simplex for every choice of μ. In particular, for $\mu = 0$ we have that $G(\mu)_{\text{sonc}} = f_{\text{sonc}} = 0$, since the Motzkin polynomial is a SONC polynomial; see Example 2.4.6. It follows that

$$-\infty \ = \ f_{\text{sos}}^{(3)} \ < \ s(f, g_1) \ = \ 0 \ = \ f_K^*.$$

Thus, $s(f, g_1)$ yields the exact solution compared to the Lasserre relaxation. This is in sharp contrast to the geometric programming approach proposed in [GM13] where $f_{\text{sos}}^{(d)} \geqslant s(f, \mathbf{g})$ holds in general. ○

Example 4.3.2. Let $f = 1 + x^4y^2 + xy$ and $g_1 = \frac{1}{2} + x^2y^4 - x^2y^6$. The feasible set K is a non-compact set depicted in Figure 4.1. Using GLOPTIPOLY, one can check that $-\infty = f_{\text{sos}}^{(4)}$ and the optimal solution is given for $d = 8$ with $f_{\text{sos}}^{(8)} \approx 0.4474$. In this case one can extract the minimizers $(-0.557, 1.2715)$ and $(0.557, -1.2715)$.

We compare these results to our approach via geometric programming instead of Lasserre's relaxation. From f and g_1 we get $G(\mu) = (1 - \frac{1}{2}\mu) + x^4y^2 + \mu x^2y^6 + xy - \mu x^2y^4$. Note that $\text{New}(G(\mu))$ is a two-dimensional simplex if $\mu \notin \{0, 2\}$. Then, we have

$\Delta(G) = \{\boldsymbol{\beta}, \tilde{\boldsymbol{\beta}}\} = \{(1,1),(2,4)\}$. Therefore, we introduce the seven variables $(a_{\boldsymbol{\beta},1}, a_{\boldsymbol{\beta},2}, a_{\tilde{\boldsymbol{\beta}},1}, a_{\tilde{\boldsymbol{\beta}},2}, b_{\boldsymbol{\beta}}, b_{\tilde{\boldsymbol{\beta}}}, \mu)$. Hence, the geometric program (4.2.2) reads as follows:

$$\inf\left\{\frac{1}{2}\mu + \frac{7}{10}\cdot b_{\boldsymbol{\beta}}^{\frac{10}{7}}\cdot\left(\frac{1}{5}\right)^{\frac{2}{7}}\cdot\left(\frac{1}{10}\right)^{\frac{1}{7}}\cdot(a_{\boldsymbol{\beta},1})^{-\frac{2}{7}}\cdot(a_{\boldsymbol{\beta},2})^{-\frac{1}{7}} + \frac{1}{5}\cdot b_{\tilde{\boldsymbol{\beta}}}^{5}\cdot\left(\frac{1}{5}\right)^{1}\cdot\left(\frac{3}{5}\right)^{3}\cdot(a_{\tilde{\boldsymbol{\beta}},1})^{-1}\cdot(a_{\tilde{\boldsymbol{\beta}},2})^{-3}\right\}$$

such that the variables satisfy

$$a_{\boldsymbol{\beta},1} + a_{\tilde{\boldsymbol{\beta}},1} \leqslant 1,\ a_{\boldsymbol{\beta},2} + a_{\tilde{\boldsymbol{\beta}},2} \leqslant \mu, \ \text{ and } \ 1 \leqslant b_{\boldsymbol{\beta}},\ \mu \leqslant b_{\tilde{\boldsymbol{\beta}}}\ .$$

We use the Matlab solver CVX to solve this program. The optimal solution is given by

$$(a_{\boldsymbol{\beta},1}, a_{\boldsymbol{\beta},2}, a_{\tilde{\boldsymbol{\beta}},1}, a_{\tilde{\boldsymbol{\beta}},2}, b_{\boldsymbol{\beta}}, b_{\tilde{\boldsymbol{\beta}}}, \mu) = (0.9105, 0.0540, 0.0895, 0.0319, 1.0000, 0.0859, 0.0859)$$

and leads to

$$\gamma_{\mathrm{sonc}} \approx 0.5526,$$

which yields $f_{\boldsymbol{\alpha}(0)} - \gamma_{\mathrm{sonc}} \approx 0.4474$. Thus, we have

$$f_{\mathrm{sos}}^{(8)} = f_{\boldsymbol{\alpha}(0)} - \gamma_{\mathrm{sonc}} = s(f, g_1).$$

The equality $f_{\boldsymbol{\alpha}(0)} - \gamma_{\mathrm{sonc}} = s(f, g_1)$ is not surprising, since the assumptions of Theorem 4.2.5 are satisfied. Observe that we get the optimal solution immediately via geometric programming whereas one needs 5 relaxation steps via Lasserre's relaxation. In this example both geometric programing and the Lasserre approach have a runtime below 1 second. However, if we multiply all exponents in f and g_1 by 10, then the approaches differ significantly. By multiplying the exponents by 10 we have made a severe change to the problem since the term $x^{10}y^{10}$ is now a monomial square such that the exponent is a lattice point in the interior of the Newton polytope of the adjusted $G(\mu)$. Hence, we have to ignore this term when running the constrained optimization program (4.2.2). The adjusted program yields with CVX an output "`NaN`" in below one second. However, the reason is that it computes $\mu = 0$, which *is* the correct answer. Namely, after multiplying the exponents by 10, the only non-monomial square terms are given by g_1. Thus, the optimal choice is $\mu = 0$, and we can see that the minimal value is attained at $(0,0)$ and $f_K^* = 1$ is given by the constant term of f.

In comparison, we have a runtime of approximately 1110 seconds, i.e., approximately 18.5 minutes with Gloptipoly. After this time Gloptipoly provides an output "`Run` `into numerical problems.`". It claims, however, to have solved the problem and provides the correct minimum $f_K^* = 1$ at a minimizer $10^{-7} \cdot (-0.1057, 0.1711)$, which, of course, is the origin up to a numerical error. ○

Example 4.3.3. Let $f = 1 + x^2z^2 + y^2z^2 + x^2y^2 - 8xyz$ and $g_1 = x^2yz + xy^2z + x^2y^2 - 2 + xyz$. Using GLOPTIPOLY, we get the following sequence of lower bounds:

$$f_{\text{sos}}^{(2)} = f_{\text{sos}}^{(3)} = f_{\text{sos}}^{(4)} = -\infty < f_{\text{sos}}^{(5)} \approx -14.999.$$

However, one cannot certify the optimality via GLOPTIPOLY in this case. Additionally, the sequence $f_{\text{sos}}^{(d)}$ is not guaranteed to converge to f_K^*, since K is unbounded. Symbolically, we were able to prove a global minimum of $f_K^* = -15$ with four global minimizers $(2,2,2), (-2,-2,2), (-2,2,-2), (2,-2,-2)$ using the quantifier elimination software SYNRAC, see [AY03]. Now, we consider the approach via geometric programming instead of Lasserre relaxations. We have

$$G(\mu) \;=\; (1+2\mu) + x^2z^2 + y^2z^2 + (1-\mu)x^2y^2 + (-8-\mu)xyz - \mu x^2yz - \mu xy^2z.$$

Obviously, $G(\mu)$ is an ST-polynomial for $\mu \in [0,1)$, and we have $\Delta(G) = \{\boldsymbol{\beta}, \overline{\boldsymbol{\beta}}, \hat{\boldsymbol{\beta}}\} = \{(1,1,1),(2,1,1),(1,2,1)\}$. Thus, our geometric program has the following 13 variables

$$(a_{\boldsymbol{\beta},1}, a_{\boldsymbol{\beta},2}, a_{\boldsymbol{\beta},3}, a_{\overline{\boldsymbol{\beta}},1}, a_{\overline{\boldsymbol{\beta}},2}, a_{\overline{\boldsymbol{\beta}},3}, a_{\hat{\boldsymbol{\beta}},1}, a_{\hat{\boldsymbol{\beta}},2}, a_{\hat{\boldsymbol{\beta}},3}, b_{\boldsymbol{\beta}}, b_{\overline{\boldsymbol{\beta}}}, b_{\hat{\boldsymbol{\beta}}}, \mu).$$

Hence, program (4.2.2) is of the form

$$\inf\left\{0\cdot\mu + \frac{1}{4}\cdot b_{\boldsymbol{\beta}}^4 \cdot \left(\frac{1}{4}\right)\cdot\left(\frac{1}{4}\right)\cdot\left(\frac{1}{4}\right)\cdot (a_{\boldsymbol{\beta},1})^{-1}\cdot(a_{\boldsymbol{\beta},2})^{-1}\cdot(a_{\boldsymbol{\beta},3})^{-1}\right\}$$

such that

$$\begin{array}{ll}
(1) & a_{\boldsymbol{\beta},1} + a_{\overline{\boldsymbol{\beta}},1} + a_{\hat{\boldsymbol{\beta}},1} \leqslant 1,\ a_{\boldsymbol{\beta},2} + a_{\overline{\boldsymbol{\beta}},2} + a_{\hat{\boldsymbol{\beta}},2} \leqslant 1,\ a_{\boldsymbol{\beta},3} + a_{\overline{\boldsymbol{\beta}},3} + a_{\hat{\boldsymbol{\beta}},3} + \mu \leqslant 1,\\
(2) & \frac{1}{2}\cdot b_{\overline{\boldsymbol{\beta}}}\cdot\left(a_{\overline{\boldsymbol{\beta}},1}\right)^{-\frac{1}{2}}\cdot\left(a_{\overline{\boldsymbol{\beta}},3}\right)^{-\frac{1}{2}} \leqslant 1,\\
 & \frac{1}{2}\cdot b_{\hat{\boldsymbol{\beta}}}\cdot\left(a_{\hat{\boldsymbol{\beta}},2}\right)^{-\frac{1}{2}}\cdot\left(a_{\hat{\boldsymbol{\beta}},3}\right)^{-\frac{1}{2}} \leqslant 1,\\
(3) & 8\cdot b_{\boldsymbol{\beta}}^{-1} \leqslant 1,\ \mu\cdot b_{\boldsymbol{\beta}}^{-1} \leqslant 1,\ \mu\cdot b_{\overline{\boldsymbol{\beta}}}^{-1} \leqslant 1,\ \mu\cdot b_{\hat{\boldsymbol{\beta}}}^{-1} \leqslant 1\,.
\end{array}$$

This leads to $\gamma_{\text{sonc}} = \frac{1}{256}\cdot 8^4 = 16$ and so $f_{\boldsymbol{\alpha}(0)} - \gamma_{\text{sonc}} = -15$. The runtime for this example is below 1 second. Multiplying the exponents of f and g_1 by 10 yields the same results; the runtime for the geometric program remains below 1 second. In comparison, Gloptipoly yields

$$f_{\text{sos}}^{(d)} \;=\; -\infty \text{ for } d \leqslant 19,$$

and provides a bound

$$f_{\text{sos}}^{(20)} \;\approx\; -14.999$$

in the 20th relaxation after 36563 seconds, i.e., approximately 10.16 hours. Moreover, although this bound is numerically equal to f_K^*, Gloptipoly was not able to certify that the correct bound was found. ○

Example 4.3.4. Let $f = z^6 + x^4y^2 + x^2y^4 - 3x^2y^2z^2$ and $g_1 = x^2 + y^2 + z^2 - 1$. We obtain $G(\mu) = f - \mu g_1$. This problem is infeasible in the sense of program (4.2.2). Namely, condition (♣) is never satisfied since for any $\mu > 0$ we have a vertex $(2, 0, 0)$ or $(0, 2, 0)$ of $\text{New}(G(\mu))$ with a negative coefficient. Therefore, one can immediately conclude that $s(f, g_1)$ has to be obtained for $\mu = 0$. Thus, we have $s(f, g_1) = f_{\text{sonc}}$. Since f is the homogenized Motzkin polynomial we obtain immediately $f_{\text{sonc}} = f_K^* = 0$. An analogous argumentation holds for the variation $\tilde{G}(\mu) = f + \mu g_1$.

It is well-known that SDP solvers have serious issues with optimizing f for $g_1 \geqslant 0$ or $g_1 \leqslant 0$. For further information see [Nie13b, Examples 5.3 and 5.4]. ○

It is an obvious question for which classes of polynomial optimization problems the geometric programming bound developed here is better than the one given by Lasserre's relaxation. One can answer this question combinatorially. A nonnegative circuit polynomial f is a sum of squares if and only if $\text{New}(f)$ has a special lattice point structure, see Section 2.4.1, especially Theorem 2.4.14. In particular, a nonnegative circuit polynomial f cannot be a SOS if $\text{New}(f)$ is an M-simplex. Though, f is always a sum of squares if it is supported on an H-simplex. Hence, if $\text{New}(G(\boldsymbol{\mu}))$ is an M-simplex (and in more cases), then our geometric programming bounds will be better than the ones obtained by Lasserre's relaxation. However, whether a simplex is an M-simplex, an H-simplex, or something in between, is not easy to decide, see [IdW16a]. Therefore, the quality of geometric programming bounds compared to semidefinite programming bounds is very closely related to understanding these combinatorial aspects of lattice point structures in Newton polytopes.

In the last example in this section we show that for special simplices our geometric programming approach coincides with the one in [GM13].

Example 4.3.5. Suppose that $\mathrm{New}(G(\boldsymbol{\mu})) = \mathrm{conv}\{\mathbf{0}, 2d\,\mathbf{e}_1, \ldots, 2d\,\mathbf{e}_n\}$. Hence, the Newton polytope is a $2d$-scaled standard simplex in $\mathbb{R}^n$, which is the case if the pure powers x_j^{2d} for $j = 1, \ldots, n$ are present in the polynomial f or in the constrained polynomials g_i. The corresponding polynomial $G(\boldsymbol{\mu})$ is an ST-polynomial; see Section 4.1.1. Indeed, all examples in [GM13, Example 4.8] are of that form and thus all of them are ST-polynomials.

In this case the program (4.1.5) coincides with the program (3) in [GM13]. One drawback of this setting is that the GP bounds obtained from (4.1.5) are at most as good as the bound $f_{\text{sos}}^{(d)}$ itself. Namely, if the Newton polytope of a circuit polynomial is a scaled standard simplex, then it is an H-simplex. Consequently by Corollary 2.4.15 the circuit polynomial is nonnegative if and only if it is a sum of squares. Thus, if we are in the setting of Ghasemi and Marshall and $G(\boldsymbol{\mu})$ is nonnegative, then it is a sum of squares of degree at most $2d$ which guarantees the existence of a decomposition in the sense of $f_{\text{sos}}^{(d)}$; see (2.3.3).

However, as we have shown in the previous examples, in the case of our program (4.1.5) there exist also cases where the geometric programming bounds are better than $f_{\text{sos}}^{(d)}$, since our approach is more general than in [GM13]. The reason is that the cones of sums of nonnegative circuit polynomials and sums of squares do not contain each other (but both of them are contained in the cone of nonnegative polynomials), see Theorem 2.4.8 and Theorem 3.1.2. ○

We point out that we make *no* assumption about the feasible set K. In particular, it is not assumed to be compact as it is in the classical setting via Lasserre relaxations in order to guarantee convergence of the relaxations. However, the crucial point in our setting so far is that $G(\boldsymbol{\mu})$ has to be an ST-polynomial. In the following Section 4.4 we lay the foundation for the usage of our geometric programming approach also for non-ST-polynomials.

But even if $G(\boldsymbol{\mu})$ is not an ST-polynomial, then we can enforce it to be an ST-polynomial in the case of a compact K. This can be achieved by adding a redundant constraint $g_{s+1} = x_1^{2d} + \ldots + x_n^{2d} + c$ for $c \in \mathbb{R}$ to the feasible set K. In consequence $\mathrm{New}(G(\boldsymbol{\mu}))$ is a $2d$-scaled standard simplex and by the previous example our approach coincides with the one in [GM13]. Hence, the Lasserre relaxation cannot be outperformed in quality anymore. However, our approach can and will still have the better runtime. It would be interesting to add other redundant inequalities to K such that the corresponding bounds are better than the ones obtained via Lasserre relaxations. Unfortunately, no systematic way is known so far.

4.4 Optimization for Non-ST-Polynomials

The goal of this section is to provide a first approach to tackle optimization problems (both unconstrained and constrained) which cannot be expressed as a single ST-polynomial using the methods developed in the previous sections of this chapter. We also provide examples illustrating our method.

We start with the case of global nonnegativity for arbitrary polynomials via SONC certificates. Recall the following statement, which immediately follows from Section 2.4 by Definition 2.4.7 and Theorem 2.4.8.

Fact 4.4.1. *Let $f \in \mathbb{R}[\mathbf{x}]$ and assume that there exist SONC polynomials $g_1, \ldots, g_k$ and positive real numbers $\mu_1, \ldots, \mu_k$ such that $f = \sum_{i=1}^k \mu_i g_i$. Then f is nonnegative.*

Of course, if a SONC decomposition exists, then it is not obvious how to find it in general. For ST-polynomials we know that we can find a SONC decomposition via the geometric optimization problem described in Theorem 4.1.4. Thus, we investigate a general polynomial $f \in \mathbb{R}[\mathbf{x}]$ supported on a set $A \subset \mathbb{N}^n$ satisfying ($\clubsuit$). We denote

$$f \quad = \quad \sum_{j=0}^{d} f_{\boldsymbol{\alpha}(j)} \mathbf{x}^{\boldsymbol{\alpha}(j)} + \sum_{\boldsymbol{\beta} \in \Delta(f)} f_{\boldsymbol{\beta}} \mathbf{x}^{\boldsymbol{\beta}}$$

such that $f_{\boldsymbol{\alpha}(j)} \mathbf{x}^{\boldsymbol{\alpha}(j)}$ are monomial squares. By ($\clubsuit$), $V(A)$ is the set of all vertices of $\text{New}(f)$ and we have $V(A) \subseteq \{\boldsymbol{\alpha}(0), \ldots, \boldsymbol{\alpha}(d)\}$; equality, however, is not required here: that is, the set $\{\boldsymbol{\alpha}(0), \ldots, \boldsymbol{\alpha}(d)\}$ can also contain exponents of monomial squares in $\Delta(A) \backslash \Delta(f)$ which are not vertices of $\text{conv}(A)$. For simplicity we assume in what follows that the affine span of A is n-dimensional. We proceed as follows:

(1) Choose a triangulation $T_1, \ldots, T_k$ of exponents $\boldsymbol{\alpha}(0), \ldots, \boldsymbol{\alpha}(d) \in A$ corresponding to the monomial squares.

(2) Compute the induced covering $A_1, \ldots, A_k$ of A given by $A_i = A \cap T_i$ for $1 \leqslant i \leqslant k$.

(3) Assume that $\boldsymbol{\beta} \in \Delta(f) \subset A$ is contained in more than one of the A_i's. Let without loss of generality $\boldsymbol{\beta} \in A_1, \ldots, A_l$ with $1 < l \leqslant k$. Then we choose $f_{\boldsymbol{\beta},1}, \ldots, f_{\boldsymbol{\beta},l} \in \mathbb{R}$ such that $\sum_{i=1}^l f_{\boldsymbol{\beta},i} = f_{\boldsymbol{\beta}}$ and $\text{sign}(f_{\boldsymbol{\beta},i}) = \text{sign}(f_{\boldsymbol{\beta}})$ for all $1 \leqslant i \leqslant l$. We proceed analogously for $\boldsymbol{\alpha}(0), \ldots, \boldsymbol{\alpha}(d)$.

