The strength of pseudo-expectations
A detailed analysis of the work of Lee, Raghavendra and Steurer on the psd rank of the family of correlation polytopes

Doctoral thesis for obtaining the academic degree Doctor of Natural Sciences

submitted by
Gruler, Sebastian

at the
University of Konstanz
Faculty of Sciences
Department of Mathematics and Statistics

Date of the oral examination: June 15, 2018

1. Reviewer: Prof. Dr. Markus Schweighofer (University of Konstanz)
2. Reviewer: Prof. Dr. Rekha Thomas (University of Washington)
Acknowledgments

First of all, my special thanks go to my adviser Markus Schweighofer not only for informative discussions and supervising this work but also for the encouragement and motivation he gave me in less productive phases and the good working atmosphere we always had. Furthermore, I am grateful to him for giving me the opportunity to attend enlightening workshops and conferences.

I am very grateful to Rekha Thomas, on the one hand for reviewing this dissertation, but mainly for her invitation to the Leibniz Center for Informatics at Schloss Dagstuhl in 2015, where I got valuable insights leading to certain parts of this work. I also want to thank Troy Lee for inviting me to the follow-up workshop at Singapore in 2016 and giving me the opportunity to talk about my work and to capture valuable ideas and information. In this connection, I want to thank the Institute for Mathematical Sciences at the National University of Singapore for giving me financial support for my travel.

Furthermore, I want to thank James Lee for some helpful notes and his quick response to my questions concerning his common work with Prasad Raghavendra and David Steurer. Without their work, this dissertation would perhaps not exist.

I want to express my deepest gratitude to all my family, especially my wife Sabrina, my parents Albert and Caroline, my sisters Ricarda and Amélie, and my brother Benedikt, for their love and the encouragement and the support I always received.

I extend my thanks to all of my former colleagues at the math department of the University of Konstanz. I always appreciated the cooperative atmosphere there. In particular, I want to thank my former office mates Aaron Kunert and Christoph Schulze for their helpful suggestions and discussions. Of course, my thanks include the non-scientific staff as Ute Otterbeck, Rainer Janssen and Gisela Cassola who always assisted me in administration matters whenever help was needed.

Finally, I want to thank Mary Beth Webb and Konrad Augustin for the efforts they had with proofreading my work and for giving me some useful hints and advice concerning the language of it.
# Contents

## Introduction

5

## 1 Preliminaries - matrices and linear forms

1.1 The trace, symmetric matrices and matrix norms .......................... 11
1.2 Positive semidefinite matrices ............................................. 17

## 2 Lower degree bounds for the Real Nullstellensatz

2.1 Real Nullstellensatz refutation ........................................... 23
2.2 Multilinear polynomials ..................................................... 25
2.3 Symmetric multilinear polynomials ...................................... 28
2.4 A new proof of Grigoriev’s Theorem ................................. 36
2.5 Two further conclusions from the proof of Grigoriev’s theorem 43
   2.5.1 The Max-Cut problem for the complete graph .................. 43
   2.5.2 A lower degree bound for the iterated real square root ideal
       refutation ........................................................................ 45

## 3 Real-valued Boolean Functions

3.1 Boolean functions over $\{-1,1\}^n$ ........................................ 52
3.2 Boolean functions over $\{0,1\}^n$ ........................................ 56
3.3 Symmetric Boolean functions over $\{0,1\}^n$ ......................... 64
3.4 Nonnegativity and sums of squares of Boolean functions .......... 69
3.5 Pseudo-expectations and pseudo-densities ............................ 72

## 4 A lower bound on the psd rank of the family of correlation polytopes

4.1 Slack matrices and nonnegative, quadratic Boolean functions - basic
   steps to handle the psd rank of the correlation polytope ............... 84
4.2 A proof overview - the basic ideas of Lee, Raghavendra and Steurer 89
4.3 Degree Reduction .................................................................. 93
4.4 Matrix approximation ......................................................... 104
4.5 The main theorem .................................................................. 120
4.6 Quantities of the Grigoriev pseudo-density .......................... 129
   4.6.1 The infinity norm ....................................................... 131
   4.6.2 The degree .................................................................. 136
4.7 Applying the main theorem to the knapsack function and the Grig-
   oriev pseudo-density .......................................................... 139
5 A detailed analysis of the occurring pseudo-densities 146
5.1 Feasible pseudo-densities and symmetry 147
5.2 Grigoriev pseudo-densities are optimal - a complete analysis of symmetric pseudo-densities in the centered standard case 151
5.3 Two conjectures about double symmetric pseudo-densities in the centered non-standard case 163
5.4 A short view on the non-centered case 173

6 Outlook and open questions 180
6.1 Matrix approximation via degree truncation 180
6.2 Further open questions 183

Zusammenfassung auf Deutsch 185

Bibliography 192


Introduction

In combinatorial optimization, many problems can be modeled by optimizing a linear functional over a polyhedron, which is the feasible set of finitely many linear inequalities. The standard tool to manage such a problem is the well-known linear programming. In this connection, the number of linear inequalities describing a polyhedron is a measure for the complexity of the linear program. Since the occurring polyhedra of combinatorial optimization problems are often defined by a large number of linear inequalities, it is hard to directly solve such problems with a linear program.

In this context, note the following observation. Let $P \subseteq \mathbb{R}^n, Q \subseteq \mathbb{R}^N$ be two polyhedra with $n \leq N$, and let $l : \mathbb{R}^n \to \mathbb{R}$ be a linear functional and $\pi : \mathbb{R}^N \to \mathbb{R}^n$ a linear map with $\pi(Q) = P$. Then, note that

$$\min_{x \in P} l(x) = \min_{y \in Q} l \circ \pi(y)$$

holds, and the right-hand side optimization problem is also a linear program.

So, it is a useful tool to write a polyhedron $P$ as the projection or more general the image under a linear map of another polyhedron $Q$ in higher dimension and described by less inequalities than $P$. Then, we can optimize a linear functional over $Q$ instead of $P$ in order to reduce the complexity of the linear program.

A simple but comprehensible and well imaginable because low dimensional example that this might work is the following: an octahedron, a polygon described by eight linear inequalities, is the projection of a three-dimensional polyhedron described by just six linear inequalities.

A more general example, which also shows the power of the above mentioned tool, is the family of permutation polyhedra, which we mention without going into detail.
To every $n \in \mathbb{N}$, the permutation polytope $P_n$ is defined as the convex hull of the set $\{(\sigma(1), \ldots, \sigma(n))^T \mid \sigma \in S_n\}$ of all permutations of the vector $(1, \ldots, n)^T$. By a result of Rado [Rad52], at least $2^n - 2$ linear inequalities are necessary to describe $P_n$. But it is easy to see that $P_n$ is the linear image of the so-called Birkhoff polytope $B_n \subseteq \mathbb{R}^{n \times n}$, the convex hull of all $n \times n$ permutation matrices. In contrast to the permutation polytope, only $n^2$ many linear inequalities are needed to describe $B_n$. This follows by a famous result of Birkhoff and von Neumann. A proof of this result as well as a more detailed explanation of the above mentioned can be found for example in [Bar02, Sections II.5,II.6].

Overall, much research has been done in this area, and a lot of positive results exist. On the other hand, there are also negative ones as we will see soon.

Note that a bounded polyhedron is called a polytope. In combinatorial optimization, the family of the correlation polytopes forms an important family of polytopes. For $n \in \mathbb{N}$, we define

$$\text{CORR}_n := \text{conv}\{xx^T \mid x \in \{0,1\}^n\} \subseteq \mathbb{R}^{n \times n}.$$ 

These polytopes are an example where the above described method reaches its limits. Fiorini et al. [FMP+12] were the first who showed that an exponential number of inequalities is needed to describe any polytope that maps linearly onto CORR$_n$. Meanwhile, concrete lower bounds were proven. In [KW15], Kaibel and Weltge gave a simple proof that at least $1.5^n$ linear inequalities are necessary, which to our knowledge is the currently best known lower bound.

A possibility to provide relief here is to use semidefinite programming, a broad generalization of linear programming. The feasible set of a semidefinite program (SDP) is a so-called spectrahedron, a generalization of a polyhedron as we will see soon. A set $S \subseteq \mathbb{R}^n$ is a spectrahedron if it is of the form

$$S = S_A := \{x \in \mathbb{R}^n \mid A(x) := A_0 + x_1 A_1 + \ldots + x_n A_n \succeq 0\}$$

for some $k \in \mathbb{N}$ and symmetric matrices $A_0, \ldots, A_n \in \mathbb{R}^{k \times k}$. In this connection, $\succeq 0$ means that the symmetric matrix on the left-hand side is positive semidefinite (psd) and $A(x) \succeq 0$ is called a linear matrix inequality (LMI). We call $k$ the size of the LMI and say $S$ is defined by the LMI $A(x) \succeq 0$. Now, a semidefinite program is optimizing a linear functional over a spectrahedron. Here, the size of the LMI that defines the spectrahedron is a measure for the complexity of the SDP.

Note that it is easy to see that every polyhedron is always a spectrahedron. If a polyhedron $P$ is defined by $m$ linear inequalities $l_1(x) \geq 0, \ldots, l_m(x) \geq 0$, we choose $A$ as a diagonal matrix with the $i$-th diagonal entry $l_i$. Then, $A(x) \succeq 0$ is a linear matrix inequality and since a diagonal matrix is psd if and only if its diagonal entries are nonnegative, we obtain $P = S_A$. Therefore, semidefinite programming is indeed a generalization of linear programming.
The power of semidefinite programming lies in the fact that there are many more problems that can be modeled by an SDP than by a linear program. On the other hand, as a special case of convex programming, there are still effective methods to solve a SDP. Besides in combinatorial optimization, semidefinite programming has many applications for example in control theory. More information and examples to applications of semidefinite programming can be found in [BV96, Section 2] and [BPT13, Section 2.2].

In general, [BV96] as well as [BPT13] are both recommended to get an overview of semidefinite programming. They provide definitions, properties and results as well as historical comments and information about the underlying algorithms of semidefinite programming. In addition, the textbook of Blekherman, Parrilo and Thomas is much more detailed and more up-to-date, and it also provides some geometric properties of spectrahedra.

To come back to our original problem, we generalize the idea of (i). For a polyhedron $P \subseteq \mathbb{R}^n$ defined by many linear inequalities, it is useful to have a spectrahedron $S$ defined by a LMI of small size $k$ and a linear map $\pi : \mathbb{R}^N \to \mathbb{R}^n$ such that

$$P = \pi(S).$$

In this case, we say $P$ admits a psd lift of size $k$. Now, let $l : \mathbb{R}^n \to \mathbb{R}$ be a linear functional. Then, we have

$$\min_{x \in P} l(x) = \min_{y \in S} l \circ \pi(y),$$

and the right-hand side optimization problem is a SDP of lower complexity than the left-hand side linear program.

The following is an example for a psd lift of size three of the square, a polyhedron defined by four linear inequalities.

$$S = \left\{ x \in \mathbb{R}^3 \left| \begin{pmatrix} 1 & x_1 & x_2 \\ x_1 & 1 & x_3 \\ x_2 & x_3 & 1 \end{pmatrix} \succeq 0 \right. \right\}$$

*image credit:* [AKW16]
More general results about polyhedra that admits psd lifts of low size can be found for example in [GPRT17].

Now, the question if the correlation polytopes admit psd lifts of low size is of great interest. For this purpose, we define the **psd rank** of a polytope $P$, denoted by $\text{rk}_{\text{psd}}(P)$, as the minimal $k$ such that $P$ admits a psd lift of size $k$.

In their recent breakthrough work [LRS15], Lee, Raghavendra and Steurer negated this question. They proved that $\text{rk}_{\text{psd}}(\text{CORR}_n)$ grows faster than any polynomial in $n$. More precisely, they claimed that there is a constant $\alpha > 0$ and a natural number $N$ such that for every $n \geq N$ it holds

$$\text{rk}_{\text{psd}}(\text{CORR}_n) \geq 2^\alpha \left( \frac{n}{\ln n} \right)^{2/11}.$$  \hspace{1cm} (ii)

In other words, every spectrahedron that linearly maps onto $\text{CORR}_n$ is defined by a LMI that has at least size $2^\alpha \left( \frac{n}{\ln n} \right)^{2/11}$. Their work showed the first-known super-polynomial lower bounds on the size of psd lifts of any explicit family of polytopes.

This work of [LRS15] plays the main role in this dissertation. Mainly, we want to provide their proof of (ii). This will be done in Chapter 4.

Essentially, we want to give some motivation and explanations of the different basic steps they use (Sections 4.1, 4.2), and we will elaborate their proof and embellish it with details they omitted (Sections 4.3, 4.4, 4.5).

Our main intention is to present a detailed version of the work of [LRS15] and in the ideal case a version that makes it easier to understand their brilliant work. In this connection, it is our motivation to bring their work to a broader number of people without much previous knowledge.

On the other hand, we present some of our own ideas and thoughts how the proof could be simplified and how it can be seen from a more real algebraic geometry point of view. At some points, with our ideas we even obtain some slightly stronger results than [LRS15].

By the way, we illustrate why [LRS15] actually show (ii), while their original work just claimed the weaker lower bound $2^\alpha \left( \frac{n}{\ln n} \right)^{2/13}$ (Section 4.6).\footnote{Independently from us, this was also noticed by Troy Lee et al. [LPWY16].}

Overall, our work can be divided into three main parts, where the just mentioned work of [LRS15] is the second one, and it is structured as follows. In Chapter 1, we will just make preparations for the further course of the work, and we will provide some basics mainly from linear algebra.

The first of the three main parts follows in Chapter 2. We will provide a new and easier proof of a famous and often cited theorem of Grigoriev [Gri01a], where he proved lower bounds on the degree of the Real Nullstellensatz (RNS). The Real Nullstellensatz yields a certificate for the infeasibility of a real polynomial equation
system, where the certificate relies on sums of squares. Grigoriev’s Theorem is the first known result stating a lower bound for the RNS in the case that the underlying polynomial equation system contains all the quadratic polynomials $X_1^2 - X_1, \ldots, X_n^2 - X_n$, but this case is very important in combinatorial optimization.

On the one hand, Grigoriev’s Theorem is on its own of high significance for questions concerning the power and complexity of algorithms relying on sums of squares and solving problems as they appear frequently in combinatorial optimization. On the other hand, this result yields a corollary that is essential for the work of [LRS15]. For this reasons, we consider it important to have a more simple proof than the original one.

We will present our new proof in Section 2.4, where we mainly work in the ring $\mathbb{R}[x] := \mathbb{R}[X_1, \ldots, X_n]/\langle X_i^2 - X_i \mid i \in [n]\rangle$, where $\langle X_i^2 - X_i \mid i \in [n]\rangle$ is the ideal generated by the $X_i^2 - X_i$’s. We will introduce some basic notation and facts of $\mathbb{R}[x]$ in Section 2.2.

The main reason for the simplicity of our proof lies in exploiting symmetries and applying a recent but unpublished result of Blekherman\footnote{A proof can be found in [LPWY16, Theorem B.11]. Note that Troy Lee et al. also gave a new proof of Grigoriev’s Theorem [LPWY16, Theorem C.1]. Completely independent from us, they also noticed that Blekherman’s result can be used to simplify it.}. Therefore, in Section 2.3 we have a closer view on symmetric polynomials in $\mathbb{R}[x]$.

We will give a more detailed introduction and motivation to the work of Grigoriev at the beginning of Chapter 2 and Section 2.1, where all relevant notions will also be defined and explained. Additionally, we present two applications of Grigoriev’s proof in Section 2.5.

Chapter 3 can be seen as a preparation for its subsequent chapters. An important component of the work of [LRS15] are Boolean functions, in particular so-called pseudo-densities and pseudo-expectations. Boolean functions are functions from $\{-1,1\}^n$ respectively $\{0,1\}^n$ to $\mathbb{R}$. We will provide some basic facts of them in the first three sections. These sections are mainly elementary and can be omitted by a proficient reader. In Section 3.3, we will especially look on symmetric Boolean functions, where we will mainly build on results from Section 2.3.

In Section 3.4, we will deduce the above mentioned corollary from Grigoriev’s Theorem. It shows the existence of a family of nonnegative Boolean functions $f_n : \{0,1\}^n \to \mathbb{R}_+$ of degree two that are not a sum of squares of Boolean functions of degree at most $n$. Finally, we will motivate and introduce the notion of pseudo-densities and pseudo-expectations in Section 3.5. We will explain the relation between them, and we will apply results from Section 3.3 in order to show how they behave if symmetries are existent.

Altogether, the proofs and explanations of this chapter will possibly seem simple and too detailed at some point, but we also want to present a plausible and understandable work for people without any knowledge of Boolean functions.
Chapter 5 consists of a more detailed analysis of the pseudo-densities that are relevant for the proof of [LRS15], which is the third main part of our work. Here, we will explain why we always can expect and exploit symmetry in our context. Under some assumptions, we will give a characterization of the relevant symmetric pseudo-densities, where the result of Blekherman from Chapter 2 is again of enormous help for our computations.

With our observations from this chapter, we will be able to state some optimality results for the lower bounds on the psd rank of the correlation polytopes obtained by [LRS15].

Note that we will provide more explanations, introductions and motivation at the beginning of each chapter.
1 Preliminaries - matrices and linear forms

As already mentioned in the introduction, in this chapter we just make some preparations for the further course of the work. Therefore, we introduce some notation and mention some notions and simple facts from linear algebra that we will use later. Some of the results will be proven, while the most simplest ones and the ones that are obvious are provided without a proof. For other stated well-known results, we refer to [HJ13] for a proof. Note that these results have not been originally proven by these authors, but the book is a standard work for matrix theory.

At some point, some statements might seem too simple to mention, but we decide to do it for reasons of convenience and for readers with fewer background knowledge.

While most parts of this chapter might be omitted by a proficient reader, mainly the notion of the different matrix norms are not in everyday use. Additionally, the last stated result, Proposition 1.32, is not well-known in every community as we noted. But especially this result is of great benefit for some parts of our work.

We start with some general notation, valid for the rest of this work.

**Notation 1.1.**
Let \( n \in \mathbb{N} \).

(a) We write \([n] := \{1, \ldots, n\}\).

(b) We write \(I_n\) for the identity matrix of size \(n\).

(c) By \(\mathbb{R}_+\) we denote the nonnegative real numbers.

### 1.1 The trace, symmetric matrices and matrix norms

In this work, the trace of a quadratic matrix \(A\), a basic concept in linear algebra and denoted by \(\text{Tr}(A)\), plays a substantial role. Although it is well-known, nevertheless we specify its definition.

Let \(r \in \mathbb{N}\) and \(A \in \mathbb{R}^{r \times r}\). Then, we have

\[
\text{Tr}(A) = \sum_{i \in [r]} A_{i,i}.
\]

Note that the trace is a linear map from the vector space \(\mathbb{R}^{r \times r}\) to \(\mathbb{R}\). Most of the time, we need the trace of matrix products. We state an easy and well-known but
Corollary 1.2. Let \( r, s \in \mathbb{N} \) and let \( A \in \mathbb{R}^{r \times s} \) and \( B \in \mathbb{R}^{s \times r} \). Then, we have
\[
\text{Tr}(AB) = \text{Tr}(BA).
\]

A second result immediately obtained by the above mentioned definitions is the following.

Corollary 1.3. Let \( r, s \in \mathbb{N} \). Then for every \( A, B \in \mathbb{R}^{r \times s} \), we have
\[
\text{Tr}(A^T B) = \sum_{i \in [s]} \sum_{j \in [r]} A_{i,j} B_{i,j}.
\]

Therefore, the term \( \text{Tr}(A^T B) \) is just the standard inner product of \( \mathbb{R}^{rs} \) if we identify \( \mathbb{R}^{r \times s} \) with \( \mathbb{R}^{rs} \). So, the inner product introduced below is actually an inner product on the vector space of real matrices.

Definition 1.4. Let \( r, s \in \mathbb{N} \). On the vector space \( \mathbb{R}^{r \times s} \) of real \( r \times s \)-matrices, we introduce an inner product \( \langle \cdot, \cdot \rangle \) by
\[
\langle A, B \rangle := \text{Tr}(A^T B).
\]

This inner product is called the Frobenius inner product. In the rest of this work, only this inner product is used in connection with matrices.

With this definition and the latter corollary, we conclude the following statement, which we will use in the further course of this work.

Corollary 1.5. Let \( r \in \mathbb{N} \). Then, for every square matrix \( A \in \mathbb{R}^{r \times r} \) and every \( x \in \mathbb{R}^r \), it holds
\[
\langle A, xx^T \rangle = \sum_{i,j \in [r]} A_{i,j} x_i x_j.
\]

\[\text{3}\]Sometimes, additionally to the usual norm axioms, a matrix norm is required to be submultiplicative, what is not important for us. In our case, a matrix norm is just a norm on the vector space of matrices. Nevertheless, the Frobenius norm actually is submultiplicative \[\text{[HJ13]}\] p. 342.\]
Definition 1.6.
Let \( r, s \in \mathbb{N} \) and \( A \in \mathbb{R}^{r \times s} \). We define

the Frobenius norm of \( A \) by \( \|A\|_F := \sqrt{\text{Tr}(A^T A)} \),

and

the infinity norm of \( A \) by \( \|A\|_\infty := \max_{1 \leq i \leq r, 1 \leq j \leq s} |A_{i,j}| \).

Remark 1.7.
Let \( M \) be a nonempty finite set and \( f \in \mathbb{R}^M \) a real-valued function. Then, analog to the above definition we define

\[ \|f\|_\infty := \max_{x \in M} |f(x)|, \]

and we will also call it the infinity norm of \( f \). This is reasonable and coincides with the above definition because \( \mathbb{R}^M \) is isomorphic to \( \mathbb{R}^{r \times 1} \) with \( r := |M| \).

Remark 1.8.
Let \( r, s \in \mathbb{N} \) and \( A \in \mathbb{R}^{r \times s} \). Note that \( \|A\|_F = \sqrt{\langle A, A \rangle} \) holds. So, \( \|\cdot\|_F \) is induced by the Frobenius inner product \( \langle \cdot, \cdot \rangle \), and thus \( \|\cdot\|_F \) is an inner product norm. Hence, the following equality holds for all \( A, B \in \mathbb{R}^{r \times s} \):

\[ \|A + B\|_F^2 = \|A\|_F^2 + \|B\|_F^2 + 2\langle A, B \rangle. \]

The previous definition together with Corollary 1.3 immediately implies the following statement.

Corollary 1.9.
Let \( r, s \in \mathbb{N} \). Then for every \( A \in \mathbb{R}^{r \times s} \), we have

\[ \|A\|_F^2 = \sum_{i \in [r], j \in [s]} A_{i,j}^2. \]

Symmetric matrices, matrices with the property \( A = A^T \), play a special role in linear algebra as well as in this work.

Notation 1.10.
Let \( r \in \mathbb{N} \). We denote by

\[ S^r := \{ A \in \mathbb{R}^{r \times r} \mid A = A^T \} \]

the vector space of all symmetric \( r \times r \)-matrices.
Their significance lies in the fact that they are always orthogonal diagonalizable and that they have only real eigenvalues. Thus, for every symmetric $A \in \mathbb{R}^{r \times r}$ with eigenvalues $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$ there exists $U \in \mathbb{R}^{r \times r}$ with $U^T = U^{-1}$ such that $A = U\Lambda U^{-1}$, where $\Lambda = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_r \end{pmatrix}$.

Therefore, we obtain
\[
\text{Tr}(A) = \sum_{i \in [r]} \lambda_i \tag{1.2}
\]
as a simple but useful and well-known consequence of Corollary 1.2.

Another conclusion from the fact that symmetric matrices are diagonalizable is that we can easily compute powers or more general polynomial expressions or their exponential.

**Definition 1.11.**
Let $r \in \mathbb{N}$ and $A \in \mathbb{R}^{r \times r}$ a quadratic matrix. Then, we define
\[
e^A := \sum_{k=0}^{\infty} \frac{1}{k!} A^k
\]
and call it the exponential of $A$.

**Remark 1.12.**
With the submultiplicativity of $\|\cdot\|_F$ (see footnote 3) we get
\[
\sum_{k=0}^{\infty} \frac{\|A^k\|_F}{k!} \leq \sum_{k=0}^{\infty} \frac{\|A\|^k}{k!} = e^{\|A\|_F} < \infty
\]
in the above expression. Therefore, the above series is absolute convergent and hence always convergent. Accordingly, the exponential of $A$ is well-defined.

**Lemma 1.13.**
Let $r \in \mathbb{N}$ and let $A \in \mathbb{R}^{r \times r}$ be symmetric with eigenvalues $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$. Furthermore, let $p = \sum_{k=0}^{m} \alpha_k x^k$ for some $m \in \mathbb{N}$ and $\alpha \in \mathbb{R}^m$ be any real, univariate polynomial. Then, $p(A)$ and $e^A$ are also symmetric with eigenvalues $p(\lambda_1), \ldots, p(\lambda_r)$ respectively $e^{\lambda_1}, \ldots, e^{\lambda_r}$.

**Proof.** Since $A$ is symmetric, there exist $U, \Lambda \in \mathbb{R}^{r \times r}$ as above. Note that we have
\[
A^k = (U\Lambda U^{-1})^k = U\Lambda^k U^{-1} \quad \text{and} \quad \Lambda^k = \begin{pmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_r^k \end{pmatrix}
\]
for every $k \in \mathbb{N}$.

Therefore, we obtain

$$p(A) = \sum_{k=0}^{m} \alpha_k A^k = \sum_{k=0}^{m} \alpha_k U^k U^{-1} = U \left( \sum_{k=0}^{m} \alpha_k \Lambda^k \right) U^{-1}$$

$$= U \begin{pmatrix} \sum_{k=0}^{m} \alpha_k \lambda_1^k & 0 & \cdots & 0 \\
0 & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & 0 \\
0 & \cdots & 0 & \sum_{k=1}^{m} \alpha_k \lambda_r^k \end{pmatrix} U^{-1}$$

$$= U \begin{pmatrix} p(\lambda_1) & 0 & \cdots & 0 \\
0 & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & 0 \\
0 & \cdots & 0 & p(\lambda_r) \end{pmatrix} U^{-1}$$

and analog

$$e^A = U \begin{pmatrix} e^{\lambda_1} & 0 & \cdots & 0 \\
0 & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & 0 \\
0 & \cdots & 0 & e^{\lambda_r} \end{pmatrix} U^{-1}.$$  

The symmetry follows from $U^{-1} = U^T$.  

Another identity we will sometimes use is the following.

**Corollary 1.14.**

Let $r, s \in \mathbb{N}$, and let $A \in \mathbb{S}^r$ and $B \in \mathbb{R}^{r \times s}$. Then, it holds

$$\|AB\|_F^2 = \text{Tr} \left( A^2 BB^T \right).$$

**Proof.**

$$\|AB\|_F^2 = \text{Tr} \left( (AB)^T (AB) \right) = \text{Tr} \left( B^T A A^T B \right) \text{Cor}_2 \text{Tr} \left( A^2 BB^T \right).$$

For symmetric matrices, we introduce two further types of matrix norms.

**Definition 1.15.**

Let $r \in \mathbb{N}$ and $A \in \mathbb{S}^r$ with eigenvalues $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$. We define

the _operator norm_ of $A$ by $$\|A\| := \max_{i \in [r]} |\lambda_i|,$$

and

the _trace norm_ of $A$ by $$\|A\|_* := \sum_{i \in [r]} |\lambda_i|.$$
Remark 1.16.
Let $r \in \mathbb{N}$ and $A \in \mathbb{R}^{r \times r}$.

(a) Usually, the operator norm is defined by
\[ \|A\| := \max_{\|x\|_2 = 1} \|Ax\|_2, \]
where $\|\cdot\|_2$ is the Euclidean norm on $\mathbb{R}^r$. But for symmetric matrices, both definitions coincide (cf. [HJ13, p. 346]). So, the operator norm is indeed a matrix norm\(^4\) (cf. [HJ13, Theorem 5.6.2]).

(b) As well, the trace norm is indeed a matrix norm\(^4\) (cf. [HJ13, Theorem 5.6.42]). The name of this norm is motivated in the next subsection.

(c) Like the operator norm and the trace norm, also the Frobenius norm of a symmetric matrix $A \in \mathbb{S}^r$ is representable by the eigenvalues $\lambda_1, \ldots, \lambda_r$ of $A$. It holds
\[ \|A\|_F = \sqrt{\sum_{i=1}^{r} \lambda_i^2}. \]
Comparing (1.2) with the definition of the trace norm immediately yields the simple fact
\[ \text{Tr}(A) \leq \|A\|_*, \quad (1.3) \]
for every symmetric $A \in \mathbb{R}^{r \times r}$.

Definition 1.17.
Let $r \in \mathbb{R}$. We say two matrices $A, B \in \mathbb{R}^{r \times r}$ are simultaneously diagonalizable if there exists an invertible matrix $U \in \mathbb{R}^{r \times r}$ such that $UAU^{-1}$ and $UBU^{-1}$ are both diagonal matrices.

From the proof of Lemma 1.13, we immediately obtain the following fact.

Corollary 1.18.
For every symmetric matrix $A \in \mathbb{R}^{r \times r}$ and every real, univariate polynomial $p$, the matrices $A, p(A)$ and $e^A$ are pairwise simultaneously diagonalizable.

The following statement yields an inequality involving matrix norms.

Lemma 1.19.
Let $r \in \mathbb{N}$ and let $A, B \in \mathbb{S}^r$ be simultaneously diagonalizable. Then, it holds
\[ \|AB\|_* \leq \|A\| \|B\|_. \]

\(^4\)Moreover, the operator norm as well as the trace norm are even submultiplicative (cf. [HJ13, Theorem 5.6.2, Theorem 5.6.42]).
Proof. Let \( \lambda_1, \ldots, \lambda_r \) be the eigenvalues of \( A \) and \( \gamma_1, \ldots, \gamma_r \) the eigenvalues of \( B \). By the assumption, there exists a matrix \( U \in \mathbb{R}^{r \times r} \) with \( U^{-1} = U^T \) such that

\[
A = U \begin{pmatrix}
\lambda_1 & 0 & \cdots & 0 \\
0 & \ddots & & \\
\vdots & & \ddots & 0 \\
0 & \cdots & 0 & \lambda_r
\end{pmatrix} U^{-1} \quad \text{and} \quad B = U \begin{pmatrix}
\gamma_1 & 0 & \cdots & 0 \\
0 & \ddots & & \\
\vdots & & \ddots & 0 \\
0 & \cdots & 0 & \gamma_r
\end{pmatrix} U^{-1},
\]

which implies

\[
AB = U \begin{pmatrix}
\lambda_1 \gamma_1 & 0 & \cdots & 0 \\
0 & \ddots & & \\
\vdots & & \ddots & 0 \\
0 & \cdots & 0 & \lambda_r \gamma_r
\end{pmatrix} U^{-1}.
\]

Because of \( U^{-1} = U^T \), \( AB \) is symmetric, and its eigenvalues are \( \lambda_1 \gamma_1, \ldots, \lambda_r \gamma_r \). By definition, this again implies

\[
\|AB\|_* = \sum_{i \in [r]} |\lambda_i \gamma_i| = \sum_{i \in [r]} |\lambda_i| |\gamma_i| 
\leq \max_{i \in [r]} |\lambda_i| \sum_{i \in [r]} |\gamma_i| = \|A\| \|B\|_*.
\]

\( \square \)

1.2 Positive semidefinite matrices

Definition 1.20.

Let \( r \in \mathbb{N} \). A symmetric matrix \( A \in \mathbb{R}^{r \times r} \) is called \emph{positive semidefinite} (psd) if and only if

\[
v^T Av \geq 0 \quad \text{for all} \quad x \in \mathbb{R}^r \quad \text{with} \quad x \neq 0.
\]

We use \( \succeq \) for the standard Loewner order on \( \mathbb{S}^r \) which means

\[
A \succeq 0 \quad \text{if and only if} \quad A \text{ is positive semidefinite.}
\]

For two matrices \( A, B \in \mathbb{S}^r \) we say \( A \succeq B \) if \( A - B \succeq 0 \).

Remark 1.21.

Note the following fact that easily follows from the above definition and that we will use in the further course of the work without saying anything.

Let \( s, t \in \mathbb{N} \), \( A \in \mathbb{S}^s \) and \( B \in \mathbb{S}^t \). Furthermore, let \( r := s + t \) and \( M \in \mathbb{S}^r \) defined by \( M := \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \). Then, \( M \) is psd if and only if \( A, B \) are psd.

Directly connected to the notion of psd matrices is the notion of a spectrahedron. We remind its definition from the introduction.
Definition 1.22.
Let \( n \in \mathbb{N} \). A set \( S \subseteq \mathbb{R}^n \) is called a *spectrahedron* if there exist some \( k \in \mathbb{N} \) and some symmetric matrices \( A_0, \ldots, A_n \in \mathbb{R}^{k \times k} \) such that
\[
S = \{ x \in \mathbb{R}^n \mid A_0 + x_1 A_1 \ldots x_n A_n \succeq 0 \}.
\]

The expression
\[
A_0 + x_1 A_1 \ldots x_n A_n \succeq 0
\]
is called a *linear matrix inequality* (lmi) and \( k \) is the *size* of the lmi.

Remark 1.23.

(a) With the above remark, we immediately see that the intersection of two spectrahedra is again a spectrahedron. In particular, by choosing \( s = t = 1 \) in the above remark and applying it iteratively, we see that every polyhedron is a spectrahedron.

As well, we see that a spectrahedron remains a spectrahedron if we add some linear inequalities and also linear equalities because they can be expressed by two inequalities.

(b) Let \( n \in \mathbb{N} \) and let \( S \subseteq \mathbb{R}^n \) be a spectrahedron. That means there exist \( m \in \mathbb{N} \) and \( A_0, \ldots, A_n \in \mathbb{S}^k \) with \( S = \{ x \in \mathbb{R}^n \mid A(x) := A_0 + x_1 A_1 \ldots x_n A_n \succeq 0 \} \).

Note that by definition we have
\[
S = \{ x \in \mathbb{R}^n \mid v^T A(x) v \geq 0 \text{ for all } v \in \mathbb{R}^m \}.
\]

Therefore, every spectrahedron is the intersection of infinitely many closed half-spaces. Consequently, a spectrahedron is always a closed, convex set.

We remind the following well-known characterization and the subsequent useful property of psd matrices.

**Proposition 1.24** (cf. [HJ13, Observation 7.1.4]).
A symmetric matrix is positive semidefinite if and only if all its eigenvalues are nonnegative.

**Proposition 1.25** (cf. [HJ13, Corollary 7.1.5]).
Let \( r \in \mathbb{N} \), and let \( A \in \mathbb{S}^r \) be positive semidefinite. Then, it holds
\[
\det(A) \geq 0.
\]

The notion of psd matrices is sufficient for most of the work. Only once, in Section 4.4, we also need the notion of positive definiteness.
**Definition 1.26.**
Let \( r \in \mathbb{N} \) and let \( A \in S^r \) be positive semidefinite with eigenvalues \( \lambda_1, \ldots, \lambda_r \geq 0 \). We call \( A \) **positive definite** and denote it by
\[
A > 0
\]
if all its eigenvalues are even positive, that means \( \lambda_1, \ldots, \lambda_r > 0 \).

Like positive real numbers, psd matrices have a unique positive square root in the following sense.

**Proposition/Definition 1.27** (cf. \([HJ13, \text{Theorem 7.2.6}]\)).
Let \( r \in \mathbb{N} \) and let \( A \in S^r \) be positive semidefinite. Then, there is a unique psd matrix \( B \in S^r \) with
\[
B^2 = A.
\]
We denote this matrix by \( \sqrt{A} \) and call it the psd square root of \( A \).

**Remark 1.28.**
Let \( r,s \in \mathbb{N} \). For a general matrix \( A \in \mathbb{R}^{r \times s} \) the trace norm is defined by
\[
\|A\|_* := \text{Tr} \left( \sqrt{A^T A} \right),
\]
which explains its name. This is well-defined because \( A^T A \) is positive semidefinite and it is not hard to see that it coincides with our definition in the case of symmetric matrices.

We proceed with a further inequality involving matrix norms, that we will use once in the course of this work. In order to derive it, we first provide a result, which follows as a special case of von Neumann’s trace inequality.

**Lemma 1.29** (cf. \([HJ13, \text{Theorem 8.7.6}]\)).
Let \( r \in \mathbb{N} \), and let \( A,B \in S^r \). Let \( \lambda_1 \leq \ldots \leq \lambda_r \) be the eigenvalues of \( A \), and let \( \gamma_1 \leq \ldots \leq \gamma_r \) be the eigenvalues of \( B \). Then, we have
\[
\text{Tr}(AB) \leq \sum_{i \in [r]} |\lambda_i| |\gamma_i|.
\]

**Corollary 1.30.**
Let \( r,s,t \in \mathbb{N} \), and let \( A \in \mathbb{R}^{r \times s} \), \( B \in \mathbb{R}^{s \times t} \). Then, it holds
\[
\|AB\|_F^2 \leq \|A^T A\| \cdot \|B\|_F^2.
\]
Proof. Note that $A^T A \in S^s$ as well as $BB^T \in S^s$ are symmetric, moreover positive semidefinite, for that reason we can apply the previous lemma. In order to do this, let $0 \leq \lambda_1 \leq \ldots \leq \lambda_s$ be the eigenvalues of $A^T A$, and let $0 \leq \gamma_1 \leq \ldots \leq \gamma_s$ be the eigenvalues of $B$. Then, we obtain
\[
\|AB\|_F^2 = \text{Tr}((AB)^T(AB)) = \text{Tr}(B^T A^T AB) \leq \sum_{i \in [s]} |\lambda_i| |\gamma_i|
\]
\[
= |\lambda_s| \sum_{i \in [s]} \gamma_i = \|A^T A\| \text{Tr}(BB^T)
\]
\[
= \|A^T A\| \|B\|_F^2.
\]

Next, we prove two results dealing with linear forms. Although they are not hard to prove, these results will be very useful in our work. In particular, the second one will turn out to be very powerful in our setting of Section 2.5. But we also make frequent use of it in Chapter 3.

Lemma 1.31.
Let $r \in \mathbb{N}$, and let $R$ be an algebra over $\mathbb{R}$. Furthermore, let $f_1, \ldots, f_r, g \in R$, and $L : R \to \mathbb{R}$ be a linear form. If $L \left( g \cdot \left( \sum_{i \in [r]} a_i f_i \right)^2 \right) \geq 0$ for all $a \in \mathbb{R}^r$, then it holds
\[
\begin{pmatrix}
L(gf_1, f_1) & \cdots & L(gf_1, f_r) \\
\vdots & \ddots & \vdots \\
L(gf_r, f_1) & \cdots & L(gf_r, f_r)
\end{pmatrix} \succeq 0.
\]

Proof. For every $a \in \mathbb{R}^r$ and all $f_1, \ldots, f_r \in R$ we have
\[
L \left( g \cdot \left( \sum_{i \in [r]} a_i f_i \right)^2 \right) = L \left( g \sum_{i,j \in [r]} a_i a_j f_i f_j \right)
\]
\[
= \sum_{i,j \in [r]} a_i a_j L(gf_i f_j)
\]
\[
= \sum_{i \in [r]} a_i \sum_{j \in [r]} L(g f_j f_i) a_j
\]
\[
= \begin{pmatrix} a_1 & \cdots & a_r \end{pmatrix} \begin{pmatrix}
L(gf_1, f_1) & \cdots & L(gf_1, f_r) \\
\vdots & \ddots & \vdots \\
L(gf_r, f_1) & \cdots & L(gf_r, f_r)
\end{pmatrix} \begin{pmatrix} a_1 \\
\vdots \\
a_r \end{pmatrix}.
\]
Hence, if \( L \left( g \cdot \left( \sum_{i \in [r]} a_i f_i \right)^2 \right) \geq 0 \) for all \( a \in \mathbb{R}^r \), then
\[
\begin{pmatrix}
L(gf_1 f_1) & \cdots & L(gf_1 f_r) \\
\vdots & \ddots & \vdots \\
L(gf_r f_1) & \cdots & L(gf_r f_r)
\end{pmatrix} \succeq 0
\]
by definition.

\[\square\]

**Proposition 1.32.**

Let \( R \) be an algebra over \( \mathbb{R} \) and \( L : R \to \mathbb{R} \) a linear form. Furthermore, let \( f \in R \) with
\[L(f^2) = 0,
\]
and let \( U \) be a subspace (not necessary a subalgebra) of \( R \) with \( f \in U \) and the additional property
\[L(g^2) \geq 0 \text{ for all } g \in U.
\]
Then, it holds
\[L(fg) = 0 \text{ for all } g \in U.
\]

In particular, if \( 1 \in U \), we obtain
\[L(f) = 0.
\]

**Proof.** Since \( U \) is a subspace, we have \( \alpha f + \beta g \in U \) for all \( \alpha, \beta \in \mathbb{R} \) and all \( g \in R \). Therefore, because of \( L(h^2) \geq 0 \) for all \( h \in U \), it follows
\[L((\alpha f + \beta g)^2) \geq 0
\]
for all \( \alpha, \beta \in \mathbb{R} \) and all \( g \in R \). This yields
\[
\begin{pmatrix}
L(f^2) & L(fg) \\
L(fg) & L(g^2)
\end{pmatrix} \succeq 0
\]
for all \( g \in R \) with the above lemma. By Proposition 1.25, we have
\[L(f^2)L(g^2) - L(fg)^2 \geq 0.
\]
Because of \( L(f^2) = 0 \), this implies
\[-L(fg)^2 \geq 0,
\]
and therefore
\[L(fg)^2 \leq 0
\]
for all \( g \in R \). Since a square is nonnegative, the statement follows. \[\square\]

We finish this chapter with some common notation concerning the ring of real polynomials.
Notation 1.33.

Let \( n \in \mathbb{N} \).

(a) We denote by \( \mathbb{R} [X_1, \ldots, X_n] \) the ring of real polynomials in \( n \) variables. Rarely, we write \( \mathbb{R} [X] := \mathbb{R} [X_1, X_2, \ldots, X_n] \) in order to save space and only if it is clear which \( n \) is meant.

(b) For \( d \in \mathbb{N} \), we define \( \mathbb{R} [X]^d := \{ f \in \mathbb{R} [X_1, \ldots, X_n] \mid \deg(f) \leq d \} \) the \( d \)-truncation of \( \mathbb{R} [X_1, \ldots, X_n] \), that means the vector space of all real polynomials in \( n \) variables with degree at most \( d \).

(c) By \( \sum \mathbb{R} [X]^2 := \left\{ f \in \mathbb{R} [X_1, \ldots, X_n] \mid f = \sum_{i=1}^{s} f_i^2 \text{ for some } s \in \mathbb{N}, f_1, \ldots, f_s \in \mathbb{R} [X] \right\}, \)
we denote the cone of sums of squares of real polynomials in \( n \) variables. For \( d \in \mathbb{N} \), we write \( \sum \mathbb{R} [X]^2_d := \left\{ f \in \mathbb{R} [X_1, \ldots, X_n] \mid f = \sum_{i=1}^{s} f_i^2 \text{ for some } s \in \mathbb{N}, f_1, \ldots, f_s \in \mathbb{R} [X]^d \right\} \).

Note that we equivalently can write \( \sum \mathbb{R} [X]^2 := \sum \mathbb{R} [X]^2 \cap \mathbb{R} [X]_{2d} \).

(d) In contrast, in order to avoid confusion we denote by \( \mathbb{R} [T] \) the ring of univariate, real polynomials.

(e) We denote by \( \mathbb{R} [T]^n := \{ p \in \mathbb{R} [T] \mid \deg(p) \leq n \} \)
the vector space of univariate, real polynomials of degree at most \( n \).

(f) By \( \sum \mathbb{R} [T]^2 := \left\{ f \in \mathbb{R} [T] \mid f = \sum_{i=1}^{s} f_i^2 \text{ for some } s \in \mathbb{N}, f_1, \ldots, f_s \in \mathbb{R} [T] \right\}, \)
we denote the cone of sums of squares of univariate, real polynomials of degree at most \( n \).

Note that \( \sum \mathbb{R} [T]^2_n \subseteq \mathbb{R} [T]_{2n} \) holds.
2 Lower degree bounds for the Real Nullstellensatz

In this chapter, our goal is to provide a new and easier proof of a result of Grigoriev [Gri01a], where he showed linear lower degree bounds for the Real Nullstellensatz. It turns out that this result is essential for the work of [LRS15].

Together with a second work of Grigoriev from the same year [Gri01b], these bounds were the first known lower degree bounds for the Real Nullstellensatz in the Boolean case, which is the interesting case for applications in combinatorial optimization. We introduce the Real Nullstellensatz and some small motivation as well as the notion of Boolean and other appearing notions in the subsequent Section 2.1.

But Grigoriev’s result is not only important for the work of [LRS15]. It also plays an important role for questions concerning the power of algebraic certificates relying on sums of squares and also for the complexity of algorithms relying on sums of squares, which occurs frequently in combinatorial optimization.

Therefore, Grigoriev’s work is often cited by various people, but the proof is not quite understood by most of them. So, we consider it important to have an easier proof that is known by more people than the one in the original work.

This is done in Section 2.4. In order to do this, we will mainly work in the ring \( \mathbb{R}[\underline{x}] := \mathbb{R}[X_1, \ldots, X_n] / \langle X_i^2 - X_i \mid i \in [n] \rangle \), where \( \langle X_i^2 - X_i \mid i \in [n] \rangle \) is the ideal generated by the \( X_i^2 - X_i \)'s. In Section 2.2 we will formally introduce this notion together with some basic properties of this ring.

The main reason for the simplicity of our new proof is exploiting the symmetry of the problem and applying a recent theorem of Blekherman for symmetric sums of squares in \( \mathbb{R}[\underline{x}] \). Therefore, in Section 2.3 we will have a closer view on symmetric polynomials in \( \mathbb{R}[\underline{x}] \) and their properties.

In Section 2.5 we present two conclusions of the proof of Grigoriev’s theorem.

2.1 Real Nullstellensatz refutation

Consider the polynomial equation system

\[
 f_1 = 0, \ldots, f_s = 0.
\]

for some \( s \in \mathbb{N} \) and given \( f_1, \ldots, f_s \in \mathbb{R}[X_1, \ldots, X_n] \). For \( s \geq n \), such a system is called Boolean if the set \( \{f_1, \ldots, f_s\} \) includes the set

\[
 \{X_i^2 - X_i \mid i \in [n]\} \text{ or } \{X_i^2 - 1 \mid i \in [n]\}
\]

respectively.
Furthermore, we denote by
\[ V(f_1, \ldots, f_s) := \{ x \in \mathbb{R}^n \mid f_1(x) = \ldots = f_s(x) = 0 \} \]
the set of common real zeros of \( f_1, \ldots, f_s \).

In general, it is hard to decide whether such a system (2.1) is feasible over the reals or not. Therefore, it is desirable to have a preferably easy but sufficient powerful certificate for feasibility or infeasibility respectively.

Because real squares are always nonnegative, the following representation
\[ 1 + \sum_{i \in [t]} h_i^2 = \sum_{i \in [s]} f_i g_i \tag{2.2} \]
for some \( t \in \mathbb{N} \) and \( h_1, \ldots, h_t, g_1, \ldots, g_s \in \mathbb{R}[X_1, \ldots, X_n] \) refutes the existence of a common real zero of the \( f_i \)'s, and therefore it certifies \( V(f_1, \ldots, f_s) = \emptyset \).

We formally define this notion of refutation.

**Definition 2.1.**
Let \( n, s \in \mathbb{N} \) and \( f_1, \ldots, f_s \in \mathbb{R}[X_1, \ldots, X_n] \). We say the polynomial equation system (2.1) has a **Real Nullstellensatz (RNS) refutation** if there exist \( t \in \mathbb{N} \) and \( g_1, \ldots, g_s, h_1, \ldots, h_t \in \mathbb{R}[X_1, \ldots, X_n] \) such that (2.2) holds.

The name of this refutation refers to the Real Nullstellensatz, first proven by Krivine [Kri64], which can be seen as a starting point of modern real algebra.

**Theorem 2.2** (Real Nullstellensatz, [Kri64]).
Let \( n, s \in \mathbb{N} \) and \( f, f_1, \ldots, f_s \in \mathbb{R}[X_1, \ldots, X_n] \). Then, it holds
\[ f(x) = 0 \text{ for every } x \in V(f_1, \ldots, f_s) \]
if and only if
\[ f^{2k} + \sum_{i \in [t]} h_i^2 = \sum_{i \in [s]} f_i g_i \]
for some \( k, t \in \mathbb{N} \) and \( h_1, \ldots, h_t, g_1, \ldots, g_s \in \mathbb{R}[X_1, \ldots, X_n] \).

In particular, we have \( V(f_1, \ldots, f_r) = \emptyset \) if and only if
\[ 1 + \sum_{i \in [t]} h_i^2 = \sum_{i \in [s]} f_i g_i \]
for some \( r \in \mathbb{N} \) and \( h_1, \ldots, h_t, g_1, \ldots, g_s \in \mathbb{R}[X_1, \ldots, X_n] \).

**Remark 2.3.**
Note that in [Gri01a] and in other work related to Grigoriev’s theorem (for example [Gri01b], [GV01], [LPWY16]), they usually talk about Positivstellensatz refutation relying on the Positivstellensatz usually referred to Stengle [Ste74]. Of course, that is not wrong because the Real Nullstellensatz can be seen as a specialization of
the Positivstellensatz. But on the one hand, we do not use inequalities here so we actually use the Real Nullstellensatz, which has its own eligibility in Real Algebra. On the other hand, even the Positivstellensatz was proven by Krivine in the same work [Kri64], but this work remained unknown for a while and was rediscovered later by several people, amongst them was Stengle.

By the Real Nullstellensatz, we know that every polynomial equation system (2.1) with \( V(f_1, \ldots, f_s) = \emptyset \) always has some RNS refutation, but nothing is said about the degree that is needed for the polynomials occurring in the refutation.

But this is the most interesting point not only for applications and if we want to compute a refutation efficiently, but also in order to say something about the power and the complexity of the relying Real Nullstellensatz and more general algebraic certificates relying on sums of squares.

Especially for the second point, to know something about lower degree bounds is essential. Therefore, the following definition is reasonable.

**Definition 2.4.**
Let \( n, s \in \mathbb{N} \) and \( f_1, \ldots, f_s \in \mathbb{R}[x] \). Consider the polynomial equation system (2.1) and a RNS refutation (2.2) of this system. Then, the *degree* of the refutation is defined as

\[
\max_{i \in [s], j \in [t]} \{ \deg(f_i g_i), \deg(h_j) \}.
\]

Exponential lower bounds on the degree of any RNS refutation for a certain polynomial equation system are proven in [GV01] and even earlier, but only known in the real algebra community, in [Sch98], where both use almost the same polynomials. Indeed, both works have in common that their systems are not Boolean, but this is the interesting case for applications in combinatorial optimization.

In 2001, Grigoriev proved the first known lower degree bounds in the Boolean case [Gri01a, Gri01b]. We will give a new proof of his lower bounds from [Gri01a], where the following Boolean system, that is obviously infeasible, is considered:

\[
f_i = X_i^2 - X_i = 0, \ 1 \leq i \leq n, \ f_{n+1} = \sum_{i \in [n]} X_i - r = 0,
\]

where \( r \) is a non-integer real number with \( 0 < r < n \).

### 2.2 Multilinear polynomials

**Notation 2.5.**
Let \( n \in \mathbb{N} \).
(a) We denote by
\[ \mathcal{J} := \langle X_i^2 - X_i \mid i \in [n] \rangle \]
the ideal of \( \mathbb{R}[X_1, \ldots, X_n] \) generated by the set \( \{ X_i^2 - X_i \mid i \in [n] \} \), and we denote by
\[ \mathbb{R}[x_1, \ldots, x_n] := \mathbb{R}[X_1, \ldots, X_n] / \mathcal{J} \]
the quotient ring of \( \mathbb{R}[X_1, \ldots, X_n] \) modulo \( \mathcal{J} \), where \( x_i := X_i := X_i + \mathcal{J} \).

(b) As quite common, for \( f \in \mathbb{R}[X_1, \ldots, X_n] \) we sometimes write
\[ \bar{f} := f(x_1, \ldots, x_n) \in \mathbb{R}[x_1, \ldots, x_n]. \]

(c) In order to save space and only if it is clear which \( n \in \mathbb{N} \) is meant, we sometimes write
\[ \mathbb{R}[x] := \mathbb{R}[x_1, \ldots, x_n]. \]

(d) Furthermore, for \( I \subseteq [n] \) we write
\[ x^I := \prod_{i \in I} x_i. \]

By definition, we have \( x_i^2 = x_i \) for every \( i \in [n] \), and therefore, we have
\[ x_i^k = x_i \quad \text{for every } k \in \mathbb{N} \text{ and every } i \in [n] \quad (2.3) \]
and
\[ x^I \cdot x^J = x^{I \cup J} \quad \text{for every } I, J \subseteq [n]. \quad (2.4) \]
Because of (2.3), it is obvious that for every \( p \in \mathbb{R}[x_1, \ldots, x_n] \) there exists a family \((\alpha_I)_{I \subseteq [n]}\) of real numbers such that
\[ p = \sum_{I \subseteq [n]} \alpha_I x^I. \quad (2.5) \]

Next, we want to show that such a representation is unique, and for that we first need some preparation.

**Lemma 2.6.**
Let \( n \in \mathbb{N} \) and let \((\alpha_I)_{I \subseteq [n]}\) be a family of real numbers with
\[ \sum_{I \subseteq [n]} \alpha_I \prod_{i \in I} x_i = 0 \]
for every \( x \in \{0,1\}^n \). Then, it holds
\[ \alpha_I = 0 \text{ for every } I \subseteq [n]. \]
Proof. Assume there exists a subset $I \subseteq [n]$ with $\alpha_I \neq 0$. Then, among all subsets satisfying this property, we can find a $J$ which is minimal with respect to the cardinality. We fix such a minimal $J \subseteq [n]$ and consider $y \in \{0,1\}^n$ with

$$y_i := \begin{cases} 1 & i \in J \\ 0 & \text{otherwise}. \end{cases}$$

Note that this definition implies

$$\prod_{i \in I} y_i = \begin{cases} 1 & I \subseteq J \\ 0 & \text{otherwise}. \end{cases}$$

Then, we have

$$0 = \sum_{I \subseteq [n]} \alpha_I \prod_{i \in I} y_i = \sum_{I \subseteq J} \alpha_I = \alpha_J,$$

where we used the minimality of $J$. But $\alpha_J = 0$ is a contradiction to the choice of $J$.

\[\square\]

Corollary 2.7.

Let $n \in \mathbb{N}$, and let $f = \sum_{I \subseteq [n]} \alpha_I \prod_{i \in I} X_i \in \mathbb{R}[X_1, \ldots, X_n]$, where $(\alpha_I)_{I \subseteq [n]}$ is a family of real numbers. Furthermore, let $f \in \mathcal{J}$. Then, it holds $f = 0$.

Proof. Since $f \in \mathcal{J}$, it obviously follows $f(x) = 0$ for every $x \in \{0,1\}^n$. But then, the coefficients $(\alpha_I)_{I \subseteq [n]}$ of $f$ satisfy the assumption of the previous lemma, and therefore we have

$$\alpha_I = 0 \text{ for every } I \subseteq [n],$$

which implies the desired statement. \[\square\]

Corollary 2.8.

Let $n \in \mathbb{N}$, $p \in \mathbb{R}[x_1, \ldots, x_n]$, and let $(\alpha_I)_{I \subseteq [n]}$, $(\beta_I)_{I \subseteq [n]}$ be two families of real numbers with

$$\sum_{I \subseteq [n]} \alpha_I x^I = p = \sum_{I \subseteq [n]} \beta_I x^I.$$

Then, it holds

$$\alpha_I = \beta_I \text{ for every } I \subseteq [n].$$

Proof. Consider $f \in \mathbb{R}[X_1, \ldots, X_n]$ with $f := \sum_{I \subseteq [n]} (\alpha_I - \beta_I) \prod_{i \in I} X_i$. With this definition, we obtain

$$\bar{f} = \sum_{I \subseteq [n]} (\alpha_I - \beta_I) x^I = \sum_{I \subseteq [n]} \alpha_I x^I - \sum_{I \subseteq [n]} \beta_I x^I = 0,$$

where the last step follows by the premise. But this implies $f \in \mathcal{J}$, and applying the previous corollary yields $f = 0$. This implies

$$\alpha_I - \beta_I = 0 \text{ for every } I \subseteq [n],$$

which implies the statement. \[\square\]
Together with \([2.5]\), the previous corollary implies that \(\{x^I\}_{I \subseteq [n]}\) forms a basis of the real vector space \(\mathbb{R}[x_1, \ldots, x_n]\). Therefore, for every \(p \in \mathbb{R}[x_1, \ldots, x_n]\) there exists a unique family \((\alpha_I)_{I \subseteq [n]}\) of real numbers such that

\[
p = \sum_{I \subseteq [n]} \alpha_I x^I.
\]

This representation motivates the notion of multilinear polynomials, which we sometimes will use for elements of \(\mathbb{R}[x_1, \ldots, x_n]\).

Even without saying it, we will only use this unique representation as multilinear polynomials for elements of \(\mathbb{R}[x_1, \ldots, x_n]\) in the further course of this work.

The above representation also motivates the following definition.

**Definition 2.9.**

Let \(n \in \mathbb{N}\), and let \(f \in \mathbb{R}[x_1, \ldots, x_n]\) with \(f = \sum_{I \subseteq [n]} \alpha_I x^I\). Then, the degree of \(f\) is denoted by \(\deg(f)\), and it is defined by

\[
\deg(f) := \begin{cases} 
\max \{|I| \mid I \subseteq [n], \alpha_I \neq 0\} & \text{if } f \neq 0 \\
-\infty & \text{if } f = 0.
\end{cases}
\]

Furthermore, for \(k \in [n]\) we denote the degree-\(k\) truncation of \(\mathbb{R}[x_1, \ldots, x_n]\) by

\[
\mathbb{R}[x]^k := \left\{ p \in \mathbb{R}[x_1, \ldots, x_n] \mid \deg(p) \leq k \right\} = \left\{ \sum_{I \subseteq [n]} \alpha_I x^I \in \mathbb{R}[x] \mid \alpha_I = 0 \text{ for all } I \subseteq [n] \text{ with } |I| > k \right\}.
\]

### 2.3 Symmetric multilinear polynomials

Symmetry of polynomials and functions will play an important role not only in this section, but also in other parts of the following work. So, we want to formally define the notion of symmetry in \(\mathbb{R}[x]\) and want to derive some nice results and properties of symmetric, multilinear polynomials.

But first, we go back to the ordinary polynomial ring \(\mathbb{R}[X_1, \ldots, X_n]\). Remember the notion of symmetry in this ring.

**Definition 2.10.**

Let \(n \in \mathbb{N}\), and as usual we denote by \(S_n\) the symmetric group on \([n]\). A polynomial \(p \in \mathbb{R}[X_1, \ldots, X_n]\) is called symmetric if

\[
p(X_{\sigma(1)}, \ldots, X_{\sigma(n)}) = p(X_1, \ldots, X_n) \text{ for every } \sigma \in S_n.
\]

By

\[
\mathbb{R}[X_1, \ldots, X_n]^{S_n} := \{ p \in \mathbb{R}[X_1, \ldots, X_n] \mid p \text{ symmetric} \}
\]

we denote the ring of all symmetric polynomials in \(\mathbb{R}[X_1, \ldots, X_n]\).
Next, we introduce two special types of symmetric polynomials.

**Definition 2.11.**

Let $n \in \mathbb{N}$. For $k \in \{0, \ldots, n\}$, let

$$e_{n,k} := \sum_{I \subseteq \{1, \ldots, n\}} \prod_{i \in I} X_i \in \mathbb{R} [X_1, \ldots, X_n]^{S_n}$$

be the *elementary symmetric polynomials* in $n$ variables, and for $k \in [n]$ let

$$p_{n,k} := \sum_{i=1}^{n} X_i^k \in \mathbb{R} [X_1, \ldots, X_n]^{S_n}$$

be the *power sum symmetric polynomials* in $n$ variables.

**Remark 2.12.**

Let $n \in \mathbb{N}$ and let $k \in \{0, \ldots, n\}$. Then, we have

$$e_{n,k} = \sum_{I \subseteq \{1, \ldots, n\}} \prod_{i \in I} X_i = \frac{n!}{k!} \sum_{\sigma \in S_n} \prod_{i=1}^{k} X_{\sigma(i)} = \frac{1}{k!(n-k)!} \sum_{\sigma \in S_n} \prod_{i=1}^{k} X_{\sigma(i)}.$$

There is the following relation between these two types of symmetric polynomials, also known as *Newton’s identities*. A proof of it can be found for example in [Mea92].

**Theorem 2.13** (Newton’s identities).

Let $n \in \mathbb{N}$. Then, it holds

$$k \cdot e_{n,k} = \sum_{i=1}^{k} (-1)^{i-1} e_{n,k-i} p_{n,i}$$

for every $k \in [n]$.

Moreover, the elementary symmetric polynomials play a special role because they generate the algebra of symmetric polynomials. This is a well-known result in algebra, called the *fundamental theorem of symmetric polynomials*. A proof can be found for example in [Bos06 4.3/Satz 5].

**Theorem 2.14** (The fundamental theorem of symmetric polynomials).

Let $n \in \mathbb{N}$. Then, we have

$$\mathbb{R} [X_1, \ldots, X_n]^{S_n} = \mathbb{R} [e_{n,1}, \ldots, e_{n,n}],$$

and $e_{n,1}, \ldots, e_{n,n}$ are algebraically independent over $\mathbb{R}$.

In particular, for every symmetric $p \in \mathbb{R} [X_1, \ldots, X_n]$, there exists a unique $f \in \mathbb{R} [X_1, \ldots, X_n]$ with

$$p = f(e_{n,1}, \ldots, e_{n,n}).$$
Next, we introduce the notion of symmetry in $\mathbb{R}[x_1, \ldots, x_n]$, where the notion is defined according to the case $\mathbb{R}[X_1, \ldots, X_n]$.

**Definition 2.15.**
Let $n \in \mathbb{N}$. The symmetric group $S_n$ acts on $\mathbb{R}[x_1, \ldots, x_n]$ by permuting the indices of the multilinear monomials. Formally, for every $p \in \mathbb{R}[x_1, \ldots, x_n]$ with $p = \sum_{I \subseteq [n]} \alpha_I x^I = \sum_{I \subseteq [n]} \alpha_I \prod_{i \in I} x_i$ and every $\sigma \in S_n$ we define

$$p^\sigma := p(x_{\sigma(1)}, \ldots, x_{\sigma(n)})$$

$$:= \sum_{I \subseteq [n]} \alpha_I \prod_{i \in I} x_{\sigma(i)}.$$ 

We call $p \in \mathbb{R}[x_1, \ldots, x_n]$ symmetric if

$$p = p^\sigma$$

for every $\sigma \in S_n$.

