Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Alex Dunbar Date

Leveraging Algebraic and Geometric Structures in Optimization

By

Alex Dunbar
Doctor of Philosophy

Mathematics

Grigoriy Blekherman, Ph.D.
Advisor

Lars Ruthotto, Ph.D.
Committee Co-chair

David Borthwick, Ph.D.
Committee Member

Elizabeth Newman, Ph.D.
Committee Member

Accepted:

Kimberly Jacob Arriola, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date

Leveraging Algebraic and Geometric Structures in Optimization

By

Alex Dunbar
B.A., Rice University, TX, 2020
B.S., Rice University, TX, 2020
M.Sc., Emory University, GA, 2023

Advisor: Grigoriy Blekherman, Ph.D.

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Mathematics
2025

Abstract

Leveraging Algebraic and Geometric Structures in Optimization
By Alex Dunbar

Mathematical optimization has been a highly active field in recent years. Typically, one seeks to leverage structure, such as convexity, when studying optimization problems. We focus on several optimization problems for which structures in the objective function or constraints are naturally phrased in the language of algebraic geometry.

The first problem we consider is regression over the space of rational functions in tropical (max-plus) algebra. Such functions form a widely expressive class of function approximators and have recently proven useful in the theoretical analysis of ReLU neural networks. We develop an alternating heuristic to solve regression problems over tropical rational functions by leveraging known results from tropical linear systems and polynomial regression. Numerical experiments show the strengths and weaknesses of the heuristic and we provide a connection between our method and geometric aspects of the loss function.

The second problem we consider is semidefinite programming in the $\star_M$ tensor-tensor product framework. We demonstrate that the choice of matrix M corresponds to the representation theory of an underlying group action. This connection lends the $*_M$ product to be a natural framework for certain invariant semidefinite programs. We demonstrate the M-SDP framework on certain invariant sums of squares polynomials and low rank tensor completion problems

The final problem we consider is the expression of the convex hull of a set defined by three quadratic inequality constraints using nonnegative linear combinations (aggregations) of the constraints. Our approach relates the problem to the topology of the spectral curve, defined as the zero set of the determinant of linear combinations of the defining quadratics. In particular, we characterize the nonexistence of solutions to systems of inequalities in terms of the hyperbolicity of the spectral curve. Hyperbolic curves are well-studied in real algebraic geometry, as their zero sets consist of maximally nested ovals, the innermost of which bounds a convex cone. By studying (non)intersections of polyhedral cones of aggregations with hyperbolicity cones of the spectral curve, we provide a sufficient condition for the convex hull to be given by aggregations and characterize when finitely many aggregations suffice for such a description.

Leveraging Algebraic and Geometric Structures in Optimization

By

Alex Dunbar
B.A., Rice University, TX, 2020
B.S., Rice University, TX, 2020
M.Sc., Emory University, GA, 2023

Advisor: Grigoriy Blekherman, Ph.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Mathematics
2025

Acknowledgments

I would like to start by acknowledging the wonderful support of Greg Blekherman as my advisor for the last few years. I have learned a lot from Greg, from technical mathematics to navigating mathjobs. I am incredibly grateful for his willingness to advise me from across town and am looking forward to continued collaborations in the future.

I was lucky to be surrounded by wonderful faculty at Emory, but I am especially grateful for my committee members. Lars Ruthotto and Elizabeth Newman were consistently willing to reach across disciplines and work with me on exciting projects. I am also grateful for David Borthwick for his consistent support, starting from first-year analysis and through his willingness to hop in and help as DGS when needed.

I would also like to acknowledge the help of all of the other wonderful graduate students at Emory University: Santiago Arango, Roberto Hernandez, Jasmine Camero, Ben Yellin, Rohan Nair, Shilpi Mandal, Ariana Brown, Jiaqi Yang, Elle Buser, Emma Hart, Griffin Johnston, Katie Keegan, Sean Longbrake, Ayush Basu, Vishwanath Seshagiri, Ylli Andoni, Mitchell Scott, Christopher Keyes, Mike Cerchia, Marcelo Sales, and last but certainly not least, Kelvin Kan. My time at Emory would not have been the same without your support.

Beyond my academic support network, my family was constantly there for me, helping out where they could. I would also like to thank my friends from Atlanta Trail Runners for making sure I had a life outside of math. In particular, I want to thank Gabi and Ozzie Seplovich for everything they did to support me in my last year of writing.

Finally, nothing in this dissertation would have been possible without the support of Vicki Powers. Vicki helped to spark my interest in real algebraic geometry and provided wonderful guidance as an advisor in my early years of graduate study. Her effect on the field of applied algebraic geometry and the associated research community cannot be overstated and I am deeply grateful for the opportunity to have worked with her. She will be missed.

Contents

List of Figures

List of Tables

List of Algorithms

Chapter 1

Introduction

Many optimization problems have underlying structure in their objective function or constraint set. In this dissertation, we explore several optimization problems whose structure is naturally phrased using a blend of convex and algebraic geometry.

As a starting point, consider *linear programming*–the problem of maximizing a linear function over a polyhedron:

$$\max \langle c, x \rangle \text{ s.t. } \langle a^{(i)}, x \rangle = b_i \text{ for } i \in [m], \;\; x \in \mathbb{R}^n_+.$$

All problem data in a linear program is linear, and one can use linear algebraic techniques to study and solve such programs. A natural generalization of linear programs arises by replacing the cone $\mathbb{R}^n_+$ with another convex cone. An important example is *semidefinite programming*–the problem of maximizing a linear function over an affine slice of the cone of positive semidefinite matrices $\mathcal{S}^n_+$:

$$\max \langle C, X \rangle \text{ s.t. } \langle A^{(i)}, X \rangle = b_i \text{ for } i \in [m], \;\; X \in \mathcal{S}^n_+.$$

While solving semidefinite programs is more difficult than solving linear programs, there are practical interior-point algorithms for their solution. One important application of

semidefinite programming is in the solution of polynomial optimization problems. Algebraic geometry is inseparable from this application. To make this connection, we consider the set $\mathcal{P}_{n,2d}$ of nonnegative homogeneous real polynomials of degree $2d$ in n variables and the set $\Sigma_{n,2d}$ of homogeneous real polynomials of degree $2d$ in n variables which admit a decomposition as a sum of squares of polynomials of degree d. Both $\mathcal{P}_{n,2d}$ and $\Sigma_{n,2d}$ are convex cones in the real vector space of homogeneous polynomials of degree $2d$ in n variables. Moreover, $\Sigma_{n,2d} \subseteq \mathcal{P}_{n,2d}$. This inner approximation has been leveraged to construct *semidefinite relaxations* of polynomial optimization problems. A crucial observation in the construction of such relaxations is that checking if a degree $2d$ real multivariate polynomial $p(x)$ is a sum of squares (SOS) amounts to finding a positive semidefinite matrix Q, called a *Gram matrix* for p, such that $[x]_d^\top Q[x]_d = p(x)$, where $[x]_d$ is a vector of all monomials of degree d. Moreover,

$$\inf\{p(x) \mid x \in \mathbb{R}^n\} = \max\{\gamma \in \mathbb{R} \mid p(x) - \gamma \geqslant 0 \text{ for all } x \in \mathbb{R}^n\} \geqslant \max\{\gamma \in \mathbb{R} \mid p(x) - \gamma \text{ is SOS}\}.$$

A classical result of Hilbert [Hil88] shows that $\Sigma_{n,2d} = \mathcal{P}_{n,2d}$ if and only if $n = 2$, $d = 1$, or $(n, 2d) = (3, 4)$. However, each element of $\mathcal{P}_{n,2d}$ admits a decomposition as a sum of squares of rational functions, motivating a hierarchy of semidefinite programming problems with multipliers of increasing degree. This approach, called the *Moment-SOS hierarchy*, has found great success in applications [Las01, Par03].

As a further generalization of semidefinite programming, one considers *hyperbolic programming*. A real homogeneous polynomial p is hyperbolic with respect to a point $e \in \mathbb{R}^n$ if $p(e) \neq 0$ and for all $a \in \mathbb{R}^n$, the univariate polynomial $p(te - a) \in \mathbb{R}[t]$ is real rooted. A canonical example is the polynomial $p(x_1, x_2, \ldots, x_n) = \det(x_1 I + x_2 A^{(2)} + x_3 A^{(3)} + \ldots + x_n A^{(n)})$ for some fixed symmetric matrices $A^{(2)}, A^{(3)}, \ldots, A^{(n)}$, which is hyperbolic with respect to $(1, 0, 0, \ldots, 0)$. A fundamental fact about hyperbolic polynomials is that the connected com-

ponent of e in $\mathbb{R}^n \backslash \mathcal{V}_{\mathbb{R}}(p)$ is an (open) convex cone. If $\Lambda(p, e)$ denotes the closure of this cone, then a hyperbolic program is

$$\max \langle c, x \rangle \text{ s.t. } \langle a^{(i)}, x \rangle = b_i \text{ for } i \in [m], \quad x \in \Lambda(p, e).$$

In this dissertation, we investigate some problems at this interface of algebra, geometry, and optimization. We will see that a blend of algebraic and convex structure can be leveraged for new algorithms and analysis for a variety of optimization problems.

Contributions and Organization The contributions of this dissertation are organized across three distinct projects, each one occupying a chapter. Before presenting the results of these projects, we provide the relevant preliminaries from optimization and algebraic geometry needed in the subsequent chapters in Chapter 2.

In Chapter 3, we study regression problems over the space of *tropical rational functions*, continuous piecewise linear functions of the form

$$r(x) = \max_{i \in [D]}(\langle w^{(i)}, x \rangle + p_i) - \max_{i \in [D]}(\langle w^{(i)}, x \rangle + q_i),$$

for fixed $w^{(1)}, w^{(2)}, \ldots, w^{(D)} \in \mathbb{R}^n$. Such functions are the rational functions over the tropical (max-plus) semiring. A recent line of work, initiated by Zhang, Naitzat, and Lim [ZNL18] has connected tropical rational functions and *ReLU Neural Networks*. We study the optimization aspect of this connection, focusing on the ℓ^∞ regression problem: For fixed $(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \ldots, (x^{(N)}, y^{(N)}) \in \mathbb{R}^n \times \mathbb{R}$ and fixed $w^{(1)}, w^{(2)}, \ldots, w^{(D)} \in \mathbb{R}^n$, we study the problem

$$\arg \min_{p_1, p_2, \ldots, p_D, q_1, q_2, \ldots, q_D} \left(\max_{j \in [N]} \left| \max_{i \in [D]}(\langle w^{(i)}, x^{(j)} \rangle + p_i) - \max_{i \in [D]}(\langle w^{(i)}, x^{(j)} \rangle + q_i) - y_j \right| \right). \tag{1.1}$$

It is known that the tropical *polynomial* regression problem has a closed form solution which

only involves (tropical) matrix vector products and vector additions [MT20]. So, we propose an alternating heuristic which alternately solves for the p_i and the q_i. We demonstrate the heuristic numerically on synthetic datasets.

We then connect the heuristic to the geometry of the loss function, showing that the iterates are contained in the nondifferentiability locus of the loss function. Additionally, we discuss a connection between regression with tropical rational functions and *tropical linear programming*.

In Chapter 4, we study structured semidefinite programming problems which are connected to the $\star_M$ tensor-tensor product. Specifically, we use the $\star_M$ product structure to define cones of *M-positive semidefinite* third-order tensors. These M-PSD cones share many properties with PSD matrices and allow for familar matrix semidefinite programming problems, such as minimum nuclear norm matrix completion, to be translated to the third-order tensor case.

Additionally, we study the algebraic structure of the $\star_M$ product on tubes (tensors of format $1 \times 1 \times n_3$), connecting these products to the representation theory of finite groups. When the matrix M is chosen compatibly with a representation $\rho : G \to GL_{n_3}(\mathbb{R})$, there is an explicit subspace of tubes $\mathbf{a}$ for which the multiplication map $\mathbf{b} \mapsto \mathbf{a} \star_M \mathbf{b}$ is ρ-equivariant. Building on this representation theoretic interpretation, we show that there is a natural connection between semidefinite programming problems over cones of M-PSD tensors and the *invariant semidefinite programs* studied by Gaterman and Parrilo [GP04].