(4) Define new polynomials $g_1, \ldots, g_k$ such that

$$g_i = \sum_{\boldsymbol{\beta} \in A_i} f_{\boldsymbol{\beta},i} \mathbf{x}^{\boldsymbol{\beta}}.$$

Note that by (1) and (2) the covering A_i is a set of integer tuples such that $\mathrm{conv}(A_i)$ is a simplex with even vertices and A_i contains no even points corresponding to monomial squares except for the vertices of $\mathrm{conv}(A_i)$. Thus, by (2) – (4) we see that all g_i are ST-polynomials, since the signs of the $f_{\boldsymbol{\beta},i}$ are identical with the signs of the coefficients of f. Therefore, monomial squares $f_{\boldsymbol{\alpha}(j)} \mathbf{x}^{\boldsymbol{\alpha}(j)}$ of f get decomposed into a sum of monomial squares $\sum_{i=1}^{k} f_{\boldsymbol{\alpha}(j),i} \mathbf{x}^{\boldsymbol{\alpha}(j)}$ such that each individual monomial square $f_{\boldsymbol{\alpha}(j),i} \mathbf{x}^{\boldsymbol{\alpha}(j)}$ is a term of exactly one g_i. We proceed analogously for the terms $f_{\boldsymbol{\beta}} \mathbf{x}^{\boldsymbol{\beta}}$. Additionally, it follows by construction that $f = \sum_{i=1}^{k} g_i$. We apply the GP proposed in Corollary 4.1.5 on each of the g_i with respect to a monomial square $f_{\boldsymbol{\alpha}(j),i} \mathbf{x}^{\boldsymbol{\alpha}(j)}$, which is a vertex of $\mathrm{New}(g_i) = \mathrm{conv}(A_i)$ (not necessarily the same $\boldsymbol{\alpha}(j)$ for every g_i); we denote the minimizer by m_i^*. We make the following observation about these minimizers which was similarly already pointed out in [IdW16a, Section 3]:

Lemma 4.4.2. *Let $f \in \mathbb{R}[\mathbf{x}]$ be a nonnegative circuit polynomial. Let $b_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}}$ be a monomial square. Then $b_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}} \cdot f$ is also a nonnegative circuit polynomial.*

Note particularly that if $\mathbf{v} \in (\mathbb{R}^*)^n$ satisfies $f(\mathbf{v}) = 0$, then $(b_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}} \cdot f)(\mathbf{v}) = 0$.

Proof. It is easy to see that all conditions for (♣) as well as the conditions (ST1) and (ST2) remain valid for $b_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}} \cdot f$. Thus, $b_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}} \cdot f$ still is a circuit polynomial and since $b_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}} \geqslant 0$ it is also nonnegative. □

Proposition 4.4.3. *Let f, $g_1, \ldots, g_k$, and m_i^* be as explained above. Assume for $i = 1, \ldots, k$ that the minimizer m_i^* corresponds to the monomial square $f_{\boldsymbol{\alpha}(j),i} \mathbf{x}^{\boldsymbol{\alpha}(j_i)}$ with $\boldsymbol{\alpha}(j_i) \in \{\boldsymbol{\alpha}(1), \ldots, \boldsymbol{\alpha}(d)\} \cap V(A_i)$. Then $f - \sum_{i=1}^{k} m_i^* \mathbf{x}^{\boldsymbol{\alpha}(j_i)}$ is a SONC and hence nonnegative. Thus, the m_i^* provide bounds for the coefficients $f_{\boldsymbol{\alpha}(j),i}$ for f to be non-negative. Particularly, if for $i = 1, \ldots, l$ with $l \leqslant k$ the exponents $\boldsymbol{\alpha}(j_i)$ are the origin, then $f_{\boldsymbol{\alpha}(0)} - \sum_{i=1}^{l} m_i^*$ is a lower bound for $f^* = \sup\{\gamma \in \mathbb{R} : f - \gamma \geqslant 0\}$.*

Proof. By construction, we know that $g_i - m_i^* \mathbf{x}^{\boldsymbol{\alpha}(j_i)}$ is a SONC. Thus, the polynomial $f - \sum_{i=1}^{k} m_i^* \mathbf{x}^{\boldsymbol{\alpha}(j_i)} = \sum_{i=1}^{k} g_i - m_i^* \mathbf{x}^{\boldsymbol{\alpha}(j_i)}$ is a SONC, too. The last part follows by the definitions of the m_i^*'s and f^*. □

Note that the decomposition of f into the g_i's is not unique. First, the triangulation in (1) is not unique in general. And, second, the decomposition of the terms in (3) is arbitrary. Note also that there exist several monomial squares which appear in more than one g_i, since membership in A_i is given by the chosen triangulation and every simplex T_1 intersects at least one other simplex T_2 in an $n-1$-dimensional face, which means that $A_1 \cap A_2$ contains at least n even elements.

It would be interesting to study the problem of identifying an optimal triangulation and an optimal decomposition of coefficients, which would certainly enhance the proposed method.

We provide some examples to show how this generalized approach can be used in practice.

Example 4.4.4. Let $f = 6 + x_1^2 x_2^6 + 2x_1^4 x_2^6 + x_1^8 x_2^2 - 1.2x_1^2 x_2^3 - 0.85x_1^3 x_2^5 - 0.9x_1^4 x_2^3 - 0.73x_1^5 x_2^2 - 1.14x_1^7 x_2^2$. We choose a triangulation

$$\{\mathbf{(0,0)}, \mathbf{(2,6)}, \mathbf{(4,6)}, (2,3), (3,5)\}, \{\mathbf{(0,0)}, \mathbf{(4,6)}, \mathbf{(8,2)}, (2,3), (4,3), (5,2), (7,2)\}.$$

Here and in what follows the vertices of each simplex are printed in red. For the corresponding Newton polytope see Figure 4.2. We split the coefficients equally among the two triangulations and obtain two ST-polynomials

$$\begin{aligned} g_1 &= 3 + x_1^2 x_2^6 + x_1^4 x_2^6 - 0.6x_1^2 x_2^3 - 0.85x_1^3 x_2^5, \text{ and} \\ g_2 &= 3 + x_1^4 x_2^6 + x_1^8 x_2^2 - 0.6x_1^2 x_2^3 - 0.9x_1^4 x_2^3 - 0.73x_1^5 x_2^2 - 1.14x_1^7 x_2^2. \end{aligned}$$

Using CVX, we apply the GP from Corollary 4.1.5 and obtain optimal values $m_1^* = 0.2121$, $m_2^* = 2.5193$, and a SONC decomposition

$$\begin{array}{lll} 0.173 + \varepsilon x_1^2 x_2^6 + 0.522x_1^4 x_2^6 - 0.6x_1^2 x_2^3 & + & 0.04 + x_1^2 x_2^6 + 0.478x_1^4 x_2^6 - 0.85x_1^3 x_2^5 \quad + \\ 0.427 + 0.211x_1^4 x_2^6 + \varepsilon x_1^8 x_2^2 - 0.6x_1^2 x_2^3 & + & 0.663 + 0.436x_1^4 x_2^6 + 0.085x_1^8 x_2^2 - 0.9x_1^4 x_2^3 \quad + \\ 0.753 + 0.186x_1^4 x_2^6 + 0.177x_1^8 x_2^2 - 0.73x_1^5 x_2^2 & + & 0.676 + 0.167x_1^4 x_2^6 + 0.738x_1^8 x_2^2 - 1.14x_1^7 x_2^2, \end{array}$$

with $\varepsilon < 10^{-10}$, i.e., ε is numerically zero. Namely, $(2,3)$ is located on the segment given by $\mathbf{(0,0)}$ and $\mathbf{(4,6)}$ and thus $\mathbf{(2,6)}$ and $\mathbf{(8,2)}$ have coefficients zero in the convex combinations of the point $(2,3)$.

Thus, the optimal value f_{sonc}, which provides us a lower bound for f^*, is $f_{\text{sonc}} \approx 6 - 2.731 = 3.269$. In comparison, via Lasserre's relaxation one obtains an only slightly better optimal value $f^* = 3.8673$.

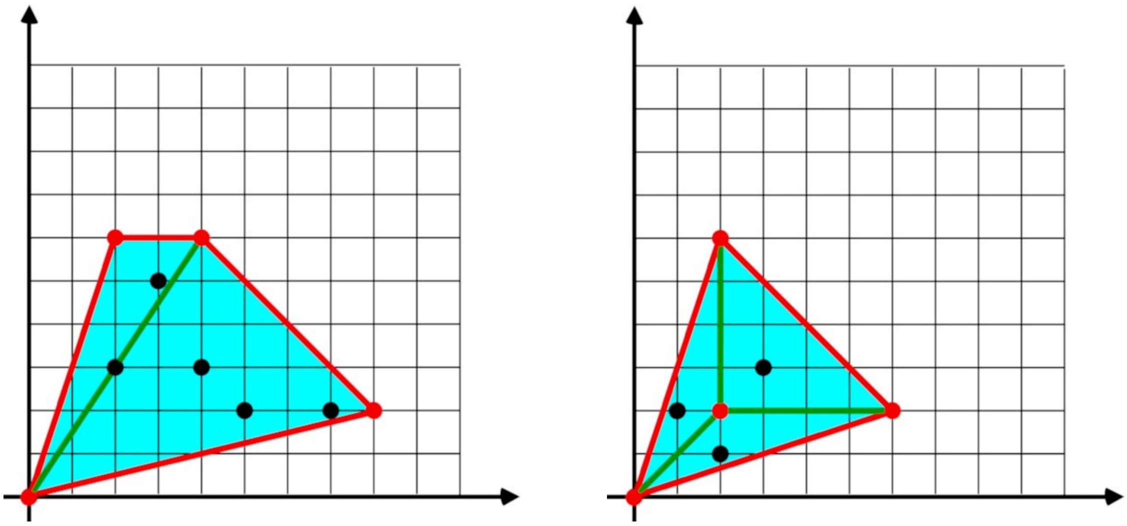

Figure 4.2: The Newton polytopes of the polynomials in the Examples 4.4.4 and 4.4.5 and their triangulations.

Our GP-based bound can be improved significantly via making small changes in the distribution of the coefficients. For example, if one decides not to split the coefficient of the term $x_1^2x_2^3$ among g_1 and g_2 equally, but to put the entire weight of the coefficient into g_1, i.e.,

$$\begin{aligned}\tilde{g}_1 &= 3 + x_1^2x_2^6 + x_1^4x_2^6 - 1.2x_1^2x_2^3 - 0.85x_1^3x_2^5, \text{ and}\\ \tilde{g}_2 &= 3 + x_1^4x_2^6 + x_1^8x_2^2 - 0.9x_1^4x_2^3 - 0.73x_1^5x_2^2 - 1.14x_1^7x_2^2,\end{aligned}$$

then this yields an improved bound $\tilde{f}_{\text{sonc}} \approx 3.572$. ○

The next example shows that we can use the approach of this section to take monomial squares into account, which are not vertices of the Newton polytope of the polynomial which we intend to minimize.

Example 4.4.5. Let $f = 1 + 3x_1^2x_2^6 + 2x_1^6x_2^2 + 6x_1^2x_2^2 - x_1x_2^2 - 2x_1^2x_2 - 3x_1^3x_2^3$. We choose a triangulation

$$\{(\mathbf{0},\mathbf{0}),(\mathbf{2},\mathbf{2}),(\mathbf{2},\mathbf{6}),(1,2)\},\{(\mathbf{0},\mathbf{0}),(\mathbf{2},\mathbf{2}),(\mathbf{6},\mathbf{2}),(2,1)\},\{(\mathbf{2},\mathbf{2}),(\mathbf{2},\mathbf{6}),(\mathbf{6},\mathbf{2}),(3,3)\}.$$

For the corresponding Newton polytope see Figure 4.2. First, we split the coefficients equally among the three triangulations such that we obtain

$$\begin{aligned} g_1 &= 0.5 + 1.5x_1^2x_2^6 + 2x_1^2x_2^2 - x_1x_2^2, \\ g_2 &= 0.5 + x_1^6x_2^2 + 2x_1^2x_2^2 - 2x_1^2x_2, \\ g_3 &= 1.5x_1^2x_2^6 + x_1^6x_2^2 + 2x_1^2x_2^2 - 3x_1^3x_2^3. \end{aligned}$$

All three polynomials g_i have a joint monomial $x_1^2x_2^2$. For all $i \in \{1,2,3\}$ we compute the maximal $b_i > 0$ such that $g_i - b_ix_1^2x_2^2$ is a nonnegative circuit polynomial. This yields a bound for the coefficient of $x_1^2x_2^2$ certifying that f is a SONC and hence nonnegative. We could apply the GP from Corollary 4.1.5, but since all g_i are circuit polynomials we can compute the corresponding circuit numbers symbolically. We obtain with Theorem 2.4.4:

$$\begin{aligned} \Theta_{g_1}(1,2) &= \left(\frac{1/2}{1/2}\right)^{\frac{1}{2}} \cdot \left(\frac{3/2}{1/4}\right)^{\frac{1}{4}} \cdot \left(\frac{2-b_1}{1/4}\right)^{\frac{1}{4}} = \sqrt[4]{4 \cdot 4 \cdot 3/2 \cdot (2-b_1)} = 2\sqrt[4]{3/2 \cdot (2-b_1)}, \\ \Theta_{g_2}(2,1) &= \left(\frac{1/4}{1/2}\right)^{\frac{1}{2}} \cdot \left(\frac{1/2}{1/4}\right)^{\frac{1}{4}} \cdot \left(\frac{1-1/2 \cdot b_2}{1/4}\right)^{\frac{1}{4}} = \sqrt[4]{1/4 \cdot 2 \cdot 4(1-1/2 \cdot b_2)} = \sqrt[4]{2-b_2}, \text{ and} \\ \Theta_{g_3}(3,3) &= \left(\frac{1/2}{1/4}\right)^{\frac{1}{4}} \cdot \left(\frac{1/3}{1/4}\right)^{\frac{1}{4}} \cdot \left(\frac{1/3(2-b_3)}{1/2}\right)^{\frac{1}{2}} = \sqrt[4]{2 \cdot 4/3} \cdot \sqrt{2/3 \cdot (2-b_3)} = 2\sqrt[4]{2/27}\sqrt{2-b_3}. \end{aligned}$$

This provides solutions:

$$\begin{aligned} 2\sqrt[4]{3/2 \cdot (2-b_1)} &\geqslant 1 \Leftrightarrow 3/2 \cdot (2-b_1) \geqslant 1/16 \Leftrightarrow b_1 \leqslant 47/24, \\ \sqrt[4]{2-b_2} &\geqslant 1 \Leftrightarrow b_2 \leqslant 1, \\ 2\sqrt[4]{2/27}\sqrt{2-b_3} &\geqslant 1 \Leftrightarrow \sqrt{2/27} \cdot (2-b_3) \geqslant 1/4 \Leftrightarrow b_3 \leqslant 2 - \sqrt{27}/(2\sqrt{2}). \end{aligned}$$

Hence, we obtain the following bound for the coefficient of $x_1^2x_2^2$:

$$6 - (47/24 + 1 + 2 - \sqrt{27}/(4\sqrt{2})) \approx 6 - 4.03977468 \approx 1.96.$$

A double check with the CVX solver for GPs yields the same value in approximately 0.753 seconds.

We want to compute a bound for f^*. We choose the same triangulation and the same split of coefficients as before, but now we optimize the constant term in g_1 and g_2, and we optimize the coefficient of $x_1^2x_2^6$ in g_3. After a runtime of approximately 0.6657 seconds we obtain optimal values $0.0722, 0.3536$, and 0.3164. Thus, we found a lower bound for the constant term given by

$$m_1^* + m_2^* \approx 0.0722 + 0.3536 = 0.4268.$$

The corresponding optimal SONC decomposition is given by

$$0.0722 + 1.5x_1^2x_2^6 + 2x_1^2x_2^2 - x_1x_2^2 \quad + \quad 0.3536 + x_1^6x_2^2 + 2x_1^2x_2^2 - x_1^2x_2 \quad + \\ 0.3164x_1^2x_2^6 + x_1^6x_2^2 + 2x_1^2x_2^2 - 3x_1^3x_2^3.$$

Hence, we obtain a bound for f^* given by

$$f_{\text{sonc}} \;=\; 1 - 0.4268 \;=\; 0.5732.$$

We make a comparison and optimize f with Lasserre's relaxation. This yields an optimal value

$$f_{\text{sos}} \;=\; f^* \;\approx\; 0.8383.$$

Therefore, we want to improve our bound. We keep the triangulation, but we use another distribution of the coefficients among the polynomials g_1, g_2, and g_3 and define instead

$$\begin{aligned} \tilde{g}_1 &= 0.25 + 2x_1^2x_2^6 + 1.217x_1^2x_2^2 - 2x_1x_2^2, \\ \tilde{g}_2 &= 0.75 + x_1^6x_2^2 + 3.652x_1^2x_2^2 - x_1^2x_2, \\ \tilde{g}_3 &= x_1^2x_2^6 + x_1^6x_2^2 + 1.13x_1^2x_2^2 - 3x_1^3x_2^3. \end{aligned}$$

Again, we optimize $\tilde{g}_1$ and $\tilde{g}_2$ with respect to the constant term and $\tilde{g}_3$ with respect to $x_1^2x_2^6$. We obtain optimal values $0.0801, 0.2616$, and 0.9912. Thus, we are able to improve our bound for f^* to

$$\tilde{f}_{\text{sonc}} \;\approx\; 1 - (0.0801 + 0.2616) \;=\; 0.6583.$$

The corresponding optimal SONC decomposition is given by

$$0.0801 + 2x_1^2x_2^6 + 1.205x_1^2x_2^2 - 2x_1x_2^2 \quad + \quad 0.2616 + x_1^6x_2^2 + 3.615x_1^2x_2^2 - x_1^2x_2 \quad + \\ 0.991x_1^2x_2^6 + x_1^6x_2^2 + 2x_1^2x_2^2 - 3x_1^3x_2^3.$$

⬡

We discuss a third example which shows that, in the case of global optimization, for the SONC/GP approach it is not necessary to optimize the constant term to obtain a bound for nonnegativity on the coefficients, but that in some cases it can be informative to focus on other vertices of the Newton polytope or on other monomial squares instead.

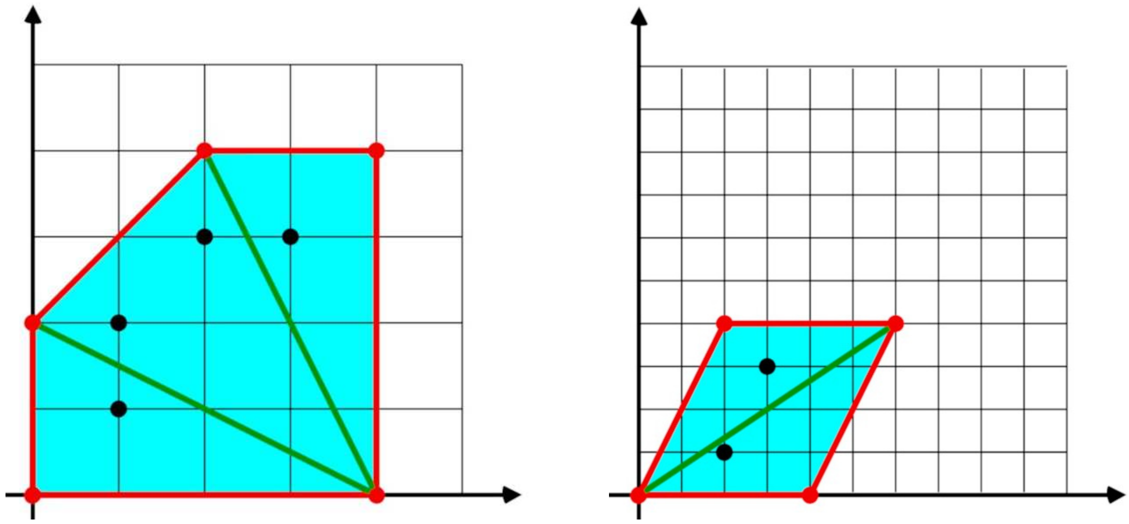

Figure 4.3: The Newton polytopes of the polynomials in the Examples 4.4.6 and 4.4.7 and their triangulations.