By

$$\mathbb{R}[x]^{S_n} := \{ p \in \mathbb{R}[x_1, \ldots, x_n] \mid p \text{ symmetric} \}$$

$$= \{ p \in \mathbb{R}[x_1, \ldots, x_n] \mid \forall \sigma \in S_n : p = p^\sigma \}$$

we denote the set of all symmetric polynomials of $\mathbb{R}[x_1, \ldots, x_n]$.

**Definition 2.16.**
Let $n \in \mathbb{N}$. The map $\text{Sym} : \mathbb{R}[x_1, \ldots, x_n] \to \mathbb{R}[x]^{S_n}$ defined by

$$\text{Sym}(p) := \frac{1}{n!} \sum_{\sigma \in S_n} p^\sigma$$

is called the symmetrization map and $\text{Sym}(p)$ is called the symmetrization of $p$.

We immediately obtain the following property of the above defined map.

**Lemma 2.17.**
Let $n \in \mathbb{N}$. Furthermore, let $p, q \in \mathbb{R}[x_1, \ldots, x_n]$, and let $p$ be symmetric. Then the following holds

$$\text{Sym}(pq) = p \cdot \text{Sym}(q).$$

In particular, with $q = 1$ we obtain

$$\text{Sym}(p) = p.$$
Proof. In order to save space, we just write \( f(x_\sigma) := f(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) \) for every \( f \in \mathbb{R}[x_1, \ldots, x_n] \) and every \( \sigma \in S_n \).

Since \( p \) is symmetric, it holds

\[
p = p(x_\sigma) \quad \text{for every } \sigma \in S_n. \tag{2.6}
\]

Therefore,

\[
\begin{align*}
\text{Sym}(pq) &= \frac{1}{n!} \sum_{\sigma \in S_n} (pq)^\sigma = \frac{1}{n!} \sum_{\sigma \in S_n} (pq)(x_\sigma) \\
&= \frac{1}{n!} \sum_{\sigma \in S_n} p(x_\sigma)q(x_\sigma) \tag{2.6} = \frac{1}{n!} \sum_{\sigma \in S_n} p \cdot q(x_\sigma) \\
&= p \cdot \frac{1}{n!} \sum_{\sigma \in S_n} q(x_\sigma) = p \cdot \frac{1}{n!} \sum_{\sigma \in S_n} q^\sigma \\
&= p \cdot \text{Sym}(q).
\end{align*}
\]

\[
\square
\]

Very important for our further work is the fact that for every symmetric polynomial \( f \in \mathbb{R}[x_1, \ldots, x_n] \) there exists a unique, univariate polynomial \( \tilde{f} \in \mathbb{R}[T]_n \) with \( f = \tilde{f}\left(\sum_{i \in [n]} x_i\right) \). Deriving this is the goal in the further course of this section.

Consider \( \bar{e}_{n,1} = e_{n,1}(x_1, \ldots, x_n) = \sum_{i \in [n]} x_i \in \mathbb{R}[x_1, \ldots, x_n] \). Note that it satisfies

\[
\bar{e}_{n,1} = \sum_{i \in [n]} x_i^k = \bar{p}_{n,k}
\]

for every \( k \in [n] \). Hence, this observation immediately yields the following corollary of Theorem 2.13. This corollary can be seen as Newton’s identities in \( \mathbb{R}[x_1, \ldots, x_n] \).

**Corollary 2.18.**

Let \( n \in \mathbb{N} \). Then, it holds

\[
k \cdot \bar{e}_{n,k} = \bar{e}_{n,1} \sum_{i=1}^{k} (-1)^{i-1} \bar{e}_{n,k-i}
\]

for every \( k \in [n] \).

In order to show that every symmetric \( f \in \mathbb{R}[x] \) can be represented as \( \tilde{f}(\sum_{i \in [n]} x_i) \) for a univariate polynomial \( \tilde{f} \in \mathbb{R}[T]_n \), we will first prove it for the \( \bar{e}_{n,i} \)'s.

**Notation 2.19.**

For every \( k \in \mathbb{N} \), we define \( \tau_k \in \mathbb{R}[T] \) by

\[
\tau_k := \frac{1}{k!} \prod_{i=0}^{k-1} (T - i).
\]
Theorem 2.20.  

Let $n \in \mathbb{N}$. Then, it holds

$$
\bar{e}_{n,k} = e_{n,k}(x_1, \ldots, x_n) = \tau_k \left( \sum_{i \in [n]} x_i \right),
$$

for every $k \in [n]$.

Proof. We prove this theorem by induction over $k$.

Base case: For $k = 1$, the statement is obvious because of $\tau_1 = T$. Therefore, we have

$$
\tau_1 \left( \sum_{i \in [n]} x_i \right) = \sum_{i \in [n]} x_i = \bar{e}_{n,1}.
$$

Induction step: We assume that the statement of our theorem is true for some $k \in [n - 1]$. That means

$$
\bar{e}_{n,k} = e_{n,k}(x_1, \ldots, x_n) = \tau_k \left( \sum_{i \in [n]} x_i \right). \quad (2.7)
$$

Now, we show that this is also true for $k + 1$. We have

$$
(k + 1) \cdot \bar{e}_{n,k+1} \underset{\text{Cor. 2.18}}{=} \bar{e}_{n,1} \sum_{i=1}^{k+1} (-1)^{i-1} \bar{e}_{n,k+1-i}
$$

$$
= \bar{e}_{n,1} \bar{e}_{n,k} + \bar{e}_{n,1} \sum_{i=2}^{k+1} (-1)^{i-1} \bar{e}_{n,k+1-i}
$$

$$
= \bar{e}_{n,1} \bar{e}_{n,k} + \bar{e}_{n,1} \sum_{j=1}^{k} (-1)^{j} \bar{e}_{n,k-j}
$$

$$
= \bar{e}_{n,1} \bar{e}_{n,k} - \bar{e}_{n,1} \sum_{j=1}^{k} (-1)^{j-1} \bar{e}_{n,k-j}
$$

$$
\underset{\text{Cor. 2.18}}{=} \bar{e}_{n,1} \bar{e}_{n,k} - k \cdot \bar{e}_{n,k} = (\bar{e}_{n,1} - k) \bar{e}_{n,k}
$$

$$
= \left( \sum_{i \in [n]} x_i - k \right) \bar{e}_{n,k}.
$$

Dividing the equation by $(k + 1)$ and applying the induction hypothesis (2.7) yields

$$
\bar{e}_{n,k+1} = \frac{1}{k + 1} \left( \sum_{i \in [n]} x_i - k \right) \cdot \tau_k \left( \sum_{i \in [n]} x_i \right)
$$

$$
= \frac{1}{k + 1} \left( \sum_{i \in [n]} x_i - k \right) \frac{1}{k!} \prod_{j=0}^{k-1} \left( \sum_{i \in [n]} x_i - j \right)
$$

$$
= \frac{1}{(k + 1)!} \prod_{j=0}^{k} \left( \sum_{i \in [n]} x_i - j \right)
$$

$$
= \tau_{k+1} \left( \sum_{i \in [n]} x_i \right).
$$

\qed
Now, we are able to prove the already mentioned and for the further course of this work a very important fact, that symmetric multilinear polynomials can be uniquely represented by univariate polynomials.

Theorem 2.21.
Let $n \in \mathbb{N}$. For every symmetric $p \in \mathbb{R}[x_1, \ldots, x_n]$, there exists a unique polynomial $\tilde{p} \in \mathbb{R}[T]$ with
\[ p = \tilde{p} \left( \sum_{i \in [n]} x_i \right). \]
Additionally, $\tilde{p}$ satisfies $\deg \tilde{p} \leq \deg p$.

Proof. Given any symmetric $p \in \mathbb{R}[x_1, \ldots, x_n]$. Thus, there is a family $(a_I)_{I \subseteq [n]}$ of real numbers such that $p = \sum_{I \subseteq [n]} a_I x^I = \sum_{I \subseteq [n]} a_I \prod_{i \in I} x_i$. We define $f := \sum_{I \subseteq [n]} a_I \prod_{i \in I} x_i \in \mathbb{R}[X_1, \ldots, X_n]$. Note that
\[ p = \bar{f} \text{ and } \deg(p) = \deg(f). \tag{2.8} \]

Next, we want to show that $f$ is also symmetric. In order to do this, consider $f^\sigma$ for any arbitrary $\sigma \in S_n$. By definition, we have $f^\sigma = \sum_{I \subseteq [n]} a_I \prod_{i \in I} x_{\sigma(i)}$, and therefore
\[ \bar{f}^\sigma = \sum_{I \subseteq [n]} a_I \prod_{i \in I} x_{\sigma(i)} = p^\sigma = p, \tag{2.9} \]
where we used the symmetry of $p$ in the last step.

Combining (2.8) and (2.9) yields
\[ \bar{f} - \bar{f}^\sigma = \bar{f} - \bar{f}^\sigma = p - p = 0, \]
which implies $f - f^\sigma \in \mathcal{J}$. Note that $f - f^\sigma$ is also multilinear, and therefore we can apply Corollary 2.7, which implies $f - f^\sigma = 0$. This is equivalent to
\[ f = f^\sigma. \]

But this shows that $f$ is symmetric.

Then, by the fundamental theorem of symmetric polynomials (cf. Theorem 2.14), there exists some $g \in \mathbb{R}[X_1, \ldots, X_n]$ with $f = g(e_{n,1}, \ldots, e_{n,n})$. Overall, we have
\[ p = \tilde{f} = g(e_{n,1}, \ldots, e_{n,n}) = g(\bar{e}_{n,1}, \ldots, \bar{e}_{n,n}) \]
\[ = g(\tau_1, \ldots, \tau_n) \left( \sum_{i \in [n]} x_i \right). \tag{2.10} \]
We define $\tilde{p} \in \mathbb{R}[T]$ by
\[ \tilde{p} := g(\tau_1, \ldots, \tau_n) \]
and obtain
\[ p = \tilde{p} \left( \sum_{i \in [n]} x_i \right). \]
as desired. It remains to show
\[
\deg(\tilde{p}) \leq \deg(p). \tag{2.11}
\]
Because of \(\deg(p) \leq n\), this also implies \(\tilde{p} \in \mathbb{R}[T]^n\).

If \(p = 0\), we also have \(f = g = 0\), and therefore \(\tilde{p} = 0\). In this case everything is fine. Hence, it remains to show (2.11) for \(p \neq 0\).

In order to do this, first we define
\[
|\alpha| := \sum_{i \in [n]} \alpha_i
\]
for every \(\alpha \in \mathbb{N}^n\) as usual. Note that \(g\) can be written as
\[
g = \sum_{\alpha \in \mathbb{N}^n \atop |\alpha| \leq N} b_{\alpha} \prod_{i \in [n]} X_i^{\alpha_i} \tag{2.12}
\]
for some \(N \in \mathbb{N}\) and some family of real numbers \((b_{\alpha})_{\alpha \in \mathbb{N}^n \atop |\alpha| \leq N}\). With this representation, we obtain
\[
f = \sum_{\alpha \in \mathbb{N}^n \atop |\alpha| \leq N} b_{\alpha} \prod_{i \in [n]} e_i^{\alpha_i}.
\]
Because of \(\deg(e_i) = i\), we obviously obtain
\[
\deg(f) \leq \max_{\alpha \in \mathbb{N}^n \atop |\alpha| \leq N} \left\{ \sum_{i \in [n]} \alpha_i i \mid b_{\alpha} \neq 0 \right\}. \tag{2.13}
\]
Since \(p \neq 0\), we also have \(f \neq 0\), and therefore not all \(b_{\alpha}\)’s are zero. Hence, the set on the right hand side of (2.13) is not empty. Additionally it is finite, and therefore the maximum is attained. For this reason, there exists a \(\beta \in \mathbb{N}^n\) with \(|\beta| \leq N\), \(b_{\beta} \neq 0\) and \(\sum_{i \in [n]} \beta_i i = M\), where \(M := \max_{\alpha \in \mathbb{N}^n \atop |\alpha| \leq N} \left\{ \sum_{i \in [n]} \alpha_i i \mid b_{\alpha} \neq 0 \right\}\).

Assume \(\deg(f)\) is strict lower than \(M\). Because of this assumption, the homogeneous degree-\(M\) part of \(f\) vanishes. That means
\[
\sum_{\alpha \in \mathbb{N}^n \atop |\alpha| \leq N} b_{\alpha} \prod_{i \in [n]} e_{n,i}^{\alpha_i} = 0.
\]
Because of \(b_{\beta} \neq 0\), at least one of the \(b_{\alpha}\)’s in this equation is not zero. This implies that the \(e_{n,i}\)’s are not algebraic independent, but this is a contradiction to the fundamental theorem of symmetric polynomials. Therefore, we obtain equality in (2.13). Thus, we have
\[
\deg(f) = \max_{\alpha \in \mathbb{N}^n \atop |\alpha| \leq N} \left\{ \sum_{i \in [n]} \alpha_i i \mid b_{\alpha} \neq 0 \right\}. \tag{2.14}
\]
Combining (2.10) and (2.12) results in
\[ p = \sum_{\alpha \in \mathbb{N}^n, |\alpha| \leq N} b_\alpha \prod_{i \in [n]} \tau_i^{\alpha_i}. \]

Because of \( \deg(\tau_i) = i \), we obviously obtain
\[ \deg(\tilde{p}) \leq \max_{\alpha \in \mathbb{N}^n, |\alpha| \leq N} \left\{ \sum_{i \in [n]} \alpha_i i \mid b_\alpha \neq 0 \right\}. \]

This, together with \( \deg(f) = \deg(p) \) and (2.14), yields the desired statement
\[ \deg(\tilde{p}) \leq \deg(p). \]

For once, the uniqueness of \( \tilde{p} \) is shown later in Corollary 3.32. \( \square \)

**Notation 2.22.**

For the rest of this chapter, for every symmetric \( p \in \mathbb{R}[x_1, \ldots, x_n] \) we denote by \( \tilde{p} \in \mathbb{R}[T]_n \) the unique polynomial from the previous theorem.

**Corollary 2.23.**

Let \( n \in \mathbb{N} \), and let \( p, q \in \mathbb{R}[x] \) be symmetric. Then, we have
\[ \tilde{p} + \tilde{q} = \tilde{p} + \tilde{q}. \]

**Proof.** By definition, we have
\[ p + q = \tilde{p}(\sum_{i \in [n]} x_i) + \tilde{q}(\sum_{i \in [n]} x_i) = (\tilde{p} + \tilde{q})(\sum_{i \in [n]} x_i). \]

Furthermore, note that \( \tilde{p}, \tilde{q} \in \mathbb{R}[T]_n \) implies
\[ \tilde{p} + \tilde{q} \in \mathbb{R}[T]_n. \]

Then, the statement follows by the uniqueness of the previous theorem. \( \square \)

In the case of products, we obtain the same result only with an additional condition on the degree.

**Corollary 2.24.**

Let \( n \in \mathbb{N} \), and let \( p, q \in \mathbb{R}[x_1, \ldots, x_n] \) be symmetric with \( \deg p + \deg q \leq n \). Then, we have
\[ \tilde{p} \tilde{q} = \tilde{pq}. \]

**Proof.** Because of \( \tilde{p}(\sum_{i \in [n]} x_i) = p \) and \( \tilde{q}(\sum_{i \in [n]} x_i) = q \), we have
\[ (\tilde{p} \tilde{q})(\sum_{i \in [n]} x_i) = \tilde{p} \left( \sum_{i \in [n]} x_i \right) \tilde{q} \left( \sum_{i \in [n]} x_i \right) = pq. \]
The previous theorem implies
\[
\deg (\tilde{pq}) = \deg \tilde{p} + \deg \tilde{q} \leq \deg p + \deg q \leq n,
\]
and therefore
\[
\tilde{pq} \in \mathbb{R}[T]_n.
\]
But also with the previous theorem, \(\tilde{pq}\) is the unique polynomial in \(\mathbb{R}[T]_n\) with \(\tilde{pq} \left(\sum_{i \in [n]} x_i\right) = pq\). Hence, we obtain
\[
\tilde{pq} = \tilde{p}q.
\]
\[
\square
\]

We give a small example that shows that the degree condition in the previous corollary cannot be omitted.

Example 2.25.
Let \(n \in \mathbb{N}\), and consider \(p := e_{n,1}(x_1, \ldots, x_n), q := e_{n,n}(x_1, \ldots, x_n) \in \mathbb{R}[x]\). We have
\[
pq = e_1(x_1, \ldots, x_n) \cdot e_n(x_1, \ldots, x_n) \\
= \left(\sum_{i \in [n]} x_i\right) \prod_{i \in [n]} x_i = \sum_{i \in [n]} \left(x_i \prod_{j \in [n]} x_j\right) \\
= \sum_{i \in [n]} \prod_{j \in [n]} x_j = n \prod_{i \in [n]} x_i \\
= n \cdot e_{n,n}(x_1, \ldots, x_n),
\]
and therefore
\[
\tilde{pq} \overset{\text{Th. 2.20}}{=} n \tau_n.
\]
In contrast, we have
\[
\tilde{pq} \overset{\text{Th. 2.20}}{=} \tau_1 \tau_n = T \cdot \tau_n.
\]

2.4 A new proof of Grigoriev’s Theorem

In the case of symmetric polynomials, Theorem 2.21 allows us to work in the more simple ring \(\mathbb{R}[T]\) of univariate polynomials instead of \(\mathbb{R}[x_1, \ldots, x_n]\). But there is still one difficulty in the work with sums of squares. A symmetric sum of squares in \(\mathbb{R}[x_1, \ldots, x_n]\) is not necessarily a sum of symmetric squares, and so \(\tilde{p}\), for a symmetric sum of squares \(p\), is not necessarily a sum of squares in \(\mathbb{R}[T]\).

Here the following recent theorem of Blekherman, where he characterizes symmetric sums of low degree squares in \(\mathbb{R}[x]\), helps. This theorem has a high significance to our work. First of all, it is not only an important ingredient for our new and more simple proof of Grigoriev’s result, but it is even the main reason for the simplicity of
the proof. Secondly, we will frequently use it in Chapter 5 for our detailed analysis of the so-called pseudo-densities.

Blekherman’s theorem is not published yet by himself, but a proof of it can be found in [LPWY16, Theorem B.11].

**Theorem 2.26** (Blekherman 2015). Let \( n,d \in \mathbb{N} \) with \( d \leq \frac{n}{2} \). Furthermore, let \( q \in \mathbb{R}[x_1, \ldots, x_n] \) be symmetric and \( \tilde{q} \in \mathbb{R}[T] \) the corresponding polynomial from Theorem 2.21. Then, it holds

\[
q \in \sum \mathbb{R}[x]^2_d,
\]

which means \( q \) is a sum of squares of multilinear polynomials of degree at most \( d \), if and only if

\[
\tilde{q} = \sum_{k=0}^{d} p_k \prod_{i=0}^{k-1} (T - i)(n - i - T)
\]

for some \( p_0, \ldots, p_d \in \mathbb{R}[T] \) with \( p_k \in \sum \mathbb{R}[T]_{d-k} \) for every \( k \in \{0, \ldots, d\} \).

Now, we are able to give a new proof of Grigoriev’s theorem on a lower degree bound for every RNS refutation of the Boolean polynomial equation system

\[
\sum_{i \in [n]} X_i - r = 0, \quad X_i^2 - X_i = 0, \quad 1 \leq i \leq n,
\]

where \( r \) is a non-integer real number with \( 0 < r < n \).

**Theorem 2.27** (Grigoriev, [Gri01a]). Let \( n \in \mathbb{N} \) and \( r \in \mathbb{R} \) with \( 0 < r < n \) and \( r \notin \mathbb{N} \). Furthermore, let \( k \in \mathbb{N} \) with \( 0 \leq k < r < n - k \). Then, the degree of every Real Nullstellensatz refutation of the polynomial equation system (2.15) is at least

\[
d := \min\{2k + 4, n + 1\}.
\]

**Proof.** We assume that (2.15) has any RNS refutation of degree \( d_0 < d \). Note that the latter implies \( d_0 \leq \min\{2k + 3, n\} \) The assumption means there exists a representation

\[
1 + \sum_{i \in [t]} h_i^2 = \left( \sum_{i \in [n]} X_i - r \right) g + \sum_{i \in [n]} (X_i^2 - X_i) g_i
\]

for some \( t \in \mathbb{N} \) and some \( h_1, \ldots, h_t, g, g_1, \ldots, g_n \in \mathbb{R}[X_1, \ldots, X_n] \) with

\[
\deg(h_i^2) \leq d_0 \leq \min\{2k + 3, n\} \text{ for every } i \in [t],
\]

\[
\deg(g) \leq d_0 - 1 \leq n - 1,
\]

\[
\deg(g_j) \leq d_0 - 2 \text{ for every } j \in [n].
\]
Chapter 2. Lower degree bounds for the Real Nullstellensatz

Note that (2.16) implies
\[ \deg(h_i) \leq \min \left\{ k + \frac{3}{2}, \frac{n}{2} \right\}, \]
which again implies
\[ \deg(h_i) \leq \min \left\{ k + 1, \frac{n}{2} \right\} \] \hspace{1cm} (2.18)
for every \( i \in [t] \) because the degree is always an integer.

Considering this representation in \( \mathbb{R}[x_1, \ldots, x_n] \) implies
\[ 1 + \sum_{i \in [t]} \bar{h}_i^2 = \left( \sum_{i \in [n]} x_i - r \right) \bar{g}, \] \hspace{1cm} (2.19)
with
\[ \deg(\bar{h}_i) \leq \deg(h_i) \leq \min \left\{ k + \frac{3}{2}, \frac{n}{2} \right\} \] \hspace{1cm} (2.18)
for every \( i \in [t], \) and
\[ \deg(\bar{g}) \leq \deg(g) \leq n - 1. \] \hspace{1cm} (2.17)

Now, we introduce the linear form
\[ L_r : \mathbb{R}[x_1, \ldots, x_n] \to \mathbb{R}, \ p \mapsto \tilde{\text{Sym}}(p)(r), \] \hspace{1cm} (2.21)
where \( \tilde{\text{Sym}}(p) \) is the unique, univariate polynomial from Theorem 2.21. Note that \( L_r \) is indeed a linear form because of Corollary 2.23 and because the symmetrization map is obviously linear by its definition.

In order to save space, we define \( f := \left( \sum_{i \in [n]} x_i - r \right). \) Note that we have \( \tilde{f} = T - r \) and therefore \( \tilde{f}(r) = 0. \) Hence, we have
\[ L_r \left( \left( \sum_{i \in [n]} x_i - r \right) p \right) = \tilde{\text{Sym}}(fp)(r) \overset{\text{Lem. 2.17}}{=} \left( \tilde{f} \tilde{\text{Sym}}(p) \right)(r) \overset{\text{Cor. 2.24}}{=} \tilde{f} \tilde{\text{Sym}}(p)(r) = \tilde{f}(r) \tilde{\text{Sym}}(p)(r) = 0 \] \hspace{1cm} (2.22)
for every \( p \in \mathbb{R}[x_1, \ldots, x_n] \) with \( \deg(p) \leq n - 1, \) where we explicitly need this degree condition in order to apply Corollary 2.24.

With this observation and (2.20), we obtain
\[ L_r \left( \left( \sum_{i \in [n]} x_i - r \right) \bar{g} \right) = 0. \]
Additionally, we obviously have
\[ L_r(1) = 1. \] \hspace{1cm} (2.23)

In order to yield a contradiction in the representation (2.19), it remains to show
\[ L_r(h^2) \geq 0 \text{ for every } h \in \mathbb{R}[x] \text{ with } \deg(h) \leq \min \left\{ k + 1, \frac{n}{2} \right\}. \] \hspace{1cm} (2.24)
Indeed, this would yield a contradiction because $L_r$ applied to the right-hand side of (2.19) would equal 0 while $L_r$ applied to the left-hand side of (2.19) would be at least 1.

Therefore, in order to finish the proof, we show (2.24). Let $h \in \mathbb{R}[x]$ with $\deg(h) \leq \min \{ k + 1, \lfloor \frac{n}{2} \rfloor \}$.

First, we consider the case $k+1 \leq \lfloor \frac{n}{2} \rfloor$. Then, by the definition of the symmetrization map, we have $\text{Sym}(h^2) \in \mathbb{R}[x]_{k+1}$, and we can apply Blekherman’s Theorem 2.26 which implies

$$\text{Sym}(h^2) = \sum_{i=0}^{k+1} p_i \prod_{j=0}^{i-1} (T - j)(n - j - T)$$

for some $p_0, \ldots, p_{k+1} \in \mathbb{R}[T]$ with $p_i \in \sum \mathbb{R}[T]_{k+1-i}^2$ for every $i \in \{0, \ldots, k+1\}$. Since the $p_i$’s are sums of squares, we obviously have

$$p_i(r) \geq 0$$

for every $i \in \{0, \ldots, k+1\}$.

Because of $k < r < n - k$, we additionally have

$$r - j > 0$$

and $n - j - r > 0$ for every $j \in \{0, \ldots, k\}$.

Overall, this implies

$$L_r(h^2) = \text{Sym}(h^2)(r) = \sum_{i=0}^{k+1} p_i(r) \prod_{j=0}^{i-1} (r - j)(n - j - r) \geq 0,$$

In the case $k+1 > \lfloor \frac{n}{2} \rfloor$, which by the way implies $k \geq \lfloor \frac{n}{2} \rfloor$, we obtain $\text{Sym}(h^2) \in \mathbb{R}[x]_{\lfloor \frac{n}{2} \rfloor}$, and we again apply Blekherman’s Theorem. This implies

$$\text{Sym}(h^2) = \sum_{i=0}^{\lfloor \frac{n}{2} \rfloor} p_i \prod_{j=0}^{i-1} (T - j)(n - j - T)$$

for some $p_0, \ldots, p_{\lfloor \frac{n}{2} \rfloor} \in \mathbb{R}[T]$ with $p_i \in \sum \mathbb{R}[T]_{\lfloor \frac{n}{2} \rfloor-i}^2$ for every $i \in \{0, \ldots, \lfloor \frac{n}{2} \rfloor\}$.

Then, analog to the first case, we obtain

$$L_r(h^2) = \text{Sym}(h^2)(r) = \sum_{i=0}^{\lfloor \frac{n}{2} \rfloor} p_i(r) \prod_{j=0}^{i-1} (r - j)(n - j - r) \geq 0,$$

where we this time use

$$r - j > 0$$

and $n - j - r > 0$ for every $j \in \{0, \ldots, \lfloor \frac{n}{2} \rfloor - 1\}$,

which is valid because of $\lfloor \frac{n}{2} \rfloor \leq k < r < n - k \leq n - \lfloor \frac{n}{2} \rfloor$. \qed
The previous theorem not only yields a lower degree bound for every RNS refutation of the system (2.15), but its proof also allows us to draw conclusions where the high degree is needed and even how this high degree is obtained. We explain this below.

Let \( n, k \in \mathbb{N} \) and \( r \in \mathbb{R} \) with \( 0 \leq k < r < n - k \) and \( r \not\in \mathbb{N} \) be given, and let

\[
1 + \sum_{j \in [t]} h_j^2 = \left( \sum_{i \in [n]} X_i - r \right) g + \sum_{i \in [n]} (X_i^2 - X_i) g_i
\]

for some \( t \in \mathbb{N} \) and some \( g, g_1, \ldots, g_n, h_1, \ldots, h_t \in \mathbb{R}[X_1, \ldots, X_n] \) be any RNS refutation of (2.15). Then, Grigoriev’s Theorem shows

\[
\max_{j \in [t]} \{1 + \deg(g), 2 + \deg(g_i), 2 \deg(h_j)\} \geq \max\{2k + 4n + 1\}.
\]

But by the proof of the theorem, we can conclude even more. Since we considered the refutation in \( \mathbb{R}[x_1, \ldots, x_n] \) (cf. (2.19)) and we deduced a contradiction to this representation where the \( g_i \)’s no more appear, we know that they are irrelevant for the lower degree bounds. In other words, if \( \max_{j \in [t]} \{1 + \deg(g), 2 \deg(h_j)\} < \max\{2k + 4n + 1\} \), there is no representation (2.25) not even for some \( g_i \)’s of arbitrary high degree. So, we can omit the \( g_i \)’s in (2.25). Thus, the high degree is obtained in \( g \) or in at least one of the \( h_j \)’s.

For the same reason, we can conclude that the exponents of variables are irrelevant for the lower degree bounds. By considering (2.19) and (2.20), we see that high degree is not only necessary for some \( \bar{p} \in \{\bar{g}, \bar{h}_1, \ldots, \bar{h}_t\} \subseteq \mathbb{R}[x_1, \ldots, x_n] \). In other words, not the existence of variables with high exponents is responsible for our lower degree bounds, but the existence of occurring monomials with a high number of different variables.

We state our observations in the following corollary.

**Corollary 2.28.**

Let \( n, k \in \mathbb{N} \) and \( r \in \mathbb{R} \) with \( 0 \leq k < r < n - k \) and \( r \not\in \mathbb{N} \) be given. Furthermore, let \( g, g_1, \ldots, g_n, h_1, \ldots, h_t \in \mathbb{R}[X_1, \ldots, X_n] \) for some \( t \in \mathbb{N} \) satisfying (2.25). Then, it holds

\[
\max_{j \in [t]} \{1 + \deg(\bar{g}), 2 \deg(\bar{h}_j)\} \geq \max\{2k + 4n + 1\}.
\]

Note that in particular for \( \left\lfloor \frac{n-3}{2} \right\rfloor < r < \left\lceil \frac{n+3}{2} \right\rceil \) and \( r \not\in \mathbb{N} \), Grigoriev’s Theorem shows that every RNS refutation of the system (2.15) has degree of at least \( n + 1 \). So, the degree grows linearly in \( n \).

In contrast, independent from \( n \) the system obviously has no solution because of \( r \not\in \mathbb{N} \). Moreover, the system is very easy and all occurring polynomials only have degree 2 also independent from \( n \).

So, Grigoriev’s Theorem shows the weakness of the Real Nullstellensatz. Even to refute such an easy polynomial equation system, the refutation relying on the Real Nullstellensatz needs polynomials of high degree.
Note that the system (2.15) can be interpreted as a special and easy instance of the \textit{knapsack problem}. Therefore, the Real Nullstellensatz also turns out to be weak for some easy instances of problems in combinatorial optimization. This fact makes Grigoriev’s result very interesting for people working in computational complexity.

**Remark 2.29.**

(a) Troy Lee et al. also figured out that Blekherman’s characterization of symmetric sums of squares perfectly matches the result of Grigoriev, and they simplify its proof for their part. Completely independent from us, in [LPWY16, Theorem C.1] they gave their own proof of Grigoriev’s result using Blekherman’s theorem. Indeed, for even \( n \) and \( r \in \left( \frac{n-2}{2}, \frac{n+2}{2} \right) \), their result is slightly weaker.

(b) Note that the lower bounds in Grigoriev’s Theorem are tight. For quite some time an upper bound \( n + 1 \) is known by Beame et al. [BIK+96], where they even found a RNS refutation without a sum of squares. In their recent work, Troy Lee et al. showed that the system (2.15) always has a RNS refutation of degree \( 2k + 4 \) [LPWY16, Theorem 4.3]. Both results together yield the tightness of Grigoriev’s lower bounds.

Our proof of Grigoriev’s Theorem follows the basic ideas of Grigoriev’s original proof (cf. [Gri01a, p.145]). He also considered the problem in \( \mathbb{R}[x_1, \ldots, x_n] \), and he also constructed a linear form on \( \mathbb{R}[x_1, \ldots, x_n] \), which we denote by \( L_{\text{Grig}} \), with the same properties than our linear form \( L_r \) in order to induce a contradiction. The difficult part of his proof was to show \( L_{\text{Grig}}(p^2) \geq 0 \) for every \( p \in \mathbb{R}[x] \) with \( \deg(h) \leq \min\{k+1, \frac{n}{2}\} \) which was not very hard in our proof because of Blekherman’s theorem.

To finish this section, we want to compare our linear form \( L_r \) with the one Grigoriev used in his original proof. He also defined \( L_{\text{Grig}} \) only on the monomials \( (x^I)_{I \subseteq [n]} \) and his linear form also depend only on the degree of the monomials. For \( I \subseteq [n] \) with \( |I| = k \), he defined

\[
L_{\text{Grig}}(x^I) := \prod_{i=0}^{k-1} \frac{r - i}{n - i},
\]

where \( r \) is the given real number from Theorem 2.27. In order to compare \( L_{\text{Grig}} \) with our linear form \( L_r \), we have to know which value we get if we evaluate \( L_r \) in the monomials. For this, we need the following lemma, which is also stated in [LPWY16, Lemma B.7], but there the proof needs more notations and definitions and seems more complicated. Our proof is very elementary and just needs Theorem 2.20 which follows from Newton’s identities.

**Lemma 2.30.**

Let \( n \in \mathbb{N} \), and let \( \text{Sym} : \mathbb{R}[x_1, \ldots, x_n] \to \mathbb{R}[x]^{S_n} \) as defined above. Then, for \( I \subseteq [n] \)
with $|I| = k$ we have

$$\tilde{\text{Sym}}(x^I) = \prod_{i=0}^{k-1} \frac{(T-i)}{(n-i)}$$

**Proof.**

$$\text{Sym}(x^I) = \frac{1}{n!} \sum_{\sigma \in S_n} \prod_{i \in I} x_{\sigma(i)}$$

$\overset{\text{Rem.2.12}}{=} \frac{k!(n-k)!}{n!} \cdot e_k(x_1, \ldots, x_n)$

$\overset{\text{Th.2.20}}{=} \frac{k!(n-k)!}{n!} \cdot \tau_k \left( \sum_{i \in [n]} x_i \right)$

$$= \frac{(n-k)!}{n!} \prod_{j=0}^{k-1} \left( \sum_{i \in [n]} x_i - j \right)$$

$$= \prod_{j=0}^{k-1} \left( \sum_{i \in [n]} x_i - j \right) \left( n-j \right)$$

Therefore, $\prod_{i=0}^{k-1} \frac{(T-i)}{(n-i)}$ satisfies the conditions of Theorem $2.21$, and the statement follows from the uniqueness.  

From the lemma, we immediately obtain that the linear form used in our proof of Theorem $2.27$ is actually the same as the linear form used by Grigoriev in his original proof.

**Corollary 2.31.**

Let $n \in \mathbb{N}$, and let $L_r : \mathbb{R}[x_1, \ldots, x_n] \to \mathbb{R}$ be our linear form from (2.21). Then, $L_r$ is the same linear form as $L_{\text{Grig}} : \mathbb{R}[x_1, \ldots, x_n] \to \mathbb{R}$ defined in (2.27).

**Proof.** It is enough to show $L_r(x^I) = L_{\text{Grig}}(x^I)$ for every $I \subseteq [n]$. For this, let $r$ be the real number from the setting of Theorem $2.27$. For every $k \in \{0, \ldots, n\}$ and every $I \subseteq [n]$ with $|I| = k$ we have

$$L_r(x^I) \overset{(2.21)}{=} \text{Sym}(x^I)(r) \overset{(2.20)}{=} \prod_{i=0}^{k-1} \frac{(r-i)}{(n-i)} L_{\text{Grig}}(x^I).$$
2.5 Two further conclusions from the proof of Grigoriev’s theorem

In this section, we present two conclusions from Grigoriev’s Theorem 2.27 or more precisely from its proof.

The first one is a further example for a linear lower bound on the degree of every RNS refutation of another Boolean polynomial equation system. The second one is a lower degree bound for a stronger infeasibility certificate than the Real Nullstellensatz refutation.

2.5.1 The Max-Cut problem for the complete graph

We consider the Boolean polynomial equation system

\[
\left( \frac{n}{2} \right)^2 - \sum_{1 \leq i < j \leq n} (X_i - X_j)^2 = 0, \ X_i^2 - X_i = 0, \ 1 \leq i \leq n \tag{2.28}
\]

for \( n \in \mathbb{N} \) odd.

First, we reflect that this is indeed an infeasible system. The last \( n \) equations show that only \( x \in \{0,1\}^n \) is a possible common zero. Thus, let \( x \in \{0,1\}^n \). For such a \( x \), \( \sum_{1 \leq i < j \leq n} (x_i - x_j)^2 \) is always an integer in contrast to \( \left( \frac{n}{2} \right)^2 \). But then, the first equation has no solution in \( \{0,1\}^n \), and therefore the system is infeasible.

Furthermore, we denote by \( |x| := |\{i \in [n] \mid x_i = 1\}| \in \{0,\ldots,n\} \) the number of 1-entries of \( x \). We obtain

\[
\sum_{1 \leq i < j \leq n} (x_i - x_j)^2 = \sum_{i \in [n]} \sum_{j \in [n]} 1 = |x|(n - |x|) \\
\leq \left\lfloor \frac{n}{2} \right\rfloor \left\lceil \frac{n}{2} \right\rceil \tag{2.29}
\]

The reason why Grigoriev’s Theorem immediately yields a lower degree bound for every RNS refutation of the system (2.28) lies in the fact that, considered in \( \mathbb{R}[x_1,\ldots,x_n] \), this system is only a special case of the system (2.15), and because the proof of Grigoriev’s theorem only runs in \( \mathbb{R}[x_1,\ldots,x_n] \).

Therefore, we can easily show the following result.

**Corollary 2.32.**

*Let \( n \in \mathbb{N} \) be odd. Then, the degree of every Real Nullstellensatz refutation of the polynomial equation system (2.28) is at least \( n + 1 \).*

**Proof.** As in the proof of Grigoriev’s Theorem 2.27, we assume that (2.28) has a RNS refutation of degree \( d_0 \leq n \). Considering this representation in \( \mathbb{R}[x_1,\ldots,x_n] \),
we have
\[ 1 + \sum_{i \in [t]} h_i^2 = \left( \frac{n}{2} \right)^2 - \sum_{1 \leq i < j \leq n} (x_i - x_j)^2 \tag{2.30} \]
for some \( t \in \mathbb{N} \) and some \( h_1, \ldots, h_t, g \in \mathbb{R}[x_1, \ldots, x_n] \) with \( \deg(h_i) \leq \left\lfloor \frac{n}{2} \right\rfloor \) for every \( i \in [t] \) and \( \deg(g) = n - 2 \).

Initially, note the useful identity
\[
\left( \sum_{i \in [n]} x_i \right)^2 = \sum_{i \in [n]} x_i + 2 \sum_{1 \leq i < j \leq n} x_i x_j,
\]
which implies
\[
2 \sum_{1 \leq i < j \leq n} x_i x_j = \left( \sum_{i \in [n]} x_i \right)^2 - \sum_{i \in [n]} x_i^2. \tag{2.31}
\]

Therefore, we have
\[
\left( \frac{n}{2} \right)^2 - \sum_{1 \leq i < j \leq n} (x_i - x_j)^2 = \left( \frac{n}{2} \right)^2 - \sum_{1 \leq i < j \leq n} x_i - 2x_i x_j + x_j
\]
\[
= \left( \frac{n}{2} \right)^2 - (n - 1) \sum_{i \in [n]} x_i + 2 \sum_{1 \leq i < j \leq n} x_i x_j.
\]
\[
\stackrel{(2.31)}{=} \left( \sum_{i \in [n]} x_i \right)^2 - n \sum_{i \in [n]} x_i + \left( \frac{n}{2} \right)^2
\]
\[
= \left( \sum_{i \in [n]} x_i - \frac{n}{2} \right)^2. \tag{2.32}
\]

But with this, (2.30) is a special case of (2.19) in the proof of Grigoriev’s Theorem with \( r = \frac{n}{2} \) and \( k = \left\lfloor \frac{n}{2} \right\rfloor \). Therefore, the linear form \( L_{\frac{n}{2}} \) defined in this proof (cf. (2.21)) evaluates to 0 when applied to the right-hand side of (2.30), while applying \( L_{\frac{n}{2}} \) to the left-hand side of (2.30) is at least 1.

Overall, this yields a contradiction to (2.30) and implies that the assumption was false. \( \square \)

The interesting point on the just considered system (2.28) is that it also can be interpreted as an instance for an important combinatorial optimization problem. It is an instance for the so-called, famous Max-Cut problem. More precisely, it is an instance for Max-Cut over the complete graph \( K_n \), which is maximizing the function \( \sum_{1 \leq i < j \leq n} (X_i - X_j)^2 \) over \( \{0,1\}^n \). (2.29) easily shows that this maximum is \( \left\lfloor \frac{n}{2} \right\rfloor \left\lceil \frac{n}{2} \right\rceil \) which equals \( \left( \frac{n}{2} \right)^2 - \frac{1}{4} \) for odd \( n \).

Otherwise, the previous corollary of Grigoriev’s Theorem shows that to certify that there is no cut of value \( \left( \frac{n}{2} \right)^2 \) with the Real Nullstellensatz needs at least degree \( n + 1 \). Roughly speaking, this shows the weakness of handling the Max-Cut problem over the complete graph for certificates relying on sums of squares. In other words, without going into details, this implies that finding a minimal upper bound for the
Max-Cut problem over the complete graph with the help of sums of squares needs a high degree.

Note that this is an already known result referred to Laurent [Lau03, Theorem 5]. But as we just saw, it already follows by Grigoriev’s Theorem. And because we found an easy proof for that, we automatically found an easy proof for [Lau03, Theorem 5].

2.5.2 A lower degree bound for the iterated real square root ideal refutation

Let us now turn to our second application of Grigoriev’s theorem. We will introduce a more powerful certificate for the infeasibility of a polynomial equation system than the RNS refutation, and we will prove lower degree bounds for this new certificate by using the proof of Grigoriev’s Theorem 2.27. The main reason why this proof is also suitable for the more powerful certificate is that we contradicted the existence of a RNS refutation of a certain low degree by a linear form. It turns out that this linear form is strong enough to also yield lower bounds for the new certificate, where we only need Proposition 1.32, which seems very powerful in this context.

Before motivating and introducing the new certificate, we need some definition.

Definition 2.33.
Let \( n \in \mathbb{N} \), and let \( I \subseteq \mathbb{R}[X_1, \ldots, X_n] \) be an ideal. Following the notation of Schweighofer, we define

\[
\sqrt[2]{I} := \{ f \in \mathbb{R}[X_1, \ldots, X_n] \mid \exists \sigma \in \sum \mathbb{R}[X]^2 : f^2 + \sigma \in I \}
\]

and call it the real square root ideal of \( I \).

Indeed, \( \sqrt[2]{I} \) is an ideal.

Lemma 2.34.
Let \( n \in \mathbb{N} \), and let \( I \subseteq \mathbb{R}[X_1, \ldots, X_n] \) be an ideal. Then, \( \sqrt[2]{I} \) is also an ideal.

Proof. Let \( h \in \mathbb{R}[X_1, \ldots, X_n] \) and \( f, g \in \sqrt[2]{I} \). Then, there exist \( \sigma_1, \sigma_2 \in \sum \mathbb{R}[X]^2 \) and \( f_1, f_2 \in I \) with

\[
f^2 + \sigma_1 = f_1 \quad \text{and} \quad g^2 + \sigma_2 = f_2.
\]

(2.33)

First, we show \( hf \in \sqrt[2]{I} \). Note that

\[
(hf)^2 + (h\sigma_1)^2 = h^2(f^2 + \sigma_1) \overset{(2.33)}{=} h^2f_1 \in I
\]

is valid, which implies the desired statement.

Secondly, we show \( f + g \in \sqrt[2]{I} \). It holds

\[
(f + g)^2 + (f - g)^2 + 2\sigma_1 + 2\sigma_2 = 2(f^2 + \sigma_1) + 2(g^2 + \sigma_2) \overset{(2.33)}{=} 2(f_1 + f_2) \in I.
\]

Together with \( (f - g)^2 + 2\sigma_1 + 2\sigma_2 \in \sum \mathbb{R}[X]^2 \), this shows the desired statement. \( \square \)
Now, we motivate our new certificate. In order to do this, consider a given polynomial equation system
\[ f_1 = 0, \ldots, f_s = 0. \]  
(2.34)
for some \( s \in \mathbb{N} \) and \( f_1, \ldots, f_s \in \mathbb{R}[X_1, \ldots, X_n] \). Furthermore, let \( I := \langle f_1, \ldots, f_s \rangle \) be the ideal generated by the \( f_i \)'s, and let \( d_i := \deg(f_i) \). Then, by the Real Nullstellensatz and by definition we know that (2.34) is infeasible if and only if
\[ 1 \in \sqrt[2]{I}. \]  
(2.35)
Note that to find a degree-\( d \) certificate for (2.35), which means to find \( g_i \in \mathbb{R}[X] \) for \( i \in [s] \) and \( \sigma \in \sum \mathbb{R}[X]^2 \) with \( 1 + \sigma = \sum_{i \in [s]} f_i g_i \), is exactly to find a RNS refutation of degree \( d \).

Now, if there is no such degree-\( d \) representation to certify (2.35) directly, the idea is to find a polynomial \( p \in \mathbb{R}[X_1, \ldots, X_n]_d \) with
\[ p \in \sqrt{I}, \text{ and} \]
\[ 1 \in \sqrt{\langle I, p \rangle}, \]  
(2.36)
where both ideal affiliations can be certified by polynomial identities of degree at most \( d \). This means there exist \( g_{1,i}g_{2,i} \in \mathbb{R}[X]_d, \) for every \( i \in [s], \) \( \sigma_1, \sigma_2 \in \sum \mathbb{R}[X]^2 \) and \( q \in \mathbb{R}[X]_q \) with
\[ p^2 + \sigma_1 = \sum_{i \in [s]} f_i g_{1,i}, \]  
(2.37)
\[ 1 + \sigma_2 = pq + \sum_{i \in [s]} f_i g_{2,i}. \]  
(2.38)
Note that by squaring (2.38), this implies
\[ 1 + 2\sigma_2 + \sigma_2^2 = p^2 q^2 + 2pq \left( \sum_{i \in [s]} f_i g_{2,i} \right) + \left( \sum_{i \in [s]} f_i g_{2,i} \right)^2, \]
which again implies
\[ 1 + 2\sigma_2 + \sigma_2^2 + q^2 \sigma_1 = q^2 \left( \sum_{i \in [s]} f_i g_{1,i} \right) + 2pq \left( \sum_{i \in [s]} f_i g_{2,i} \right) + \left( \sum_{i \in [s]} f_i g_{2,i} \right)^2 \in I. \]
Therefore, this polynomial identity is a degree-2d certificate for \( 1 \in \sqrt{I} \) or equivalently a RNS refutation of degree-2d.
In summary, by the two identities (2.36) of degree at most $d$, we implied $1 \in \sqrt[2]{I}$ although there is no direct degree-$d$ certificate for $1 \in \sqrt[2]{I}$. Or in other words, we implied a RNS refutation of degree-$2d$ respectively a degree-$2d$ certificate for $1 \in \sqrt[2]{I}$ just by polynomial identities of degree $d$.

If there is neither a degree-$d$ certificate for (2.36), we can iterate our idea. We can look for an additional $p_2 \in \mathbb{R}[X]^d$ with degree-$d$ certificates for $p \in \sqrt[2]\langle I,p \rangle$ and $p_2 \in \sqrt[2]\langle I,p,pp_2 \rangle$ and so on.

We want to formalize this thought and idea. For this purpose, we introduce the following.

**Definition 2.35.**
Let $n, s \in \mathbb{N}$ and $f_1, \ldots, f_s \in \mathbb{R}[X_1, \ldots, X_n]$. Furthermore, let $d_i := \deg(f_i)$ for every $i \in [s]$, and let $d \in \mathbb{N}$ with $d \geq \max\{d_1, \ldots, d_s\}$. We say the polynomial equation system (2.34) has an iterated real square root refutation of degree $d$ if there exist $t \in \mathbb{N}$, $p_1, \ldots, p_{t-1} \in \mathbb{R}[X]^d$, $\sigma_1, \ldots, \sigma_t \in \sum \mathbb{R}[X]^2$, $g_{i,j} \in \mathbb{R}[X]^{d-d_i}$ for every $i \in [s]$ and every $j \in [t]$, and $q_{i,j} \in \mathbb{R}[X]^d$ for every $i \in \{2, \ldots, t\}$ and every $j \in [i-1]$ with

$$p_k^2 + \sigma_k = \sum_{i=1}^{k-1} p_i q_{k,i} + \sum_{i=1}^{s} f_i g_{i,k} \text{ for every } k \in [t-1] \quad (2.39)$$

and

$$1 + \sigma_t = \sum_{i=1}^{t-1} p_i q_{t,i} + \sum_{i=1}^{s} f_i g_{i,t}. \quad (2.40)$$

**Remark 2.36.**
Note that the notion of refutation is reasonable because these polynomial identities (2.39) and (2.40) indeed certify the infeasibility of (2.34), which we show in the following:

First, we assume there exists $a \in \mathbb{R}^n$ with

$$f_1(a) = \ldots = f_s(a) = 0. \quad (2.41)$$

Initially, we show $p_k(a) = 0$ by induction over $k \in [t-1]$. In order to show the base case, we consider (2.39) for $k = 1$ and evaluate the equation in $a$. This yields

$$p_1(a)^2 + \sigma_1(a) = \sum_{i=1}^{s} f_i(a) g_{i,1}(a) \quad (2.41)$$

0,
which implies $p_1(a) = 0$ because squares are always nonnegative.

The induction step is just as easy. We assume

$$p_1(a) = \ldots = p_l(a) = 0 \text{ for some } l \in [t-2]$$

and show $p_{l+1}(a) = 0$. In order to do this, we consider (2.39) for $k = l+1$ and again evaluate the equation in $a$. This yields

$$p_{l+1}(a)^2 + \sigma_{l+1}(a) = \sum_{i=1}^l p_i(a)q_{l+1,i} + \sum_{i=1}^s f_i(a)g_{i,l+1}(a) = 0,$$

where we used (2.41) together with the induction hypothesis (2.42). With the same argument as above, this implies the desired statement $p_{l+1}(a) = 0$. Overall, we have

$$p_i(a) = 0 \text{ for every } i \in [t-1].$$

Evaluating (2.40) in $a$ and applying (2.41) and (2.43) yields the contradiction

$$1 \leq 1 + \sigma_t(a) = \sum_{i=1}^{t-1} p_i(a)q_{t,i}(a) + \sum_{i=1}^s f_i(a)g_{i,t}(a) = 0.$$

Therefore, the polynomials $f_1, \ldots, f_s$ have no common real root and the system (2.34) is infeasible.

**Remark 2.37.**

It is not hard to see that the iterated real square root ideal refutation is indeed stronger than the RNS refutation. If a polynomial equation system has a RNS refutation of degree $d$, it always has a iterated real square root ideal refutation of degree $d$ by choosing $t = 1$ in the above definition. Then, the RNS refutation yields (2.40).

Concluding this chapter, we show that the proof of Grigoriev’s Theorem 2.27 or more precisely the used linear form is strong enough to also yield a lower degree bound for this more powerful refutation. For this reason, we restate the polynomial equation system

$$\sum_{i \in [n]} X_i - r = 0, \; X_i^2 - X_i = 0, \; 1 \leq i \leq n,$$

where $r$ is a non-integer real number with $0 < r < n$.

**Theorem 2.38.**

Let $n \in \mathbb{N}$ and $r \in \mathbb{R}$ with $0 < r < n$ and $r \notin \mathbb{N}$. Furthermore, let $m \in \mathbb{Z}$ with $0 \leq m < r < n - m$. Then, the polynomial equation system (2.44) has no iterated real square root ideal refutation of degree $d_0 := \min\{2m + 3, n\}$.

**Proof.** Assume there exist such a refutation of degree $d_0$. 

Then, by definition, there exist \( t \in \mathbb{N}, p_1, \ldots, p_{t-1} \in \mathbb{R}[X]_{\frac{d_0}{2}}, \sigma_1, \ldots, \sigma_t \in \sum \mathbb{R}[X]_{\frac{d_0}{2}}, g_{i,j} \in \mathbb{R}[X]_{d_0-2} \) for every \( i \in [n] \) and every \( j \in [t], h_j \in \mathbb{R}[X]_{d_0-1} \) for every \( j \in [t], \) and \( q_{i,j} \in \mathbb{R}[X]_{\frac{d_0}{2}} \) for every \( i \in \{2, \ldots, t\} \) and every \( j \in [i-1], \) such that

\[
p_k^2 + \sigma_k = \sum_{i=1}^{k-1} p_i q_k,i + \left( \sum_{i\in[n]} X_i - r \right) h_k + \sum_{i=1}^{n} (X_i^2 - X_i) g_{i,k} \tag{2.45}
\]

for every \( k \in [t-1], \) and

\[
1 + \sigma_i = \sum_{i=1}^{t-1} p_i q_{i,i} + \left( \sum_{i\in[n]} X_i - r \right) h_t + \sum_{i=1}^{n} (X_i^2 - X_i) g_{i,t}. \tag{2.46}
\]

Remind the linear form \( L_r : \mathbb{R}[x] \to \mathbb{R} \) from the proof of Grigoriev’s Theorem introduced in (2.21). We naturally extend it to \( \mathbb{R}[X_1, \ldots, X_n] \) by

\[
\tilde{L} : \mathbb{R}[X_1, \ldots, X_n] \to \mathbb{R}, \quad p \mapsto L_r(p).
\]

Then, the properties of \( L_r \) obtained and proved in the proof of Grigoriev’s Theorem obviously yield the following properties of \( \tilde{L} \). It holds

\[
\tilde{L}(1) = 1 \text{ by (2.23)},
\]

\[
\tilde{L}\left((X_i^2 - X_i)g\right) = 0 \text{ for every } g \in \mathbb{R}[X_1, \ldots, X_n] \text{ by the definition of } \tilde{L},
\]

\[
\tilde{L}\left(\left(\sum_{i\in[n]} X_i - r\right)g\right) = 0 \text{ for every } g \in \mathbb{R}[X]_{n-1} \text{ by (2.22)},
\]

and

\[
\tilde{L}(h^2) \geq 0 \text{ for all } h \in \sum \mathbb{R}[X]_{\frac{d_0}{2}} \text{ by (2.24)}.
\]

Now, we proceed similar to Remark 2.36 in order to show

\[
\tilde{L}(p_k g) = 0 \text{ for every } g \in \mathbb{R}[X]_{\frac{d_0}{2}} \tag{2.51}
\]

by induction over \( k \in [t-1]. \)

Applying \( \tilde{L} \) to (2.45) with \( k = 1 \) yields

\[
\tilde{L}(p_1^2) + \tilde{L}(\sigma_1) \tilde{L}\left(\left(\sum_{i\in[n]} X_i - r\right)h_1\right) \geq 0.
\]

Because of (2.50), this implies \( \tilde{L}(p_1^2) = 0. \) Now, the crucial step is applying Proposition 1.32. Its assumptions are satisfied because of (2.50), and applying it implies the base case \( \tilde{L}(p_1 g) = 0 \) for every \( g \in \mathbb{R}[X]_{\frac{d_0}{2}} \).

For the induction step, we assume

\[
\tilde{L}(p_i g) = 0 \text{ for every } g \in \mathbb{R}[X]_{\frac{d_0}{2}} \text{ and for every } i \in [t-1] \tag{2.52}
\]
for some \( l \in [t-2] \) and prove that the statement also holds for \( l+1 \). In order to do this, we apply \( \tilde{L} \) to (2.45) with \( k = l + 1 \) and obtain

\[
\tilde{L}(p_{l+1}^2) + \tilde{L}(\sigma_{l+1}) \sum_{i=1}^{l} \tilde{L}(p_{i}q_{l+1}).
\]

Then, the induction hypothesis (2.52) together with (2.50) implies \( \tilde{L}(p_{l+1}^2) = 0 \), and again applying Proposition 1.32 implies the desired statement. Overall, (2.51) is shown.

Finally, we apply \( \tilde{L} \) to (2.46) and obtain the contradiction

\[
1 \overset{(2.47)}{=} \tilde{L}(1) \leq \tilde{L}(1) + \tilde{L}(\sigma_t) \tilde{L}\left(\sum_{i=1}^{t-1} p_i q_{t,i}\right) \overset{(2.51)}{=} 0.
\]

Similar as in the previous subsection, we easily obtain an analog result for the Max-Cut system (2.28).

Let \( \tilde{L} \) be the extension of \( L_{2^n} \). Then, we have

\[
\tilde{L}\left(\left(\frac{n}{2}\right)^2 - \sum_{1 \leq i < j \leq n} (X_i - X_j)^2\right)g \overset{(2.32)}{=} \tilde{L}\left(\sum_{i \in [n]} X_i - \frac{n}{2}\right)^2 g = 0.
\]

for every \( g \in \mathbb{R}[X_1, \ldots, X_n] \) with \( \deg(g) \leq n - 2 \).

The properties (2.47), (2.48) and (2.50) are independent from the considered system, where (2.50) is valid for \( d_0 = n \).

Thus, the above proof also works for (2.28), and we immediately obtain the following corollary.

**Corollary 2.39.**

Let \( n \in \mathbb{N} \) be odd. Then, the polynomial equation system (2.28) has no iterated real square root ideal refutation of degree \( n \).
3 Real-valued Boolean Functions

In this chapter, we have a closer view on the so called real-valued Boolean functions, which will play an important role in the work of [LRS15] and in our analysis of it.

Definition 3.1.
A real-valued Boolean function is either defined as a function $f : \{-1,1\}^n \rightarrow \mathbb{R}$ or as a function $f : \{0,1\}^n \rightarrow \mathbb{R}$, where $n \in \mathbb{N}$ in both cases. We call $n$ the number of bits of $f$. The set of all Boolean functions with $n$ bits is denoted as usual by $\mathbb{R}^{\{-1,1\}^n}$ or $\mathbb{R}^{\{0,1\}^n}$.

Note that the set $\mathbb{R}^{\{-1,1\}^n}$ or $\mathbb{R}^{\{0,1\}^n}$ not only forms a real vector space but also a commutative ring with element-wise addition and multiplication.

Actually, there is a third class of real-valued Boolean functions that occurs frequently in the literature, namely functions from $\mathbb{F}_2^n$ to $\mathbb{R}$. In the standard book of the theory of Boolean functions [O'D14] they play the second most important role after the functions with domain $\{-1,1\}^n$. One of the main reasons is that $\{-1,1\}^n$ as well as $\mathbb{F}_2^n$ build a finite group, which allows applying representation theory of finite groups and yields some nice results.

But for us this will not be important; we will just consider the case of the above definition. Of course, $\mathbb{R}^{\{-1,1\}^n}$ and $\mathbb{R}^{\{0,1\}^n}$ are obviously isomorphic as real vector spaces. Nevertheless, there are some noticeable differences in the two types of real-valued Boolean functions in their behavior as multilinear polynomials, which is the essential thing for our work.

First of all, in the first two sections we will just provide some basic facts about real-valued Boolean functions, and we will prove some small results which are helpful for our later work.

In the third section, we will look at symmetric Boolean functions from $\{0,1\}^n$ to $\mathbb{R}$. Similar to the symmetric multilinear polynomials in Section 2.3 they can be represented as univariate polynomials. This fact simplifies their use, which is very helpful for our analysis in Chapter 5, where, it turns out that considering symmetric pseudo-densities is sufficient. By the way, the nice behavior in the symmetric case is the main reason why we give Boolean functions with domain $\{0,1\}^n$ much more priority than for example [O'D14]. Another reason is that these functions are the natural ones to work with in our context of Chapter 4 because CORR$_n$ is defined by elements from $\{0,1\}^n$.

In Chapter 4, we are mainly interested in nonnegative Boolean functions. Therefore, in Section 3.4 we will introduce a natural nonnegativity certificate for such
functions relying on sums of squares, and we will adapt Blekherman’s Theorem from the last chapter to Boolean functions.

The so-called pseudo-densities, which we already mentioned several times, are specific Boolean functions that play an important role in the work of [LRS15]. We will introduce and motivate them in Section 3.5 together with so-called pseudo-expectations. In this section, we also explain the relation between these two types of functions.

3.1 Boolean functions over $\{-1,1\}^n$

We start with real-valued Boolean functions from $\{-1,1\}^n$ to $\mathbb{R}$. We only need some basics in the further course of this work. First, we consider two special types of such functions, the monomial functions and the needle functions.

Definition 3.2.
Let $n \in \mathbb{N}$.

(a) For $I \subseteq [n]$, we define the monomial function
\[
\chi_I : \{-1,1\}^n \to \mathbb{R}, \ x \mapsto \prod_{i \in I} x_i.
\]

(b) For $a \in \{-1,1\}^n$, we define the needle function
\[
\delta_a : \{-1,1\}^n \to \mathbb{R}, \ x \mapsto \begin{cases} 1 & \text{if } x = a, \\ 0 & \text{otherwise}. \end{cases}
\]

(c) For $I,J \subseteq [n]$, we denote by
\[
I \triangle J := I \cup J \setminus (I \cap J)
\]
their symmetric difference.

Remark 3.3.
Let $n \in \mathbb{N}$.

(a) For $I,J \subseteq [n]$, we have
\[
\chi_I \cdot \chi_J = \chi_{I \triangle J}.
\]

(b) It is obvious that the family $(\delta_a)_{a \in \{-1,1\}^n}$ of the needle functions is a basis of the vector space $\mathbb{R}^{\{-1,1\}^n}$. For every $f \in \mathbb{R}^{\{-1,1\}^n}$ the unique representation of $f$ as a linear combination of needle functions is
\[
f = \sum_{a \in \{-1,1\}^n} f(a) \delta_a.
\]
As well, the family \((\chi_I)_{I \subseteq [n]}\) of the monomial functions is a basis of \(\mathbb{R}^{\{-1,1\}^n}\) (cf. [O'D14, Theorem 1.1]). So, for every \(f \in \mathbb{R}^{\{-1,1\}^n}\) and every \(I \subseteq [n]\) there exist a unique real number \(\hat{f}(I)\) such that
\[
 f = \sum_{I \subseteq [n]} \hat{f}(I) \chi_I. \tag{3.1}
\]

With this remark and the definition of the monomial functions, every real-valued Boolean function can be seen as a multilinear polynomial. Therefore, the following Definition is reasonable.