In Chapter 5, we study *aggregations* of quadratic inequalities using algebraic topology. Given three linearly independent symmetric matrices $Q_1, Q_2, Q_3 \in \mathcal{S}^{n+1}$, we are interested in two related questions:

1. Let f_i^h be the quadratic form associated to Q_i for $i \in [3]$. When is the real projective variety $\mathcal{V}_{\mathbb{R}}(f_1^h, f_2^h, f_3^h) \subseteq \mathbb{RP}^n$ empty?

2. Set $S = \left\{x \in \mathbb{R}^n \mid (x^\top, 1)Q_i(x^\top, 1)^\top \leqslant 0, i \in [3]\right\}$ to be the set defined by the (affine) quadratic inequalities. When is there a finite set $\Lambda \subseteq \mathbb{R}^3_+$ such that

$$\overline{\text{conv}}(S) = \bigcap_{\lambda \in \Lambda} \left\{ x \in \mathbb{R}^n \;\middle|\; (x^\top, 1)(\sum_{i=1}^{3} \lambda_i Q_i)(x^\top, 1)^\top \leqslant 0 \right\}?$$

We take a unified perspective to answer these two questions by studying the topology of the *spectral curve*, defined in $\mathbb{RP}^2$ by the vanishing of $g(\lambda) = \det(\sum_{i=1}^{3} \lambda_i Q_i)$. Our main technical tool is a spectral sequence developed in [AL12] which relates the topology of (projective) solution sets to systems of quadratic inequalities to the topology of combinations of the defining quadratics with specified number of positive eigenvalues.

As an answer to the first question, we show that, aside from exceptional small n cases, the projective variety $\mathcal{V}_{\mathbb{R}}(f_1^h, f_2^h, f_3^h)$ is empty if and only if the spectral curve is hyperbolic and there is a linear combination of the defining quadratics which has n positive and one negative eigenvalue. We then leverage the hyperbolicity of the polynomial g to provide a sufficient condition for a description of $\overline{\text{conv}}(S)$ in terms of aggregations. Specifically, such a description exists when the polyhedral cone of aggregations does not intersect a hyperbolicity cone of g. This condition generalizes conditions based on hidden convexity results derived in [BDS24, DMnS22].

Chapter 2

Preliminaries

This chapter provides the mathematical background for the results in subsequent chapters, focusing on relevant topics in optimization and algebraic geometry. Since the methods and results in this dissertation involve a broad range of areas of mathematics, we include thorough reviews in an attempt to keep the dissertation self-contained and accessible to a wider audience.

2.1 Convex Geometry

A unifying theme throughout this dissertation will be the presence of convexity. We review the needed ideas here. A complete introduction can be found in [Bar02, BV04].

Definition 2.1.1 (Convex Set)**.** *Let V be a real vector space. A subset $C \subseteq V$ is* convex *if for all $x, y \in C$ and all $t \in [0,1]$, we have*

$$tx + (1-t)y \in C.$$

The set C is called a convex cone *if it is convex and in addition, for any $x \in C$ and $\lambda \geqslant 0$, we have $\lambda x \in C$.*

Example 2.1.1. An immediate observation is that affine subspaces of a real vector space are convex. Moreover, it follows from the definition of a convex set that the intersection of any number of convex sets is convex, leading to the following important examples:

- *Half-spaces*: Sets of the form $\{x \in \mathbb{R}^n | \ell(x) \geqslant b\}$ for some linear functional $\ell \in (\mathbb{R}^n)^*$ and $b \in \mathbb{R}$.
- *Polyhedra*: The intersection of finitely many half-spaces.
- *Spectrahedra*: The intersection of the convex cone of positive definite matrices and an affine subspace.

$\diamond$

A closed convex set comes with distinguished convex subsets, known as faces.

Definition 2.1.2 (Face, Extreme Point, Extreme Ray)**.** *Let $C \subseteq \mathbb{R}^n$ be a closed convex set. A subset $F \subseteq C$ is a* face *of C if for all $\lambda \in [0,1]$ and all $x, y \in C$,*

$$\lambda x + (1-\lambda)y \in F \implies x, y \in F.$$

A point $v \in C$ is an extreme point *of C if $\{v\}$ is a face of C. If C is a convex cone and $v \in C$ satisfies the property that $v = \lambda_1 x + \lambda_2 y$ for $\lambda_1, \lambda_2 \geqslant 0$ and $x, y \in C$, then v* spans an extreme ray *of C.*

Given an arbitrary set, we construct the smallest convex set (cone) which contains it.

Definition 2.1.3 (Convex Hull, Conical Hull)**.** *Let $S \subseteq \mathbb{R}^n$. The* convex hull *of S is the set*

$$\operatorname{conv}(S) = \bigcap_{S \subseteq C \text{ convex}} C = \left\{ \sum_{i=1}^{r} \lambda_i x^{(i)} \;\middle|\; \lambda_i \geqslant 0, x^{(i)} \in S, r \in \mathbb{N} \right\}.$$

The conical hull *of S is the set*

$$\operatorname{cone}(S) = \left\{ \sum_{i=1}^{r} \lambda_i x^{(i)} \;\middle|\; \sum_{i=1}^{r} \lambda_i = 1, \lambda_i \geqslant 0, x^{(i)} \in S, r \in \mathbb{N} \right\}.$$

Theorem 2.1.4 (Krein-Milman (see e.g.; [Bar02]))**.** *Let $K \subseteq \mathbb{R}^n$ be a compact convex set and* $\mathrm{ex}(K)$ *be the set of extreme points of K. Then, $K = \mathrm{conv}(\mathrm{ex}(K))$.*

One important property of convex sets is that it is straightforward to certify nonmembership of an element in a convex set.

Theorem 2.1.5 (Separating Hyperplane Theorem (see e.g. [BPT13, Appendix A.3]))**.** *Let $A, B \subseteq \mathbb{R}^n$ be two convex sets.*

- *If $A \cap B = \varnothing$, then there is a linear functional $\ell \in (\mathbb{R}^n)^*$ and a constant $\gamma \in \mathbb{R}$ such that $\ell(a) \leqslant \gamma$ for all $a \in A$ and $\ell(b) \geqslant \gamma$ for all $b \in B$. In this case, we say that A and B are* separated *by an affine hyperplane.*
- *If A is compact and B is closed, then we can further conclude that there exists $\ell \in (\mathbb{R}^n)^*$ and $\gamma \in \mathbb{R}$ with $\ell(a) < \gamma$ and $\ell(b) > \gamma$ for all $a \in A$ and $b \in B$. In this case, we say that A and B are* strictly separated *by an affine hyperplane.*

Remark 2.1.1. Note in particular that if B is a closed convex set and $x \notin B$, then since $\{x\}$ is compact, we can strictly separate x from B using an affine hyperplane.

Many of our applications of convex geometry will come from conic optimization problems– the optimization of a linear functional over an affine slice of a proper cone.

Definition 2.1.6 (Proper Cone)**.** *A convex cone $K \subseteq \mathbb{R}^n$ is* proper *if it is closed, full dimensional and contains no lines.*

Definition 2.1.7 (Dual Cone)**.** *Let $K \subseteq \mathbb{R}^n$. The* dual *cone to K is the cone*

$$K^* = \{y \in \mathbb{R}^n \mid \langle y, x \rangle \geqslant 0 \text{ for all } x \in K\}.$$

Note that if K is proper, then so is K^* and we have that $(K^*)^* = K$, though this last statement only needs that K is closed.

Definition 2.1.8 (Conic Optimization Primal-Dual Pair)**.** *Let $c \in \mathbb{R}^n$ be a vector, $K \subseteq \mathbb{R}^n$ a convex cone, $L : \mathbb{R}^n \to \mathbb{R}^m$ a linear transformation, and $b \in \mathbb{R}^m$ a vector.*

A conic optimization problem *has the* primal *problem*

$$\min \langle c, x \rangle \quad s.t.\ L(x) = b, \ \ x \in K. \tag{2.1}$$

and dual problem

$$\max \langle y, b \rangle \quad s.t.\ c - L^*(y) \in K^* \tag{2.2}$$

Where L^* is the adjoint linear transformation of L.

A crucial example of conic optimization for us will be semidefinite programming problems.

Example 2.1.2 (Semidefinite Programming (SDP))**.** Let $\mathcal{S}^n \simeq \mathbb{R}^{\binom{n}{2}}$ be the vector space of symmetric $n \times n$ real matrices, equipped with the inner product $\langle A, B \rangle = \text{trace}(AB)$. Recall that a matrix $X \in \mathcal{S}^n$ is called positive semidefinite (PSD) if $v^\top X v \geqslant 0$ for all $v \in \mathbb{R}^n$, or equivalently, every eigenvalue of X is nonnegative. The set of PSD matrices is denoted $\mathcal{S}^n_+$ and is a proper cone in $\mathcal{S}^n$. We typically write $X \succeq 0$ if $X \in \mathcal{S}^n_+$ and $X \succ 0$ if $X \in \text{int}(\mathcal{S}^n_+)$.

Fix matrices $C, A^{(1)}, A^{(2)}, \ldots, A^{(m)} \in \mathcal{S}^n$ and a vector $b = (b_1, b_2, \ldots, b_m)^\top \in \mathbb{R}^m$. A *semidefinite program* in primal form is the problem

$$\min \langle C, X \rangle \text{ s.t. } \langle A^{(i)}, X \rangle = b_i \text{ for all } i \in [m], X \succeq 0.$$

The corresponding dual problem is

$$\max \langle y, b \rangle \text{ s.t. } C - \sum_{i=1}^{m} y_i A^{(i)} \succeq 0.$$

$\diamond$

The conic optimization problem in Definition 2.1.8 is given as a primal-dual pair. It is always the case that the optimal value to (2.1) is greater than the optimal value to (2.2).

This property is known as *weak duality*. The two problems have equal optimal values when they are both strictly feasible, a property known as *strong duality*.

Example 2.1.3 (Semidefinite Program with Strong Duality)**.** Consider the SDP

$$\min x_2 \quad \text{s.t.} \quad \begin{bmatrix} x_1 & x_2 \\ x_2 & 1 - x_2 \end{bmatrix} \succeq 0.$$

The feasible set of this problem is the circle in the plane centered at $(x_1, x_2) = (1/2, 0)$ with radius $1/2$ so that the optimal value is $-1/2$.

The dual to this SDP is

$$\max y \quad \text{s.t.} \quad \begin{bmatrix} -y & 1/2 \\ 1/2 & -y \end{bmatrix} \succeq 0.$$

Note that this problem is feasible only for $y \leqslant -1/2$ and that the optimal value of the dual is also $-1/2$. ◇

We will also need the notion of the polar dual to a convex cone.

Definition 2.1.9 (Polar Cone)**.** *Let $K \subseteq \mathbb{R}^n$ be a convex cone. The* polar cone *to K is*

$$K^\circ = \{y \in \mathbb{R}^n | \langle y, x \rangle \leqslant 0 \text{ for all } x \in K\}.$$

Note that if K is closed, then $(K^\circ)^\circ = K$.

2.2 Tropical Algebra

In this section, we provide a brief summary of the necessary results in tropical algebra and geometry. A more thorough introduction to the subject can be found in [MS15].

The (max-plus) *tropical semiring* is the set $\mathbb{T} = \mathbb{R} \cup \{-\infty\}$ together with the operations of tropical addition $a \oplus b = \max(a, b)$ and tropical multiplication $a \odot b = a + b$. Tropical

operations are commutative and associative, and tropical multiplication distributes over tropical addition. For an integer n, we set $a^{\odot n} = a \odot a \odot \ldots \odot a = na$; i.e, tropical exponentiation is standard multiplication. We can formally adjoin variables to the tropical semiring to obtain tropical polynomials and tropical rational functions.