Example 4.4.6. Let $f = 1 + x_1^4 + x_2^2 + x_1^2x_2^4 + x_1^4x_2^4 - x_1x_2 - x_1x_2^2 - x_1^2x_2^3 - x_1^3x_2^3$. We choose a triangulation

$$\{\mathbf{(0,0)}, \mathbf{(0,2)}, \mathbf{(4,0)}, (1,1)\}, \{\mathbf{(0,2)}, \mathbf{(2,4)}, \mathbf{(4,0)}, (1,2), (2,3)\}, \{\mathbf{(2,4)}, \mathbf{(4,0)}, \mathbf{(4,4)}, (3,3)\},$$

and again, we choose a decomposition of coefficients such that their values split equally. We obtain the following ST-polynomials

$$\begin{aligned} g_1 &= 1 + 1/3 \cdot x_1^4 + 1/2 \cdot x_2^2 - x_1x_2, \\ g_2 &= 1/3 \cdot x_1^4 + 1/2 \cdot x_1^2x_2^4 + 1/2 \cdot x_2^2 - x_1x_2^2 - x_1^2x_2^3, \\ g_3 &= 1/3 \cdot x_1^4 + 1/2 \cdot x_1^2x_2^4 + x_1^4x_2^4 - x_1^3x_2^3. \end{aligned}$$

Actually g_1 and g_3 are circuit polynomials while g_2 contains two negative terms. For the corresponding Newton polytope see Figure 4.3. Note that $\mathbf{(4,0)}$ is the only exponent contained in the support of all three ST-polynomials. Since $\mathbf{(4,0)}$ is a monomial square which is a vertex of the convex hull of the three support sets, we optimize the corresponding coefficient in g_1, g_2, and g_3. Applying the GP from Corollary 4.1.5 yields optimal values

$$m_1^* = 0.0625, \quad m_2^* = 4.2867, \quad \text{and} \quad m_3^* = 0.0625.$$

Since $m_2^* = 4.2867 > 1/3$ we found no certificate of nonnegativity for f. However, we can identify a SONC decomposition for f provided that the coefficient $b_{(4,0)}$ of x_1^4 is at least $m_1^* + m_2^* + m_3^* = 4.412$. For this minimal choice of $b_{(4,0)}$ a SONC decomposition is given by

$$\begin{array}{lcll} 0.063x_1^4 + 1 + 0.5x_2^2 - x_1x_2 & + & 2.143x_1^4 + 0.4x_2^2 + 0.1x_1^2x_2^4 - x_1x_2^2 & + \\ 2.143x_1^4 + 0.1x_2^2 + 0.4x_1^2x_2^4 - x_1^2x_2^3 & + & 0.063x_1^4 + 0.5x_1^2x_2^4 + x_1^4x_2^4 - x_1^3x_2^3. & \end{array}$$

○

Finally, we apply the new method to a constrained optimization problem using the methods developed in Section 4.2.

Example 4.4.7. Let $f = 1 + x^4 + x^2y^4$ and $g = \frac{1}{2} + x^2y - x^6y^4 - x^3y^3$. Hence, we obtain $G(\mu) = (1 - \frac{1}{2}\mu) + x^4 + x^2y^4 + \mu x^6y^4 - \mu x^2y + \mu x^3y^3$. Choosing the triangulation

$$\{(\mathbf{0},\mathbf{0}),(\mathbf{4},\mathbf{0}),(\mathbf{6},\mathbf{4}),(2,1)\},\{(\mathbf{0},\mathbf{0}),(\mathbf{6},\mathbf{4}),(\mathbf{2},\mathbf{4}),(3,3)\},$$

we split the coefficients again, such that their values are equal. For the corresponding Newton polytope see Figure 4.3. We obtain the ST-polynomials

$$\begin{aligned} G_1(\mu) &= \left(\frac{1}{2} - \frac{1}{4}\mu\right) + x^4 + \frac{1}{2}\mu x^6y^4 - \mu x^2y, \\ G_2(\mu) &= \left(\frac{1}{2} - \frac{1}{4}\mu\right) + x^2y^4 + \frac{1}{2}\mu x^6y^4 + \mu x^3y^3. \end{aligned}$$

Therefore, we see that the possible μ values to obtain ST-polynomials are $\mu \in [0, 2)$. We optimize both polynomials with respect to the constant term and obtain $m_1^* = m_2^* = 0$. The CVX solver yields "`NaN`" as an optimal value, since 0 is not positive. However, it solves the problem and computes values 0 or $\varepsilon < 10^{-200}$ for all variables, such that $m_1^* = m_2^* = 0$ follows. Hence, $f_{\boldsymbol{\alpha}(0)} - m^* = 1 - 0 = 1$ and because all of the assumptions in Theorem 4.2.5 are satisfied we know that $s(f, g) = 1$ holds.

Checking this optimization problem with Lasserre's relaxation, we get $f_{\text{sos}} = f_K^* = 1$, which approves the optimal value. Both, for the SDP and the GP we have runtimes below 1 second.

Now, we tackle the same problem, but we multiply every exponent by 10, and we compare the runtimes again. For the GP we obtain the same result and the runtime remains below 1 second. For the SDP we obtain with GloptiPoly $f_{\text{sos}} = f_K^* = 1$ in approximately 5034.5 seconds, i.e., approximately 1.4 hours.

In a third approach we tackle the same problem, but we multiply the originally given exponents by 20. In this case GLOPTIPOLY is not able to handle the given matrices anymore. In comparison, we still have a runtime below 1 second for our GP providing the same bound as before. ○

4.5 Conclusion

In this chapter we have focused on tackling polynomial optimization problems via the SONC/GP method, which was introduced for the unconstrained case for ST-polynomials. Particularly, we extended this method for those polynomials to (CPOPs) where we traced back the constrained problem to the unconstrained ones by defining a new polynomial, which incorporates both the objective function and the constrained polynomials.

Furthermore, we were able to apply both unconstrained and constrained polynomial optimization methods based on SONC and GPs efficiently beyond the class of ST-polynomials. In this case we found GP-based lower bounds via triangulations of the support sets of the involved polynomials. These general types of nonnegativity problems have to be investigated more carefully.

In sum, the current status of the question, whether a given a polynomial $f \in \mathbb{R}[\mathbf{x}]$ is a SONC polynomial, can be decided as follows:

- If f is an ST-polynomial this is a geometric program.
- For an arbitrary f by triangulation of $\mathrm{New}(f)$ the GP-based method provides a first approach to answer this question. To decide the problem properly it is not sufficient to study only one triangulation but we have to take all possible even simplices into account which cover $\mathrm{New}(f)$.

Moreover, we provided a comparison of our SONC/GP method to Lasserre's relaxation. We observed that the new approach comes with the benefit that GPs can be solved much faster than SDPs, especially for high-degree polynomials. Since the SOS and the SONC cone are not contained in each other, both approaches handle different types of polynomials differently well.

However, the key strength of Lasserre's relaxation is that it yields a converging hierarchy of lower bounds which allows to approximate the optimal value arbitrarily close, whereas the methods for SONC only use a single geometric optimization program to derive a lower bound for f_K^*. In the next chapter we close this gap.

Chapter 5

Hierarchical Approach to Constrained Optimization Problems via SONC and REP

Motivated by the results of the preceding chapter we now provide an hierarchical approach to constrained polynomial optimization problems on a compact set K. Based on a Positivstellensatz involving SONC polynomials we establish a hierarchy of lower bounds converging to the minimum of a polynomial on K. Positivstellensätze are an essential tool from real algebraic geometry to tackle (CPOPs), for an introduction to this topic see Section 2.3.3. Moreover, we show that these bounds can be computed efficiently via *relative entropy programming (REP)*. Particularly, all results are independent of sums of squares and semidefinite programming.

We begin with studying the cone of *sums of nonnegative arithmetic geometric exponentials (SAGE)*, an interesting cone related to the SONC cone introduced by Chandrasekaran and Shah [CS16] as well as relative entropy programming. This is a convex optimization program, which is more general than a geometric program, but still efficiently solvable via interior point methods; see [CS17, NN94]. Afterwards we discuss the relation between the SAGE and the SONC cone in more detail, see Section 5.2. In the subsequent Section 5.3 we formulate the Positivstellensatz using SONC polynomials; see Theorem 5.3.5. The following statement is a rough version.

Theorem 5.0.1 (Positivstellensatz for SONC; rough version). *Let $f \in \mathbb{R}[\mathbf{x}]$ be a real polynomial which is strictly positive on a given compact, basic closed semialgebraic set K defined by polynomials $g_1, \ldots, g_s \in \mathbb{R}[\mathbf{x}]$. Then there exists an explicit representation of f as a sum of products of the g_i's and SONC polynomials.*

The Positivstellensatz yields a *hierarchy* of lower bounds $f_{\text{sonc}}^{(d,q)}$ for f_K^* based on the maximal allowed degree of the representing polynomials in the Positivstellensatz. We show in Theorem 5.4.2 that the bounds $f_{\text{sonc}}^{(d,q)}$ converge to f_K^* for $d, q \to \infty$.

Finally, we provide in (5.4.3) an optimization program for the computation of $f_{\text{sonc}}^{(d,q)}$. We prove in Theorem 5.4.3 that our program (5.4.3) is a relative entropy program, and hence is efficiently solvable. Concluding we discuss an example, see Section 5.4.3.

5.1 Preliminaries

In this section we recall key results about sums of nonnegative arithmetic geometric exponentials (SAGE), and relative entropy programing (REP), which are used in this chapter.

5.1.1 Relative Entropy and the SAGE Cone

There exists an important concept related to the SONC cone, which was introduced by Chandrasekaran and Shah in [CS16], namely the *cone of sums of nonnegative arithmetic geometric exponentials*. In what follows, we introduce relative entropy programs and the SAGE cone. Later, in Section 5.2, we discuss its relationship to SONC polynomials and how we can use relative entropy programming for our results.

Recall that we denote by $\langle \cdot, \cdot \rangle$ the standard inner product. Following [CS16], a *signomial* is a sum of exponentials

$$f(\mathbf{x}) \;=\; \sum_{j=0}^{l} f_{\boldsymbol{\alpha}(j)} e^{\langle \boldsymbol{\alpha}(j), \mathbf{x} \rangle}$$

with $f_{\boldsymbol{\alpha}(j)} \in \mathbb{R}, \mathbf{x} \in \mathbb{R}^n$, and real vectors $\boldsymbol{\alpha}(0), \ldots, \boldsymbol{\alpha}(l) \in \mathbb{R}^n$. A signomial with at most one negative coefficient is called an *AM/GM-exponential*. Thus, an AM/GM-exponential has the following form

$$f(\mathbf{x}) \;=\; \sum_{j=0}^{l} f_{\boldsymbol{\alpha}(j)} e^{\langle \boldsymbol{\alpha}(j), \mathbf{x} \rangle} + f_{\boldsymbol{\beta}} \cdot e^{\langle \boldsymbol{\beta}, \mathbf{x} \rangle},$$

where $f_{\boldsymbol{\beta}} \in \mathbb{R}, f_{\boldsymbol{\alpha}(j)} \in \mathbb{R}_{>0}$, and $\boldsymbol{\beta}, \boldsymbol{\alpha}(j) \in \mathbb{R}^n$ for $j = 0, \ldots, l$. Note that $l > n$ is possible.

As shown in [CS16], testing whether an AM/GM-exponential is nonnegative is possible via the *relative entropy function*. This function is defined as follows for

$\boldsymbol{\nu} = (\nu_0, \ldots, \nu_l)$ and $\boldsymbol{\zeta} = (\zeta_0, \ldots, \zeta_l)$ in the nonnegative orthant $\mathbb{R}^{l+1}_{\geqslant 0}$:

$$D(\boldsymbol{\nu}, \boldsymbol{\zeta}) \;=\; \sum_{j=0}^{l} \nu_j \log\left(\frac{\nu_j}{\zeta_j}\right).$$

By convention, we define $0 \log \frac{0}{\zeta_j} = 0$ for any $\zeta_j \in \mathbb{R}_{\geqslant 0}$, and $\nu_j \log \frac{\nu_j}{0} = 0$ if $\nu_j = 0$ and $\nu_j \log \frac{\nu_j}{0} = \infty$ if $\nu_j > 0$. Furthermore, let $\boldsymbol{f_\alpha} = (f_{\boldsymbol{\alpha}(0)}, \ldots, f_{\boldsymbol{\alpha}(l)}) \in \mathbb{R}^{l+1}_{>0}$. Then the following lemma holds.

Lemma 5.1.1 ([CS16], Lemma 2.2)**.** *Let $f(\mathbf{x})$ be an AM/GM-exponential. Then $f(\mathbf{x})$ is nonnegative for all $\mathbf{x} \in \mathbb{R}^n$ if and only if there exists a $\boldsymbol{\nu} \in \mathbb{R}^{l+1}_{\geqslant 0}$ satisfying the conditions*

$$(5.1.1) \qquad D(\boldsymbol{\nu}, e\boldsymbol{f_\alpha}) - f_{\boldsymbol{\beta}} \leqslant 0\ ,\ \boldsymbol{Q\nu} = \langle \mathbf{1}, \boldsymbol{\nu} \rangle \boldsymbol{\beta} \ \text{ with } \ \boldsymbol{Q} = (\boldsymbol{\alpha}(0) \cdots \boldsymbol{\alpha}(l)) \in \mathbb{R}^{n \times (l+1)}.$$

Checking whether such a vector $\boldsymbol{\nu} \in \mathbb{R}^{l+1}_{\geqslant 0}$ exists is a convex optimization problem by means of the joint convexity of the relative entropy function $D(\boldsymbol{\nu}, \boldsymbol{\zeta})$. More specifically, the corresponding problem is a *relative entropy program*; see [CS17].

Definition 5.1.2. Let $\boldsymbol{\nu}, \boldsymbol{\zeta} \in \mathbb{R}^{l+1}_{\geqslant 0}$ and $\boldsymbol{\delta} \in \mathbb{R}^{l+1}$. A *relative entropy program (REP)* is of the form:

$$(5.1.2) \qquad \begin{cases} \text{minimize} & p_0(\boldsymbol{\nu}, \boldsymbol{\zeta}, \boldsymbol{\delta}), \\ \text{subject to:} & \begin{array}{ll} (1) & p_i(\boldsymbol{\nu}, \boldsymbol{\zeta}, \boldsymbol{\delta}) \leqslant 1 \ \text{ for all } \ i = 1, \ldots, m, \\ (2) & \nu_j \log\left(\frac{\nu_j}{\zeta_j}\right) \leqslant \delta_j \ \text{ for all } \ j = 0, \ldots, l, \end{array} \end{cases}$$

where $p_0, \ldots, p_m$ are linear functionals and the constraints (2) are jointly convex functions in $\boldsymbol{\nu}, \boldsymbol{\zeta}$, and $\boldsymbol{\delta}$ defining the relative entropy cone. ○

Relative entropy programs are convex and can be solved efficiently via interior point methods [NN94]. Geometric programs, a prominent class of convex optimization programs [BKVH07, BV04, DPZ67] which were discussed in Chapter 4, comprise a subclass of REPs; see [CS17] for further information.

If a signomial consists of more than one negative term, then a natural and sufficient condition for certifying nonnegativity is to express the signomial as a sum of nonnegative

AM/GM-exponentials. For a finite set of exponents $M \subset \mathbb{R}^n$, one denotes by

$$SAGE(M) \;=\; \left\{ f = \sum_{i=1}^{m} f_i \;:\; \begin{array}{c} \text{every } f_i \text{ is a nonnegative AM/GM-exponential} \\ \text{with exponents in } M \end{array} \right\}$$

the set of *sums of nonnegative AM/GM-exponentials (SAGE)* with respect to M; see [CS16].

5.1.2 Signomials and Polynomials

The connection between signomials and polynomials is given by the bijective componentwise exponential function

$$\exp : \mathbb{R}^n \to \mathbb{R}^n_{>0}, \quad (x_1, \ldots, x_n) \mapsto (e^{x_1}, \ldots, e^{x_n}).$$

Via this mapping a signomial

$$f(\mathbf{x}) \;=\; \sum_{j=0}^{l} f_{\boldsymbol{\alpha}(j)} e^{\langle \boldsymbol{\alpha}(j), \mathbf{x} \rangle}$$

is transformed into

$$f(\mathbf{x}) \;=\; \sum_{j=0}^{l} f_{\boldsymbol{\alpha}(j)} \mathbf{x}^{\boldsymbol{\alpha}(j)},$$

which is a polynomial if $\boldsymbol{\alpha}(0), \ldots, \boldsymbol{\alpha}(l) \in \mathbb{N}^n$. Hence, checking nonnegativity of such signomials corresponds to checking nonnegativity of *a polynomial on the positive orthant.* Note that it is sufficient to consider the positive orthant to certify nonnegativity, since the positive orthant is dense in the nonnegative orthant. We call such a polynomial $f(\mathbf{x}) = \sum_{j=0}^{l} f_{\boldsymbol{\alpha}(j)} \mathbf{x}^{\boldsymbol{\alpha}(j)}$ a *SAGE polynomial*, and we call it an *AM/GM-polynomial* if it has at most one negative coefficient.

5.2 A Comparison of SAGE and SONC

The concept of SAGE polynomials explicitly addresses the question of nonnegativity of polynomials on $\mathbb{R}^n_{>0}$. However, Iliman and de Wolff showed already before the development of the SAGE class that for circuit polynomials global nonnegativity coincides with nonnegativity on $\mathbb{R}^n_{>0}$ assuming that its inner term is negative; see [IdW16a, particularly Section 3.1] and also Section 2.4. This fact was, next to the circuit number, the key motivation to consider the class of circuit polynomials. Hence, in what follows we can use results from the analysis of the SAGE cone applied to circuit polynomials as a certificate for *global* nonnegativity rather than just nonnegativity on $\mathbb{R}^n_{>0}$.

Let $f(\mathbf{x}) = \sum_{j=0}^r f_{\boldsymbol{\alpha}(j)}\mathbf{x}^{\boldsymbol{\alpha}(j)} + f_{\boldsymbol{\beta}}\mathbf{x}^{\boldsymbol{\beta}}$ be a proper circuit polynomial, i.e., f is not a sum of monomial squares. We can assume without loss of generality that $f_{\boldsymbol{\beta}} < 0$ after a possible transformation of variables $x_j \mapsto -x_j$. In this case, we have

$$(5.2.1) \qquad f(\mathbf{x}) \geqslant 0 \text{ for all } \mathbf{x} \in \mathbb{R}^n \Longleftrightarrow f(\mathbf{x}) \geqslant 0 \text{ for all } \mathbf{x} \in \mathbb{R}^n_{>0};$$

see [IdW16a, Section 3.1]. Using this fact, we can characterize the corresponding AM/GM-exponential coming from a circuit polynomial under the exp-map. We call this a *simplicial AM/GM-exponential.*

Proposition 5.2.1. *Let f be a nonnegative simplicial AM/GM-exponential with interior point $\boldsymbol{\beta}$. Then* (5.1.1) *is always satisfied for the probability measure $\nu_j = \lambda_j$ for $j = 0, \ldots, r$ where λ_j is the j-th coefficient in the convex combination of the interior point $\boldsymbol{\beta} \in \mathbb{N}^n$ with respect to the vertices $\boldsymbol{\alpha}(0), \ldots, \boldsymbol{\alpha}(r) \in (2\mathbb{N})^n$.*

Proof. By (5.2.1) it is sufficient to investigate circuit polynomials. The proof follows from Theorem 2.4.4 where nonnegativity of circuit polynomials is explicitly characterized via the circuit number and hence by the convex combination of the interior point $\boldsymbol{\beta}$ in terms of the vertices $\boldsymbol{\alpha}(0), \ldots, \boldsymbol{\alpha}(r)$. The coefficients $\lambda_0, \ldots, \lambda_r$ in the convex combination form a probability measure by definition. □

The circuit number is defined via barycentric coordinates; see Section 2.4. This parametrization for nonnegativity corresponds to the geometric programming literature; see [CS16, (2.2), page 1151] and also [DPZ67]:

$$(5.2.2) \qquad D(\boldsymbol{\nu}, \boldsymbol{f_\alpha}) + \log(-f_{\boldsymbol{\beta}}) \leqslant 0, \ \boldsymbol{\nu} \in \mathbb{R}^{l+1}_{\geqslant 0}, \ \boldsymbol{Q\nu} = \boldsymbol{\beta}, \ \langle \mathbf{1}, \boldsymbol{\nu} \rangle = 1.$$

Note that we assume $f_{\boldsymbol{\beta}} < 0$ here. Chandrasekaran and Shah showed that the conditions (5.1.1) and (5.2.2) are equivalent (this is non-obvious); see [CS16]. However, they also point out therein that restricting $\boldsymbol{\nu}$ to a probability measure as in (5.2.2) comes with the drawback that the parametrization in (5.2.2) is not *jointly convex* in $\boldsymbol{\nu}, \boldsymbol{f_\alpha}$, and $f_{\boldsymbol{\beta}}$. This is in sharp contrast to the parametrization (5.1.1), which *is* jointly convex in $\boldsymbol{\nu}, \boldsymbol{f_\alpha}$, and $f_{\boldsymbol{\beta}}$ and yields a convex relative entropy program, which can be solved efficiently. Thus, the chosen parametrization has a significant impact from the perspective of optimization.