**Definition 3.4.**
Let \(n \in \mathbb{N}\) and \(f \in \mathbb{R}^{\{-1,1\}^n}\). The representation (3.1) of \(f\) is called the *Fourier expansion* of \(f\). For \(I \subseteq [n]\), the real number \(\hat{f}(I)\) is called the *Fourier coefficient* of \(f\) on \(I\), and the *degree* of \(f\) is defined by
\[
 \deg f := \begin{cases} 
 \max \{|I| \mid I \subseteq [n], \hat{f}(I) \neq 0\} & \text{if } f \neq 0 \\
 -\infty & \text{if } f = 0.
 \end{cases}
\]

**Lemma 3.5.**
Let \(n \in \mathbb{N}\), and \(f,g \in \mathbb{R}^{\{-1,1\}^n}\) with \(f = \sum_{I \subseteq [n]} \hat{f}(I) \chi_I\) and \(g = \sum_{I \subseteq [n]} \hat{g}(I) \chi_I\). Then it holds
\[
 fg = \sum_{I \subseteq [n]} \sum_{J \subseteq [n]} \hat{f}(J) \hat{g}(I \triangle J) \chi_I.
\]
In particular, for all \(I \subseteq [n]\) it holds \(\hat{f}g(I) = \sum_{J \subseteq [n]} \hat{f}(J) \hat{g}(I \triangle J)\).

**Proof.**
\[
 fg = \left( \sum_{I \subseteq [n]} \hat{f}(I) \chi_I \right) \left( \sum_{I \subseteq [n]} \hat{g}(I) \chi_I \right)
 = \sum_{I \subseteq [n]} \sum_{J \subseteq [n]} \hat{f}(I) \hat{g}(J) \chi_{I \triangle J}
 = \sum_{S \subseteq [n]} \left( \sum_{I,J \subseteq [n]} \hat{f}(I) \hat{g}(J) \chi_{I \triangle J} \right) \chi_S
 = \sum_{S \subseteq [n]} \left( \sum_{K \subseteq [n]} \hat{f}(K) \hat{g}(S \triangle K) \right) \chi_S.
\]

**Notation 3.6.**
We write \(x \sim \{-1,1\}^n\) to denote that \(x\) is a uniformly chosen random element from
\(-1,1\)^n. Expectations over \(-1,1\)^n will always be with respect to this uniform distribution. Therefore, we write the shorter \(\mathbb{E}_x\) instead of \(\mathbb{E}_{x \sim \{-1,1\}^n}\). So for \(f \in \mathbb{R}^{-1,1}^n\), it holds
\[
\mathbb{E}_x f(x) = \frac{1}{2^n} \sum_{x \in \{-1,1\}^n} f(x).
\]
This notion will be used frequently in the rest of the work.

The following proposition and the subsequent corollary, an immediate consequence of the proposition, give an important property of the monomial functions. This property is one of the main reasons why it is often more practical to consider functions from \(-1,1\)^n to \(\mathbb{R}\) than from \(\{0,1\}^n\) to \(\mathbb{R}\) and why their use is more frequent in general. We will make use of this result mostly in Section 4.3.

**Proposition 3.7** ([O'D14, Fact 1.7]).

Let \(n \in \mathbb{N}\). For \(I \subseteq [n]\) with \(I \neq \emptyset\) it holds
\[
\mathbb{E}_x \chi_I(x) = \begin{cases} 0 & I \neq \emptyset \\ 1 & I = \emptyset \end{cases}.
\]

Because of the linearity of the expectation, we immediately obtain the following statement.

**Corollary 3.8** ([O'D14, Fact 1.12]).

Let \(n \in \mathbb{N}\), \(f \in \mathbb{R}^{-1,1}^n\) and let \(f = \sum_{I \subseteq [n]} \hat{f}(I) \chi_I\) be the Fourier expansion of \(f\). Then,
\[
\mathbb{E}_x f(x) = \hat{f}(\emptyset).
\]

From the previous proposition it also follows that the monomial functions even build an orthogonal basis of \(\mathbb{R}^{-1,1}^n\) with respect to the standard inner product (cf. [O'D14, Theorem 1.5]). Another consequence is the following, well-known Fourier expansion of the needle functions.

**Proposition 3.9** ([O'D14, Proposition 1.8]).

Let \(n \in \mathbb{N}\), \(a \in \{-1,1\}^n\) and \(I \subseteq [n]\). Then the Fourier coefficient of \(\delta_a\) on \(I\) satisfies \(\hat{f}(I) = \frac{1}{2} \prod_{i \in I} a_i\). In particular, we have
\[
\delta_a = \frac{1}{2^n} \sum_{I \subseteq [n]} \prod_{i \in I} a_i \chi_I.
\]
We want to generalize the notation of this section to so called \textit{matrix-valued Boolean functions} that map from $\{-1,1\}^n$ to the space of real $s \times t$-matrices, where $s,t \in \mathbb{N}$. We will use such functions in Section 4.3.

**Notation 3.10.**
Let $n,s,t \in \mathbb{N}$ and let $M : \{-1,1\}^n \rightarrow \mathbb{R}^{s \times t}$ be given. Then, we can decompose $M$ into $st$ many real-valued Boolean functions in the following natural way. For $i \in [s], j \in [t]$ let $M_{i,j} : \{-1,1\}^n \rightarrow \mathbb{R}$ defined by $M_{i,j}(x) := (M(x))_{i,j}$. Therefore, for every $I \subseteq [n]$, we define $\widehat{M}(I) \in \mathbb{R}^{s \times t}$ by $(\widehat{M}(I))_{i,j} := \widehat{M}_{i,j}(I)$ for every $i \in [s], j \in [t]$.

This yields the following unique Fourier expansion of $M$

$$M = \sum_{I \subseteq [n]} \widehat{M}(I) \chi_I.$$ 

The \textit{degree} of such a matrix-valued Boolean function $M$ is defined by

$$\deg M := \max_{i \in [s], j \in [t]} \{\deg M_{i,j}\}.$$ 

For $M$ we define the matrix-valued Boolean function $M^T : \{-1,1\}^n \rightarrow \mathbb{R}^{t \times s}$ by $M^T(x) := (M(x))^T$. With this definition, $\widehat{M}^T(I) = \widehat{M}(I)^T$ is obvious for all $I \subseteq [n]$, and in particular we have

$$M^T = \sum_{I \subseteq [n]} \widehat{M}(I)^T \chi_I.$$ 

**Definition 3.11.**
Let $n,s,t,u \in \mathbb{N}$ and let $M$ be a matrix-valued Boolean functions $M : \{-1,1\}^n \rightarrow \mathbb{R}^{s \times t}$. The \textit{support} of $M$ is defined as the set of all subsets $I \subseteq [n]$ with $\widehat{M}(I) \neq 0$.

Proposition 3.7 and the above definition yields the following result.

**Lemma 3.12.**
Let $n,s,t,u \in \mathbb{N}$. For $M : \{-1,1\}^n \rightarrow \mathbb{R}^{s \times t}$ and $N : \{-1,1\}^n \rightarrow \mathbb{R}^{t \times u}$ with disjoint support it holds

$$\mathbb{E}_x M(x)N(x) = 0.$$ 

**Proof.** By definition, there are $P_1,P_2 \subseteq \mathcal{P}([n])$ with $M = \sum_{I \in P_1} \widehat{M}(I) \chi_I$ and $N = \sum_{J \in P_2} \widehat{N}(J) \chi_J$ and $P_1 \cap P_2 = \emptyset$, where the latter is equivalent to

$$I \triangle J \neq \emptyset \text{ for all } I \in P_1, J \in P_2.$$ 

Therefore, by Proposition 3.7 we have

$$\mathbb{E}_x \chi_{I \triangle J}(x) = 0 \quad (3.2)$$
for all \( I \in P_1, J \in P_2 \). For every \( x \in \{-1,1\}^n \), we have
\[
M(x)N(x) = \left( \sum_{I \in P_1} \tilde{M}(I) \chi_I(x) \right) \left( \sum_{J \in P_2} \tilde{N}(J) \chi_J(x) \right)
= \sum_{I \in P_1} \sum_{J \in P_2} \tilde{M}(I) \tilde{N}(J) \chi_{I \triangle J}(x).
\]

Now, taking the expectation over all \( x \in \{-1,1\}^n \) and applying (3.2) yields
\[
\mathbb{E}_x M(x)N(x) = \mathbb{E}_x \sum_{I \in P_1} \sum_{J \in P_2} \tilde{M}(I) \tilde{N}(J) \chi_{I \triangle J}(x)
= \sum_{I \in P_1} \sum_{J \in P_2} \tilde{M}(I) \tilde{N}(J) \mathbb{E}_x \chi_{I \triangle J}(x)
\overset{(3.2)}{=} 0.
\]

\[
3.2 \text{ Boolean functions over } \{0,1\}^n
\]

In this section, we have a look at real-valued Boolean functions with domain \( \{0,1\}^n \), where most notations and definitions from the last section are kept but repeated. In the subsequent section, it turns out that especially in the symmetric case they are much nicer to work with.

According to Boolean functions over \( \{-1,1\}^n \), we start by defining the monomial and needle functions.

**Definition 3.13.**
Let \( n \in \mathbb{N} \).

(a) For \( I \subseteq [n] \), we define the monomial function
\[
\chi_I : \{0,1\}^n \to \mathbb{R}, \ x \mapsto \prod_{i \in I} x_i.
\]

(b) For \( a \in \{0,1\}^n \), we define the needle function
\[
\delta_a : \{0,1\}^n \to \mathbb{R}, \ x \mapsto \begin{cases} 1 & \text{if } x = a, \\ 0 & \text{otherwise}. \end{cases}
\]

**Remark 3.14.**
Let \( n \in \mathbb{N} \).

(a) For \( I,J \subseteq [n] \), we have
\[
\chi_I \cdot \chi_J = \chi_{I \cup J}.
\]
(b) Just as in the last section, it is obvious that the family \( \{ \delta_a \}_{a \in \{0,1\}^n} \) of the needle functions is a basis of the real vector space \( \mathbb{R}^{(0,1)^n} \). For every \( f \in \mathbb{R}^{(0,1)^n} \) the unique representation of \( f \) as a linear combination of needle functions is

\[
f = \sum_{a \in \{0,1\}^n} f(a) \delta_a.
\]

(c) And again, the family \( \{ \chi_I \}_{I \subseteq [n]} \) of the monomial functions is also a basis of \( \mathbb{R}^{(0,1)^n} \). The linear independence of the monomial functions is exactly the statement of Lemma 2.6. Since we have \( 2^n \) linearly independent monomial functions, what matches the dimension of \( \mathbb{R}^{(0,1)^n} \), these functions also generate the whole space.

So, for every \( f \in \mathbb{R}^{(0,1)^n} \) and every \( I \subseteq [n] \) there exist a unique real number \( \lambda_I \) such that

\[
f = \sum_{I \subseteq [n]} \lambda_I \chi_I. \tag{3.3}
\]

According to the last section, we also define the expectation over \( \{0,1\}^n \).

**Notation 3.15.**

We write \( x \sim \{0,1\}^n \) to denote that \( x \) is a uniformly chosen random element from \( \{0,1\}^n \). Expectations over \( \{0,1\}^n \) will always be with respect to this uniform distribution. Therefore, we write the shorter \( \mathbb{E}_x \) instead of \( \mathbb{E}_{x \sim \{0,1\}^n} \) if it is clear that we take the expectation over \( \{0,1\}^n \). So for \( f \in \mathbb{R}^{(0,1)^n} \), it holds

\[
\mathbb{E}_x f(x) = \frac{1}{2^n} \sum_{x \in \{0,1\}^n} f(x).
\]

According to the case \( \{-1,1\}^n \) in the last section, also this notion will be used frequently in the rest of the work.

**Definition 3.16.**

Let \( n \in \mathbb{N} \) and \( f \in \mathbb{R}^{(0,1)^n} \). For every \( I \subseteq [n] \), let \( \lambda_I \in \mathbb{R} \) such that the situation of (3.3) is given. Then, the degree of \( f \) is defined by

\[
\deg f := \begin{cases} 
\max \{|I| \mid I \subseteq [n], \lambda_I \neq 0 \} & \text{if } f \neq 0 \\
-\infty & \text{if } f = 0
\end{cases}.
\]

Furthermore, for every \( x \in \{0,1\}^n \) we denote by

\[
|x| := \sum_{i \in [n]} x_i
\]

the number of 1-entries of an element \( x \) and call it the hamming weight of \( x \).
Chapter 3. Real-valued Boolean Functions

Especially the quadratic Boolean functions are related to the correlation polytope which we are working with in the next Chapter. Therefore, we state the following small result.

**Lemma 3.17.**
Let $n \in \mathbb{N}$, and let $f \in \mathbb{R}^{\{0,1\}^n}$ be a quadratic Boolean function, that means $\deg f = 2$. Then, there exist a matrix $B \in \mathbb{R}^{n \times n}$ and a real number $b$ such that

$$f(x) = \langle B, xx^T \rangle + b$$

for all $x \in \{0,1\}^n$.

**Proof.** Since $\deg f = 2$, we have

$$f = \sum_{I \in [n], |I| = 2} \lambda_I \chi_I$$

for some $\lambda_I \in \mathbb{R}$ for all $I \in [n]$ with $|I| = 2$. Therefore, we have

$$f(x) = \sum_{I \in [n], |I| = 2} \lambda_I \chi_I(x) = \sum_{1 \leq i \leq j \leq n} \lambda_{\{i,j\}} x_i x_j + \lambda_\emptyset,$$

(3.4)

for all $x \in \{0,1\}^n$, where we used $x_i = x_i^2$ for all $x \in \{0,1\}^n$ and all $i \in [n]$.

We define $b := \lambda_\emptyset$ and $B \in \mathbb{R}^{n \times n}$ by $B_{i,j} := \begin{cases} \lambda_{\{i,j\}} & i \leq j \\ 0 & \text{otherwise} \end{cases}$. Then, we obtain

$$f(x) = \sum_{1 \leq i \leq j \leq n} \lambda_{\{i,j\}} x_i x_j + \lambda_\emptyset$$

$$= \sum_{1 \leq i \leq j \leq n} B_{i,j} x_i x_j + b = \sum_{i,j \in [n]} B_{i,j} x_i x_j + b$$

(3.4)

$$\overset{\text{Lem.1.5}}{=} \langle B, xx^T \rangle + b.$$

\qed

In the case of $\mathbb{R}^{\{-1,1\}^n}$, the nice theory about Fourier expansion and coefficients, called *Fourier analysis of Boolean functions* and based on Proposition 3.7, gives a nice representation for the needle functions as a linear combination of monomial functions (cf. Proposition 3.9). In contrast, in the case of $\mathbb{R}^{\{0,1\}^n}$ it is not so easy. Deriving such a representation of the needle functions is our next goal, where we want to resort to the results from the case $\mathbb{R}^{\{-1,1\}^n}$. For this purpose, we first need some preparation.

**Notation 3.18.**
Let $n \in \mathbb{N}$. Then we define

$$\psi : \{0,1\}^n \to \{-1,1\}^n, \ (x_i)_{i \in [n]} \mapsto (1 - 2x_i)_{i \in [n]}.$$ 

Obviously, $\psi$ is a bijection.
We put $\mathbb{R}^{(-1,1)^n}$ and $\mathbb{R}^{(0,1)^n}$ in context via the following map.

**Notation 3.19.**
Let $n \in \mathbb{N}$. Then, we define

$$\hat{\psi} : \mathbb{R}^{(-1,1)^n} \to \mathbb{R}^{(0,1)^n}, \ f \mapsto f \circ \psi.$$  

Note that $\hat{\psi}$ preserves the degree of the Boolean functions. Additionally, it turns out that $\hat{\psi}$ is not only an isomorphism but it is also a vector space isomorphism with the property that it maps needle functions onto needle functions, as we will see in the next lemma.

**Lemma 3.20.**
Let $n \in \mathbb{N}$ and let $\hat{\psi}$ and $\psi$ defined as above. Then, for every $a \in \{-1,1\}^n$ we have

$$\hat{\psi}(\delta_a) = \delta_{\psi^{-1}(a)}.$$  

In particular, $\hat{\psi}$ is a vector space isomorphism.

**Proof.** Let $a \in \{-1,1\}^n$. For every $x \in \{0,1\}^n$, we have

$$\hat{\psi}(\delta_a)(x) = \delta_a(\psi(x)) = \begin{cases} 1 & \psi(x) = a \\ 0 & \text{otherwise} \end{cases} = \begin{cases} 1 & x = \psi^{-1}(a) \\ 0 & \text{otherwise} \end{cases}.$$  

By definition and because $\psi^{-1}(a)$ is a single element, we obtain

$$\hat{\psi}(\delta_a) = \delta_{\psi^{-1}(a)}. \quad (3.5)$$

Note that

$$(\delta_{\psi^{-1}(a)})_{a \in \{-1,1\}^n} = (\delta_b)_{b \in \{0,1\}^n} \quad (3.6)$$

because $\psi$ is a bijection.

The linearity of $\hat{\psi}$ easily follows by the pointwise definition of the addition and multiplication by a scalar in $\mathbb{R}^{(0,1)^n}$. $\hat{\psi}$ is additionally bijective because it maps the basis $(\delta_a)_{a \in \{-1,1\}^n}$ of $\mathbb{R}^{(-1,1)^n}$ onto the basis $(\delta_b)_{b \in \{0,1\}^n}$ of $\mathbb{R}^{(0,1)^n}$ which follows by combining (3.5) and (3.6).

Applying the map $\hat{\psi}$ on the Fourier expansion of needle functions in $\mathbb{R}^{(-1,1)^n}$ from Proposition 3.9 yields the following almost monomial representation of needle functions in $\mathbb{R}^{(0,1)^n}$. The final representation is then obtained in the subsequent proposition.
Lemma 3.21.
Let $n \in \mathbb{N}$. Then, for all $b, x \in \{0,1\}^n$ it holds

$$\delta_b(x) = \frac{1}{2^n} \sum_{I \subseteq [n]} \prod_{i \in I} (1 - 2b_i)(1 - 2x_i).$$

Proof. Let $b \in \{0,1\}^n$. Then, we have

$$\delta_b \overset{\text{Lem 3.20}}{=} \hat{\psi}(\delta_{\psi(b)}) = \delta_{\psi(b)} \circ \psi,$$

where $\delta_{\psi(b)} \in \mathbb{R}^{|-1,1|^n}$. The Fourier expansion of $\delta_{\psi(b)}$ yields

$$\delta_{\psi(b)} \overset{\text{Prop 3.19}}{=} \frac{1}{2^n} \sum_{I \subseteq [n]} \prod_{i \in I} \psi(b)_i \chi_I,$$

where $\chi_I \in \mathbb{R}^{|-1,1|^n}$ for every $I \subseteq [n]$. By combining the last two equations and the definition of $\psi$, we obtain

$$\delta_b(x) = \delta_{\psi(b)} \circ \psi(x) = \delta_{\psi(b)}(\psi(x)) = \frac{1}{2^n} \sum_{I \subseteq [n]} \prod_{i \in I} \psi(b)_i \chi_I(\psi(x)) = \frac{1}{2^n} \sum_{I \subseteq [n]} \prod_{i \in I} \psi(b)_i \psi(x)_i$$

$$= \frac{1}{2^n} \sum_{I \subseteq [n]} \prod_{i \in I} (1 - 2b_i)(1 - 2x_i).$$

Proposition 3.22.
Let $n \in \mathbb{N}$ and $a \in \{0,1\}^n$. We define the set $I_a := \{i \in [n] \mid a_i = 1\}$. Then, the unique representation of $\delta_a$ as a linear combination of monomial functions is

$$\delta_a = \sum_{I \subseteq [n]} (-1)^{|I| - |a|} \chi_I.$$ 

Proof. We prove the proposition by induction over $n$.

**Base case**: For the base case let $n = 1$, and therefore let $a \in \{0,1\}$. We distinguish the cases $a = 0$ and $a = 1$.

**Case 1**: $a = 0$. 

Note that $I_a = \emptyset$ and $|a| = 0$. For all $x \in \{0,1\}$ we have
\[
\delta_a(x) = \begin{cases} 
1 & x = a \\
0 & \text{otherwise} 
\end{cases} = \begin{cases} 
1 & x = 0 \\
0 & x = 1 
\end{cases} = 1 - x \\
= \chi_\emptyset(x) - \chi_{\{1\}}(x).
\]
Therefore
\[
\delta_a = \chi_\emptyset - \chi_{\{1\}} = \sum_{I \subseteq \{1\}} (-1)^{|I| - 0} \chi_I.
\]

Case 2: $a = 1$.
Note that $I_a = \{1\}$ and $|a| = 1$. For all $x \in \{0,1\}$ we have
\[
\delta_a(x) = \begin{cases} 
1 & x = a \\
0 & \text{otherwise} 
\end{cases} = \begin{cases} 
1 & x = 1 \\
0 & x = 0 
\end{cases} = x \\
= \chi_{\{1\}}(x)
\]
Therefore,
\[
\delta_a = \chi_{\{1\}} = \sum_{I \subseteq \{1\}} (-1)^{|I| - 1} \chi_I.
\]

Induction step: We assume that the statement of our proposition is true for some $k \in \mathbb{N}$. That means for all $a \in \{0,1\}^k$,
\[
\delta_a = \sum_{I \subseteq [k]} (-1)^{|I| - |a|} \chi_I,
\]
where $\delta_a \in \mathbb{R}^{\{0,1\}^k}$ and $\chi_I \in \mathbb{R}^{\{0,1\}^k}$ for every $I \subseteq [k]$. This implies
\[
\delta_a(x) = \sum_{I \subseteq [k]} (-1)^{|I| - |a|} \prod_{i \in I} x_i \quad (3.7)
\]
for all $x \in \{0,1\}^k$. 

In order to show that our statement is also true for \( k + 1 \), let \( b \in \{0,1\}^{k+1} \). For \( \delta_b \in \mathbb{R}^{(0,1)^{k+1}} \) and for all \( x \in \{0,1\}^{k+1} \), we obtain by the identity of the previous lemma

\[
\delta_b(x) = \frac{1}{2^{k+1}} \sum_{I \subseteq [k+1]} \prod_{i \in I} (1 - 2b_i)(1 - 2x_i)
= \frac{1}{2^{k+1}} \left( \sum_{I \subseteq [k+1]} \prod_{i \in I} (1 - 2b_i)(1 - 2x_i) + \sum_{k+1 \in I} \prod_{i \in I} (1 - 2b_i)(1 - 2x_i) \right)
= \frac{1}{2^{k+1}} \left( \sum_{I \subseteq [k]} \prod_{i \in I} (1 - 2b_i)(1 - 2x_i) + \sum_{I \subseteq [k+1]} \prod_{i \in I \cup \{k+1\}} (1 - 2b_i)(1 - 2x_i) \right)
= \frac{1}{2^{k+1}} \left( \sum_{I \subseteq [k]} \prod_{i \in I} (1 - 2b_i)(1 - 2x_i) + (1 - 2b_{k+1})(1 - 2x_{k+1}) \sum_{I \subseteq [k]} \prod_{i \in I} (1 - 2b_i)(1 - 2x_i) \right)
= \frac{1}{2} \left( 1 + (1 - 2b_{k+1})(1 - 2x_{k+1}) \right) \left( \sum_{I \subseteq [k]} \prod_{i \in I} (1 - 2b_i)(1 - 2x_i) \right)
= \frac{1}{2} \left( 1 + (1 - 2b_{k+1})(1 - 2x_{k+1}) \right) \left( \frac{1}{2^k} \sum_{I \subseteq [k]} \prod_{i \in I} (1 - 2b_i)(1 - 2x_i) \right). \tag{3.8}
\]

We define \( \hat{b} := (b_1, \ldots, b_k) \in \{0,1\}^k \), and analog for every \( x \in \{0,1\}^{k+1} \) we define \( \hat{x} := (x_1, \ldots, x_k) \in \{0,1\}^k \). With this definition and by applying the previous lemma, we obtain

\[
\delta_{\hat{b}}(\hat{x}) = \frac{1}{2^k} \sum_{I \subseteq [k]} \prod_{i \in I} (1 - 2b_i)(1 - 2x_i).
\]

Then, putting this into (3.8) yields

\[
\delta_b(x) = \frac{1}{2} \left( 1 + (1 - 2b_{k+1})(1 - 2x_{k+1}) \right) \delta_{\hat{b}}(\hat{x})
\]

for every \( x \in \{0,1\}^{k+1} \). Since \( \delta_b \in \mathbb{R}^{(0,1)^{k}} \) and \( \hat{x} \in \{0,1\}^k \), we can apply the induction hypothesis (3.7). Overall, we obtain

\[
\delta_b(x) = \frac{1}{2} \left( 1 + (1 - 2b_{k+1})(1 - 2x_{k+1}) \right) \sum_{I \subseteq [k]} (-1)^{|I| - |\hat{b}|} \prod_{i \in I} x_i \tag{3.9}
\]

for all \( x \in \{0,1\}^{k+1} \).

This time, we distinguish the cases \( b_{k+1} = 0 \) and \( b_{k+1} = 1 \).
Chapter 3. Real-valued Boolean Functions

63

Case 1: $b_{k+1} = 0$.

Note that $b_{k+1} = 0$ implies $\hat{b} = |b|$ and $I_\hat{b} = I_b$. In this case, the previous equation results in

$$\delta_b(x) = (1 - x_{k+1}) \sum_{I \subseteq [k]} (-1)^{|I| - |b|} \prod_{i \in I} x_i$$

$$= \sum_{I \subseteq [k]} (-1)^{|I| - |b|} \prod_{i \in I} x_i - x_{k+1} \sum_{I \subseteq [k]} (-1)^{|I| - |b|} \prod_{i \in I} x_i$$

$$= \sum_{I \subseteq [k]} (-1)^{|I| - |b|} \prod_{i \in I} x_i + \sum_{I \subseteq [k]} (-1)^{|I| - |b|} \prod_{i \in I \cup \{k+1\}} x_i$$

$$= \sum_{I \subseteq [k+1]} (-1)^{|I| - |b|} \prod_{i \in I} x_i$$

for all $x \in \{0,1\}^{k+1}$. This implies the desired statement

$$\delta_b = \sum_{I \subseteq [k+1]} (-1)^{|I| - |b|} \chi_I.$$  

Case 2: $b_{k+1} = 1$.

Note that $b_{k+1} = 1$ implies $|\hat{b}| = |b| - 1$ and

$$I_b = I_\hat{b} \cup \{k + 1\}.$$  

(3.10)

In this case, (3.9) results in

$$\delta_b(x) = x_{k+1} \sum_{I \subseteq [k]} (-1)^{|I|+1-|b|} \prod_{i \in I} x_i$$

$$= \sum_{I \subseteq [k]} (-1)^{|I|+1-|b|} \prod_{i \in I \cup \{k+1\}} x_i = \sum_{I \subseteq [k+1]} (-1)^{|I| - |b|} \prod_{i \in I} x_i$$

$$= \sum_{I \subseteq [k+1]} (-1)^{|I| - |b|} \prod_{i \in I} x_i$$  

(3.11)

for all $x \in \{0,1\}^{k+1}$. In the last step, we used that for all $I \subseteq [k + 1]$ the properties $I \supseteq I_b$ and $k + 1 \in I$ together with (3.10) are equivalent to the property $I \supseteq I_b$. Now, (3.11) implies the desired statement

$$\delta_b = \sum_{I \subseteq [k+1]} (-1)^{|I| - |b|} \chi_I.$$  

□
3.3 Symmetric Boolean functions over \( \{0,1\}^n \)

As already mentioned in the introduction of the chapter, we want to obtain a representation of symmetric Boolean functions as univariate polynomials similar as for the multilinear polynomials in Section 2.3. Therefore, deriving an analogon of Theorem 2.21 for symmetric Boolean functions is the goal of this section.

The notion of a symmetric Boolean function will be clear, but we want to define it formally.

**Definition 3.23.**
Let \( n \in \mathbb{N} \). The symmetric group \( S_n \) acts on \( \{0,1\}^n \) by permuting the components. Formally, for all \( x \in \{0,1\}^n \) and all \( \sigma \in S_n \) we define

\[
x^\sigma := (x_{\sigma(1)}, \ldots, x_{\sigma(n)}).
\]

Additionally, the group \( S_n \) also acts on \( \mathbb{R}^{\{0,1\}^n} \). For every \( f \in \mathbb{R}^{\{0,1\}^n} \) and all \( \sigma \in S_n \) we define

\[
f^\sigma : \{0,1\}^n \to \mathbb{R}, \ x \mapsto f(x^\sigma).
\]

We call a \( f \in \mathbb{R}^{\{0,1\}^n} \) symmetric if

\[
f(x) = f(x^\sigma) \text{ for all } x \in \{0,1\}^n \text{ and all } \sigma \in S_n,
\]

or equivalent, if

\[
f = f^\sigma \text{ for all } \sigma \in S_n.
\]

**Remark 3.24.**
Let \( n \in \mathbb{N} \).

(a) By definition, we immediately obtain

\[
(x^\sigma)^\tau = x^{\sigma \circ \tau}
\]

for every \( x \in \{0,1\}^n \) and all \( \sigma, \tau \in S_n \).

(b) For every fixed \( \sigma \in S_n \), \( x \mapsto x^\sigma \) is a bijection, that means

\[
\{x^\sigma \mid x \in \{0,1\}^n\} = \{0,1\}^n.
\]

In particular, as an immediate consequence, we obtain

\[
\mathbb{E}_xf(x^\sigma) = \mathbb{E}_xf(x) \quad (3.12)
\]

for every \( f \in \mathbb{R}^{\{0,1\}^n} \) and every \( \sigma \in S_n \), what we will use several times in the further course of our work.
(c) If \( f = \sum_{I \subseteq [n]} \lambda_I \chi_I \), then \( f^\sigma \) satisfies
\[
 f^\sigma(x) = f(x^\sigma) = \sum_{I \subseteq [n]} \lambda_I \prod_{i \in I} x_{\sigma(i)}
\]
for all \( x \in \{0,1\}^n \) by definition. Therefore, we have
\[
 f^\sigma = \sum_{I \subseteq [n]} \lambda_I \prod_{i \in I} \chi_{\{\sigma(i)\}}.
\]
In particular, the action of \( S_n \) does not affect the degree, or in other words
\[
 \deg(f) = \deg(f^\sigma)
\]
for every \( f \in \mathbb{R}\{0,1\}^n \) and every \( \sigma \in S_n \).

(d) Note that \( |x| = |x^\sigma| \) for all \( x \in \{0,1\}^n \) and all \( \sigma \in S_n \). Furthermore, it is not hard to see that for all \( x,y \in \mathbb{R}\{0,1\}^n \) with \( |x| = |y| \) there exists a \( \sigma \in S_n \) with \( x^\sigma = y \). Together, this yields an additional equivalent definition of symmetry for a \( f \in \mathbb{R}\{0,1\}^n \). Thus, \( f \) is symmetric if and only if
\[
 f(x) = f(y) \text{ for all } x,y \in \{0,1\}^n \text{ with } |x| = |y|.
\]

**Definition 3.25.**
Let \( n \in \mathbb{N} \). We denote by \( \text{Sym} \left( \mathbb{R}\{0,1\}^n \right) \) the *ring of symmetric functions* of \( \mathbb{R}\{0,1\}^n \). That means
\[
 \text{Sym} \left( \mathbb{R}\{0,1\}^n \right) := \{ f \in \mathbb{R}\{0,1\}^n \mid f \text{ symmetric} \}.
\]
Analog to Section 2.3 we define the map \( \text{Sym} : \mathbb{R}\{0,1\}^n \to \text{Sym} \left( \mathbb{R}\{0,1\}^n \right) \) by
\[
 \text{Sym}(f) := \frac{1}{n!} \sum_{\sigma \in S_n} f^\sigma.
\]
For every \( f \in \mathbb{R}\{0,1\}^n \), we call \( \text{Sym}(f) \) the *symmetrization* of \( f \).

Now, the following map formally links the ring \( \mathbb{R}[x_1, \ldots, x_n] \) of multilinear polynomials from the previous chapter with the ring \( \mathbb{R}\{0,1\}^n \).

**Notation 3.26.**
Let \( n \in \mathbb{N} \). Then, we introduce the canonical vector space isomorphism
\[
 \Phi : \mathbb{R}\{0,1\}^n \to \mathbb{R}[x_1, \ldots, x_n].
\]
We define it on the basis of monomial functions and extend it linearly. So, for all \( I \subseteq [n] \) we define
\[
 \Phi(\chi_I) := x^I.
\]
with \( x^I = \prod_{i \in I} x_i \), as defined in Section 2.2. Therefore, \( \Phi \) maps the basis of monomial functions onto the basis \( (x^I)_{I \subseteq \{n\}} \) of \( \mathbb{R} [x] \). For an arbitrary \( f \in \mathbb{R}^{(0,1)^n} \) with \( f = \sum_{I \subseteq \{n\}} \lambda_I \chi_I \), this yields

\[
\Phi(f) = \sum_{I \subseteq \{n\}} \lambda_I x^I.
\]

**Remark 3.27.**

(a) Because of \( \chi_I \cdot \chi_J = \chi_{I \cup J} \) and \( x^I \cdot x^J = x^{I \cup J} \), \( \Phi \) is even a ring isomorphism.

(b) Note that \( \Phi \) preserves the degree. That means,

\[
\deg f = \deg \Phi(f)
\]

for all \( f \in \mathbb{R}^{(0,1)^n} \).

The notion of symmetry is preserved under \( \Phi \) with respect to the definition of symmetry in \( \mathbb{R} [x_1, \ldots, x_n] \).

**Lemma 3.28.**

Let \( n \in \{n\} \) and \( \Phi \) defined as above. Let \( f \in \mathbb{R}^{(0,1)^n} \) be symmetric. Then, \( \Phi(f) \) is symmetric as well.

**Proof.** Let \( f = \sum_{I \subseteq \{n\}} \lambda_I \chi_I \), where \( \lambda_I \in \mathbb{R} \) for all \( I \subseteq \{n\} \). Then we have

\[
\Phi(f) = \sum_{I \subseteq \{n\}} \lambda_I x^I,
\]

and therefore

\[
\Phi(f)^\sigma = \sum_{I \subseteq \{n\}} \lambda_I \prod_{i \in I} x_{\sigma(i)}
\]

for all \( \sigma \in S_n \) by Definition 2.15.

By Remark 3.24 and because \( \Phi \) is a ring homomorphism, we also obtain

\[
\Phi(f^\sigma) = \sum_{I \subseteq \{n\}} \lambda_I \prod_{i \in I} x_{\sigma(i)}
\]

for all \( \sigma \in S_n \). Comparing the last two equations, we obtain

\[
\Phi(f^\sigma) = \Phi(f^\sigma) = \Phi(f)
\]

for all \( \sigma \in S_n \), where we used the assumption \( f = f^\sigma \) in the last step .

Overall, by Definition 2.15 this yields the statement. \( \square \)

With the help of the isomorphism \( \Phi \), we are now able to apply Theorem 2.21 in order to obtain the same result for symmetric functions in \( \mathbb{R}^{(0,1)^n} \) as for symmetric polynomials in \( \mathbb{R} [x_1, \ldots, x_n] \).
Notation 3.29.
Let \( n \in \mathbb{N} \). Then, we introduce the map
\[
\Psi : \mathbb{R}[T]_n \to \text{Sym} \left( \mathbb{R}^{\{0,1\}^n} \right)
\]
\[
p \mapsto p \left( \sum_{i \in [n]} \chi_{\{i\}} \right).
\]
Note that for every \( p \in \mathbb{R}[T]_n \) the Boolean function \( \Psi(p) \) satisfies
\[
\Psi(p)(x) = p \left( \sum_{i \in [n]} \chi_{\{i\}} \right)(x) = p(|x|) \quad \text{for all } x \in \{0,1\}^n.
\] (3.13)
Therefore, for every \( x,y \in \{0,1\}^n \) with \(|x| = |y|\), we have
\[
\Psi(p)(x) = p(|x|) = p(|y|) = \Psi(p)(y),
\]
which shows the symmetry of \( \Psi(p) \), why \( \Psi \) is well-defined.

Now, we prove that \( \Psi \) is even a vector space isomorphism. In particular, this shows our desired analogon of Theorem 2.21.

Proposition 3.30.
Let \( n \in \mathbb{N} \) and \( \Psi \) defined as above. Then, \( \Psi \) is a vector space isomorphism. In particular, for every \( f \in \text{Sym} \mathbb{R}^{\{0,1\}^n} \) there is a unique univariate polynomial \( \tilde{f} \in \mathbb{R}[T]_n \) with
\[
f = \tilde{f} \left( \sum_{i \in [n]} \chi_{\{i\}} \right).
\]

Proof. The linearity of \( \Psi \) is obvious by the definition of the addition and the multiplication by a scalar in the vector spaces \( \mathbb{R}[T]_n \) and \( \text{Sym} \left( \mathbb{R}^{\{0,1\}^n} \right) \).

Let \( p,q \in \mathbb{R}[T]_n \) with \( \Psi(p) = \Psi(q) \). Then, we have
\[
(p - q)(|x|) = p(|x|) - q(|x|) = \Psi(p)(x) - \Psi(q)(x) = 0
\]
for every \( x \in \{0,1\}^n \). This implies
\[
(p - q)(k) = 0 \quad \text{for every } k \in \{0,\ldots,n\}.
\]
So, \( p - q \) is a univariate polynomial with at least \( n + 1 \) zeros and degree at most \( n \). Hence, \( p - q \) is the zero polynomial, and therefore we have \( p = q \), which implies that \( \Psi \) is injective.

In order to show that \( \Psi \) is surjective, we take any \( f \in \text{Sym} \left( \mathbb{R}^{\{0,1\}^n} \right) \), and consider \( \Phi(f) \in \mathbb{R}[x_1, \ldots, x_n] \). Lemma 3.28 shows that \( \Phi(f) \) is also symmetric. Then, by Theorem 2.21 there is a \( p \in \mathbb{R}[T]_n \) with
\[
p \left( \sum_{i \in [n]} x_i \right) = \Phi(f),
\]
and therefore
\[ f = \Phi^{-1} \left( p \left( \sum_{i \in [n]} x_i \right) \right) = p \left( \Phi^{-1} \left( \sum_{i \in [n]} x_i \right) \right) \]
\[ = p \left( \sum_{i \in [n]} \chi_{\{i\}} \right) = \Psi(p), \]
where we used that \( \Phi^{-1} \) is a ring homomorphism. Overall, \( f \) has a preimage in \( \mathbb{R}[T]_n \), which implies that \( \Psi \) is surjective.

\[ \square \]

**Remark 3.31.**
If we would replace \( \mathbb{R}[T]_n \) by \( \mathbb{R}[T]/\langle \prod_{i=0}^{n-1} (T - i) \rangle \) in \( \Psi \), we would even obtain a ring isomorphism. But for us, a vector space isomorphism will be sufficient, why we do without it for more convenience.

As an immediate consequence of the vector space isomorphism \( \Psi \), we obtain the still missing proof of the uniqueness in Theorem 2.21.

**Corollary 3.32.**
Let \( n \in \mathbb{N} \) and let \( p,q \in \mathbb{R}[T]_n \) with \( p \neq q \). Then, we also have
\[ p \left( \sum_{i \in [n]} x_i \right) \neq q \left( \sum_{i \in [n]} x_i \right). \]

**Proof.** Let \( p,q \in \mathbb{R}[T]_n \) with \( p \neq q \). Since \( \Psi \) is injective, we also obtain
\[ p \left( \sum_{i \in [n]} \chi_{\{i\}} \right) = \Psi(p) \neq \Psi(q) = q \left( \sum_{i \in [n]} \chi_{\{i\}} \right). \]
Applying \( \Phi \) results in
\[ p \left( \Phi \left( \sum_{i \in [n]} \chi_{\{i\}} \right) \right) = \Phi \left( p \left( \sum_{i \in [n]} \chi_{\{i\}} \right) \right) \neq \Phi \left( q \left( \sum_{i \in [n]} \chi_{\{i\}} \right) \right) = q \left( \Phi \left( \sum_{i \in [n]} \chi_{\{i\}} \right) \right), \]
where we used that \( \Phi \) is a ring homomorphism. But this implies the desired statement
\[ p \left( \sum_{i \in [n]} x_i \right) \neq q \left( \sum_{i \in [n]} x_i \right). \]

\[ \square \]

**Remark 3.33.**
Indeed, we used the map \( \Psi \) to finish the proof of Theorem 2.21 and we proved the isomorphism of \( \Psi \) in Proposition 3.30 on his part with the use of Theorem 2.21, but we avoid a circular reasoning. In order to finish the proof of Theorem 2.21 with the previous corollary, we just used that \( \Phi \) is injective, what we obtained independent from Theorem 2.21. Additionally, for the proof of Proposition 3.30, we just used the already proven part of Theorem 2.21.
3.4 Nonnegativity and sums of squares of Boolean functions

In the further course of this work, we are mainly interested in nonnegative real-valued Boolean functions. Since real squares are always nonnegative, the representation of a function \( f \) as a sum of squares of functions is a natural certificate for the nonnegativity of \( f \). Analogue to the Real Nullstellensatz certificate in Section 2.1, we will additionally be interested in the degree of such a certificate. We formally define this as follows.

**Definition 3.34.**
Let \( n \in \mathbb{N} \) and let \( f : \{0,1\}^n \to \mathbb{R}_+ \) be a nonnegative function. We say \( f \) has a sum of squares (sos) certificate of degree \( d \) if there exist a natural number \( s \) and functions \( g_1,\ldots,g_s \in \mathbb{R}\{0,1\}^n \) with \( \deg(g_1),\ldots,\deg(g_s) \leq \frac{d}{2} \) such that
\[
f = \sum_{i \in [s]} g_i^2.
\]
We define the sos degree of \( f \), denoted by \( \deg_{sos}(f) \), as the smallest \( d \) such that \( f \) has a sos certificate of degree \( d \).

**Remark 3.35.**
Note that a function having a sos certificate of degree \( 2k + 1 \) also always has a sos certificate of degree \( 2k \) because the degree of a Boolean function is integral. Therefore, the sos degree of a Boolean function is always even.

In contrast to polynomial functions from \( \mathbb{R}^n \) to \( \mathbb{R} \), where nonnegative or even positive functions without any sos representation exist (cf. the famous work of David Hilbert from 1888 \[Hil88\]), every nonnegative real-valued Boolean function has a sos certificate as follows.

Let \( f : \{0,1\}^n \to \mathbb{R}_+ \). Then it holds
\[
f = \sum_{x \in \{0,1\}^n} f(x)\delta_x
= \sum_{x \in \{0,1\}^n} \left( \sqrt{f(x)}\delta_x \right)^2,
\]
where we used \( f(x) \geq 0 \) and \( \delta_x = \delta_x^2 \) for all \( x \in \{0,1\}^n \). In particular, because of \( \deg(\delta_x) \leq n \), every nonnegative \( f \in \mathbb{R}\{0,1\}^n \) has a sos certificate of degree \( 2n \), and therefore
\[
\deg_{sos}(f) \leq 2n.
\]
So, not the existence but only the minimal degree of a sos certificate is interesting for nonnegative real-valued Boolean functions.

\[6\]In his proof in \[Hil88\], Hilbert only proved the existence of such functions, without giving a concrete example. The first known concrete example was detected by Motzkin \[Mot67\].
For them, a second phenomenon that does not occur in the continuous case can happen. While a $p \in \mathbb{R}[X_1, \ldots, X_n]$ with $\deg(p) = d$ can only have a sos certificate of degree $d$ as well, because degree cancellation cannot happen, for a nonnegative $f \in \mathbb{R}^{\{0,1\}^n}$ it only holds

$$\deg(f) \leq \deg_{\mathrm{sos}}(f).$$

So, even for nonnegative functions of low degree it is completely unclear if they also have a sos certificate of low degree.

In particular, the proof of Grigoriev’s Theorem 2.27 in connection with the map $\Phi$ from the previous section shows that for every $n \in \mathbb{N}$ there exists a quadratic, nonnegative function $f \in \mathbb{R}^{\{0,1\}^n}$ with $\deg_{\mathrm{sos}}(f) > n$.

Although this is just a corollary of Grigoriev’s Theorem or more precisely of its proof, we state it as an theorem on its own because of the significance for our further work.

**Theorem 3.36.**

Let $n \in \mathbb{N}$ with $n \geq 2$ and $r \in \left(\left[\frac{n-3}{2}, \frac{n+3}{2}\right]\right)$ with $r \notin \mathbb{N}$, and let $b \in \mathbb{R}$ be defined as $b := \min_{k \in \mathbb{N}}(k - r)^2$. Let $c \in (0, b]$, and let $f \in \mathbb{R}^{\{0,1\}^n}$ be the quadratic function

$$f = \left(\sum_{i \in [n]} \chi_{\{i\}} - r\right)^2 - c.$$

Then, $f$ is nonnegative and

$$\deg_{\mathrm{sos}}(f) > n.$$

Moreover, for even $n$ we actually have $\deg_{\mathrm{sos}}(f) \geq n + 2$.

**Proof.** By the definition of $b$, we have

$$b \leq \left(\sum_{i \in [n]} x_i - r\right)^2$$

for every $x \in \{0,1\}^n$. Because of $c \leq b$, this implies

$$f(x) = \left(\sum_{i \in [n]} x_i - r\right)^2 - c \geq b - c \geq 0$$

for every $x \in \{0,1\}^n$, and therefore $f$ is nonnegative.

Assume $\deg_{\mathrm{sos}}(f) \leq n$. Then, there exist $s \in \mathbb{N}$ and $g_1, \ldots, g_s \in \{0,1\}^n$ with $\deg(g_1), \ldots, \deg(g_s) \leq \left\lfloor \frac{n}{2}\right\rfloor$ such that

$$\left(\sum_{i \in [n]} \chi_{\{i\}} - r\right)^2 - c = \sum_{i \in [s]} g_i^2. \tag{3.14}$$
Applying the ring isomorphism $\Phi : \mathbb{R}^{(0,1)^n} \to \mathbb{R}[x_1, \ldots, x_n]$ defined in the previous section (cf. Notation 3.26) yields
\[
\left(\sum_{i \in [n]} x_i - r\right)^2 - c = \sum_{i \in [s]} \Phi(g_i)^2, \tag{3.15}
\]
where $\deg \Phi(g_i) \leq \lfloor \frac{n}{2} \rfloor$ for all $i \in [s]$ because $\Phi$ preserves the degree.

But this is a contradiction to the proof of Grigoriev’s Theorem 2.27 because our linear form $L_r$ we defined there equals $0$ evaluated to $\left(\sum_{i \in [n]} x_i - r\right)^2$, and therefore it equals $-c$ evaluated to the left-hand side of (3.15), while $L_r$ evaluated to the right-hand side of (3.15) is greater or equal $0$, what follows by (2.24). In particular, note that the linear form $L_r \circ \Phi$ evaluates to a nonnegative value if evaluated in the nonnegative function $f$.

So, nonnegativity is not sufficient in order to certify a low degree sos representation. However, for a symmetric $f \in \mathbb{R}^{(0,1)^n}$, applying the map $\Phi : \mathbb{R}^{(0,1)^n} \to \mathbb{R}[x_1, \ldots, x_n]$ from the last section to Blekherman’s Theorem 2.26 yields a sufficient and even necessary condition for a low degree sos certificate of $f$ via the representation of $\Psi^{-1}(f)$, where $\Psi$ is the recently introduced isomorphism from $R[T]^n$ to $\text{Sym}(\mathbb{R}^{(0,1)^n})$.

We formally state such an analogon of Blekherman’s Theorem for real-valued Boolean functions. This theorem completely characterizes the preimage of symmetric sums of low degree squares under $\Psi$.

**Theorem 3.37** (Blekherman 2015).

Let $n,d \in \mathbb{N}$ with $d$ even and $d \leq n$. Furthermore, let $f : \{0,1\}^n \to \mathbb{R}_+$ be symmetric. Then, it holds
\[
\deg_{\text{sos}}(f) \leq d,
\]
meaning $f$ is a sum of squares of functions of degree at most $\frac{d}{2}$, if and only if
\[
\Psi^{-1}(f) = \sum_{k=0}^{\frac{d}{2}} p_k \prod_{i=0}^{k-1} (T - i)(n - i - T)
\]
for some $p_0, \ldots, p_{\frac{d}{2}} \in \mathbb{R}[T]$ with $p_k \in \sum \mathbb{R}[T]_{2-k}$ for every $k \in \{0, \ldots, \frac{d}{2}\}$.

**Proof.** By definition and because $\Phi$ is a ring isomorphism that preserves the degree, we obtain
\[
\deg_{\text{sos}}(f) \leq d \text{ if and only if } \Phi(f) \in \sum R[x]_{\frac{d}{2}}^2.
\]
Because of Blekherman’s Theorem 2.26, which we can apply because $\Phi(f)$ is also symmetric by Lemma 3.28, the latter holds if and only if there exist $p_0, \ldots, p_{\frac{d}{2}} \in \mathbb{R}[T]^n$ with $p_k \in \sum \mathbb{R}[T]_{2-k}$ for every $k \in \{0, \ldots, \frac{d}{2}\}$ such that
\[
\Phi(f) = \sum_{k=0}^{\frac{d}{2}} p_k \prod_{i=0}^{k-1} (T - i)(n - i - T),
\]
where \( \Phi(f) \in \mathbb{R}[T]^n \) is the unique polynomial from Theorem 2.21 satisfying
\[
\Phi(f) \left( \sum_{i \in [n]} x_i \right) = \Phi(f).
\] (3.16)

Now, it remains to show
\[
\Phi(f) = \Psi^{-1}(f).
\]
But this is valid because applying \( \Phi^{-1} \) to (3.16) yields
\[
f = \Phi(f) \left( \sum_{i \in [n]} \chi_{\{i\}} \right) = \Psi \left( \Phi(f) \right).
\]
Since \( \Psi \) is injective (cf. Theorem 3.30), the statement follows.

Later, we will apply this theorem several times, and therefore we introduce the following plausible notation.

**Notation 3.38.** Let \( n \in \mathbb{N} \) and \( d \in [n] \). Then, we write
\[
\sum_{k=0}^{d} \prod_{j=0}^{k-1} (T-j)(n-j-T) \sum \mathbb{R}[T]_{d-k}^2 := \left\{ \sum_{k=0}^{d} p_k \prod_{j=0}^{k-1} (T-j)(n-j-T) \right| p_k \in \sum \mathbb{R}[T]_{d-k}^2 \right\}.
\]

Note that \( \deg(f) \leq 2d \), for every \( f \in \sum_{k=0}^{d} \prod_{j=0}^{k-1} (T-j)(n-j-T) \sum \mathbb{R}[T]_{d-k}^2 \).

### 3.5 Pseudo-expectations and pseudo-densities

In the situation of Theorem 3.36, the linear form \( L := L_r \circ \Phi \) caused a contradiction in (3.14). It is a linear form \( L : \mathbb{R}^{[0,1]^n} \rightarrow \mathbb{R} \) with \( L(g^2) \geq 0 \) for all \( g \in \mathbb{R}^{[0,1]^n} \) with \( \deg(g) \leq \frac{k}{2} \) and \( L(f) < 0 \).

More general, linear forms from \( \mathbb{R}^{[0,1]^n} \) to \( \mathbb{R} \) with the property
\[
L(g^2) \geq 0 \quad \text{for all} \quad g \in \mathbb{R}^{[0,1]^n} \quad \text{with} \quad \deg(g) \leq \frac{k}{2}
\] (3.17)
for some fixed \( k \in [2n] \) will play an important role for us.

Every function \( f \in \mathbb{R}^{[0,1]^n} \) can be seen as a real random variable on \( \{0,1\}^n \). Therefore, for such a linear form \( L : \mathbb{R}^{[0,1]^n} \rightarrow \mathbb{R} \) with the properties (3.17) and \( L(1) = 1 \), the quantity \( L(f) \) can be interpreted as a kind of an expectation of \( f \).

Of course, in Theorem 3.36 we saw that even for a nonnegative \( f \) the quantity \( L(f) \) can be negative. Therefore, such a linear form is not a real expectation. It surely coincides with the behavior of an expectation only on nonnegative functions \( g \) with a low-degree sos certificate, strictly speaking only for \( g \in \mathbb{R}^{[0,1]^n} \) with \( \deg_{sos}(g) \leq k \). For this reason, we will speak of a pseudo-expectation in this case.
Definition 3.39.
Let $n \in \mathbb{N}$ and $k \in [2^n]$. A linear form $L : \mathbb{R}^{\{0,1\}^n} \to \mathbb{R}$ that satisfies $L(1) = 1$ and

$$L(g) \geq 0 \text{ for all } g \in \mathbb{R}^{\{0,1\}^n} \text{ with } \text{deg}_{\text{sos}}(g) \leq k,$$

is called a degree-$k$ pseudo-expectation.

Remark 3.40.
Note that because of the linearity, for every degree-$k$ pseudo-expectation $L$ it is equivalent to require only (3.17), which in general is a weaker statement than the above one.

Note that for an arbitrary $f \in \mathbb{R}^{\{0,1\}^n}$, the existence of a degree-$k$ pseudo-expectation $L$ with the additional property $L(f) < 0$ yields a certificate for $\text{deg}_{\text{sos}}(f) > k$.

Even more, the converse also holds. If $\text{deg}_{\text{sos}}(f) > d$, there exists a degree-$d$ pseudo-expectation $L : \mathbb{R}^{\{0,1\}^n} \to \mathbb{R}$ with $L(f) < 0$.

This follows by the Separation Theorem, a well-known result in convex geometry that can be found for example in [Bar02, Theorem III.1.3].

Definition/Example 3.41.
Let $n \in \mathbb{N}$ and $r \in \mathbb{R}$. Furthermore, let $L_r : \mathbb{R}[x_1, \ldots, x_n] \to \mathbb{R}$ the linear form from our proof of Grigoriev's Theorem 2.27 defined in (2.21), and let $\Phi : \mathbb{R}^{\{0,1\}^n} \to \mathbb{R}[x_1, \ldots, x_n]$ the ring isomorphism as before. Remember that $L_r$ is exactly the linear form Grigoriev used in his original proof (cf. Corollary 2.31). This motivates the following definition.

We define $L_{\text{Gr}, r} : \mathbb{R}^{\{0,1\}^n} \to \mathbb{R}$ by

$$L_{\text{Gr}, r} := L_r \circ \Phi,$$

and call it the Grigoriev pseudo-expectation.

For example, for $r \in \left(\left\lfloor \frac{n-3}{2} \right\rfloor, \left\lceil \frac{n+3}{2} \right\rceil \right)$, $L_{\text{Gr}, r}$ is actually a degree-$n$ pseudo-expectation, what immediately follows by our proof of Grigoriev's Theorem and because $\Phi$ preserves the degree.

In the case $r = \frac{n}{2}$, we just write

$$L_{\text{Gr}} := L_{\text{Gr}, r} = L_{\text{Gr}, \frac{n}{2}}$$

in the further course of the work.

An actual expectation only depends on an underlying distribution. On the other hand, a distribution on $\{0,1\}^n$ corresponds to a density $D$ with respect to the uniform measure on $\{0,1\}^n$. That means the expectation of a random variable $f$ is calculated from

$$\mathbb{E}_x D(x) f(x).$$
We will see that there is the same relation in the case of pseudo-expectations. The corresponding densities will be called pseudo-densities. Instead of working with pseudo-expectations, it turns out that sometimes it is more convenient and practical to use their corresponding pseudo-densities. Especially in Chapter 4, they turn out to be very useful. We introduce their notion formally.

**Definition 3.42.**

Let $n \in \mathbb{N}$ and $d \in [2^n]$. A function $D : \{0,1\}^n \rightarrow \mathbb{R}$ is called *degree-$d$ pseudo-density* if it satisfies the following two conditions.

(i) \[ \mathbb{E}_x D(x) = 1, \]

(ii) \[ \mathbb{E}_x D(x)g(x)^2 \geq 0 \]

for all $g \in \mathbb{R}^{\{0,1\}^n}$ with $\deg(g) \leq \frac{d}{2}$.

**Remark 3.43.**

(a) Similar to the sos degree it is sufficient to consider degree-$d$ pseudo-densities for even $d$ because every degree-$2k$ pseudo-density is also a degree-$(2k + 1)$ pseudo-density.

(b) Because of the linearity of the expectation, condition (ii) in the above definition implies

\[ \mathbb{E}_x D(x)g(x) \geq 0 \]

for all $g \in \mathbb{R}^{\{0,1\}^n}$ with $\deg_{\text{sos}}(g) \leq d$.

Therefore, every degree-$d$ pseudo-density $D \in \mathbb{R}^{\{0,1\}^n}$ obviously induces a degree-$d$ pseudo-expectation $L : \mathbb{R}^{\{0,1\}^n} \rightarrow \mathbb{R}$ via

\[ L(f) := \mathbb{E}_x D(x)f(x). \]

Note that for every function $D : \{0,1\}^n \rightarrow \mathbb{R}$ and every $a \in \{0,1\}^n$ the following is valid.

\[ \mathbb{E}_x D(x)\delta_a(x) = \frac{1}{2^n} D(a). \tag{3.18} \]

This simple but nice relation shows that the mapping $D \mapsto \left\{ f \mapsto \mathbb{E}_x D(x)f(x) \right\}$ is injective. In particular, two different degree-$d$ pseudo-densities induce two different linear forms.

The above equation also shows how to obtain a corresponding density from a given linear from, and therefore we can show that there is a one-to-one correspondence between pseudo-densities and pseudo-expectations.
Proposition 3.44.
Let $n \in \mathbb{N}$, $d \in [2n]$, and let $M := \{ D \in \mathbb{R}^{(0,1)^n} \mid D \text{ is a degree-}d \text{ pseudo-density} \}$ and $N := \{ L : \mathbb{R}^{(0,1)^n} \rightarrow \mathbb{R} \mid L \text{ is a degree-}d \text{ pseudo-expectation} \}$. Then, the mapping

$$M \rightarrow N, \; D \mapsto \left\{ L : \mathbb{R}^{(0,1)^n} \rightarrow \mathbb{R} \bigg\vert \sum_{x} D(x) f(x) \right\}$$

is a bijection.

Proof. As already mentioned, for a degree-$d$ pseudo-density it obviously follows that $L : \mathbb{R}^{(0,1)^n} \rightarrow \mathbb{R}, f \mapsto \mathbb{E}D(x)f(x)$ is a degree-$d$ pseudo-expectation. Thus, the above defined mapping is well-defined. The injectivity follows by (3.18).

The interesting part is that every $L \in N$ is induced by a degree-$d$ pseudo-density. In order to show this, to a given $L \in N$ we define

$$D : \{0,1\}^n \rightarrow \mathbb{R}, \; x \mapsto 2^n L(\delta_x).$$

With this definition, we obtain

$$L(f) = L \left( \sum_{x \in \{0,1\}^n} f(x) \delta_x \right) = \sum_{x \in \{0,1\}^n} f(x) L(\delta_x)$$

$$= \sum_{x \in \{0,1\}^n} f(x) \frac{1}{2^n} D(x)$$

$$= \mathbb{E}D(x)f(x),$$

for every $f \in \mathbb{R}^{(0,1)^n}$. Therefore, it remains to show that $D$ is indeed a degree-$d$ pseudo-density.

We have

$$\mathbb{E}D(x) = \mathbb{E}2^n L(\delta_x) = \sum_{x \in \{0,1\}^n} L(\delta_x)$$

$$= L \left( \sum_{x \in \{0,1\}^n} \delta_x \right) = L(1)$$

$$= 1$$

and

$$\mathbb{E}D(x)g(x)^2 = \sum_{x \in \{-1,1\}^n} L(\delta_x)g(x)^2$$

$$= L \left( \sum_{x \in \{0,1\}^n} g(x)^2 \delta_x \right) = L(g^2)$$

$$\geq 0$$

for all $g \in \{0,1\}^n$ with $\deg(g) \leq \frac{d}{2}$, which implies that $D$ is a degree-$d$ pseudo-density.
Together with our observations after Remark 3.40, we immediately obtain the following small, nice result.

**Corollary 3.45.**

Let \( n \in \mathbb{N} \), \( d \in [2n] \) and \( f \in \mathbb{R}^{\{0,1\}^n} \). Then,

\[
\deg_{\text{sos}}(f) > d
\]

if and only if there exists a degree-\( d \) pseudo density \( D \in \mathbb{R}^{\{0,1\}^n} \) with

\[
\mathbb{E}_x D(x)f(x) < 0.
\]

The above proposition allows to switch arbitrarily between pseudo-expectations and pseudo-densities depending on what is more practical for us in the current situation. Especially in Chapter 5, we will make frequent use of this.

We formally define the notion of correspondence, due to the above proposition.

**Definition 3.46.**

Let \( n \in \mathbb{N} \) and \( d \in [2n] \). To a given degree-\( d \) pseudo-Density \( D \in \mathbb{R}^{\{0,1\}^n} \), we call \( L : \mathbb{R}^{\{0,1\}^n} \rightarrow \mathbb{R}, f \mapsto \mathbb{E}_x D(x)f(x) \) the corresponding degree-\( d \) pseudo-expectation. Vice versa, to a given degree-\( d \) pseudo-expectation \( L : \mathbb{R}^{\{0,1\}^n} \rightarrow \mathbb{R} \), we call the function \( D : \{0,1\}^n \rightarrow \mathbb{R}, x \mapsto 2^n L(\delta_x) \) or equivalent the unique function \( D : \{0,1\}^n \rightarrow \mathbb{R} \) with

\[
L(f) = \mathbb{E}_x D(x)f(x)
\]

for all \( f \in \mathbb{R}^{\{0,1\}^n} \)

the corresponding degree-\( d \) pseudo-density.

**Remark 3.47.**

(a) Note that a nonnegative function \( D : \{0,1\}^n \rightarrow \mathbb{R}_+ \) with \( \mathbb{E} D(x) = 1 \) is always a degree-\( d \) pseudo-density for all \( d \in [n] \). In this case, it is even an actual density with respect to the uniform measure on \( \{0,1\}^n \) and the corresponding pseudo-expectation is an actual expectation.

(b) The notion of a pseudo-density is also valid for real-valued Boolean functions \( D : \{-1,1\}^n \rightarrow \mathbb{R} \), where in condition (ii) of Definition 3.42 the function \( g \) has to be in \( \mathbb{R}^{\{-1,1\}^n} \) of course. This case is not as important for us than the case of the domain \( \{0,1\}^n \), but we will use it in Section 4.3.