Definition 2.2.1 (Tropical polynomial)**.** *A* tropical polynomial *is a function of the form*

$$p(x) = \bigoplus_{j\in[D]} p_j \odot x^{\odot w^{(j)}} = \max_{j\in[D]}(p_j + \langle w^{(j)}, x\rangle),$$

Where $p_j \in \mathbb{R}$ *and* $w^{(j)} \in \mathbb{R}^n$ *for* $j \in [D]$*. The set of tropical polynomials in* n *variables is denoted* $\mathbb{T}[x_1, x_2, \ldots, x_n]$*.*

Definition 2.2.2 (Tropical rational function)**.** *A* tropical rational function *is a function of the form* $r(x) = p(x) - q(x)$ *for some tropical polynomials* $p, q \in \mathbb{T}[x_1, x_2, \ldots, x_n]$*. More explicitly,*

$$r(x) = \bigoplus_{j\in[D]} p_j \odot x^{\odot w^{(j)}} - \bigoplus_{j\in[D]} q_j \odot x^{\odot w^{(j)}} = \max_{j\in[D]}(p_j + \langle w^{(j)}, x\rangle) - \max_{j\in[D]}(q_j + \langle w^{(j)}, x\rangle)$$

Note that tropical polynomials are continuous piecewise linear convex functions and tropical rational functions are continuous and piecewise linear.

2.2.1 Tropical Hypersurfaces

In this subsection, we discuss the geometric objects associated to tropical algebraic objects. The theory of tropical geometry is rapidly developing and we provide only the basics we will need in Chapter 3.

Definition 2.2.3 (Tropical Hypersurface)**.** *Let* $p(x) = \bigoplus_{j\in[D]} p_j \odot x^{\odot w^{(j)}}$ *be a tropical polynomial. The tropical hypersurface of* p *is the set of points where the maximum is achieved (at least) twice.*

$$\begin{aligned}\mathcal{V}(p) &= \{x \in \mathbb{R}^n \mid p(x) = p_i \odot x^{w^{(i)}} = p_j \odot x^{w^{(j)}} \text{ for some } i \neq j\} \\ &= \{x \in \mathbb{R}^n \mid p \text{ is not differentiable at } x\}.\end{aligned}$$

Note that the complement of $\mathcal{V}(p)$ in $\mathbb{R}^n$ is a union of (open) polyhedra defined by the inequalities $p_i + \langle w^{(i)}, x\rangle > p_j + \langle w^{(j)}, x\rangle$ for all $j \neq i$.

Definition 2.2.4 (Polyhderal Complex)**.** *A* polyhedral complex Σ *is a collection of polyhedra such that if* $\sigma \in \Sigma$*, then every face of* σ *is an element of* Σ *and if* $\sigma, \tau \in \Sigma$*, then* $\sigma \cap \tau \in \Sigma$*.*

Tropical hypersurfaces are polyhedral complexes where every top-dimensional cell has dimension $n-1$ [MS15].

It is a fundamental fact of tropical geometry that the tropical hypersurface of a tropical polynomial is determined by the polyhedral geometry of the coefficients.

Remark 2.2.1. Let $p(x) = \sum_{j\in[D]} p_j \oplus x^{\oplus w^{(j)}}$ be a tropical polynomial in the variables $x = (x_1, x_2, \ldots, x_n)$. The *Newton Polytope* of p is the set

$$\mathcal{N}(p) = \text{conv}(\{w^{(j)} \mid j \in [D]\}).$$

One obtains a *polyhedral subdivision* of $\mathcal{N}(p)$ by constructing the polyhedral set

$$\text{conv}(\{(w^{(j)}, p_j) \mid j \in [D]\}) \subseteq \mathbb{R}^n \times \mathbb{R}$$

and projecting the *upper faces*, those faces whose outward normal vectors have positive last component, onto $\mathcal{N}(p)$. The tropical hypersurface $\mathcal{V}(p)$ is then dual to this subdivision (i.e. k dimensional regions in the subdivision correspond to $n-k$ dimensional polyhedra in the hypersurface and conversely).

Example 2.2.1 (Some Tropical Hypersurfaces)**.** 1. Consider the tropical line $p(x, y) = 0 \oplus x \oplus y$. The tropical line $\mathcal{V}(p)$ is shown in the first row of Figure 2.1

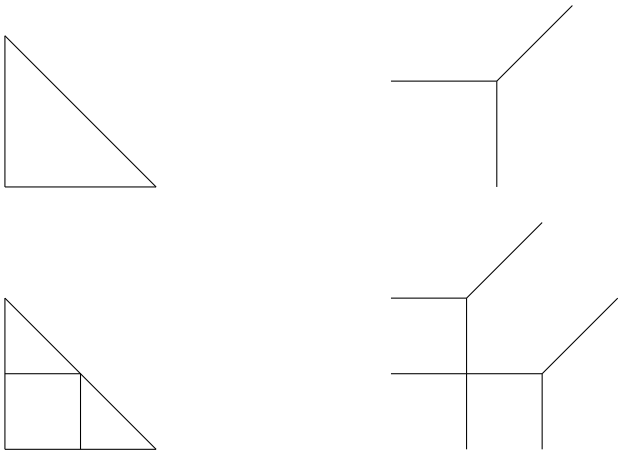


Figure 2.1: Newton Polytopes (left) and Tropical Hypersurfaces (right) for Example 2.2.1

2. Consider the tropical quadratic $p(x, y) = 1 \oplus (1 \odot x) \oplus (1 \odot y) \oplus (1 \odot x \odot y) \oplus (x^{\odot 2}) \oplus (y^{\odot 2})$. It's Newton Polytope and Hypersurface are shown in the second row of Figure 2.1.

$\diamond$

Remark 2.2.2. We will also work with with the nondifferentiability locus of tropical rational functions. Let $f(x) = p(x) - q(x)$ be a tropical rational function for some tropical polynomials p, q and set $X \subseteq \mathbb{R}^n$ to be the set of points for which f is nondifferentiable. Then $X \subseteq \mathcal{V}(p) \cup \mathcal{V}(q)$ and this containment can be proper. In particular, the complement $\mathbb{R}^n \backslash X$ need not consist of (open) polyhedra.

2.2.2 Tropical Linear Algebra

The set $\mathbb{T}^n$ of n-vectors with entries in $\mathbb{T}$ carries many properties analogous to linear algebraic properties of $\mathbb{R}^n$. First, $\mathbb{T}^n$ inherits an additive semigroup structure from $\mathbb{T}$ and a natural scalar multiplication.

Definition 2.2.5. *(Addition and scalar multiplication in $\mathbb{T}^n$) Let $a = (a_1, a_2, \ldots, a_n)^\top, b = (b_1, b_2, \ldots, b_n)^\top \in \mathbb{T}^n$ and $\lambda \in \mathbb{T}$. We define* vector addition

$$a \oplus b = (a_1 \oplus b_1, a_2 \oplus b_2, \ldots, a_n \oplus b_n)^\top = (\max(a_1, b_1), \max(a_2, b_2), \ldots, \max(a_n, b_n))^\top$$

and scalar multiplication

$$\lambda \odot a = (\lambda \odot a_1, \lambda \odot a_2, \ldots, \lambda \odot a_n)^\top.$$

Definition 2.2.6 ((max, +)-linear transformation)**.** *A function* $f : \mathbb{T}^n \to \mathbb{T}^m$ *is* (max, +)-linear *if for all* $a, b \in \mathbb{T}^n$ *and* $\lambda \in \mathbb{T}$*,*

$$f(a \oplus b) = f(a) \oplus f(b) \quad \text{and} \quad f(\lambda \odot a) = \lambda \odot f(a).$$

Remark 2.2.3. In linear algebra, linear maps $\mathbb{R}^n \to \mathbb{R}^m$ can be represented by matrix-vector multiplication. Similarly, (max, +)-linear maps can be represented by (max, +)-matrix vector products.

Definition 2.2.7 ((max, +) and (min, +) matrix-vector products)**.** *Let* $A = (a_{i,j}) \in \mathbb{T}^{m \times n}$ *be an* $m \times n$ *matrix with entries in* $\mathbb{T}$ *and* $x \in \mathbb{T}^n$*. We define the* (max, +) matrix-vector product *to be*

$$\begin{aligned} A \boxplus x &= \left(\bigoplus_{j=1}^{n} a_{1,j} \odot x_j, \bigoplus_{j=1}^{n} a_{2,j} \odot x_j, \ldots, \bigoplus_{j=1}^{n} a_{m,j} \odot x_j \right)^\top \\ &= \left(\max_{j \in [n]} (a_{1,j} + x_j), \max_{j \in [n]} (a_{2,j} + x_j), \ldots, \max_{j \in [n]} (a_{m,j} + x_j) \right)^\top. \end{aligned}$$

The dual notion is (min, +) matrix-vector multiplication*, denoted* $\boxplus'$*:*

$$A \boxplus' x = \left(\min_{j \in [n]} (a_{1,j} + x_j), \min_{j \in [n]} (a_{2,j} + x_j), \ldots, \min_{j \in [n]} (a_{m,j} + x_j) \right)^\top.$$

Theorem 2.2.8 ([CG79])**.** *Let* $A \in \mathbb{R}^{m \times n}$ *and* $b \in \mathbb{R}^m$ *and fix* $p \in \mathbb{N}$*. The optimal solution to the optimization problem*

$$\arg\min_x \|A \boxplus x - b\|_p \quad \text{s.t. } A \boxplus x \leqslant b$$

is given by $x^* = (-A)^\top \boxplus' b$*. The solution* x^* *is called the* principal subsolution *to the*

$(\max,+)$-*linear system* $A \boxplus x = b$. *The unconstrained* ∞*-norm problem*

$$\arg\min_x \|A \boxplus x - b\|_\infty$$

has optimal solution $x^* + \frac{1}{2}\|A \boxplus x^* - b\|_\infty$.

Remark 2.2.4. The principal solution $x^* = (-A)^\top \boxplus' b$ in Theorem 2.2.8 is analogous to the solution of the ordinary least squares problem in standard linear algebra using the normal equations. However, note that the principal solution $x^* = (-A)^\top \boxplus' b$ can be computed in time linear in m and n and does not suffer from the ill-conditioning of the normal equations.

2.3 Real Algebraic Geometry

In this section we provide the necessary results from real algebraic geometry, focusing on plane curves, sums of squares, and hyperbolic polynomials.

2.3.1 Real Plane Curves

Here we give a brief overview of real plane curves.

Definition 2.3.1. *A subset* $C \subseteq \mathbb{RP}^2$ *is a* real algebraic plane curve *of degree* d *if there exists a homogeneous polynomial* $f \in \mathbb{R}[x,y,z]_d$ *such that* $C = \mathcal{V}_\mathbb{R}(f)$.

Even if the polynomial f is irreducible, the corresponding real plane curve C may have multiple connected components.

Definition 2.3.2 (Oval)**.** *Let* $C_1 \subseteq C$ *be a connected component of a real plane curve. If* $\mathbb{RP}^2 \backslash C_1$ *has two connected components, exactly one of which is contractible, then* C_1 *is called an* oval. *If* C_1 *is an oval, then the contractible component of* $\mathbb{RP}^2 \backslash C_1$ *is called the* interior *of* C_1. *If* $C_1, C_2, \ldots, C_k$ *are ovals such that* C_i *is contained in the interior of* C_j *for* $1 \leqslant j < i \leqslant k$, *and this is a maximal set of such ovals, then* C_k *is an oval of* depth k *and* $C_1, C_2, \ldots, C_k$ *is a* nest *of ovals.*

Remark 2.3.1. There is a body of work in the real algebraic geometry literature which studies the numbers and arrangements of ovals; see e.g., [DIK12, PSV11]. We will be primarily concerned with the extremal case of degree d curves with $\lfloor \frac{d}{2} \rfloor$ nested ovals. Such curves correspond to *hyperbolic polynomials* (see Section 2.3.3).