However, while this fact is a serious problem for *arbitrary* AM/GM-exponentials, it turns out that this problem is much simpler for *circuit polynomials* and the corresponding *simplicial* AM/GM-exponentials as we show in what follows.

For a simplicial AM/GM-exponential we have that $l = r$ in (5.1.1). Moreover, since the support is a circuit, $\boldsymbol{Q}$ is a full-rank matrix. Therefore, $\boldsymbol{\nu}$ is unique up to a scalar multiple. By the definition of circuit polynomials, Definition 2.4.1, we know that the barycentric coordinates $(\lambda_0, \ldots, \lambda_r)$ of $\boldsymbol{\beta}$ with respect to the vertices $\boldsymbol{\alpha}(0), \ldots, \boldsymbol{\alpha}(r)$ of $\text{New}(f)$ are the unique solution of (5.2.2). It follows that the barycentric coordinates $(\lambda_0, \ldots, \lambda_r)$ are also a solution of (5.1.1). Hence, we obtain for every solution $\boldsymbol{\nu}$ that $\boldsymbol{\nu} = d \cdot (\lambda_0, \ldots, \lambda_r)$ for some $d \in \mathbb{R}^*$. We can now conclude the following theorem.

Theorem 5.2.2. *Let $f(\mathbf{x}) = \sum_{j=0}^{r} f_{\boldsymbol{\alpha}(j)} \mathbf{x}^{\boldsymbol{\alpha}(j)} + f_{\boldsymbol{\beta}} \mathbf{x}^{\boldsymbol{\beta}}$ be a proper circuit polynomial. Then $f(\mathbf{x})$ is nonnegative on $\mathbb{R}^n$ if and only if a particular relative entropy program is feasible, which is jointly convex in $\boldsymbol{\nu}$, the $f_{\boldsymbol{\alpha}(j)}$, $|f_{\boldsymbol{\beta}}|$, and an additional vector $\boldsymbol{\delta} \in \mathbb{R}^{r+1}$.*

Note that the question of whether a given $f(\mathbf{x})$ is a sum of monomial squares is computationally trivial such that these circuit polynomials can safely be excluded.

Proof. By Theorem 2.4.4 we know that the circuit polynomial $f(\mathbf{x})$ is nonnegative if and only if $|f_{\boldsymbol{\beta}}| \leqslant \Theta_f$.

$$
\begin{aligned}
|f_{\boldsymbol{\beta}}| \leqslant \Theta_f \quad &\Leftrightarrow \quad |f_{\boldsymbol{\beta}}| \cdot \prod_{j=0}^{r} \left(\frac{\lambda_j}{f_{\boldsymbol{\alpha}(j)}} \right)^{\lambda_j} \leqslant 1 \Leftrightarrow \prod_{j=0}^{r} \left(\frac{|f_{\boldsymbol{\beta}}| \cdot \lambda_j}{f_{\boldsymbol{\alpha}(j)}} \right)^{\lambda_j} \leqslant 1 \\
&\Leftrightarrow \quad \prod_{j=0}^{r} \left(\frac{|f_{\boldsymbol{\beta}}| \cdot \lambda_j}{f_{\boldsymbol{\alpha}(j)}} \right)^{|f_{\boldsymbol{\beta}}| \cdot \lambda_j} \leqslant 1^{|f_{\boldsymbol{\beta}}|} = 1 \\
&\Leftrightarrow \quad \sum_{j=0}^{r} |f_{\boldsymbol{\beta}}| \cdot \lambda_j \cdot \log \left(\frac{|f_{\boldsymbol{\beta}}| \cdot \lambda_j}{f_{\boldsymbol{\alpha}(j)}} \right) \leqslant 0 \\
&\Leftrightarrow \quad \begin{cases} \text{minimize} & 1 \\ & (1) \quad \nu_j = |f_{\boldsymbol{\beta}}| \cdot \lambda_j \quad \text{for all} \quad j = 0, \ldots, r, \\ \text{subject to:} & (2) \quad \nu_j \cdot \log \left(\frac{\nu_j}{f_{\boldsymbol{\alpha}(j)}} \right) \leqslant \delta_j \quad \text{for all} \quad j = 0, \ldots, r, \\ & (3) \quad \sum_{j=0}^{r} \delta_j \leqslant 0. \end{cases}
\end{aligned}
$$

□

Observe that $|f_{\boldsymbol{\beta}}|$ is redundant in the REP given in the proof of Theorem 5.2.2 since one can leave out the constraint (1) e.g., for $j = 0$ and replace $|f_{\boldsymbol{\beta}}|$ by ν_0/λ_0.

There exists another important difference between SAGE and SONC next to the characterization of nonnegativity on $\mathbb{R}^n_{>0}$ (SAGE) and nonnegativity on $\mathbb{R}^n$ (SONC). In the SONC cone we decompose a polynomial f in a sum of nonnegative circuit polynomials f_i with simplex Newton polytopes. However, in SAGE we decompose a polynomial f in a sum of nonnegative AM/GM-polynomials f_i such that the Newton polytopes of the f_i are not simplices in general and the supports of the f_i have several points in the interior of $\mathrm{New}(f_i)$ in general. If a polynomial f can be decomposed in SAGE, then this certifies nonnegativity of f on $\mathbb{R}^n_{>0}$, but not globally on $\mathbb{R}^n$. However, as we showed, circuit polynomials are special since they are nonnegative on $\mathbb{R}^n$ if and only if they are nonnegative on $\mathbb{R}^n_{>0}$.

In the following example, which was discussed by Chandrasekaran and Shah, we demonstrate how our explicit characterization of circuit polynomials yields an explicit convex, semialgebraic description for special nonnegativity sets compared to SDP methods.

Example 5.2.3 ([CS16], page 1167)**.** Let

$$S_d \;=\; \{(a,b) \in \mathbb{R}^2 : x^{2d} + ax^2 + b \geqslant 0\}.$$

The set S_d is a convex, semialgebraic set for each $d \in \mathbb{N}^*$. Since a univariate polynomial is nonnegative if and only if it is a sum of squares, S_d is also SDP representable, i.e., a projection of a slice of the cone of quadratic, positive semidefinite matrices of some size $w_d \in \mathbb{N}^*$. As noted in [CS16], the algebraic degree of the boundary of S_d grows with d, and hence the size w_d of the smallest SDP description of S_d must also grow with d. In [CS16], the authors use the corresponding relative entropy description (5.1.1) of S_d (note that here nonnegativity on $\mathbb{R}$ is the same as nonnegativity on $\mathbb{R}_{>0}$):

$$S_d \;=\; \{(a,b) \in \mathbb{R} \times \mathbb{R}_{\geqslant 0} : \exists\, \boldsymbol{\nu} \in \mathbb{R}^2_{\geqslant 0} \text{ such that } D(\boldsymbol{\nu}, e \cdot (1,b)^T) \leqslant a, (d-1)\nu_1 = \nu_2\}.$$

A major advantage of this description compared to the SDP method is that the size of S_d does not grow with d. However, we can do even better and use circuit polynomials

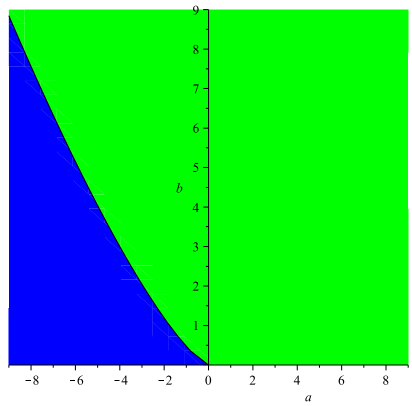

Figure 5.1: The set S_4 is shown in the green area.

and our Theorem 2.4.4 to describe the convex, semialgebraic set S_d directly:

$$S_d = \left\{ (a, b) \in \mathbb{R} \times \mathbb{R}_{\geqslant 0} : a + (d)^{\frac{1}{d}} \cdot \left(\frac{d \cdot b}{d - 1} \right)^{\frac{d-1}{d}} \geqslant 0 \right\}.$$

For $d = 4$ the set S_4 is given as the green area in Figure 5.1. ○

5.3 The Positivstellensatz using SONC Polynomials

In this section we formulate and prove the Positivstellensatz for sums of nonnegative circuit polynomials; see Theorem 5.3.5.

First, we give some basic definitions and recall a representation theorem from real algebraic geometry, which we consult to prove the Positivstellensatz for SONC polynomials. We use Marshall's book [Mar08] as a general source, making some very minor adjustments.

Definition 5.3.1. A *preprime* P is a subset of $\mathbb{R}[\mathbf{x}]$ that contains $\mathbb{R}_{\geqslant 0}$, and that is closed under addition and multiplication. A preprime P is called *Archimedean* if for every $f \in \mathbb{R}[\mathbf{x}]$ there exists an integer $N \geqslant 1$ such that $N - f \in P$. ○

Let P be a preprime. We define the corresponding *ring of P-bounded elements* of $\mathbb{R}[\mathbf{x}]$ as follows:

$$H_P = \{f \in \mathbb{R}[\mathbf{x}] : \text{ there exists an integer } N \geqslant 1 \text{ such that } N \pm f \in P\}.$$

The set H_P is an indicator of how close a given preprime P is to being Archimedean.

In particular, a preprime P is Archimedean if and only if $H_P = \mathbb{R}[\mathbf{x}]$.

Note that H_P is actually a ring [Mar08, Proposition 5.1.3 (1), page 72], which immediately implies the following lemma; see, e.g., [Sch09].

Lemma 5.3.2. *Let $P \subseteq \mathbb{R}[\mathbf{x}]$ be a preprime. Then the following are equivalent:*

(1) *P is Archimedean.*

(2) *There exists an integer $N \geqslant 1$ such that $N \pm x_i \in P$ for all $i = 1, \ldots, n$.*

For convenience of the reader, we give a proof here.

Proof. Implication (1) ⇒ (2) is clear. Let $f, g \in \mathbb{R}[\mathbf{x}]$ with

$$N \pm f \in P \text{ and } M \pm g \in P$$

for some $N, M \in \mathbb{N}^*$, so f and g are P-bounded elements. Since P is closed under addition and multiplication we have

$$(N \pm f) + (M \pm g) \;=\; (N + M) \pm (f + g) \in P,$$

and

$$\frac{1}{2}\left((N \pm f) \cdot (M - g) + (N \pm f) \cdot (M + g)\right) \;=\; N \cdot M \pm f \cdot g \in P.$$

This means, products and sums of P-bounded elements are P-bounded; in fact H_P is a subring of $\mathbb{R}[\mathbf{x}]$. By assumption (2) the variables x_i are P-bounded elements and therefore every polynomial expression in the variables x_i is also P-bounded. Thus, P is Archimedean. □

Given $f_1, \ldots, f_s \in \mathbb{R}[\mathbf{x}]$, we denote by $\text{Prep}(f_1, \ldots, f_s)$ the preprime generated by the $f_1, \ldots, f_s$, i.e., the set of finite sums of elements in $\mathbb{R}[\mathbf{x}]$ of the form $a_{\boldsymbol{i}} f_1^{i_1} \cdots f_s^{i_s}$, where $\boldsymbol{i} = (i_1, \ldots, i_s) \in \mathbb{N}^s$ and $a_{\boldsymbol{i}} \in \mathbb{R}_{\geqslant 0}$:

$$\text{Prep}(f_1, \ldots, f_s) \;=\; \left\{ \sum_{\text{finite}} a_{\boldsymbol{i}} f_1^{i_1} \cdots f_s^{i_s} \;:\; \boldsymbol{i} \in \mathbb{N}^s, a_{\boldsymbol{i}} \in \mathbb{R}_{\geqslant 0} \right\}.$$

The final algebraic structure, which we need to formulate the statements in this section, is a *module* over a preprime:

Definition 5.3.3. Let $P \subseteq \mathbb{R}[\mathbf{x}]$ be a preprime. Then $M \subseteq \mathbb{R}[\mathbf{x}]$ is a *P-module* if it is closed under addition, if it is closed under multiplication by an element of P, and if it contains 1. Analogous to preprimes, a P-module M is *Archimedean* if for each $f \in \mathbb{R}[\mathbf{x}]$ there exists an integer $N \geqslant 1$ such that $N - f \in M$. ○

Note that $1 \in M$ for a P-module M implies that $P \subseteq M$. Obviously, P itself is a P-module.

Now, we state the theorem, which provides the foundation for the proof of the Positivstellensatz including SONC polynomials. There exist various different variations of this statement. For example, one prominent special case is by Krivine [Kri64a, Kri64b]. In fact, the Positivstellensatz for SONC polynomials is basically an implication of this special case, see Remark 5.3.6. Since his statement and proof is rather abstract and for the sake of clearness, we provide a proof of the SONC Positivstellensatz based on the classical subsequent representation theorem.

We follow Marshall's book where the reader can find an overview of the different versions; see [Mar08, page 79].

Theorem 5.3.4 ([Mar08], Theorem 5.4.4)**.** *Let $P \subseteq \mathbb{R}[\mathbf{x}]$ be an Archimedean preprime and let M be an Archimedean P-module. Let $\mathcal{K}_M$ denote the semialgebraic set of points in $\mathbb{R}^n$ on which every element of M is nonnegative:*

$$\mathcal{K}_M \;=\; \{\mathbf{x} \in \mathbb{R}^n : g(\mathbf{x}) \geqslant 0 \text{ for all } g \in M\}.$$

Let $f \in \mathbb{R}[\mathbf{x}]$. If $f(\mathbf{x}) > 0$ for all $\mathbf{x} \in \mathcal{K}_M$, then $f \in M$.

Observe that if a preprime P is Archimedean, then every P-module M is also Archimedean since $P \subseteq M$.

We consider the constrained polynomial optimization problem $f_K^* = \inf_{\mathbf{x} \in K} f(\mathbf{x})$, where $K = \{\mathbf{x} \in \mathbb{R}^n : g_i(\mathbf{x}) \geqslant 0,\ i = 1, \ldots, s\}$.

In what follows we have to assume that K is compact. Namely, in order to use Theorem 5.3.4, we need the involved preprime to be Archimedean. We ensure this by enlarging the definition of K by the $2n$ many redundant constraints $N \pm x_i \geqslant 0$ with $N \in \mathbb{N}$ sufficiently large. We denote these constraints by $l_j(\mathbf{x})$ for $j = 1, \ldots, 2n$. Geometrically speaking, we know that if K is a compact set, then it is contained in some cube $[-N, N]^n$. Hence, if we know the edge length N of such a cube, then we can add the redundant cube constraints l_j to the description of K. We obtain:

$$(5.3.1) \quad K = \{\mathbf{x} \in \mathbb{R}^n : g_i(\mathbf{x}) \geqslant 0 \text{ for } i = 1, \ldots, s \text{ and } l_j(\mathbf{x}) \geqslant 0 \text{ for } j = 1, \ldots, 2n\}.$$

Furthermore, we consider for the given compact K the set of polynomials defined as products of the enlarged set of constraints

$$(5.3.2) \qquad R_q(K) = \left\{ \prod_{k=1}^{q} h_k : h_k \in \{1, g_1, \ldots, g_s, l_1, \ldots, l_{2n}\} \right\}.$$

Moreover, we define $\rho_q = |R_q(K)|$ and $\tau_q = \max_{i=1,\ldots,s} \{\deg(g_i), 1\} \cdot q$.

Now we state the Positivstellensatz for sums of nonnegative circuit polynomials.

Theorem 5.3.5 (Positivstellensatz for SONC). *Let $f, g_1, \ldots, g_s \in \mathbb{R}[\mathbf{x}]$, let K be a compact, basic closed semialgebraic set as in* (5.3.1)*, and let $R_q(K)$ be defined as in* (5.3.2)*. If $f(\mathbf{x})$ is strictly positive for all $\mathbf{x} \in K$, then there exist $d, q \in \mathbb{N}^*$, SONC polynomials $s_j(\mathbf{x}) \in C_{n,2d}$, and polynomials $H_j(\mathbf{x}) \in R_q(K)$ indexed by $j = 1, \ldots, \rho_q$ such that*

$$f(\mathbf{x}) = \sum_{j=1}^{\rho_q} s_j(\mathbf{x}) H_j(\mathbf{x}).$$

Note that the sum $\sum_{j=1}^{\rho_q} s_j(\mathbf{x}) H_j(\mathbf{x})$ is of degree at most $2d + \tau_q$, and that it contains a summand $s_0 \cdot 1 \in C_{n,2d}$, which is analogous to the structure of various SOS-based Positivstellensätze.

Proof. Let $f, g_1, \ldots, g_s \in \mathbb{R}[\mathbf{x}]$ and $P \subseteq \mathbb{R}[\mathbf{x}]$ be the preprime generated by all polynomials $g_1, \ldots, g_s$ and the redundant linear constraints $l_1, \ldots, l_{2n}$, which we were allowed to add since K is compact, i.e.,

$$P = \operatorname{Prep}(g_1, \ldots, g_s, l_1, \ldots, l_{2n}).$$

The preprime P is Archimedean since it contains the cube inequalities; see Lemma 5.3.2. In what follows we consider the set

$$(5.3.3) \quad M = \left\{ \sum_{\text{finite}} s(\mathbf{x}) H(\mathbf{x}) : \exists\, d, q \in \mathbb{N}^* \text{ such that } s(\mathbf{x}) \in C_{n,2d}, H(\mathbf{x}) \in R_q(K) \right\}.$$

Claim 1: M is an Archimedean P-module.

This follows immediately from (5.3.3), Definition 5.3.3, and the fact that P is Archimedean.

Claim 2: The nonnegativity set $\mathcal{K}_M = \{\mathbf{x} \in \mathbb{R}^n : g(\mathbf{x}) \geqslant 0 \text{ for all } g \in M\}$ equals K.

On the one hand, we have that $\mathcal{K}_M \subseteq K$ since M is a P-module. Thus, the polynomials defining K are contained in M. On the other hand, a polynomial in M has the form $\sum_{\text{finite}} s(\mathbf{x})H(\mathbf{x})$, such that every $s(\mathbf{x}) \in C_{n,2d}$. So, every $s(\mathbf{x})$ is nonnegative on $\mathbb{R}^n$. Thereby, the nonnegativity of polynomials in M depends only on the polynomials $H(\mathbf{x}) \in R_q(K)$. But these polynomials are exactly products of the constraint polynomials in K. Thus, we can conclude that $K \subseteq \mathcal{K}_M$ and hence $K = \mathcal{K}_M$.

With Claims 1 and 2 satisfied, we can apply Theorem 5.3.4 to conclude that $f \in M$. By (5.3.3) the expression of the Positivstellensatz is of the desired form. □

For a fixed q, the number of elements in the set $R_q(K)$ is at most $\binom{s+2n+q}{q}$; thus, its cardinality is exponential in q. One may ask whether it is possible to formulate a Positivstellensatz involving only a linear number of terms, like Putinar's Positivstellensatz, Theorem 2.3.11, based on sums of squares decompositions for polynomial optimization problems. It would be desirable to define an object like a quadratic module of the constraint polynomials. The main difficulty in carrying out such a construction is that the product of two SONC polynomials is not a SONC polynomial in general, in contrast to the product of two SOS, which is an SOS; see Lemma 3.4.1, and also Chapter 6.