(c) Note that for the bijection \( \psi : \{0,1\}^n \rightarrow \{-1,1\}^n \) from Notation 3.18 and every degree-\( d \) pseudo-density \( D \in \mathbb{R}^{\{0,1\}^n} \), the function \( \hat{D} : \{-1,1\}^n \rightarrow \mathbb{R}, x \mapsto D(\psi^{-1}(x)) \) is also a degree-\( d \) pseudo-density. This is valid because \( \psi \) is a bijection and because \( \hat{\psi} \) as defined in Notation 3.19 preserves the degree of a Boolean function.
We will finish this section with some observations about symmetric pseudo-densities. We will show that in this case also the corresponding linear forms have some nice property. This will allow us to restrict the linear form on symmetric Boolean functions, what in many cases, due to Proposition 3.30, will be a nice simplification. In this connection, in Chapter 5 it turns out that the requirement for symmetric pseudo-densities is not an excessive restriction.

First, we show that symmetric pseudo-densities correspond to pseudo-expectations that have constant values on orbits of $\mathbb{R}^{(0,1)^n}$ under the action of $S_n$ and vice versa.

**Lemma 3.48.**
Let $n \in \mathbb{N}$ and $d \in [2n]$.

(i) Let $D \in \mathbb{R}^{(0,1)^n}$ be a symmetric degree-$d$ pseudo density. Furthermore, let $L : \mathbb{R}^{(0,1)^n} \to \mathbb{R}$ be the corresponding pseudo-expectation. Then, $L$ satisfies

$$L(f) = L(f^\sigma)$$

for every $f \in \mathbb{R}^{(0,1)^n}$ and every $\sigma \in S_n$.

(ii) Let $L : \mathbb{R}^{(0,1)^n} \to \mathbb{R}$ be a degree-$d$ pseudo-expectation with $L(f) = L(f^\sigma)$ for every $f \in \mathbb{R}^{(0,1)^n}$ and every $\sigma \in S_n$. Then the corresponding degree-$d$ pseudo-density is symmetric.

**Proof.**

(i) For every $f \in \mathbb{R}^{(0,1)^n}$ and every $\sigma \in S_n$, we have

$$L(f^\sigma) = \mathbb{E}_x D(x) f^\sigma(x) = \mathbb{E}_x D(x) f(x^\sigma)$$

$$= \mathbb{E}_x D(x^\sigma) f(x^\sigma) = \mathbb{E}_x D(x) f(x)$$

$$= L(f).$$

The first and the last step is valid because $L$ is the corresponding linear form of $D$, the second step is just the definition of $f^\sigma$. The third step is valid because $D$ is symmetric, and in the second last step we used (3.12).

(ii) Note that $\delta_x^\sigma = \delta_{x^\sigma}$ for all $x \in \{0,1\}^n$ and every $\sigma \in S_n$. Besides, by the assumption we have $L(\delta_x^{\sigma^{-1}}) = L(\delta_x)$. Overall, for all $x \in \{0,1\}^n$ and every $\sigma \in S_n$ it holds

$$D(x^\sigma) = 2^n L(\delta_x^\sigma) = 2^n L(\delta_x^{\sigma^{-1}})$$

$$= 2^n L(\delta_x) = D(x).$$

A linear form with the property of the above lemma is determined by its values on the symmetric functions.
Lemma 3.49.
Let $n \in [n]$ and let $L : \mathbb{R}^{(0,1)^n} \to \mathbb{R}$ be a linear form with $L(f) = L(f^\sigma)$ for all $f \in \mathbb{R}^{(0,1)^n}$ and all $\sigma \in S_n$. Then,

$$L(f) = L(\text{Sym}(f))$$

for all $f \in \mathbb{R}^{(0,1)^n}$.

Proof. For every $f \in \mathbb{R}^{(0,1)^n}$, it holds

$$L(\text{Sym}(f)) = L \left( \frac{1}{n!} \sum_{\sigma \in S_n} f^\sigma \right) = \frac{1}{n!} \sum_{\sigma \in S_n} L(f^\sigma) = \frac{1}{n!} \sum_{\sigma \in S_n} L(f) = L(f).$$

Taking Proposition 3.44, Lemma 3.48 and Lemma 3.49 together yields the following corollary.

Corollary 3.50.
Let $n \in \mathbb{N}$ and $d \in [2n]$. Then the map $D \mapsto \left\{ L : \text{Sym} \left( \mathbb{R}^{(0,1)^n} \right) \to \mathbb{R} \right\}$ yields a one-to-one correspondence between symmetric degree-$d$ pseudo densities and linear forms from $\text{Sym} \left( \mathbb{R}^{(0,1)^n} \right)$ to $\mathbb{R}$ with

$$L(1) = 1$$

and

$$L(f) \geq 0 \text{ for all } f \in \text{Sym} \mathbb{R}^{(0,1)^n} \text{ with } \deg_{\text{SOS}}(f) \leq d. \quad (3.19)$$

Remark 3.51.
In contrast to Remark 3.40, in this case we cannot reduce the property to $L(g^2) \geq 0$ for all $g \in \text{Sym} \left( \mathbb{R}^{(0,1)^n} \right)$ with $\deg(g) \leq \frac{d}{2}$ because not every symmetric sum of squares of Boolean functions is a sum of symmetric squares.

Because of the isomorphism $\Psi : \mathbb{R}[T]_n \to \text{Sym} \left( \mathbb{R}^{(0,1)^n} \right)$ from section 3.3, we additionally have a canonical one-to-one correspondence between linear forms from $\text{Sym} \left( \mathbb{R}^{(0,1)^n} \right)$ to $\mathbb{R}$ and linear forms from $\mathbb{R}[T]_n$ to $\mathbb{R}$ as follows. Every linear form $L : \text{Sym} \left( \mathbb{R}^{(0,1)^n} \right) \to \mathbb{R}$ induces a linear form $\mathcal{L} : \mathbb{R}[T]_n \to \mathbb{R}$ via

$$\mathcal{L} = L \circ \Psi,$$
and analog every linear form $\mathcal{L} : \mathbb{R}[T]_n \to \mathbb{R}$ induces a linear form $L : \text{Sym} \left( \mathbb{R}^{(0,1)^n} \right) \to \mathbb{R}$ via

$$L = \mathcal{L} \circ \Psi^{-1}.$$  

Combining these observations with the previous corollary yields a one-to-one correspondence between symmetric pseudo-densities in $\mathbb{R}^{(0,1)^n}$ and linear forms from $\mathbb{R}[T]_n$ to $\mathbb{R}$. How property (3.19) is translated in this connection is exactly given by Theorem 3.37.

**Corollary 3.52.**

Let $n \in \mathbb{N}$ and $d \in [2n]$. Then the map $D \mapsto \left\{ \mathcal{L} : \mathbb{R}[T]_n \to \mathbb{R} \right\}$ yields a one-to-one correspondence between symmetric degree-$d$ pseudo-densities in $\mathbb{R}^{(0,1)^n}$ and linear forms from $\mathbb{R}[T]_n$ to $\mathbb{R}$ with

$$\mathcal{L}(1) = 1$$

and

$$\mathcal{L}(f) \geq 0 \text{ for all } f \in \sum_{k=0}^{d} \prod_{j=0}^{k-1} (T - j)(n - j - T) \sum \mathbb{R}[T]^2_{2-k}. \quad (3.20)$$

**Remark 3.53.**

The corresponding pseudo-density to a pseudo-expectation $L \in \text{Sym} \left( \mathbb{R}^{(0,1)^n} \right)$ (cf. Corollary 3.50) is obviously also the corresponding pseudo-density to $L \circ \Psi$ from the above Corollary.

The previous result allows us to work with linear forms from the univariate polynomial ring if we are interested in symmetric pseudo-densities what will be the case in Chapter 5. Therefore, it is very important for the further course of our work because this facilitates the analysis of symmetric pseudo-densities and yields some nice results as we will see later.

At first sight, condition (3.20) seems to be a disadvantage compared to (3.19) because it seems more complicated, but the opposite will be the case. If we start with a linear form satisfying (3.20), the structure of the condition yields many different single conditions and we will benefit from its strength.

Since we will permanent use such linear forms in Chapter 5, we state the following definition.

**Definition 3.54.**

Let $n \in \mathbb{N}$, $r \in \mathbb{R}$, and let $L_{\bar{G},r} : \text{Sym} \left( \mathbb{R}^{(0,1)^n} \right) \to \mathbb{R}$ be the restriction of the Grigoriev pseudo-expectation, defined in Definition 3.41, to $\text{Sym} \left( \mathbb{R}^{(0,1)^n} \right)$. Then we define the linear form $\mathcal{L}_{\bar{G},r} : \mathbb{R}[T]_n \to \mathbb{R}$ by

$$\mathcal{L}_{\bar{G},r} := L_{\bar{G},r} \circ \Psi,$$
and call it *Grigoriev linear form*.

The similar name is justified because the two linear forms only differ in an isomorphism. Confusions are excluded because we will only consider linear forms from \( \mathbb{R}[T]_n \) to \( \mathbb{R} \) in future.

Furthermore, in the case \( r = \frac{n}{2} \), we just write

\[
\mathcal{L}_g := \mathcal{L}_{g,r} = \mathcal{L}_{g,\frac{n}{2}}
\]

in the further course of the work.

It turns out that the just defined Grigoriev linear form \( \mathcal{L}_{g,r} \) is a very easy linear form. Actually, \( \mathcal{L}_{g,r} \) is just the evaluation in \( r \).

**Lemma 3.55.**

Let \( n \in \mathbb{N} \), \( r \in \mathbb{R} \) and let \( \mathcal{L}_{g,r} : \mathbb{R}[T]_n \rightarrow \mathbb{R} \) the Grigoriev linear form. Then, we have

\[
\mathcal{L}_{g,r}(p) = p(r)
\]

for every \( p \in \mathbb{R}[T]_n \).

**Proof.** Note that \( p \left( \sum_{i \in [n]} x_i \right) \) is symmetric for every \( p \in \mathbb{R}[T]_n \), and therefore we have

\[
\text{Sym} \left( p \left( \sum_{i \in [n]} x_i \right) \right) = p \left( \sum_{i \in [n]} x_i \right) \quad \text{for every } p \in \mathbb{R}[T]_n. \quad (3.21)
\]

Additionally, it holds

\[
p \left( \sum_{i \in [n]} x_i \right) = p \quad (3.22)
\]

by definition and the uniqueness of Theorem 2.21.

Overall, we have

\[
\mathcal{L}_{g,r}(p) = L_{g,r} \left( \Psi(p) \right) = L_{g,r} \left( p \left( \sum_{i \in [n]} \chi_{\{i\}} \right) \right) \\
= L_r \circ \Phi \left( p \left( \sum_{i \in [n]} \chi_{\{i\}} \right) \right) = L_r \left( p \left( \sum_{i \in [n]} x_i \right) \right) \\
= \text{Sym} \left( p \left( \sum_{i \in [n]} x_i \right) \right) (r) \overset{3.21}{=} p \left( \sum_{i \in [n]} x_i \right) (r) \overset{3.22}{=} p(r)
\]

for every \( p \in \mathbb{R}[T]_n \). \( \square \)

After having an easy representation of the Grigoriev linear form, we are now interested in a representation of the corresponding symmetric pseudo-density. [LRST15] give a representation of such a pseudo-density en route of the proof of [LRST15, Theorem 5.3] for the special case \( r = \frac{n}{2} \) without motivating it. They just constructed a function out of the blue and showed that it satisfies the condition of a corresponding pseudo-density without deducing it from the linear form.
In contrast, we want to deduce a representation of the corresponding symmetric pseudo-density from the linear form and for every \( r \in \mathbb{R} \). Therefore, we first show how to obtain the corresponding symmetric pseudo-density for a general linear form \( L : \mathbb{R}[T]_n \to \mathbb{R} \).

Note that in our initial situation \( L : \mathbb{R}^{(0,1)^n} \to \mathbb{R} \), the trick to obtain the corresponding pseudo-density was evaluating \( L \) in the needle functions. This does not work in our setting now. But there are again special univariate polynomials that yields the values of the corresponding pseudo-density if we insert them into the linear form.

**Lemma 3.56.**

Let \( n \in \mathbb{N} \) and let \( L : \mathbb{R}[T]_n \to \mathbb{R} \) be a linear form. Then, the corresponding symmetric pseudo-density \( D \in \mathbb{R}^{(0,1)^n} \) is given by

\[
D(x) = \frac{2^n}{\binom{n}{|x|}} \cdot \frac{L \left( \prod_{j=0 \atop j \neq |x|}^{n} (T - j) \right)}{\prod_{j=0 \atop j \neq |x|}^{n} (|x| - j)}
\]

for every \( x \in \{0,1\}^n \).

**Proof.** Let \( D \in \mathbb{R}^{(0,1)^n} \) be the corresponding symmetric pseudo-density to \( L \). By previous observations and definitions, \( D \) satisfies

\[
\mathbb{E} D(x)f(|x|) = L(f)
\]

for every \( f \in \mathbb{R}[T]_n \).

For \( k \in \{0, \ldots, n\} \) consider the polynomial \( f_k = \prod_{j=0 \atop j \neq k}^{n} (T - j) \in \mathbb{R}[T]_n \). Note that we have

\[
f_k(|x|) = 0 \text{ for every } x \in \{0,1\}^n \text{ with } |x| \neq k. \tag{3.23}
\]

Now, we fix \( x_0 \in \mathbb{R}^{(0,1)^n} \). For \( k := |x_0| \), we have

\[
L(f_k) = \mathbb{E} D(x)f_k(|x|) \stackrel{\text{3.23}}{=} \frac{1}{2^n} \sum_{x \in \{0,1\}^n \atop |x| = k} D(x)f_k(k)
\]

\[
= \frac{1}{2^n} \binom{n}{k} D(x_0)f(|x_0|),
\]

where we use the symmetry of \( D \) in the last step. Then, we obtain

\[
D(x_0) = \frac{2^n}{\binom{n}{|x_0|}} \cdot \frac{L(f_{|x_0|})}{\prod_{j=0 \atop j \neq |x_0|}^{n} (|x_0| - j)},
\]

which implies the statement. \( \square \)
**Definition 3.57.**
Let \( n \in \mathbb{N}, r \in \mathbb{R} \), and let \( L_{\mathcal{E},r} : \mathbb{R}^{\{0,1\}^n} \to \mathbb{R} \) be the Grigoriev pseudo-expectation, and let \( L_{\mathcal{E},r} : \mathbb{R}[T]_n \to \mathbb{R} \) be the Grigoriev linear form. Then, we denote its common corresponding symmetric pseudo-density in \( \mathbb{R}^{\{0,1\}^n} \) by \( D_{\mathcal{E},r} \), and call it the **Grigoriev pseudo-density**.

Similar as before, in the case \( r = \frac{n}{2} \), we just write
\[
D_{\mathcal{E}} := D_{\mathcal{E},r} := D_{\mathcal{E}, \frac{n}{2}}.
\]

**Remark 3.58.**
Let \( n \in \mathbb{N} \) and \( r \in \left( \left\lceil \frac{n-3}{2} \right\rceil, \left\lfloor \frac{n+3}{2} \right\rfloor \right) \). Then, the Grigoriev pseudo-density \( D_{\mathcal{E},r} \in \mathbb{R}^{\{0,1\}^n} \) is a degree-\( n \) pseudo-density. This easily follows by combining Corollary 3.52 with Lemma 3.55. Another possibility is to apply Example 3.41 together with Corollary 3.50 and the proof of Grigoriev’s Theorem.

The previous lemma, together with Lemma 3.55, immediately yields the following corollary.

**Corollary 3.59.**
Let \( n \in \mathbb{N} \) and \( r \in \mathbb{R} \). Then, we have
\[
D_{\mathcal{E},r}(x) = 2^n \frac{n}{|x|} \prod_{j \neq |x|} (r - j)
\]
for every \( x \in \{0,1\}^n \).
4 A lower bound on the psd rank of the family of correlation polytopes

First of all, we remind the definition of the psd rank of a polytope from the introduction.

**Definition 4.1.**
Let \( n,k \in \mathbb{N} \), and let \( P \subseteq \mathbb{R}^n \) be a polytope. We say \( P \) admits a *psd lift* of size \( k \) if there exists \( N \in \mathbb{N} \) with \( N \geq n \), a spectrahedron \( S \subseteq \mathbb{R}^N \) defined by a linear matrix inequality of size \( k \) and a linear map \( \pi : R^N \rightarrow \mathbb{R}^n \) such that
\[
P = \pi(S).
\]

We define the *psd rank* of \( P \), denoted by \( \text{rk}_{\text{psd}}(P) \) as the minimal \( k \) such that \( P \) admits a psd lift of size \( k \).

Initially, it seems unnatural or even wrong to use the word rank in this case, but we will see later in Theorem 4.4 that this actually makes sense.

In their recent breakthrough work [LRS15], Lee, Raghavendra and Steurer showed that \( \text{rk}_{\text{psd}}(\text{CORR}_n) \) grows faster than any polynomial in \( n \). In fact, that means
\[
\lim_{n \to \infty} \frac{\text{rk}_{\text{psd}}(\text{CORR}_n)}{n^k} = \infty \tag{4.1}
\]
for every \( k \in \mathbb{N} \).

In a more precise analysis, they showed that there exist a universal constant \( \alpha > 0 \) and some \( N \in \mathbb{N} \) with
\[
\text{rk}_{\text{psd}}(\text{CORR}_n) \geq 2^{\alpha \left( \frac{n}{\ln n} \right)^{2/13}}, \tag{4.2}
\]
for every \( n > N \). Their work proved the first-known super-polynomial lower bounds on the size of psd lifts of any explicit family of polytopes.

To provide a detailed proof of (4.1) and (4.2) is the goal of this chapter. In this context, we illustrate why we can improve the exponent in their bound from \( \frac{2}{13} \) to \( \frac{2}{11} \). That means, instead of (4.2) we even prove the existence of a universal constant \( \alpha > 0 \) and some \( N \in \mathbb{N} \) with
\[
\text{rk}_{\text{psd}}(\text{CORR}_n) \geq 2^{\alpha \left( \frac{n}{\ln n} \right)^{2/11}}, \tag{4.3}
\] 7Independently from us, this was also noticed by Troy Lee et al. [LPWY16]
for every $n > N$.

As already mentioned in the introduction, our main intention is to present a detailed version of the work of [LRS15] and in the ideal case a version that makes it easier to understand their brilliant work. In this connection, it is our motivation to bring their work to a broader number of people without much previous knowledge. On the other hand, we present some own ideas and implementations how their proof could be simplified and how it can be seen from a more real algebraic geometry point of view. At some points, with our ideas we even obtain some slightly stronger results than [LRS15].

We will first provide some general ideas how to handle the psd rank of polytopes (Section 4.1), and afterwards, we will give some motivation and explanations of the different basic steps and ideas used by [LRS15] (Section 4.2). The actual proof is provided in the Sections 4.3 - 4.7 where we will elaborate the proof of [LRS15] and embellish it with details they omitted.

### 4.1 Slack matrices and nonnegative, quadratic Boolean functions - basic steps to handle the psd rank of the correlation polytope

First of all, it eventually seems unclear how to handle the geometric property of the psd rank of a polytope. For this reason, we introduce the definition of the psd rank of a (nonnegative) matrix, the object for what the word rank is used in the normal case. This is a more algebraic concept, which is easier to work with. Afterwards we cite a theorem which connects the two types of psd rank.

**Definition 4.2.**

Let $s, t \in \mathbb{N}$, and let $M \in \mathbb{R}^{s \times t}$ be a matrix with nonnegative entries. We say $M$ admits a rank-$r$ psd factorization if for every $i \in [s]$ and every $j \in [t]$ there exist positive semidefinite matrices $A_i, B_j \in \mathbb{S}^r$ such that $M_{i,j} = \langle A_i, B_j \rangle = \text{Tr}(A_i B_j)$ for all $i \in [s], j \in [t]$. We define the psd rank of $M$ as the smallest $r$ such that $M$ admits a rank-$r$ psd factorization, and we denote it by $\text{rk}_{\text{psd}}(M)$.

After introducing the psd rank of a matrix, in order to apply this definition we need a connection from a polytope to a matrix. There is such a natural connection called the slack matrix of a polytope and motivated by two different types of representing a polytope.

For every polytope $P \subseteq \mathbb{R}^n$, by definition, there always exist a number $t \in \mathbb{N}$ and a finite subset $V = \{v_1, \ldots, v_t\} \subseteq P$ such that $P = \text{conv}(V)$ is the convex hull of $V$. We call this representation of $P$ as the convex hull of a finite set an inner representation of $P$, also known as vertex representation in
the literature\footnote{For example in \cite{griepert1967}, where it is also abbreviated to V-representation}. In contrast, an outer representation of $P$, also called half-space representation, is the representation of $P$ as an intersection of a finite number of half-spaces. This means there exist $s \in \mathbb{N}$, $b \in \mathbb{R}^s$ and $a_i \in \mathbb{R}^n$ for every $i \in [s]$ such that

$$P = \{ x \in \mathbb{R}^n \mid \forall i \in [s] : \langle a_i, x \rangle \geq b_i \}.$$ 

Of course, both representations are not unique, on the contrary, every polytope has infinitely many representations.

**Definition 4.3.**

Let $n \in \mathbb{N}$. For a polytope $P \in \mathbb{R}^n$ fix both an inner representation and an outer representation as above. Then, the slack matrix of $P$ associated to the given representations is the matrix $S \in \mathbb{R}^{s \times t}$ defined by $S_{i,j} := \langle a_i, v_j \rangle - b_i$. Consider that $S$ has only nonnegative entries.

Now, the following theorem yields the desired connection between our two definitions of psd rank, and at once it explains why it also make sense to speak about the psd rank of a polytope.

**Theorem 4.4** (\cite{fmp+12, gpt13}).

Let $n \in \mathbb{N}$. Then, for every polytope $P \in \mathbb{R}^n$ and every slack matrix $S$ of $P$, it holds

$$\text{rk}_{psd}(P) = \text{rk}_{psd}(S).$$

In order to find a lower bound for $\text{rk}_{psd}({\text{CORR}}_n)$, with this theorem it is equivalent to lower bound the psd rank of any slack matrix of ${\text{CORR}}_n$ associated to any representation. Besides, it is clear from its definition that the psd rank of a matrix $M$ is greater or equal than the psd rank of every submatrix of $M$. For this reason, it is even enough to find a lower bound for the psd rank of some submatrix of a slack matrix of ${\text{CORR}}_n$.

Here the nonnegative, quadratic Boolean functions come in. They play an important role in view of the submatrices of the slack matrices of the correlation polytope. For this reason the correlation polytope is also known as the Boolean quadric polytope.\footnote{This notion dates back to \cite{pad89}.} In order to connect the nonnegative, quadratic Boolean functions and the correlation polytope, we still need some definition and notation. Especially, the next definition will often be used in the further course of the work.

**Definition 4.5.**

Let $n,m \in \mathbb{N}$ with $n \geq m$. Then, we denote by $\binom{[n]}{m}$ the set of all the subsets of $[n]$.\footnote{For example in \cite{gruip1967}, where it is also abbreviated to V-representation} \footnote{For example in \cite{gruip1967}, where it is also abbreviated to H-representation}
Chapter 4. A lower bound on the psd rank of the family of correlation polytopes

with \( m \) elements, that means
\[
\binom{[n]}{m} := \{ S \subseteq [n] \mid |S| = m \} \subseteq \mathcal{P}([n]).
\]

Furthermore, let \( x \in \{-1,1\}^n \) or \( x \in \{0,1\}^n \) respectively, and let \( S \in \binom{[n]}{m} \) with \( S = \{i_1, \ldots, i_m\} \) and \( 1 \leq i_1 < \cdots < i_m \leq n \). Then, \( x_S \in \{-1,1\}^m \) or \( x_S \in \{0,1\}^m \) respectively is defined by
\[
x_S := (x_{i_1}, \ldots, x_{i_m}),
\]
and it is called the projection of \( x \) onto the variables of \( S \).

Notation 4.6.
Let \( n,m \in \mathbb{N} \) with \( n \geq m \), and let \( f \in \mathbb{R}^{\{0,1\}^m} \). Furthermore, let \( \{S_1, \ldots, S_{\binom{n}{m}}\} = \binom{[n]}{m} \) and \( \{x_1, \ldots, x_{2^n}\} = \{0,1\}^n \) respectively. Then, we define \( \mathcal{M}_n^f \in \mathbb{R}^{\binom{n}{m} \times 2^n} \) by
\[
(\mathcal{M}_n^f)_{i,j} := f((x_j)_{S_i}).
\]

Now, it turns out that for quadratic and nonnegative \( f \), \( \mathcal{M}_n^f \) is a submatrix of some slack matrix of \( \text{CORR}_n \), what is shown in the next theorem. A fragmentary sketch of a proof is given in [LRS15, Proposition 5.1]. But the statement is very essential for the further work because it allows us to connect the lower bound of \( \text{rk}_{psd}(\text{CORR}_n) \) to a lower bound of a concrete matrix. Because of this importance and for a good understanding of the proposition, we give an own, more precise and detailed proof of it here. The proof is not hard, but it is a bit technical and needs some notation and caution with the indices.

Theorem 4.7 ([LRS15] Proposition 5.1).
Let \( n,m \in \mathbb{N} \) with \( n \geq m \), and let \( f : \{0,1\}^m \to \mathbb{R}_+ \) be a nonnegative, quadratic function. Furthermore, let \( \{S_1, \ldots, S_{\binom{n}{m}}\} \) and \( \{x_1, \ldots, x_{2^n}\} \) be enumerations of \( \binom{[n]}{m} \) and \( \{0,1\}^n \) respectively. Then, the above defined matrix \( \mathcal{M}_n^f \in \mathbb{R}^{\binom{n}{m} \times 2^n} \) is a submatrix of some slack matrix of \( \text{CORR}_n \).

Proof. Let \( B \in \mathbb{R}^{m \times m} \) be a matrix and \( b_0 \) a real number such that \( f(x) = \langle B, xx^T \rangle + b_0 \) for all \( x \in \{0,1\}^m \) (cf. Proposition 3.17). Furthermore, for an arbitrary \( S \in \binom{[n]}{m} \) let \( \{s_1, \ldots, s_m\} = S \) with \( s_1 < \cdots < s_m \). Then, we define \( A_S \in \mathbb{R}^{n \times n} \) by
\[
(A_S)_{i,j} := \begin{cases} B_{k,l} & i = s_k, j = s_l \\ 0 & i \notin S \text{ or } j \notin S. \end{cases}
\]
For every $x \in \{0,1\}^n$, such a matrix satisfies

$$\langle A_S, xx^T \rangle + b_0 \overset{\text{Corr} \mbox{ of } A}{=} \sum_{i,j \in [n]} (A_S)_{i,j} x_i x_j + b_0 = \sum_{i,j \in S} (A_S)_{i,j} x_i x_j + b_0$$

$$= \sum_{k,l \in [m]} B_{k,l} x_k x_l + b_0 = \sum_{k,l \in [m]} B_{k,l}(x_{S_k}) x_{S_l} + b_0$$

$$= \langle B, xx^T_S \rangle + b_0 = f(x_S) \quad (4.5)$$

$$\geq 0. \quad (4.6)$$

We remember the natural inner representation

$$\text{CORR}_n = \text{conv} \left( \{ xx^T \mid x \in \{0,1\}^n \} \right) = \text{conv} \left( \{ x_1 x_1^T, \ldots, x_{2^n} x_{2^n}^T \} \right)$$

of CORR$_n$. For $1 \leq i \leq \binom{n}{m}$, we define $A_i := A_{S_i}$ and $b_i := -b_0$, where $A_{S_i}$ is the above defined matrix for $S = S_i$. Now, the just showed inequality yields

$$\langle A_i, xx^T \rangle = \langle A_{S_i}, xx^T \rangle \overset{4.6}{=} -b_0 = b_i,$$

for every $1 \leq i \leq \binom{n}{m}$ and every $x \in \{0,1\}^n$. Since this inequality is valid for all the vertices, it is also valid for all $X \in \text{CORR}_n$, and we get

$$\text{CORR}_n \subseteq \{ X \in \mathbb{R}^{n \times n} \mid \forall 1 \leq i \leq \binom{n}{m} : \langle A_i, X \rangle \geq b_i \}.$$  

Thus, there exists an outer representation that contains $\langle A_i, X \rangle \geq b_i$ for all $1 \leq i \leq \binom{n}{m}$. Hence, for the slack matrix $\mathcal{S}$ of CORR$_n$ associated to this outer representation and the above mentioned inner representation holds

$$\mathcal{S}_{i,j} = \langle A_i, x_j x_j^T \rangle - b_i = \langle A_{S_i}, x_j x_j^T \rangle + b_0 \overset{4.6}{=} f((x_j)_{S_i}) \overset{\mbox{Not} \mbox{ of } \mathcal{M}_n}{=} (\mathcal{M}_n^i)_{i,j},$$

for all $1 \leq i \leq \binom{n}{m}$ and all $j \in [2^n]$. Therefore, $\mathcal{M}_n^i$ is a submatrix of $\mathcal{S}$. \hfill \Box

A disadvantage of the notion of $\mathcal{M}_n^i$ is that we always need an enumeration of $\binom{[n]}{m}$ and $\{0,1\}^n$ respectively. To avoid this in the future, we give index-free versions of Definition 4.2 and Notation 4.6. Thereby, we have to take care about enumerations no longer. Moreover, it is more comfortable.

**Definition 4.8.**

Let $n,m \in \mathbb{N}$ with $n \geq m$. We consider any nonnegative function $N : \binom{[n]}{m} \times \{0,1\}^n \rightarrow \mathbb{R}_+$. According to Definition 4.2 we say $N$ admits a rank-$r$ psd factorization if there exist two mappings $A : \binom{[n]}{m} \rightarrow \mathbb{S}^r$ and $B : \{0,1\}^n \rightarrow \mathbb{S}^r$ such that

$$A(S), B(x) \mbox{ are psd and } N(S,x) = \langle A(S), B(x) \rangle = \text{Tr}(A(S)B(x))$$
for all $S \in \binom{[n]}{m}$, $x \in \{0,1\}^n$.

We define the \textit{psd rank of $N$} as the smallest $r$ such that $N$ admits a rank-$r$ psd factorization, and we denote it by $\text{rk}_{\text{psd}}(N)$. The set of all functions $N : \binom{[n]}{m} \times \{0,1\}^n \to \mathbb{R}$ is denoted by $\mathbb{R}^{(n)\times (0,1)^n}$ as usual.

Again, there is a connection from such functions to Boolean functions. According to Notation 4.6, every Boolean function $f \in \mathbb{R}^{\{0,1\}^m}$ yields a special element of $\mathbb{R}^{\binom{[n]}{m} \times \{0,1\}^n}$ for every $n \geq m$.

\textbf{Notation 4.9.}
Let $n,m \in \mathbb{N}$ with $n \geq m$, and let $f \in \mathbb{R}^{\{0,1\}^m}$. We define the function $M^f_n : \binom{[n]}{m} \times \{0,1\}^n \to \mathbb{R}$ by
\[
M^f_n(S,x) := f(x_S).
\]

Now, we obviously obtain the following consequence.

\textbf{Corollary 4.10.}
Let $n,m \in \mathbb{N}$ with $n \geq m$, and let $f : \{0,1\}^m \to \mathbb{R}_+$ be a nonnegative, quadratic function. Furthermore, let $\{S_1,\ldots,S_{\binom{n}{m}}\}$ and $\{x_1,\ldots,x_{2^n}\}$ be arbitrary but fixed enumerations of $\binom{[n]}{m}$ and $\{0,1\}^n$ respectively, and let $\mathcal{M}^f_n$ the concerning nonnegative real matrix from Notation 4.6. Then, it holds
\[
\text{rk}_{\text{psd}}(\mathcal{M}^f_n) = \text{rk}_{\text{psd}}(M^f_n).
\]

Now, combining Theorem 4.4, Theorem 4.7 and the previous corollary immediately yields the following consequence.

\textbf{Corollary 4.11.}
Let $n,m \in \mathbb{N}$ with $n \geq m$, and let $f : \{0,1\}^m \to \mathbb{R}_+$ be a nonnegative, quadratic function. Furthermore, let $M^f_n : \binom{[n]}{m} \times \{0,1\}^n \to \mathbb{R}_+$ defined as above. Then, it holds
\[
\text{rk}_{\text{psd}}(\text{CORR}_n) \geq \text{rk}_{\text{psd}}(M^f_n).
\]

This shows that proving a lower bound for the psd rank of $M^f_n$ in terms of $n$ for some nonnegative, quadratic Boolean function $f \in \mathbb{R}^{\{0,1\}^m}$ with $m \leq n$ is enough in order to get a lower bound for the psd rank of CORR$_n$.

To prove such a lower bound for the psd rank of $M^f_n$ is now the crucial step. For that, we take advantage of the fact that in a psd factorization of every $N \in \mathbb{R}^{(n)\times (0,1)^n}$ a matrix-valued Boolean function $B : \{0,1\}^n \to \mathbb{R}^{r \times r}$ occurs. So, we are able to use results and techniques from Boolean functions, what we will do in the next sections.
4.2 A proof overview - the basic ideas of Lee, Raghavendra and Steurer

After the results in the last section rather described general ideas to work with the psd rank of a polytope, we want to explain and motivate the special ideas of [LRS15] now. In this section, we want to give a sketchy overview on the steps to do as well to illustrate how our own ideas integrate in the proof of [LRS15], that formally will be given in the following sections.

In Section 3.5, we used linear forms to separate a Boolean function from the cone of sums of squares of low degree in order to certify that the sos degree of this function is greater than a certain lower bound.

We want to apply this idea to our setting. Roughly speaking, the goal is to find a quadratic, nonnegative function $f \in \mathbb{R}^{\{0,1\}^m}$, a suitable lower bound $b \in \mathbb{N}$, $\alpha \in \mathbb{R}$ and a function $L : \mathbb{R}^{\binom{n}{m}} \times \{0,1\}^n \to \mathbb{R}$ with $L(M^f_n) < \alpha$ and $L(N) \geq \alpha$ for all $N \in \mathbb{R}^{\binom{n}{m}} \times \{0,1\}^n$ with $\text{rk}_{\text{psd}}(N) \leq b$. This would certify $\text{rk}_{\text{psd}}(M^f_n) > b$.

It turns out that such a separating function can indeed be a linear form, moreover a linear form defined by pseudo-densities.

**Notation 4.12.** Let $n,m,d \in \mathbb{N}$ with $n \geq m$ and $d \in [2m]$. For a degree-$d$ pseudo-density $D \in \mathbb{R}^{\{0,1\}^m}$, we define the function $L_D : \mathbb{R}^{\binom{n}{m}} \times \{0,1\}^n \to \mathbb{R}$ by

$$L_D(N) = \mathbb{E}_{S \sim \binom{n}{m}} \mathbb{E}_x D(x_S) N(S,x),$$

where $S \sim \binom{n}{m}$ denotes that $S$ is a uniformly chosen random element from $\binom{n}{m}$. In this connection, note that expectations over $\binom{n}{m}$ will always be with respect to this uniform distribution. Therefore, we write the shorter $\mathbb{E}_S$ instead of $\mathbb{E}_{s \sim \binom{n}{m}}$ and $\mathbb{E}_{S,x}$ instead of $\mathbb{E}_{s \sim \binom{n}{m}} \mathbb{E}_x$.

One important reason why this just defined functions will be suitable for our purpose is the following statement:

**Proposition 4.13.** Let $n,m \in \mathbb{N}$ with $n \geq m$ and $d \in [2m]$, and let $f \in \mathbb{R}^{\{0,1\}^m}$. Furthermore, let $D \in \mathbb{R}^{\{0,1\}^m}$ be a degree-$d$ pseudo-density and $\delta \in \mathbb{R}$ with $\mathbb{E}_x D(x) f(x) = \delta$, and let $L_D : \mathbb{R}^{\binom{n}{m}} \times \{0,1\}^n \to \mathbb{R}$ be defined as above. Then, we have

$$L_D(M^f_n) = \delta.$$
Proof. For every fixed $S \in \binom{n}{m}$, we have
\[
\mathbb{E}_{y \sim \{0,1\}^n} D(yS)f(yS) = \frac{1}{2^n} \sum_{y \in \{0,1\}^n} D(yS)f(yS) = \frac{1}{2^n} \sum_{x \in \{0,1\}^m} \sum_{y \in \{0,1\}^n \mid y_i = x_i} D(x)f(x)
\]
\[
= \frac{1}{2^n} \sum_{x \in \{0,1\}^m} 2^{n-m} D(x)f(x) = \frac{1}{2^m} \sum_{x \in \{0,1\}^m} D(x)f(x)
\]
\[
= \mathbb{E}_x D(x)f(x) = \delta. \tag{4.7}
\]
Therefore, we obtain
\[
L_D(M^f_n) = \mathbb{E}_{S \sim \binom{[n]}{m}} \mathbb{E}_{y \sim \{0,1\}^n} D(yS)M^f_n(S,y) = \mathbb{E}_{S \sim \binom{[n]}{m}} \mathbb{E}_{y \sim \{0,1\}^n} D(yS)f(yS) = \mathbb{E}_{S \sim \binom{[n]}{m}} \delta = \delta.
\]

We see that $L_D$ returns the same value applied to $M^f_n$ than $L$ applied to $f$, if $L$ is the degree-$d$ pseudo-expectation corresponding to $D$. The second important thing is that $L_D$ also preserves the behavior of $L$ related to low degree squares. While $L$ only returns nonnegative values applied to squares of functions of degree at most \(\frac{d}{2}\), we will show that $L_D$ does the same if applied to functions that admit a psd factorization in terms of squares of degree $\frac{d}{2}$.

Concretely, this means $L_D(N) \geq 0$ for all $N \in \mathbb{R}^{\binom{[n]}{m} \times \{0,1\}^n}$ satisfying
\[
N(S,x) = \text{Tr} \left( A(S)^2 R(x)^2 \right) \quad \text{for all } S \in \binom{[n]}{m}, x \in \{0,1\}^n \tag{4.8}
\]
for some $r \in \mathbb{N}$, some $A : \binom{[n]}{m} \to \mathbb{S}^r$ and some $R : \{0,1\}^n \to \mathbb{S}^r$ with $\text{deg}(R) \leq \frac{d}{2}$. Before we will prove this, we first need some preparation.

**Notation 4.14.**

Let $n,m \in \mathbb{N}$ with $n \geq m$, and let $S \in \binom{[n]}{m}$. We introduce the following obvious bijection
\[
\varphi_S : \{0,1\}^n \to \{0,1\}^m \times \{0,1\}^{n-m}, \quad z \mapsto (z_S, z_{[n]\setminus S}).
\]

Furthermore, for every $f \in \mathbb{R}^{\{0,1\}^n}$ and every $y \in \{0,1\}^{n-m}$, we define
\[
f_{S,y} : \{0,1\}^m \to \mathbb{R}, \quad x \mapsto f \left( \varphi_S^{-1}(x,y) \right).
\]

**Proposition 4.15.**

Let $n,m \in \mathbb{N}$ with $m \leq m$, and let $d \in [2m]$. Let $S \in \binom{[n]}{m}$, and let $D \in \mathbb{R}^{\{0,1\}^n}$ be a degree-$d$ pseudo-density. Furthermore, let $f \in \mathbb{R}^{\{0,1\}^n}$ with
\[
\text{deg}_{\text{sos}}(f_{S,y}) \leq d \quad \text{for every } y \in \{0,1\}^{n-m}.
\]

Then, we have
\[
\mathbb{E}_{z \sim \{0,1\}^n} D(z_S)f(z) \geq 0.
\]
Proof. Note that by the assumption and because \( D \) is a degree-\( d \) pseudo-density, we have
\[
\mathbb{E} D(x) f_{S,y}(x) \geq 0
\]
for every \( y \in \{0,1\}^{n-m} \). This implies
\[
\mathbb{E}_{z \sim \{0,1\}^{n}} D(z_S) f(z) = \frac{1}{2^n} \sum_{x \in \{0,1\}^m} \sum_{y \in \{0,1\}^{n-m}} D(x) f(\varphi_S^{-1}(x,y)) \geq 0
\]
(4.9)

Note that \( \deg_{\text{sos}}(f_{S,y}) \leq \deg_{\text{sos}}(f) \) obviously holds. Therefore, the previous result is particularly valid for every \( f \in \mathbb{R}^{\{0,1\}^n} \) with \( \deg_{\text{sos}}(f) \leq d \).

Furthermore, note that for every \( n,r,s,t,k \in \mathbb{N} \), every \( A \in \mathbb{R}^{r \times s} \) and every \( R : \{0,1\}^n \to \mathbb{R}^{s \times t} \) with \( \deg(R) \leq k \), the function
\[
x \mapsto \|AR(x)\|_F^2
\]
is a sum of squares of degree at most \( k \). This follows by
\[
\|AR(x)\|_F^2 \overset{\text{Cor. 4.14}}{=} \sum_{i \in [r]} \sum_{j \in [t]} (AR(x)_{ij})^2 = \sum_{i \in [r]} \sum_{j \in [t]} \sum_{k \in [s]} A_{ik} R_{kj}(x)^2
\]
(4.10)

With this two observations, we can state the following.

**Corollary 4.16.**
Let \( m,r \in \mathbb{N} \) with \( n \geq m \), and let \( d \in [2m] \). Furthermore, let \( D \in \mathbb{R}^{\{0,1\}^m} \) be a degree-\( d \) pseudo-density and \( N \in \mathbb{R}^{(n-m) \times \{0,1\}^n} \) such that there exist \( A : \binom{[n]}{m} \to \mathbb{S}^r \) and \( R : \{0,1\}^n \to \mathbb{S}^r \) with \( \deg(R) \leq \frac{d}{2} \) and \( N(S,x) = \text{Tr}(A(S)^2 R(x)^2) \). Then, it holds
\[
L_D(N) \geq 0.
\]

**Proof.** It holds
\[
L_D(N) = \mathbb{E}_{S,x} D(x_S) N(S,x) = \mathbb{E}_{S,x} D(x_S) \text{Tr}(A(S)^2 R(x)^2)
\]
\[
= \mathbb{E}_{S,x} D(x_S) \text{Tr}(A(S)^2 R(x)^2) \overset{\text{Cor. 4.14}}{=} \mathbb{E}_{S,x} \|A(S)R(x)\|_F^2 \overset{\text{Prop. 4.15}}{\geq} 0.
\]
\[\square\]
Based on the hitherto results, we now present the basic ideas of [LRS15] how to lower bound the psd rank of $M_n^f$ for some suitable quadratic and nonnegative function $f \in \mathbb{R}^{(0,1)^m}$. Note that this is just a short and informal draft, while the formal and detailed proofs follow in the subsequent sections.

For every $m \in \mathbb{N}$, combining Theorem 3.36, Corollary 3.45 and Proposition 4.13 yields some concrete quadratic and nonnegative $f \in \mathbb{R}^{(0,1)^m}$ and the existence of a real number $\varepsilon > 0$ and a degree-$d$ pseudo-density $D \in \mathbb{R}^{(0,1)^m}$ such that

$$L_D(M_n^f) < -\varepsilon$$

(4.11)

for every $n \geq m$. Hence, the previous corollary implies that $M_n^f$ has no degree-$\frac{d}{2}$ psd factorization in the sense of (4.8).

Thus, in order to lower bound the psd rank of $M_n^f$, it would be enough to have that a low-rank psd factorization would also imply a low-degree factorization. Of course, we cannot assume this.

On the other hand, we do not need $L_D(N) \geq 0$ for all $N \in \mathbb{R}^{(n)} \times \{0,1\}^n$ with low psd rank, $L_D(N) \geq -\varepsilon$ is sufficient for us.

Therefore, based on the previous Corollary and the fact

$$L_D \left( \text{Tr} \left( A(S)^2 R(x)^2 \right) \right) \geq 0$$

for every $R : \{0,1\}^n \rightarrow \mathbb{S}$ of low degree, concretely that means $\deg(R) \leq \frac{d}{2}$, [LRS15] showed

$$L_D \left( \text{Tr} \left( A(S)^2 R(x)^2 \right) \right) \geq -\frac{\varepsilon}{2}$$

(4.12)

for every $R : \{0,1\}^n \rightarrow \mathbb{S}$ of 'medium' degree at most $K$. They call this first essential step degree reduction.

In their second essential step, they showed that every $N \in \mathbb{R}^{(n)} \times \{0,1\}^n$ with low psd rank can be approximated with respect to $L_D$ by a function that admits a psd factorization of 'medium' degree $K$, that means there exist $A : \left( \begin{bmatrix} n \\ m \end{bmatrix} \right) \rightarrow \mathbb{S}$ and $R : \{0,1\}^n \rightarrow \mathbb{S}$ with $\deg(R) \leq K$ such that

$$L_D(N) - L_D \left( \text{Tr} \left( A(S) R(x)^2 \right) \right) \geq -\frac{\varepsilon}{2}.$$  

(4.13)

This is called matrix approximation by [LRS15]. But note that actually it is not a really approximation but only a bound in one direction. Combining (4.12) and (4.13) implies

$$L_D(N) \geq -\varepsilon,$$

for every $N \in \mathbb{R}^{(n)} \times \{0,1\}^n$ with low psd rank, and together with (4.11), this certifies that $M_n^f$ admits no low-rank psd factorization as desired.
Chapter 4. A lower bound on the psd rank of the family of correlation polytopes

From a more real algebraic point of view, which we adopt, it is unusual to use just a square of \( R : \{0,1\}^n \to \mathbb{S}^r \) in the situation of (4.13). It would be much more natural to also allow non-symmetric squares \( R : \{0,1\}^n \to \mathbb{R}^{r \times s} \) for some \( s \in \mathbb{N} \) and replace \( R(x)^2 \) by \( R(x)R(x)^T \) in (4.13) or even allow sums of squares \( \sum_{i \in [l]} R_i(x)^2 \) with \( R_i : \{0,1\}^n \to \mathbb{S}^r \) for every \( i \in [l] \) for some \( l \in \mathbb{N} \).

Note that because of

\[
\left( A_1 \ldots A_l \right) \cdot \left( A_1 \ldots A_l \right)^T = A_1^2 + \ldots + A_l^2 \quad (4.14)
\]

for every \( l \in \mathbb{N} \) and all \( A_1, \ldots, A_l \in \mathbb{R}^{r \times r} \), our first case includes the second one.

Due to our own considerations, we will generalize the degree reduction result of [LRS15], and we will show that (4.12) is still valid if we replace \( R : \{0,1\}^n \to \mathbb{S}^r \) and \( R(x)^2 \) by \( R : \{0,1\}^n \to \mathbb{R}^{r \times s} \) and \( R(x)R(x)^T \). This is done in Section 4.3.

For this reason, in the proof of the matrix approximation result, we are allowed to use a non-symmetric \( R : \{0,1\}^n \to \mathbb{R}^{r \times s} \) to show (4.13). Therefore, we will obtain some slightly better bounds on the degree of such an \( R \) than [LRS15] in their original work, which is the content of Section 4.4. There, we will also give some more explanation and motivation about our two slightly different approaches and the resulting different bounds.

Note that with (4.12) and (4.13), we just presented a rough idea and a large simplification of the actual results of [LRS15], in reality it is more complicated. Therefore, there is still a lot of work to do in bringing these two results together and stating the main theorem. This is done in Section 4.5. In the two further Sections 4.6 and 4.7 we apply this main theorem to the family of correlation polytopes in order to obtain the concrete lower bound (4.1).

### 4.3 Degree Reduction

Based on (4.12), we want to lower bound

\[
\mathbb{E}_{S,x} D(x_S) \operatorname{Tr} \left( (A(S))^2 B(x) B(x)^T \right)
\]

with respect to the degree of \( B \) in this section, where we actually use the equivalent formulation

\[
\mathbb{E}_{S,x} D(x_S) \left\| A(S)B(x) \right\|^2_F
\]

(cf. Corollary 1.14). In this connection, \( A : \left( \frac{n}{m} \right) \to \mathbb{S}^r \) and \( B : \{0,1\}^n \to \mathbb{R}^{s \times t} \) are two matrix-valued functions, and \( D \in \mathbb{R}^{\{0,1\}^m} \) is a degree-\( d \) pseudo-density. The goal is to show that this expression cannot adopt arbitrary negative values if \( B \) is a function of "medium" degree.
In order to use the nice properties of the Fourier expansion of Boolean functions in \( \mathbb{R}^{(-1,1)^n} \), in particular Proposition 3.7, we have to work in the ring \( \mathbb{R}^{(-1,1)^n} \) instead of \( \mathbb{R}^{(0,1)^n} \). We show the desired result for functions \( B : \{-1,1\}^n \to \mathbb{R}^{s \times t} \). At the end of the section, we transform our result back to the case \( \{0,1\}^n \).

First, we introduce some notation.

**Notation 4.17.**
Let \( n, m, r, s \in \mathbb{N} \) with \( n \geq m \). Furthermore, let \( B : \{-1,1\}^n \to \mathbb{R}^{r \times s} \) be a matrix-valued Boolean function with its Fourier expansion (cf. Notation 3.10)

\[
B = \sum_{I \subseteq [n]} \hat{B}(I) \chi_I,
\]

where \((\chi_I)_{I \subseteq [n]}\) denotes the family of the monomial functions in \( \{-1,1\}^n \). Then, for every \( S \in \binom{[n]}{m} \) and every \( d \in [2m] \), we define

\[
B_{S,d}^{\text{low}} := \sum_{\substack{I \subseteq [n] \\
|I \setminus S| \leq \frac{d}{2}}} \hat{B}(I) \chi_I \quad \text{and} \quad B_{S,d}^{\text{high}} := \sum_{\substack{I \subseteq [n] \\
|I \setminus S| > \frac{d}{2}}} \hat{B}(I) \chi_I.
\]

It obviously holds

\[
B = B_{S,d}^{\text{low}} + B_{S,d}^{\text{high}}.
\]

Note that the results from the previous section as well as Notation 4.14 are obviously still valid if we replace \( \{0,1\}^n \) by \( \{-1,1\}^n \). Therefore, for every fixed \( S \in \binom{[n]}{m} \), every \( A : \binom{[n]}{m} \to \mathbb{R}^{r \times s} \) and every \( B : \{-1,1\}^n \to \mathbb{R}^{s \times t} \), the function

\[
x \mapsto \left\| A(S) B_{S,d}^{\text{low}}(x) \right\|_F^2
\]

is a sum of squares by (4.10).

The crucial thing of the above notation is that \( B_{S,d}^{\text{low}} \) is the part of \( B \) with degree at most \( \frac{d}{2} \) in the variables \( x_S \). Therefore, following Notation 4.14, it is easy to see

\[
\deg_{\text{sos}} \left( x \mapsto \left\| A(S) B_{S,d}^{\text{low}}(x) \right\|_F^2 \right) \leq d
\]

for every \( y \in \{-1,1\}^{n-m} \).

Thus, by Proposition 4.15, we have

\[
\mathbb{E}_x D(x_S) \left\| A(S) B_{S,d}^{\text{low}}(x) \right\|_F^2 \geq 0
\]

for every degree-\( d \) pseudo-density \( D \in \mathbb{R}^{(-1,1)^m} \), which again immediately yields the following statement.
Corollary 4.18.
Let \( n,m,r,s,t \in \mathbb{N} \) with \( n \geq m \), and let \( d \in [2m] \). Furthermore, let \( D \in \mathbb{R}^{(-1,1)^m} \) be a degree-\( d \) pseudo-density, and let \( A : \left( \begin{bmatrix} n \end{bmatrix} \right) \to \mathbb{R}^{r \times s} \) and \( B : \{-1,1\}^n \to \mathbb{R}^{s \times t} \) be two matrix-valued functions. Then, we have
\[
\mathbb{E}_{S,x} D(x_S) \| A(S) B_{S,d}^\text{low}(x) \|_F^2 \geq 0.
\]

In the proof of the main theorem of this section, it will turn out that we need to lower bound \( \mathbb{E}_{S,x} \| B_{S,d}^\text{high}(x) \|_F^2 \) what we will do in the next Proposition 4.20. For this, the following combinatorial lemma is relevant. [LRS15] claim a slightly weaker version of our lemma en route in their proof of [LRS15, Lemma 3.7] without proving it. It is not hard to prove, but because of the significance for the proposition and therefore for the main theorem, we state it as an independent lemma and give a proof of it.

Lemma 4.19.
Let \( n \in \mathbb{N} \), \( A \subseteq \left[ n \right] \), \( m \in \left[ n \right] \) and \( t \in \mathbb{N} \) with \( t \leq m \). We set \( k := |A| \). Then, we have
\[
\mathbb{P}_{S \sim \left( \begin{bmatrix} n \end{bmatrix} \right)} \{ |S \cap A| \geq t \} \leq \left( \frac{k}{t} \right) \left( \frac{n-t}{m-t} \right) \frac{n!}{m!}.
\]

Proof. Let \( M := \left\{ S \in \left( \begin{bmatrix} n \end{bmatrix} \right) \mid |S \cap A| \geq t \right\} \). Thus,
\[
\mathbb{P}_{S \sim \left( \begin{bmatrix} n \end{bmatrix} \right)} \{ |S \cap A| \geq t \} = \frac{|M|}{\binom{n}{m}}. \tag{4.15}
\]

The set \( M \) satisfies
\[
M = \left\{ S \in \left( \begin{bmatrix} n \end{bmatrix} \right) \mid \exists B \subseteq A, |B| = t, B \subseteq S \right\} = \bigcup_{B \subseteq A, |B| = t} \left\{ S \in \left( \begin{bmatrix} n \end{bmatrix} \right) \mid B \subseteq S \right\},
\]
and therefore
\[
|M| = \left| \bigcup_{B \subseteq A, |B| = t} \left\{ S \in \left( \begin{bmatrix} n \end{bmatrix} \right) \mid B \subseteq S \right\} \right|
\leq \sum_{B \subseteq A, |B| = t} \left| \left\{ S \in \left( \begin{bmatrix} n \end{bmatrix} \right) \mid B \subseteq S \right\} \right| = \sum_{B \subseteq A, |B| = t} \binom{n-t}{m-t} = \left( \frac{k}{t} \right) \left( \frac{n-t}{m-t} \right) \tag{4.16}
\]

Combining (4.15) and (4.16) yields the statement. \( \square \)
Now we can prove the already mentioned proposition. Compared to [LRS15 Lemma 3.7.] the upper bound is slightly stronger because our bound of the previous lemma is slightly stronger, and it is also more general because it is not only valid for a symmetric and therefore quadratic matrix-valued function $B$. Besides, our proof is much more detailed.

**Proposition 4.20** (cf. [LRS15 Lemma 3.7]).

Let $n,k,s,t \in \mathbb{N}$ with $k \leq n$, $m \in [n]$, and let $d \in [2m-2]$ be even. Furthermore, let $B : \{-1,1\}^n \rightarrow \mathbb{R}^{s \times t}$ with $\deg(B) \leq k$. Then, it holds

$$\mathbb{E}_{x \sim \binom{[n]}{m}} \mathbb{E}_{I \subseteq [n] \atop |I \cap S| > \frac{d}{2}} \mathbb{E}_{J \subseteq [n] \atop |J \cap S| > \frac{d}{2}} \left\| B_{S,d}(x) \right\|_F^2 \leq \left( \frac{km}{n - \frac{d}{2}} \right)^{\frac{d}{2} + 1} \mathbb{E}_{x} \left\| B(x) \right\|_F^2.$$ 

**Proof.** For a fixed $S \in \binom{[n]}{m}$, we have

$$\mathbb{E}_{x} \left\| B_{S,d}(x) \right\|_F^2 = \mathbb{E}_{x} \text{Tr} \left( B_{S,d}(x)^T B_{S,d}(x) \right)$$

$$= \mathbb{E}_{x} \text{Tr} \left( \sum_{I \subseteq [n] \atop |I \cap S| > \frac{d}{2}} \hat{B}(I)^T \chi_I(x) \cdot \sum_{J \subseteq [n] \atop |J \cap S| > \frac{d}{2}} \hat{B}(J) \chi_J(x) \right)$$

$$= \mathbb{E}_{x} \sum_{I \subseteq [n] \atop |I \cap S| > \frac{d}{2}} \sum_{J \subseteq [n] \atop |J \cap S| > \frac{d}{2}} \text{Tr} \left( \hat{B}(I)^T \hat{B}(J) \right) \chi_{I \Delta J}(x)$$

$$= \sum_{I \subseteq [n] \atop |I \cap S| > \frac{d}{2}} \sum_{J \subseteq [n] \atop |J \cap S| > \frac{d}{2}} \text{Tr} \left( \hat{B}(I)^T \hat{B}(J) \right) \mathbb{E}_{x} \chi_{I \Delta J}(x)$$

$$= \sum_{I \subseteq [n] \atop |I \cap S| > \frac{d}{2}} \sum_{J \subseteq [n] \atop |J \cap S| > \frac{d}{2}} \text{Tr} \left( \hat{B}(I)^T \hat{B}(J) \right) = \sum_{I \subseteq [n] \atop |I \cap S| > \frac{d}{2}} \left\| \hat{B}(I) \right\|_F^2.$$ 

(4.17)

Note that with exactly the same steps, we can congruently show

$$\mathbb{E}_{x} \left\| B(x) \right\|_F^2 = \sum_{I \subseteq [n] \atop |I \cap S| > \frac{d}{2}} \left\| \hat{B}(I) \right\|_F^2.$$ 

(4.18)

By (4.17), for the expectation over all $S \in \binom{[n]}{m}$ holds

$$\mathbb{E}_{S \sim \binom{[n]}{m}} \mathbb{E}_{x} \left\| B_{S,d}(x) \right\|_F^2 = \mathbb{E}_{S} \sum_{I \subseteq [n] \atop |I \cap S| > \frac{d}{2}} \left\| \hat{B}(I) \right\|_F^2.$$ 

(4.19)
For a fixed $S \in \binom{[n]}{m}$, let
\[
\mathbb{1}_{|J \cap S| > \frac{d}{2}} := \mathbb{1}_{\{J \subseteq [n] \mid |J \cap S| > \frac{d}{2}\}} : \mathcal{P}([n]) \to \{0,1\}
\]
be the indicator function of $\{J \subseteq [n] \mid |J \cap S| > \frac{d}{2}\}$; and for a fixed $I \subseteq [n]$, let
\[
\mathbb{1}_{|T \cap I| > \frac{d}{2}} := \mathbb{1}_{\{T \in \binom{[n]}{m} \mid |T \cap I| > \frac{d}{2}\}} : \binom{[n]}{m} \to \{0,1\}
\]
be the indicator function of $\{T \in \binom{[n]}{m} \mid |T \cap I| > \frac{d}{2}\}$. Then, for every $S \in \binom{[n]}{m}$ and every $I \subseteq [n]$ we have
\[
\mathbb{1}_{|J \cap S| > \frac{d}{2}}(I) = \mathbb{1}_{|T \cap I| > \frac{d}{2}}(S). \tag{4.20}
\]
Furthermore, the second indicator function satisfies
\[
\mathbb{E}_{S \sim \binom{[n]}{m}} \mathbb{1}_{|T \cap I| > \frac{d}{2}}(S) = \mathbb{P}_{S \sim \binom{[n]}{m}}(\{S \cap I\}) \tag{4.21}
\]
for every fixed $I \subseteq [n]$.

Putting the last three equations together leads to
\[
\mathbb{E}_{S \sim \binom{[n]}{m}} \mathbb{E}_{x} \left\| B_{S,d}^{\text{high}}(x) \right\|_F^2 \overset{(4.19)}{=} \mathbb{E}_{S \sim \binom{[n]}{m}} \sum_{I \subseteq [n]} \left\| \hat{B}(I) \right\|_F^2 \mathbb{1}_{|J \cap S| > \frac{d}{2}}(I) = \mathbb{E}_{S \sim \binom{[n]}{m}} \sum_{I \subseteq [n]} \left\| \hat{B}(I) \right\|_F^2 \mathbb{1}_{|T \cap I| > \frac{d}{2}}(S) \overset{(4.20)}{=} \sum_{I \subseteq [n]} \left\| \hat{B}(I) \right\|_F^2 \mathbb{E}_{S \sim \binom{[n]}{m}} \mathbb{1}_{|T \cap I| > \frac{d}{2}}(S)
\]
\[
= \sum_{I \subseteq [n]} \left\| \hat{B}(I) \right\|_F^2 \mathbb{P}_{S \sim \binom{[n]}{m}}(\{\{S \cap I\} > \frac{d}{2}\}) \tag{4.22}
\]
For every $I \subseteq [n]$ we can bound $\mathbb{P}_{S \sim \binom{[n]}{m}}(\{|S \cap I| > \frac{d}{2}\})$ with Lemma 4.19 by
\[
\mathbb{P}_{S \sim \binom{[n]}{m}}(\{|S \cap I| > \frac{d}{2}\}) = \mathbb{P}_{S \sim \binom{[n]}{m}}(\{|S \cap I| \geq \frac{d}{2} + 1\}) \overset{\text{Lem. 4.19}}{\leq} \binom{|I|}{\frac{d}{2} + 1} \binom{n - (\frac{d}{2} + 1)}{m - (\frac{d}{2} + 1)} \binom{n}{m}.
\]
Hence, for $I \subseteq [n]$ with $|I| \leq k$ this yields to

$$
\mathbb{P}_{S \sim \binom{[n]}{m}} \left( \left\{ |S \cap I| > \frac{d}{2} \right\} \right) \leq \frac{k!}{\left( \frac{d}{2} + 1 \right)! (k - \frac{d}{2} - 1)! (m - \frac{d}{2} - 1)!(n - m)!} \frac{(n - \frac{d}{2} - 1)!}{n!} \frac{m!(n - m)!}{n!} 
\leq \frac{1}{\left( \frac{d}{2} + 1 \right)!} \frac{k!}{(k - \frac{d}{2} - 1)! (m - \frac{d}{2} - 1)!} \frac{m!}{n!} \frac{(n - \frac{d}{2} - 1)!}{n!} 
\leq \frac{1}{\left( \frac{d}{2} + 1 \right)!} \prod_{i=0}^{\frac{d}{2}} \frac{(k - i)(m - i)}{n - i} \leq \left( \frac{km}{n - \frac{d}{2}} \right)^{\frac{d}{2} + 1}
$$

In the situation of (4.22), for every $I \subseteq [n]$ either we have $|I| \leq k$ or we have $\|\hat{B}(I)\|_F = 0$. This holds because $\deg B \leq k$. So, we either have

$$
\mathbb{P}_{S \sim \binom{[n]}{m}} \left( \left\{ |S \cap I| > \frac{d}{2} \right\} \right) \leq \left( \frac{km}{n - \frac{d}{2}} \right)^{\frac{d}{2} + 1}
$$

or we have

$$
\|\hat{B}(I)\|_F = 0.
$$

Applying this fact to (4.22) yields

$$
\mathbb{E}_{S \sim \binom{[n]}{m}} \mathbb{E}_{x} \left\| B_{S,d}^{\text{high}}(x) \right\|_F^2 \leq \sum_{I \subseteq [n]} \|\hat{B}(I)\|_F^2 \left( \frac{km}{n - \frac{d}{2}} \right)^{\frac{d}{2} + 1}
\leq \left( \frac{km}{n - \frac{d}{2}} \right)^{\frac{d}{2} + 1} \mathbb{E} \left\| B(x) \right\|_F^2.
$$

Before proving the main theorem, we need one more small result, stated in Corollary 4.22 and obtained by the following lemma.