We will be particularly interested in cases for which the real plane curve C has a nice description as the vanishing set of the determinant of linear combinations of matrices.

Definition 2.3.3 (Symmetric Determinantal Representation)**.** *A* symmetric determinantal representation *of a real algebraic plane curve C of degree d is an expression of C as $\mathcal{V}_{\mathbb{R}}(f)$ where f is given by*

$$f(x, y, z) = \det(xA + yB + zC)$$

for $A, B, C \in \mathcal{S}^d$. A symmetric determinantal representation is called definite *if* $\operatorname{span}_{\mathbb{R}}(A, B, C) \cap \operatorname{int}(\mathcal{S}^d_+) \neq \varnothing$.

2.3.2 Sums of Squares Polynomials

Central to real algebraic geometry is the theory of sums of squares polynomials. There are many excellent references for the topic, including [Mar08, Pow21, BPT13].

Definition 2.3.4 (Sum of Squares)**.** *A polynomial $f \in \mathbb{R}[x_1, x_2, \ldots, x_n]_{\leqslant 2d}$ of degree $2d$ is a sum of squares (SOS) if there exist $h_1, h_2, \ldots, h_r \in \mathbb{R}[x_1, x_2, \ldots, x_n]_{\leqslant d}$ such that $f = \sum_{i=1}^{r} h_i^2$.*

Remark 2.3.2. It is often more convenient to work with homogeneous polynomials. Recall that if $f \in \mathbb{R}[x_1, x_2, \ldots, x_n]_{\leqslant 2d}$ is a polynomial of degree $2d$, its homogenization is given by

$$\hat{f} = x_0^{2d} f\left(\frac{x_1}{x_0}, \frac{x_2}{x_0}, \ldots, \frac{x_n}{x_0}\right) \in \mathbb{R}[x_0, x_1, \ldots, x_n]_{2d}.$$

The polynomial f is a sum of squares if and only if its homogenization $\hat{f}$ is a sum of squares.

Definition 2.3.5 (Nonnegative and SOS cones)**.** *Let* $\mathbb{R}[x_1, x_2, \ldots, x_n]_{2d}$ *be the vector space of homogeneous polynomials of degree* $2d$ *in* n *variables. The* cone of nonnegative forms *is*

$$\mathcal{P}_{n,2d} = \{f \in \mathbb{R}[x_1, x_2, \ldots, x_n]_{2d} \mid f(x) \geqslant 0 \text{ for all } x \in \mathbb{R}^n\},$$

The cone of sums of squares *is*

$$\Sigma_{n,2d} = \left\{ f \in \mathbb{R}[x_1, x_2, \ldots, x_n] \;\middle|\; f = \sum_{i=1}^{r} h_i^2 \text{ for some } h_1, h_2, \ldots, h_r \in \mathbb{R}[x_1, x_2, \ldots, x_n]_d \right\}.$$

It is clear from the definitions that $\Sigma_{n,2d} \subseteq \mathcal{P}_{n,2d}$ for all $(n, 2d)$. Hilbert's Theorem resolves the cases where equality holds.

Theorem 2.3.6 (Hilbert)**.** $\mathcal{P}_{n,2d} = \Sigma_{n,2d}$ *if and only if*

- $d = 1$,
- $n = 2$, *or*
- $(n, 2d) = (3, 4)$.

There is also a relative version of nonnegativity and sums of squares. Let $X \subseteq \mathbb{P}^{n-1}$ be a variety defined over the real numbers with homogeneous coordinate ring $R = \mathbb{R}[x_1, x_2, \ldots, x_n]/I$.

Definition 2.3.7 (Nonnegative and SOS cones on a variety)**.** *The* cone of nonnegative quadratic forms *on a variety* $X \subseteq \mathbb{P}^{n-1}$ *with coordinate ring* R *is*

$$\mathcal{P}_X = \{f \in R_2 \mid f([x]) \geqslant 0 \text{ for all } [x] \in X(\mathbb{R})\}.$$[1]

The cone of SOS quadratic forms *on* X *is*

$$\Sigma_X = \left\{ f \in R_2 \;\middle|\; f = \sum_{i=1}^{r} h_i^2, \text{ for some } h_1, h_2, \ldots, h_r \in R_1 \right\}.$$

[1] Note that evaluation of quadratic forms on points $[x] \in X(\mathbb{R})$ is well-defined up to sign.

Remark 2.3.3. Note that it suffices to consider quadratic forms on a variety, as we can replace the variety X with the d-uple veronese embedding $\nu_d(X)$ if necessary.

Hilbert's Theorem is generalized to the case of varieties as follows.

Theorem 2.3.8 ([BSV16])**.** *Let $X \subseteq \mathbb{RP}^{n-1}$ be a nondegenerate totally real variety, that is, X is not contained in a hyperplane and $X(\mathbb{R})$ is Zariski dense in X. Then, $\mathcal{P}_X = \Sigma_X$ if and only if X is a variety of minimal degree:* $\deg(X) = \operatorname{codim}(X) + 1$.

Sums of squares polynomials are intimately related to semidefinite programming. In particular, the certification that a homoegneous polynomial $f \in \mathbb{R}[x_1, x_2, \ldots, x_n]_{2d}$ is a semidefinite programming feasibility question.

Proposition 2.3.9. *A homogeneous polynomial $f \in \mathbb{R}[x_1, x_2, \ldots, x_n]_{2d}$ is a sum of squares if and only if*

$$f(x) = [x]_d^\top Q [x]_d, \tag{2.3}$$

for some positive semidefinite matrix Q and $[x]_d$ a vector enumerating all monomials of degree d in the $x_1, x_2, \ldots, x_n$.

Similarly, a form $f \in R_2$ is a sum of squares if and only if there is a symmetric matrix Q and an element $g(x) = x^\top Z x \in I_2$ such that $Q + Z \succeq 0$ and

$$f(x) = x^\top (Q + Z) x. \tag{2.4}$$

Remark 2.3.4. Note that the equalities in (2.3) and (2.4) impose affine conditions on the entries of Q and therefore the existence of such a certificate is a semidefinite feasibility problem.

We conclude this subsection by connecting the classical S-lemma in optimization theory (see [PT07] for a survey) to the real algebraic geometry framework of Theorem 2.3.8. The S-lemma asserts that if Q_1 is a quadratic polynomial on $\mathbb{R}^n$ such that $Q_1(x) > 0$ for some

$x \in \mathbb{R}^n$, and if Q_2 is a quadratic polynomial on $\mathbb{R}^n$ such that $Q_2(x)$ is nonnegative for all x such that $Q_1(x) \geqslant 0$, then there is a positive semidefinite Z and a nonnegative constant c such that $Q_2 = Z + cQ_1$. Similarly, if $X = \mathcal{V}_{\mathbb{R}}(Q_1)$ is the real variety defined by a quadric, then X is a variety of minimal degree, so that any quadratic Q_2 which is nonnegative on X is a sum of squares modulo the ideal generated by Q_1. That is, there is a positive semidefinite Z such that $Q_2 = Z + cQ_1$. Note that since we are dealing with the variety defined by Q_1, the constant c is not restricted to be nonnegative. In chapter 5, we will see a similar situation for statements involving three quadratics.

2.3.3 Hyperbolic Polynomials

An important class of polynomials in real algebraic geometry is the class of hyperbolic polynomials, which share many geometric features with characteristic polynomials.

Definition 2.3.10 (Hyperbolic Polynomial)**.** *A homogeneous polynomial $p \in \mathbb{R}[x_1, x_2, \ldots, x_n]_d$ is* hyperbolic *with respect to a point $e \in \mathbb{R}^n$ if $p(e) \neq 0$ and for all $a \in \mathbb{R}^n$, the univariate polynomial $p(te - a) \in \mathbb{R}[t]$ is real-rooted.*

Example 2.3.1. The polynomial

$$p(x_1, x_2, \ldots, x_n) = \det(x_1 I + x_2 A^{(2)} + x_3 A^{(3)} + \ldots + x_n A^{(n)})$$

for some fixed symmetric matrices $A^{(2)}, A^{(3)}, \ldots, A^{(n)}$ is hyperbolic with respect to the point $(1, 0, 0, \ldots, 0)$. ◇

Topologically, the hypersurfaces corresponding to smooth hyperbolic polynomials are extremal in the sense that they contain the maximum number of nested ovaloids.

Proposition 2.3.11 (See e.g. [KPV15])**.** *If p is a smooth hyperbolic polynomial, then the real zero set $\mathcal{V}_{\mathbb{R}}(p) \subseteq \mathbb{RP}^{n-1}$ consists of $\lfloor \frac{\deg p}{2} \rfloor$ maximally nested ovaloids.*

Additionally, the zero sets of hyperbolic polynomials possess convex structure.

Theorem 2.3.12 ([Går59]). *If p is hyperbolic with respect to $e \in \mathbb{R}^n$, then the connected component of e in $\mathbb{R}^n \backslash \mathcal{V}_{\mathbb{R}}(p)$ is an (open) convex cone.*

We will be particularly interested in the case of hyperbolic plane curves. In this case, due to a theorem of Helton and Vinnikov [HV07], there is always a definite determinantal representation for a hyperbolic plane curve.

Theorem 2.3.13 (Helton Vinnikov Theorem [HV07]). *Let $p \in \mathbb{R}[x, y, z]$ be hyperbolic polynomial of degree d with respect to $(1, 0, 0)$. Then, there exist symmetric matrices $A, B \in \mathcal{S}^d$ such that*

$$p(x, y, z) = \det(xI + yA + zB).$$

Though the Helton-Vinnikov theorem ensures the existence of a definite determinantal representation of hyperbolic plane curves, there are three dimensional spaces of symmetric matrices which do not contain a positive definite matrix but have a hyperbolic determinant.

Example 2.3.2 ([PSV12, Example 5.2]). The polynomial

$$p(x, y, z) = \det\left(\begin{bmatrix} 25x & 0 & 12y - 32x & -60z \\ 0 & 25x & 10z & 24x + 16y \\ 12y - 32x & 10z & 6x + 4y & 0 \\ -60z & 24x + 16y & 0 & 6x + 4y \end{bmatrix}\right)$$

is hyperbolic with respect to $(1, 0, 0)$. However, there are no values of (x, y, z) which result in a positive definite matrix. The real algebraic curve corresponding to p is shown in Figure 2.2 ◇

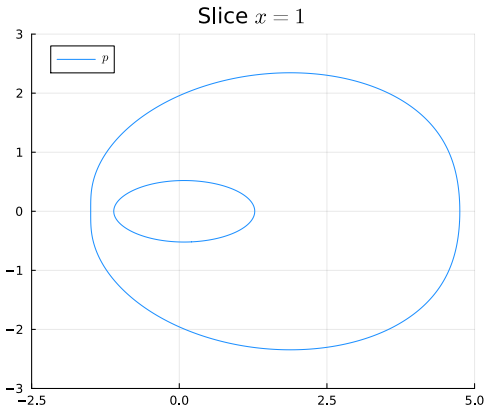


Figure 2.2: Hyperbolic curve from Example 2.3.2

2.4 Algebraic Topology and Homological Algebra

In this section, we review the necessary results from algebraic topology and homological algebra with a partiuclar emphasis on computations using spectral sequences. A standard refence text for algebraic topology is [Hat02] and a thorough introduction to the theory of spectral sequences is [McC01]. In this dissertation, all (co)homology will be taken over $\mathbb{Z}_2$.

We begin by recalling the definitions of the primary objects in homological algebra.