Remark 5.3.6. Schweighofer pointed out that since $P = \text{Prep}(g_1, \ldots, g_s, l_1, \ldots, l_{2n})$ is an Archimedean preprime we can apply Theorem 5.3.4 with choosing $M = P$, which corresponds to Krivine's special case. Hence, if $f(\mathbf{x}) > 0$ for all $\mathbf{x} \in \mathcal{K}_P$, then f is in the preprime P. Therefore f can be represented as

$$f(\mathbf{x}) \;=\; \sum_{j=1}^{\rho_q} a_j H_j(\mathbf{x}),$$

where $a_j \in \mathbb{R}_{\geqslant 0}$, and $H_j(\mathbf{x}), \rho_q$ as given in Theorem 5.3.5, see also [Sch02] for this representation. Since SONC polynomials are nonnegative, we can replace the a_j with SONC polynomials and the representation of f in Theorem 5.3.5 follows immediately.

5.4 Application of the SONC Positivstellensatz to Constrained Polynomial Optimization Problems

In this section we establish a hierarchy of lower bounds $f_{\text{sonc}}^{(d,q)}$ given by the SONC Positivstellensatz, Theorem 5.3.5, for the solution f_K^* of a (CPOP) on a compact, semialgebraic set, and we formulate an optimization problem to compute these bounds. As main results we show first that the bounds $f_{\text{sonc}}^{(d,q)}$ converge to f_K^* for $d, q \to \infty$, Theorem 5.4.2, and second we show that the corresponding optimization problem is a *relative entropy program* and hence efficiently solvable with interior point methods, Theorem 5.4.3. We also discuss an example in Section 5.4.3.

5.4.1 A Converging Hierarchy for Constrained Polynomial Optimization

Recall that minimizing a polynomial $f(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]$ on a semialgebraic set $K \subseteq \mathbb{R}^n$ is equivalent to maximizing a lower bound of this polynomial. Thus, we have:

$$f_K^* \;=\; \inf_{\mathbf{x} \in K} f(\mathbf{x}) \;=\; \sup\{\gamma \in \mathbb{R} \;:\; f(\mathbf{x}) - \gamma \geqslant 0 \text{ for all } \mathbf{x} \in K\}.$$

To obtain a general lower bound for f_K^*, which is efficiently computable, we relax the nonnegativity condition to find the real number:

$$f_{\text{sonc}}^{(d,q)} \;=\; \sup\left\{\gamma \in \mathbb{R} \;:\; f(\mathbf{x}) - \gamma \;=\; \sum_{j=1}^{\rho_q} s_j(\mathbf{x}) H_j(\mathbf{x})\right\},$$

where $s_j(\mathbf{x}) \in C_{n,2d}$ are SONC polynomials and $H_j(\mathbf{x}) \in R_q(K)$ with $R_q(K)$ being defined as in (5.3.2). Indeed, the number $f_{\text{sonc}}^{(d,q)}$ is a lower bound for f_K^* and grows monotonically in d and q as the following lemma shows.

Lemma 5.4.1. *Let $f, g_1, \ldots, g_s \in \mathbb{R}[\mathbf{x}]$, and let K be a semialgebraic set. Then we have the following:*

(1) $f_{\text{sonc}}^{(d,q)} \leqslant f_K^*$ *for all* $d, q \in \mathbb{N}^*$.

(2) $f_{\text{sonc}}^{(d,q)} \leqslant f_{\text{sonc}}^{(\tilde{d},\tilde{q})}$ *for all* $d \leqslant \tilde{d}, q \leqslant \tilde{q}$ *with* $d, \tilde{d}, q, \tilde{q} \in \mathbb{N}^*$.

Lemma 5.4.1 yields a sequence $\left\{f_{\text{sonc}}^{(d,q)}\right\}_{d,q \in \mathbb{N}^*}$ of lower bounds of f_K^* which is increasing both in d and q.

Proof.

(1) For every $s_j(\mathbf{x}) \in C_{n,2d}$ and every $H_j(\mathbf{x}) \in R_q(K)$ the polynomial $s_j(\mathbf{x})H_j(\mathbf{x})$ is nonnegative on K. Thus, the sum $\sum_{j=1}^{\rho_q} s_j(\mathbf{x})H_j(\mathbf{x})$ is nonnegative on K and we have for every $\gamma \in \mathbb{R}$ and every $\mathbf{x} \in K$ that

$$f(\mathbf{x}) - \gamma \;=\; \sum_{j=1}^{\rho_q} s_j(\mathbf{x})H_j(\mathbf{x}) \;\Rightarrow\; f(\mathbf{x}) - \gamma \;\geqslant\; 0.$$

Hence, we have $f_{\text{sonc}}^{(d,q)} \leqslant f_K^*$ for every $d, q \in \mathbb{N}^*$.

(2) We have $C_{n,2d} \subseteq C_{n,2\tilde{d}}$ and $R_q(K) \subseteq R_{\tilde{q}}(K)$ for all $d \leqslant \tilde{d}, q \leqslant \tilde{q}$ with $d, \tilde{d}, q, \tilde{q} \in \mathbb{N}^*$. Therefore, the hierarchy of the bounds follows. □

Observe that Lemma 5.4.1 does not require K to be compact. An analogous statement and proof can be given literally without involving the redundant cube constraints $l_1, \ldots, l_{2n}$ in the definition of $R_q(K)$.

For a *compact* constraint set K, however, we have an asymptotic convergence to the optimum f_K^* of the sequence $\left\{ f_{\text{sonc}}^{(d,q)} \right\}_{d,q \in \mathbb{N}^*}$. Thus, for compact K the provided hierarchy is *complete.*

Theorem 5.4.2. *Let everything be defined as in Lemma* 5.4.1. *In addition, let K be compact. Then*

$$f_{\text{sonc}}^{(d,q)} \uparrow f_K^* \quad \textit{for } d, q \to \infty.$$

Note that q is bounded from above by the chosen d. Therefore, it is sufficient to investigate $d \to \infty$ and choose for every d the corresponding maximal q.

Proof. Let $\varepsilon > 0$ be arbitrary. Then $f(\mathbf{x}) - (f_K^* - \varepsilon)$ is strictly positive on K for all $\mathbf{x} \in \mathbb{R}^n$. According to Theorem 5.3.5, there exist sufficiently large $d, q \in \mathbb{N}^*$ such that $f(\mathbf{x}) - f_K^* + \varepsilon = \sum_{j=1}^{\rho_q} s_j(\mathbf{x})H_j(\mathbf{x})$. Thus,

$$f_K^* - \varepsilon \;\leqslant\; f_{\text{sonc}}^{(d,q)}, \tag{5.4.1}$$

by definition of $f_{\text{sonc}}^{(d,q)}$. Since $d, q \to \infty$, (5.4.1) holds for all $\varepsilon \downarrow 0$ for sufficiently large d, q. By Lemma 5.4.1 (2) the values $f_{\text{sonc}}^{(d,q)}$ are monotonically increasing in d, q, and the result follows. □

5.4.2 Computation of the new Hierarchy via Relative Entropy Programming

Let $n, 2d, q$ be fixed. We intend to compute $f_{\text{sonc}}^{(d,q)}$ via a suitable optimization program. This means for $f \in \mathbb{R}[\mathbf{x}]$ and a compact set K we are looking for the maximal $\gamma \in \mathbb{R}$ such that

$$f(\mathbf{x}) - \gamma \;=\; \sum_{\text{finite}} H_\ell(\mathbf{x}) s_\ell(\mathbf{x}), \tag{5.4.2}$$

where $H_\ell(\mathbf{x}) \in R_q(K)$ and $s_\ell(\mathbf{x}) \in C_{n,2d}$. We formulate such a program in (5.4.3) and show in Theorem 5.4.3 that this program is a relative entropy program and hence efficiently solvable.

In what follows it is sufficient to consider nonnegative circuit polynomials instead of general SONC polynomials. Namely, since every $s_\ell(\mathbf{x}) \in C_{n,2d}$ in (5.4.2) is of the form $\sum_{\text{finite}} p_{i,\ell}(\mathbf{x})$ where every $p_{i,\ell}(\mathbf{x})$ is a nonnegative circuit polynomial, we can split every term $H_\ell(\mathbf{x}) s_\ell(\mathbf{x})$ into $\sum_{\text{finite}} H_\ell(\mathbf{x}) p_{i,\ell}(\mathbf{x})$ by distribution law.

Recall from Sections 2.1 and 2.4 that Circ_A denotes the set of all circuit polynomials with support $A \subset \mathbb{Z}^n$, that $\Delta_{n,2d}$ describes the standard simplex in n variables of edge length $2d$, and that we define $\mathcal{L}_{n,2d} = \Delta_{n,2d} \cap \mathbb{Z}^n$. The support of every circuit polynomial is contained in a sufficiently large scaled standard simplex $\Delta_{n,2d}$. We define

$$\text{Circ}_{n,2d} \;=\; \{p \in \text{Circ}_A \;:\; A \subseteq \mathcal{L}_{n,2d}\},$$

that is the set of all circuit polynomials with a support A which is contained in $\Delta_{n,2d}$.

Let $f(\mathbf{x}) = f_{\mathbf{0}} + \sum_{\boldsymbol{\eta} \in \mathcal{L}_{n,2d+\tau_q} \setminus \{\mathbf{0}\}} f_{\boldsymbol{\eta}} \mathbf{x}^{\boldsymbol{\eta}} \in \mathbb{R}[\mathbf{x}]$. Note that we allow $f_{\boldsymbol{\eta}} = 0$. Furthermore, let K be a compact, semialgebraic set given by a list of constraints $g_1, \ldots, g_s$. Here, we simplify the notation by assuming that the g_i's already contain the linear constraints $l_1, \ldots, l_{2n}$, which we added in Section 5.3. Let

$$\text{Circ}_{n,2d} \;=\; \text{Circ}_{A(1)} \sqcup \cdots \sqcup \text{Circ}_{A(t)},$$

where $A(1), \ldots, A(t) \subseteq \mathcal{L}_{n,2d}$ is the finite list of possible support sets of circuit polynomials in $\Delta_{n,2d}$. We use the notation

$$\text{Circ}_{A(i)} \;=\; \left\{ \sum_{j=0}^{r_i} c_{\boldsymbol{\alpha}(j,i)} \mathbf{x}^{\boldsymbol{\alpha}(j,i)} + \varepsilon \cdot c_{\boldsymbol{\beta}(i)} \mathbf{x}^{\boldsymbol{\beta}(i)} \;:\; \begin{array}{c} c_{\boldsymbol{\alpha}(j,i)},\; c_{\boldsymbol{\beta}(i)} \in \mathbb{R}_{\geqslant 0}, \\ \text{and } \varepsilon \in \{1, -1\} \end{array} \right\}.$$

We denote by $\lambda_{0,i}, \ldots, \lambda_{r_i,i}$ the barycentric coordinates satisfying $\sum_{j=0}^{r_i} \lambda_{j,i}\boldsymbol{\alpha}(j,i) = \boldsymbol{\beta}(i)$. Let $R_q(K) = \{H_1, \ldots, H_{\rho_q}\}$ such that $H_\ell(\mathbf{x}) = \sum_{j=1}^{k_\ell} H_{\boldsymbol{\gamma}(j,\ell)}\mathbf{x}^{\boldsymbol{\gamma}(j,\ell)}$ with $H_{\boldsymbol{\gamma}(j,\ell)} \in \mathbb{R}$. Moreover, we define the following support vectors

$$\begin{aligned} \operatorname{supp}(\mathrm{Circ}_{n,2d}) &= [\boldsymbol{\alpha}(j,i), \boldsymbol{\beta}(i) \ : \ i = 1, \ldots, t, \ j = 0, \ldots, r_i], \\ \operatorname{supp}(R_q(K)) &= [\boldsymbol{\gamma}(j,\ell) \ : \ \ell = 1, \ldots, \rho_q, \ j = 0, \ldots, k_\ell]. \end{aligned}$$

This means that $\operatorname{supp}(\mathrm{Circ}_{n,2d})$ is the vector which contains all exponents contained in $A(1), \ldots, A(t)$ with repetition. Similarly, $\operatorname{supp}(R_q(K))$ is the vector which contains all exponents contained in the supports of $H_1, \ldots, H_{\rho_q}$ with repetition. By construction, we have that $\operatorname{supp}(\mathrm{Circ}_{n,2d})$ is contained in $\mathcal{L}_{n,2d}$, and every entry of $\operatorname{supp}(R_q(K))$ is contained in $\mathcal{L}_{n,\tau_q}$.

By (5.4.2) we have to construct an optimization program which guarantees that for every exponent $\boldsymbol{\eta} \in \mathcal{L}_{n,2d+\tau_q}$, we have that the term $f_{\boldsymbol{\eta}}\mathbf{x}^{\boldsymbol{\eta}}$ of the given polynomial f, which has to be minimized, equals the sums of a term with exponent $\boldsymbol{\eta}$ in $\sum_{\text{finite}} H_\ell s_\ell$ with $H_\ell \in R_q(K)$ and $s_\ell \in C_{n,2d}$. Thus, we have to (1) guarantee that the involved functions are indeed SONC polynomials and (2) add a linear constraint for every $\boldsymbol{\eta} \in \mathcal{L}_{n,2d+\tau_q}$ to match the coefficients of the terms with exponent $\boldsymbol{\eta}$ in f with the coefficients of the terms with exponent $\boldsymbol{\eta}$ in $\sum_{\text{finite}} H_\ell s_\ell$; see (5.4.2).

Let R be the subset of a real space given by

$$R = \left\{ c^{(\ell,\varepsilon)}_{\boldsymbol{\alpha}(j,i)}, c^{(\ell,\varepsilon)}_{\boldsymbol{\beta}(i)}, \nu^{(\ell,\varepsilon)}_{j,i} \in \mathbb{R}_{\geqslant 0}, \delta^{(\ell,\varepsilon)}_{j,i} \in \mathbb{R} \ : \ \begin{array}{c} \text{for every } \ell = 1, \ldots, \rho_q, \ \varepsilon \in \{1,-1\}, \\ \text{and } \boldsymbol{\alpha}(j,i), \boldsymbol{\beta}(i) \in \operatorname{supp}(\mathrm{Circ}_{n,2d}) \end{array} \right\}.$$

Note that we are constructing a relative entropy program. The $\nu^{(\ell,\varepsilon)}_{j,i} \in \mathbb{R}_{\geqslant 0}$, and $\delta^{(\ell,\varepsilon)}_{j,i} \in \mathbb{R}$ in R form the vectors $\boldsymbol{\nu}$ and $\boldsymbol{\delta}$ of variables in the general form of an REP as defined in Definition 5.1.2.

In order to match the coefficients of f with a representing polynomial coming from the SONC Positivstellensatz, we define for every $\boldsymbol{\eta} \in \mathcal{L}_{n,2d+\tau_q} \backslash \{\mathbf{0}\}$ the following linear functions from R to $\mathbb{R}$:

$$\Gamma_1(\boldsymbol{\eta}) = \sum_{\substack{\boldsymbol{\beta}(i)+\boldsymbol{\gamma}(j,\ell)=\boldsymbol{\eta} \\ \boldsymbol{\beta}(i)\in\operatorname{supp}(\mathrm{Circ}_{n,2d}) \\ \boldsymbol{\gamma}(j,\ell)\in\operatorname{supp}(R_q(K)) \\ \varepsilon\in\{1,-1\}}} \varepsilon \cdot c^{(\ell,\varepsilon)}_{\boldsymbol{\beta}(i)} \cdot H_{\boldsymbol{\gamma}(j,\ell)}, \quad \Gamma_2(\boldsymbol{\eta}) = \sum_{\substack{\boldsymbol{\alpha}(j,i)+\boldsymbol{\gamma}(j,\ell)=\boldsymbol{\eta} \\ \boldsymbol{\alpha}(j,i)\in\operatorname{supp}(\mathrm{Circ}_{n,2d}) \\ \boldsymbol{\gamma}(j,\ell)\in\operatorname{supp}(R_q(K)) \\ \varepsilon\in\{1,-1\}}} c^{(\ell,\varepsilon)}_{\boldsymbol{\alpha}(j,i)} \cdot H_{\boldsymbol{\gamma}(j,\ell)},$$

where the $H_{\boldsymbol{\gamma}(j,\ell)}$ are constants given by the coefficients of the functions $H_1, \ldots, H_{\rho_q}$.

We define an optimization program to compute $f_{\text{sonc}}^{(d,q)}$. In what follows, the variables $\nu_{j,i}^{(\ell,\varepsilon)}$ and $\delta_{j,i}^{(\ell,\varepsilon)}$ are completely redundant for the actual optimization process; see (1a), (1b), and (1c). We only have to introduce them to guarantee that the program (5.4.3) has the form of a relative entropy program.

$$
(5.4.3)\quad \begin{cases}
\text{minimize} & \displaystyle\sum_{\substack{\boldsymbol{\alpha}(j,i)+\boldsymbol{\gamma}(\jmath,\ell)=\mathbf{0}\\ \boldsymbol{\alpha}(j,i)\in\operatorname{supp}(\mathrm{Circ}_{n,2d})\\ \boldsymbol{\gamma}(\jmath,\ell)\in\operatorname{supp}(R_q(K))\\ \varepsilon\in\{1,-1\}}} c_{\boldsymbol{\alpha}(j,i)}^{(\ell,\varepsilon)} \cdot H_{\boldsymbol{\gamma}(\jmath,\ell)} \\
& \text{over the subset } R' \text{ of } R \text{ defined by:} \\
(1a) & \nu_{j,i}^{(\ell,\varepsilon)} = c_{\boldsymbol{\beta}(i)}^{(\ell,\varepsilon)} \cdot \lambda_{j,i} \quad \text{for all } \ell = 1,\ldots,\rho_q,\ \varepsilon \in \{1,-1\},\ j = 0,\ldots,r_i,\ i = 1,\ldots,t, \\
(1b) & \nu_{j,i}^{(\ell,\varepsilon)} \cdot \log\left(\frac{\nu_{j,i}^{(\ell,\varepsilon)}}{c_{\boldsymbol{\alpha}(j,i)}^{(\ell,\varepsilon)}}\right) \leqslant \delta_{j,i}^{(\ell,\varepsilon)} \quad \text{for all } \ell = 1,\ldots,\rho_q,\ \varepsilon \in \{1,-1\},\ j = 0,\ldots,r_i,\ i = 1,\ldots,t, \\
(1c) & \sum_{j=0}^{r_i} \delta_{j,i}^{(\ell,\varepsilon)} \leqslant 0 \quad \text{for all } \ell = 1,\ldots,\rho_q,\ \varepsilon \in \{1,-1\},\ i = 1,\ldots,t, \\
(2) & \Gamma_1(\boldsymbol{\eta}) + \Gamma_2(\boldsymbol{\eta}) = f_{\boldsymbol{\eta}} \quad \text{for every } \boldsymbol{\eta} \in \mathcal{L}_{n,2d+\tau_q}\backslash\{\mathbf{0}\}.
\end{cases}
$$

Theorem 5.4.3. *The program* (5.4.3) *is a relative entropy program and hence efficiently solvable, and its output coincides with* $f_{\mathbf{0}} - f_{\text{sonc}}^{(d,q)}$.

Proof. First, we show that (5.4.3) is indeed an REP, i.e., we need to show that it is of the form (5.1.2) in Definition 5.1.2. Constraint (1b) in (5.4.3) is a constraint of the form (2) in (5.1.2) satisfying $\nu_{j,i}^{(\ell,\varepsilon)}, c_{\boldsymbol{\alpha}(j,i)}^{(\ell,\varepsilon)} \geqslant 0$ and $\delta_{j,i}^{(\ell,\varepsilon)} \in \mathbb{R}$ as required. The constraints (1a),(1c), and (2) in (5.4.3) are linear constraints since all $\lambda_{j,i}$, $H_{\boldsymbol{\gamma}(\jmath,\ell)}$, and ε are constants; note that linear equalities can be represented by two linear inequalities. Thus, these constraints are of the form (1) in (5.1.2). Finally, the objective function is also linear as required by (5.1.2). Hence, (5.4.3) is an REP by Definition 5.1.2.