**Lemma 4.21.**

Let $n,r,s,t \in \mathbb{N}$. Furthermore, let $A \in \mathbb{R}^{r \times s}$, and let $B : \{-1,1\}^n \rightarrow \mathbb{R}^{s \times t}$ be matrix-valued Boolean functions. Then, for every $S \subseteq [n]$ and every $d \in [n]$ holds

$$
\mathbb{E}_x \left\langle AB_{S,d}^{\text{low}}(x), AB_{S,d}^{\text{high}}(x) \right\rangle = 0.
$$
Proof. For every $S \subseteq [n]$ and every $d \in [n]$, we have

$$
\mathbb{E}_x \left\langle AB_{S,d}^\text{low}(x), AB_{S,d}^\text{high}(x) \right\rangle = \mathbb{E}_x \text{Tr} \left( \left( AB_{S,d}^\text{low}(x) \right)^T \left( AB_{S,d}^\text{high}(x) \right) \right)
$$

$$
= \mathbb{E}_x \text{Tr} \left( B_{S,d}^\text{low}(x) A^T B_{S,d}^\text{high}(x) \right)
$$

$$
= \mathbb{E}_x \text{Tr} \left( A^T B_{S,d}^\text{high}(x) B_{S,d}^\text{low}(x)^T \right)
$$

$$
= \text{Tr} \left( A^T \mathbb{E}_x B_{S,d}^\text{high}(x) B_{S,d}^\text{low}(x)^T \right).
$$

(4.23)

The matrix-valued Boolean functions $B_{S,d}^\text{high}$ and $B_{S,d}^\text{low}$ have disjoint support by its definition. Hence, the same holds for $B_{S,d}^\text{high}$ and $B_{S,d}^\text{low}^T$ (cf. Notation 3.10). Therefore, by applying Lemma 3.12 we obtain

$$
\mathbb{E}_x B_{S,d}^\text{low}(x) B_{S,d}^\text{low}^T(x) = 0.
$$

Putting this into (4.23) yields the statement.

Corollary 4.22.

Let $n,r,s,t \in \mathbb{N}$, and let $m \in [n]$ and $d \in [m]$. Furthermore, let $A : \binom{[n]}{m} \to \mathbb{R}^{r \times s}$ and $B : \{-1,1\}^n \to \mathbb{R}^{s \times t}$ matrix-valued functions.

$$
\mathbb{E}_{s \sim \binom{[n]}{m}} \mathbb{E}_x \| A(S) B(x) \|_F^2 \geq \mathbb{E}_{s \sim \binom{[n]}{m}} \mathbb{E}_x \| A(S) B_{S,d}^\text{low}(x) \|_F^2
$$

Proof. Since the previous lemma holds for every $S \in [n]$, it particular holds for every $S \in \binom{[n]}{m}$. Therefore, we have

$$
\mathbb{E}_x \left\langle A(S) B_{S,d}^\text{low}(x), A(S) B_{S,d}^\text{high}(x) \right\rangle = 0
$$

for every $S \in \binom{[n]}{m}$, and hence

$$
\mathbb{E}_{s \sim \binom{[n]}{m}} \mathbb{E}_x \left\langle A(S) B_{S,d}^\text{low}(x), A(S) B_{S,d}^\text{high}(x) \right\rangle = 0.
$$

(4.24)

Due to $\| A(S) B(x) \|_F^2 = \| A(S) B_{S,d}^\text{low}(x) + A(S) B_{S,d}^\text{high}(x) \|_F^2$ and Remark 1.8 it holds

$$
\| A(S) B(x) \|_F^2 = \| A(S) B_{S,d}^\text{low}(x) \|_F^2 + \| A(S) B_{S,d}^\text{high}(x) \|_F^2 + 2 \left\langle A(S) B_{S,d}^\text{low}(x), A(S) B_{S,d}^\text{high}(x) \right\rangle.
$$
Therefore, we have
\[
\mathbb{E}_{S \sim \binom{[m]}{n}} \mathbb{E}_{x} \| A(S)B(x) \|_F^2 = \mathbb{E}_{S \sim \binom{[m]}{n}} \mathbb{E}_{x} \left\| A(S)B_{S,d}^{\text{low}}(x) \right\|_F^2 + \mathbb{E}_{S \sim \binom{[m]}{n}} \mathbb{E}_{x} \left\| A(S)B_{S,d}^{\text{high}}(x) \right\|_F^2 + 2 \mathbb{E}_{S \sim \binom{[m]}{n}} \mathbb{E}_{x} \left\langle A(S)B_{S,d}^{\text{low}}(x), A(S)B_{S,d}^{\text{high}}(x) \right\rangle \]
\[
\geq \mathbb{E}_{S \sim \binom{[m]}{n}} \mathbb{E}_{x} \left\| A(S)B_{S,d}^{\text{low}}(x) \right\|_F^2.
\]

Now, we are able to prove the main theorem of this section. For the same reasons as for Proposition 4.20, our theorem is slightly stronger and slightly more general than the corresponding result in [LRS15]. Again, our proof is more detailed.

Let \( n, k, s, t \in \mathbb{N} \) with \( k \leq n \), \( m \in \binom{[n]}{k} \), and let \( d \in [2m - 2] \) be even. Let \( A : \binom{[m]}{n} \rightarrow \mathbb{S}^* \), and \( B : \{-1,1\}^n \rightarrow \mathbb{R}^{s \times t} \) be matrix-valued functions with \( \deg(B) \leq k \), and let \( D : \{-1,1\}^m \rightarrow \mathbb{R} \) be a degree-\( d \) pseudo-density. Furthermore, we define 
\[
\lambda := \max_{S \in \binom{[n]}{m}} \| A(S)^2 \|.
\]
Then,
\[
\mathbb{E}_{S,x} D(x_S) \| A(S)B(x) \|_F^2 \geq -2\sqrt{\lambda} \| D \|_{\infty} \left( \frac{km}{n - d} \right)^{\frac{d}{2} + \frac{1}{2}} \left( \mathbb{E}_x \| B(x) \|_F^2 \right)^{\frac{1}{2}} \left( \mathbb{E}_{x} \| A(S)B(x) \|_F^2 \right)^{\frac{1}{2}}.
\]

**Proof.** First, note that Corollary 4.18 yields
\[
\mathbb{E}_{S,x} D(x_S) \left\| A(S)B_{S,d}^{\text{low}}(x) \right\|_F^2 = \mathbb{E}_{S,x} D(x_S) \left\| A(S)B_{S,d}^{\text{low}}(x) \right\|_F^2 \geq 0,
\]
which implies
\[
\mathbb{E}_{S,x} D(x_S) \| A(S)B(x) \|_F^2 \geq \mathbb{E}_{S,x} D(x_S) \| A(S)B(x) \|_F^2 - \mathbb{E}_{S,x} D(x_S) \left\| A(S)B_{S,d}^{\text{low}}(x) \right\|_F^2.
\]

Because we are interested in a negative lower bound for \( \mathbb{E}_{S,x} D(x_S) \| A(S)B(x) \|_F^2 \),
it is sufficient to upper bound the absolute value of the right-hand side of (4.25).

\[
\left| \mathbb{E}_{S,x} D(x_S) \left\| A(S)B(x) \right\|_F^2 - \mathbb{E}_{S,x} D(x_S) \left\| A(S)B_{S,d}^{\text{low}}(x) \right\|_F^2 \right|
\]

\[
= \mathbb{E}_{S,x} D(x_S) \left( \left\| A(S)B(x) \right\|_F^2 - \left\| A(S)B_{S,d}^{\text{low}}(x) \right\|_F^2 \right)
\]

\[
\leq \mathbb{E}_{S,x} D(x_S) \left( \left\| A(S)B(x) \right\|_F^2 - \left\| A(S)B_{S,d}^{\text{low}}(x) \right\|_F^2 \right)
\]

\[
\leq \left\| D \right\|_\infty \mathbb{E}_{S,x} \left( \left\| A(S)B(x) \right\|_F + \left\| A(S)B_{S,d}^{\text{low}}(x) \right\|_F \right) \left( \left\| A(S)B(x) \right\|_F - \left\| A(S)B_{S,d}^{\text{low}}(x) \right\|_F \right) \left( \left\| A(S)B(x) \right\|_F + \left\| A(S)B_{S,d}^{\text{low}}(x) \right\|_F \right) \frac{1}{2}
\]

\[
\leq \left( \mathbb{E}_{S,x} \left\| A(S)B(x) \right\|_F + \left\| A(S)B_{S,d}^{\text{low}}(x) \right\|_F \right)^2 \frac{1}{2},
\] (4.26)

where the last step follows by the Cauchy-Schwarz inequality.

Now, we have to upper bound the two factors in the brackets of the right-hand side of (4.26).

Again by applying Cauchy-Schwarz, we get

\[
\mathbb{E}_{S,x} \left\| A(S)B(x) \right\|_F \left\| A(S)B_{S,d}^{\text{low}}(x) \right\|_F \leq \sqrt{\mathbb{E}_{S,x} \left\| A(S)B(x) \right\|_F^2 \mathbb{E}_{S,x} \left\| A(S)B_{S,d}^{\text{low}}(x) \right\|_F^2}
\]

\[
\leq \sqrt{\mathbb{E}_{S,x} \left\| A(S)B(x) \right\|_F^2 \mathbb{E}_{S,x} \left\| A(S)B_{S,d}^{\text{low}}(x) \right\|_F^2}
\]

\[
= \mathbb{E}_{S,x} \left\| A(S)B(x) \right\|_F ,
\] (4.27)

and thereby

\[
\mathbb{E}_{S,x} \left( \left\| A(S)B(x) \right\|_F + \left\| A(S)B_{S,d}^{\text{low}}(x) \right\|_F \right)^2 \leq 2 \mathbb{E}_{S,x} \left\| A(S)B(x) \right\|_F^2 + 2 \mathbb{E}_{S,x} \left\| A(S)B_{S,d}^{\text{low}}(x) \right\|_F^2
\]

\[
+ 2 \mathbb{E}_{S,x} \left\| A(S)B(x) \right\|_F \left\| A(S)B_{S,d}^{\text{low}}(x) \right\|_F
\]

\[
\leq 2 \mathbb{E}_{S,x} \left\| A(S)B(x) \right\|_F^2
\]

\[
+ 2 \mathbb{E}_{S,x} \left\| A(S)B(x) \right\|_F \left\| A(S)B_{S,d}^{\text{low}}(x) \right\|_F
\]

\[
\leq 4 \mathbb{E}_{S,x} \left\| A(S)B(x) \right\|_F^2 .
\] (4.28)
For every $S \in \left( \binom{[n]}{m} \right)$ and every $x \in \{-1,1\}^n$, we have
\[
\left\| A(S)B(x) \right\|_F - \left\| A(S)B_{S,d}^{low}(x) \right\|_F \leq \left\| A(S)B(x) - A(S)B_{S,d}^{low}(x) \right\|_F
\]
\[
= \left\| A(S) \left( B(x) - B_{S,d}^{low}(x) \right) \right\|_F
\]
\[
= \left\| A(S)B_{S,d}^{high}(x) \right\|_F
\]
by the reverse triangle inequality, and hence
\[
\left\| A(S)B(x) \right\|_F - \left\| A(S)B_{S,d}^{low}(x) \right\|_F \leq \left\| A(S)B_{S,d}^{high}(x) \right\|_F^2.
\]
This implies
\[
\mathbb{E}_{S \in \mathcal{E}} \left\| A(S)B(x) \right\|_F - \left\| A(S)B_{S,d}^{low}(x) \right\|_F \leq \mathbb{E}_{S \in \mathcal{E}} \left\| A(S)B_{S,d}^{high}(x) \right\|_F^2.
\]
(4.29)

Overall, substituting (4.28) and (4.29) into (4.26) leads to
\[
\left| \mathbb{E}_{S \in \mathcal{E}} D(x_S) \left\| A(S)B(x) \right\|_F^2 - \mathbb{E}_{S \in \mathcal{E}} D(x_S) \left\| A(S)B_{S,d}^{low}(x) \right\|_F^2 \right|
\]
\[
\leq 2 \left\| D \right\|_{\infty} \left( \mathbb{E}_{S \in \mathcal{E}} \left\| A(S)B(x) \right\|_F^2 \right)^{\frac{1}{2}} \left( \mathbb{E}_{S \in \mathcal{E}} \left\| A(S)B_{S,d}^{high}(x) \right\|_F^2 \right)^{\frac{1}{2}}.
\]
(4.30)

Combining (4.25) and (4.30), we obtain
\[
\mathbb{E}_{S \in \mathcal{E}} D(x_S) \left\| A(S)B(x) \right\|_F^2
\]
\[
\geq -2 \left\| D \right\|_{\infty} \left( \mathbb{E}_{S \in \mathcal{E}} \left\| A(S)B(x) \right\|_F^2 \right)^{\frac{1}{2}} \left( \mathbb{E}_{S \in \mathcal{E}} \left\| A(S)B_{S,d}^{high}(x) \right\|_F^2 \right)^{\frac{1}{2}}.
\]
(4.31)

Finally, it remains to lower bound the term \( \left( \mathbb{E}_{S \in \mathcal{E}} \left\| A(S)B_{S,d}^{high}(x) \right\|_F^2 \right)^{\frac{1}{2}} \). In order to do this, note that we have
\[
\left\| A(S)B_{S,d}^{high}(x) \right\|_F^2 \leq \left\| A(S)^2 \right\| \cdot \left\| B_{S,d}^{high}(x) \right\|_F^2
\]
for every $S \in \left( \binom{[n]}{m} \right)$ by Corollary 1.30. Additionally, we used that $A$ maps to symmetric matrices. If we take the expectation over all $S \in \left( \binom{[n]}{m} \right)$, we obtain
\[
\mathbb{E}_{S \in \mathcal{E}} \left\| A(S)B_{S,d}^{high}(x) \right\|_F^2 \leq \mathbb{E}_{S \in \mathcal{E}} \left\| A(S)^2 \right\| \cdot \left\| B_{S,d}^{high}(x) \right\|_F^2
\]
\[
\leq \max_{S \in \left( \binom{[n]}{m} \right)} \left\| A(S)^2 \right\| \cdot \mathbb{E}_{S \in \mathcal{E}} \left\| B_{S,d}^{high}(x) \right\|_F^2
\]
Prop. 4.20
\[
\leq \max_{S \in \left( \binom{[n]}{m} \right)} \left\| A(S)^2 \right\| \left( \frac{km}{n - 2} \right)^{\frac{d+1}{2}} \mathbb{E} \left\| B(x) \right\|_F^2.
\]
Substituting the previous inequality into (4.31) yields the desired statement. \qed
Finally, we translate the theorem back to the case \( \{0,1\}^n \).

**Theorem 4.24** (cf. [LRS15, Theorem 3.5, Lemma 3.6]).

Let \( n,k,s,t \in \mathbb{N} \) with \( k \leq n \), \( m \in [n] \), and let \( d \in [2m-2] \) be even. Let \( A : \left( \begin{array}{c} n \\ m \end{array} \right) \to \mathbb{S}^s \) and \( B : \{0,1\}^n \to \mathbb{R}^{s \times t} \) be matrix-valued functions with \( \deg(B) \leq k \), and let \( D : \{0,1\}^m \to \mathbb{R} \) be a degree-\( d \) pseudo-density. Furthermore, we define \( \lambda := \max_{S \in \left( \begin{array}{c} n \\ m \end{array} \right)} \|A(S)\|^2 \).

Then,

\[
\mathbb{E}_{S,x} D(x_S) \|A(S)B(x)\|_F^2 \geq -2\sqrt{\lambda} \|D\|_\infty \left( \frac{km}{n^2 - \frac{d}{2}} \right)^{\frac{d}{2} + \frac{1}{2}} \left( \frac{\mathbb{E}_x \|B(x)\|_F^2}{\mathbb{E}_{S,x} \|A(S)B(x)\|_F^2} \right)^{\frac{1}{2}}.
\]

**Proof.** Let \( \psi_1 : \{0,1\}^n \to \{-1,1\}^n \) and \( \psi_2 : \{0,1\}^m \to \{-1,1\}^m \) be the bijections from Notation 3.18. An obvious but important fact is that these bijections preserves the degree what means

\[ \deg(f) = \deg(f \circ \psi_i) \text{ for } i = 1,2 \]

and every \( f \in \mathbb{R}^{\{-1,1\}^n} \) respectively every \( f \in \mathbb{R}^{\{-1,1\}^m} \). Therefore,

\[ \hat{D} : \{-1,1\}^m \to \mathbb{R}, x \mapsto D \circ \psi_2^{-1}(x) \]

is also a degree-\( d \) pseudo-density with

\[ \|\hat{D}\|_\infty = \|D\|_\infty, \]

and

\[ \hat{B} : \{-1,1\}^n \to \mathbb{R}^{s \times t}, x \mapsto B \circ \psi_1^{-1}(x) \]

has also degree at most \( k \).

Additionally, note that

\[ (\psi_1(y))_S = \psi_2(y_S) \]

is also obvious for every \( y \in \{0,1\}^n \) and every \( S \in \left( \begin{array}{c} n \\ m \end{array} \right) \) by the underlying definitions.
Now, we obtain
\[
\mathbb{E}_{S \sim \{0,1\}^n} D(y_S) \|A(S)B(y)\|_F^2 = \mathbb{E}_{S,x} D \circ \psi^{-1}_2(x_S) \|A(S)B \circ \psi^{-1}_1(x)\|_F^2 \\
= \mathbb{E}_{S,x} \tilde{D}(x_S) \|A(S)\tilde{B}(x)\|_F^2
\]

Th.4.23 \[\geq -2\sqrt{\mathbb{E}_{S,x} \|A(S)\tilde{B}(x)\|_F^2} \left( \frac{km}{n - \frac{d}{2}} \right)^{\frac{d}{2} + \frac{1}{2}} \left( \mathbb{E}_{y \sim \{0,1\}^n} \|B \circ \psi_1(y)\|_F^2 \right)^{\frac{1}{2}}
\]

\[= -2\sqrt{\mathbb{E}_{S \sim \{0,1\}^n} \|A(S)\tilde{B}(y)\|_F^2} \left( \frac{km}{n - \frac{d}{2}} \right)^{\frac{d}{2} + \frac{1}{2}} \left( \mathbb{E}_{y \sim \{0,1\}^n} \|B(y)\|_F^2 \right)^{\frac{1}{2}}
\]

\[\cdot \left( \mathbb{E}_{S,x} \|A(S)\tilde{B}(x)\|_F^2 \right)^{\frac{1}{2}}.
\]

\[\square
\]

4.4 Matrix approximation

Based on (4.13), where we presented a rough idea of what to do in the matrix approximation section, we first show what we will actually prove in this section. For given \(n,r \in \mathbb{N}, \varepsilon > 0\) and matrix-valued Boolean functions \(F : \{0,1\}^n \to S^r\) and \(Q : \{0,1\}^n \to S^r\), where \(Q\) additionally maps onto PSD matrices, we want to show the existence of some \(s \in \mathbb{N}\) and some matrix-valued Boolean function \(B : \{0,1\}^n \to \mathbb{R}^{r \times s}\) of suitable 'medium' degree such that

\[\frac{1}{\tau} \mathbb{E}_{x} \text{Tr}(F(x)Q(x)) \geq \frac{1}{\gamma} \mathbb{E}_{x} \text{Tr}(F(x)B(x)B(x)^T) - \varepsilon \max_{x \in \{0,1\}^n} \|F(x)\|, \tag{4.32}\]

with \(\tau := \mathbb{E}_{x} \text{Tr}(Q(x))\) and \(\gamma := \mathbb{E}_{x} \text{Tr}(B(x)B(x)^T)\).

In the original proof of [LRS15], the so-called duality formula for quantum entropy plays an essential role. This is a result from quantum information theory, a field not used to everyone, especially not to the people in real algebraic geometry. Therefore, to manage a proof without this was a large motivation and one of the main goals of this part of the work, but unfortunately, at this time we are not able to do this. So, as an insertion, we have to provide the mentioned result and some notation from
quantum information theory first. At the end of this work, in Section 6.1 we will present a conjecture that would yield a more natural and more algebraic approach to the main theorem of this section and that should normally be provable without quantum information theory if it is true.

In order to introduce the central concept of quantum relative entropy, we need a reasonable definition and interpretation of the logarithm of a matrix. For positive definite matrices there is the following natural one.

**Definition 4.25.**
Let \( r \in \mathbb{N} \) and let \( A \in S^r \) be positive definite with eigenvalues \( \lambda_1, \ldots, \lambda_r > 0 \). Then, there exist a matrix \( U \in \mathbb{R}^{r \times r} \) with \( U^T = U^{-1} \) such that

\[
A = \begin{pmatrix}
\lambda_1 & 0 & \cdots & 0 \\
0 & \ddots & \ddots & \\
\vdots & \ddots & \ddots & 0 \\
0 & \cdots & 0 & \lambda_r
\end{pmatrix} U^{-1}.
\]

We define \( \ln A \in S^r \) by

\[
\ln A := U \begin{pmatrix}
\ln \lambda_1 & 0 & \cdots & 0 \\
0 & \ddots & \ddots & \\
\vdots & \ddots & \ddots & 0 \\
0 & \cdots & 0 & \ln \lambda_r
\end{pmatrix} U^{-1}
\]

and call it the (natural) logarithm of \( A \).

**Remark 4.26.**
(a) Together with the definition of the matrix exponential on page 14, it is easy to see that

\[
e^{\ln A} = A
\]

holds for every positive definite matrix \( A \).

(b) The natural logarithm has a representation as a power series in the interval \((0,2]\). More precisely, for all real numbers \( t \in (0,2] \) holds

\[
\ln t = -\sum_{k=1}^{\infty} \frac{(1-t)^k}{k}.
\]

Note that our definition of the logarithm of a matrix coincides with this representation in the following sense. Let \( A \in S^r \) be positive definite with eigenvalues \( 0 < \lambda_1, \ldots, \lambda_r \leq 2 \). Then, easy computation analog to Lemma 1.13 shows

\[
\ln A = -\sum_{k=1}^{\infty} \frac{1}{k} (I_r - A)^k.
\]
With this definition, the expression $A \ln A$ is defined for positive definite matrices. We want to extend it also to psd matrices. Therefore, we define the function

$$f_{\ln} : \mathbb{R}_+ \to \mathbb{R}, \ x \mapsto \begin{cases} x \cdot \ln x & x > 0 \\ 0 & x = 0 \end{cases}.$$ 

Note that this definition is reasonable because of $\lim_{x \to 0} x \cdot \ln x = 0$, what follows by L'Hospital's rule.

Now, we are able to proceed with our desired definitions.

**Definition 4.27.**
Let $r \in \mathbb{N}$.

(i) Let $A \in \mathbb{S}^r$ be positive semidefinite with eigenvalues $\lambda_1, \ldots, \lambda_r \geq 0$. Then, there exist a matrix $U \in \mathbb{R}^{r \times r}$ with $U^T = U^{-1}$ such that $A = U \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_r \end{pmatrix} U^{-1}$.

We define

$$A \ln A := U \begin{pmatrix} f_{\ln}(\lambda_1) & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & f_{\ln}(\lambda_r) \end{pmatrix} U^{-1}.$$

(ii) A matrix $A \in \mathbb{S}^r$ with

$$A \succeq 0 \text{ and } \text{Tr}(A) = 1$$

is called a density matrix.

(iii) Now, let $A, B \in \mathbb{S}^r$ be two density matrices. Furthermore, let $B$ be positive definite. Then we define

$$S(A\|B) := \text{Tr}(A \ln A - A \cdot \ln B)$$

and call it the quantum relative entropy of $A$ with respect to $B$. Furthermore, we call $S(A\|\frac{1}{r} I_r)$ the quantum relative entropy with respect to the uniform density.

**Lemma 4.28.**
Let $r \in \mathbb{N}$, and let $P, Q \in \mathbb{S}^r$ with $P \succeq Q \succeq 0$ and $P \succ 0$. Furthermore, let $P$ and $Q$ be simultaneously diagonalizable. Then, it holds

$$\text{Tr}(Q \ln Q) \leq \text{Tr}(Q \ln P).$$
Proof. Let \( \lambda_1, \ldots, \lambda_r \geq 0 \) and \( \gamma_1, \ldots, \gamma_r > 0 \) be the eigenvalues of \( Q \) and \( P \). Then, there exists a matrix \( U \in \mathbb{R}^{r \times r} \) with

\[
Q = U \begin{pmatrix}
\lambda_1 & 0 & \cdots & 0 \\
0 & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & 0 \\
0 & \cdots & 0 & \lambda_r
\end{pmatrix} \quad \text{and} \quad P = U \begin{pmatrix}
\gamma_1 & 0 & \cdots & 0 \\
0 & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & 0 \\
0 & \cdots & 0 & \gamma_r
\end{pmatrix} U^{-1}.
\]

This implies

\[
P - Q = U \begin{pmatrix}
\gamma_1 - \lambda_1 & 0 & \cdots & 0 \\
0 & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & 0 \\
0 & \cdots & 0 & \gamma_r - \lambda_r
\end{pmatrix} U^{-1}
\]

and

\[
Q \ln P = U \begin{pmatrix}
\lambda_1 \ln(\gamma_1) & 0 & \cdots & 0 \\
0 & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & 0 \\
0 & \cdots & 0 & \lambda_r \gamma_r
\end{pmatrix} U^{-1}.
\]

Since all the eigenvalues of \( P - Q \) are nonnegative, we obtain

\[
\lambda_i \leq \gamma_i \quad \text{for every } i \in [r],
\]

which implies

\[
f_{\ln}(\lambda_i) \leq \lambda_i \ln(\gamma_i) \quad \text{for every } i \in [r],
\]

because \( \ln \) is a monotone function.

Overall, we obtain

\[
\text{Tr}(Q \ln Q) = \sum_{i \in [r]} f_{\ln}(\lambda_i) \leq \sum_{i \in [r]} \lambda_i \ln(\gamma_i) = \text{Tr}(Q \ln P)
\]

Although the quantum relative entropy seems to be a complicated expression in general, it is easy to compute in the uniform density case, and it is also easy to show \( S \left( A \| \frac{1}{r} I_r \right) \geq 0 \) for every density matrix \( A \in \mathcal{S}^r \), a well-known result in quantum information theory. In order to do this, we initially prove a useful inequality, also known as Gibbs’ inequality\(^{11}\).

\(^{11}\) named after J.W. Gibbs (1839-1903)
**Lemma 4.29** (Gibbs’ inequality).

Let \( r \in \mathbb{N} \) and let \( \lambda, \gamma \in \mathbb{R}_+^r \) with \( \sum_{i \in [r]} \lambda_i = \sum_{i \in [r]} \gamma_i = 1 \), where \( \gamma \) additionally satisfies \( \gamma_1, \ldots, \gamma_r > 0 \). Furthermore, let \( I := \{ i \in [r] \mid \lambda_i > 0 \} \). Then we have

\[
\sum_{i \in I} \lambda_i \ln \frac{\lambda_i}{\gamma_i} \geq 0,
\]

and

\[
\sum_{i \in I} \lambda_i \ln \frac{\lambda_i}{\gamma_i} = 0
\]

if and only if \( \lambda_i = \gamma_i \) for every \( i \in [r] \).

**Proof.** Note that we have

\[
1 - x \leq -\ln x \quad \text{for every real number } x > 0,
\]

where equality holds only for \( x = 1 \).

It holds

\[
\sum_{i \in I} \lambda_i \ln \frac{\lambda_i}{\gamma_i} = \sum_{i \in I} \lambda_i \cdot \left( -\ln \frac{\gamma_i}{\lambda_i} \right) \geq \sum_{i \in I} \lambda_i \left( 1 - \frac{\gamma_i}{\lambda_i} \right) = \sum_{i \in I} \lambda_i - \sum_{i \in I} \gamma_i
\]

\[
= \sum_{i \in [r]} \lambda_i - \sum_{i \in I} \gamma_i = 1 - \sum_{i \in I} \gamma_i \geq 1 - \sum_{i \in [r]} \gamma_i = 0.
\]

We only have equality in the case \( \frac{\lambda_i}{\gamma_i} = 1 \) for every \( i \in I \). This implies \( \lambda_i = \gamma_i \) for every \( i \in I \). Additionally, we have \( \sum_{i \in [r]} \gamma_i = 1 = \sum_{i \in I} \lambda_i = \sum_{i \in I} \gamma_i \). Therefore, we have \( \gamma_i = 0 \) for every \( i \in [r] \setminus I \). This implies \( \lambda_i = \gamma_i \) for every \( i \in [r] \). \( \square \)

**Lemma 4.30.**

Let \( r \in \mathbb{N} \) and let \( A \in S^r \) be a density matrix. Then we have

(i) \( S \left( A \left\| \frac{1}{r} I_r \right. \right) \geq 0 \),

(ii) \( S \left( A \left\| \frac{1}{r} I_r \right. \right) = 0 \) if and only if \( A = \frac{1}{r} I_r \), and

(iii) \( S \left( A \left\| \frac{1}{r} I_r \right. \right) = \frac{1}{r} \text{Tr} \left( rA \ln(rA) \right) \).

**Proof.** Let \( \lambda_1, \ldots, \lambda_r \geq 0 \) be the eigenvalues of \( A \) and let \( I := \{ i \in [r] \mid \lambda_i > 0 \} \). Since \( A \) is a density matrix, we have \( \sum_{i \in [r]} \lambda_i = 1 \). Therefore, by choosing \( \gamma_i = \frac{1}{r} \) in the previous lemma, we obtain

\[
\sum_{i \in I} \lambda_i \ln(\lambda_i r) \geq 0,
\]

where equality only holds if \( \lambda_i = \frac{1}{r} \) for every \( i \in I \).
The first crucial thing why the quantum relative entropy with respect to the uniform density is much simpler than the general case is the obvious fact

\( \ln \left( \frac{1}{r} I_r \right) = \ln (1 / r) \cdot I_r \).

This implies

\[
S \left( A \left\| \frac{1}{r} I_r \right\| \right) = \operatorname{Tr} \left( A \ln A - A \cdot \ln \left( \frac{1}{r} \right) \cdot I_r \right) = \operatorname{Tr} \left( A \ln A + (\ln r) A \right).
\]

The second crucial thing is that \( A \ln A \) and \( (\ln r) A \) are simultaneously diagonalizable and their eigenvalues are known. Analog to Lemma 1.13, we obtain that their sum is also diagonalizable with eigenvalues \( f \ln (\lambda_i) + \ln(r) \cdot \lambda_i \) for \( i \in [r] \). With (1.2), we obtain

\[
\operatorname{Tr} \left( A \ln A + (\ln r) A \right) = \sum_{i \in [r]} f \ln (\lambda_i) + \ln(r) \cdot \lambda_i.
\]

Combining the last two equation yields

\[
S \left( A \left\| \frac{1}{r} I_r \right\| \right) = \sum_{i \in [r]} f \ln (\lambda_i) + \ln(r) \cdot \lambda_i = \sum_{i \in I} \lambda_i \ln(\lambda_i r) \geq 0,
\]

which shows the desired statement (i).

If equality holds above, we have \( \lambda_i = \frac{1}{r} \) for every \( i \in I \) again by (4.34). Therefore, we obtain

\[
1 = \sum_{i \in [r]} \lambda_i = \sum_{i \in I} \lambda_i = \frac{|I|}{r},
\]

which implies \( I = [r] \). Thus, it follows \( \lambda_i = \frac{1}{r} \) for every \( i \in [r] \), which again implies \( A = \frac{1}{r} I_r \).

In order to show statement (iii) we proceed with the above identity. We obtain

\[
S \left( A \left\| \frac{1}{r} I_r \right\| \right) = \sum_{i \in I} \lambda_i \ln(\lambda_i r) = \frac{1}{r} \sum_{i \in I} \lambda_i r \ln(\lambda_i r)
\]

\[
= \frac{1}{r} \sum_{i \in [r]} f \ln(r \lambda_i) = \frac{1}{r} \operatorname{Tr} (r A \ln (r A)).
\]

Finally, we state the already mentioned result from quantum information theory. A proof of it can be found for example in [Car10].

**Theorem 4.31** (Duality formula for quantum entropy, cf. [Car10]).

Let \( r \in \mathbb{N}, \lambda \in \mathbb{R}_+ \) and let \( F, Q \in S^r \) where \( Q \) is additionally a density matrix. Then, we have

\[
\lambda \operatorname{Tr} \left( F \frac{e^{-\lambda F}}{\operatorname{Tr}(e^{-\lambda F})} \right) + S \left( \frac{e^{-\lambda F}}{\operatorname{Tr}(e^{-\lambda F})} \left\| \frac{1}{r} I_r \right\| \right) \leq \lambda \operatorname{Tr} (F Q) + S \left( Q \left\| \frac{1}{r} I_r \right\| \right).
\]
By Lemma 4.30(i), we immediately obtain the following corollary.

**Corollary 4.32.**

Let $r \in \mathbb{N}, \lambda \in \mathbb{R}$ with $\lambda > 0$, and let $F, Q \in \mathbb{S}^r$, where $Q$ is additionally a density matrix. Then, we have

$$\text{Tr} \left( F \frac{e^{-\lambda F}}{\text{Tr}(e^{-\lambda F})} \right) \leq \text{Tr}(FQ) + \frac{S(Q \frac{1}{\lambda} I_r)}{\lambda}.$$

Now, we proceed with the actual preliminaries for our main theorem.

**Notation 4.33.**

For $k \in \mathbb{N}$, we write

$$p_k := \sum_{j=0}^{k} \frac{T^j}{j!}$$

for the $k$-th order Taylor polynomial of the exponential function.

With the previous Corollary and with some trick and more preparation that we will see later, it turns out that $x \mapsto e^{\alpha F(x)}$ for some suitable $\alpha \in \mathbb{R}$ is a good candidate to approximate $Q$ with respect to the expectation of the trace in the sense of (4.32). Unfortunately we have no control about its degree. Additionally, we want to approximate $Q$ by a square. Now, the idea of [LRS15] is to approximate $e^t$ by its Taylor polynomial of suitable order. For sufficiently large $k$ we have

$$e^t \approx p_k(t). \quad (4.35)$$

Based on this, they use

$$e^t = \left(e^{\frac{t}{2}}\right)^2 \approx p_k \left(\frac{t}{2}\right)^2 \quad (4.36)$$

in order to show that

$$B = \left(x \mapsto p_k \left(\alpha \frac{F(x)}{2}\right)\right) \quad (4.37)$$

is a suitable candidate for (4.32).

As already mentioned, our idea is to allow not only a symmetric square but also a square in the sense of $BB^T$ (and therefore a sum of symmetric squares by (4.14)) in (4.32), what we are actually allowed to do because of our more general results from the previous section. For our approach, the following result, where a proof for example can be found in [KLS15, Lemma 1], helps.

**Proposition 4.34** (cf. [KLS15, Lemma 1]).

For $k \in \mathbb{N}$, let

$$p_k := \sum_{j=0}^{k} \frac{T^j}{j!}$$


the $k$-th order Taylor polynomial of the exponential function. Then, if $k$ is even, we have
\[ p_k(t) > 0 \text{ for all } t \in \mathbb{R}. \]

Thus, for example Hilbert’s famous theorem from 1888 [Hil88] shows that $p_k$ is a sum of squares if $k$ is even. Moreover, it is a small exercise to prove that in the univariate case every nonnegative real polynomial is even a sum of two squares. A very simple proof of this fact can be found for example in [Sch14].

Corollary 4.35 (cf. [Sch14]).
Let $k \in \mathbb{N}$ be even, and let $p_k$ the $k$-th order Taylor polynomial. Then, $p_k$ is a sum of two squares.

Hence, we can directly use (4.35) and $p_k$ instead of $p^2_k$, which is a sum of two squares if $k$ is even by Corollary 4.35. So, there exist $p_1, p_2 \in \mathbb{R}[T]_2$ with $p_k = p_1^2 + p_2^2$. Now, our approach is to vary the results of [LRS15] in order to show that $B : \{0,1\}^n \to \mathbb{R}^{r \times 2r}, x \mapsto (p_1(\alpha F(x)) \ p_2(\alpha F(x)))$ (4.38)
is also a suitable candidate for (4.32), what is reasonable because of $B(x)B(x)^T = p_k(\alpha F(x))$ (cf. (4.14)). Comparing (4.37) and (4.38), we see that our approach yields a function $B$ with half of the degree than in the approach of [LRS15]. Additionally, it seems that our approximation is even better. Both is the reason for why we will obtain slightly better degree bounds in the main theorem of this section. We will say something more to these small differences in the further course of the work.

We proceed with our variation of the first result [LRS15] used as a preparation for the main result.

Proposition 4.36 (cf. [LRS15] Lemma 4.3).
Let $r \in \mathbb{N}$ and $\delta \in (0,1]$. Furthermore, let $F \in S^r$ with $\beta := \|F\|$ and let $p \in \mathbb{R}[T]$ be a univariate polynomial with $\text{Tr}(p(F)) \neq 0$ and
\[ |e^t - p(t)| \leq \delta e^t \]
for every $t \in [-\beta,\beta]$. Then, $p$ satisfies
\[ \left\| \frac{e^F}{\text{Tr}(e^F)} - \frac{p(F)}{\text{Tr}(p(F))} \right\| \leq 2\delta. \]

Proof. Let $\lambda_1, \ldots, \lambda_r$ be the eigenvalues of $F$. By the definition of $\beta$, we have $\lambda_1, \ldots, \lambda_r \in [-\beta,\beta]$, and therefore the assumption implies
\[ |e^{\lambda_i} - p(\lambda_i)| \leq \delta e^{\lambda_i}. \]
for every $i \in [r]$.

Next, we want to show that $p(t)$ is nonnegative for every $t \in [-\beta, \beta]$. In order to do this, note that if $p(t) \geq e^t$, we are done. So, let $p(t) < e^t$. This in turn implies

$$e^t - p(t) = |e^t - p(t)| \leq \delta e^t,$$

which is equivalent to

$$p(t) \geq (1 - \delta)e^t \geq 0.$$ 

Overall, we can conclude

$$p(\lambda_i) \geq 0$$

for every $i \in [r]$.

Furthermore, note that

$$\frac{a}{b} - \frac{c}{d} = \frac{ad - cd + cd - bc}{bd} = \frac{a - c}{b} + \frac{c(d - b)}{bd}$$

for every $a, b, c, d \in \mathbb{R}$ with $b, d \neq 0$.

By Lemma 1.13, $e^F$ as well as $p(F)$ are symmetric with eigenvalues $e^{\lambda_1}, \ldots, e^{\lambda_r}$ respectively $p(\lambda_1), \ldots, p(\lambda_r)$. Therefore, we have $\text{Tr}(e^F) = \sum_{i \in [r]} e^{\lambda_i}$ and $\text{Tr}(p(F)) = \sum_{i \in [r]} p(\lambda_i)$ by (1.2). Applying the same methods as in the proof of this lemma also implies that the matrix $\frac{e^F}{\text{Tr}(e^F)} - \frac{p(F)}{\text{Tr}(p(F))}$ is symmetric as well with eigenvalues $\frac{e^{\lambda_i} - p(\lambda_i)}{\sum_{i = 1}^{r} e^{\lambda_i} - \sum_{i = 1}^{r} p(\lambda_i)}$ for $i \in [r]$. Thus, we have

$$\left\| \frac{e^F}{\text{Tr}(e^F)} - \frac{p(F)}{\text{Tr}(p(F))} \right\| = \sum_{i = 1}^{r} \left| \frac{e^{\lambda_i} - p(\lambda_i)}{\sum_{i = 1}^{r} e^{\lambda_i} - \sum_{i = 1}^{r} p(\lambda_i)} \right|$$

by definition, and it remains to lower bound this second expression.

$$\sum_{i = 1}^{r} \left| \frac{e^{\lambda_i} - p(\lambda_i)}{\sum_{i = 1}^{r} e^{\lambda_i} - \sum_{i = 1}^{r} p(\lambda_i)} \right| \leq \sum_{i = 1}^{r} \left( \left| \frac{e^{\lambda_i} - p(\lambda_i)}{\sum_{i = 1}^{r} e^{\lambda_i}} \right| + \left| \frac{p(\lambda_i)}{\sum_{i = 1}^{r} e^{\lambda_i}} \right| \right)$$

$$\leq \sum_{i = 1}^{r} \left( \left| \frac{e^{\lambda_i}}{p(\lambda_i)} \right| + \left| \frac{p(\lambda_i)}{e^{\lambda_i}} \right| \right)$$

$$\leq \sum_{i = 1}^{r} \left| \frac{e^{\lambda_i}}{p(\lambda_i)} \right| + \sum_{i = 1}^{r} \left| \frac{p(\lambda_i)}{e^{\lambda_i}} \right|$$

$$\leq \sum_{i = 1}^{r} \left| \frac{e^{\lambda_i} - p(\lambda_i)}{\sum_{i = 1}^{r} e^{\lambda_i} - \sum_{i = 1}^{r} p(\lambda_i)} \right|$$

$$\leq \delta + \frac{\sum_{i = 1}^{r} |p(\lambda_i)|}{\sum_{i = 1}^{r} p(\lambda_i)}$$

$$= 2\delta.$$
Before proceeding, we provide a result yielding an estimation for the factorial, known as Stirling’s formula, which we use in the proof of the subsequent proposition. For this purpose, a weak lower bound for the factorial is sufficient. But in Section 4.6, a much more precise version of Stirling’s formula is necessary for which reason we already provide this stronger version now. A proof of this more precise version can be found for example in [Rob55].

**Proposition 4.37 (Stirling’s formula).**

For every positive integer \( n \) holds

\[
 n! \leq \sqrt{2\pi n} \left( \frac{n}{e} \right)^n e^{\frac{1}{12n}}
\]

and

\[
 n! \geq \sqrt{2\pi n} \left( \frac{n}{e} \right)^n e^{\frac{1}{12n+1}}.
\]

In particular, we have

\[
 n! \geq \sqrt{2\pi n} \left( \frac{n}{e} \right)^n \geq \left( \frac{n}{e} \right)^n.
\]

For the next proposition, we will use Taylor’s Theorem amongst others at the beginning of the proof. Here, we are much more detailed than [LRS15], but in their original work is an error at this point\(^\text{[12]}\), so we have to be very careful here.

As a consequence of replacing the square of a polynomial in the original work by a sum of squares, we obtain better bounds than [LRS15], around a half of the upper bound on the degree as well as a third of the upper bound of the trace norm.

**Proposition 4.38 (cf. [LRS15, Corollary 4.4]).**

Let \( \varepsilon \in (0,1], r \in \mathbb{N} \) and \( F \in S^r \). Then, there exist a sos polynomial \( p \in \sum \mathbb{R}[T]^2 \) with \( \deg(p) \leq \varepsilon^2 \| F \| + \ln \frac{6}{\varepsilon} + 1 \) and \( \text{Tr}(p(F)) \neq 0 \) such that

\[
 \left\| \frac{e^F}{\text{Tr}(e^F)} - \frac{p(F)}{\text{Tr}(p(F))} \right\|_* \leq \frac{\varepsilon}{3}.
\]

In particular, \( p \) is a sum of two squares.

**Proof.** For \( k \in \mathbb{N} \), let

\[
 p_k := \sum_{j=0}^{k} \frac{T^j}{j!}
\]

the \( k \)-th order Taylor polynomial of the exponential function. Furthermore, let \( \beta \in \mathbb{R}_+ \) and \( t \in (-\beta, \beta) \).

\(^{12}\)Meanwhile, there exists a revised but not published version of [LRS15] in which the error is fixed.
By Taylor’s Theorem or more precisely with the Lagrange form of the remainder, both is explained and proven for example in [Kön04, Section 14.1], we obtain

\[ |e^t - p_k(t)| = e^\zeta |t|^{k+1} \]

for some \( \zeta \in \mathbb{R} \) with \( \zeta \in \{ [0,t] \mid t \geq 0 \} \cup \{ [t,0] \mid t < 0 \} \). This implies

\[ |e^t - p_k(t)| \leq e^\zeta \frac{\beta^{k+1}}{(k+1)!} \leq \frac{\beta^{k+1}}{(k+1)!} e^{\zeta-t} e^t. \]

In the case \( t \geq 0 \), we have \( e^{\zeta-t} \leq e^\zeta \leq e^\beta \). Otherwise, if \( t < 0 \), we have \( e^{\zeta-t} \leq e^{-t} \leq e^\beta \).

Overall, we have

\[ |e^t - p_k(t)| \leq \beta^{k+1} e^\beta e^t \]

Now, we choose \( \beta := \|F\| \) and set

\[ k := \begin{cases} \left\lfloor e^2 \|F\| + \ln \frac{6}{\varepsilon} \right\rfloor & \text{if \( e^2 \|F\| + \ln \frac{6}{\varepsilon} \) is even} \\
\left\lfloor e^2 \|F\| + \ln \frac{6}{\varepsilon} \right\rfloor + 1 & \text{if \( e^2 \|F\| + \ln \frac{6}{\varepsilon} \) is odd} \end{cases} \]

This choice implies

\[ k + 1 \geq e^2 \beta \tag{4.46} \]

and

\[ k + 1 \geq \beta + \ln \frac{6}{\varepsilon}. \tag{4.47} \]

Then, for every \( t \in [-\beta, \beta] \) holds

\[ |e^t - p_k(t)| \leq \beta \left( \frac{e \beta}{k+1} \right)^{k+1} e^t = e^\beta \left( \frac{e \beta}{k+1} \right)^{k+1} e^t. \tag{4.45} \]

Furthermore, let \( \lambda_1, \ldots, \lambda_r \in \mathbb{R} \) be the eigenvalues of \( F \). Then, we have

\[ \text{Tr}(p_k(F)) \overset{\text{Lem. 4.13}}{=} \sum_{i \in [r]} \text{tr}(p_k(\lambda_i)) \overset{\text{Prop. 4.39}}{>} 0 \]
because \( k \) is even. Thus, together with (4.48), applying Proposition 4.36 yields the desired statement

\[
\left\| \frac{e^F}{\text{Tr}(e^F)} - \frac{p_k(F)}{\text{Tr}(p_k(F))} \right\|_* \leq \frac{\varepsilon}{3}.
\]

Now, Corollary 4.35 shows that \( p_k \) is a sum of two squares, again because \( k \) is even.

We proceed with our variation of the main theorem of matrix approximation. In contrast to [LRS15] and as already mentioned, one difference is that we allow a non-symmetric matrix \( B \). So, the result of [LRS15] is stronger in some sense, but therefore we obtain a better degree bound than the original result. We will specify this after the proof of the theorem.

Another difference is the formulation of the theorem. [LRS15] stated the matrix approximation in a much broader and general way, while we adapt the theorem to the setting where it is applied later. So, again our formulation is weaker in some sense, but we think that this gives a better understanding of the result. For this reason, our formulation is also homogenized with respect to the occurring functions \( F, B, Q \).

Additionally, our formulation do not require a matrix \( B \) that is a polynomial in \( F \), what we think is very unnatural. Indeed, we are not able to omit it in the proof, but in our opinion it is important for further research, because it potentially allows a more natural way of matrix approximation (cf. Conjecture 4.40).

Besides, our proof is more accurate, and we fill some steps [LRS15] just made sparsely.

**Theorem 4.39 (cf. [LRS15, Theorem 4.1]).**

Let \( n, r \in \mathbb{N} \) and let \( F : \{0,1\}^n \to S^r \) and \( Q : \{0,1\}^n \to S^r \) be matrix-valued Boolean functions with \( F, Q \neq 0 \), where \( F \) maps onto symmetric matrices and \( Q \) even onto psd matrices. Furthermore, let \( \tau := \mathbb{E}_x \text{Tr}(Q(x)) \). Then, for all \( \varepsilon \in (0,1] \) there exist \( s \in \mathbb{N} \) and a map \( B : \{0,1\}^n \to \mathbb{R}^{r \times s} \) with \( \gamma := \mathbb{E}_x \text{Tr}(B(x)B(x)^T) \neq 0 \) such that

\[
\frac{1}{\tau} \mathbb{E}_x \text{Tr}(F(x)Q(x)) \geq \frac{1}{\gamma} \mathbb{E}_x \text{Tr}(F(x)B(x)B(x)^T) - \varepsilon \max_{x \in \{0,1\}^n} \|F(x)\|
\]

and

\[
\deg B \leq \frac{\deg(F)}{2} \left( \frac{1.2e^2}{\varepsilon} \frac{1}{\tau} \mathbb{E}_x \left( \frac{1}{\tau} \mathbb{E}_x \text{Tr}(Q(x) \ln \left( \frac{Q(x)}{\tau} \right)) \right) + \ln \left( \frac{12}{\varepsilon} \right) + 1 \right).
\]

**Proof.** As already mentioned in the introduction of this section, we want to apply a result from quantum information theory, more precisely Corollary 4.32, why we have to encode the maps \( Q \) and \( F \), or to be more precise its \( 2^n \) many matrix values, each by one single matrix. More specific, the single matrix that encodes \( Q \) is required to be a density matrix, a psd matrix with trace 1. This encoding can be done by an easy trick. First, we fix an arbitrary enumeration

\[
\{x_1, \ldots, x_{2^n}\} = \{0,1\}^n.
\]
Then, we define the matrices
\[
\bar{F} := \sum_{i \in [2^n]} e_i e_i^T \otimes F(x_i)
\]
and
\[
\bar{Q} := \frac{1}{2^n} \sum_{i \in [2^n]} e_i e_i^T \otimes Q(x_i),
\]
where \((e_i)_{i \in [2^n]}\) denotes the standard basis for \(\mathbb{R}^{2^n}\) and \(\otimes\) denotes the tensor product of matrices also known as Kronecker product. With this notation, we have \(\bar{F}, \bar{Q} \in \mathbb{R}^{r \times 2^n \times r \times 2^n}\). Concretely, these matrices look like
\[
\bar{F} = \begin{pmatrix} F(x_1) & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & F(x_{2^n}) \end{pmatrix} \quad \text{and} \quad \bar{Q} = \frac{1}{2^n} \begin{pmatrix} Q(x_1) & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & Q(x_{2^n}) \end{pmatrix}.
\]

As an immediate consequence of the structure of \(\bar{F}\), we obtain
\[
\|\bar{F}\| = \max_{x \in \{0,1\}^n} \|F(x)\|. \tag{4.49}
\]
As already mentioned, the matrix encoding \(Q\) have to be a density matrix, and therefore we have to divide \(\bar{Q}\) by its trace. Since \(\bar{Q}\) is in particular a block matrix, the trace is easy to compute.
\[
\text{Tr}(\bar{Q}) = \frac{1}{2^n} \sum_{i \in [2^n]} \text{Tr}(Q(x_i)) = \frac{1}{2^n} \sum_{x \in \{0,1\}^n} \text{Tr}(Q(x)) = \mathbb{E}_x \text{Tr}(Q(x)) = \tau.
\]
Hence, we define
\[
\hat{Q} := \frac{1}{\tau} \bar{Q}.
\]
So, \(\bar{F}\) and \(\hat{Q}\) are block diagonal matrices with \(2^n\) many blocks of \(r \times r\)-matrices. Therefore, by simple matrix multiplication the product of them is easy to compute and the block diagonal structure is preserved. We obtain
\[
\bar{F} \hat{Q} = \frac{1}{2^n \tau} \begin{pmatrix} F(x_1)Q(x_1) & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & F(x_{2^n})Q(x_{2^n}) \end{pmatrix},
\]
and therefore it holds
\[
\text{Tr}(\bar{F} \hat{Q}) = \frac{1}{2^n \tau} \sum_{i \in [2^n]} \text{Tr}(F(x_i)Q(x_i)) = \frac{1}{2^n \tau} \sum_{x \in \{0,1\}^n} \text{Tr}(F(x)Q(x))
\]
\[
= \frac{1}{\tau} \mathbb{E}_x \text{Tr}(F(x)Q(x)), \tag{4.50}
\]
and we note that the expression $\text{Tr}(\hat{F}\hat{Q})$ is exactly the one we want to lower bound.

Now, we are able to apply the quantum information theory result from Corollary 4.32. We set
\[
\lambda := \frac{6}{5} S \left( \frac{1}{2^{2r}} I_{2^n} \right) \overset{\text{Lem. 4.30}}{\geq} 0 \quad \text{and} \quad \hat{F} := -\lambda e^{\hat{F}},
\]
and if $\lambda > 0$, we obtain
\[
\text{Tr} \left( \hat{F}\hat{Q} \right) \overset{\text{Cor. 4.32}}{\geq} \text{Tr} \left( \hat{F} \frac{e^F}{\text{Tr}(e^F)} \right) - \frac{5}{6} \varepsilon \|\hat{F}\|.
\]

Note that by Lemma 4.30(ii), $\lambda = 0$ is only possible in the case $\hat{Q} = \frac{1}{2^{2r}} I_{2^n}$. But in this case, we also have $\hat{F} = 0$, which implies $\frac{e^F}{\text{Tr}(e^F)} = \frac{1}{2^{2r}} I_{2^n}$. Thus, (4.52) obviously holds for $\lambda = 0$.

Next, we apply Proposition 4.38 to $\hat{F}$ with $\varepsilon' = \frac{\varepsilon}{2}$. Therefore, we obtain $p, p_1, p_2 \in \mathbb{R}[T]$ with
\[
\text{Tr}(p(\hat{F})) \neq 0,
\]
\[
p = p_1^2 + p_2^2
\]
and
\[
\deg(p) \leq e^2 \|\hat{F}\| + \ln \frac{6}{\varepsilon} + 1
\]
such that
\[
\left\| \frac{e^F}{\text{Tr}(e^F)} - \frac{p(\hat{F})}{\text{Tr}(p(\hat{F}))} \right\| \leq \frac{\varepsilon'}{3} = \frac{\varepsilon}{6}.
\]

We define
\[
B : \{0,1\}^n \rightarrow \mathbb{R}^{r \times 2r}, x \mapsto \begin{pmatrix} p_1(-\lambda F(x)) & p_2(-\lambda F(x)) \end{pmatrix}
\]
and
\[
\gamma := \mathbb{E}_x \text{Tr} B(x)B(x)^T.
\]

By this definition and simple matrix multiplication, we have
\[
B(x)B(x)^T = p_1(-\lambda F(x))^2 + p_2(-\lambda F(x))^2 \overset{\text{4.54}}{=} p(-\lambda F(x))
\]
for every $x \in \{0,1\}^n$, and
\[
\text{Tr} \left( p(\hat{F}) \right) = \sum_{x \in \{0,1\}^n} \text{Tr}(p(-\lambda F(x))) \overset{\text{4.57}}{=} \sum_{x \in \{0,1\}^n} \text{Tr} \left( B(x)B(x)^T \right)
\]
\[
= 2^n \mathbb{E}_x \text{Tr} \left( B(x)B(x)^T \right) = 2^n \gamma.
\]
Simple matrix multiplication shows that powers of a block diagonal matrix are still block diagonal. Therefore, the block diagonal structure of $\hat{F}$ is preserved when substituting it into $p$, what we used here in the first step.

Note that by (4.53), we additionally have $\gamma \neq 0$.

We have
\[
\text{Tr} \left( \hat{F} \frac{p(\hat{F})}{\text{Tr}(p(\hat{F}))} \right) - \text{Tr} \left( \hat{F} \bar{Q} \right) \overset{(4.52)}{\leq} \text{Tr} \left( \hat{F} \frac{p(\hat{F})}{\text{Tr}(p(\hat{F}))} \right) - \text{Tr} \left( \hat{F} \frac{e^F}{\text{Tr}(e^F)} \right) + \frac{5}{6} \| \hat{F} \| \\
= \text{Tr} \left( \hat{F} \left( \frac{p(\hat{F})}{\text{Tr}(p(\hat{F}))} - \frac{e^F}{\text{Tr}(e^F)} \right) \right) + \frac{5}{6} \| \hat{F} \|.
\]

By the proof of Lemma 1.13, we see that $\frac{p(\hat{F})}{\text{Tr}(p(\hat{F}))} - \frac{e^F}{\text{Tr}(e^F)}$ is obviously symmetric, and because $\hat{F}$ is just a multiple of $\bar{F}$, $\hat{F}$ and $\frac{p(\hat{F})}{\text{Tr}(p(\hat{F}))} - \frac{e^F}{\text{Tr}(e^F)}$ are simultaneously diagonalizable. Therefore, we can apply (1.3) and Lemma 1.19 and based on the above inequality, we obtain
\[
\text{Tr} \left( \hat{F} \frac{p(\hat{F})}{\text{Tr}(p(\hat{F}))} \right) - \text{Tr} \left( \hat{F} \bar{Q} \right) \overset{(4.59)}{=} \text{Tr} \left( \hat{F} \left( \frac{p(\hat{F})}{\text{Tr}(p(\hat{F}))} - \frac{e^F}{\text{Tr}(e^F)} \right) \right) + \frac{5}{6} \| \hat{F} \| \\
\overset{\text{Lem} (1.19)}{\leq} \| \hat{F} \| \left\| \frac{p(\hat{F})}{\text{Tr}(p(\hat{F}))} - \frac{e^F}{\text{Tr}(e^F)} \right\|_* + \frac{5}{6} \| \hat{F} \| \\
\overset{(4.56)}{\leq} \| \hat{F} \| \frac{\varepsilon}{6} + \frac{5}{6} \| \hat{F} \| = \varepsilon \| \hat{F} \|.
\]

This implies
\[
\frac{1}{\tau} \mathbb{E}_x \text{Tr}(F(x)Q(x)) \overset{(4.59)}{=} \text{Tr} \left( \hat{F} \bar{Q} \right) \overset{(4.59)}{=} \text{Tr} \left( \hat{F} \frac{p(\hat{F})}{\text{Tr}(p(\hat{F}))} \right) - \varepsilon \| \hat{F} \| \\
= \frac{1}{2^{n\gamma}} \text{Tr} \left( \hat{F} p(\hat{F}) \right) - \varepsilon \| \hat{F} \| \\
= \frac{1}{2^{n\gamma}} \sum_{x \in \{0,1\}^n} \text{Tr}(F(x)p(-\lambda F(x))) - \varepsilon \| \hat{F} \| \\
= \frac{1}{\gamma} \mathbb{E}_x \text{Tr} \left( F(x)B(x)B(x)^T \right) - \varepsilon \| \hat{F} \| \\
= \frac{1}{\gamma} \mathbb{E}_x \text{Tr} \left( F(x)B(x)B(x)^T \right) - \varepsilon \max_{x \in \{0,1\}^n} \| F(x) \|,
\]
what yields the desired statement.
Now, it remains to show that $B$ satisfies the required degree condition. We have
\[
\deg(B) \leq \max\{\deg(p_1) \deg(F), \deg(p_2) \deg(F)\} = \deg(F) \max\{\deg(p_1), \deg(p_2)\}
\]
\[
= \frac{1}{2} \deg(F) \deg(p) \leq \frac{1}{2} \deg(F) \left( e^2 \|\hat{F}\| + \ln \frac{\delta}{\varepsilon} + 1 \right)
\]
\[
= \frac{1}{2} \deg(F) \left( \lambda e^2 \|\hat{F}\| + \ln \frac{12}{\varepsilon} + 1 \right)
\]
\[
= \frac{1}{2} \deg(F) \left( \frac{6 \varepsilon^2}{5 \varepsilon} S \left( \hat{Q} \|r^{-1} I_{2^n r} \right) + \ln \frac{12}{\varepsilon} + 1 \right) \quad (4.54)
\]
Computing $S \left( \frac{1}{2^n} I_{2^n r} \right)$ yields
\[
S \left( \frac{1}{2^n} I_{2^n r} \right) = \frac{1}{2^n} \sum_{x \in \{0,1\}^n} \text{Tr} \left( \xi Q(x) \ln (\xi Q(x)) \right) = \frac{1}{r} \mathbb{E} \text{Tr} \left( \xi Q(x) \ln (\xi Q(x)) \right).
\]
Combining the latter with (4.60) yields the desired bound on the degree of $B$. \hfill \Box

Note that our bound on the degree of $B$ is
\[
\deg B \leq \frac{1}{2} \deg(F) \left( \frac{1.2 e^2}{\varepsilon} \mathbb{E} \text{Tr} \left( \xi Q(x) \ln (\xi Q(x)) \right) + \ln \left( \frac{12}{\varepsilon} \right) + 1 \right)
\]
while [LRST15] shows the following one
\[
\deg B \leq \deg(F) \left( \frac{2 e^2}{\varepsilon} \mathbb{E} \text{Tr} \left( \xi Q(x) \ln (\xi Q(x)) \right) + \ln \left( \frac{12}{\varepsilon} \right) \right).
\]
Therefore, our upper bound is almost the half of their bound. The first addend is improved even by a factor smaller than a third.

Unfortunately, our better degree bound do not affect the results of the next sections and in particular the bounds on the psd rank of the family of correlation polytopes as we will see later.

As already mentioned, there is a more natural approach to this theorem, but until now, we have no way to handle it why we just claim it as a conjecture.

**Conjecture 4.40.**

Let $n,r \in \mathbb{N}$, and let $F : \{0,1\}^n \rightarrow S^r$ and $Q : \{0,1\}^n \rightarrow S^r$ be matrix-valued Boolean functions with $F,Q \neq 0$, where $F$ maps onto symmetric matrices and $Q$ even onto
psd matrices. Furthermore, let \( \tau := \mathbb{E}_x \text{Tr}(Q(x)) \), and let \( P : \{0,1\}^n \to \mathbb{S}^r \) be the well-defined map defined by \( P(x) \mapsto \sqrt{Q(x)} \), and let

\[
P = \sum_{I \subseteq [n]} P_I \chi_I
\]

its Fourier-expansion. Furthermore, let

\[
k := \frac{\deg F}{2} \left( \frac{1.2e^2}{\varepsilon} \mathbb{E}_x \text{Tr} \left( \tau Q(x) \ln \left( \frac{\tau Q(x)}{\varepsilon} \right) \right) + \ln \left( \frac{12}{\varepsilon} \right) + 1 \right),
\]

and let

\[
B := \sum_{|I| \leq k} P_I \chi_I
\]

be the degree-\( \lfloor k \rfloor \) truncation of \( P \) and \( \gamma := \mathbb{E}_x \text{Tr}(B(x)^2) \). Then, we have

\[
\frac{\gamma}{\tau} \mathbb{E}_x \text{Tr}(F(x)Q(x)) \geq \mathbb{E}_x \text{Tr}(F(x)B(x)^2) - \varepsilon \gamma \max_{x \in \{0,1\}^n} \|F(x)\|.
\]

We will discuss this conjecture, and we will say something more about it in Section 6.1.