Definition 2.4.1 ((co)Chain Complex)**.** *A* chain complex *is a sequence of $\mathbb{Z}_2$ vector spaces V_i indexed by $i \in \mathbb{Z}$ and differentials $d_i : V_i \to V_{i-1}$ such that $d_{i-1} \circ d_i = 0$.*

$$\cdots \xleftarrow{d_{i-1}} V_{i-1} \xleftarrow{d_i} V_i \xleftarrow{d_{i+1}} V_{i+1} \xleftarrow{d_{i+2}} \cdots \qquad (V_i \xleftarrow{0} V_{i+1})$$

A cochain complex *is a sequence of $\mathbb{Z}_2$ vector spaces V_i indexed by $i \in \mathbb{Z}$ and differentials $d^i : V_i \to V_{i+1}$ such that $d^{i+1} \circ d^i = 0$.*

$$\cdots \xrightarrow{d^{i-2}} V_{i-1} \xrightarrow{d^{i-1}} V_i \xrightarrow{d^i} V_{i+1} \xrightarrow{d^{i+1}} \cdots \qquad (V_{i-1} \xrightarrow{0} V_{i+1})$$

Definition 2.4.2 ((co)homology of a (co)chain complex)**.** *The* homology *groups of a chain complex are $H_i = \frac{\ker(d_i)}{\mathrm{im}(d_{i+1})}$. The* cohomology *groups of a cochain complex are $H^i = \frac{\ker(d^i)}{\mathrm{im}(d^{i-1})}$.*

A primary motivation for homological algebra is to compute the homology of topological spaces. Loosely speaking, these are $\mathbb{Z}_2$ vector spaces whose dimensions count the number of "holes" in a topological space with n dimensional boundary.

Example 2.4.1 (Singular (co)homology of a topological space (see e.g. [Hat02]))**.** Let X be a topological space. For each $n \geqslant 0$, let Δ^n be the standard n-simplex and let C_n be the $\mathbb{Z}_2$ vector space[2] generated by continuous maps $\sigma : \Delta^n \to X$. Using the C_n, we get a chain complex by setting

$$d_n(\sum_{i=1}^{m} \alpha_i \sigma_i) = \sum_{i=1}^{m} \left(\alpha_i \sum_{j=0}^{n} \sigma_i([v_0, v_1, \ldots, \widehat{v_j}, \ldots, v_n]) \right) \in C_{n-1}.$$

where $[v_0, v_1, \ldots, \widehat{v_j}, \ldots, v_n]$ denotes the $n-1$ simplex obtained by removing the vertex v_j. The singular homology of the space X is then the homology of the resulting chain complex, i.e. $H_n(X) = \frac{\ker(d_n)}{\mathrm{im}(d_{n+1})}$.

Singular cohomology of a topological space, on the other hand, is constructed from the cochain complex which is dual to the chain complex which defines singular homology. That is, the singluar cohomology groups $H^n(X)$ are obtained by taking the cohomology of the cochain complex

$$\ldots \xrightarrow{d_{n-2}^\top} \mathrm{Hom}(C_{n-1}, \mathbb{Z}_2) \xrightarrow{d_{n-1}^\top} \mathrm{Hom}(C_n, \mathbb{Z}_2) \xrightarrow{d_n^\top} \mathrm{Hom}(C_{n+1}, \mathbb{Z}_2) \xrightarrow{d_{n+1}^\top} \ldots ,$$

where $d_i^\top$ is the dual map to d_i. It follows by the Universal Coefficient Theorem (see e.g. [Hat02]) that since we are working over the field $\mathbb{Z}_2$, $H_n(X) \simeq H^n(X)$. ◇

In Chapter 5, we will make heavy use of a *spectral sequence* to compute singular homology groups of topological spaces.

Definition 2.4.3 (Spectral Sequence)**.** *A* (first quadrant, cohomology) spectral sequnce

[2] if we were working over $\mathbb{Z}$, then we would take C_n to be the free abelian group generated by the σ

(E_r, d_r) *is a collection of* pages E_r *of* $\mathbb{Z}_2$*-vector spaces* $E_r^{i,j}$ *indexed by* $(i,j) \in \mathbb{Z}^2$ *and* differentials $d_r : E_r \to E_r$ *satisfying the following:*

- $E_r^{i,j} = 0$ *if* $i < 0$ *or* $j < 0$.
- *The differentials satisfy* $d_r^2 = 0$*; that is, we have the following cochain complex*

$$\cdots \longrightarrow E_r^{i-r,j+r-1} \xrightarrow{d_r^{i-r,j+r-1}} E_r^{i,j} \xrightarrow{d_r^{i,j}} E_r^{i+r,j-r+1} \longrightarrow \cdots$$

$$0$$

- *The* E_{r+1} *page has entries isomorphic to the homology of the differentials* d_r*. That is,*

$$E_{r+1}^{i,j} \simeq \frac{\ker(d_r^{i,j})}{\mathrm{im}(d_r^{i-r,j+r-1})}$$

Spectral sequences are usually displayed graphically in terms of their pages.

Example 2.4.2. The following E_2 page depicts the labeling and differential information for the E_2 page of a spectral sequence.

E_2

1	$E_2^{0,1}$	$E_2^{1,1}$	$E_2^{2,1}$
		$d_2^{0,1}$ ↘	
0	$E_2^{0,0}$	$E_2^{1,0}$	$E_2^{2,0}$
	0	1	2

If the map $d_2^{0,1} : E_2^{0,1} \to E_2^{2,0}$ is an isomporhism, and all other $d_2^{i,j}$ are zero, then the E_3 page of the sequence would have the form

E_3

1	0	$E_2^{1,1}$	$E_2^{2,1}$
0	$E_2^{0,0}$	$E_2^{1,0}$	0
	0	1	2

$\diamond$

Note that for a first quadrant cohomology spectral sequence (E_r, d_r), and fixed $i, j \geqslant 0$, there exists $k \geqslant 0$ such that $i-k$ and $j-k+1$ are both negative. In particular, the differential $d_k^{i,j}$ will have codomain 0 and the differential $d_k^{i-k,j+k-1}$ will have domain 0. Because $E_{k+1}^{i,j} \simeq \frac{\ker(d_k^{i,j})}{\mathrm{im}(d_k^{i-k,j+k-1})}$, this implies that $E_r^{i,j} \simeq E_k^{i,j}$ for all $r \geqslant k$. We therefore denote

$$E_k^{i,j} = E_{k+1}^{i,j} = \ldots =: E_\infty^{i,j}$$

Definition 2.4.4 (Convergence of Spectral Sequence)**.** *We say that a spectral sequence* (E_r, d_r) converges *to some graded* $\mathbb{Z}_2$ *vector space* H_**, denoted* $E_r \implies H_*$*, if for each* n*,*

$$H_n \simeq \bigoplus_{i+j=n} E_\infty^{i,j}.$$

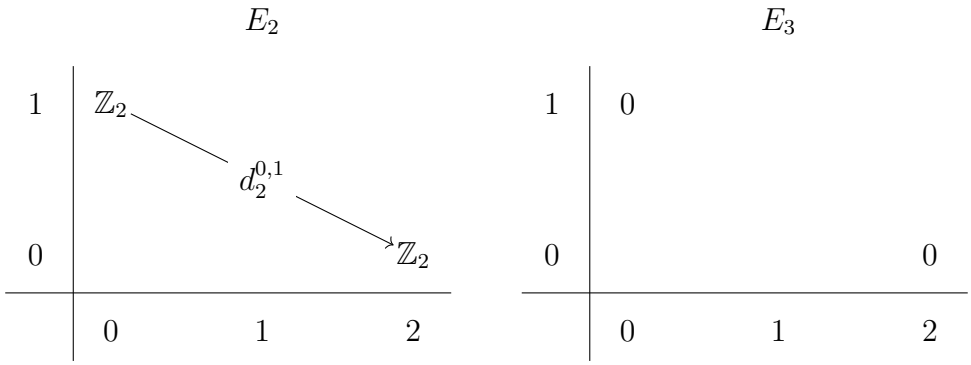
Example 2.4.3 ([BD25])**.** Consider a spectral sequence (E_r, d_r) with $E_r \implies$ such that

$$E_2^{i,j} \simeq \begin{cases} \mathbb{Z}_2 & (i,j) = (0,1) \text{ or } (1,0) \\ 0 & \text{otherwise} \end{cases}$$

Suppose that $d_2^{0,1}$ is an isomorphism. Then, we have that $E_3^{i,j} = 0$ for all $(i,j) \in \mathbb{Z}_2$ and therefore $E_\infty^{i,j} = 0$ for all $(i,j) \in \mathbb{Z}^2$. So,

$$H_1 \simeq E_\infty^{0,1} \oplus E_\infty^{1,0} \simeq 0 \oplus 0 \simeq 0.$$

Pictorally,



◇

2.5 Representation Theory

In this section, we review the necessary standard results about the representation theory of finite groups. A standard reference with complete proofs is [FH91].

Definition 2.5.1 (Group Representation)**.** *A* (linear) representation *of a group G is a group homomorphism $\rho : G \to GL(V)$ for some finite dimensional complex vector space V. The representation ρ is said to be* faithful *if the map ρ is injective.*

It is common to drop ρ from the notation and refer to the vector space V as a representation. Note also that if we fix a basis for the vector space V, then we can instead consider $GL_n(\mathbb{C})$, the group of invertible $n \times n$ matrices with complex entries, as the target of the map ρ. We can also consider maps between representations:

Definition 2.5.2 (Equivariant Map)**.** *Let $\rho : G \to GL(V)$ and $\eta : G \to GL(W)$ be two representations of a group G. A linear map $T : V \to W$ is said to be* equivariant (or G-linear) *if for all $v \in V$ and all $g \in G$, $T(\rho(g)v) = \eta(g)T(v)$. That is, the following diagram*

commutes:

$$\begin{array}{ccc} V & \xrightarrow{T} & W \\ \downarrow{\scriptstyle \rho(g)} & & \downarrow{\scriptstyle \eta(g)} \\ V & \xrightarrow{T} & W \end{array}$$

An invertible equivariant map whose inverse is also equivariant is called an *isomorphism* of representations.

Definition 2.5.3 (Subrepresentation, Irreducible Representation)**.** *Let $\rho : G \to GL(V)$ be a representation of G. A vector subspace $W \subseteq V$ is called a* subrepresentation *if $\rho(g)W \subseteq W$ for all $g \in G$. A representation V is called* irreducible *if there are no nontrivial subrepresentaions.*

Irreducible representations form the building blocks of representation theory. Precisely, every representation can be written as the direct sum of irreducible representations $V \simeq \bigoplus_{i=1}^{s} V_i$. Note that it may be the case that some irreducible representations appear with multiplicity greater than one. That is,

$$V \simeq \bigoplus_{i=1}^{s} \bigoplus_{j=1}^{\alpha_s} V_{i,j},$$

where for fixed i, we have $V_{i,j_1} \simeq V_{i,j_2}$ for any $1 \leqslant j_1, j_2 \leqslant \alpha_i$.

Example 2.5.1 (Decomposition into irreducibles)**.** Consider the symmetric group $S_3 = \langle \sigma, \tau | \sigma^2 = \tau^3 = 1, \sigma\tau = \tau^{-1}\sigma \rangle$ and the representation $\rho : S_3 \to GL_3(\mathbb{C})$ with

$$\rho(\sigma) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \rho(\tau) = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

The decomposition of $\mathbb{C}^3$ into irreducibles is then given by

$$\mathbb{C}^3 \simeq \text{span}\{(1,1,1)^\top\} \oplus \text{span}\{(1,-1,0)^\top, (0,1,-1)^\top\}$$

◇

Example 2.5.2 (Decomposition into irreducibles with multiplicity 2)**.** Consider the group $\mathbb{Z}_2 = \{1, \sigma\}$, and the representation $\rho : \mathbb{Z}_2 \to GL_3(\mathbb{C})$ with

$$\rho(\sigma) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Now, the decomposition of $\mathbb{C}^3$ into irreducibles is given by

$$\mathbb{C}^3 \simeq \text{span}\{(1,0,0)^\top\} \oplus \text{span}\{(0,1,1)^\top\} \oplus \text{span}\{(0,1,-1)^\top\}$$

Note that $\text{span}\{(1,0,0)^\top\}$ and $\text{span}\{(0,1,1)^\top\}$ are both the trivial representation, as σ acts on the identity on both of these spaces, and $\text{span}\{(0,1,-1)^\top\}$ is the alternating representation, since σ acts as mutliplication by -1 on this space. ◇

Lemma 2.5.4 (Schur's Lemma (see e.g., [FH91]))**.** *Let V and W be irreducible representations of a finite group G and $T : V \to W$ an equivariant map. Then,*

- *If V is not isomorphic to W then T is the zero map.*
- *If $V = W$, then $T = \lambda I$ for some $\lambda \in \mathbb{C}$ and I the identity map.*

Remark 2.5.1. It follows from Schur's Lemma and the existence of decompositions into irreducibles that if V is a representation, $V = \bigoplus_{i=1}^{s} \bigoplus_{j=1}^{\alpha_i} V_{i,j}$ a decomposition into irreducibles with $V_{i,j_1} \simeq V_{i,j_2}$ for all $i \in [s]$ and $j_1, j_2 \in [\alpha_i]$, and $T : V \to V$ is an equivariant map, then there exists a basis of V such that a matrix representative of T has a block structure. Moreover, the only nonzero blocks are multiples of the identity and correspond to blocks $V_{i,j_1} \to V_{i,j_2}$.