Second, we need to show that the program provides the correct output. Note that the program is infeasible if there exist i, j, ℓ such that $c_{\boldsymbol{\alpha}(j,i)}^{(\ell,\varepsilon)} = 0$ and $c_{\boldsymbol{\beta}(i)}^{(\ell,\varepsilon)} \cdot \lambda_{j,i} > 0$. Hence, we can omit this case. By Theorem 5.2.2 the union of the constraints (1a),(1b), and (1c) is equivalent to a constraint

$$
(3) \quad c_{\boldsymbol{\beta}(i)}^{(\ell,\varepsilon)} \prod_{j=0}^{r_i} \left(\frac{\lambda_{j,i}}{c_{\boldsymbol{\alpha}(j,i)}^{(\ell,\varepsilon)}}\right)^{\lambda_{j,i}} \leqslant 1 \quad \text{for every } i = 1,\ldots,t,\ \varepsilon \in \{1,-1\}.
$$

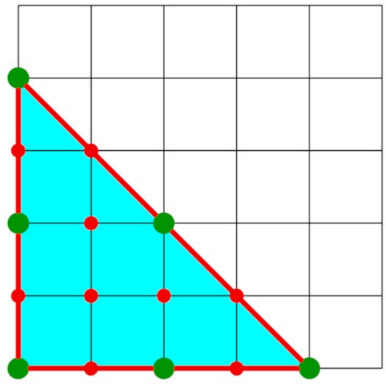

Figure 5.2: $\Delta_{2,4}$ with the lattice points $\mathcal{L}_{2,4}$. The even points are the green ones.

The variables $c_{\boldsymbol{\alpha}(j,i)}^{(\ell,\varepsilon)}$ and $c_{\boldsymbol{\beta}(i)}^{(\ell,\varepsilon)}$ in the program (5.4.3) are by construction the coefficients of circuit polynomials. For the purpose of the program, these circuit polynomials need to be nonnegative; see (5.4.2). This is guaranteed by constraint (3).

For every $\boldsymbol{\eta} \in \mathcal{L}_{n,2d+\tau_q} \backslash \{\mathbf{0}\}$, constraint (2) guarantees that every coefficient $f_{\boldsymbol{\eta}}$ equals $\Gamma_1(\boldsymbol{\eta}) + \Gamma_2(\boldsymbol{\eta})$, which are exactly all polynomials of the form $\sum_{\text{finite}} H_\ell s_\ell$, where $H_\ell \in R_q(K)$ and $s_\ell \in C_{n,2d}$. Particularly, it is sufficient to consider (nonnegative) circuit polynomials in $\Gamma_1(\boldsymbol{\eta})$ and $\Gamma_2(\boldsymbol{\eta})$ instead of SONC polynomials. Namely, for every term $H_\ell s_\ell$ with $s_\ell \in C_{n,2d}$ we can write $s_\ell = \sum_{\text{finite}} p_{i,\ell}$, where $p_{i,\ell}$ are nonnegative circuit polynomials. Thus, on the one hand, we obtain an expression $H_\ell s_\ell = \sum_{\text{finite}} H_\ell p_{i,\ell}$ which depends only on circuit polynomials. On the other hand, we can guarantee that the representation (5.4.2) is satisfied, which we need to show. Finally, the program minimizes the constant term of the function $\sum_{\text{finite}} H_\ell s_\ell$, where $H_\ell \in R_q(K)$, which is equivalent to maximizing γ. □

5.4.3 An Example

We consider the polynomial $f = x_1^3 + x_2^3 - x_1 x_2 + 4$ and a semialgebraic set K given by constraints $g_1 = -x_1 + 1$, $g_2 = x_1 + 1$, $g_3 = -x_2 + 1$, and $g_4 = x_2 + 1$. It is easy to see that f is positive on K. We want to represent f with the SONC Positivstellensatz, Theorem 5.3.5. We consider $C_{2,4}$; see Figure 5.2 for $\Delta_{2,4}$ and the lattice points $\mathcal{L}_{2,4}$.

$\text{Circ}_{2,4}$ is a union of 28 different support sets. There exist:

- six even lattice points in $\mathcal{L}_{2,4}$ and thus six zero-dimensional circuit polynomials,
- $\binom{6}{2} = 15$ circuit polynomials with one-dimensional Newton polytope, and

- $\binom{6}{3}$ even 2-simplices, which are contained in $\Delta_{2,4}$. One simplex contains three lattice points in the interior, four contain one lattice point in the interior, and the remaining ones contain no lattice point in the interior. Thus, we need only to consider seven circuit polynomials with two-dimensional Newton polytope.

The number of elements in $R_q(K)$ is $\rho_q = \binom{4+q}{q}$; see Section 5.3. That is, we have in this example $\rho_1 = 5$, $\rho_2 = 15$, $\rho_3 = 35$.

Let us assume that we want to compute $f_{\text{sonc}}^{(2,1)}$. We are looking for the maximal γ such that $f(\mathbf{x}) - \gamma$ can be represented as a sum $s_j(\mathbf{x})H_j(\mathbf{x})$ with $s_j(\mathbf{x}) \in C_{2,4}$ and $H_j(\mathbf{x}) \in R_1(K)$. We would not, however, consider all these polynomials in practice. First, the circuit polynomials with one-dimensional Newton polytope are sufficient to construct every lattice point in $\mathcal{L}_{2,4}$ and thus it makes sense to disregard all 2-simplices. Second, f does not contain every lattice point in $\mathcal{L}_{2,4}$ as an exponent, and hence it is not surprising that several further circuit polynomials can be omitted. Indeed, we find a decomposition according to the Positivstellensatz, Theorem 5.3.5, of the form

$$f(\mathbf{x}) = (x_1+1)\cdot(x_1^2 - 2x_1 + 1) + (x_2+1)\cdot(x_2^2 - 2x_2 + 1) + 1\cdot\left(\frac{1}{2}x_1^2 - x_1x_2 + \frac{1}{2}x_2^2\right) + 1\cdot\left(\frac{1}{2}x_1^2 + x_1 + \frac{1}{2}\right) + 1\cdot\left(\frac{1}{2}x_2^2 + x_2 + \frac{1}{2}\right) + 1,$$

which involves only 3 of the 15 one-dimensional circuit polynomials, one zero-dimensional circuit polynomial, and no two-dimensional one.

5.5 Conclusion

In this chapter we have shown that the SONC cone yields a new way to attack (CPOPs), independent of the SOS/SDP approach. Namely, we provided a converging hierarchy of lower bounds, which can be computed efficiently via relative entropy programming. Hence, the new results establish SONC polynomials as a promising alternative or rather extension to SOS certificates. SONC certificates are an alternative for SOS certificates particularly for those problems, where the SOS approach has its difficulties. The new results of the hierarchical approach lead to many future tasks and open problems, which will be addressed in aggregated form in the subsequent Chapter 6.

Moreover, we compared in this chapter the related concepts of SAGE and SONC for the first time yielding some interesting observations regarding their relation.

Chapter 6

Final Remarks and Open Problems

We conclude this thesis with final remarks, open problems, and some future tasks arising from the observations of our work.

In this thesis we studied sums of nonnegative circuit polynomials and their related cone $C_{n,2d}$. We investigated our key objects geometrically as well as in application to polynomial optimization problems. The results and observations of this thesis provide important new advancements in the area of both pure and applied real and convex algebraic geometry.

Geometric Analysis of the SONC Cone. The first part of this work discussed SONC polynomials and the SONC cone from the theoretical point of view. We observed general properties of the SONC cone and provided an explicit and complete characterization of the number of zeros of SONCs. Based upon these observations, we provided a first approach to the exposed faces of the SONC cone, which have to be analyzed in more detail. Finally, we showed that $C_{n,2d}$ is full-dimensional in $P_{n,2d}$.

However, there are many more open problems regarding the SONC cone itself which need to be addressed, e.g., convex geometric structures of $C_{n,2d}$ such as its boundary and its extreme rays as well as its dual cone. The understanding of these structures is crucial for the study of the relation between $C_{n,2d}$ and $P_{n,2d}$ and the knowledge of its dual cone is highly desirable for the application of SONC certificates.

Moreover, the connection of the SONC and the SAGE cone need to be analyzed more carefully. An interesting question is whether there is some kind of primal/dual relation between these cones.

Another important problem concerns the relation between the cones $C_{n,2d}$, $\Sigma_{n,2d}$, and

$P_{n,2d}$. First, it would be interesting to discuss the set theoretic difference $P_{n,2d} \backslash C_{n,2d}$, i.e., to investigate the quantitative relationship between the SONC and the nonnegativity cone. Note that in this context also the quantitative relationship between $\Sigma_{n,2d}$ and $P_{n,2d}$ is not completely understood. For instance, the exact quantitative relationship between these cones in small dimension is still an open problem, see Section 2.2.3. Second, the relation between the SOS and the SONC cone need to be explored in more detail. Third and maybe the most important task is to study the *convex hull* of the SONC and the SOS cone. Many counterexamples for polynomials being nonnegative but not SOS are in fact SONC polynomials, see for example the famous Motzkin polynomial. Therefore, it would be interesting to know the approximation quality of $\mathrm{conv}(C_{n,2d}, \Sigma_{n,2d})$ in $P_{n,2d}$. This analysis would also have an immense impact in applications, among others to polynomial optimization problems.

SONC Polynomials in Application to Optimization. The second part of this thesis focused on applying SONC polynomials to polynomial optimization problems, mostly in the constrained case. Initially, we derived a single lower bound for the optimal value f_K^* of (CPOP) for the class of ST-polynomials based on the SONC/GP approach. This approach has the significant advantage over the SOS/SDP-based approach that the runtime is much shorter and not sensitive to increasing the degree. Moreover, we extended the SONC/GP approach both in the unconstrained and the constrained case to non-ST-polynomials. This general case needs to be studied in more detail. Then, we established a hierarchy of lower bounds converging to f_K^* of a (CPOP) on a compact set K which is efficiently computable by an REP.

Particularly resulting from the hierarchical approach for (CPOP) there are many obvious tasks and questions, whose answers would be very useful for practical applications.

First, it is important to implement the program (5.4.3), test it for various instances of constrained polynomial optimization problems, and compare the runtime and optimal values with the counterparts from SDP results using Lasserre's relaxation. Given the runtime comparison of the SONC and the SOS approach in [DIdW18, GM12, IdW16a] using geometric programming, there is reasonable hope that our relative entropy programs are faster than semidefinite programming in several cases.

Second, we have seen in Section 5.4.3 that it can (and likely will often) happen that many of the circuit polynomials in $\mathrm{supp}(\mathrm{Circ}_{n,2d})$ are redundant for finding a representation of a given polynomial with respect to the SONC Positivstellensatz, Theorem 5.3.5.

Hence, the corresponding optimization problem (5.4.3) can be reduced in these cases. For practical applications, we have to develop strategies to restrict ourselves to useful subsets of circuit polynomials to reduce the runtime of our programs via reducing the number of variables.

Third, an important problem is to study the convergence given by our provided hierarchy in more detail. Since the optimal value f_K^* is typically unknown, two practical questions are raised immediately:

(a) How do we check *exactness* of our relaxations?

(b) How do we certify *finite convergence* of our hierarchy?

Unfortunately, there is no obvious way to attack this problem, since similar statements for the Lasserre relaxation (see, e.g., [Las10, Lau09]) cannot be proved with analogous methods for the SONC Positivstellensatz straightforwardly. Namely, the existing theory for Lasserre's relaxation is based on the dual optimization problem whose description uses localizing matrices and the theory of moments, see Section 2.3. Therefore, once more it would be interesting to study the dual perspective of the SONC theory.

Fourth, in the SONC Positivstellensatz, Theorem 5.3.5, the representation of f is exponential in the number of polynomial constraints. Hence, a delicate open problem is to analyze whether there always exists a decomposition which is linear in the constraints, i.e., a representation with $q = 1$, which corresponds to a Putinar-equivalent Positivstellensatz for SONC polynomials. If such a representation does not exist in general, then it would also be interesting to investigate modified or stronger assumptions or to search for certain instances of polynomials for which there exist such a minimal representation. Note that proofs of the analogous statement for SOS polynomials, Putinar's Positivstellensatz, are based on classical ideas from real algebraic geometry and often make use of some properties of quadratic modules. In contrast to SOS polynomials the set of SONC polynomials does not form a quadratic module, see Section 3.4. Thus, again we have to use other techniques to address this question.

Beyond the questions arising from the hierarchical approach for (CPOPs) it would be very interesting to investigate the application of SONC certificates to optimization over the (constrained) hypercube. A special case of (CPOPs), which is particularly relevant to applications in combinatorial optimization, concerns *optimization over the Boolean hypercube* $\mathcal{H}$, mostly $\mathcal{H} = \{0,1\}^n$ or $\mathcal{H} = \{\pm 1\}^n$. Hence, one often speaks of 0/1-optimization or binary (CPOPs) in this case. For instance, one can use those binary (CPOPs) to attack the maximum cut problem [GW95], the maximum stable set problems [Her95], and maximum weighted independent set problems [HLZ13]. Those problems are well studied using the SOS/SDP-based approach, see also Section 2.2.1. Therefore, it would be interesting to study Boolean (CPOPs) by means of SONC polynomials and their (computational) complexity.

Finally, we hope to find a way to combine SOS and SONC certificates in theory and in practice. We already outlined the theoretical motivation for this endeavor, applied to optimization problems such a combined certificate is expected to be extremely powerful. Since the SONC and the SOS cone intersect but not contain each other, the Lasserre relaxation applies to cases where the SONC approach may not work (properly) and vice versa. Moreover, one should take advantage of the different runtimes of SDPs and GPs/REPs for various problems. Therefore a joint method taking all this into account should be highly promising and rewarding.

Bibliography

[AM17] A.A. Ahmadi and A. Majumdar, *DSOS and SDSOS optimization: more tractable alternatives to sum of squares and semidefinite optimization*, 2017, Preprint, arXiv:1706.02586.

[AOPT13] A.A. Ahmadi, A. Olshevsky, P.A. Parrilo, and J.N. Tsitsiklis, *NP-hardness of deciding convexity of quartic polynomials and related problems*, Math. Program. **137** (2013), no. 1-2, Ser. A, 453–476.

[Art27] E. Artin, *Über die Zerlegung definiter Funktionen in Quadrate*, Abh. Math. Sem. Univ. Hamburg **5** (1927), no. 1, 100–115.

[AY03] H. Anai and H. Yanami, *SyNRAC: a Maple-package for solving real algebraic constraints*, Computational science—ICCS 2003. Part I, Lecture Notes in Comput. Sci., vol. 2657, Springer, Berlin, 2003, pp. 828–837.

[Bar02] A. Barvinok, *A Course in Convexity*, Graduate Studies in Mathematics, vol. 54, American Mathematical Society, Providence, RI, 2002.

[BCR98] J. Bochnak, M. Coste, and M.-F. Roy, *Real Algebraic Geometry*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3), vol. 36, Springer-Verlag, Berlin, 1998.

[BCSS98] L. Blum, F. Cucker, M. Shub, and S. Smale, *Complexity and Real Computation*, Springer-Verlag, New York, 1998.

[Ber87] C. Berg, *The multidimensional moment problem and semigroups*, Moments in mathematics, Proc. Sympos. Appl. Math., vol. 37, Amer. Math. Soc., Providence, RI, 1987, pp. 110–124.

[BG08] S. Boyd and M. Grant, *Graph implementations for nonsmooth convex programs*, Recent Advances in Learning and Control, Lect. Notes Control Inf. Sci., vol. 371, Springer, London, 2008, pp. 95–110.

[BHO+12] G. Blekherman, J. Hauenstein, J.C. Ottem, K. Ranestad, and B. Sturmfels, *Algebraic boundaries of Hilbert's SOS cones*, Compos. Math. **148** (2012), no. 6, 1717–1735.

[BIK15] G. Blekherman, S. Iliman, and M. Kubitzke, *Dimensional differences between faces of the cones of nonnegative polynomials and sums of squares*, Int. Math. Res. Not. IMRN (2015), no. 18, 8437–8470.

[BKVH07] S. Boyd, S.J. Kim, L. Vandenberghe, and A. Hassibi, *A tutorial on geometric programming*, Optim. Eng. **8** (2007), no. 1, 67–127.

[Ble06] G. Blekherman, *There are significantly more nonnegative polynomials than sums of squares*, Israel J. Math. **153** (2006), 355–380.

[Ble12] ———, *Nonnegative polynomials and sums of squares*, J. Amer. Math. Soc. **25** (2012), no. 3, 617–635.

[BPT13] G. Blekherman, P.A. Parrilo, and R.R. Thomas, *Semidefinite Optimization and Convex Algebraic Geometry*, MOS-SIAM Series on Optimization, vol. 13, SIAM and the Mathematical Optimization Society, Philadelphia, 2013.

[BŠ16] A. Buckley and K. Šivic, *Nonnegative biquadratic forms with maximal number of zeros*, 2016, Preprint, arXiv:1611.09513.

[BV04] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, UK, 2004.

[CKLR82] M.D. Choi, M. Knebusch, Y.T. Lam, and B. Reznick, *Transversal zeros and positive semidefinite forms*, Real algebraic geometry and quadratic forms, Lecture Notes in Math., vol. 959, Springer, Berlin-New York, 1982, pp. 273–298.

[CL77] M.D. Choi and Y.T. Lam, *An old question of Hilbert*, Queen's Papers in Pure and Appl. Math. **46** (1977), 385–405.

[CL78] ———, *Extremal positive semidefinite forms*, Math. Ann. **231** (1977/78), no. 1, 1–18.

[CLR80] M.D. Choi, Y.T. Lam, and B. Reznick, *Real zeros of positive semidefinite forms. I*, Math. Z. **171** (1980), no. 1, 1–26.

[CS16] V. Chandrasekaran and P. Shah, *Relative entropy relaxations for signomial optimization*, SIAM J. Optim. **26** (2016), no. 2, 1147–1173.

[CS17] ———, *Relative entropy optimization and its applications*, Math. Program. **161** (2017), no. 1-2, Ser. A, 1–32.

[DIdW17] M. Dressler, S. Iliman, and T. de Wolff, *A Positivstellensatz for sums of nonnegative circuit polynomials*, SIAM J. Appl. Algebra Geom. **1** (2017), no. 1, 536–555.

[DIdW18] ———, *An approach to constrained polynomial optimization via nonnegative circuit polynomials and geometric programming*, Preprint available at arXiv:1602.06180, to appear in J. Symbolic Comput., 2018.

[dK02] E. de Klerk, *Aspects of Semidefinite Programming*, Applied Optimization, vol. 65, Kluwer Academic Publishers, Dordrecht, 2002, Interior point algorithms and selected applications.

[dKL10] E. de Klerk and M. Laurent, *Error bounds for some semidefinite programming approaches to polynomial minimization on the hypercube*, SIAM J. Optim. **20** (2010), no. 6, 3104–3120.

[dKS10] E. de Klerk and R. Sotirov, *Exploiting group symmetry in semidefinite programming relaxations of the quadratic assignment problem*, Math. Program. **122** (2010), no. 2, Ser. A, 225–246.

[DPZ67] R.J. Duffin, E.L. Peterson, and C. Zener, *Geometric Programming: Theory and Application*, John Wiley & Sons, Inc., New York-London-Sydney, 1967.

[dW15] T. de Wolff, *Amoebas, nonnegative polynomials and sums of squares supported on circuits*, Oberwolfach Rep. (2015), no. 23, 53–56.

[FK11] C. Fidalgo and A. Kovacec, *Positive semidefinite diagonal minus tail forms are sums of squares*, Math. Z. **269** (2011), no. 3-4, 629–645.

[GBY06] M. Grant, S. Boyd, and Y. Ye, *Disciplined convex programming*, Global optimization, Nonconvex Optim. Appl., vol. 84, Springer, New York, 2006, pp. 155–210.

[GLM14] M. Ghasemi, J.B. Lasserre, and M. Marshall, *Lower bounds on the global minimum of a polynomial*, Comput. Optim. Appl. **57** (2014), no. 2, 387–402.