### 4.5 The main theorem

Before proving our main theorem, we firstly need a result that guarantees that every nonnegative function \( M \in \mathbb{R}^{(n)}_{m \times \{0,1\}^n} \) approximately admits a well-behaved psd factorization. This will be of great help for our proof of the main theorem, and it is called psd factorization scaling by [LRS15]. In order to prove this result, we provide a theorem of Briët, Dadush and Pokutta [BDP13], or to be more precise, we state an index-free version of it what is more useful for our purpose.

**Theorem 4.41** ([BDP13]).

Let \( n,m \in \mathbb{N} \) with \( n \geq m \), and let \( M \in \mathbb{R}^{(n)}_{m \times \{0,1\}^n} \) be a nonnegative function with \( \text{rk}_{\text{psd}}(M) \leq r \). We define

\[
\gamma_r(M) := \sup \left\{ \max_{S \in \{0,1\}^n} \|A_S\| \cdot \|B_x\| : \|M(S,x) = \text{Tr}(A_S B_x)\| \text{ for all } S \in \{0,1\}^n, x \in \{0,1\}^n \right\}.
\]

Then, it holds

\[
\gamma_r(M) \leq r^2 \|M\|_\infty.
\]

Next, we state an index-free version of the mentioned factorization scaling theorem of [LRS15], and we give a detailed proof of it.

---

13This quantity is sometimes called analytic psd rank.
Theorem 4.42 (cf. [LRS15, Theorem 3.3]).

Let $n, m \in \mathbb{N}$ with $m \leq n$. Let $M \in \mathbb{R}^{(n, m)}_{\supseteq 0}$ be a nonnegative function with $M \neq 0$ and $r := \text{rk}_{\text{psd}}(M)$. Then, for every $\eta \in (0, 1]$ there exist matrix-valued functions $P : \binom{n}{m} \rightarrow \mathbb{S}^r$ and $Q : \{0,1\}^n \rightarrow \mathbb{S}^r$ with the following properties:

(i) $P(S)$ and $Q(x)$ are positive semidefinite for all $S \in \binom{n}{m}$, $x \in \{0,1\}^n$,

(ii) $M(S, x) \leq \text{Tr}(P(S)Q(x)) \leq M(S, x) + \eta \|M\|_{\infty}$ for all $S \in \binom{n}{m}$, $x \in \{0,1\}^n$,

(iii) $\mathbb{E}_S P(S) = I_r$,

(iv) $\|P(S)\| \leq \frac{2r^2}{\eta}$ for all $S \in \binom{n}{m}$,

(v) $Q(x) \leq (\eta + r^2) \|M\|_{\infty} I_r$ for all $x \in \{0,1\}^n$.

Proof. Following the notation of the previous theorem, we set

$$\gamma := \gamma_r(M) \leq \frac{r^2}{\eta} \|M\|_{\infty},$$

and we fix a rank-$r$ psd factorization of $M$. That means, we fix $A : \binom{n}{m} \rightarrow \mathbb{S}^r$ and $B : \{0,1\}^n \rightarrow \mathbb{S}^r$ with $A(S), B(x)$ psd and $\text{Tr}(A(S)B(x)) = M(S, x)$ for every $S \in \binom{n}{m}$, $x \in \{0,1\}^n$. By the previous theorem, we obtain

$$\max_{S \in \binom{n}{m}, x \in \{0,1\}^n} \|A(S)\| \cdot \|B(x)\|_* \leq \gamma.$$

Note that for every $c > 0$, $S \mapsto cA(S)$ and $x \mapsto \frac{1}{c}B(x)$ obviously yields a rank-$r$ factorization. Therefore, by an appropriate scaling, we can assume

$$\|A(S)\| \leq \gamma$$

and $\|B(x)\|_* \leq 1$ for all $S \in \binom{n}{m}$, $x \in \{0,1\}^n$. (4.61)

Next, we define

$$N := \eta \|M\|_{\infty} I_r + \mathbb{E}_S A(S) \in \mathbb{S}^r.$$

Note that we have $\|M\|_{\infty} > 0$ because of $M \neq 0$. Let $0 \leq \lambda_1 \leq \ldots \leq \lambda_r$ be the eigenvalues of the psd matrix $\mathbb{E}_S A(S)$. Then, $0 < \eta \|M\|_{\infty} + \lambda_1 \leq \ldots \leq \eta \|M\|_{\infty} + \lambda_r$ are the eigenvalues of $N$ (cf. Lemma 1.13). On the one hand, this implies that $N$ is psd and invertible, where $N^{-1}$ is also psd, which again implies that $N^{-\frac{1}{2}}$ is well-defined (cf. Proposition 1.27). On the other hand, we obtain

$$\|N\| = \eta \|M\|_{\infty} + \lambda_r \geq \eta \|M\|_{\infty}.$$

(4.62)

Now, we are able to construct the desired functions $P$ and $Q$. We define $P : \binom{n}{m} \rightarrow \mathbb{S}^r$ and $Q : \{0,1\}^n \rightarrow \mathbb{S}^r$ by

$$P(S) := N^{-\frac{1}{2}}(\eta \|M\|_{\infty} I_r + A(S))N^{-\frac{1}{2}}.$$
Note that we have
\[ v^T Q(x)v = v^T N^{\frac{1}{2}} B(x) N^{\frac{1}{2}} v = (N^{\frac{1}{2}} v)^T B(x) (N^{\frac{1}{2}} v) \geq 0 \]
for every \( x \in \{0,1\}^n \) and every \( v \in \mathbb{R}^r \), where we used the assumption of \( B \) and that \( N^{\frac{1}{2}} \) is symmetric by definition. Similar, we can show \( P(S) \geq 0 \) for every \( S \in \binom{[n]}{m} \), where we use \( \eta \| M \|_\infty I_r + A(S) \geq 0 \) for every \( S \in \binom{[n]}{m} \). Overall, this shows item (i).

By definition, we have
\[
\text{Tr}(P(S)Q(x)) = \text{Tr} \left(N^{-\frac{1}{2}} (\eta \| M \|_\infty B(x) + A(S)B(x))N^{\frac{1}{2}} \right)
\]
\[
= \text{Tr}(\eta \| M \|_\infty B(x) + A(S)B(x)) = \eta \| M \|_\infty \text{Tr}(B(x)) + M(S,x)
\]
for every \( S \in \binom{[n]}{m}, x \in \{0,1\}^n \). This implies \( \text{Tr}(P(S)Q(x)) \geq M(S,x) \) because of \( \text{Tr}(B(x)) \geq 0 \) as well as
\[
\text{Tr}(P(S)Q(x)) \leq M(S,x) + \eta \| M \|_\infty \| B(x) \|_*, \quad (4.61) 
\]
for every \( S \in \binom{[n]}{m}, x \in \{0,1\}^n \), what verifies item (ii).

In order to show item (iii), note that it holds
\[
\mathbb{E}_S P(S) = N^{-\frac{1}{2}} \left( \eta \| M \|_\infty I_r + \mathbb{E}_S A(S) \right) N^{-\frac{1}{2}} 
\]
\[
= N^{-\frac{1}{2}} N N^{-\frac{1}{2}} = I_r.
\]

For every \( S \in \binom{[n]}{m} \), we have
\[
\| P(S) \| = \| N^{-\frac{1}{2}} (\eta \| M \|_\infty I_r + A(S))N^{-\frac{1}{2}} \| = \| \eta \| M \|_\infty N^{-1} + N^{-\frac{1}{2}} A(S)N^{-\frac{1}{2}} \|
\]
\[
\leq \eta \| M \|_\infty \left\| N^{-1} \right\| + \left\| N^{-\frac{1}{2}} A(S)N^{-\frac{1}{2}} \right\| \leq \eta \| M \|_\infty \left\| N \right\|^{-1} + \| A(S) \| \left\| N^{-\frac{1}{2}} \right\|^2
\]
\[
= \eta \| M \|_\infty \left\| N \right\|^{-1} + \| A(S) \| \left\| N \right\|^{-1} = \left\| N \right\|^{-1} (\eta \| M \|_\infty + \| A(S) \|)
\]
\[
\leq \frac{1}{\eta \| M \|_\infty} \left( \eta \| M \|_\infty + \| A(S) \| \right) \leq 1 + \frac{\gamma}{\eta \| M \|_\infty}
\]
\[
\leq 1 + \frac{r^2}{\eta} \leq \frac{1 + r^2}{\eta} \leq \frac{2r^2}{\eta},
\]
verifying item (iv), where we once used that the operator norm is submultiplicative (cf. the footnote on page [16]).
Finally, for every $x \in \{-1,1\}^n$, we obtain

$$
\|Q(x)\|_* = \text{Tr} \left( N^{1/2} B(x) N^{1/2} \right) \overset{[4.23]}{=} \text{Tr}(NB(x)) \overset{[4.3]}{\leq} \|NB(x)\|_* \overset{\text{Lem.}[4.10]}{\leq} \|N\| \|B(x)\|_*,
$$

(4.61)

$$
\|N\| \leq \eta \|M\|_\infty + \frac{1}{\varepsilon} \|A(S)\| \overset{[4.61]}{\leq} \eta \|M\|_\infty + \gamma.
$$

(4.63)

For a fixed $x \in \{0,1\}^n$, let $\lambda_1,\ldots,\lambda_r$ be the eigenvalues of $Q(x)$. Then, the eigenvalues of $(\eta \|M\|_\infty + \gamma) I_r - Q(x)$ are $\eta \|M\|_\infty + \gamma - \lambda_1,\ldots,\eta \|M\|_\infty + \gamma - \lambda_r$ (cf. Lemma [1.13]). With the just derived inequality, for every $i \in [r]$ we obtain

$$
\eta \|M\|_\infty + \gamma - \lambda_i \geq \eta \|M\|_\infty + \gamma - \sum_{j \in [r]} |\lambda_j|
$$

$$
= \eta \|M\|_\infty + \gamma - \|Q(x)\|_*. \quad \overset{(4.63)}{\geq} \quad 0.
$$

But this implies $Q(x) \preceq (\eta \|M\|_\infty + \gamma) I_r$.

Overall, we obtain

$$
Q(x) \preceq (\eta \|M\|_\infty + \gamma) I_r \preceq (\eta + r^2) \|M\|_\infty I_r
$$

for every $x \in \{0,1\}^n$, what proves item (v).

Now, we are able to state the main theorem of this chapter, where we mainly bring together the main results from degree reduction (Theorem 4.23), matrix approximation (Theorem 4.39) and the just proven factorization scaling theorem, for what still some work is to do. Roughly speaking, we show in this theorem that a nonnegative function $N \in \mathbb{R}^{(m) \times (0,1)^n}$ of suitable low psd rank cannot admit large negative values with respect to $L_D$.

Compared to the original result of [LRS15], our result is slightly stronger what is an immediate consequence of our slightly stronger Theorem 4.23 compared to the corresponding result in [LRS15]. Another important difference is that we extract some role of the degree of $D$ from the proof of [LRS15], what is not obvious in the formulation of their original result. Additionally and similar to our whole previous work, our proof is much more detailed.

**Theorem 4.43 ([LRS15, Theorem 3.1]).**

Let $m \in \mathbb{N}$, $\varepsilon \in (0,1]$, and let $d \in [2m]$ be even. In particular, it holds $d \geq 2$. Furthermore, let $D \in \mathbb{R}^{(0,1)^m}$ be a degree-$d$ pseudo-density with $\deg(D) \geq 1$, and let $n \in \mathbb{N}$ with $n \geq \max\{m+1,3\}$. Then, every nonnegative function $N : \binom{n}{m} \times \{0,1\}^n \to \mathbb{R}_+$ with $\|N\|_\infty \leq 1$ and

$$
\text{rk}_{\text{psd}}(N) \leq \left( \frac{1}{dm \deg(D) \|D\|_\infty \ln(n)} \right)^{\frac{d+1}{2}} \left( \frac{\varepsilon}{9 \|D\|_\infty} \right)^{\frac{1}{2}} \sqrt{\frac{\text{E} N(S,x)}}{S,x},
$$

satisfies

$$
L_D(N) \geq -\varepsilon.
$$
Proof. First, note that for $N = 0$ we are done because this obviously implies $L_D(N) = 0 > -\varepsilon$. Therefore, without loss of generality we can assume $N \neq 0$.

Let $r := \text{rk}_{\text{psd}}(N)$, and

$$\eta := \frac{\varepsilon}{9 \|D\|_{\infty}}.$$  \hfill (4.64)

By the assumptions on $n, m, d$ and $\varepsilon$, $r$ in particular satisfies

$$r \leq \left( \frac{n}{374 \|D\|_{\infty}} \right)^{\frac{d}{4} + \frac{1}{2}} \frac{1}{\|D\|_{\infty}^2} \sqrt{\mathbb{E} N(S, x)}.$$  \hfill (4.66)

Note that $D$ satisfies $\mathbb{E} D(x) = 1$ as a pseudo-density, which implies

$$\|D\|_{\infty} \geq 1.$$  \hfill (4.65)

Therefore, we have

$$r \leq \left( \frac{n}{374} \right)^{\frac{d}{4} + \frac{1}{2}} \sqrt{\mathbb{E} N(S, x)},$$  \hfill (4.67)

and

$$\eta \leq \frac{\varepsilon}{9} \leq 1.$$  \hfill (4.68)

Let $P : {[n]\choose m} \to \mathbb{S}^r$ and $Q : \{0,1\}^n \to \mathbb{S}^r$ be the two functions obtained by applying Theorem 4.42 to $N$ and $\eta$. Hence, we have

$$\tau := \mathbb{E}_x \text{Tr}(Q(x)) \overset{\text{Th 4.42 (iii)}}{=} \mathbb{E}_x \text{Tr} \left( \mathbb{E}_S P(S) Q(x) \right)$$

$$= \mathbb{E}_{S,x} \text{Tr}(P(S)Q(x)) \overset{\text{Th 4.42 (ii)}}{\geq} \mathbb{E}_{S,x} N(S, x)$$  \hfill (4.69)

and

$$\tau = \mathbb{E}_{S,x} \text{Tr}(P(S)Q(x)) \overset{\text{Th 4.42 (ii)}}{\leq} \mathbb{E}_{S,x} N(S, x) + \eta \leq 1 + \eta \overset{4.67}{\leq} 2,$$  \hfill (4.70)

where we used

$$\mathbb{E}_{S,x} N(S, x) \leq \|N\|_{\infty} \leq 1$$  \hfill (4.70)

in the second last step.

Since we have $N \neq 0$, Theorem 4.42 (ii) implies

$$Q \neq 0.$$  \hfill (4.71)
Furthermore, we define \( I := \{(S,x) \subseteq \binom{[n]}{m} \times \{0,1\}^n | D(x_S) \geq 0\} \). Therefore, we obtain

\[
L_D(N) = \mathbb{E}_{S,x} D(x_S) N(S,x) = \frac{1}{2^n \binom{m}{n}} \left( \sum_{(S,x) \in I} D(x_S) N(S,x) + \sum_{(S,x) \notin I} D(x_S) N(S,x) \right)
\]

\[
\geq \frac{1}{2^n \binom{m}{n}} \left( \sum_{(S,x) \in I} D(x_S) (\text{Tr}(P(S)Q(x)) - \eta) + \sum_{(S,x) \notin I} D(x_S) \text{Tr}(P(S)Q(x)) \right)
\]

\[
= \mathbb{E}_{S,x} D(x_S) \text{Tr}(P(S)Q(x)) - \frac{1}{2^n \binom{m}{n}} \sum_{(S,x) \in I} D(x_S) \eta
\]

\[
\geq \mathbb{E}_{S,x} D(x_S) \text{Tr}(P(S)Q(x)) - \mathbb{E}_{S,x} \|D(x_S)\| \eta
\]

\[
\geq \mathbb{E}_{S,x} D(x_S) \text{Tr}(P(S)Q(x)) - \eta \|D\|_\infty
\]

\[
\geq \mathbb{E}_{S,x} D(x_S) \text{Tr}(P(S)Q(x)) - \frac{\varepsilon}{9}
\]

(4.72)

Now, we define

\[
F : \{0,1\}^n \to \mathcal{S}, \ x \mapsto \mathbb{E}_S D(x_S) P(S).
\]

This function satisfies

\[
\deg(F) \leq \deg(D),
\]

\[
\mathbb{E}_x \text{Tr}(F(x)Q(x)) = \mathbb{E}_x \text{Tr}(\mathbb{E}_S D(x_S) P(S)Q(x)) = \mathbb{E}_{S,x} D(x_S) \text{Tr}(P(S)Q(x)),
\]

and

\[
\max_{x \in \{0,1\}^n} \|F(x)\| = \max_{x \in \{0,1\}^n} \|\mathbb{E}_S D(x_S) P(S)\| \leq \|D\|_\infty \|\mathbb{E}_S P(S)\| \|D\|_\infty.
\]

(4.75)

Putting (4.74) into (4.72) yields

\[
L_D(N) \geq \mathbb{E}_x \text{Tr}(F(x)Q(x)) - \frac{\varepsilon}{9}.
\]

(4.76)

Note that if \( F = 0 \), we are already done because this implies \( L_D(N) \geq -\frac{\varepsilon}{9} > -\varepsilon \).

Hence, we can assume \( F \neq 0 \) without loss of generality. Therefore and because of (4.71), we can now apply the main theorem from matrix approximation with

\[
\varepsilon' := \frac{\varepsilon}{9(1 + \max_{x \in \{0,1\}^n} \|F(x)\|)} \leq \varepsilon \leq 1
\]

in order to obtain some \( s \in \mathbb{N} \) and a function \( B : \{0,1\}^n \to \mathbb{R}^{r \times s} \) with

\[
\gamma := \mathbb{E}_x \text{Tr} \left( B(x)B(x)^T \right) \neq 0,
\]

\[
\frac{1}{\gamma} \mathbb{E}_x \text{Tr}(F(x)Q(x)) \geq \frac{1}{\gamma} \mathbb{E}_x \text{Tr} \left( F(x)B(x)B(x)^T \right) - \varepsilon' \max_{x \in \{0,1\}^n} \|F(x)\|
\]

(4.77)
and
\[
\deg(B)^{\text{Th. 4.30}} \leq \frac{1}{2} \deg(F) \left( \frac{1.2 e^2}{\varepsilon} \frac{1}{r_x} \mathbb{E} \text{Tr} \left( \frac{x Q(x)}{\tau} \ln \left( \frac{x Q(x)}{\tau} \right) \right) + \ln \left( \frac{12}{\varepsilon} \right) + 1 \right)
\]
\[
\leq \frac{1}{2} \deg(F) \left( \frac{1.2 e^2}{\varepsilon} \frac{1}{r_x} \mathbb{E} \text{Tr} \left( \frac{x Q(x)}{\tau} \ln \left( \frac{x Q(x)}{\tau} \right) \right) + \frac{12}{\varepsilon} + 1 \right)
\]
\[
= \frac{\deg(F)}{\varepsilon'} \left( \frac{0.6 e^2}{r_x} \mathbb{E} \text{Tr} \left( \frac{x Q(x)}{\tau} \ln \left( \frac{x Q(x)}{\tau} \right) \right) + 6 + 0.5\varepsilon' \right)
\]
\[
\leq 9 \deg(D)(1 + \|D\|_{\infty}) \left( \frac{0.6 e^2}{r_x} \mathbb{E} \text{Tr} \left( \frac{x Q(x)}{\tau} \ln \left( \frac{x Q(x)}{\tau} \right) \right) + 6.5 \right)
\]
\[
\leq 18 \deg(D) \|D\|_{\infty} \left( \frac{0.6 e^2}{r_x} \mathbb{E} \text{Tr} \left( \frac{x Q(x)}{\tau} \ln \left( \frac{x Q(x)}{\tau} \right) \right) + 6.5 \right). \tag{4.79}
\]

Our chosen \(\varepsilon'\) satisfies \(\varepsilon' \leq \frac{\varepsilon}{g_{\max x \in [-1,1]^n}\|F(x)\|}\), why (4.78) implies
\[
\frac{1}{\tau} \mathbb{E} \text{Tr} \left( F(x)Q(x) \right) \geq \frac{1}{\gamma} \mathbb{E} \text{Tr} \left( F(x)B(x)B(x)^T \right) - \frac{\varepsilon}{9}. \tag{4.80}
\]

Note that we have
\[
\gamma = \mathbb{E} \text{Tr} \left( B(x)B(x)^T \right) = \mathbb{E} \|B(x)\|_F^2. \tag{4.81}
\]

Then, applying (4.80) to (4.76) implies
\[
L_D(N) \geq \frac{\tau}{\gamma} \mathbb{E} \text{Tr} \left( F(x)B(x)B(x)^T \right) - \frac{\tau \varepsilon}{9} - \frac{\varepsilon}{9}
\]
\[
\geq \frac{\tau}{\gamma} \mathbb{E} \text{Tr} \left( F(x)B(x)B(x)^T \right) - \frac{\varepsilon}{3}
\]
\[
= \frac{\tau}{\gamma} \mathbb{E} D(x_S) \text{Tr} \left( P(S)B(x)B(x)^T \right) - \frac{\varepsilon}{3}
\]
\[
= \frac{\tau}{\gamma} \mathbb{E} D(x_S) \text{Tr} \left( \sqrt{P(S)} B(x)B(x)^T \right) - \frac{\varepsilon}{3}
\]
\[
= \mathbb{E} D(x_S) \left\| \sqrt{P(S)} B(x) \right\|_F^2 - \frac{\varepsilon}{3}. \tag{4.82}
\]

We set \(k := \deg(B)\) and
\[
\lambda := \max_{S \in \binom{[n]}{d}} \left\| \sqrt{P(S)} \right\| = \max_{S \in \binom{[n]}{d}} \|P(S)\| \overset{\text{Th. 4.12 iv}}{\leq} \frac{2 \tau^2}{\eta} 18 \varepsilon \|D\|_{\infty}, \tag{4.83}
\]

and we apply Theorem 4.24, the main result from degree reduction. This yields
\[
L_D(N) \geq -2\frac{\tau}{\gamma} \sqrt{\lambda} \|D\|_{\infty} \left( \frac{km}{n-d} \right)^{\frac{d+1}{2}} \left( \mathbb{E} \|\sqrt{P(S)} B(x)\|_F^2 \mathbb{E} \|B(x)\|_F^2 \right)^{\frac{1}{2}} - \frac{\varepsilon}{3}
\]
\[
\geq -12\sqrt{2} \frac{\tau}{\gamma} \sqrt{\varepsilon} \|D\|_{\infty} \left( \frac{km}{n-d} \right)^{\frac{d+1}{2}} \left( \mathbb{E} \|\sqrt{P(S)} B(x)\|_F^2 \mathbb{E} \|B(x)\|_F^2 \right)^{\frac{1}{2}} - \frac{\varepsilon}{3} \tag{4.83}
\]
Chapter 4. A lower bound on the psd rank of the family of correlation polytopes

It holds
\[
\left( \mathbb{E}_{S,x} \left\| \sqrt{P(S)} B(x) \right\|^2_F \mathbb{E}_{x} \left\| B(x) \right\|^2_F \right)^{\frac{1}{2}} \geq \operatorname{Cor}(4.1) \left( \mathbb{E}_{S,x} \operatorname{Tr} \left( P(S) B(x) B(x)^T \right) \mathbb{E}_{x} \left\| B(x) \right\|^2_F \right)^{\frac{1}{2}}
\]
\[
\geq \operatorname{Th}(4.42) \left( \mathbb{E}_{x} \operatorname{Tr} \left( B(x) B(x)^T \right) \right)^{\frac{1}{2}} = \gamma.
\]

Putting this into (4.83) yields
\[
L_D(N) \geq -12 \sqrt{2} \frac{r}{\sqrt{\varepsilon}} \|D\|_\infty^2 \left( \frac{km}{n - \frac{d}{2}} \right)^{\frac{d}{4} + \frac{1}{2}} - \frac{\varepsilon}{3}.
\]

Now, we have to specify \( k \). In order to do this, note that
\[
1 r x \mathbb{E} \operatorname{Tr} \left( \frac{1}{\tau} Q(x) \ln \left( \frac{1}{\tau} Q(x) \right) \right) \geq \frac{1}{r x} \mathbb{E} \operatorname{Tr} \left( \frac{1}{\tau} Q(x) \ln \left( \frac{r(n+\tau^2)}{\tau} I_r \right) \right)
\]
\[
= \frac{1}{r x} \mathbb{E} \operatorname{Tr} \left( \frac{1}{\tau} \ln \left( \frac{r(n+\tau^2)}{\tau} \right) Q(x) \right)
\]
\[
= \ln \left( \frac{r(n+\tau^2)}{\tau} \right) - \ln \left( \mathbb{E}_{S,x} N(S,x) \right)
\]
\[
\geq \ln \left( \frac{(1.26r)^3}{\mathbb{E}_{S,x} N(S,x)} \right) = 3 \ln \left( \frac{1.26r}{\sqrt{\mathbb{E}_{S,x} N(S,x)}} \right),
\]
holds. Therefore, we have
\[
k = \deg(B) \leq 18 \frac{\deg(D) \|D\|_\infty}{\varepsilon} \left( 0.6e^2 \frac{1}{r x} \mathbb{E} \operatorname{Tr} \left( \frac{1}{\tau} Q(x) \ln \left( \frac{1}{\tau} Q(x) \right) \right) + 6.5 \right)
\]
\[
\leq 18 \frac{\deg(D) \|D\|_\infty}{\varepsilon} \left( 1.8e^2 \ln \left( \frac{1.26r}{\sqrt{\mathbb{E}_{S,x} N(S,x)}} \right) + 6.5 \right).
\]

Since we have \( 1.8e^2 x + 6.5 \leq 41.5x \) for every \( x \in \mathbb{R} \) with \( x \geq \ln(1.26) \), this implies
\[
k \leq 747 \frac{\deg(D) \|D\|_\infty}{\varepsilon} \ln \left( \frac{1.26r}{\sqrt{\mathbb{E}_{S,x} N(S,x)}} \right) \leq 747 \frac{\deg(D) \|D\|_\infty}{\varepsilon} \ln \left( 1.26 \left( \frac{1}{374} n \right)^{\frac{d}{4} + \frac{1}{2}} \right)
\]
\[
\leq 747 \frac{\deg(D) \|D\|_\infty}{\varepsilon} \ln \left( \left( 1.26 \left( \frac{1}{374} n \right)^{\frac{d}{4} + \frac{1}{2}} \right)^{\frac{d}{4} + \frac{1}{2}} \right) \ln (n)
\]
\[
\leq 747 \frac{\deg(D) \|D\|_\infty}{\varepsilon} 0.5d \ln (n) \leq 374 \frac{d \deg(D) \|D\|_\infty}{\varepsilon} \ln (n),
\]
where we used \( d \geq 2 \) respectively \( \frac{d}{4} + \frac{1}{2} \geq 1 \) two times.
By applying this bound to (4.84), we obtain
\[ L_D(N) \geq -12\sqrt{2} \frac{r}{\sqrt{\varepsilon}} \|D\|_{\infty}^{3/2} \left( \frac{374dm \deg(D) \|D\|_{\infty} \ln(n)}{\varepsilon \left( n - \frac{d}{2} \right)} \right)^{4+1/2} - \frac{\varepsilon}{3}. \]

Now, the upper bound on \( r \) from the assumption implies
\[ L_D(N) \geq -12\sqrt{2} \frac{\varepsilon}{9} \sqrt{\mathbb{E} N(S,x)} - \frac{\varepsilon}{3} \geq -\frac{4\sqrt{2}}{9} \varepsilon - \frac{\varepsilon}{3} \geq -\frac{6}{9} \varepsilon - \frac{1}{3} \varepsilon = -\varepsilon. \]

Combining the previous theorem with Proposition 4.13 immediately yields the following consequence. Roughly speaking, it shows that for a nonnegative function \( f \in \mathbb{R}^{\{0,1\}^m} \) that admits no sos certificate of low-degree, the function \( M^f_n \in \mathbb{R}^{(\{0,1\}^n)\times\{0,1\}^n} \) admits no low-rank psd factorization. Although this result is actually just a corollary of the previous one and Proposition 4.13, we state it as a theorem because it will be the one we apply in the next sections in order to obtain a lower bound on the psd rank of the family of correlation polytopes. Therefore, it is of high significance for us.

**Theorem 4.44 (cf. [LRS15, Theorem 3.8]).**

Let \( n, m \in \mathbb{N} \) with \( n \geq \max\{m + 1, 3\} \), \( \varepsilon \in (0,1] \) and \( d \in [2m] \) with \( d \) even. Furthermore, let \( f : \{0,1\}^m \to [0,1] \) be a nonnegative function and \( D \in \mathbb{R}^{\{0,1\}^m} \) a degree-\( d \) pseudo-density with \( \deg(D) \geq 1 \) and \( \mathbb{E} D(x) f(x) < -\varepsilon \). Then, we have

\[
\text{rk}_{\text{psd}}(M^f_n) \geq \left( \frac{1}{374} \varepsilon \left( n - \frac{d}{2} \right) \right)^{4+1/2} \left( \frac{\varepsilon}{9 \|D\|_{\infty}} \right)^{3} \sqrt{\mathbb{E} f(x)}. 
\]

**Proof.** By definition and because \( f \) is nonnegative, \( M^f_n \) is also a nonnegative function in \( \mathbb{R}^{(\{0,1\}^n)\times\{0,1\}^n} \). Additionally, we have

\[
\|M^f_n\|_{\infty} = \max_{S \in \binom{[m]}{m}} \max_{x \in \{0,1\}^n} M^f_n(S,x) = \max_{S \in \binom{[m]}{m}} f(x_S) \leq 1. 
\]

Therefore, \( M^f_n \) satisfies the conditions of Theorem 4.43.

In addition, note that by Proposition 4.13, we have

\[ L_D(M^f_n) < -\varepsilon. \]
Therefore, applying the contrapositive of Theorem 4.43 implies
\[
\text{rk}_\text{psd}(M'_f) \geq \left( \frac{\frac{1}{374}\varepsilon (n - \frac{d}{2})^{\frac{d+1}{2}}}{dn^2 \|D\|_\infty \ln n} \right)^\frac{1}{2} \sqrt{\mathbb{E} M'_f (S,x)}.
\]

The desired statement follows because
\[
\mathbb{E} M'_f (S,x) = \mathbb{E}_{S,x} f(x_S) = \mathbb{E}_{S} \frac{1}{2^n} \sum_{x \in \{0,1\}^n} f(x_S) = \mathbb{E}_{S} \frac{1}{2^n} \sum_{y \in \{0,1\}^m} \sum_{x_{S=y}} f(y)
\]
\[
= \mathbb{E}_{S} \frac{1}{2^n} \sum_{y \in \{0,1\}^m} 2^{n-m} f(y) = \mathbb{E}_{S} \mathbb{E}_{y} f(y) = \mathbb{E}_{y} f(y).
\]

\[\Box\]

### 4.6 Quantities of the Grigoriev pseudo-density

In order to apply the last theorem from the previous section, we first need a nonnegative function \( f \in \mathbb{R}^{\{0,1\}^m} \), some \( d \in [2m] \) and a degree-\( d \) pseudo density \( D \in \mathbb{R}^{\{0,1\}^m} \) such that
\[
\mathbb{E}_{x} D(x)f(x) < 0.
\]
By Corollary 3.45, this is equivalent to \( \text{deg}_{\text{sos}}(f) > d \). Because of Corollary 4.11, such a function \( f \) should additionally be quadratic. But this is exactly the statement of Theorem 3.36, a consequence of Grigoriev’s Theorem. Applying Corollary 3.45 to this Theorem immediately yields the following.

**Corollary 4.45.**

Let \( m \in \mathbb{N} \) with \( m \geq 2 \) and \( r \in \left( \left[ \frac{m-3}{2}, \frac{m+3}{2} \right] \right) \) with \( r \notin \mathbb{N} \), and let \( b \in \mathbb{R} \) be defined as \( b := \min_{k \in \mathbb{N}} (k - r)^2 \). Let \( c \in (0,b] \), and let \( f \in \mathbb{R}^{\{0,1\}^m} \) be the quadratic and nonnegative function
\[
f = \left( \sum_{i \in [m]} \chi(i) - r \right)^2 - c.
\]
Then, there exists some degree-\( m \) pseudo-density \( D \in \mathbb{R}^{\{0,1\}^m} \) with
\[
\mathbb{E}_{x} D(x)f(x) < 0. \tag{4.86}
\]

Of course, there are many degree-\( m \) pseudo-densities satisfying (4.86), but the proof of Theorem 3.36 yields a concrete one. The theorem is proven by applying the Grigoriev pseudo-expectation, and combining the proof of this theorem with Definition 3.57 respectively the proof of Grigoriev’s Theorem 2.27 shows that the degree-\( m \) Grigoriev pseudo-density \( D_{\varphi,r} \in \mathbb{R}^{\{0,1\}^m} \) is suitable for proving (4.86). In particular, we obtain
\[
\mathbb{E}_{x} D_{\varphi,r}(x)f(x) = -c < 0. \tag{4.87}
\]
For an arbitrary degree-$m$ pseudo-density $D \in \mathbb{R}^{\{0,1\}^m}$ satisfying (4.86), we set

$$\varepsilon := -\mathbb{E}_x D(x)f(x) = -\left(\mathbb{E}_x D(x) \left(\sum_{i \in [n]} \chi_{\{i\}} - r\right)^2\right) + c.$$ 

Considering Theorem 4.44, we see the larger $\varepsilon$ the better our lower bound on the psd rank. In order to maximize $\varepsilon$, it is enough to consider $c = b$ in the above Corollary. On the other hand, we have

$$b = \min_{k \in \mathbb{N}} (k - r)^2 \leq \frac{1}{4}.$$ 

Therefore, $b$ is maximized on his part for $r \in \left\{\frac{m-2}{2}, \frac{m}{2}, \frac{m+2}{2}\right\}$ and $m$ odd, respectively for $r \in \left\{\frac{m-1}{2}, \frac{m+1}{2}\right\}$ and $m$ even.

Here, the case $r = \frac{m}{2}$ plays a particular role. Since it is centered between 0 and $m$, it gives the function $f$ some kind of symmetry, and because we benefited of symmetry during this work so far, we choose this $r$, and therefore we just consider odd $m$.

Note that we have $\mathbb{E}_x D(x) \left(\sum_{i \in [n]} \chi_{\{i\}} - \frac{m}{2}\right)^2 \geq 0$ because $D$ is a degree-$m$ pseudo-density, and therefore we obtain

$$\varepsilon \leq c = b = \frac{1}{4}.$$ 

In order to maximize $\varepsilon$ again and because of (4.87), we see that taking for example the Grigoriev pseudo-density yields the maximum possible value, why we think it is a good choice for a pseudo-density using for applying Theorem 4.44 in the next section. Besides, it is also the only pseudo-density on that we have complete information (cf. Corollary 3.59). Note that we just write $D_G$ instead of $D_{G,\frac{m}{2}}$.

Looking on Theorem 4.44, we see that additionally to $\varepsilon$ also the infinity norm and the degree of the used pseudo-density affect the lower bound on the psd rank. Therefore, we will compute $\|D_G\|_\infty$ and $\deg(D_G)$ for $D_G \in \mathbb{R}^{\{0,1\}^m}$ and $m$ odd in the further course of this section.

Note that [LRS15] also computed $\|D_G\|_\infty$, but on the one hand, they only proved an upper bound. Indeed, this is enough for our purpose, nevertheless we will also present a lower bound to give a more exact impression on the behavior of $\|D_G\|_\infty$. On the other hand, we will fix a computational error that improves the upper bound of $\|D_G\|_\infty$ as well as the lower bound on the psd rank of the family of the correlation polytopes.

[LRS15] did not compute the degree of $D_G \in \mathbb{R}^{\{0,1\}^m}$, they just used the trivial upper bound $\deg(D_G) \leq m$. 
4.6.1 The infinity norm

We start with the computation of the infinity norm of the Grigoriev pseudo-density. In order to upper bound it, we give another representation than the one obtained by Corollary \[3.59\]. We follow the computation that [LRS15] did by the way in their proof of [LRS15] Theorem 5.3. Our computation is more detailed, but more importantly, it fixes a computational error what will later yield to a better upper bound.

Let \( m \in \mathbb{N} \) be odd, \( D_\varphi \in \mathbb{R}^{(0,1)^m} \) and \( x \in \{0,1\}^m \), then we have

\[
D_\varphi(x) \overset{\text{Cor.} \[3.59\]}{=} \frac{2^m}{\binom{m}{|x|}} \cdot \frac{\prod_{j=0,j\neq |x|}^m \left( \frac{m}{2} - j \right)}{\prod_{j=0,j\neq |x|}^m (|x| - j)} = \frac{1}{\binom{m}{|x|}} \cdot \frac{\prod_{j=0,j\neq |x|}^m (m - 2j)}{\prod_{j=0,j\neq |x|}^m (|x| - j)}
\]

\[
= \frac{1}{\binom{m}{|x|}} \cdot \prod_{j=0,j\neq |x|}^m (m - 2j) \cdot \prod_{j=0,j\neq |x|}^m (|x| - j)
\]

\[
= \frac{1}{\binom{m}{|x|}} \cdot \prod_{j=0,j\neq |x|}^m (m - 2j) \cdot \prod_{j=0,j\neq |x|}^m (|x| - j)
\]

\[
= \frac{1}{(m - 2|x|)} \cdot \frac{\prod_{j=0,j\neq |x|}^m (m - 2j)}{|x|!(m - |x|)!(-1)^{m-|x|}} \cdot \frac{1}{(m - 2|x|)}
\]

\[
= \frac{(-1)^{m+1}+|x|}{(m - 2|x|)} \cdot \frac{\prod_{j=0,j\neq |x|}^m (m - 2j)}{|x|!(m - |x|)!(-1)^{m-|x|}} \cdot \frac{1}{(m - 2|x|)}
\]

\[
= \frac{(-1)^{m+1}}{(m - 2|x|)} \cdot \frac{\prod_{j=0,j\neq |x|}^m (m - 2j)}{|x|!(m - |x|)!(-1)^{m-|x|}} \cdot \frac{1}{(m - 2|x|)}
\]

By applying the factorial formula of the binomial coefficient, we obtain

\[
D_\varphi(x) = \frac{(-1)^{m+1}}{(m - 2|x|)} \cdot \frac{m}{2^{m-1}} \cdot \left( \frac{m - 1}{m - 2} \right) \cdot \left( \frac{m - 1}{2} \right).
\]  \( (4.88) \)

In order to get an upper bound for \( \|D_\varphi\|_\infty \) in terms of \( m \), we have to estimate \( \left( \frac{m - 1}{m - 2} \right) \), for which we need the more precise version of Stirling’s formula, which we already stated in Proposition \[4.3\] for this purpose.

With it, we get the following corollary.
Corollary 4.46.
For every positive, even integer \( n \) holds
\[
\binom{n}{n/2} \leq \sqrt{\frac{2}{\pi}} \cdot \frac{2^n}{\sqrt{n}} e^{\frac{-18n+1}{12n+2}}, \tag{4.89}
\]
\[
\binom{n}{n/2} \geq \sqrt{\frac{2}{\pi}} \cdot \frac{2^n}{\sqrt{n}} e^{\frac{-9n-1}{36n^2+3n}}. \tag{4.90}
\]

In particular, this implies
\[
\lim_{n \to \infty} \left( \frac{n}{2} \right)^n \frac{1}{n^{n+1}} = \sqrt{\frac{2}{\pi}}.
\] \tag{4.91}

\textbf{Proof.} If we apply the inequalities of the precise version of Stirling's formula \[4.37\], we obtain
\[
\binom{n}{n/2} \leq \sqrt{\frac{2}{\pi}} \sqrt{n} \frac{n^n e^{\frac{1}{12n+1}}}{e^n (\sqrt{\frac{n}{2}})^{\frac{n}{2}} e^{\frac{1}{6n+1}}} \leq e^n \left( \sqrt{\frac{2}{\pi}} \sqrt{n} \right)^n \frac{2^n}{\sqrt{n} e^{\frac{1}{12n+1}}} e^{\frac{-18n+1}{12n+2}},
\]
and
\[
\binom{n}{n/2} \geq \sqrt{\frac{2}{\pi}} \sqrt{n} \frac{n^n e^{\frac{1}{12n+1}}}{e^n (\sqrt{\frac{n}{2}})^{\frac{n}{2}} e^{\frac{1}{6n+1}}} \geq e^n \left( \sqrt{\frac{2}{\pi}} \sqrt{n} \right)^n \frac{2^n}{\sqrt{n} e^{\frac{1}{12n+1}}} e^{\frac{-9n-1}{36n^2+3n}}.
\]

Since for odd \( m \geq 3 \), \( m - 1 \) is even and positive, we can use the above corollary, and we are now able to prove an upper bound for \( \|D_\varphi\|_\infty \).

In this context, note that in their original work, [LRST13] stated
\[
\|D_\varphi\|_\infty \leq m^\frac{3}{2},
\]
also by using an alternative representation of \( D_\varphi \) as we derived in \[4.88\] and by applying Stirling's formula. This bound is not wrong, but as already mentioned they had a small error in their computation of their alternative representation that...
yields to a weaker bound. Actually, without the computational error their method even shows
\[ \|D_\varphi\|_\infty \leq m^{\frac{1}{2}}. \] (4.92)

As we will see in the next section, this immediately improves the lower bound on the psd rank of the family of the correlation polytopes from \( \text{rk}_{\text{psd}}(\text{CORR}_n) \geq 2^\alpha \left( \frac{n}{\ln n} \right)^{2/13} \) to \( \text{rk}_{\text{psd}}(\text{CORR}_n) \geq 2^\alpha \left( \frac{n}{\ln n} \right)^{2/11} \).

Note that Troy Lee et al. \cite{LPWY16} independently from us also recognized that there has to be a computational error in \cite{LRS15} in the computation of \( \|D_\varphi\|_\infty \) and that (4.92) is valid, but they did not specify the error nor fixed it or gave a new proof of (4.92).

Now, in our representation (4.88) the error is fixed, and as already mentioned, we will also use Stirling’s formula in order to compute an upper bound on \( \|D_\varphi\|_\infty \), but in contrast to \cite{LRS15}, we will use a more precise version of it. This slightly improves (4.92) once more, but this time, it has no more effect on further results.

**Proposition 4.47.**
Let \( m \in \mathbb{N} \) be odd and \( m \geq 3 \). Then, the Grigoriev pseudo-density \( D_\varphi \in \mathbb{R}^{\{0,1\}^m} \) satisfies
\[ \|D_\varphi\|_\infty < 0.87 m^{\frac{1}{2}}. \]

**Proof.** By (4.88), for \( x \in \{0,1\}^m \) with \( |x| = k \) we have
\[ D_\varphi(x) = \frac{(-1)^{m-1}}{(m-2k)} \frac{m}{2^{m-1}} \left( \frac{m-1}{2} \right), \]
and therefore
\[ \|D_\varphi\|_\infty = \max_{x \in \{0,1\}^m} |D_\varphi(x)| \]
\[ = \max_{k \in [m]} \frac{1}{|m-2k|} \frac{m}{2^{m-1}} \left( \frac{m-1}{2} \right) \]
\[ = \frac{m}{2^{m-1}} \left( \frac{m-1}{2} \right). \] (4.93)

Now, let \( m \geq 5 \). Note, in this case the following is valid.
\[ \frac{\pi}{2} (m-1)e^{-\frac{-18(m-1)+1}{36(m-1)^2+6(m-1)}} \geq \frac{\pi}{2} (m-1)e^{\frac{71}{1000}} > \frac{1}{0.87^2 m}. \] (4.94)
This results in
\[
\|D_g\|_\infty \leq \frac{m}{2^{m-1}} \left( \frac{m-1}{2} \right) \leq \frac{2^{m-1}}{\sqrt{\pi}} \sqrt{m-1} e^{-\frac{18(m-1)+1}{36(m-1)^2+6(m-1)}}
\]
\[
\leq \frac{m}{\sqrt{0.87m}} = 0.87 \sqrt{m} = 0.87m^{1/2}.
\]

Now, we check the remaining case \(m = 3\) by hand. In this case, we obtain
\[
\|D_g\|_\infty = \frac{3}{2^{3/2}(1)} = \frac{3}{2} = 1.5 < 0.87\sqrt{3}.
\]

In contrast to [LRS15], who just stated an upper bound, in the following proposition we will also prove a lower bound which shows that our bound is optimal in some sense and cannot be improved once more. Actually, we show that our bound is tight in the size of \(m\) or more precise in the exponent of \(m\). That means, for every pair \(\alpha < \frac{1}{2}\) and \(C > 0\) there exist a number \(N \in \mathbb{N}\) such that for every \(m \in \mathbb{N}\) with \(m \geq N\) and \(D_g \in \{0,1\}^m\) holds
\[
\|D_g\|_\infty > Cm^\alpha.
\]

**Proposition 4.48.**

Let \(m \in \mathbb{N}\) be odd. Then, for \(D_g \in \mathbb{R}^{\{0,1\}^m}\) we have
\[
\|D_g\|_\infty > 0.773m^{1/2}.
\]

**Proof.** For \(m = 1\), we have \(D_g(0) = D_g(1) = 1\), and therefore we obviously obtain \(\|D_g\|_\infty = 1 > 0.773\).

Now, we consider the case \(m \geq 9\). In this case, we have
\[
e^{-\frac{9(m-1)+1}{36(m-1)^2+6(m-1)}} \geq e^{-\frac{9(9-1)+1}{36(9-1)^2+6(9-1)}} = e^{-\frac{73}{273}}.
\]

In particular, \(m \geq 9\) implies \(m-1 \geq 0\), and therefore we can apply the lower bound
for \((\frac{m-1}{2})\) obtained in (4.90). Overall, we have

\[
\|D_g\|_\infty \mathbf{4.93} \geq \frac{m}{2^{m-1}} \left(\frac{m-1}{2}\right) \mathbf{4.90} \\
\geq \frac{m}{2^{m-1}} \sqrt{\frac{2}{\pi}} \frac{2^{m-1}}{\sqrt{m-1}} e^{-\frac{9(m-1)^2}{36(m-1)^2+3(m-1)}} \\
\geq \sqrt{\frac{2}{\pi}} \frac{m}{\sqrt{m-1}} e^{-\frac{73}{2328}} \geq \sqrt{\frac{2}{\pi}} e^{-\frac{73}{2328}} \frac{m}{\sqrt{m}} > 0.773m^2.
\]

Finally, we check the remaining cases \(m \in \{3, 5, 7\}\) by hand.

Let \(m = 3\). Then, we obtain

\[
\|D_g\|_\infty \mathbf{4.93} \geq \frac{3}{2^2} \left(\frac{2}{1}\right) = \frac{3}{2} = 1.5 > 0.773\sqrt{3}.
\]

Let \(m = 5\). Then, we obtain

\[
\|D_g\|_\infty \mathbf{4.93} \geq \frac{5}{2^4} \left(\frac{4}{2}\right) = \frac{15}{8} = 1.875 > 0.773\sqrt{5}.
\]

Let \(m = 7\). Then, we obtain

\[
\|D_g\|_\infty \mathbf{4.93} \geq \frac{7}{2^6} \left(\frac{6}{2}\right) = \frac{35}{16} = 2.1875 > 0.773\sqrt{7}.
\]

It has no use for the lower bound of \(\text{rk}_{\text{psd}}(\text{CORR}_m)\), but with the limit obtained from Corollary 4.46 we can still analyze \(\|D_g\|_\infty\) more exactly.

**Proposition 4.49.**

For \(m \in \mathbb{N}\), let \(D_g \in \mathbb{R}^{\{0,1\}^m}\) be the the Grigoriev pseudo-density. Then, it holds

\[
\lim_{\text{odd }} m \to \infty \frac{\|D_g\|_\infty}{\sqrt{m}} = \sqrt{\frac{2}{\pi}} \approx 0.798.
\]

**Proof.** By (4.93) we have \(\|D_g\|_\infty = \frac{m}{2^{m-1}} \left(\frac{m-1}{2}\right)\). Therefore, we obtain

\[
\lim_{m \to \infty} \frac{\|D_g\|_\infty}{\sqrt{m}} = \lim_{m \to \infty} \frac{\sqrt{m}}{2^{m-1}} \frac{m}{\sqrt{m-1}} \left(\frac{m-1}{2}\right) = \lim_{m \to \infty} \frac{\sqrt{m}}{2^{m-1}} \frac{m}{\sqrt{m-1}} \left(\frac{m-1}{2}\right) = \lim_{n \to \infty} \frac{\sqrt{n}}{2^{n}} \left(\frac{n}{2}\right) \mathbf{4.91} \sqrt{\frac{2}{\pi}}.
\]

\(\square\)
In fact, this convergence is fast. Computing the ratio for small \( m \) yields a ratio smaller than 0.83 already for \( m \geq 7 \) and a ratio even smaller than 0.81 for \( m \geq 17 \).

**Remark 4.50.**
Due to the previous proposition, it should be possible to prove a stronger lower bound than in Proposition 4.48. Indeed, the number 0.773 is a bit arbitrary. We could improve this bound to an arbitrary number \( C < \sum \frac{2}{\pi} \approx 0.798 \). Analog to (4.95), we find some odd \( M \in \mathbb{N} \) such that

\[
\frac{e^{-9(m-1)-1}}{36(m-1)^2+3(m-1)} > C \sqrt{\frac{\pi}{2}}
\]

for every odd \( m \geq M \), what would imply \( \| D_g \|_\infty > C m^{\frac{1}{2}} \) for \( D_g \in \mathbb{R}\{0,1\}^m \) and every odd \( m \geq M \) following the proof of Proposition 4.48. But then, we would have to check the remaining cases \( m \in \{1,3,\ldots,M-2\} \) by hand, what eventually could be many cases.

Note that for our purpose, the bound in Proposition 4.48 is sufficient, and a better bound \( 0.773 < C < \sum \frac{2}{\pi} \approx 0.798 \) would not have any effect to other results.

### 4.6.2 The degree

Now, we have a look on the degree of \( D_g \). Let \( m \in \mathbb{N} \) be odd, and let \( D_g \in \mathbb{R}\{0,1\}^m \). We first remind some notation and some results from Section 3.2. Let \( (\chi_I)_{I \subseteq [m]} \) and \( (\delta_a)_{a \in \{0,1\}^m} \) be the family of the monomial respectively the needle functions in \( \mathbb{R}\{0,1\}^m \). Furthermore, for \( a \in \{-1,1\}^m \) we define the set \( I_a := \{ i \in [n] \mid a_i = 1 \} \).

Because of (3.3), there exist a unique family of real numbers \( (\lambda_I)_{I \subseteq [m]} \) such that

\[
D_g = \sum_{I \subseteq [m]} \lambda_I \chi_I. \tag{4.96}
\]

So, by computing the \( \lambda_I \)'s, we can compute the degree of \( D_g \). We additionally remind the obvious identity

\[
D_g = \sum_{a \in \{0,1\}^m} D_g(a) \delta_a. \tag{4.97}
\]

In Proposition 3.22, we computed the representation of the needle functions as monomial functions. Inserting this representation in (4.97) yields

\[
D_g \overset{(4.97)}{=} \sum_{a \in \{0,1\}^m} D_g(a) \delta_a \overset{\text{Prop 3.22}}{=} \sum_{a \in \{0,1\}^m} D_g(a) \sum_{I \subseteq [m]} (-1)^{|I| - |a|} \chi_I
\]

\[
= \sum_{I \subseteq [m]} \left( \sum_{a \in \{0,1\}^m \atop I_a \subseteq I} (-1)^{|I| - |a|} D_g(a) \right) \chi_I.
\]
Therefore, the unique family \((\lambda_I)_{I \subseteq [m]}\) from (4.96) satisfies

\[
\lambda_I = \sum_{a \in \{0,1\}^m \atop I_a \subseteq I} (-1)^{|I|-|a|} D_{\varphi}(a)
\]  

(4.98)

**Proposition 4.51.**
Let \(m \in \mathbb{N}\) be odd, and let \(D_{\varphi} \in \{0,1\}^m\). Then, it holds

\[
\deg(D_{\varphi}) = m - 1.
\]

**Proof.** For \(I \subseteq [m]\), we define

\[
\lambda_I := \sum_{a \in \{0,1\}^m \atop I_a \subseteq I} (-1)^{|I|-|a|} D_{\varphi}(a) \quad \text{and}
\]

\[
c_I := (-1)^{m-1+|I|} \frac{m}{2^{m-1}} \left(\frac{m-1}{m-1}\right).
\]

Note that \(c_I \neq 0\) holds for every \(I \subseteq [m]\).

By (4.96) and (4.98), it remains to show

\[
\lambda_{[m]} = 0, \quad \text{and}
\]

\[
\lambda_I \neq 0 \text{ for some } I \subseteq [m] \text{ with } |I| = m - 1.
\]
It holds

\[ \lambda_I = \sum_{a \in \{0,1\}^m} (-1)^{|I|-|a|} D_g(a) \]

\[ \sum_{a \in \{0,1\}^m} (-1)^{|I|-|a|} (-1)^{m-1-|a|} \frac{m}{m-2|a|} \frac{1}{2^{m-1} \left(\frac{m-1}{2}\right)} \]

\[ = c_I \sum_{a \in \{0,1\}^m} \frac{1}{m-2|a|} = c_I \sum_{k=0}^{m} \sum_{a \in \{0,1\}^m \mid |a|=k} \frac{1}{m-2k} \]

\[ = c_I \left( \sum_{k=0}^{m-1} \sum_{a \in \{0,1\}^m \mid |a|=k} \frac{1}{m-2k} + \sum_{k=0}^{m-1} \sum_{a \in \{0,1\}^m \mid |a|=m-k} \frac{1}{m-2k} \right) \]

\[ = c_I \sum_{k=0}^{m-1} \frac{1}{m-2k} \left( \sum_{a \in \{0,1\}^m \mid |a|=k} 1 - \sum_{a \in \{0,1\}^m \mid |a|=m-k} 1 \right) \]

(4.99)

Therefore, we have

\[ \lambda_{[m]} = c_{[m]} \sum_{k=0}^{m-1} \frac{1}{m-2k} \left( \sum_{a \in \{0,1\}^m \mid |a|=k} 1 - \sum_{a \in \{0,1\}^m \mid |a|=m-k} 1 \right) \]

\[ = c_{[m]} \sum_{k=0}^{m-1} \frac{1}{m-2k} \binom{m}{k} - \binom{m}{m-k} = 0. \]

Now, consider \( I = [m-1] \). We will show \( \lambda_{[m-1]} \neq 0 \). But firstly, note the following
useful identity
\[
\binom{m-1}{k} - \binom{m-1}{m-k} = \frac{(m-1)!}{(m-1-k)!k!} - \frac{(m-1)!}{(m-k)!(m-1-k)!(k-1)!} = \frac{(m-k)!(m-1)!(m-1-k)!}{(m-k)!k!} = \frac{(m-1)!(m-k-k)}{(m-k)!k!} = \frac{(m-2k)(m-1)!}{(m-k)(m-k-1)!(k-1)!} = \frac{m-2k}{m-k} \binom{m-1}{k},
\]
valid for every \( k \in [m] \). Then, we obtain
\[
\lambda_{[m-1]} = c_{[m-1]} \sum_{k=0}^{m-1} \frac{1}{m-2k} \left( \sum_{a \in \{0,1\}^m} 1 - \sum_{a \in \{0,1\}^m \atop |a|=k} 1 \right) = c_{[m-1]} \sum_{k=0}^{m-1} \frac{1}{m-2k} \binom{m-1}{k} - \binom{m-1}{m-k}) \neq 0.
\]

Therefore, on the one hand we slightly improved the trivial upper bound \( \deg(D_{\mathcal{G}}) \leq m \) for \( D_{\mathcal{G}} \in \mathbb{R}^{\{0,1\}^m} \) used by [LRS15], but on the other hand our lower bound \( \deg(D_{\mathcal{G}}) \geq m-1 \) shows that there are no real numbers \( C > 0, \alpha < 1 \) such that there exists some \( N \in \mathbb{N} \) with \( \deg(D_{\mathcal{G}}) > Cm^\alpha \) for every \( m \in \mathbb{N} \) with \( m \geq N \) and every \( D_{\mathcal{G}} \in \mathbb{R}^{\{0,1\}^m} \). But only this case would improve the lower bound on the psd rank of the family of correlation polytopes.

### 4.7 Applying the main theorem to the knapsack function and the Grigoriev pseudo-density

Now, we are finally able to derive a lower bound on the psd rank of the family of the correlation polytopes, what we will do by applying Theorem [4.44] to the knapsack-type function
\[
f : \{0,1\}^m \to \mathbb{R}_+, \ x \mapsto \left( |x| - \frac{m}{2} \right)^2 - \frac{1}{4}
\]
and to the Grigoriev pseudo-density \( D_{\mathcal{G}} \in \mathbb{R}^{\{0,1\}^m} \) as explained at the beginning of the previous section.
Looking on Theorem 4.44, we see that after computing the infinity norm and the degree of the Grigoriev pseudo-density only the expectation of $f$ is the remaining quantity. Hence, we first compute this expectation, for which we need the following well-known and easy to be proven identity

\[
\binom{m+1}{k+1} = \binom{n}{k} + \binom{n}{k+1},
\]

valid for all $m,k \in \mathbb{Z}$ with $0 \leq k \leq m$.

The expectation of $f$ easily follows from the following lemma, whose proof is not hard but a bit cumbersome. Perhaps one can prove it in a more elegant way than we do.

**Lemma 4.52.**

Let $m \in \mathbb{N}$ be odd. Then, we have

\[
\mathbb{E}_{x \sim \{0,1\}^m} \left( |x| - \frac{m}{2} \right)^2 = \frac{m}{4}.
\]

**Proof.** First, note that we have

\[
\mathbb{E}_{x \sim \{0,1\}^m} \left( |x| - \frac{m}{2} \right)^2 = \frac{1}{2^m} \sum_{k=0}^{m} \left( \binom{m}{k} \left( k - \frac{m}{2} \right)^2 \right).
\]

Therefore, it remains to show

\[
\sum_{k=0}^{m} \binom{m}{k} \left( k - \frac{m}{2} \right)^2 = 2^m \frac{m}{4},
\]

for all odd $m \in \mathbb{N}$. We prove this statement by induction over $m$.

**Base case:** We set $m = 1$, and we obtain

\[
\sum_{k=0}^{1} \left( \binom{1}{k} \left( k - \frac{1}{2} \right)^2 \right) = \left( \left( 0 - \frac{1}{2} \right)^2 + \left( 1 - \frac{1}{2} \right)^2 \right) = \left( \frac{1}{4} + \frac{1}{4} \right) = 2^{1 \frac{1}{2}}.
\]

**Induction step:** We assume that (4.102) is true for some odd $m \in \mathbb{N}$, and we show that it is also true for $m + 2$. Therefore, we assume the induction hypothesis

\[
\sum_{k=0}^{m} \binom{m}{k} \left( k - \frac{m}{2} \right)^2 = 2^m \frac{m}{4}.
\]
Now, we have
\[
\sum_{k=0}^{m+2} \binom{m+2}{k} \left( k - \frac{m+2}{2} \right)^2 = \left( \frac{m+2}{2} \right)^2 + \sum_{k=1}^{m+2} \binom{m+1}{k-1} \binom{m+1}{k} \left( k - \frac{m+2}{2} \right)^2
\]
correlation polytopes.

**Theorem 4.53.**
Let \( m \in \mathbb{N} \) be odd with \( m \geq 3 \), and let \( n \in \mathbb{N} \) with \( n \geq m + 1 \). Then, we have

\[
\text{rk}_{\text{psd}}(\text{CORR}_n) \geq \frac{1}{5} \left( \frac{0.0015 \cdot n}{m^{\frac{11}{2}} \cdot \ln n} + \frac{1}{m^{\frac{11}{2}}} \cdot \frac{1}{m^{\frac{11}{2}}} \right).
\]

**Proof.** Let \( m \in \mathbb{N} \) be odd with \( m \geq 3 \). Then, we define \( f \in \mathbb{R}^{\{0,1\}^m} \) by

\[
f := \frac{4}{m^2 - 1} \left( \sum_{i \in [n]} \chi(i) - \frac{m}{2} \right)^2 - \frac{1}{4}.
\]

Note that \( f \) is nonnegative (cf. Theorem 3.36), and because of \( \max_{x \in \{0,1\}^m} f(x) = \frac{4}{m^2 - 1} \left( \left( \frac{m}{2} \right)^2 - \frac{1}{3} \right) = 1 \), it satisfies the conditions of Theorem 4.44.

Note that we have

\[
\mathbb{E}_x f(x) = \frac{4}{m^2 - 1} \left( \mathbb{E}_x \left( |x| - \frac{m}{2} \right)^2 - \frac{1}{4} \right) \overset{\text{Lem.4.52}}{=} \frac{4}{m^2 - 1} \cdot \frac{m}{4} - \frac{4}{m^2 - 1} \cdot \frac{1}{4} = \frac{m - 1}{m^2 - 1} = \frac{1}{m + 1},
\]

(4.104)

Furthermore, let \( \varepsilon := \frac{4}{m^2 - 1} \), and let \( D_\varphi = D_{\varphi, \frac{m}{2}} \in \mathbb{R}^{\{0,1\}^m} \) be the Grigoriev pseudo-density, that is in particular a degree-\( m \) pseudo-density (cf. Remark 3.58) satisfying

\[
\mathbb{E}_x D_\varphi(x) f(x) = \frac{4}{m^2 - 1} \mathbb{E}_x D_\varphi(x) \left( \left( \sum_{i \in [n]} \chi(i) - \frac{m}{2} \right)^2 - \frac{1}{4} \right) \overset{4.87}{=} - \frac{4}{m^2 - 1} \cdot \frac{1}{4} = - \varepsilon < 0.
\]

Together with Proposition 4.51 we see that \( D_\varphi \) also satisfies the conditions of Theorem 4.44.
Then, for every \(n \in \mathbb{N}\) with \(n \geq m + 1\), applying Theorem 4.44 yields

\[
\text{rk}_{\text{psd}}(M_n) \geq \left( \frac{\frac{1}{374} (n - \frac{m}{2})}{(m^2 - 1)m^2 \deg(D_{\theta}) \|D_{\theta}\|_\infty \ln n} \right)^{\frac{m}{2} + \frac{1}{2}} \left( \frac{1}{9(m^2 - 1) \|D_{\theta}\|_\infty} \right)^{\frac{3}{2}} \sqrt{\mathbb{E}f(x)}
\]

\[
\geq \left( \frac{\frac{1}{737} n}{(m^2 - 1)m^2 \deg(D_{\theta}) \ln n} \right)^{\frac{m}{2} + \frac{1}{2}} \left( \frac{1}{8(m^2 - 1)m^{\frac{3}{2}}} \right)^{\frac{3}{2}} \sqrt{\mathbb{E}f(x)}
\]

Prop. 4.7

\[
\geq \left( \frac{0.0015 n}{(m^2 - 1)m^{\frac{5}{2}} (m - 1) \ln n} \right)^{\frac{m}{2} + \frac{1}{2}} \left( \frac{1}{8(m^2 - 1)m^{\frac{3}{2}}} \right)^{\frac{3}{2}} \frac{1}{\sqrt{m + 1}}
\]

where we used \(1.1m^{7} \geq ((m^2 - 1)^3 (m + 1))\) for all \(m \in \mathbb{N}\) in the second last step.