2.6 ReLU Neural Networks

As a final preliminary section, we provide a brief overview of and fix notation for fully connected ReLU neural networks. Loosely speaking, a neural network is a composition of affine functions and nonlinear *activation functions*. We will be interested in the piecewise linear ReLU activation function.

Definition 2.6.1 (ReLU Activation Function)**.** *The* Rectified Linear Unit (ReLU) *activation function is* $\sigma : \mathbb{R}^n \to \mathbb{R}^n$ *given by*

$$\sigma(x_1, x_2, \ldots, x_n) = \begin{bmatrix} \max(x_1, 0) & \max(x_2, 0) & \ldots & \max(x_n, 0) \end{bmatrix}^\top$$

Definition 2.6.2 (Fully Connected ReLU Neural Network)**.** *An* L-layer fully connected ReLU Neural Network *is a function* $\nu : \mathbb{R}^n \to \mathbb{R}^n$ *expressed as the composition*

$$\nu = \rho^{(L)} \circ \sigma^{(L-1)} \circ \rho^{(L-1)} \circ \sigma^{(L-2)} \circ \ldots \circ \sigma^{(1)} \circ \rho^{(1)},$$

where $\rho^{(\ell)} : \mathbb{R}^{n_\ell} \to \mathbb{R}^{n_{\ell+1}}$ *is an affine map* $\rho^{(\ell)}(x) = W^{(\ell)}x + b^{(\ell)}$ *and* $\sigma^{(\ell)} : \mathbb{R}^{n_{\ell+1}} \to \mathbb{R}^{n_\ell+1}$ *is the ReLU activation functions. The sequence* $(n_1, n_2, \ldots, n_L)$ *is the* architecture *of the network representing* ν*, the matrices* $W^{(\ell)}$ *are called the* weights *at layer* ℓ *and the vectors* $b^{(\ell)}$ *are the* biases *at layer* ℓ*.*

By construction, fully connected ReLU networks are continuous piecewise linear functions. Conversely, any continuous piecewise linear function can be expressed using a neural network with a bounded number of layers.

Theorem 2.6.3 ([ABMM18])**.** *Any continuous piecewise linear function* $\nu : \mathbb{R}^n \to \mathbb{R}$ *can be expressed as a fully connected ReLU neural network with at most* $\lceil \log_2(n+1) \rceil + 1$ *layers.*

Chapter 3

Regression with Tropical Rational Functions

The content of this chapter is based on joint work with Lars Ruthotto and appears in [DR24].

Tropical geometry has been connected to the study of fully connected feedforward ReLU Networks, starting with the work of Zhang, Naitzat, and Lim [ZNL18]. Indeed, the authors show that the sets of tropical rational functions, ReLU neural networks with integral weights, and continuous peicewise linear functions with integral slopes are equal. Moreover, the integrality constraints are not restrictive (from a theoretical point of view), as one can approximate real weights with rational weights and clear denominators. It has been useful in the theoretical analysis of such networks, particularly in counting the number of linear regions of a network. Since the full dimensional regions of the complement of a tropical hypersurface correspond to vertices of the induced subdivision of the Newton polytope (see Remark 2.2.1), the number of linear regions of a neural network can be bounded by looking at the coefficients of the network's representation as a tropical rational function. More work using tropical geometry to understand neural networks has appeared in [CM19, MCT21, SM19, SM20, TPS21, MRZ22]. A survey article from 2021 is [MCT21].

In another direction, researchers have applied the theory of principal solutions to tropical

linear systems (Theorem 2.2.8) to solve tropical polynomial regression problems [MT19, MT20]. To formally define the regression problem, let $W = \{w^{(1)}, w^{(2)}, \dots, w^{(D)}\} \subseteq \mathbb{R}^n$ of permissible exponents and let $\mathcal{D} = \{(\mathbf{x}^{(1)}, y^{(1)}), (\mathbf{x}^{(2)}, \mathbf{x}^{(2)}), \dots, (\mathbf{x}^{(N)}, y^{(N)})\} \subset \mathbb{R}^n \times \mathbb{R}$ be a dataset. Then, one can form the Vandermonde matrix $X \in \mathbb{R}^{N \times D}$ with $X_{i,j} = \langle w^{(j)}, \mathbf{x}^{(i)} \rangle$ and the right-hand side vector $y = \begin{bmatrix} y^{(1)} & y^{(2)} & \dots & y^{(N)} \end{bmatrix}^\top$. The tropical polynomial regression problem is then

$$\arg\min_{p} \|X \boxplus p - y\|_\infty. \tag{3.1}$$

By Theorem 2.2.8, the problem (3.1) has the analytical solution

$$\mathbf{p}^* = (-X)^\top \boxplus' y + \frac{1}{2} \left\| X \boxplus ((-X)^\top \boxplus' y) - y \right\|_\infty .$$

Therefore, the ∞-norm tropical polynomial regression problem can be solved quickly. Variants of tropical polynomial regression have been studied in [TM19, TTM22, Hoo19]

In this chapter, we utilize the tools from tropical linear systems and tropical polynomial regression to develop a heuristic for regression with tropical rational functions (Algorithm 1 below). Specifically, given $W, \mathcal{D}$, and X as above, we consider the problem

$$\arg\min_{p,q} \mathcal{L}(\mathbf{p}, \mathbf{q}) = \|\mathbf{X} \boxplus \mathbf{p} - \mathbf{X} \boxplus \mathbf{q} - \mathbf{y}\|_\infty . \tag{3.2}$$

Note that problem (3.2) is a continuous piecewise linear regression problem, and that the class of tropical rational functions is a difference of convex functions [Har59]. Such problems have received attention in multiple communities; see e.g., [MB09, KL21, TV12]

The proposed heuristic provides a step towards leveraging tropical algebraic structure for the training problem for fully connected ReLU networks.

Algorithm 1: Alternating Heuristic for Regression with Tropical Rational Functions

Input: Dataset $\mathcal{D} \subseteq \mathbb{R}^n \times \mathbb{R}$, Exponents $W \subseteq \mathbb{Z}^n$
Output: Vectors $\mathbf{p}, \mathbf{q} \in \mathbb{R}^D$ of coefficients of polynomials $p, q \in \mathbb{T}[\mathbf{x}]$ such that $p(\mathbf{x}) - q(\mathbf{x}) \approx y$ for $(\mathbf{x}, y) \in \mathcal{D}$.

Form $\mathbf{X} \in \mathbb{R}^{N \times |W|}$ with $X_{i,j} = \langle \mathbf{w}^{(j)}, \mathbf{x}^{(i)} \rangle$;
$\mathbf{p}^0, \mathbf{q}^0 \leftarrow -\infty$, $\mathbf{q}_0^0 \leftarrow -\text{mean}(\mathbf{y})$;
for $k = 1, 2, \ldots, k_{max}$ **do**
  $\mathbf{p}^k \leftarrow \arg\min_{\mathbf{p}} \|\mathbf{X} \boxplus \mathbf{p} - \mathbf{X} \boxplus \mathbf{q}^{k-1} - \mathbf{y}\|_\infty$;
  $\mathbf{q}^k \leftarrow \arg\min_{\mathbf{q}} \|\mathbf{X} \boxplus \mathbf{p}^k - \mathbf{X} \boxplus \mathbf{q} - \mathbf{y}\|_\infty$;
end

3.1 Alternating Method

Here we derive the alternating heuristic. Fix $\mathbf{q}$. Then, the problem

$$\min_{\mathbf{p}} \mathcal{L}(\mathbf{q}) = \min_{\mathbf{p}} \|X \boxplus \mathbf{p} - (X \boxplus \mathbf{q} + \mathbf{y})\|_\infty$$

is a tropical polynomial regression problem. So, by Theorem 2.2.8, there is an optimal solution

$$\mathbf{p}^*(\mathbf{q}) = (-X)^\top \boxplus' (X \boxplus \mathbf{q} + \mathbf{y}) + \frac{1}{2} \left\| X \boxplus \left((-X)^\top \boxplus' (X \boxplus \mathbf{q} + \mathbf{y}) \right) - (X \boxplus \mathbf{q} + \mathbf{y}) \right\|_\infty .$$

This solution can be computed quickly, as it only involves forward $(\max, +)$ and $(\min, +)$ operations and vector addition. Similarly, for a fixed $\mathbf{p}$, we can compute

$$\begin{aligned} \mathbf{q}^*(\mathbf{p}) &= \min_{\mathbf{q}} \|X \boxplus \mathbf{q} - (X \boxplus \mathbf{p} - \mathbf{y})\|_\infty \\ &= (-X)^\top \boxplus' (X \boxplus \mathbf{p} - \mathbf{y}) + \frac{1}{2} \left\| X \boxplus \left((-X)^\top \boxplus' (X \boxplus \mathbf{p} - \mathbf{y}) \right) - (X \boxplus \mathbf{p} - \mathbf{y}) \right\|_\infty . \end{aligned}$$

Therefore, given an initialization $\mathbf{q}^0$, one can perform alternating steps $\mathbf{p}^{k+1} = \mathbf{p}^*(\mathbf{q}^k)$ and $\mathbf{q}^{k+1} = \mathbf{q}^*(\mathbf{p}^{k+1})$. The procedure is summarized in Algorithm 1.

We begin with some immediate observations about the iterates produced by the method. First, the iterates produced by Algorithm 1 lead to nonincreasing values of the loss function. Fix $e^k = \mathcal{L}(\mathbf{p}^k, \mathbf{q}^k)$ to be the residual at iteration k.

Proposition 3.1.1. *The residuals e^k are nonincreasing. That is, $e^{k+1} \leqslant e^k$.*

Proof. Since $\mathbf{p}^{k+1} = \arg\min_{\mathbf{p}} \mathcal{L}(\mathbf{p}, \mathbf{q}^k)$ and $\mathbf{q}^{k+1} = \arg\min_{\mathbf{q}} \mathcal{L}(\mathbf{p}^{k+1}, \mathbf{q})$, it follows immediately that

$$\begin{aligned} e^k &= \|X \boxplus \mathbf{p}^k - X \boxplus \mathbf{q}^k - \mathbf{y}\|_\infty \\ &\geqslant \|X \boxplus \mathbf{p}^{k+1} - X \boxplus \mathbf{q}^k - \mathbf{y}\|_\infty \\ &\geqslant \|X \boxplus \mathbf{p}^{k+1} - X \boxplus \mathbf{q}^{k+1} - \mathbf{y}\|_\infty \\ &= e^k. \end{aligned}$$

□

Additionally, the difference in residual is bounded by a constant multiple of the norm of the update step. We use this to determine an effective stopping criterion in our experiments, where we observe that this bound is nonincreasing.