[GM12] M. Ghasemi and M. Marshall, *Lower bounds for polynomials using geometric programming*, SIAM J. Optim. **22** (2012), no. 2, 460–473.

[GM13] ______, *Lower bounds for a polynomial on a basic closed semialgebraic set using geometric programming*, 2013, Preprint, arXiv:1311.3726.

[GP04] K. Gatermann and P.A. Parrilo, *Symmetry groups, semidefinite programs, and sums of squares*, J. Pure Appl. Algebra **192** (2004), no. 1-3, 95–128.

[GW95] M.X. Goemans and D.P. Williamson, *Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming*, J. Assoc. Comput. Mach. **42** (1995), no. 6, 1115–1145.

[Hab40] W. Habicht, *Über die Zerlegung strikte definiter Formen in Quadrate*, Comment. Math. Helv. **12** (1940), 317–322.

[Har99] W.R. Harris, *Real even symmetric ternary forms*, J. Algebra **222** (1999), no. 1, 204–245.

[Hav36] E.K. Haviland, *On the momentum problem for distribution functions in more than one dimension. II*, Amer. J. Math. **58** (1936), no. 1, 164–168.

[Her95] A. Hertz, *Polynomially solvable cases for the maximum stable set problem*, Discrete Appl. Math. **60** (1995), no. 1-3, 195–210.

[Hil88] D. Hilbert, *Über die Darstellung definiter Formen als Summe von Formenquadraten*, Math. Ann. **32** (1888), no. 3, 342–350.

[Hil93] ______, *Über ternäre definite Formen*, Acta Math. **17** (1893), no. 1, 169–197.

[HLL09] D. Henrion, J.B. Lasserre, and J. Löfberg, *GloptiPoly 3: moments, optimization and semidefinite programming*, Optim. Methods Softw. **24** (2009), no. 4-5, 761–779.

[HLZ13] S. He, Z. Li, and S. Zhang, *Approximation algorithms for discrete polynomial optimization*, Journal of the Operations Research Society of China **1** (2013), no. 1, 3–36.

[Hur91] A. Hurwitz, *Über den Vergleich des arithmetischen und des geometrischen Mittels*, J. Reine Angew. Math. **108** (1891), 266–268.

[IdW16a] S. Iliman and T. de Wolff, *Amoebas, nonnegative polynomials and sums of squares supported on circuits*, Res. Math. Sci. **3** (2016), 3:9.

[IdW16b] ———, *Lower bounds for polynomials with simplex Newton polytopes based on geometric programming*, SIAM J. Optim. **26** (2016), no. 2, 1128–1146.

[Kri64a] J.-L. Krivine, *Anneaux préordonnés*, J. Analyse Math. **12** (1964), 307–326.

[Kri64b] ———, *Quelques propriétés des préordres dans les anneaux commutatifs unitaires*, C. R. Acad. Sci. Paris **258** (1964), 3417–3418.

[KS18] A. Kunert and C. Scheiderer, *Extreme positive ternary sextics*, Trans. Amer. Math. Soc. **370** (2018), no. 6, 3997–4013.

[Las01] J.B. Lasserre, *Global optimization with polynomials and the problem of moments*, SIAM J. Optim. **11** (2000/01), no. 3, 796–817.

[Las06a] ———, *Convergent SDP-relaxations in polynomial optimization with sparsity*, SIAM J. Optim. **17** (2006), no. 3, 822–843.

[Las06b] ———, *A sum of squares approximation of nonnegative polynomials*, SIAM J. Optim. **16** (2006), no. 3, 751–765.

[Las07] ———, *Sufficient conditions for a real polynomial to be a sum of squares*, Arch. Math. (Basel) **89** (2007), no. 5, 390–398.

[Las08] ———, *A semidefinite programming approach to the generalized problem of moments*, Math. Program. **112** (2008), no. 1, Ser. B, 65–92.

[Las10] ———, *Moments, Positive Polynomials and Their Applications*, Imperial College Press Optimization Series, vol. 1, Imperial College Press, London, 2010.

[Las15] ———, *An Introduction to Polynomial and Semi-Algebraic Optimization*, Cambridge Texts in Applied Mathematics, Cambridge University Press, Cambridge, 2015.

[Lau08] M. Laurent, *Semidefinite programming in combinatorial and polynomial optimization*, Nieuw Arch. Wiskd. (5) **9** (2008), no. 4, 256–262.

[Lau09] ———, *Sums of squares, moment matrices and optimization over polynomials*, Emerging Applications of Algebraic Geometry, IMA Vol. Math. Appl., vol. 149, Springer, New York, 2009, pp. 157–270.

[LN07] J.B. Lasserre and T. Netzer, *SOS approximations of nonnegative polynomials via simple high degree perturbations*, Math. Z. **256** (2007), no. 1, 99–112.

[Lov79] L. Lovász, *On the Shannon capacity of a graph*, IEEE Trans. Inform. Theory **25** (1979), no. 1, 1–7.

[Lov03] ———, *Semidefinite programs and combinatorial optimization*, Recent advances in algorithms and combinatorics, CMS Books Math./Ouvrages Math. SMC, vol. 11, Springer, New York, 2003, pp. 137–194.

[LTY17] J.B. Lasserre, K.-C. Toh, and S. Yang, *A bounded degree SOS hierarchy for polynomial optimization*, EURO J. Comput. Optim. **5** (2017), no. 1-2, 87–117.

[LV12] M. Laurent and F. Vallentin, *Semidefinite optimization*, 2012, Extended lecture notes.

[Mar08] M. Marshall, *Positive Polynomials and Sums of Squares*, Mathematical Surveys and Monographs, vol. 146, American Mathematical Society, Providence, RI, 2008.

[Mir99] R. Miranda, *Linear systems of plane curves*, Notices Amer. Math. Soc. **46** (1999), no. 2, 192–201.

[MK87] K. G. Murty and S. N. Kabadi, *Some NP-complete problems in quadratic and nonlinear programming*, Math. Programming **39** (1987), no. 2, 117–129.

[Mot67] T. Motzkin, *The arithmetic-geometric inequality*, Inequalities, Academic Press, New York, 1967, pp. 205–224.

[Nes00] Y. Nesterov, *Squared functional systems and optimization problems*, High performance optimization, Appl. Optim., vol. 33, Kluwer Acad. Publ., Dordrecht, 2000, pp. 405–440.

[Nie12] J. Nie, *Discriminants and nonnegative polynomials*, J. Symbolic Comput. **47** (2012), no. 2, 167–191.

[Nie13a] ______, *Certifying convergence of Lasserre's hierarchy via flat truncation*, Math. Program. **142** (2013), no. 1-2, Ser. A, 485–510.

[Nie13b] ______, *An exact Jacobian SDP relaxation for polynomial optimization*, Math. Program. **137** (2013), no. 1-2, Ser. A, 225–255.

[Nie14] ______, *Optimality conditions and finite convergence of Lasserre's hierarchy*, Math. Program. **146** (2014), no. 1-2, Ser. A, 97–121.

[NN94] Y. Nesterov and A. Nemirovskii, *Interior-Point Polynomial Algorithms in Convex Programming*, SIAM Studies in Applied Mathematics, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1994.

[Oxl11] J. Oxley, *Matroid Theory*, second ed., Oxford Graduate Texts in Mathematics, vol. 21, Oxford University Press, Oxford, UK, 2011.

[Par00] P.A. Parrilo, *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization*, 2000, PhD Thesis, California Institute of Technology.

[Par03] ______, *Semidefinite programming relaxations for semialgebraic problems*, Math. Program. **96** (2003), no. 2, Ser. B, 293–320.

[PD01] A. Prestel and C. Delzell, *Positive Polynomials: From Hilbert's 17th Problem to Real Algebra*, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2001.

[PR01] V. Powers and B. Reznick, *A new bound for Pólya's theorem with applications to polynomials positive on polyhedra*, J. Pure Appl. Algebra **164** (2001), no. 1-2, 221–229.

[PS03] P.A. Parrilo and B. Sturmfels, *Minimizing polynomial functions*, Algorithmic and quantitative real algebraic geometry, DIMACS Ser. Discrete Math. Theoret. Comput. Sci., vol. 60, Amer. Math. Soc., Providence, RI, 2003, pp. 83–99.

[Put93] M. Putinar, *Positive polynomials on compact semi-algebraic sets*, Indiana University Mathematics Journal **42** (1993), no. 3, 969–984.

[PY17] D. Papp and S. Yıldız, *Sum-of-squares optimization without semidefinite programming*, 2017, Preprint, arXiv:1712.01792.

[Qua15] R. Quarez, *On the real zeros of positive semidefinite biquadratic forms*, Comm. Algebra **43** (2015), no. 3, 1317–1353.

[Rez78] B. Reznick, *Extremal PSD forms with few terms*, Duke Math. J. **45** (1978), no. 2, 363–374.

[Rez89] ______, *Forms derived from the arithmetic-geometric inequality*, Math. Ann. **283** (1989), no. 3, 431–464.

[Rez92] ______, *Sums of even powers of real linear forms*, Mem. Amer. Math. Soc. **96** (1992), no. 463, viii+155.

[Rez95] ______, *Uniform denominators in Hilbert's seventeenth problem*, Math. Z. **220** (1995), no. 1, 75–97.

[Rez00] ______, *Some concrete aspects of Hilbert's 17th problem*, Real algebraic geometry and ordered structures, Contemp. Math., vol. 253, Amer. Math. Soc., Providence, RI, 2000, pp. 251–272.

[Rez07] ______, *On Hilbert's construction of positive polynomials*, 2007, Preprint, arXiv:0707.2156.

[RTAL13] C. Riener, T. Theobald, L.J. Andrén, and J.B. Lasserre, *Exploiting symmetries in SDP-relaxations for polynomial optimization*, Math. Oper. Res. **38** (2013), no. 1, 122–141.

[Sch91] K. Schmüdgen, *The K-moment problem for compact semi-algebraic sets*, Math. Ann. **289** (1991), no. 2, 203–206.

[Sch02] M. Schweighofer, *An algorithmic approach to Schmüdgen's Positivstellensatz*, J. Pure Appl. Algebra **166** (2002), no. 3, 307–319.

[Sch05] ______, *Optimization of polynomials on compact semialgebraic sets*, SIAM J. Optim. **15** (2005), no. 3, 805–825.

[Sch09] C. Scheiderer, *Positivity and sums of squares: A guide to recent results*, Emerging Applications of Algebraic Geometry, IMA Vol. Math. Appl., vol. 149, Springer, New York, NY, 2009, pp. 271–324.

[Sha77] I.R. Shafarevich, *Basic Algebraic Geometry*, study ed., Springer-Verlag, Berlin-New York, 1977.

[Sho87a] N.Z. Shor, *An approach to obtaining global extrema in polynomial problems of mathematical programming*, Kibernetika (1987), no. 5, 102–106, 136.

[Sho87b] ———, *A class of estimates for the global minimum of polynomial functions*, Kibernetika (1987), no. 6, 9–11, 133.

[Ste74] G. Stengle, *A Nullstellensatz and a Positivstellensatz in semialgebraic geometry*, Math. Ann. **207** (1974), 87–97.

[Val09] F. Vallentin, *Symmetry in semidefinite programs*, Linear Algebra Appl. **430** (2009), no. 1, 360–369.

[VB96] L. Vandenberghe and S. Boyd, *Semidefinite programming*, SIAM Rev. **38** (1996), no. 1, 49–95.

[WKKM06] H. Waki, S. Kim, M. Kojima, and M. Muramatsu, *Sums of squares and semidefinite program relaxations for polynomial optimization problems with structured sparsity*, SIAM J. Optim. **17** (2006), no. 1, 218–242.

[XY17] J. Xu and Y. Yao, *A proof of Hilbert's theorem on ternary quartic forms with the ladder technique*, 2017, Preprint, arXiv:1701.08294.

[Zie95] G.M. Ziegler, *Lectures on Polytopes*, Graduate Texts in Mathematics, vol. 152, Springer-Verlag, New York, 1995.

Deutsche Zusammenfassung

Die Forschungsergebnisse der Dissertation liegen im Schnitt der reellen algebraischen Geometrie, Konvexgeometrie und Optimierung, das heißt im Gebiet der konvexen algebraischen Geometrie.

Ein zentrales Problem der reellen algebraischen Geometrie und der polynomiellen Optimierung ist es, die Nichtnegativität eines reellen Polynoms zu entscheiden. Wir stellen uns daher folgende Frage:

$$\text{Sei } f \in \mathbb{R}[\mathbf{x}] = \mathbb{R}[x_1, \ldots, x_n]. \text{ Gilt } f(\mathbf{x}) \geqslant 0 \text{ für alle } \mathbf{x} \in \mathbb{R}^n?$$

Das Ziel globaler polynomieller Optimierung ist es, ein reelles multivariates Polynom f über $\mathbb{R}^n$ zu minimieren, das heißt den Optimalwert $f^* = \inf\{f(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^n\}$ zu bestimmen. Wie man leicht sehen kann, ist die Suche nach einer globalen unteren Schranke eines Polynoms f äquivalent dazu, die größte reelle Zahl γ zu finden, sodass $f - \gamma$ nichtnegativ ist, also

$$f^* = \sup\{\gamma \in \mathbb{R} : f(\mathbf{x}) - \gamma \geqslant 0 \text{ für alle } \mathbf{x} \in \mathbb{R}^n\}.$$

Das polynomielle Optimierungsproblem kann daher auf die Frage der Nichtnegativität eines Polynoms reduziert werden.

Das Problem findet in der Entscheidungs- sowie Optimierungsversion zahlreiche Anwendungen, wie zum Beispiel in dynamischen Systemen, in der Robotertechnik, Steuerungstheorie, Computervision, Signalverarbeitung und in der Ökonomie. Für einen Überblick siehe z.B. [BPT13] und [Las10].

Die Nichtnegativität eines Polynoms zu entscheiden ist im Allgemeinen co-NP-schwer, siehe [MK87]. Folglich ist es von Interesse hinreichende Bedingungen zu finden, die die Nichtnegativität eines Polynoms zertifizieren und leichter überprüfbar sind. Solch ein Zertifikat ist durch Summen von Quadraten (sums of squares, kurz: SOS) gegeben, die offensichtlich nichtnegativ sind. Falls ein Polynom f demnach als Quadratsumme geschrieben werden kann, ist es allein aus der Darstellung ersichtlich, dass f nichtnegativ ist. Der Zusammenhang zwischen nichtnegativen Polynomen und Quadratsummen ist eine klassische Frage in der reellen algebraischen Geometrie, die ihre Anfänge Ende des neunzehnten Jahrhunderts in Hilberts Arbeiten hat. Dieser studierte intensiv die konvexen Kegel $P_{n,2d}$ der nichtnegativen Polynome und $\Sigma_{n,2d}$ der Quadratsummen in n Variablen vom Grad höchstens $2d$. Seine Untersuchungen führten schließlich zu dem wegweisenden Resultat, dass die beiden Kegel genau in drei Fällen übereinstimmen, im univariaten Fall, im quadratischen Fall und für binäre Quartiken. In allen anderen Fällen konnte Hilbert nachweisen, dass nichtnegative Polynome existieren, die keine Summen von Quadraten sind. Dieser Beweis war jedoch nicht konstruktiv und das erste Beispiel für ein solches Polynom wurde erst siebzig Jahre später von Motzkin [Mot67] angegeben. Die Tatsache, dass nicht jedes nichtnegative Polynom als SOS geschrieben werden kann, motivierte Hilbert zu der berühmten Frage: „Besitzt jedes nichtnegative Polynom eine Darstellung als Quadratsumme rationaler Funktionen?“ Diese Frage ist als Hilberts 17. Problem bekannt, welche 1927 von Emil Artin [Art27] positiv beantwortet wurde. Für einen historischen Überblick siehe [Rez00].

Der Vorteil SOS Zertifikate zu nutzen ist, dass die Frage, ob ein Polynom als SOS darstellbar ist, mittels eines semidefiniten Optimierungsproblems (SDP) gelöst werden kann. SDPs gehören zur Klasse konvexer Optimierungsprobleme [BV04, VB96] und können als Verallgemeinerung der linearen Optimierung gesehen werden. Überdies existieren gute numerische Algorithmen mit denen sich SDPs (mit beliebiger Genauigkeit) in Polynomialzeit lösen lassen, siehe [BPT13, Seite 41]. Somit kann man die Nichtnegativitätsbedingung in polynomiellen Optimierungsproblemen im globalen sowie restringierten Fall zu einer Quadratsummenbedingung relaxieren. Diese Relaxierung lässt sich nun effizient durch semidefinite Optimierung berechnen. Der SOS/SDP Ansatz der polynomiellen Optimierung geht auf Shor im Jahre 1987 zurück und wurde von Nesterov [Nes00], Parrilo [Par00, Par03] und Lasserre [Las01] weiterentwickelt. Seitdem hat diese Forschungsrichtung eine beachtliche Entwicklung erfahren und es wurden zahlreiche Relaxierungsmethoden vorgeschlagen, welche hinsichtlich verschiedener Aspekte (Laufzeit der Berechnungen, Exaktheit und Qualität der Relaxierung und Geometrie der zugrunde liegenden Strukturen) intensiv untersucht wurden. Die meisten dieser Resultate basieren auf der Lasserre-Relaxierung [Las01], die auf der SOS/SDP Methode beruht und eine Hierarchie konvergierender unterer Schranken an den Optimalwert restringierter Optimierungsprobleme liefert, siehe z.B. [Las10], [Las15].

Ein bekanntes Problem des SOS/SDP Ansatzes ist, dass die Größe des korrespondierenden semidefiniten Programms mit wachsender Variablenanzahl oder wachsendem Grad des Polynoms rapide zunimmt. Für viele Anwendungen lässt sich daher mit diesem Ansatz nur schwer eine (geeignete) Lösung finden. Weiterhin hat Blekherman [Ble06] gezeigt, dass es für einen festen Grad $2d \geqslant 4$ und wachsender Variablenanzahl signifikant mehr nichtnegative Polynome als Quadratsummen gibt. Sich diesen Problemen zu widmen ist ein aktives Forschungsgebiet, zu dem diese Dissertation von theoretischer und praktischer Seite beiträgt.

Ein Schwerpunkt der momentanen Forschung liegt darauf, die „solver" zu verbessern und zusätzliche Strukturen wie Symmetrien und Dünnbesetztheit auszunutzen, siehe z.B. [Val09] und [Las06a].

Im Gegensatz dazu verfolgt diese Arbeit den Ansatz andere Nichtnegativitätszertifikate, unabhängig der SOS Zertifikate, zu verwenden.

Kürzlich führten Iliman und de Wolff [IdW16a] Summen von nichtnegativen Kreispolynomen (sums of nonnegative circuit polynomials, kurz: SONC) als neues Nichtnegativitätszertifikat für reelle Polynome ein. Kreispolynome f haben folgende spezielle Struktur bezüglich ihrer Trägermenge: das Newtonpolytop von f formt einen Simplex mit geraden Ecken, die Koeffizienten der zu den Ecken korrespondierenden Terme sind strikt positiv und es gibt einen zusätzlichen Trägerpunkt, der im Inneren des Simplex liegt. Für jedes Kreispolynom kann man die zugehörige Kreiszahl definieren, die sich sofort aus dem gegebenen Polynom ergibt. Der entscheidende Faktor ist, dass die Nichtnegativität von Kreispolynomen mittels dieser Kreiszahl einfach getestet werden kann. Die Menge der Summen von nichtnegativen Kreispolynomen in n Variablen vom Grad höchstens $2d$ wird mit $C_{n,2d}$ bezeichnet und ist sogar ein konvexer Kegel, der zwar den SOS Kegel $\Sigma_{n,2d}$ schneidet, jedoch sind beide Kegel nicht ineinander enthalten. Somit stellen Summen von nichtnegativen Kreispolynomen in der Tat ein neues Nichtnegativitätszertifikat, unabhängig von SOS Zertifikaten, dar.