Now, the desired statement immediately follows by Corollary 4.11. \(\square\)

As an alternative formulation, we immediately obtain the following result.

**Corollary 4.54.**

For every odd \(m \in \mathbb{N}\) with \(m \geq 3\), there exists a real constant \(C_m > 0\) such that

\[
\text{rk}_{\text{psd}}(\text{CORR}_n) \geq C_m \left( \frac{n}{\ln n} \right)^{\frac{m}{2} + \frac{1}{7}}
\]

holds for every \(n \in \mathbb{N}\) with \(n \geq m + 1\).

With this result, it is easy to see that \(\text{rk}_{\text{psd}}(\text{CORR}_n)\) grows faster than any polynomial in \(n\).

**Corollary 4.55.**

For every \(k \in \mathbb{N}\), it holds

\[
\lim_{n \to \infty} \frac{\text{rk}_{\text{psd}}(\text{CORR}_n)}{n^k} = \infty.
\]
Proof. We set \( m := \begin{cases} 4k + 2 & \text{k odd} \\ 4k + 3 & \text{k even} \end{cases} \). In particular, this choice implies \( \frac{m}{2} + \frac{1}{2} \geq k + 1 \). Therefore, by applying the previous corollary, we obtain a constant \( C > 0 \) such that

\[
\lim_{n \to \infty} \frac{\text{rk}_{\text{psd}}(\text{CORR}_n)}{n^k} \geq \lim_{n \to \infty} C \left( \frac{n}{\ln n} \right)^{\frac{m}{2} + \frac{1}{2}} n^{k} \\
\geq \lim_{n \to \infty} C \left( \frac{n}{\ln n} \right)^{k + 1} n^{k} \geq \lim_{n \to \infty} C \frac{n}{(\ln n)^k} = \infty.
\]

Here, we used \( \lim_{n \to \infty} \frac{n}{(\ln n)^k} = \infty \), which can easily be shown by applying L’Hospital’s rule inductively.

Finally, we proceed with a more precise analysis of Theorem 4.53, and we will prove (4.3). This is the main statement of the work of [LRS15], and to provide a proof of it was our main goal, which we stated at the beginning of this chapter. In contrast to [LRS15, Theorem 5.4], where a lot of details were omitted, our proof is more detailed.

**Theorem 4.56** ([LRS15] Theorem 5.4).

There exist \( N \in \mathbb{N} \) and a universal, real constant \( \alpha > 0 \) such that for every \( n \in \mathbb{N} \) with \( n > N \) holds

\[
\text{rk}_{\text{psd}}(\text{CORR}_n) \geq 2^{\alpha \left( \frac{n}{\ln n} \right)^{\frac{2}{11}}}
\]

**Proof.** First, we choose \( n_0 \in \mathbb{N} \) such that \( \frac{n_0}{\ln(n_0)} \geq \frac{4000}{3} \frac{11}{22} \). Such a number \( n_0 \) always exists because of \( \lim_{n \to \infty} \frac{n}{\ln(n)} = \infty \).

Next, we define

\[
c_0 := \frac{1}{5} \left( \frac{3}{4000} \right)^{-\frac{17}{22}}
\]

and

\[
c_1 := \frac{1}{4} \left( \frac{3}{4000} \right)^{\frac{2}{11}}.
\]

Note that there exist some \( n_1 \in \mathbb{N} \) and some real number \( \alpha > 0 \) such that

\[
c_1 \left( \frac{n}{\ln n} \right)^{\frac{2}{11}} + \log_2 c_0 - \frac{17}{22} \log_2 \left( \frac{n}{\ln n} \right) \geq \alpha \left( \frac{n}{\ln n} \right)^{\frac{2}{11}} \tag{4.105}
\]

holds for every \( n \geq n_1 \).

We set

\[
N := \max\{n_0, n_1\}.
\]
Now, let \( n \in \mathbb{N} \) with \( n > N \). Then, we define
\[
m_0 := \left\lfloor \left( \frac{3}{4000} \frac{n}{\ln n} \right)^{\frac{2}{11}} \right\rfloor,
\]
and we set \( m := m_0 \) if \( m_0 \) is odd and \( m := m_0 - 1 \) if \( m_0 \) is even. Hence, note that \( m \) is always odd. Additionally, by our choice of \( N \), we have \( m \geq 3 \) and \( n \geq m + 1 \). So, we can apply Theorem 4.53. Besides, we obtain
\[
m \leq m_0 \leq \left( \frac{3}{4000} \frac{n}{\ln n} \right)^{\frac{2}{11}} \leq m_0 + 1 \leq m + 2.
\]
\[
(4.106)
\]
Applying Theorem 4.53 yields
\[
\text{rk}_{psd}(\text{CORR}_n) \geq \frac{1}{5} \left( \frac{0.0015 \ n}{m^{\frac{1}{2}} \ln n} \right)^{\frac{m}{4} + \frac{1}{2}} m^{-\frac{17}{4}}
\]
\[
\geq \frac{1}{5} \left( \frac{0.0015 \ n}{m^{\frac{1}{2}} \ln n} \right)^{\frac{m}{4} + \frac{1}{2}} m^{-\frac{17}{4}} \tag{4.106}
\]
\[
\geq \frac{1}{5} \left( \frac{0.0015 \ n}{m^{\frac{1}{2}} \ln n} \right)^{\frac{m}{4} + \frac{1}{2}} \left( \frac{3}{4000} \right)^{\frac{17}{22}} \left( \frac{n}{\ln n} \right)^{\frac{17}{22}}
\]
\[
= c_0 2^{\frac{m}{12} + \frac{1}{2} + \frac{1}{2} \log_2 \left( \frac{n}{m^{n}} \right)^{-17/22}}
\]
\[
= 2^{\frac{m}{12} + \frac{1}{2} + \log_2 c_0 - \frac{17}{22} \log_2 \left( \frac{n}{m^{n}} \right)}
\]
\[
\geq 2^{\frac{1}{4} \left( \frac{3}{4000} \right)^{2/11} \left( \frac{n}{m^{n}} \right)^{2/11} + \log_2 c_0 - \frac{17}{22} \log_2 \left( \frac{n}{m^{n}} \right)}
\]
\[
\geq 2^{\frac{1}{4} \left( \frac{3}{4000} \right)^{2/11} \left( \frac{n}{m^{n}} \right)^{2/11}}.
\]
\[
\tag{4.105}
\]
\[\square\]
5 A detailed analysis of the occurring pseudo-densities

Let $n \in \mathbb{N}$ with $n \geq 2$ and let $r \in \left(\left\lceil \frac{n-3}{2} \right\rceil, \left\lfloor \frac{n+3}{2} \right\rfloor \right)$ with $r \notin \mathbb{N}$. Furthermore, let $f := \left(\sum_{i \in [n]} \chi_{\{i\}} - r\right)^2 - b_r \in \mathbb{R}^{(0,1)^n}$, where $b_r$ is defined by

$$b_r := \min_{k \in \mathbb{N}} \{(k-r)^2\} \leq \frac{1}{4}.$$

Considering our observations from the beginning of Section 4.6, we are not limited to use an odd $n$, $r = \frac{n}{2}$ and the Grigoriev pseudo-density $D_{\delta} \in \mathbb{R}^{(0,1)^n}$ in order to apply Theorem 4.44 and derive a lower bound on the psd rank of the family of correlation polytopes as in Theorem 4.53.

Therefore, in this chapter, we want to have a closer view on the different pseudo-densities suitable for the use in the proof of Theorem 4.44 and their impact on the lower bound from this theorem.

Note that we need a degree-$n$ pseudo-density $D \in \mathbb{R}^{(0,1)^n}$ with

$$\varepsilon := -\mathbb{E}_x D(x)f(x) > 0. \quad (5.1)$$

But by linearity, $\varepsilon$ is on his part determined by

$$\delta := \mathbb{E}_x D(x)(|x| - r)^2 \geq 0,$$

which is obvious nonnegative because of the properties of a degree-$n$ pseudo-density. It holds $\varepsilon = b_r - \delta$, and because of $\varepsilon > 0$, we have

$$\delta \in [0,b_r).$$

Overall, the requirements on $D$ are uniquely determined by $r$ and $\delta$. So, for given $n \in \mathbb{N}$, $r \in \left(\left\lceil \frac{n-3}{2} \right\rceil, \left\lfloor \frac{n+3}{2} \right\rfloor \right)$ with $r \notin \mathbb{N}$ and $\delta \in [0,b_r)$, we are looking for degree-$n$ pseudo-densities $D \in \mathbb{R}^{(0,1)^n}$ satisfying

$$\mathbb{E}_x D(x)(|x| - r)^2 = \delta. \quad (5.2)$$

This motivates the following definition.
Definition 5.1.
Let $n \in \mathbb{N}$ with $n \geq 2$, and let $r \in \left(\left\lceil \frac{n-3}{2} \right\rceil, \left\lfloor \frac{n+3}{2} \right\rfloor \right)$ with $r \notin \mathbb{N}$ and $\delta \in [0, b_r)$ be given and fixed. Then, a degree-$n$ pseudo-density $D \in \mathbb{R}^{\{0,1\}^n}$ satisfying (5.2) is called a feasible pseudo-density.

In order to obtain a better lower bound in Theorem 4.53, Theorem 4.44 yields that we need a feasible pseudo-density $D$ with small degree and small infinity norm but with a large value $\varepsilon = b_r - \delta$. All these three quantities are influenced by the choice of $r$ and $\delta$, where their influence on $\varepsilon$ can be easily seen, and what we already used in Section 4.6.

Independently from the choice of $r$, we obtain the best possible value for $\varepsilon$ by setting $\delta = 0$.

In the further course of this chapter, we call this case the standard case.

In Section 4.6, we also saw that the best possible value for $\varepsilon$ can be reached with $r \in \left\{\frac{n-2}{2}, \frac{n}{2}, \frac{n+2}{2}\right\}$ if $n$ is odd, respectively with $r \in \left\{\frac{n-1}{2}, \frac{n+1}{2}\right\}$ if $n$ is even. For odd $n$ and $r = \frac{n}{2}$, $r$ is centered between 0 and $n$, why we call this case the centered case in the further course of the chapter.

Here, we have a special kind of symmetry, and it will turn out that this case is easy to treat.

In Section 5.2, we will give a complete characterization of the feasible pseudodensities in the centered standard case, which we understand very well. We will show that in this case, the Grigoriev pseudo-density indeed yields the optimal lower bound in Theorem 4.44.

In Section 5.3, we will show that this bound cannot be improved by choosing another $\delta \neq 0$ while keeping $r = \frac{n}{2}$ and staying in the centered case.

Last, we will have a short look on the non-centered case in Section 5.4. This case will turn out to be much more difficult than the centered one, and we are able to derive some results and state a conjecture only in the standard case.

5.1 Feasible pseudo-densities and symmetry

If we identify $\mathbb{R}^{\{0,1\}^n}$ with $\mathbb{R}^{2^n}$, we obtain a nice geometric property of the set of feasible pseudo-densities.

Proposition 5.2.
For fixed numbers $n \in \mathbb{N}$, $r \in \left(\left\lceil \frac{n-3}{2} \right\rceil, \left\lfloor \frac{n+3}{2} \right\rfloor \right)$ with $r \notin \mathbb{N}$ and $\delta \in [0, b_r)$, the set of
feasible pseudo-densities forms a spectrahedron in $\mathbb{R}^{2^n}$. In particular, it is a convex set.

**Proof.** In order to be a feasible pseudo-density, a Boolean function $D \in \{0,1\}^n$ has to satisfy three conditions. Two are the ones coming from the degree-$n$ pseudo-density condition (cf. Definition 3.42), the other one is \eqref{eq:5.2}.

Note that \eqref{eq:5.2} as well as the condition $\mathbb{E} D(x) = 1$ are both linear in the variables $(D(x))_{x \in \{0,1\}^n}$, so the set of functions satisfying these two conditions obviously forms a polyhedron and therefore also a spectrahedron (cf. Remark 1.23). So, we consider the remaining degree-$n$ pseudo-density condition

$$
\mathbb{E} D(x) g(x)^2 \geq 0 \quad \text{for every } g \in \{0,1\}^n \text{ with } \deg(g) \leq \frac{n}{2}. \tag{5.3}
$$

First, we fix an arbitrary enumeration

$$
\{I \subseteq [n] \mid |I| \leq \lfloor \frac{n}{2} \rfloor \} = \{I_1, \ldots, I_N\},
$$

where $N := \left( \begin{array}{c} n \\ \lfloor \frac{n}{2} \rfloor \end{array} \right)$. With this enumeration, \eqref{eq:5.3} is equivalent to

$$
\mathbb{E} D(x) \left( \sum_{i \in [N]} \alpha_i \chi_{I_i}(x) \right)^2 \geq 0 \quad \text{for every } \alpha \in \mathbb{R}^N. \tag{5.4}
$$

Now, for every $x \in \{0,1\}^n$ we define a matrix $A_x \in \mathbb{R}^{N \times N}$ by

$$
(A_x)_{i,j} = \frac{1}{2^n} \chi_{I_i \cup I_j}(x).
$$

Then, similar to Lemma 1.31, we transform the left-hand side of \eqref{eq:5.4} and obtain

$$
\mathbb{E} D(x) \left( \sum_{i \in [N]} \alpha_i \chi_{I_i}(x) \right)^2 = \mathbb{E} D(x) \sum_{i,j \in [N]} \alpha_i \alpha_j \chi_{I_i \cup I_j}(x)
$$

$$
= \mathbb{E} D(x) \sum_{i \in [N]} \alpha_i \sum_{j \in [N]} \chi_{I_i \cup I_j}(x) \alpha_j
$$

$$
= \frac{1}{2^n} \sum_{x \in \{0,1\}^n} D(x) \alpha^T \begin{pmatrix} \chi_{I_1 \cup I_1}(x) & \cdots & \chi_{I_1 \cup I_N}(x) \\ \vdots & \ddots & \vdots \\ \chi_{I_N \cup I_1}(x) & \cdots & \chi_{I_N \cup I_N}(x) \end{pmatrix} \alpha
$$

$$
= \sum_{x \in \{0,1\}^n} D(x) \alpha^T A_x \alpha = \alpha^T \left( \sum_{x \in \{0,1\}^n} D(x) A_x \right) \alpha.
$$

Therefore, \eqref{eq:5.4} is equivalent to

$$
\sum_{x \in \{0,1\}^n} D(x) A_x \succeq 0, \tag{5.5}
$$

and every $D \in \{0,1\}^n$ satisfying \eqref{eq:5.3} also satisfies \eqref{eq:5.5}. But by definition, the feasible set of \eqref{eq:5.5} forms a spectrahedron in $\{0,1\}^n \cong \mathbb{R}^n$.

Overall, the set of feasible pseudo-densities forms a spectrahedron, that is in particular a convex set (cf. Remark 1.23). \qed
Remark 5.3.
The above Proposition allows us to handle feasible pseudo-densities or pseudo-
densities in general with semidefinite programming. In particular, we can solve
certain optimization problems over the set of feasible pseudo-densities with an SDP.

We made frequent use of this in the background of this work. For example, we
made numerical experiments to fortify our conjectures or to get ideas and impressions
about the underlying pseudo-densities. For these experiments, we used YALMIP\footnote{in connection with the both commercial softwares MATLAB and the SDP-solver MOSEK} \cite{L04}, an open distributed modeling tool. We would like to sincerely thank Johan
Löfberg for providing this useful tool.

In particular, our two conjectures from Section 5.3 are build on solving a series
of SDPs with YALMIP, but for example also beforehand Theorem 5.5, the main
theorem of the subsequent section, many numerical tests were done in order to get
more insights about the topic.

Next, we show that if there exists a feasible pseudo-density, there also exists a
symmetric one. Moreover, to each feasible pseudo-density there always exists a
symmetric one with smaller or equal infinity norm and smaller or equal degree.

Proposition 5.4.
Let $n \in \mathbb{N}$ with $n \geq 2$, and let $r \in \left(\left\lceil \frac{n-3}{2} \right\rceil, \left\lfloor \frac{n+3}{2} \right\rfloor \right)$ and $\delta \in [0,b_r)$ be fixed. Furthermore, let $D \in \mathbb{R}^{(0,1)^n}$ be a feasible pseudo-density. Then, $\text{Sym}(D)$ is also a feasible pseudo-density. Additionally, we have

$$\|\text{Sym}(D)\|_\infty \leq \|D\|_\infty \text{ and } \deg(\text{Sym}(D)) \leq \deg(D).$$

Proof. Remember the definition

$$\text{Sym}(D) = \frac{1}{n!} \sum_{\sigma \in S_n} D^\sigma$$

of the symmetrization of $D$. Together with $\deg(D) = \deg(D^\sigma)$, valid for every
$\sigma \in S_n$ and what is obvious by the definition of $D^\sigma$, we immediately obtain

$$\deg(\text{Sym}(D)) \leq \deg(D).$$

Now, we show that $\text{Sym}(D)$ is a feasible pseudo-density. In order to do this, we
firstly show that $D^\sigma$ is a feasible pseudo-density for every $\sigma \in S_n$. In order to do
this, we fix an arbitrary $\sigma \in S_n$. Then, we obtain

$$\mathbb{E}_x D^\sigma(x) (|x| - r)^2 = \mathbb{E}_x D^\sigma(x) (|x^\sigma| - r)^2$$

$$= \mathbb{E}_x D(x^\sigma) (|x^\sigma| - r)^2$$

$$\overset{(3.12)}{=} \mathbb{E}_x D(x) (|x| - r)^2$$

$$= \delta,$$
where we used $|x| = |x^\sigma|$ in the first step and the assumption in the last one.

Because of this observation, it remains to show that $D^\sigma$ is a degree-$n$ pseudo-density as well. In order to show this, let $g \in \mathbb{R}^{(0,1)^n}$ with $\deg(g) \leq \frac{n}{2}$. We define $\hat{g} \in \mathbb{R}^{(0,1)^n}$ by $\hat{g} := g^{\sigma^{-1}}$. Note that

$$\hat{g}(x^\sigma) = g^{\sigma^{-1}}(x^\sigma) = g\left((x^\sigma)^{\sigma^{-1}}\right)^{\text{Rem}3.24(a)} g(x).$$  \hspace{1cm} (5.6)$$

Then, we have

$$\mathbb{E}_x D^\sigma(x)g(x)^2 = \mathbb{E}_x D(x^\sigma)g(x)^2 \overset{\text{Rem}3.12}{=} \mathbb{E}_x D(x^\sigma)\hat{g}(x)^2 \geq 0.$$

In the last step, we used Remark 3.24(c) to obtain $\deg(\hat{g}) = \deg(g^{\sigma^{-1}}) = \deg(g) \leq \frac{n}{2}$, and additionally we used the assumption that $D$ is a degree-$n$ pseudo-density. Note that $\mathbb{E}_x D^\sigma(x) = \mathbb{E}_x D(x^\sigma) = 1$ also holds, whereby altogether follows that $D^\sigma$ is a feasible pseudo-density.

Therefore, $\text{Sym}(D)$ is a convex combination of feasible pseudo-density, and because the set of feasible pseudo-densities is convex, $\text{Sym}(D)$ is also a feasible pseudo-density.

We have

$$|\text{Sym}(D)(x)| = \frac{1}{n!} \left| \sum_{\sigma \in S_n} D(x^\sigma) \right| \leq \frac{1}{n!} \sum_{\sigma \in S_n} |D(x^\sigma)|$$

$$\leq \frac{1}{n!} \sum_{\sigma \in S_n} \|D\|_{\infty} = \|D\|_{\infty}$$

for every $x \in \{0,1\}^n$, which implies

$$\|\text{Sym}(D)\|_{\infty} = \max_{x \in \{0,1\}^n} |\text{Sym}(D)(x)| \leq \|D\|_{\infty}.$$
translation is also given by this Corollary.

For \( n \in \mathbb{N} \) with \( n \geq 2 \) and fixed \( r \in \left(\lceil \frac{n-3}{2} \rceil, \left\lfloor \frac{n+3}{2} \right\rfloor \right) \) and \( \delta \in [0, b_r) \), the linear forms \( L : \mathbb{R}[T]_n \to \mathbb{R} \) corresponding to symmetric feasible pseudo-densities satisfy the following conditions.

\[
L(1) = 1, \quad (I)
\]
\[
L(g) \geq 0 \quad \text{for all } g \in \sum_{k=0}^{\left\lceil \frac{n}{2} \right\rceil} \prod_{j=0}^{k-1} (T - j)(n - j - T) \sum_{T} \mathbb{R}[T]_{\left\lceil \frac{n}{2} \right\rceil - k}, \quad (II)
\]
\[
L \left( (T - r)^2 \right) = \delta. \quad (III)
\]

**5.2 Grigoriev pseudo-densities are optimal - a complete analysis of symmetric pseudo-densities in the centered standard case**

In this section, we have a closer view on the feasible pseudo-densities respectively their corresponding linear forms from \( \mathbb{R}[T]_n \) to \( \mathbb{R} \) in the centered standard case. As we saw before, this is a case where the best possible value for \( \varepsilon \), defined as in (5.1), is reached. It will turn that we understand the feasible pseudo-densities very well in this case, and we will derive some optimality result for the Grigoriev pseudo-density.

Note that in this case condition [III] translates into

\[
L \left( \left( T - \frac{n}{2} \right)^2 \right) = 0. \quad (III')
\]

Initially, it is entirely unclear why just the Grigoriev pseudo-density should be the best choice in terms of a minimal infinity norm.

Contrariwise, if we look at the conditions which the corresponding linear form \( L : \mathbb{R}[T]_n \to \mathbb{R} \) of a feasible symmetric pseudo-density \( D \in \mathbb{R}^{(0,1)^n} \) has to satisfy, there is a reason why the choice of the Grigoriev pseudo-density could be rather specific. Because of the conditions [I], [III] and Proposition [1.32] such a linear form satisfies

\[
L \left( \left( T - \frac{n}{2} \right) g \right) = 0
\]
for every \( g \in \mathbb{R}[T]_n \) with \( \deg(g) \leq \frac{n}{2} \), while the Grigoriev linear form \( \mathcal{L}_g \) even satisfies

\[
\mathcal{L}_g \left( \left( T - \frac{n}{2} \right) g \right) = 0
\]
for every \( g \in \mathbb{R}[T]_n \) with \( \deg(g) \leq n - 1 \) what follows by Lemma [3.55].
Therefore, the Grigoriev linear form and hence also its corresponding pseudo-density has a much stronger property than required. This stronger requirement could be quite adversely for a small infinity norm. So, it seems possible to find a feasible pseudo-density with smaller infinity norm than the Grigoriev density.

Nevertheless, it turns out that the Grigoriev density is the best choice. In order to prove this, our next theorem is a big step. It shows that under the conditions (I), (II) and (III') a linear form is determined on almost all values, and it can be seen as the main theorem of this chapter.

**Theorem 5.5.**

Let $n \in \mathbb{N}$ be odd with $n \geq 3$, and let $L : \mathbb{R}[T]_n \to \mathbb{R}$ be a linear form satisfying the conditions (I), (II) and (III'). Then,

$$L(T^k) = \left(\frac{n}{2}\right)^k$$

for all $k \in \{0, \ldots, n - 1\}$. In particular, it holds

$$L(p) = p \left(\frac{n}{2}\right) \text{ for all } p \in \mathbb{R}[T]_n \text{ with } \deg(p) \leq n - 1.$$  

**Proof.** We prove the theorem by induction over $k$.

**Base case:** The base case will include all $k \in \{0, \ldots, \frac{n+1}{2}\}$. The conditions (II) and (III') together with Proposition 1.32 imply

$$L \left( \left( T - \frac{n}{2} \right) T^k \right) = 0,$$

for every $k \in \{0, \ldots, \frac{n-1}{2}\}$. This is equivalent to

$$L(T^{k+1}) = \frac{n}{2} L(T^k), \quad (5.7)$$

for every $k \in \{0, \ldots, \frac{n-1}{2}\}$.

For $k = 0$, the theorem is true because of $L(1) = 1$. Putting this into (5.7) implies $L(T) = \frac{n}{2}$. Again putting this into (5.7) implies $L(T^2) = \left(\frac{n}{2}\right)^2$. We can continue in this way up to $k = \frac{n-1}{2}$. Overall, we obtain

$$L(T^{k+1}) = \left(\frac{n}{2}\right)^{k+1}$$

for every $k \in \{0, \ldots, \frac{n-1}{2}\}$, or equivalently

$$L(T^k) = \left(\frac{n}{2}\right)^k \quad (5.8)$$
for every $k \in \{0, \ldots, \frac{n+1}{2}\}$.

Note that in the case $n = 3$ we are already done. The case $n = 5$ will be separately proved at the end of the proof. Therefore, without loss of generality let $n \geq 7$ from now on.

**Induction step:** Let $k \in \{\frac{n+3}{2}, \ldots, n-1\}$ be even and $k \geq 6$. We assume that the statement of our theorem is valid for all $j \in \{k-5, \ldots, k-2\}$, that means

$$L(T^j) = \left(\frac{n}{2}\right)^j.$$  

(5.9)

Now, we simultaneously show that this also holds for $k-1$ and $k$.

Condition (I) implies $L\left((aT^{\frac{k-1}{2}} + bT^{\frac{k+1}{2}})^2\right) \geq 0$ for all $a, b \in \mathbb{R}$. Then, Lemma 1.31 implies

$$\begin{pmatrix} L(T^{k-2}) & L(T^{k-1}) \\ L(T^{k-1}) & L(T^k) \end{pmatrix} \succeq 0,$$

and the induction hypothesis (5.9) yields

$$\begin{pmatrix} \left(\frac{n}{2}\right)^{k-2} L(T^{k-1}) \\ L(T^{k-1}) & L(T^k) \end{pmatrix} \succeq 0.$$

By Corollary 1.25, we obtain

$$\left(\frac{n}{2}\right)^{k-2} L(T^k) - L(T^{k-1})^2 \geq 0,$$

which implies

$$L(T^k) \geq \frac{L(T^{k-1})^2}{\left(\frac{n}{2}\right)^{k-2}}.$$  

(5.10)

Again by condition (I), we obtain $L\left(T(n-T)(aT^{\frac{k-3}{2}} + bT^{\frac{k-1}{2}})^2\right) \geq 0$ for all $a, b \in \mathbb{R}$. Then, Lemma 1.31 implies

$$\begin{pmatrix} L(T(n-T)T^{k-4}) & L(T(n-T)T^{k-3}) \\ L(T(n-T)T^{k-3}) & L(T(n-T)T^{k-2}) \end{pmatrix} \succeq 0,$$

or equivalent

$$\begin{pmatrix} nL(T^{k-3}) - L(T^{k-2}) & nL(T^{k-2}) - L(T^{k-1}) \\ nL(T^{k-2}) - L(T^{k-1}) & nL(T^{k-1}) - L(T^k) \end{pmatrix} \succeq 0.$$

Then, the induction hypothesis (5.9) yields

$$\begin{pmatrix} \left(\frac{n}{2}\right)^{k-3} - \left(\frac{n}{2}\right)^{k-2} & \left(\frac{n}{2}\right)^{k-2} - L(T^{k-1}) \\ \left(\frac{n}{2}\right)^{k-2} - L(T^{k-1}) & nL(T^{k-1}) - L(T^k) \end{pmatrix} \succeq 0,$$
or equivalent
\[
\begin{pmatrix}
\frac{n}{2}^{k-2} & 2\left(\frac{n}{2}\right)^{k-1} - L(T^{k-1}) \\
2\left(\frac{n}{2}\right)^{k-1} - L(T^{k-1}) & nL(T^{k-1}) - L(T^{k})
\end{pmatrix} \geq 0.
\]
By Corollary \[ \text{1.25}, \]
we obtain
\[
\left(\frac{n}{2}\right)^{k-2} \left(nL(T^{k-1}) - L(T^{k})\right) - \left(2\left(\frac{n}{2}\right)^{k-1} - L(T^{k-1})\right)^2 \geq 0.
\]
Now, multiplying out the brackets results in
\[
-\left(\frac{n}{2}\right)^{k-2} L(T^{k}) - L(T^{k-1})^2 + 6\left(\frac{n}{2}\right)^{k-1} L(T^{k-1}) - 4\left(\frac{n}{2}\right)^{2k-2} \geq 0,
\]
which implies
\[
L(T^{k}) \leq \frac{-L(T^{k-1})^2 + 6\left(\frac{n}{2}\right)^{k-1} L(T^{k-1}) - 4\left(\frac{n}{2}\right)^{2k-2}}{\left(\frac{n}{2}\right)^{k-2}}.
\]
Applying condition \[ (\text{II}) \] a third time implies
\[
L\left(T(n - T)(T - 1)(n - 1 - T)(aT^{\frac{5}{2}-3} + bT^{\frac{5}{2}-2})^2\right) \geq 0
\]
for all \(a, b \in \mathbb{R}\). We define \(p := T(n - T)(T - 1)(n - 1 - T)\). Then, analogue as above, Lemma \[ \text{1.31}, \]
implies
\[
\begin{pmatrix}
L(p \cdot T^{k-6}) & L(p \cdot T^{k-5}) \\
L(p \cdot T^{k-5}) & L(p \cdot T^{k-4})
\end{pmatrix} \geq 0.
\]
We multiply out the inner of \(L\) and use the linearity of \(L\) in order to obtain
\[
\begin{pmatrix}
a_{11} & a_{12} \\
a_{12} & a_{22}
\end{pmatrix} \geq 0,
\]
where
\[
a_{11} := L(T^{k-2}) - 2nL(T^{k-3}) + (n^2 + n - 1)L(T^{k-4}) - (n^2 - n)L(T^{k-5}),
\]
\[
a_{12} := L(T^{k-1}) - 2nL(T^{k-2}) + (n^2 + n - 1)L(T^{k-3}) - (n^2 - n)L(T^{k-4})
\]
and
\[
a_{22} := L(T^{k}) - 2nL(T^{k-1}) + (n^2 + n - 1)L(T^{k-2}) - (n^2 - n)L(T^{k-3}).
\]
By applying the induction hypothesis \[ \text{5.9}, \]
we obtain
\[
a_{11} = \left(\frac{n}{2}\right)^{k-2} - 2n\left(\frac{n}{2}\right)^{k-3} + (n^2 + n - 1)\left(\frac{n}{2}\right)^{k-4} - (n^2 - n)\left(\frac{n}{2}\right)^{k-5},
\]
\[
a_{12} = L(T^{k-1}) - 2n\left(\frac{n}{2}\right)^{k-2} + (n^2 + n - 1)\left(\frac{n}{2}\right)^{k-3} - (n^2 - n)\left(\frac{n}{2}\right)^{k-4}
\]
and
\[
a_{22} = L(T^{k}) - 2nL(T^{k-1}) + (n^2 + n - 1)\left(\frac{n}{2}\right)^{k-2} - (n^2 - n)\left(\frac{n}{2}\right)^{k-3}.
\]
After simplifying them, these expressions turn into

\[ a_{11} = \left( \frac{n}{2} - 1 \right)^2 \left( \frac{n}{2} \right)^{k-4}, \]

\[ a_{12} = L(T^{k-1}) - (n-1) \left( \frac{n}{2} \right)^{k-3} \]

and

\[ a_{22} = L(T^k) - 2nL(T^{k-1}) + (n^2 - n + 1) \left( \frac{n}{2} \right)^{k-2}. \]

Applying Corollary 1.25 to (5.12) yields

\[ a_{11}a_{22} - a_{12}^2 \geq 0, \]

and therefore

\[ \left( L(T^k) - 2nL(T^{k-1}) + (n^2 - n + 1) \left( \frac{n}{2} \right)^{k-2} \right) \left( \frac{n}{2} \right)^{k-4} \]

\[ - \left( L(T^{k-1}) - (n-1) \left( \frac{n}{2} \right)^{k-3} \right)^2 \geq 0. \]

After multiplying out, simplifying and solving the inequality for \( L(T^k) \), we obtain

\[ L(T^k) \geq \frac{L(T^{k-1})^2 + (n^2 - 6n + 6) \left( \frac{n}{2} \right)^{k-3} L(T^{k-1}) - (n^2 - 5n + 5) \left( \frac{n}{2} \right)^{2k-4}}{\left( \frac{n}{2} - 1 \right)^2 \left( \frac{n}{2} \right)^{k-4}}. \]  

(5.13)

The three inequalities (5.10), (5.11) and (5.13) build a system of inequalities with the two unknowns \( L(T^k) \) and \( L(T^{k-1}) \). For the sake of convenience, we substitute

\[ x := L(T^{k-1}) \] and \( y := L(T^k) \).

Then, combining (5.10) and (5.11) yields

\[ \frac{x^2}{\left( \frac{n}{2} \right)^{k-2}} \leq y \leq -x^2 + 6 \left( \frac{n}{2} \right)^{k-1} x - 4 \left( \frac{n}{2} \right)^{2k-2}. \]

The latter is equivalent to

\[ x^2 - 3 \left( \frac{n}{2} \right)^{k-1} x + 2 \left( \frac{n}{2} \right)^{2k-2} \leq 0, \]

which is again equivalent to

\[ (x - \left( \frac{n}{2} \right)^{k-1}) \left( x - 2 \left( \frac{n}{2} \right)^{k-1} \right) \leq 0 \]

by Vieta’s formulas. It follows

\[ \left( \frac{n}{2} \right)^{k-1} \leq x \leq 2 \left( \frac{n}{2} \right)^{k-1}. \]  

(5.14)
Combining the inequalities (5.11) and (5.13) yields
\[
x^2 + (n^2 - 6n + 6) \left( \frac{n}{2} \right)^{k-3} x - (n^2 - 5n + 5) \left( \frac{n}{2} \right)^{2k-4} \leq \frac{y}{\left( \frac{n}{2} - 1 \right)^2 \left( \frac{n}{2} \right)^{k-4}} \leq -x^2 + 6 \left( \frac{n}{2} \right)^{k-1} x - 4 \left( \frac{n}{2} \right)^{2k-2} \left( \frac{n}{2} \right)^{k-2}.
\]

By multiplying with \( \left( \frac{n}{2} - 1 \right)^2 \left( \frac{n}{2} \right)^{k-2} \), summarization of terms and some easy computation afterwards, we obtain
\[
\left( \frac{n^2}{2} - n + 1 \right) x^2 - 2 \left( \frac{n}{2} \right)^{k+1} x + (n - 1) \left( \frac{n}{2} \right)^{2k-2} \leq 0.
\]

With the help of the quadratic formula, we can factorize the left-hand side and obtain
\[
\frac{2n - 2}{n^2 - 2n + 2} \left( \frac{n}{2} \right)^{k-1} \leq x \leq \left( \frac{n}{2} \right)^{k-1}, \tag{5.15}
\]
where we used \( \frac{2n-2}{n^2-2n+2} \leq 1 \) and \( n^2 - 2n + 2 \geq 0 \), valid for all \( n \in \mathbb{N} \).

Combining (5.14) and (5.15) yields
\[
L(T^{k-1}) = x = \left( \frac{n}{2} \right)^{k-1}. \tag{5.16}
\]

Inserting \( L(T^{k-1}) = \left( \frac{n}{2} \right)^{k-1} \) in the inequalities (5.10) and (5.11) results in
\[
L(T^k) \overset{5.10}{\leq} L(T^{k-1})^2 \overset{5.16}{=} \left( \frac{n}{2} \right)^{2k-2} = \left( \frac{n}{2} \right)^k \tag{5.17}
\]
and
\[
L(T^k) \overset{5.11}{\leq} -L(T^{k-1})^2 + 6 \left( \frac{n}{2} \right)^{k-1} L(T^{k-1}) - 4 \left( \frac{n}{2} \right)^{2k-2} \overset{5.16}{=} \left( \frac{n}{2} \right)^{2k-2} - 6 \left( \frac{n}{2} \right)^{2k-2} - 4 \left( \frac{n}{2} \right)^{2k-2} = \left( \frac{n}{2} \right)^k. \tag{5.18}
\]

By (5.17) and (5.18), we obtain
\[
L(T^k) = \left( \frac{n}{2} \right)^k.
\]
Together with (5.16), this yields the desired statement.

Now, we still have to proof the remaining case \( n = 5 \). In order to do this, note that
the base case includes every \( k \in \{0, \ldots, 3\} \). So, it remains to show \( L(T^4) = \left(\frac{n}{2}\right)^4 \).
Again, note that in the induction step \( k \geq 6 \) is only required to obtain inequality
(5.13). In order to obtain the other two important inequalities (5.10) and (5.11), \( k \geq 4 \) is sufficient. Together with
\[
L(T^3) = \left(\frac{n}{2}\right)^3,
\]
these two inequalities will be sufficient to finish the proof.

We have
\[
L(T^4) \overset{(5.10)}{\geq} \frac{L(T^3)^2}{\left(\frac{n}{2}\right)^2} \overset{(5.19)}{=} \frac{\left(\frac{n}{2}\right)^6}{\left(\frac{n}{2}\right)^2} = \left(\frac{n}{2}\right)^4
\]
and
\[
L(T^4) \overset{(5.11)}{\leq} -L(T^3)^2 + 6 \left(\frac{n}{2}\right)^3 L(T^3) - 4 \left(\frac{n}{2}\right)^6
\]
\[
\overset{(5.19)}{=} \frac{-\left(\frac{n}{2}\right)^6 + 6 \left(\frac{n}{2}\right)^6 - 4 \left(\frac{n}{2}\right)^6}{\left(\frac{n}{2}\right)^2} = \left(\frac{n}{2}\right)^4,
\]
which implies
\[
L(T^4) = \left(\frac{n}{2}\right)^4
\]
as desired. \( \Box \)

Now, in the view of this theorem, the Grigoriev linear form seems much less
restrictive and specific than before. For an arbitrary linear form \( L \) satisfying (I), (II)
and (III'), \( L(p) \) is just evaluating \( p \) in \( \frac{n}{2} \) for all univariate polynomials \( p \) up to degree
\( n - 1 \), while the Grigoriev linear form \( L_G \) evaluates every polynomial \( p \in \mathbb{R}[T]_n \) in
\( \frac{n}{2} \) (cf. Lemma 3.55).

Another conclusion concerning the symmetric, feasible pseudo-densities can be
done by the previous theorem.

**Corollary 5.6.**
For a fixed, odd \( n \in \mathbb{N} \) with \( n \geq 3 \), \( r = \frac{n}{2} \) and \( \delta = 0 \), the set of symmetric, feasible
pseudo-densities is one-dimensional.

That means, there exist universal \( \alpha, \beta \in \mathbb{R}^{(0,1)^n} \) such that
\[
D = \alpha + s \cdot \beta
\]
for some \( s \in \mathbb{R} \) for every arbitrary symmetric, feasible pseudo-density \( D \in \mathbb{R}^{(0,1)^n} \).
Chapter 5. A detailed analysis of the occurring pseudo-densities

Proof. For every $k \in \{0, \ldots, n\}$, we define $a_k := \frac{2^n}{(n/k) \prod_{j=0}^{n-k} T - j} \prod_{j \neq k} T - j \in \mathbb{R}[T]_n$ and $q_k := p_k - T^n \in \mathbb{R}[T]_n$. Note that the definition of $p_k$ implies

$$\deg(q_k) \leq n - 1 \text{ for every } k \in \{0, \ldots, n\}. \quad (5.20)$$

Furthermore, we define $\alpha, \beta \in \mathbb{R}^{\{0,1\}^n}$ by

$$\alpha(x) := a_{|x|} q_{|x|} \left( \frac{n}{2} \right) \text{ and } \beta(x) = a_{|x|} \quad (5.21)$$

for every $x \in \{0,1\}^n$.

Now, let $D \in \mathbb{R}^{\{0,1\}^n}$ be any arbitrary symmetric and feasible pseudo-density and $L : \mathbb{R}[T]_n \to \mathbb{R}$ its corresponding linear form. Then, $L$ satisfies the conditions (I), (II) and (III'). Therefore, applying the previous theorem to (5.20) implies

$$L(q_k) = q_k \left( \frac{n}{2} \right) \text{ for every } k \in \{0, \ldots, n\}. \quad (5.22)$$

Then, Lemma 3.56 implies

$$D(x) = a_{|x|} L(p_{|x|}) = a_{|x|} L(q_{|x|}) + a_{|x|} L(T^n) \quad (5.22)$$

$$= a_{|x|} q_{|x|} \left( \frac{n}{2} \right) + a_{|x|} L(T^n) \quad (5.21)$$

$$= \alpha(x) + L(T^n) \beta(x)$$

for every $x \in \{0,1\}^n$. \hfill \Box

Besides the invariance under the action of $S_n$, a closer view to the Grigoriev pseudo-density $D_G \in \mathbb{R}^{\{0,1\}^n}$ shows a second kind of symmetry. It satisfies $D_G(x) = D_G(1 - x_1, \ldots, 1 - x_n)$ for every $x \in \{0,1\}^n$. We want to formalize this property.

Definition 5.7.

Let $n \in \mathbb{N}$. For $x \in \{0,1\}^n$, we define $\bar{x} \in \{0,1\}^n$ by

$$\bar{x} = (1 - x_1, \ldots, 1 - x_n).$$

Furthermore, to every function $D \in \mathbb{R}^{\{0,1\}^n}$, we define

$$\bar{D} : \{0,1\}^n \to \mathbb{R}, \ x \mapsto D(\bar{x})$$

A symmetric function $D \in \mathbb{R}^{\{0,1\}^n}$ is called double-symmetric if it satisfies

$$D = \bar{D},$$

or equivalent

$$D(x) = D(\bar{x}) \text{ for every } x \in \{0,1\}^n.$$
Remark 5.8.
Let \( n \in \mathbb{N} \). Note that the following statements are valid:

(a) \( |\bar{x}| = n - |x| \) for every \( x \in \{0,1\}^n \),

(b) \( \{\bar{x} \mid x \in \{0,1\}^n\} = \{0,1\}^n \),

(c) \( \bar{x} = x \) for every \( x \in \{0,1\}^n \).

Lemma 5.9.
Let \( n \in \mathbb{N} \) and \( r = \frac{n}{2} \). Then, the Grigoriev pseudo-density \( D_g \in \mathbb{R}^{\{0,1\}^n} \) is double-symmetric.

Proof. Let \( x \in \{0,1\}^n \). Then, it holds

\[
D_g(x) = 2^n \frac{\prod_{j \neq |x|} \binom{n}{j} \left(\frac{n}{2} - j\right) \prod_{j \neq |x|} \binom{n}{j} (|x| - j)}{\prod_{j \neq n - |x|} \binom{n}{j} (n - |x| - j)} = 2^n \frac{\prod_{i \neq |x|} \binom{n}{i} (-\frac{n}{2} + i) \prod_{i \neq |x|} \binom{n}{i} (i - |x|)}{\prod_{i \neq |x|} \binom{n}{i} (|x| - i)} = D_g(x).
\]

The next statement shows that in the centered case there is always a feasible double-symmetric pseudo-density, and they are exactly the ones with minimal infinity norm and minimal degree. In order to also use the proposition later, we prove it more general for arbitrary \( \delta \in \left[0,\frac{1}{4}\right) \) although we especially treat the case \( \delta = 0 \) during this section.

Proposition 5.10.
Let \( n \in \mathbb{N} \), \( r = \frac{n}{2} \) and \( \delta \in \left[0,\frac{1}{4}\right) \). Furthermore, let \( D \in \mathbb{R}^{\{0,1\}^n} \) be a symmetric, feasible pseudo-density. Then, there exists a double-symmetric function \( \hat{D} \in \mathbb{R}^{\{0,1\}^n} \) which is also a feasible pseudo-density and which additionally satisfies

\[
\|\hat{D}\|_\infty \leq \|D\|_\infty \quad \text{and} \quad \deg(\hat{D}) \leq \deg(D).
\]

Proof. First, we show that \( \hat{D} \) is a feasible pseudo-density. We have

\[
\frac{\mathbb{E} \hat{D}(x)}{\mathbb{E} D(x)} = \frac{\mathbb{E} D(x)}{\mathbb{E} D(x)} = 1
\]
Chapter 5. A detailed analysis of the occurring pseudo-densities

and
\[
\mathbb{E}_{\bar{x}} \bar{D}(x) \left( |x| - \frac{n}{2} \right)^2 = \mathbb{E}_{\bar{x}} D(\bar{x}) \left( |\bar{x}| - \frac{n}{2} \right)^2 = \mathbb{E}_{\bar{x}} D(\bar{x}) \left( n - |x| - \frac{n}{2} \right)^2
\]
\[
= \delta.
\]

Now, it remains to show that \( \bar{D} \) is a degree-\( n \) pseudo-density. In order to do this, let \( g \in \mathbb{R}^{\{0,1\}^n} \) with \( \deg(g) \leq \frac{n}{2} \). We define \( \bar{g} \in \mathbb{R}^{\{0,1\}^n} \) by \( \bar{g}(x) := g(\bar{x}) = g(1 - x_1, \ldots, 1 - x_n) \). Note that
\[
\deg(\bar{g}) = \deg(g) \leq \frac{n}{2}
\]
and
\[
\bar{g}(\bar{x}) = g(x),
\]
for every \( x \in \{0,1\}^n \). Then, we have
\[
\mathbb{E}_{\bar{x}} \bar{D}(x) g(x)^2 = \mathbb{E}_{\bar{x}} D(\bar{x}) g(\bar{x})^2 = \mathbb{E}_{\bar{x}} D(\bar{x}) \bar{g}(\bar{x})^2 \]
\[
\geq 0,
\]
where in the last step we used that \( D \) is a degree-\( n \) pseudo-density together with \((5.23)\). Altogether, \( \bar{D} \) is a feasible pseudo-density.

Therefore, we define
\[
\hat{D} = \frac{1}{2} D + \frac{1}{2} \bar{D}.
\]

Then, because of the convexity of the set of feasible pseudo-densities (cf. Proposition \(5.2\)), \( \hat{D} \) is also a feasible pseudo-density. Additionally, we have
\[
\hat{D}(\bar{x}) = \frac{D(\bar{x}) + \bar{D}(\bar{x})}{2} = \frac{D(\bar{x}) + D(\bar{x})}{2}
\]
\[
= \hat{D}(x),
\]
which implies that \( \hat{D} \) is double-symmetric.

Furthermore, we have
\[
\left| \hat{D}(x) \right| = \left| \frac{1}{2} D(x) + \frac{1}{2} \bar{D}(x) \right|
\]
\[
\leq \frac{1}{2} |D(x)| + \frac{1}{2} |\bar{D}(\bar{x})|
\]
\[
\leq \| D \|_{\infty},
\]
for every \( x \in \{0,1\}^n \), which implies
\[
\|\hat{D}\|_\infty = \max_{x \in \{0,1\}^n} |\hat{D}(x)| \leq \|D\|_\infty.
\]

Finally, note that because \( \bar{D} \) is just an affine linear transformation of \( D \), which implies \( \deg(D) = \deg(\bar{D}) \). Therefore, we obviously have
\[
\deg(\hat{D}) \leq \deg(D)
\]
by definition.

Combining this proposition with Proposition 5.4 shows that in the centered case it is sufficient to consider double-symmetric feasible pseudo-densities if we are looking for a feasible pseudo-density with minimal infinity norm or minimal degree. We state this observation in the following corollary.

**Corollary 5.11.**

Let \( n \in \mathbb{N} \), \( r = \frac{n}{2} \) and \( \delta \in \left[0, \frac{1}{4}\right) \). Then, to each feasible pseudo-density \( D \in \mathbb{R}^{(0,1)^n} \) there exists a double-symmetric feasible pseudo-density \( \bar{D} \in \mathbb{R}^{(0,1)^n} \) satisfying
\[
\|\bar{D}\|_\infty \leq \|D\|_\infty \quad \text{and} \quad \deg(\bar{D}) \leq \deg(D).
\]

Note that the corresponding linear form \( L \) of a double-symmetric pseudo-density \( D \in \mathbb{R}^{(0,1)^n} \) satisfies
\[
L(p) = \mathbb{E} D(x)p(|x|) \overset{\text{Rem. 5.8}}{=} \mathbb{E} \bar{D}(\bar{x})p(|\bar{x}|) = \mathbb{E} D(x)p(n - |x|)
\]
for every \( p \in \mathbb{R}[T]_n \).

Therefore, a linear form \( L : \mathbb{R}[T]_n \to \mathbb{R} \) corresponding to a double-symmetric feasible pseudo-density satisfies, in addition to (I)-(III), the following condition:
\[
L(p) = L(p(n - T)) \quad \text{for every} \quad p \in \mathbb{R}[T]_n.
\]

Now, by extending Theorem 5.5, it turns out that a linear form is uniquely determined by the conditions (I), (II), (III) and (IV).

**Theorem 5.12.**

Let \( n \in \mathbb{N} \) be odd with \( n \geq 3 \) and let \( L : \mathbb{R}[T]_n \to \mathbb{R} \) be a linear form satisfying the conditions (I), (II), (III) and (IV). Then,
\[
L(p) = p \left( \frac{n}{2} \right)
\]
for every \( p \in \mathbb{R}[T]_n \). In particular, \( L \) is the Grigoriev linear form \( \mathcal{L}_g \).
Proof. Because of Theorem 5.5, we already have $L(p) = p \left(\frac{n}{2}\right)$ for every $p \in \mathbb{R}[T]_n$ with $\deg(p) \leq n - 1$. In particular, we have

$$L(T^k) = \left(\frac{n}{2}\right)^k$$

for every $k \in \{0, \ldots, n-1\}$. (5.25)

By the linearity of $L$, it remains to show $L(T^n) = \left(\frac{n}{2}\right)^n$.

In order to do this, note that condition (IV) implies

$$L(T^n) = L((n - T)^n).$$

Therefore, we obtain

$$L(T^n) = L((n - T)^n) = \sum_{i=0}^{n-1} \binom{n}{i} n^{n-i} (-T)^i - L(T^n).$$

This implies

$$2L(T^n) = \sum_{i=0}^{n-1} \binom{n}{i} n^{n-i} (-1)^i L(T^i) \quad \text{by (5.25)}$$

$$= \sum_{i=0}^{n-1} \binom{n}{i} n^{n-i} \left(-\frac{n}{2}\right)^i = \sum_{i=0}^{n-1} \binom{n}{i} n^{n-i} \left(-\frac{n}{2}\right)^i + \left(-\frac{n}{2}\right)^n - \left(-\frac{n}{2}\right)^n$$

$$= \sum_{i=0}^{n} \binom{n}{i} n^{n-i} \left(-\frac{n}{2}\right)^i - \left(-\frac{n}{2}\right)^n = \left(n - \frac{n}{2}\right)^n + \left(\frac{n}{2}\right)^n = 2 \left(\frac{n}{2}\right)^n,$$

which implies

$$L(T^n) = \left(\frac{n}{2}\right)^n.$$

Because of the one-to-one correspondence between symmetric pseudo-densities and linear forms satisfying (I)-(III), we obtain the following corollary as an immediate consequence of the previous theorem.

**Corollary 5.13.**

For every fixed odd $n \in \mathbb{N}$ with $n \geq 3$, $r = \frac{n}{2}$ and $\delta = 0$, the Grigoriev pseudo-density $D_G \in \mathbb{R}^{\{0,1\}^n}$ is the unique double-symmetric, feasible pseudo-density.

With this statement, we obtain additional information about symmetric, feasible pseudo-densities.
Corollary 5.14.
Let \( n \in \mathbb{N} \) be odd with \( n \geq 3 \), \( r = \frac{n}{2} \) and \( \delta = 0 \). Then, for every symmetric, feasible pseudo-density \( D \in \mathbb{R}^{(0,1)^n} \) it holds
\[
D + \bar{D} = 2D_\varrho.
\]

In particular, we have
\[
D(x) + D(\bar{x}) = \frac{(-1)^{\frac{m-1}{2} - |x|}}{(m - 2|x|)} \cdot \frac{m}{2^{m-2}} \left( \frac{m - 1}{2} \right)
\]
for every \( x \in \{0,1\}^n \).

Proof. The proof of Proposition 5.10 shows that for every symmetric, feasible pseudo-density \( D \in \mathbb{R}^{(0,1)^n} \), \( \frac{1}{2} (D + \bar{D}) \) is a double-symmetric, feasible pseudo-density. Moreover, by the previous corollary it is the Grigoriev pseudo-density, which implies the first part of the statement. The definition of \( \bar{D} \) (cf. Definition 5.7) and the representation of \( D_\varrho \), obtained in (4.88), yield the second part of the statement. \( \square \)

Since the double-symmetric pseudo-densities are the ones with minimal norm and minimal degree by Corollary 5.11, Corollary 5.13 immediately implies the following statement which was our goal of this section. Although, it is rather a corollary, we state it as a proposition because of its significance.

Proposition 5.15.
Let \( n \in \mathbb{N} \) be odd with \( n \geq 3 \), \( r = \frac{1}{2} \), \( \delta = 0 \), and let \( D_\varrho \in \mathbb{R}^{(0,1)^n} \) be the Grigoriev pseudo-density. Then, there is no feasible pseudo-density \( D \in \mathbb{R}^{(0,1)^n} \) with smaller infinity norm or smaller degree than the one of the Grigoriev pseudo-density. Or equivalent, for every feasible pseudo-density \( D \in \mathbb{R}^{(0,1)^n} \) it holds
\[
\|D_\varrho\|_\infty \leq \|D\|_\infty \quad \text{and} \quad \deg(D_\varrho) \leq \deg(D).
\]
In particular, it holds
\[
0.773\sqrt{n} \leq \|D\|_\infty.
\]

5.3 Two conjectures about double symmetric pseudo-densities in the centered non-standard case

In this section, we stay in the centered case, that means we keep \( r = \frac{1}{2} \), but we want to know if there is a feasible pseudo-density with smaller infinity norm than the Grigoriev pseudo-density by taking an arbitrary \( \delta \in \left(0,\frac{1}{4}\right) \) but in expense of a smaller \( \varepsilon \) than in the standard case. By Corollary 5.11 it is again sufficient to look for double-symmetric, feasible pseudo-densities.
As in the previous section, for the analysis it is an advantage to consider the corresponding linear forms. For fixed $n \in \mathbb{N}$ and $\delta \in (0, \frac{1}{4})$, by previous results and observations a linear form $L : \mathbb{R}[T]_n \rightarrow \mathbb{R}$ corresponding to a double-symmetric, feasible pseudo-density satisfies the conditions (I)-(IV), where condition (III) translates into

$$L \left( (T - \frac{n}{2})^2 \right) = \delta$$

(III*)

in this case.

**Proposition 5.16.**

Let $n \in \mathbb{N}$ be odd with $n \geq 3$ and $\delta \in (0, \frac{1}{4})$. Furthermore, let $L : \mathbb{R}[T]_n \rightarrow \mathbb{R}$ be a linear form satisfying the conditions (I), (II), (III*), and (IV). Then, we have

$$L(T) = \frac{n}{2},$$

$$L(T^2) = \left( \frac{n}{2} \right)^2 + \delta \text{ and}$$

$$L(T^3) = \left( \frac{n}{2} \right)^3 + 3\frac{n}{2}\delta.$$

Furthermore, for $m \geq 5$ it additionally holds

$$L(T^4) \leq \frac{\left( \frac{n}{2} \right)^6 + 6 \left( \frac{n}{2} \right)^4 \delta - 7 \left( \frac{n}{2} \right)^2 \delta^2}{\left( \frac{n}{2} \right)^2 - \delta} \text{ and}$$

$$L(T^4) \geq \left( \frac{n}{2} \right)^4 + 6 \left( \frac{n}{2} \right)^2 \delta + \delta^2.$$

**Proof.** Condition (IV) yields $L(T) = L(n - T)$, which is equivalent to $L(T) = n - L(T)$. This implies

$$L(T) = \frac{n}{2}.$$

By condition (III*), we have $L \left( (T - \frac{n}{2})^2 \right) = \delta$. This implies $L(T^2) - nL(T) + \left( \frac{n}{2} \right)^2 = \delta$. Substituting $L(T) = \frac{n}{2}$ into this equation and solving it yields

$$L(T^2) = \left( \frac{n}{2} \right)^2 + \delta.$$

Again condition (IV) implies

$$L(T^3) = L((n - T)^3) = n^3 - 3n^2L(T) + 3nL(T^2) - L(T^3).$$

Substituting the already computed values for $L(T)$ and $L(T^2)$ into this equation

$$L(T^3) = \left( \frac{n}{2} \right)^3 + 3\frac{n}{2}\delta,$$
Chapter 5. A detailed analysis of the occurring pseudo-densities

implies

\[ L(T^3) = \frac{n^3 - 3n^2 \left(\frac{n}{2}\right)^2 + 3n \left(\frac{n}{2}\right)^2 + \delta}{2} = \left(\frac{n}{2}\right)^3 + \frac{3n^2}{2}. \]

Now, let \( n \geq 5 \). Then, condition (II) implies \( L \left((a + bT + cT^2)^2\right) \geq 0 \) for all \( a, b, c \in \mathbb{R} \). Then, Lemma 1.31 implies

\[
\begin{pmatrix}
1 & L(T) & L(T^2) \\
L(T) & L(T^2) & L(T^3) \\
L(T^2) & L(T^3) & L(T^4)
\end{pmatrix} \succeq 0,
\]

and putting in the already computed values yields

\[
\begin{pmatrix}
1 & \left(\frac{n}{2}\right)^2 + \delta & \left(\frac{n}{2}\right)^3 + \frac{3n^2}{2} \\
\left(\frac{n}{2}\right)^2 + \delta & \left(\frac{n}{2}\right)^3 + \frac{3n^2}{2} & L(T^4)
\end{pmatrix} \succeq 0.
\]

By Corollary 1.25, we obtain

\[
L(T^4) \left(\left(\frac{n}{2}\right)^2 + \delta - \left(\frac{n}{2}\right)^2\right) - \left(\left(\frac{n}{2}\right)^3 + \frac{3n^2}{2}\right) \left(3\frac{n^2}{2} - \frac{n}{2} \left(\left(\frac{n}{2}\right)^2 + \delta\right)\right)
+ \left(\frac{n}{2}\right)^2 + \delta \left(\frac{n}{2} \left(\left(\frac{n}{2}\right)^3 + \frac{3n^2}{2}\right) - \left(\frac{n}{2}\right)^2 + \delta\right)^2 \geq 0. \quad (5.26)
\]

After multiplying out and simplifying the inequality, we obtain

\[
\delta L(T^4) - \left(\frac{n}{2}\right)^4 \delta - 6 \left(\frac{n}{2}\right)^2 \delta^2 - \delta^3 \geq 0,
\]

which implies

\[
L(T^4) \geq \left(\frac{n}{2}\right)^4 + 6 \left(\frac{n}{2}\right)^2 \delta + \delta^2. \quad (5.27)
\]

Condition (II) also implies \( L \left(T(n - T)(a + bT)^2\right) \geq 0 \) for all \( a, b \in \mathbb{R} \). Then, again by Lemma 1.31, this implies

\[
\begin{pmatrix}
L(T(n - T)) & L(T(n - T)T) \\
L(T(n - T)T) & L(T(n - T)T^2)
\end{pmatrix} \succeq 0,
\]

or equivalent

\[
\begin{pmatrix}
nL(T) - L(T^2) & nL(T^2) - L(T^3) \\
nL(T^2) - L(T^3) & nL(T^3) - L(T^4)
\end{pmatrix} \succeq 0.
\]
Then, inserting the known values for $L(T), L(T^2)$ and $L(T^3)$ implies
\[
\begin{pmatrix}
n \left( \frac{n}{2} \right)^2 - \left( \frac{n}{2} \right)^2 + \delta \\
n \left( \frac{n}{2} \right)^3 - \left( \left( \frac{n}{2} \right)^3 + 3 \frac{n}{2} \delta \right)
\end{pmatrix}
\begin{pmatrix}
n \left( \frac{n}{2} \right)^2 - \left( \frac{n}{2} \right)^2 + \delta \\
n \left( \frac{n}{2} \right)^3 - \left( \left( \frac{n}{2} \right)^3 + 3 \frac{n}{2} \delta \right)
\end{pmatrix}
\geq 0,
\]
or equivalent
\[
\begin{pmatrix}
\left( \frac{n}{2} \right)^2 - \delta \\
\left( \frac{n}{2} \right)^3 - \frac{n}{2} \delta
\end{pmatrix}
\begin{pmatrix}
\left( \frac{n}{2} \right)^2 - \delta \\
\left( \frac{n}{2} \right)^3 - \frac{n}{2} \delta
\end{pmatrix}
\geq 0.
\]
Applying Corollary 1.25 again yields
\[
\begin{pmatrix}
\left( \frac{n}{2} \right)^2 - \delta \\
\left( \frac{n}{2} \right)^3 - \frac{n}{2} \delta
\end{pmatrix}
\begin{pmatrix}
\left( \frac{n}{2} \right)^2 - \delta \\
\left( \frac{n}{2} \right)^3 - \frac{n}{2} \delta
\end{pmatrix}
\geq 0,
\]
but this is equivalent to
\[
\left( \frac{n}{2} \right)^6 + 6 \left( \frac{n}{2} \right)^4 \delta - 7 \left( \frac{n}{2} \right)^2 \delta^2 - \left( \frac{n}{2} \right)^2 - \delta \right) L(T^4) \geq 0,
\]
which implies
\[
L(T^4) \leq \frac{\left( \frac{n}{2} \right)^6 + 6 \left( \frac{n}{2} \right)^4 \delta - 7 \left( \frac{n}{2} \right)^2 \delta^2}{\left( \frac{n}{2} \right)^2 - \delta}.
\]
(5.28)
The inequalities (5.27) and (5.28) yields the desired bounds for $L(T^4)$.