Proposition 3.1.2. *Set $\eta^k = \left\| \begin{bmatrix} \mathbf{p}^{k+1} & \mathbf{q}^{k+1} \end{bmatrix}^\top - \begin{bmatrix} \mathbf{p}^k & \mathbf{q}^k \end{bmatrix}^\top \right\|_\infty$. Then, $e^k - e^{k+1} \leqslant 2\eta^k$.*

Proof. Note that for each $j \in [D]$ and for any k, the inequalities

$$p_j^k - \eta^k \leqslant p_j^{k+1} \leqslant p_j^k + \eta^k \quad \text{and} \quad q_j^k - \eta^k \leqslant q_j^{k+1} \leqslant q_j^k + \eta^k$$

each hold. Moreover, since for each $i \in [N]$,

$$\max_{j \in [D]} (\langle \mathbf{w}^{(j)}, \mathbf{x}^{(i)} \rangle + p_j^k) \pm \eta^k = \max_{j \in [D]} (\langle \mathbf{w}^{(j)}, \mathbf{x}^{(i)} \rangle + p_j^k \pm \eta^k)$$

it follows that

$$\max_{j\in[D]}(\langle \mathbf{w}^{(j)}, \mathbf{x}^{(i)}\rangle + p_j^k) - \eta^k \leqslant \max_{j\in[D]}(\langle \mathbf{w}^{(j)}, \mathbf{x}^{(i)}\rangle + p_j^{k+1}) \leqslant \max_{j\in[D]}(\langle \mathbf{w}^{(j)}, \mathbf{x}^{(i)}\rangle + p_j^k) + \eta^k.$$

The analogous statements hold for $\mathbf{q}^k$. Denote by $P^k(\mathbf{x})$ and $Q^k(\mathbf{x})$ the tropical polynomials with coefficients $\mathbf{p}^k$ and $\mathbf{q}^k$, respectively. Set $\ell \in [N]$ to be such that the residual e^k is achieved at the datapoint $(\mathbf{x}^{(\ell)}, y^{(\ell)})$, i.e, $e^k = |P^k(\mathbf{x}^{(\ell)}) - Q^k(\mathbf{x}^{(\ell)}) - y^{(\ell)}|$. Then,

$$\begin{aligned}
e^k - e^{k+1} &= |P^k(\mathbf{x}^\ell) - Q^k(\mathbf{x}^\ell) - y^{(\ell)}| - \max_{i\in[N]} |P^{k+1}(\mathbf{x}^{(i)}) - Q^{k+1}(\mathbf{x}^{(i)}) - y^{(i)}| \\
&\leqslant |P^k(\mathbf{x}^\ell) - Q^k(\mathbf{x}^\ell) - y^{(\ell)}| - |P^{k+1}(\mathbf{x}^\ell) - Q^{k+1}(\mathbf{x}^\ell) - y^{(\ell)}| \\
&\leqslant |P^k(\mathbf{x}^{(\ell)}) - Q^k(\mathbf{x}^{(\ell)}) - P^{k+1}(\mathbf{x}^{(\ell)}) + Q^{k+1}(\mathbf{x}^{(\ell)})| \\
&\leqslant |P^k(\mathbf{x}^{(\ell)}) - P^{k+1}(\mathbf{x}^{(\ell)})| + |Q^k(\mathbf{x}^{(\ell)}) - Q^{k+1}(\mathbf{x}^{(\ell)})| \\
&\leqslant 2\eta^k.
\end{aligned}$$

□

3.2 Geometric Aspects

The iterates produced by Algorithm 1 have a nice connection to the geometry of the loss function. Specifically, we show below that the loss function $\mathcal{L}$ is a tropical rational function of the parameters $\mathbf{p}, \mathbf{q}$. Moreover, there always exists a minimizer $(\mathbf{p}^*, \mathbf{q}^*)$ in the nondifferentiability locus of $\mathcal{L}$ and the iterates $\mathbf{p}^k, \mathbf{q}^k$ produced by Algorithm 1 are such that $\nabla\mathcal{L}(\mathbf{p}^k, \mathbf{q}^k)$ does not exist.

Proposition 3.2.1. *The loss function $\mathcal{L}(\mathbf{p}, \mathbf{q})$ is a tropical rational function.*

Proof. We can expand

$$\begin{aligned}
L(\mathbf{p}, \mathbf{q}) &= \max_{i\in[N]} \left|P(\mathbf{x}^{(i)}) - Q(\mathbf{x}^{(i)}) - y^{(i)}\right| \\
&= \max_{i\in[N]} \left(\max\left(P(x^{(i)}) - Q(x^{(i)}) - y^{(i)}, -P(x^{(i)}) + Q(x^{(i)}) + y^{(i)}\right)\right).
\end{aligned}$$

Since the set of tropical rational functions form a semifield, it follows that L is a tropical rational function. □

As a tropical rational function, $\mathcal{L}$ is continuous and piecewise linear. So, we can study the optimization problem (3.2) using polyhedral geometry. To start, we show that the optimization problem (3.2) has a solution in the nondifferentiability locus of $\mathcal{L}$.

Proposition 3.2.2. *There is an optimal solution to* (3.2)*. Moreover, there is an optimal solution* $(\mathbf{p}^*, \mathbf{q}^*)$ *to* (3.2) *such that* $\nabla\mathcal{L}$ *does not exist at* $(\mathbf{p}^*, \mathbf{q}^*)$*.*

Proof. By Proposition 3.2.1, there are tropical polynomials g, h in $2D$ indeterminates such that

$$\mathcal{L}(\mathbf{p}, \mathbf{q}) = g(\mathbf{p}, \mathbf{q}) - h(\mathbf{p}, \mathbf{q}).$$

The nondifferentiability locus of $\mathcal{L}$ is then contained in the union $\mathcal{V}(g) \cup \mathcal{V}(h) \subseteq \mathbb{R}^{2D}$. The complement of $\Sigma = \mathcal{V}(g) \cup \mathcal{V}(h)$ in $\mathbb{R}^{2D}$ is a collection of open polyhedra. Label these $A_1, A_2, \ldots, A_s$.

For the first claim, note that the restriction of $\mathcal{L}$ to the closed polyhedron $\mathrm{cl}(A_i)$ is linear for each $i \in [s]$. Since $\mathcal{L}(\mathbf{p}, \mathbf{q}) \geqslant 0$, there is a minimum value $z_i = \min\{\mathcal{L}(\mathbf{p}, \mathbf{q}) \mid (\mathbf{p}, \mathbf{q}) \in \mathrm{cl}(A_i)\}$ for each $i \in [s]$. So, $\mathcal{L}$ achieves the minimum value $z = \min\{z_i \mid i \in [s]\}$.

For the second claim, note that since the restriction to of $\mathcal{L}$ to $\mathrm{cl}(A_i)$ is linear for each $i \in [s]$, there must be a minimizer $(\mathbf{p}^{(i)}, \mathbf{q}^{(i)}) \in \partial\mathrm{cl}(A_i)$ with $\mathcal{L}(\mathbf{p}^{(i)}, \mathbf{q}^{(i)}) = z_i$ for each $i \in [s]$. In particular, there is $(\hat{\mathbf{p}}, \hat{\mathbf{q}}) \in \Sigma = \bigcup_{i=1}^{s} \partial\mathrm{cl}(A_i)$ such that $\mathcal{L}(\hat{\mathbf{p}}, \hat{\mathbf{q}}) = z$. If $\nabla\mathcal{L}(\hat{\mathbf{p}}, \hat{\mathbf{q}})$ does not exist then we are done. Otherwise, $\nabla\mathcal{L}(\hat{\mathbf{p}}, \hat{\mathbf{q}}) = 0$. Relabel the A_i so that $(\hat{\mathbf{p}}, \hat{\mathbf{q}}) \in A_1$ and $(\hat{\mathbf{p}}, \hat{\mathbf{q}}) \in \bigcap_{i=1}^{k} \mathrm{cl}(A_i)$. Then, it follows that every point of $A = \bigcup_{i=1}^{k} \mathrm{cl}(A_i)$ is a minimizer of $\mathcal{L}$. Set B to be the smallest connected set containing A on which $\mathcal{L}$ is minimized. Note that $B \neq \mathbb{R}^{2D}$. Indeed if $B = \mathbb{R}^{2D}$, then $\mathcal{L}$ is constant, which is impossible: if $\mathbf{w}^{(1)} \in W$, $\mathbf{q}$ is fixed, and p_j is fixed for $j \geqslant 2$, then for sufficiently large values of p_1,

$$\mathcal{L}(\mathbf{p},\mathbf{q}) = \max_{i\in[N]} \left| \langle \mathbf{w}^{(1)}, \mathbf{x}^{(i)} \rangle + p_1 - \max_{j\in[D]} (\langle \mathbf{w}^{(j)}, \mathbf{x}^{(i)} + q_j) - y^{(j)} \right| = p_1 + C$$

for some constant C. Since $B \neq \mathbb{R}^{2D}$, any point (p^*, q^*) on $\partial B \neq \varnothing$ is a minimizer of $\mathcal{L}$ such that $\nabla\mathcal{L}(\mathbf{p}^*, \mathbf{q}^*)$ does not exist. □

To further the connection to the geometry of the loss function, we show that the iterates produced by Algorithm 1 are located in the nondifferentiability locus of $\mathcal{L}$. Combined with Propositions 3.1.1 and 3.2.2, this suggests that the alternating heuristic is a reasonable approach to searching for a global minimizer.

Proposition 3.2.3. *Let $\mathbf{p}^k, \mathbf{q}^k$ for $k \geqslant 1$ be iterates produced by Algorithm 1. Then, $\nabla\mathcal{L}(\mathbf{p}^k, \mathbf{q}^k)$ does not exist.*

Proof. We prove that if $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$, then

$$\mathbf{u}^* = (-A)^\top \boxplus' \mathbf{b} + \frac{1}{2} \| A \boxplus ((-A)^\top \boxplus' \mathbf{b}) - \mathbf{b} \|_\infty$$

is a nondifferentiable point of the residual function $\mathcal{R}(\mathbf{u}) = \|A \boxplus \mathbf{u} - \mathbf{b}\|_\infty$. The claim of the proposition will then follow since $\mathbf{p}^k = \arg\min_{\mathbf{p}} \mathcal{L}(\mathbf{p}, \mathbf{q}^{k-1})$ and $\mathbf{q}^k = \arg\min_{\mathbf{q}} \mathcal{L}(\mathbf{p}^k, \mathbf{q})$ have this form.

Suppose for the sake of a contradiction that $\nabla\mathcal{R}(\mathbf{u}^*)$ exists. Then, because $\mathbf{u}^*$ is a minimizer of $\mathcal{R}(\mathbf{u})$, it must be the case that $\nabla\mathcal{R}(\mathbf{u}) = 0$. Fix indices i, j such that

$$\mathcal{R}(\mathbf{u}^*) = \left| \max_\ell (a_{i,\ell} + u^*_\ell) - b_i \right| = \left| a_{i,j} + u^*_j - b_i \right|,$$

and let $K = \{k \mid a_{i,k} + u^*_k = \max_\ell (a_{i,\ell} + u^*_\ell)\}$ be the set of indices where the maximum is attained. Set $\mathbf{e}_K = \sum_{k\in K} e_k$ be the indicator vector with 1 in component k if $k \in K$ and 0 otherwise. Note that the previously fixed index j is an element of K. For $\epsilon > 0$ small enough that

$$a_{i,k} + u_k^* - \epsilon > a_{i,\ell} + u_\ell^* \text{ when } k \in K \text{ and } \ell \notin K,$$

there exists $c \in \{-1, 1\}$ such that

$$\mathcal{R}(\mathbf{u}^* + c\epsilon \mathbf{e}_K) \geqslant \left|a_{i,j} + u_j^* + c\epsilon - b_i\right| = \left|a_{i,j} + u_j^* - b_i\right| + \epsilon = \mathcal{R}(\mathbf{u}^*) + \epsilon.$$

That is, perturbing $\mathbf{u}^*$ along either $\pm\mathbf{e}_K$ will increase the residual by at least ϵ. But then,

$$\left|\frac{\mathcal{R}(\mathbf{u}^* + c\epsilon \mathbf{e}_K) - \mathcal{R}(\mathbf{u}^*)}{\epsilon}\right| \geqslant 1.$$

Since the difference quotient is bounded away from 0 for small $\epsilon > 0$, this contradicts the hypothesis that $\nabla\mathcal{R}(\mathbf{u}^*) = 0$. □

The geometric properties discussed above lead to concrete conditions on the dataset. In particular, the sources of nondifferentiability in the loss function are in the infinity norm and the nondifferentiability of the tropical polynomials P and Q. This allows us to connect the geometry of the optimization problem to the dataset.