In dieser Dissertation untersuchen wir Summen von nichtnegativen Kreispolynomen sowie den zugehörigen Kegel $C_{n,2d}$ und studieren diese geometrisch und betrachten ihre Anwendung in der polynomiellen Optimierung. Dies führt zu neuen Resultaten in den Gebieten der reinen und angewandten reellen algebraischen und konvexen algebraischen Geometrie. Diese Arbeit gliedert sich in zwei Teile, das theoretische Studium des SONC Kegels und das praktische Studium der Anwendung in der polynomiellen Optimierung. Die nachstehenden Abschnitte skizzieren die untersuchten Probleme und geben einen Überblick über die Resultate und Beiträge dieser Arbeit.

Der SONC Kegel genauer betrachtet. SONC Polynome und deren Kegel sind noch zu einem großen Teil unerforscht, bringen aber ein hohes Forschungspotential mit sich. Vom theoretischen Standpunkt aus ist der SONC Kegel, als konvexer

Kegel von Polynomen mit spezieller Struktur, an sich bereits interessant. Vor allem aber unter dem Aspekt, dass $C_{n,2d}$ den Nichtnegativitätskegel $P_{n,2d}$ approximiert, ist ein tieferes Verständnis des SONC Kegels aus theoretischer sowie praktischer Sicht äußerst wünschenswert. Dessen Studium reiht sich somit in die klassische Theorie der nichtnegativen Polynome und Quadratsummen ein. Daher ist es wichtig, die Struktur und die (konvexen) Eigenschaften von $C_{n,2d}$ sowie dessen Beziehung zu $P_{n,2d}$ und $\Sigma_{n,2d}$ zu erforschen.

Dadurch motiviert werden zunächst einige konvexgeometrische Aspekte des SONC Kegels studiert. Als Erstes zeigen wir in Proposition 3.1.1, dass $C_{n,2d}$ ein echter Kegel ist. In [IdW16a] werden die Fälle $(n, 2d)$ charakterisiert, in denen die Kegel $C_{n,2d}$ und $\Sigma_{n,2d}$ sich enthalten, beziehungsweise nicht enthalten, siehe Theorem 2.4.8. Zwei Fälle sind darin nicht abgedeckt: $(n, 2)$ für alle $n \geqslant 2$ und der Fall $(n, 4)$ für alle n. Wir schließen diese Lücke in Theorem 3.1.2.

In der Literatur sind Resultate für nichtnegative Polynome und SOS häufig homogen formuliert. Homogene Polynome sind allgegenwärtig in der Mathematik und ein zentrales Studienobjekt der algebraischen Geometrie. Daher wollen wir ebenfalls homogene SONC Polynome studieren. Hierfür zeigen wir zunächt die fundamentale Tatsache, dass die SONC Eigenschaft unter Homogenisierung erhalten bleibt.

Ein interessantes Forschungsobjekt für (homogene) Polynome sind deren reelle Nullstellen. Es gibt eine Vielzahl von Arbeiten, welche die reellen Nullstellen von nichtnegativen Polynomen und Quadratsummen studieren. In diesem Zusammenhang werden die Nullstellen häufig dazu genutzt die mengentheoretische Differenz der beiden Kegel zu untersuchen und einen Einblick in die Seitenstruktur von $P_{n,2d}$ und $\Sigma_{n,2d}$ zu gewinnen. Hierfür siehe z.B. [BHO+12, Ble12, CL77, CLR80, KS18, Rez78, Rez00]. Durch diese Ideen motiviert, untersuchen wir die reellen Nullstellen von (homogenen) SONC Polynomen. Der Hauptbeitrag der Dissertation zu diesem Thema ist eine vollständige und explizite Charakterisierung der reellen Nullstellen von SONC Polynomen sowie homogenen SONC Polynomen, siehe Abschnitt 3.2. Die Nullstellenresultate führen zu weiteren interessanten Beobachtungen. Zum Beispiel folgt, dass das Analogon zu Hilberts 17. Problem für SONC Polynome allgemein nicht gelten kann. Auf dem Studium der Nullstellen aufbauend wird eine erste Betrachtung der exponierten Seiten des SONC Kegels dargelegt. Insbesondere erhalten wir hierbei Schranken für die Dimensionen der exponierten Seiten von $C_{n,2d}$ und studieren den univariaten und bivariaten Fall mitsamt expliziten Beispielen genauer.

Eine grundlegende Eigenschaft von SOS ist, dass die Menge der Quadratsummen multiplikativ abgeschlossen ist. Diese Eigenschaft ist essentiell für Anwendungen der Quadratsummen in der polynomiellen Optimierung, im Speziellen für bestimme Positivstellensätze, siehe Abschnitt 2.3.4. In Lemma 3.4.1 wird gezeigt, dass die Menge der SONC Polynome hingegen nicht multiplikativ abgeschlossen ist. Ein weiterer Hauptbeitrag zur Analyse des SONC Kegels ist das Resultat, dass $C_{n,2d}$ volldimensional in dem konvexen Kegel der nichtnegativen Polynome $P_{n,2d}$ ist. Dieses Resultat ist eine notwendige Bedingung dafür, SONC Polynome als in der Praxis nützliche Zertifikate zu etablieren. Daher haben beide Beobachtungen einen direkten Einfluss auf die Anwendung von SONC Polynomen in polynomiellen Optimierungsproblemen. Diese angewandte Perspektive wird in den kommenden Abschnitten diskutiert.

Polynomielle Optimierung mittels SONC und GP. Wie bereits erwähnt, ist auch aus praktischer Sicht ein tieferes Verständnis des SONC Kegels und der SONC Polynome von großem Nutzen. Der zweite Teil dieser Arbeit widmet sich der Anwendung von SONC Polynomen in der Optimierung.

Neben dem SDP-basierten Ansatz für polynomielle Optimierungsprobleme haben Ghasemi und Marshall [GM12, GM13] kürzlich vorgeschlagen geometrische Programme für globale und restringierte Optimierungsprobleme zu nutzen. Ein geometrisches Optimierungsproblem (GP) ist konvex und kann (bis auf einen ε-Fehler) mittels Innere-Punkte-Verfahren in Polynomialzeit gelöst werden [NN94], siehe auch [BKVH07, S. 118]. Wie experimentelle Resultate zeigen, können GPs in der Praxis wesentlich schneller als SDPs berechnet werden, siehe z.B. [BKVH07, GM12, GM13, GLM14]. Ein Nachteil der Methode von Ghasemi und Marshall ist jedoch, dass die unteren Schranken, die man mittels GP erhält, nicht so gut sind wie die SDP-basierten Schranken und der Ansatz nur auf sehr spezielle Fälle anwendbar ist.

Iliman und de Wolff [IdW16b] zeigten, dass der GP-basierte Ansatz für globale Optimierung durch SONC Zertifikate für bestimmte Polynome verallgemeinert werden kann. Genauer gesagt kann mittels GP effizient entschieden werden, ob ST-Polynome eine SONC Zerlegung besitzen. Kreispolynome gehören zur Klasse der ST-Polynome, deren Newtonpolytop ein Simplex ist und die weitere Bedingungen erfüllen, siehe Definition 4.1.1. Somit ist der Zusammenhang zwischen SONC und GP analog dem zwischen SOS und SDP. Ein wesentlicher Unterschied zum Ansatz von Ghasemi und Marshall ist, dass diverse Polynomklassen existieren, für die die SONC/GP-basierte Methode nicht nur schneller ist, sondern auch bessere Schranken als der SOS/SDP Ansatz liefert, siehe hierzu [IdW16b, Korollar 3.6]. Das ist darauf zurückzuführen, dass die von Ghasemi und Marshall genutzten Zertifikate immer SOS sind, was für SONC Zertifikate im Allgemeinen nicht zutrifft, siehe Theorem 2.4.8.

Durch diese jüngsten Entwicklungen motiviert liegt der Fokus des zweiten Teils dieser Arbeit darauf, restringierte Optimierungsprobleme mit SONC Polynomen zu studieren. Restringierte Optimierungsprobleme haben die Form: $f_K^* = \inf\{f(\mathbf{x}) : \mathbf{x} \in K\} = \sup\{\gamma \in \mathbb{R} : f(\mathbf{x}) - \gamma \geqslant 0$ für alle $\mathbf{x} \in K\}$, wobei $K \subseteq \mathbb{R}^n$ die basisch abgeschlossene semialgebraische Menge der Polynome $g_1, \ldots, g_s \in \mathbb{R}[\mathbf{x}]$ ist.

Im Wesentlichen verfolgen wir unser Ziel auf zwei verschiedene Weisen. Zunächst als Verallgemeinerung des beschriebenen SONC/GP Ansatzes. Das heißt, wir erhalten eine untere Schranke an f_K^*, indem wir ein einziges konvexes Optimierungsproblem lösen, welches unter gewissen Annahmen ein GP ist. Danach wird ein erweiterter Ansatz analysiert, welcher eine Hierarchie unterer Schranken liefert, die gegen f_K^* konvergiert. Dieser hierarchische Ansatz wird im nächsten Abschnitt diskutiert.

Der erste Beitrag dieser Dissertation im Kontext polynomieller Optimierung ist eine Erweiterung der Resultate in [IdW16b] auf restringierte polynomielle Optimierungsprobleme für die Klasse der ST-Polynome. Der Ansatzpunkt hierfür ist ein allgemeines Optimierungsproblem von [IdW16b, Abschnitt 5], siehe Programm (4.1.5), welches eine untere Schranke für das restringierte Problem liefert, aber nicht durch ein GP berechnet werden kann. Das Programm (4.1.5) kann mittels Resultaten aus [GM13] zu einem GP relaxiert werden, siehe Programm (4.2.2) und Theorem 4.2.1. Überdies wird gezeigt, dass das neue relaxierte GP (4.2.2) für bestimmte Spezialfälle genauso gute Schranken wie das ursprüngliche Programm (4.1.5) liefert.

Abschnitt 4.3 enthält Beispiele, in denen das neue Programm (4.2.2) in der Praxis mit SDP verglichen wird. In den Beispielen ist deutlich zu sehen, dass unser Programm viel schneller als SDP ist. Im Gegensatz zu SDP ist unser GP unempfindlich gegenüber Erhöhung des Grades. Dadurch ist der GP-basierte Ansatz vor allem in hochgradigen Beispielen nützlich, in denen SDPs ernsthafte Probleme aufweisen.
Weiterhin kann eine von Ghasemi und Marshall in [GM13] erhaltene Schranke nie besser sein als die Schranke der d-ten Lasserre-Relaxierung für ein spezifisches d, welches von den Graden der involvierten Polynome bestimmt ist. Abschnitt 4.3 enthält Beispiele, die zeigen, dass das neue Programm (4.2.2) in der Tat Schranken liefert, die besser sind als die spezielle d-te Lasserre-Relaxierung.

Der zweite Beitrag in diesem Kontext ist die Anwendung des SONC/GP Ansatzes für polynomielle Optimierung über die Klasse der ST-Polynome hinaus. In Abschnitt 4.4 wird ein erster Ansatz vorgestellt, der auf einer Triangulierung der Trägermenge der beteiligten Polynome beruht. Dieser liefert, basierend auf der SONC/GP Methode, Nichtnegativitätsschranken für beliebige Polynome im globalen und restringierten Fall. Wieder werden einige Beispiele bereitgestellt, die diesen Ansatz mit dem SDP-basierten vergleichen. Erneut ist das Ergebnis, dass in allen Beispielen, speziell in hochgradigen Fällen, die GP-basierte Methode deutlich schneller ist.

In beiden Ansätzen werden keine Annahmen an den Zulässigkeitsbereich K getroffen. Insbesondere wird keine Kompaktheit verlangt, wie im klassischen Fall der Lasserre-Relaxierung.

Hierarchischer Ansatz in der restringierten Optimierung mittels SONC und REP. Da der SONC/GP-basierte Ansatz für restringierte Optimierungsprobleme nur eine einzige untere Schranke an den Optimalwert f_K^* liefert, wird im Folgenden ein erweiterter Ansatz untersucht, der zu einer konvergierenden Hierarchie von unteren Schranken an f_K^* führt. Der Hauptunterschied beider Ansätze ist, dass letzterer auf einem Positivstellensatz beruht. Positivstellensätze spielen in der Entwicklung der restringierten Optimierung eine zentrale Rolle und haben eine lange Geschichte. Ein Positivstellensatz liefert für ein Polynom, welches auf einer semialgebraischen Menge K strikt positiv ist, eine bestimmte algebraische Darstellung. Es gibt eine Vielzahl von Positivstellensätzen, die typischerweise auf Quadratsummen beruhen, siehe Abschnitt 2.3.3. Beispielsweise basiert die Lasserre-Relaxierung auf Putinars Positivstellensatz [Put93].

Vor Kurzem haben Chandrasekaran und Shah [CS16] Summen von nichtnegativen arithmetisch geometrischen Exponentialen (sums of nonnegative AM/GM-exponentials, kurz: SAGE) als Nichtnegativitätszertifikat für Signome eingeführt. Signome sind gewichtete Summen von Exponentialen, wodurch dieses Konzept das Problem anspricht, die Nichtnegativität eines Polynoms auf dem positiven Orthanten zu entscheiden. Ob ein AM/GM-Exponential nichtnegativ ist, kann mit Hilfe relativer Entropieprogrammierung (REP) getestet werden. Ein REP ist ein konvexes Optimierungsproblem, welches allgemeiner als ein GP ist, jedoch immer noch effizient mittels Innere-Punkte-Verfahren lösbar ist, siehe [CS17, NN94].

Die Grundlage für den hierarchischen Ansatz stellt ein Positivstellensatz für SONC Polynome dar, siehe Theorem 5.3.5, der sich als Konsequenz aus Krivines Positivstellensatz [Kri64a, Kri64b] ergibt. Der SONC Positivstellensatz sagt aus, dass ein Polynom f, wenn es strikt positiv auf der kompakten Menge K ist, durch die restringierenden Polynome gewichtet mit SONC Polynomen dargestellt werden kann.

Durch diesen Positivstellensatz können wir den Parameter $f_{\text{sonc}}^{(d,q)}$ definieren, der durch die größte reelle Zahl γ gegeben ist, sodass $f(\mathbf{x}) - \gamma$ eine SONC Darstellung besitzt. Dieser Parameter ist offensichtlich eine untere Schranke an f_K^*, die auf dem maximal erlaubten Grad der darstellenden Polynome des Positivstellensatzes basiert. Desweiteren ist die untere Schranke monoton wachsend in d und q, siehe Lemma 5.4.1 und liefert daher eine Hierarchie unterer Schranken für f_K^*. Der Hauptbeitrag zum Gebiet der polynomiellen Optimierung ist das Resultat, dass einerseits diese Hierarchie vollständig ist, das heißt, die unteren Schanken $f_{\text{sonc}}^{(d,q)}$ konvergieren für $d, q \to \infty$ gegen f_K^*, siehe Theorem 5.4.2, und andererseits, dass die Schranken $f_{\text{sonc}}^{(d,q)}$ effizient berechenbar sind. Genauer gesagt stellen wir ein Optimierungsprogramm (5.4.3) zur Berechnung von $f_{\text{sonc}}^{(d,q)}$ zur Verfügung und zeigen in Theorem 5.4.3, dass dieses Programm (5.4.3) ein REP ist. Diese Verbindung wurde durch das oben angesprochene neue Konzept des SAGE Kegels inspiriert, der einen Zusammenhang zum SONC Kegel aufweist. Daher bieten wir überdies erstmalig einen Vergleich dieser beiden Kegel an, siehe Abschnitt 5.2.

Im Abschnitt 5.4.3 illustrieren wir die neue Methode an einem Beispiel.

Gliederung der Dissertation. In Kapitel 2 geben wir einen umfangreichen Überblick über die grundlegende Theorie und die relevanten Resultate. Zunächst führen wir einige Notationen ein und wiederholen die wesentlichen Konzepte der Konvexitätstheorie und von Polynomen. Anschließend werden im Abschnitt 2.2 intensiv die Kegel der nichtnegativen Polynome und Summen von Quadraten studiert. Die angesprochenen Themen umfassen die Verbindung der Quadratsummen zur semidefiniten Optimierung, das quantitative Verhältnis der beiden Kegel und Fakten über deren duale Kegel, deren Rand sowie deren Seitenstruktur. Danach werden der Hintergrund polynomieller Optimierungsprobleme und reeller algebraischer Geometrie, wie die SOS Relaxierungen, Positivstellensätze und die berühmte Lasserre-Relaxierung für restringierte Optimierungsprobleme, vorgestellt, siehe Abschnitt 2.3. Abschließend führen wir in Abschnitt 2.4 Summen von nichtnegativen Kreispolynomen ein, das Hauptstudienobjekt der vorliegenden Dissertation, und führen die zugehörige Theorie ein.

Kapitel 3 ist dem konvexgeometrischen Studium des SONC Kegels gewidmet. Zu Beginn präsentieren wir einige Eigenschaften und allgemeine Resultate über die Struktur des SONC Kegels und seiner Beziehung zu dem SOS Kegel. In Abschnitt 3.2 legen wir unseren Fokus auf die reellen Nullstellen der (homogenen) SONC Polynome. Das führt zu einer vollständigen und expliziten Charakterisierung dieser Nullstellen, woraus sich interessante Konsequenzen ergeben. Auf dem Nullstellenwissen aufbauend, geben wir einen ersten Ansatz zum Verständnis der exponierten Seiten des SONC Kegels, siehe Abschnitt 3.3, bei dem wir den univariaten und bivariaten Fall tiefergehend analysieren und Schranken an die Dimension der exponierten Seiten angeben. Im Abschnitt 3.4 wird gezeigt, dass, im Gegensatz zu SOS, die Menge der SONC Polynome nicht unter Multiplikation abgeschlossen ist. Weiterhin wird das wichtige Resultat präsentiert, dass der SONC Kegel volldimensional im Kegel der nichtnegativen Polynome ist.

In den folgenden beiden Kapiteln wenden wir uns dem praktischen Studium des SONC Kegels in der Anwendung auf restringierte polynomielle Optimierung zu. Kapitel 4 diskutiert dieses Problem durch Erlangen einer einzigen unteren Schranke an den Optimalwert, die durch geometrische Programmierung berechnet werden kann. Dazu führen wir erst die untersuchten ST-Polynome und geometrische Optimierung ein. Anschließend rekapitulieren wir den SONC/GP-basierten Ansatz für den globalen Fall und einen initialen Ansatz zum restringierten Fall. Dieser Ansatz liefert allerdings eine untere Schranke, die nicht durch ein GP ermittelt werden kann. Wir erweitern das Resultat für den restringierten Fall in Abschnitt 4.2 und formulieren Relaxierungen, die mittels GPs berechnet werden können. Zusätzlich werden Beispiele aufgezeigt, die den neuen Ansatz in der Praxis mit SDP vergleichen, siehe Abschnitt 4.3. Schließlich verallgemeinern wir den SONC/GP Ansatz in Abschnitt 4.4 im globalen sowie restringierten Fall für Polynome, die nicht ST-Polynome sind.

Kapitel 5 beschreibt einen erweiterten Ansatz für die restringierte Optimierung, die eine Hierarchie von an den Optimalwert konvergierenden unteren Schranken liefert. Zunächst studieren wir den Kegel der Summen von nichtnegativen AM/GM-Exponentialen und führen relative Entropieoptimierung ein. Nach einem Vergleich des SONC und des SAGE Kegels in Abschnitt 5.2 formulieren wir den Positivstellensatz für SONC Polynome, siehe Abschnitt 5.3. Auf diesem Satz aufbauend etablieren wir in Abschnitt 5.4 eine Hierarchie von unteren Schranken an den Optimalwert restringierter Optimierungsprobleme und formulieren ein Optimierungsprogramm zur Berechnung dieser Schranken. Anschließend zeigen wir, dass die bereitgestellte Hierarchie auf einer kompakten Menge vollständig und durch relative Entropieprogramme effizient lösbar ist. Zum Abschluss wird die Zerlegung eines gegebenen Polynoms in die durch den SONC Positivstellensatz gegebene Form an einem Beispiel demonstriert.

Im letzten Kapitel 6 der vorliegenden Dissertation finden sich abschließende Bemerkungen und eine Diskussion offener Fragen.