**Remark 5.17.**

Let $n \in \mathbb{N}$, and let $k \in [n]$ be odd. Furthermore, let $L : \mathbb{R}[T]_n \to \mathbb{R}$ be a linear form satisfying condition (IV). Then, we have
\[
L(T^k) \overset{\text{IV}}{=} L((n-T)^k) = L \left( \sum_{i=0}^{k} \binom{k}{i} n^{k-i} (-T)^i \right)
= \sum_{i=0}^{k} \binom{k}{i} n^{k-i} (-1)^i L(T^i) = \sum_{i=0}^{k-1} \binom{k}{i} n^{k-i} (-1)^i L(T^i) - L(T^k).
\]
This implies
\[
2L(T^k) = \sum_{i=0}^{n-1} \binom{k}{i} n^{k-i} (-1)^i L(T^i),
\]
or equivalent
\[
L(T^k) = \frac{1}{2} \sum_{i=0}^{k-1} \binom{k}{i} n^{k-i} (-1)^i L(T^i).
\]
Note that the condition $k$ odd is necessary because we used $(-1)^k = -1$ in the fourth step.

Therefore, we see that for odd $k$, the value $L(T^k)$ is determined by the values $L(1), L(T), \ldots, L(T^{k-1})$. 
\[\square\]
Combining this remark with the previous proposition yields that a linear form satisfying the conditions of the previous theorem has a degree of freedom only for the values \( L(T^k) \) with even \( 4 \leq k \leq n - 1 \). Therefore, these linear forms build a set of dimension at most \( \frac{n-3}{2} \).

Hence, in contrast to the previous section, it is too complicated to compute all the possible ranges for the values \( L(T^k) \) of such a linear form, similar to \( L(T^4) \) in the previous theorem, because we have too much leeway on possibilities.

But based on the previous proposition and mainly based on numerical experiments, we have some idea concerning optimal linear forms that lead us to the two conjectures.

But first, we need some preparation. Let \( n \in \mathbb{N} \) be odd with \( n \geq 3 \), and let \( \delta \in \left(0, \frac{1}{4}\right] \) be given. Then, we consider the following two optimization problems.

\[
\text{(Opt)} \quad \min_{\delta \in \mathbb{R}, D \in \mathbb{R}^{\{0,1\}^n}} \|D\|_\infty \\
\text{subject to } 0 \leq \delta \leq \frac{1}{4}, \\
\quad \mathbb{E}_x D(x) = 1, \\
\quad \mathbb{E}_x D(x) g(x) \geq 0 \text{ for every } g \in \mathbb{R}^{\{0,1\}^n} \text{ with } \deg_{\text{sos}}(g) \leq n, \\
\quad \mathbb{E}_x D(x) \left(|x| - \frac{n}{2}\right)^2 = \hat{\delta};
\]

and

\[
\text{(Opt}_\delta) \quad \min_{D \in \mathbb{R}^{\{0,1\}^n}} \|D\|_\infty \\
\text{subject to } \mathbb{E}_x D(x) = 1, \\
\quad \mathbb{E}_x D(x) g(x) \geq 0 \text{ for every } g \in \mathbb{R}^{\{0,1\}^n} \text{ with } \deg_{\text{sos}}(g) \leq n, \\
\quad \mathbb{E}_x D(x) \left(|x| - \frac{n}{2}\right)^2 = \delta.
\]

**Remark 5.18.**

Note that the feasible sets of these optimization problems are spectrahedra (cf. Remark 1.23 and Proposition 5.2).

Note that the original objective functions are actually not linear. But by adding an additional variable \( z \), (Opt) respectively (Opt_\delta) is equivalent to

\[
\min_{z \in \mathbb{R}, \delta \in \mathbb{R}, D \in \mathbb{R}^{\{0,1\}^n}} z
\]

respectively

\[
\min_{z \in \mathbb{R}, D \in \mathbb{R}^{\{0,1\}^n}} z
\]
and adding the inequalities

\[-z \leq D(x) \leq z \text{ for every } x \in \{0,1\}^n\]

to the existing constraints. Therefore, we have a linear objective function, and because the additional constraints are also linear, the feasible sets remain spectrahedra, why these optimization problems can be treated via semidefinite programming\textsuperscript{15}.

**Conjecture 5.19.**

Let \( n \in \mathbb{N} \) be odd with \( n \geq 3 \). For a fixed \( \delta \in (0,\frac{1}{4}] \), we consider the optimization problem \((\text{Opt}_\delta)\). Then, the linear form \( L : \mathbb{R}[T]_n \to \mathbb{R} \) corresponding to the optimal pseudo-density is determined by

\[
L(T^k) = \sum_{j=0}^{\lfloor \frac{k}{2} \rfloor} \binom{k}{2j} \left( \frac{n}{2} \right)^{k-2j} \delta^j
\]

for every \( k \in \{0,\ldots,n\} \).

**Conjecture 5.20.**

Let \( n \in \mathbb{N} \) be odd with \( n \geq 3 \). Consider the optimization problem \((\text{Opt})\). Then, the optimum is attained for \( \delta = \frac{1}{4} \).

Sure, \( \delta = \frac{1}{4} \) is not suitable for our feasible pseudo-densities because only \( \delta \in (0,\frac{1}{4}] \) is allowed. However, it is interesting to know how the optimal density looks like in this limit case. It turns out that in this case the pseudo-density has a very nice form and it is easy to compute. Therefore, we are able to compute the infinity norm of it. So, assuming the two conjectures, we have a lower bound for the minimal infinity norm of every feasible degree-\( n \) pseudo density in the centered non-standard case.

**Proposition 5.21.**

Let \( n \in \mathbb{N} \) be odd with \( n \geq 3 \). Let \( L : \mathbb{R}[T]_n \to \mathbb{R} \) be the linear form from Conjecture 5.19 for \( \delta = \frac{1}{4} \), that means

\[
L(T^k) = \sum_{k=0}^{\lfloor \frac{k}{2} \rfloor} \binom{k}{2j} \left( \frac{n}{2} \right)^{k-2j} \left( \frac{1}{4} \right)^j
\]

Then, \( D \in \mathbb{R}^{\{0,1\}^n} \) with

\[
D(x) = \begin{cases} 
\frac{2n-2}{(n-1)} & \text{if } |x| \in \left\{\frac{n-1}{2}, \frac{n+1}{2}\right\} \\
0 & \text{otherwise}
\end{cases}
\]

is the double-symmetric degree-\( n \) pseudo-density corresponding to \( L \). Furthermore, it holds

\[

\|D\|_\infty > 0.574n^{\frac{1}{2}}
\]

\textsuperscript{15}Numerical experiments to these two optimization problems were be done with YALMIP \textsuperscript{L04}. \]
and

\[
\lim_{n \to \infty} \frac{\|D\|_\infty}{n^\frac{1}{2}} = \frac{\sqrt{\pi}}{2\sqrt{2}} \approx 0.627.
\]

**Proof.** Note that $D$ is double-symmetric by definition. In order to show that it is the double-symmetric degree-$n$ pseudo-density corresponding to $L$, we have to show

\[
L(f) = \mathbb{E}_x D(x)f(|x|) \text{ for every } f \in \mathbb{R}[T]_n.
\]

But it remains to show

\[
L(T^k) = \mathbb{E}_x D(x)|x|^k \text{ for every } k \in \{0, \ldots, n\}.
\]

In order to do this, let $k \in \{0, \ldots, n\}$. Then, we obtain

\[
\mathbb{E}_x D(x)|x|^k = \frac{1}{2^n} \left( \sum_{x \in \{0,1\}^n, |x| = \frac{n-1}{2}} D(x)|x|^k + \sum_{x \in \{0,1\}^n, |x| = \frac{n+1}{2}} D(x)|x|^k \right)
\]

\[
= \frac{1}{2^n} \left( \sum_{x \in \{0,1\}^n, |x| = \frac{n-1}{2}} 2^{n-1} \left( \frac{n-1}{2} \right)^k + \sum_{x \in \{0,1\}^n, |x| = \frac{n+1}{2}} 2^{n-1} \left( \frac{n+1}{2} \right)^k \right)
\]

\[
= \frac{1}{2^n} \left( \left( \frac{n-1}{2} \right)^k + 2^{n-1} \left( \frac{n+1}{2} \right)^k \right)
\]

\[
= \frac{1}{2} \left( \left( \frac{n-1}{2} \right)^k + \left( \frac{n+1}{2} \right)^k \right)
\]

\[
= \frac{1}{2} \left( \sum_{i=0}^{k} \binom{k}{i} \left( \frac{n}{2} \right)^{k-i} \left( \frac{-1}{2} \right)^i + \sum_{i=0}^{k} \binom{k}{i} \left( \frac{n}{2} \right)^{k-i} \left( \frac{1}{2} \right)^i \right)
\]

\[
= \frac{1}{2} \left( \sum_{i=0}^{k} \binom{k}{i} \left( \frac{n}{2} \right)^{k-i} \left( \frac{1}{2} \right)^i \left( (-1)^i + 1 \right) \right),
\]

where we used \( \binom{n}{\frac{n}{2}} = \binom{n}{\frac{n+1}{2}} \) in the fourth line, which is valid because of \( n - \frac{n-1}{2} = \frac{n+1}{2} \).

Note, in the last equation, we have \((-1)^i + 1 = 0\) if \(i\) is odd and \((-1)^i + 1 = 2\) if
Chapter 5. A detailed analysis of the occurring pseudo-densities

\[ i \text{ is even. Therefore, the last equation turns into} \]
\[ \mathbb{E} D(x) | x |^k = \frac{1}{2} \left( \sum_{j=0}^{\left\lfloor \frac{k}{2} \right\rfloor} \binom{k}{2j} \left( \frac{n}{2} \right)^{k-2j} \left( \frac{1}{4} \right)^{2j} \right) \]
\[ = \left( \sum_{j=0}^{\left\lfloor \frac{k}{2} \right\rfloor} \binom{k}{2j} \left( \frac{n}{2} \right)^{k-2j} \left( \frac{1}{4} \right)^{j} \right) \]
\[ = L(T^k), \]

what shows that \( D \) is actually the corresponding symmetric degree-\( n \) pseudo-density to our optimal \( L \).

Now, we want to lower bound the infinity norm of \( D \). By its definition,
\[ \| D \|_\infty = \frac{2^{n-1}}{\left( \frac{n+1}{2} \right)} \]
is obvious.

The Grigoriev pseudo-density \( D_\varphi \in \mathbb{R}^{\{0,1\}^n} \) satisfies \( \| D_\varphi \|_\infty = \frac{n}{2^{n-\tau}} \left( \frac{m-1}{m-1/2} \right) \), which is equivalent to
\[ \frac{2^{n-1}}{\left( \frac{n+1}{2} \right)} = \| D_\varphi \|_\infty. \tag{5.29} \]

In order to compare the two norms, we transform the first equation. We obtain,
\[ \| D \|_\infty = \frac{2^{n-1}}{\left( \frac{n+1}{2} \right)} = \frac{2^{n-1}}{\left( \frac{n+1}{2} \right)} \cdot \frac{n+1}{2n} \]
\[ = \frac{n+1}{2 \| D_\varphi \|_\infty}. \]

Now, we can use our results and bounds for \( \| D_\varphi \|_\infty \) computed in Section 4.6. Then, we obtain
\[ \| D \|_\infty = \frac{n+1}{2 \| D_\varphi \|_\infty} \]
\[ \geq \frac{n}{2 \| D_\varphi \|_\infty} \quad \text{Prop \ref{prop:bound}} \]
\[ \geq \frac{n}{2 \cdot 0.87n^{3/2}} \]
\[ > 0.574n^{1/2}. \]

Moreover, we have
\[ \lim_{n \to \infty} \frac{\| D \|_\infty}{n^{1/2}} = \frac{1}{2} \lim_{n \to \infty} \frac{\sqrt{n}}{n^{1/2}} = \frac{1}{2} \cdot 0.87 \]
\[ \approx \frac{1}{2} \cdot \frac{\sqrt{\pi}}{\sqrt{2}}. \]

\[ \square \]
Remark 5.22.
Note that in this limit case $\delta = \frac{1}{4}$, the optimal degree-$n$ pseudo-density is an actual density.

Corollary 5.23.
Let $n \in \mathbb{N}$ be odd with $n \geq 3$, $r = \frac{n}{2}$ and $\delta \in \left[0, \frac{1}{4}\right)$. Assuming Conjectures 5.19 and 5.20, we obtain

$$\|D\|_{\infty} > 0.574n^{\frac{1}{2}}$$

and

$$\lim_{n \to \infty} \frac{\|D\|_{\infty}}{n^{\frac{1}{2}}} \geq \frac{\sqrt{\pi}}{2\sqrt{2}} \approx 0.627$$

for every feasible pseudo-density $D \in \mathbb{R}^{(0,1)^n}$.

Proof. Let $n \in \mathbb{N}$, and let $D^* \in \mathbb{R}^{(0,1)^n}$ be the pseudo-density from Proposition 5.21. By Conjectures 5.19 and Proposition 5.21, $D^*$ is a feasible pseudo-density with minimal infinity norm in the case $\delta = \frac{1}{4}$. Now, fix an arbitrary $\delta \in (0, \frac{1}{4})$, and let $D \in \mathbb{R}^{(0,1)^n}$ be an arbitrary feasible pseudo-density. Then, Conjecture 5.20 implies

$$\|D\|_{\infty} \geq \|D^*\|_{\infty}.$$ 

Now, again applying Proposition 5.21 yields the desired statements

$$\|D\|_{\infty} \geq \|D^*\|_{\infty}^{\text{Prop. 5.21}} > 0.574n^{\frac{1}{2}}$$

and

$$\lim_{n \to \infty} \frac{\|D\|_{\infty}}{n^{\frac{1}{2}}} \geq \lim_{n \to \infty} \frac{\|D^*\|_{\infty}}{n^{\frac{1}{2}}}^{\text{Prop. 5.21}} \frac{\sqrt{\pi}}{2\sqrt{2}} \approx 0.627.$$

Remark 5.24.
(a) With the previous result, we see that by choosing $r = \frac{n}{2}$ and $\delta \neq 0$, we cannot beat the upper bound on the infinity norm of the Grigoriev pseudo-density in the size of $n$. Indeed, the previous result shows that there possibly exist some $\delta \in (0, \frac{1}{4})$ and a feasible pseudo-density with slightly smaller norm than the Grigoriev pseudo-density, but only the exponent of $n$ and not the prefixed constant is relevant for the bound of $\text{rk}_{\text{psd}}(\text{CORR}_n)$ as we saw in the proof of Theorem 4.56.

(b) Additionally, note that if we are really interested in the more detailed bound and not only up to a constant, we have to respect that compared to the standard case in the previous section, we could enable the potentially better bound on the infinity norm of a feasible pseudo-density only at the expense of a worse $\varepsilon$, but this also effects the bound on $\text{rk}_{\text{psd}}(\text{CORR}_n)$. 


So, in order to compare both cases, we have to consider the quotient $\|D\|_\infty^\varepsilon$.

In the standard case, where the Grigoriev pseudo-density is optimal, we have $\varepsilon = \frac{1}{4}$ and $\|D\|_\infty^\text{Prop 5.49} \approx 0.798n^{\frac{1}{2}}$, where we used the computed limiting value for the infinity norm what makes more sense in our opinion. Therefore, we have

$$\varepsilon \|D\|_\infty \approx \frac{1}{4} \cdot 0.798n^{\frac{1}{2}}.$$

In the non-standard case, we can choose some $\delta \in (0, \frac{1}{4})$, and we obtain $\varepsilon = \frac{1}{4} - \delta$. Now, let $D \in \mathbb{R}^{\{0,1\}^n}$ be a feasible pseudo-density for some fixed $\delta$, and let $C > 0$ with $\|D\|_\infty = Cn^{\frac{2}{3}}$. Note that assuming our two conjectures and because of the previous result, we have $C \geq 0.627$ if we also use the limiting value. Therefore, we have

$$\varepsilon \|D\|_\infty = \frac{1}{4} - \delta \frac{Cn^{\frac{2}{3}}}{4 \cdot 0.798n^{\frac{1}{2}}}.$$

So, having a larger quotient in the non-standard case is only possible if we have

$$\frac{1}{4} - \delta \frac{Cn^{\frac{2}{3}}}{4 \cdot 0.798n^{\frac{1}{2}}} \geq \frac{1}{4} \cdot 0.798n^{\frac{1}{2}}.$$

But this implies

$$\delta \leq \frac{1}{4} \left(1 - \frac{C}{0.798}\right).$$

Thus, even if we take the optimal and theoretically possible $C = 0.627$, this implies that a pseudo-density with this small norm should be feasible in the case $\delta \leq 0.054$

But this seems very unlikely to us because the minimal infinity norm is reached in the limit case $\delta = \frac{1}{4}$ if our Conjecture 5.20 is true.

**Remark 5.25.**

In their work [LPWY16], Troy Lee et al. discovered independently from us a result similar to our results Proposition 5.15 and Corollary 5.23. They also show that $\frac{1}{2}$ is the best possible value that we can obtain for the exponent of $n$ in the lower bound of $\|D\|_\infty$ for an arbitrary feasible pseudo-density $D \in \mathbb{R}^{\{0,1\}^n}$ in the centered case.

But their approach is very different from ours. They work with the notion of the so-called approximate sos-degree and obtain the optimality result on the infinity norm of a feasible pseudo-density en route on their result about an upper bound on the approximate degree of $\left(\sum_{i \in [n]} \chi(i) - \frac{n}{2}\right)^2 - \frac{1}{4}$ for odd $n \in \mathbb{N}$. But in some sense, the approximate sos-degree can be seen as the dual version to pseudo-densities.

The power of their result lies in the fact that it is independent from the correctness of any conjecture what is not the case in our Corollary 5.23.

But we have a completely different focus. While they do not say anything concrete about feasible pseudo-densities or optimality in connection with symmetry, these
symmetric pseudo-densities are the main focus of our contribution. So, we think the most advantage of our work on the one hand lies in the previous section, for example the nice characterization of symmetric pseudo-densities or the result about the uniqueness of the Grigoriev pseudo-density as a double-symmetric pseudo-density, but also in presenting some concrete pseudo-densities that are optimal in some sense (cf. Proposition 5.15 and 5.21).

Another benefit of our results is that they are very comprehensible, and their proofs are simple and elementary. They are not directly build on other work and can be derived only by applying the two nice results Lemma 1.31 and Proposition 1.32, just relying on linear algebra, Blekherman’s Theorem 3.37, and some easy calculation.

5.4 A short view on the non-centered case

In the last two sections, we only considered the centered case \( r = \frac{n}{2} \), but of course, we are not fixed to this choice. Even if we are interested in the largest possible value for \( \varepsilon \) as in (5.1), that is obtained in the case \( r = \frac{n}{2} \), our observations from the beginning of this chapter shows that there are also other possibilities for the choice of \( r \).

But it turns out that computations in the non-centered case \( r \neq \frac{n}{2} \) are more complicated, and additionally we cannot use Proposition 5.10, and therefore we have no double-symmetric pseudo-densities. This again means that we cannot use the nice property (IV) for the corresponding linear forms. In order to use at least Proposition 1.32, which was of great help in the proof of the important Theorem 5.5 we will only consider the standard case \( \delta = 0 \) in this short view on the non-centered case \( r \neq \frac{n}{2} \).

Although computations are more complicated as already mentioned, and although there is an absence of some kind of symmetry compared to the centered case, however we are able to prove an analogon of Theorem 5.5 that is anyway the main theorem of this chapter. One of the crucial reasons for that is in addition to the strength of Theorem 5.5 the high variability of Corollary 3.52. But compared to Theorem 5.5 the computations are more elaborate now.

**Theorem 5.26.**

Let \( n \in \mathbb{N} \) with \( n \geq 2 \), and let \( L : \mathbb{R}[T]_n \to \mathbb{R} \) be a linear form satisfying the conditions (I) - (III), where we set \( \delta = 0 \) in condition (III). Then, we have

\[
L(T^k) = r^k
\]

for all \( k \in \{0, \ldots, n - 1\} \) if \( n \) is odd, respectively for all \( k \in \{0, \ldots, n\} \) if \( n \) is even.

**Proof.** Analog to Theorem 5.5 we prove the theorem by induction over \( k \).
**Base case:** The base case will include all \( k \in \{0, \ldots, \lceil \frac{n+1}{2} \rceil \}.\) The conditions (II) and (III) together with Proposition 1.32 imply
\[
L \left( (T - r) T^k \right) = 0,
\]
for every \( k \in \{0, \ldots, \lceil \frac{n}{2} \rceil \}.\) This is equivalent to
\[
L(T^{k+1}) = rL(T^k),
\]
(5.31)
for every \( k \in \{0, \ldots, \lceil \frac{n}{2} \rceil \}.\)

For \( k = 0, \) the theorem is true because of \( L(1) = 1.\) Putting this into (5.31) implies \( L(T) = r.\) Again putting this into (5.31) implies \( L(T^2) = r^2.\) We can continue in this way up to \( k = \lceil \frac{n}{2} \rceil.\) Overall, we obtain
\[
L(T^{k+1}) = r^{k+1}
\]
for every \( k \in \{0, \ldots, \lceil \frac{n}{2} \rceil \},\) or equivalent
\[
L(T^k) = r^k
\]
(5.32)
for every \( k \in \{0, \ldots, \lceil \frac{n+1}{2} \rceil \}.\)

Note that in the case \( n \leq 3 \) we are already done. Therefore, without loss of generality let \( n \geq 4 \) from now on.

**Induction step:** Let \( k \in \{ \frac{n+3}{2}, \ldots, 2\lceil \frac{n}{2} \rceil \} \) be even and \( k \geq 6.\) We assume that the statement of our theorem is valid for all \( j \in \{k - 5, \ldots, k - 2\},\) that means
\[
L(T^j) = r^j.
\]
(5.33)
Now, we simultaneously show that this also holds for \( k - 1 \) and \( k.\) In this context, note that we have
\[
2 \left\lfloor \frac{n}{2} \right\rfloor = \begin{cases} 
  n & \text{n even} \\
  n - 1 & \text{n odd} 
\end{cases}
\]
Condition (II) implies \( L \left( (aT^{k-1} + bT^k)^2 \right) \geq 0 \) for all \( a,b \in \mathbb{R}.\) Then, Lemma 1.31 implies
\[
\begin{pmatrix} L(T^{k-2}) & L(T^{k-1}) \\ L(T^{k-1}) & L(T^k) \end{pmatrix} \succeq 0,
\]
and the induction hypothesis (5.33) yields
\[
\begin{pmatrix} r^{k-2} & L(T^{k-1}) \\ L(T^{k-1}) & L(T^k) \end{pmatrix} \succeq 0.
\]
By Corollary \[\text{Corollary 1.25}\] we obtain
\[r^{k-2}y - x^2 \geq 0,
\]
which implies
\[y \geq \frac{x^2}{r^{k-2}}. \tag{5.34}\]

Again by condition \[\text{(II)}\], we obtain
\[L\left(T (n - T) \left(aT^\frac{k}{2} - bT^\frac{k}{2} - 1\right)^2\right) \geq 0 \text{ for all } a, b \in \mathbb{R}.
\]
Then, Lemma \[\text{Lemma 1.31}\] implies
\[
\begin{pmatrix}
L(T (n - T) T^{k-4}) & L(T (n - T) T^{k-3}) \\
L(T (n - T) T^{k-3}) & L(T (n - T) T^{k-2})
\end{pmatrix} \succeq 0,
\]
or equivalent
\[
\begin{pmatrix}
nL(T^{k-3}) - L(T^{k-2}) & nL(T^{k-2}) - L(T^{k-1}) \\
nL(T^{k-2}) - L(T^{k-1}) & nL(T^{k-1}) - L(T^k)
\end{pmatrix} \succeq 0.
\]

Then, the induction hypothesis \[\text{(5.9)}\] yields
\[
\begin{pmatrix}
nr^{k-3} - r^{k-2} & nr^{k-2} - L(T^{k-1}) \\
r^{k-2} - L(T^{k-1}) & nr^{k-2} - L(T^k)
\end{pmatrix} \succeq 0.
\]

By Corollary \[\text{Corollary 1.25}\] we obtain
\[(nx - y)(nr^{k-3} - r^{k-2}) - (nr^{k-2} - x)^2 \geq 0,
\]
which implies
\[y \leq \frac{-x^2 + (nr^{k-2} + n^2r^{k-3})x - n^2r^{2k-4}}{r^{k-3}(n - r)}. \tag{5.35}\]

Applying condition \[\text{(II)}\] a third time implies
\[L\left(T (n - T) (T - 1) (n - 1 - T) \left(aT^\frac{k}{2} - bT^\frac{k}{2} - 2\right)^2\right) \geq 0
\]
for all \(a, b \in \mathbb{R}\). We define \(p := T (n - T) (T - 1) (n - 1 - T)\). Then, analogue as above, Lemma \[\text{Lemma 1.31}\] implies
\[
\begin{pmatrix}
L(p \cdot T^{k-6}) & L(p \cdot T^{k-5}) \\
L(p \cdot T^{k-5}) & L(p \cdot T^{k-4})
\end{pmatrix} \succeq 0.
\]

We multiply out the inner of \(L\) and use the linearity of \(L\) in order to obtain
\[
\begin{pmatrix}
a_{11} & a_{12} \\
a_{12} & a_{22}
\end{pmatrix} \succeq 0, \tag{5.36}
\]
where
\[ a_{11} := L(T^{k-2}) - 2nL(T^{k-3}) + (n^2 + n - 1)L(T^{k-4}) - (n^2 - n)L(T^{k-5}), \]
\[ a_{12} := L(T^{k-1}) - 2nL(T^{k-2}) + (n^2 + n - 1)L(T^{k-3}) - (n^2 - n)L(T^{k-4}) \]
and
\[ a_{22} := L(T^k) - 2nL(T^{k-1}) + (n^2 + n - 1)L(T^{k-2}) - (n^2 - n)L(T^{k-3}). \]

By applying the induction hypothesis (5.33) and simplifying the terms, we obtain
\[ a_{11} = r^k - (r - 1) \left( n^2 - 2nr - n + r^2 + r \right), \]
\[ a_{12} = x + r^{k-4} \left( -2nr^2 + (n^2 + n - 1)r - (n^2 - n) \right) \]
and
\[ a_{22} = y - 2nx + r^{k-3} \left( (n^2 + n - 1)r - (n^2 - n) \right). \]

Applying Corollary 1.25 to (5.36) yields
\[ a_{11}a_{22} - a_{12}^2 \geq 0. \]

After multiplying out, simplifying and solving the inequality for \( y \), we obtain
\[ y \geq \frac{x^2 - \alpha x + \beta}{r^{k-5}(r - 1)(n^2 - 2nr - n + r^2 + r)}, \]
(5.37)

where
\[ \alpha = 2r^{k-5}(-n^3r + n^3 + n^2r^2 - n^2 + nr^3 - nr^2 + r^2) \]
and
\[ \beta = r^{2k-6}(-2n^3r + 2n^3 + 3n^2r^2 - n^2r - 2n^2 - nr^2 + nr + r^2). \]

The three inequalities (5.34), (5.35) and (5.37) build a system of inequalities with the two unknowns \( y \) and \( x \).

Combining (5.34) and (5.35) yields
\[ \frac{x^2}{r^{k-2}} \leq \frac{x^2 + (nr^{k-2} + n^2r^{k-3})x - n^2r^{2k-4}}{r^{k-3}(n-r)}. \]

This is equivalent to
\[ (n - r)x^2 \leq -rx^2 + (nr^{k-1} + n^2r^{k-2})x - n^2r^{2k-3}, \]
which again is equivalent to
\[ x^2 - (r^{k-1} + nr^{k-2})x + nr^{2k-3} \leq 0. \]

Then, Vieta’s formulas imply
\[ (x - r^{k-1})(x - nr^{k-2}) \leq 0. \]
Because of $r < n$, it follows

$$r^{k-1} \leq x \leq nr^{k-2}.$$  \hspace{1cm} (5.38)

Combining the inequalities (5.35) and (5.37) yields

$$\frac{x^2 - ax + \beta}{r^{k-5}(r - 1)(n^2 - 2nr - n + r^2 + r)} \leq y \leq \frac{-x^2 + (nr^{k-2} + n^2r^{k-3})x - n^2r^{2k-4}}{r^{k-3}(n - r)}.$$  \hspace{1cm} (5.37)

By multiplying with $r^{k-3}(n - r)(r - 1)(n^2 - 2nr - n + r^2 + r)$, summarization of terms and some easy computation afterwards, we obtain

$$ax^2 + bx + c \leq 0,$$

where

$$a := nr - n + 1,$$

$$b := r^{k-3}(n^3r - n^3 - 2n^2r^2 + n^2r + n^2 - nr^3 + 2nr^2 - nr - 2r^2),$$

and

$$c := r^{2k-4}(-n^3r + n^3 + 2n^2r^2 - n^2r - n^2 - nr^2 + nr + r^2).$$

By Vieta’s formulas, the latter inequality is equivalent to

$$a(x - x_1)(x - x_2) \leq 0,$$  \hspace{1cm} (5.39)

where

$$x_1 = r^{k-1} \quad \text{and} \quad x_2 = r^{k-3} \frac{-n^3r + n^3 + 2n^2r^2 - n^2r - n^2 - nr^2 + nr + r^2}{nr - n + 1}.$$  \hspace{1cm} (5.40)

By the choose of $r$, it follows $x_1 \geq x_2$, and therefore (5.39) implies

$$x_2 \leq x \leq x_1 = r^{k-1}.$$  \hspace{1cm} (5.40)

Combining (5.38) and (5.40) implies

$$L(T^{k-1}) = x = r^{k-1}.$$  \hspace{1cm} (5.41)

Inserting this into (5.34) and (5.35) yields

$$y \geq r^k$$  \hspace{1cm} (5.42)

and

$$y \geq \frac{-r^{2k-2} + nr^{2k-3} + n^2r^{2k-4} - n^2r^{2k-4}}{r^{k-3}(n - r)} = \frac{r^{2k-3}(n - r)}{r^{k-3}(n - r)} = r^k.$$  \hspace{1cm} (5.43)
The inequalities (5.42) and (5.43) imply
\[ L(T^k) = y = r^k. \]
Together with (5.41), this yields our desired statement.

Analog to Theorem 5.5, a careful look on the proof shows that for the induction step an even \( k \geq 6 \) is needed. But the case \( k \leq 4 \) is covered by the base case whenever we have \( 4 \leq \lceil \frac{n+1}{2} \rceil \). But this is always the case except for \( n \in \{4,5\} \) because we assume \( n \geq 4 \). So, we still have to proof these cases separately.

In order to do this, note that the base case includes every \( k \in \{0, \ldots , 3\} \) by (5.32). In particular, we have
\[ L(T^3) = r^3. \tag{5.44} \]
So, it remains to show
\[ L(T^4) = r^4. \]

We have
\[ L(T^4) \leq \frac{L(T^3)^2}{r^2} \tag{5.44}, \]
and
\[ L(T^4) \leq -L(T^3)^2 + 6r^3L(T^3) - 4r^6 \tag{5.44}, \]
which implies
\[ L(T^4) = r^4 \]
as desired.

As we see, this theorem yields not only a generalization of Theorem 5.5 for even \( n \) it actually yields to a stronger statement.

With our observations from the end of Section 5.1, we see that for even \( n \), the Grigoriev pseudo-density is the only symmetric feasible pseudo-density in the non-centered standard case. Therefore, Proposition 5.4 immediately implies the following consequence.

**Corollary 5.27.**
Let \( n \in \mathbb{N} \) be even, \( r \in \left( \lfloor \frac{n-3}{2} \rfloor, \lceil \frac{n+3}{2} \rceil \right) \) with \( r \notin \mathbb{N} \) and \( \delta = 0 \). Furthermore, let \( D_{\varphi,r} \in \mathbb{R}_{\{0,1\}^n} \) be the Grigoriev pseudo-density. Then, for every feasible pseudo-density \( D \in \mathbb{R}_{\{0,1\}^n} \) it holds
\[ \|D_{\varphi,r}\|_\infty \leq \|D\|_\infty. \]

As already mentioned, in the non-centered case we cannot expect any double-symmetric pseudo-densities, but for those \( r \) who yield the best possible value for \( \varepsilon \), however we are able to show another kind of symmetry for the Grigoriev pseudo-density. Actually, we prove a generalization of Lemma 5.9.
Lemma 5.28.
Let \( n \in \mathbb{N} \) and let \( r \in \left\{ \frac{n-2}{2}, \frac{n}{2}, \frac{n+2}{2} \right\} \) if \( n \) is odd, respectively \( r \in \left\{ \frac{n-1}{2}, \frac{n+1}{2} \right\} \) if \( n \) is even. Furthermore, let \( D_{g,r} \in \mathbb{R}^{(0,1)} \) be the Grigoriev pseudo-density and \( x \in \{0,1\}^n \). Then, we have
\[
D_{g,r}(x) = D_{g,r}(y)
\]
for every \( y \in \{0,1\}^n \) with
\[
|y| = 2r - |x|.
\]

Proof. Note that in the case \( n \) odd and \( r = \frac{n}{2} \), the statement is exactly the one from Lemma 5.9. We extend the proof from this lemma to the remaining cases, where we exemplarily show the case \( n \) even and \( r = \frac{n-1}{2} \). The other cases run analog.

Let \( y \in \{0,1\}^n \) with \( |y| = 2r - |x| = n - 1 - |x| \). Therefore, we can assume \(|x| \neq n\). Then, we have
\[
D_{g,r}(y) \overset{\text{Cor}3.59}{=} \frac{2^n}{\binom{n}{|y|}} \cdot \frac{\Pi_{j=0}^{n-1} \left( \frac{n-1}{2} - j \right)}{\Pi_{j \neq |y|}^{n-1} \left( \frac{n-1}{2} - j \right)} = \frac{2^n}{\binom{n}{|x|+1}} \cdot \frac{\Pi_{i=0}^{n-1} \left( \frac{n-1}{2} - i \right)}{\Pi_{i \neq |x|}^{n-1} \left( \frac{n-1}{2} - i \right)} = \frac{2^n}{\binom{n}{|x|}} \cdot \frac{\Pi_{i=0}^{n-1} \left( \frac{n-1}{2} - i \right)}{\Pi_{i \neq |x|}^{n-1} \left( |x| - i \right)} = \frac{2^n}{\binom{n}{|x|}} \cdot \frac{\Pi_{i=0}^{n-1} \left( \frac{n-1}{2} - i \right)}{\Pi_{i \neq |x|}^{n-1} \left( |x| - i \right)} \overset{\text{Cor}3.59}{=} D_{g}(x).
\]

The above result is besides to Theorem 5.26 another indication of the special significance of the Grigoriev pseudo-densities also in the non-centered case.

However, we think that in order to find feasible pseudo-densities with small infinity norm, it is better to consider the centered case \( r = \frac{1}{2} \) and the double-symmetric Grigoriev pseudo-densities occurring there. Based on some computations of the infinity norm of the Grigoriev pseudo-densities (cf. Corollary 3.59) for several values of \( n \) and \( r \), we finish this chapter with the following conjecture.

Conjecture 5.29.
Let \( n \in \mathbb{N} \) be odd with \( n \geq 3 \). Let \( r \in \left( \left\lfloor \frac{n-3}{2} \right\rfloor, \left\lceil \frac{n+3}{2} \right\rceil \right) \) and \( D_{g,r} \in \mathbb{R}^{(0,1)} \) the Grigoriev pseudo-density. Furthermore, let \( D_{g} = D_{g,r}^{\frac{1}{2}} \in \mathbb{R}^{(0,1)} \) the Grigoriev pseudo-density occurring in the centered case. Then, it holds
\[
\|D_{g}\|_{\infty} \leq \|D_{g,r}\|_{\infty}.
\]
6 Outlook and open questions

Our work can be divided into three main parts. The new proof of Grigoriev’s Theorem and its applications in Chapter 2, the proof of \([\text{LRS15}]\) of a lower bound on \(\text{rk}_{\text{psd}}(\text{CORR}_n)\) in Chapter 4, and in the previous chapter, the analysis of pseudo-densities suitable for proving this bound. Of course, for all these three topics there are still some open problems respectively some interesting questions that can be asked.

We finish this work with an outlook on some of these problems and questions, where we will treat problems from all three parts.

In this context, the main question in our opinion is if there exists a more natural and more algebraic proof of Theorem 4.39, the matrix approximation theorem and perhaps the most essential theorem in the proof of \([\text{LRS15}]\) (cf. Conjecture 4.40). We focus on this topic in the first section.

In the second section, we formulate some interesting open questions concerning the two other parts of the work.

6.1 Matrix approximation via degree truncation

As just mentioned, we are interested in an alternative approach to Theorem 4.39 or to be more precise, a proof that can be done without quantum information theory. Sure, that does not make the proof automatically more simple and shorter but more algebraic, and we think our approach is much more natural.

We state the conjecture not only because we think it would be natural that it is true, it is additionally based on numerical experiments. For this purpose, we transform the conjecture from the setting \(\{0,1\}^n\) into the setting \(\{-1,1\}^n\). But there are not only numerical reasons for this transformation. In the case \(\{-1,1\}^n\), we can use the nice properties of the Fourier analysis and in particular Proposition 3.7 that simplifies computation and therefore perhaps also a possible proof if it exists.

Conjecture 6.1.
Let \(n,r \in \mathbb{N}, \varepsilon > 0\), and let \(F : \{-1,1\}^n \to \mathbb{S}^r\) and \(P : \{-1,1\}^n \to \mathbb{S}^r\) be matrix-valued Boolean functions with \(F \neq 0\). Furthermore, let \(\tau := \mathbb{E}_x \text{Tr}(P(x)^2)\), and let

\[
P = \sum_{I \subseteq [n]} \hat{P}(I) \chi_I
\]
be the Fourier expansion of \( P \). Furthermore, let
\[
k := \deg F \left( \frac{0.6e^2}{\varepsilon r} - \mathbb{E} \frac{\tau}{x} P(x)^2 \ln \left( \frac{\tau P(x)^2}{\varepsilon} \right) + \ln \left( \frac{\sqrt{12}}{\varepsilon} \right) + 0.5 \right),
\]
and let
\[
B := \sum_{I \subseteq [n]} \hat{P}(I) \chi_I
\]
be the degree-\( \lfloor k \rfloor \) truncation of \( P \) and \( \gamma := \mathbb{E} \frac{\tau}{x} \text{Tr} (B(x)^2) \). Then, we have
\[
\frac{\gamma}{\tau} \mathbb{E} \frac{\tau}{x} \text{Tr} \left( F(x) P(x)^2 \right) - \mathbb{E} \frac{\tau}{x} \text{Tr} \left( F(x) B(x)^2 \right) \geq -\varepsilon \gamma \max_{x \in \{0,1\}^n} \| F(x) \|.
\]

Note that this conjecture is equivalent to Conjecture 4.40, what can be easily shown similar to the translation in Theorem 4.24.

Now, we roughly introduce our numerical experiments.

For them, we used MuPAD, a computer algebra system contained in the symbolic math toolbox of MATLAB. Since we wanted to have preferably general examples, the goal was to generated preferably random functions \( F \) and \( P \). Therefore, we used that every matrix \( M : \{-1,1\}^r \to \mathbb{S}^r \) is uniquely determined by the \( r^2 + r \) many Boolean functions
\[
M_{i,j} : \{-1,1\}^n \to \mathbb{R}, \ x \mapsto M(x)_{i,j},
\]
for every \( i,j \in [r] \) together with \( M_{i,j} = M_{j,i} \) what have to be respected. Thus, in order to generate a random matrix-valued Boolean function, it is enough to generate \( r^2 + r \) many Boolean functions. These, in turn, are determined by its Fourier coefficients. That means, in order to generate a random Boolean function \( f \in \mathbb{R}^{\{-1,1\}^n} \), it is enough to generate \( 2^n \) many real numbers \( \hat{f}(I) \) for every \( I \subseteq [n] \) and set \( f = \sum_{I \subseteq [n]} \hat{f}(I) \chi_I \). The random numbers \( \hat{f}(I) \) were chosen by the internal function \text{stats::uniformRandom}(-1,1), that chooses uniformly distributed random numbers in the interval \((-1,1)\). Note that with \( P \), also \( Q \) is determined via \( Q(x) = P(x)^2 \). Hereby, it is also ensured that \( Q \) maps onto psd matrices.

In order to compute \( \frac{\tau}{\varepsilon} Q(x) \ln \left( \frac{\tau}{\varepsilon} Q(x) \right) \) for the degree bound, we directly computed \( \ln \left( \frac{\tau}{\varepsilon} Q(x) \right) \) with the internal \text{numeric::fn(ln)} function and matrix multiplication if \( \frac{\tau}{\varepsilon} Q(x) \) is strictly positive definite. For the case that it is only psd, we wrote a function that computes \( \frac{\tau}{\varepsilon} Q(x) \ln \left( \frac{\tau}{\varepsilon} Q(x) \right) \) according to Definition 4.27 with the help of some internal functions from the library \text{linalg}, that for example computes eigenvectors of a matrix. This is also used for computing \( \| F(x) \| \).

The rest of our program, for example constructing the monomials, constructing \( \{0,1\}^n \) or evaluating the functions \( F,P,Q \) can be done with some simple programming tools as for example loops, together with some standard commands in MuPAD.
But note that there are several steps that require a large running time, for example constructing \(\{0,1\}^n\) or the \(2^n\) many monomials \((\chi_I)_{I \subseteq [n]}\) as well as \(\max_{x \in \{0,1\}^n} \|F\|\), where we have to compute the eigenvalues of \(2^n\) many \(r \times r\)-matrices, or the expectations where we evaluate the matrix polynomials in \(2^n\) many points and sum up their values. Therefore, our program grows exponentially in \(n\) (and polynomially in \(r\)), and we are only able to run the program for small \(n\). For example, for \(n = 13\), even in the case \(r = 1\), the program did not terminate after one hour of computation.

In order to give a better impression of the running times, we present a short table with some rough running times\(^{16}\) of our program with selected combinations of \(n\) and \(r\).

<table>
<thead>
<tr>
<th>(r) (\backslash n)</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1800</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>10</td>
<td>60</td>
<td>180</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2000</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td></td>
<td>120</td>
<td>900</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1400</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>40</td>
<td>150</td>
<td>700</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1000</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
<td>40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>400</td>
</tr>
<tr>
<td>15</td>
<td>70</td>
<td>250</td>
<td>1000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>20</td>
<td>90</td>
<td>400</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>60</td>
<td>200</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 6.1: running time in seconds

**Remark 6.2.**

Our numerical experiments not only validate our conjecture, they also indicate two more things.

(a) In all of our computed examples, the left-hand side of (6.2) is much larger than the right-hand side. This indicates that we could possibly improve and simplify the bounds in our conjecture. If the conjecture is true, it is eventually still true even if we replace (6.1) by

\[
k = \deg F \left( \frac{1}{\varepsilon^{\tau}} \mathbb{E} \text{Tr} \left( \frac{\varepsilon}{\tau} Q(x) \ln \left( \frac{\varepsilon}{\tau} Q(x) \right) \right) + \ln \left( \frac{1}{\sqrt{\varepsilon}} \right) \right).
\]

\(^{16}\)We run the program on our laptop with a 2.5 GHz processor (2 cores) and 4 GB RAM.
(b) All of our examples also indicate that we even have an actual approximation instead of only a lower bound in \((6.2)\). That means, our experiments indicate that even the stronger statement

\[
\left| \frac{\gamma}{\tilde{\tau}} \mathbb{E}_x \text{Tr}(F(x)Q(x)) - \mathbb{E}_x \text{Tr}(F(x)B(x)^2) \right| \leq \varepsilon \gamma \max_{x \in \{0,1\}^n} \|F(x)\|
\]

is true.

### 6.2 Further open questions

We conclude this work with some open problems concerning the other two main parts of our work.

(i) First, it would be interesting to know if there exist other Boolean polynomial equation systems than \((2.15)\) with a linear lower bound on the degree of any Real Nullstellensatz refutation as Grigoriev’s Theorem \(2.27\) shows it for \((2.15)\). \((2.28)\) yields such another example, but as we saw in Corollary \(2.32\) considered in the residue ring \(\mathbb{R} \langle x \rangle\) both equation systems are equivalent. Therefore, it would be interesting to have an example completely independent from \((2.15)\).

(ii) The main reason for the simplicity of our new proof of Grigoriev’s Theorem lies in exploiting symmetry and applying Blekherman’s Theorem \(2.26\). We would be not surprised if this nice characterization of symmetric sums of squares could be applied to other problems or theorems where sums of squares of Boolean functions occurs and where symmetry is present. Then, as in our work, applying Blekherman’s Theorem could eventually yield to significant simplifications of the proofs of such results.

(iii) Blekherman’s Theorem, or more precisely its Boolean version Theorem \(3.37\) gives a complete characterization of symmetric sums of squares of Boolean functions of degree at most \(\frac{n}{2}\). Of course, it would be very nice if one could extend this result and could obtain a similar characterization for all symmetric sums of squares, that means for symmetric sums of squares of Boolean functions of degree up to \(n\).

(iv) Considering our third main part, the analysis of the feasible pseudo-densities in Chapter \(5\) the most important thing would be to prove our two conjectures from Section \(5.3\).

In this connection, it would be interesting to know how an optimal, feasible pseudo-density could look like for some \(\delta = \frac{1}{4} - \varepsilon\) for some small \(\varepsilon > 0\). For example, one approach to solve this problem would be to derive such a pseudo-density from the actual density from Proposition \(5.21\) by perturbing it a bit. Of course, at the moment it is completely unclear how such a perturbation
could look under the condition to keep the required properties of a feasible pseudo-density.

(v) Last, it would be preferable to know more about the non-centered case. In this connection, it would be nice to have a closed formula for the Grigoriev pseudo-density $D_g,r \in \mathbb{R}^{\{0,1\}^n}$ for every $r \in \left(\lceil \frac{n-3}{2} \rceil, \lfloor \frac{n+3}{2} \rfloor \right)$ similar to the case $r = \frac{n}{2}$ in (4.88). We think such a formula should not be too hard to prove, but for lack of time, it is not part of this work.

Then, with such a formula, it should be not too hard to prove Conjecture 5.29. Finally, it would be nice to know more about the non-centered and non-standard case, that is not treated in this work and about which nothing else is known so far to our knowledge.
Zusammenfassung auf Deutsch

In der kombinatorischen Optimierung lassen sich sehr viele Probleme als lineare Optimierungsprobleme modellieren. Dies bedeutet, dass sich das jeweilige Problem auf das Optimieren einer linearen Zielfunktion über einem Polyeder zurückführen lässt. Ein Polyeder ist dabei die Lösungsmenge endlich vieler linearer Ungleichungen.

Deshalb ist die allseits bekannte lineare Programmierung auch ein Standardwerkzeug zum Lösen solcher Probleme. Die Anzahl der benötigten linearen Ungleichungen, die benötigt werden um das jeweilige Polyeder zu beschreiben, ist dabei ein Maß für die Komplexität des linearen Programms. Da die in den kombinatorischen Optimierungsproblemen auftretenden Polyeder aber sehr häufig durch eine sehr große Anzahl linearer Ungleichungen definiert sind, ist es oft sehr schwer solche Probleme direkt durch ein lineares Programm zu lösen.

In diesem Zusammenhang machen wir die folgende Beobachtung. Seien $P \subseteq \mathbb{R}^n$ und $Q \subseteq \mathbb{R}^N$ zwei Polyeder mit $n \leq N$ und sei $l : \mathbb{R}^n \to \mathbb{R}$ ein lineares Funktional und $\pi : \mathbb{R}^N \to \mathbb{R}^n$ eine lineare Abbildung mit $\pi(Q) = P$. Dann gilt

$$\min_{x \in P} l(x) = \min_{y \in Q} l \circ \pi(y),$$

wobei das rechte Optimierungsproblem ebenfalls ein lineares Optimierungsproblem darstellt.

Es kann deshalb hilfreich sein ein Polyeder $P$ als Projektion oder allgemeiner als Bild einer linearen Abbildung eines anderen höherdimensionalen Polyeders $Q$ darzustellen, das durch weniger Ungleichungen als $P$ beschrieben werden kann, um dann ein lineares Funktional über $Q$ anstatt über $P$ zu optimieren, um die Komplexität des linearen Programms zu reduzieren.

Dass dies tatsächlich funktionieren kann zeigt die folgende Abbildung auf der nächsten Seite, welches ein einfaches, verständliches und gut vorstellbares Beispiel liefert. Hier wird ein Oktaeder, ein Polygon beschrieben durch acht lineare Ungleichungen, als Projektion eines dreidimensionalen Polyeders, das durch nur sechs lineare Ungleichungen beschrieben wird, dargestellt.

Ein allgemeineres auch in höheren Dimensionen gültiges Beispiel, welches die Stärke des oben genannten Verfahrens aufzeigt, liefert die Familie der Permutationspolytope. Hierbei sei daran erinnert, dass ein beschränktes Polyeder auch Polytop genannt wird. Zu jedem $n \in \mathbb{N}$ ist das Permutationspolytop $P_n$ definiert als die konvexe Hülle von $\{(\sigma(1),\ldots,\sigma(n))^T \mid \sigma \in S_n\}$, die Menge aller Permutationen des Vektors $(1,\ldots,n)^T$. Nach einem Resultat von Rado [Rad52] sind mindestens $2^n - 2$ lineare...
Ungleichungen notwendig, um $P_n$ zu beschreiben. Es ist hingegen leicht zu sehen, dass $P_n$ das Bild unter einer linearen Abbildung des sogenannten Birkhoff-Polytops $B_n \subseteq \mathbb{R}^{n \times n}$, die konvexe Hülle aller $n \times n$ Permutationsmatrizen, ist. Im Gegensatz zum Permutationspolytop sind nur $n^2$-viele lineare Ungleichungen nötig um $B_n$ zu beschreiben, was aus dem bekannten Satz von Birkhoff und von Neumann folgt. Ein Beweis dieses Satzes sowie weitere Erläuterungen zu den beiden eben genannten Familien von Polytopen finden sich zum Beispiel in [Bar02, Sections II.5, II.6].

Insgesamt ist zu diesem Thema bereits ausführliche Forschung betrieben worden und es ist sehr viel darüber bekannt – sowohl positive aber auch negative Resultate, wie wir in Kürze sehen werden.

In der kombinatorischen Optimierung spielt die Familie der sogenannten Korrelationspolytope eine wichtige Rolle. Zu jedem $n \in \mathbb{N}$ definieren wir

$$\text{CORR}_n := \text{conv}\{xx^T \mid x \in \{0,1\}^n\} \subseteq \mathbb{R}^{n \times n}.$$ 

Diese Polytope sind ein Beispiel dafür, dass das oben genannte Verfahren auch an seine Grenzen stoßen kann. Fiorini et al. [FMP+12] waren die ersten die zeigten, dass die Anzahl der Ungleichungen, die nötig sind um ein beliebiges Polytop, das linear auf $\text{CORR}_n$ abgebildet werden kann, zu beschreiben, exponentiell in $n$ wächst. Mittlerweile sind konkrete untere Schranken für diese Anzahl bekannt. Kaibel und Weltge geben in [KW15] einen einfachen Beweis dafür an, dass mindestens $1.5^n$ viele solcher linearen Ungleichungen nötig sind, was nach unserem Wissen die aktuell beste bekannte untere Schranke ist.

Abhilfe könnte hier die semidefinite Programmierung liefern, eine mächtigere Verallgemeinerung der linearen Programmierung. Die zulässige Menge eines semidefiniten Programms (SDP) bildet dabei ein sogenanntes Spektraeder; das eine Verallgemeinerung eines Polyeders darstellt. Eine Menge $S \in \mathbb{R}^n$ wird Spektraeder genannt, wenn sie von der Form

$$S = S_A := \{x \in \mathbb{R}^n \mid A(x) := A_0 + x_1A_1 + \ldots + x_nA_n \succeq 0\}$$

ist für gewisse $k \in \mathbb{N}$ und symmetrische Matrizen $A_0, \ldots, A_n \in \mathbb{R}^{k \times k}$. Hierbei bedeutet $\succeq 0$, dass die Matrix positiv semidefinit (psd) ist, und $A(x) \succeq 0$ wird auch lineare Matrixungleichung genannt. Weiter bezeichnen wir $k$ als die Größe der linearen Matrixungleichung und sagen $S$ ist definiert durch die lineare Matrixungleichung
A(x) \geq 0. Ein semidefinites Programm ist dann die Optimierung einer linearen Zielfunktion über einem Spektraeder. Hierbei ist die Größe der linearen Matrixungleichung, die das Spektraeder definiert, das entscheidende Maß für die Komplexität des SDPs.

Es ist einfach zu sehen, dass jedes Polyeder auch immer ein Spektraeder ist. Für ein Polyeder P, das durch m lineare Ungleichungen \( l_1(x) \geq 0, \ldots, l_m(x) \geq 0 \) definiert ist, wählen wir A als die Diagonalmatrix mit \( i \)-tem Diagonaleintrag \( l_i \). Dann ist \( A(x) \geq 0 \) eine lineare Matrixungleichung und da eine Diagonalmatrix genau dann psd ist, wenn jeder ihrer Diagonaleinträge nichtnegativ ist, erhalten wir \( P = S_A \). Somit sehen wir, dass semidefinite Programmierung tatsächlich eine Verallgemeinerung der linearen Programmierung ist.


Um zu unserem Ausgangsproblem zurückzukommen, wollen wir die Idee von (i) verallgemeinern. Für ein Polyeder \( P \subseteq \mathbb{R}^n \), das durch sehr viele lineare Ungleichungen definiert ist, kann es sehr nützlich sein, ein Spektraeder S, das durch eine lineare Matrixungleichung von kleiner Größe k definiert ist, und eine lineare Abbildung \( \pi : R^N \rightarrow \mathbb{R}^n \) mit
\[
P = \pi(S)
\]
zu finden. In diesem Fall sagen wir, dass P eine positiv semidefinite Erweiterung der Größe k besitzt. Sei nun \( l : \mathbb{R}^n \rightarrow \mathbb{R} \) ein lineares Funktional. Dann gilt
\[
\min_{x \in P} l(x) = \min_{y \in S_A} l \circ \pi(y),
\]
wobei das rechte Optimierungsproblem ein SDP von niedrigerer Komplexität ist als das linke lineare Programm.

Im folgenden zeigen wir ein Beispiel für eine positiv semidefinite Erweiterung der Größe drei eines Quadrats, was wiederum ein durch vier lineare Ungleichungen
Zusammenfassung auf Deutsch

definiertes Polyeder ist.

\[ S = \left\{ x \in \mathbb{R}^3 \left| \begin{array}{ccc} 1 & x_1 & x_2 \\ x_1 & 1 & x_3 \\ x_2 & x_3 & 1 \end{array} \right\} \geq 0 \right\} \]

Quelle: [AKW16]

Allgemeinere Resultate über Polyeder, die positiv semidefinite Erweiterungen von kleiner Größe besitzen, finden sich in [GPRT17].

Nun ist natürlich die Frage, ob die Korrelationspolytope positiv semidefinite Erweiterungen von kleiner Größe besitzen, von großem Interesse. Zu diesem Zweck definieren wir den \( \text{psd Rang} \) eines Polytops \( P \) als das kleinste \( k \), sodass \( P \) eine positiv semidefinite Erweiterung der Größe \( k \) besitzt und schreiben hierfür \( r_{\text{psd}}(P) \).

In ihrer aktuellen und preisgekrönten Arbeit [LRS15] konnten Lee, Raghavendra und Steurer diese Frage mit Nein beantworten. Sie bewiesen, dass \( r_{\text{psd}}(\text{CORR}_n) \) schneller wächst als jedes Polynom in \( n \). Konkret bedeutet dies

\[
\lim_{n \to \infty} \frac{r_{\text{psd}}(\text{CORR}_n)}{n^k} = \infty
\]

für jedes beliebige \( k \in \mathbb{N} \).

In einer genaueren Analyse ihres Beweises von (ii), zeigen sie sogar die Existenz einer natürlichen Zahl \( N \) und einer Konstanten \( \alpha > 0 \), sodass für jedes \( n \geq N \)

\[
r_{\text{psd}}(\text{CORR}_n) \geq 2^\alpha \left( \frac{n}{\ln n} \right)^{2/11}
\]

(iii)

gilt. Anders gesagt kann jedes Spektraeder, das linear auf \( \text{CORR}_n \) abgebildet werden kann, nur durch lineare Matrixungleichungen, die mindestens von Größe \( 2^\alpha \left( \frac{n}{\ln n} \right)^{2/11} \) sind, definiert werden. Dieses Resultat ist die erste bekannte super-polynomiale untere Schranke an die Größe positiv semidefiniter Erweiterungen einer konkreten Familie von Polytopen.


Zusätzlich präsentieren wir auch eigene Ideen und Gedanken wie der Beweis verbreiteter werden könnte und wie er von einem eher reell algebraischen Standpunkt aus gesehen werden kann. Hierbei erreichen wir mit unseren Ansätzen an manchen Stellen sogar etwas stärkere Aussagen als [LRS15].

Unter anderem erläutern wir warum [LRS15] wirklich (iii) beweisen, während sie die Aussage in ihrer Originalarbeit nur mit der kleineren Potenz $\frac{2}{13}$ behaupten (Section 4.6).  

Im Gesamten ist unsere Arbeit folgendermaßen strukturiert. In Kapitel 1 treffen wir nur Vorbereitungen für den weiteren Verlauf der Arbeit und stellen einige Grundlagen bereit, die hauptsächlich aus dem Bereich der linearen Algebra stammen.

Der erste von drei Hauptteilen, in die unsere Arbeit aufgeteilt werden kann, folgt dann in Kapitel 2. Hier liefert wir einen neuen und einfacheren Beweis eines berühmten und oft zitierten Theorems von Grigoriev [Gri01a], in dem er untere Gradeschranken für den Reellen Nullstellensatz (RNS) beweist. Der Reelle Nullstellensatz liefert ein Zertifikat für die Unlösbarkeit eines polynomialen Gleichungssystems, wobei das Zertifikat auf Quadratsummen von Polynomen beruht. Grigorievs Theorem ist dabei das erste bekannte Resultat, das eine untere Schranke für den RNS liefert für den Fall, dass das zugrundeliegende polynomiale Gleichungssystem all die quadratischen Polynome $X_1^2 - X_1, \ldots, X_n^2 - X_n$ enthält. Genau dieser Fall ist aber von großer Bedeutung in der kombinatorischen Optimierung.


17 Unabhängig von uns wurde dies auch von Troy Lee et al. bemerkt [LPWY16]
Wir präsentieren unseren neuen Beweis in Abschnitt 2.4. Dabei werden wir hauptsächlich im Ring

\[ \mathbb{R}[x] := \mathbb{R}[X_1, \ldots, X_n] / \langle X_i^2 - X_i \mid i \in [n] \rangle \]

arbeiten, wobei \( \langle X_i^2 - X_i \mid i \in [n] \rangle \) das von den \( X_i^2 - X_i \)'s erzeugte Ideal bezeichnet. Hierfür werden wir in Abschnitt 2.2 einige grundlegende Notationen und Fakten von \( \mathbb{R}[x] \) vorstellen.

Der wesentliche Grund für die Einfachheit unseres Beweises beruht auf dem Ausnutzen von Symmetrien und dem Anwenden eines relativ aktuellen aber unveröffentlichten Resultats von Blekherman\(^{18}\). Deshalb werden wir uns in Abschnitt 2.3 etwas näher mit symmetrischen Polynomen in \( \mathbb{R}[x] \) beschäftigen.


Kapitel 2 kann als Vorbereitung auf die darauffolgenden Kapitel gesehen werden. Ein wichtiger Bestandteil in der Arbeit von [LRS15] sind die Booleschen Funktionen, insbesondere die sogenannten Pseudo-Dichten und Pseudo-Erwartungen. Boolesche Funktionen sind Funktionen von \( \{-1,1\}^n \) beziehungsweise \( \{0,1\}^n \) nach \( \mathbb{R} \). In den ersten drei Abschnitten stellen wir einige grundlegenden Fakten dieser Funktionen zur Verfügung. Diese Abschnitte sind im wesentlichen sehr elementar und können von in diesem Bereich erfahrenen Lesern problemlos ausgelassen werden. In Abschnitt 3.3 werden wir vor allem symmetrische Boolesche Funktionen betrachten, wobei wir hauptsächlich auf Resultate aus Abschnitt 2.3 aufbauen werden.

In Abschnitt 3.4 leiten wir dann das oben erwähnte Korollar aus Grigorievs Theorem her. Es liefert die Existenz einer Familie von nichtnegativen Booleschen Funktionen \( f_n : \{0,1\}^n \rightarrow \mathbb{R}_+ \) von Grad zwei, die sich nicht als Summe von Quadraten von Booleschen Funktionen mit höchstens Grad \( n \) schreiben lassen. Anschließend werden wir in Abschnitt 3.5 den Begriff der Pseudo-Dichten und Pseudo-Erwartungen einführen und motivieren. Wir werden die Beziehung zwischen diesen beiden speziellen Klassen Boolescher Funktionen erläutern und wir werden Resultate aus Abschnitt 3.3 anwenden, um aufzuzeigen, wie sich diese Funktionen verhalten, wenn Symmetrien vorliegen.

Möglicherweise erscheinen einem die Beweise und Erläuterungen dieses Kapitels an gewissen Stellen als relativ einfach und zu detailliert, aber es ist unser Ziel auch Lesern ohne irgendwelche Vorkenntnisse über Boolesche Funktionen eine plausible und verständliche Arbeit zu liefern.

Kapitel 5 beinhaltet dann eine genauere Analyse derjenigen Pseudo-Dichten, die für den Beweis von \cite{LRS15} relevant sind, was den dritten Hauptteil unserer Arbeit darstellt. Hier werden wir unter anderem ausführen, wieso wir in unserem Kontext immer Symmetrie erwarten und ausnutzen können. Unter gewissen Annahmen geben wir eine Charakterisierung der relevanten, symmetrischen Pseudo-Dichten, wobei das Resultat von Blekherman aus Kapitel 2 erneut eine große Hilfe für unsere Ausführungen ist.

Mit unseren Beobachtungen aus diesem Kapitel sind wir in der Lage gewisse Optimalitätsaussagen für die unteren Schranken für den psd Rang der Familie der Korrelationspolytope, wie sie \cite{LRS15} gefunden haben, zu treffen.

Wir weisen darauf hin, dass wir jeweils zu Beginn eines Kapitels weitere Motivation, Einführungen und Erklärungen zu den einzelnen Inhalten des Kapitels geben werden.
Bibliography