Proposition 3.2.4. *There exists a minimizer $(\mathbf{p}^*, \mathbf{q}^*)$ of $\mathcal{L}$ such that at least one of the following holds:*

1. *There is an index $i \in [N]$ such that $\mathbf{x}^{(i)} \in \mathcal{V}(\mathbf{p}^*) \cup \mathcal{V}(\mathbf{q}^*)$.*
2. *There are $i, j \in [N]$ with $i \neq j$ such that*

$$\mathcal{L}(\mathbf{p}^*, \mathbf{q}^*) = \left|P^*(\mathbf{x}^{(i)}) - Q^*(\mathbf{x}^{(i)}) - y^{(i)}\right| = \left|P^*(\mathbf{x}^{(j)}) - Q^*(\mathbf{x}^{(j)}) - y^{(j)}\right|.$$

Proof. We will prove the contrapositive. Let $(\mathbf{p}^*, \mathbf{q}^*)$ be a minimizer of $\mathcal{L}$ such that $\nabla\mathcal{L}(\mathbf{p}^*, \mathbf{q}^*)$ does not exist and suppose that neither condition holds. Then, there is $i \in [N]$ such that

$$\mathcal{L}(\mathbf{p}^*, \mathbf{q}^*) = \left|P^*(\mathbf{x}^{(i)}) - Q^*(\mathbf{x}^{(i)}) - y^{(i)}\right| > \left|P^*(\mathbf{x}^{(j)}) - Q^*(\mathbf{x}^{(j)}) - y^{(j)}\right| \text{ for all } j \neq i.$$

There is an open neighborhood $U \subseteq \mathbb{R}^{2D}$ of $(\mathbf{p}^*, \mathbf{q}^*)$ such that

$$\mathcal{L}(\mathbf{p}, \mathbf{q}) = \left|\max_{j\in[D]}(\mathbf{p}_j + \langle \mathbf{w}^{(j)}, \mathbf{x}^{(i)}\rangle) - \max_{j\in[D]}(q_j + \langle \mathbf{w}^{(j)}, \mathbf{x}^{(i)}\rangle) - y^{(i)}\right|$$

for all $(\mathbf{p}, \mathbf{q}) \in U$. Since $\mathcal{L}(\mathbf{p}^*, \mathbf{q}^*) > 0$, we show that if $\mathbf{x}^{(i)} \notin \mathcal{V}(\mathbf{p}^*) \cup \mathcal{V}(\mathbf{q}^*)$, then the evaluation map $(\mathbf{p}, \mathbf{q}) \mapsto P(\mathbf{x}^{(i)}) - Q(\mathbf{x}^{(i)})$ is differentiable, contradicting the construction of $(\mathbf{p}^*, \mathbf{q}^*)$. If $\mathbf{x}^{(i)} \notin \mathcal{V}(\mathbf{p}^*) \cup \mathcal{V}(\mathbf{q}^*)$, then there are $j, k \in [D]$ such that

$$p_j^* + \langle \mathbf{w}^{(j)}, \mathbf{x}^{(i)}\rangle > p_\ell^* + \langle \mathbf{w}^{(\ell)}, \mathbf{x}^{(i)}\rangle \text{ for } \ell \neq j$$

and

$$q_k^* + \langle \mathbf{w}^{(k)}, \mathbf{x}^{(i)}\rangle > q_\ell^* + \langle \mathbf{w}^{(\ell)}, \mathbf{x}^{(i)}\rangle \text{ for } \ell \neq k.$$

So, restricting U to a smaller open neighborhood of $(\mathbf{p}^*, \mathbf{q}^*)$ if necessary, we see that

$$\mathcal{L}|_U(\mathbf{p}, \mathbf{q}) = |p_j + \langle \mathbf{w}^{(j)}, \mathbf{x}^{(i)}\rangle - q_k + \langle \mathbf{w}^{(k)}, \mathbf{x}^{(i)}\rangle - y^{(i)}|.$$

Since $\mathcal{L} > 0$, this is an affine function of $(\mathbf{p}, \mathbf{q})$ near $(\mathbf{p}^*, \mathbf{q}^*)$, and therefore differentiable.

□

Additionally, the sublevel sets of the loss function $\mathcal{L}$ provide a connection to *tropical convexity*.

Proposition 3.2.5. *Let $\delta > 0$. Then, determining the existence of $\mathbf{p}, \mathbf{q} \in \mathbb{R}^D$ with $\mathcal{L}(\mathbf{p}, \mathbf{q}) \leqslant \delta$ is a tropical linear programming feasibility problem.*

Proof. This is an application of the standard technique of linear programming for solving minmax problems (see e.g. [BV04, Section 1.2.2]). The problem $\min_{\mathbf{p},\mathbf{q}} \mathcal{L}(\mathbf{p}, \mathbf{q})$ is equivalent

to the problem

$$\begin{aligned} &\min t \\ \text{s.t.}\ &\max_{i=1,2,\dots,D}(\langle \mathbf{w}^{(i)}, \mathbf{x}^{(j)} \rangle + p_i) - \max_{i=1,2,\dots,D}(\langle \mathbf{w}^{(i)}, \mathbf{x}^{(j)} \rangle + q_i) - y^{(j)} \leqslant t \ \forall j \\ &- \max_{i=1,2,\dots,D}(\langle \mathbf{w}^{(i)}, \mathbf{x}^{(j)} \rangle + p_i) + \max_{i=1,2,\dots,D}(\langle \mathbf{w}^{(i)}, \mathbf{x}^{(j)} \rangle + q_i) + y^{(j)} \leqslant t \ \forall j. \end{aligned} \tag{3.3}$$

Now, there is a feasible solution $(\mathbf{p}, \mathbf{q}, t)$ to (3.3) with $t \leqslant \delta$ if and only if the following system of tropical linear inequalities in the variables $\mathbf{p}, \mathbf{q}$ has a solution.

$$\begin{cases} \max_{i=1,2,\dots,D}(\langle \mathbf{w}^{(i)}, \mathbf{x}^{(j)} \rangle + p_i) \leqslant \max_{i=1,2,\dots,D}(\langle \mathbf{w}^{(i)}, \mathbf{x}^{(j)} \rangle + q_i + y^{(j)} + \delta) & \forall j \\ \max_{i=1,2,\dots,D}(\langle \mathbf{w}^{(i)}, \mathbf{x}^{(j)} \rangle + q_i + y^{(j)}) \leqslant \max_{i=1,2,\dots,D}(\langle \mathbf{w}^{(i)}, \mathbf{x}^{(j)} \rangle + p_i + \delta) & \forall j \end{cases}$$

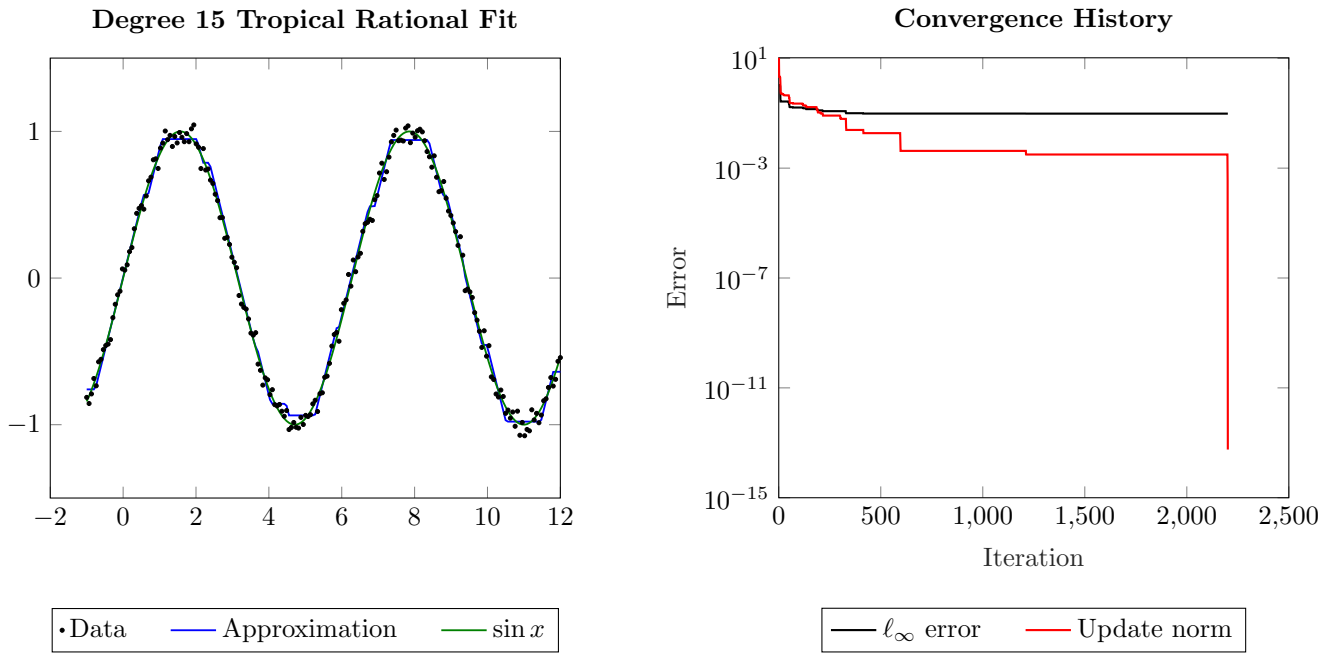
□

3.3 Numerical Results

This section is reproduced from [DR24]. In this section, we use Algorithm 1 for regression tasks and examine its convergence behavior empirically. We provide univariate, bivariate, and higher dimensional examples. In the univariate case we analyze the relationship between the degree hyperparameter and the error in the computed fit. In the bivariate case, we analyze the effect of precomposition with a scaling parameter c, that is, we study functions of the form $f(c\mathbf{x}) = p(c\mathbf{x}) - q(c\mathbf{x})$. For six variable functions, we examine the use of Algorithm 1 on data generated from tropical rational functions. We then present the performance of our approach on the existing datasets used by [KL21, MT20, RK15]. Finally, we present preliminary experiments using the output of Algorithm 1 to initialize ReLU neural networks. All Matlab and Python codes to reproduce our experiments can be found at

`https://github.com/Alex-Dunbar/Tropical-Data.git`.

Throughout this section, we say that an n-variate tropical rational function has degree d



(a) Tropical rational approximation

(b) Training loss and update norm

Figure 3.1: Results of applying Algorithm 1 with degree 15 tropical rational functions to noisy data from a sine curve. Figure 3.1a shows the training data, the approximation by a tropical rational function, and the function $\sin x$. The approximating function captures the general behavior of the dataset. Figure 3.1b shows the ℓ_∞ error $e^k = \|\mathbf{X} \boxplus \mathbf{p}^k - \mathbf{X} \boxplus \mathbf{q}^k - \mathbf{y}\|_\infty$ and the update norm $\eta^k = \| \begin{bmatrix} \mathbf{p}^{k+1} & \mathbf{q}^{k+1} \end{bmatrix}^\top - \begin{bmatrix} \mathbf{p}^k & \mathbf{q}^k \end{bmatrix}^\top \|_\infty$. Both the training loss and the update norm are nonincreasing and contain intervals on which they are nearly constant.

if $W = \{0, 1, \ldots, d\}^n$.

3.3.1 Univariate Data

We apply Algorithm 1 to a dataset consisting of 200 equally spaced points $x^{(i)} \in [-1, 12]$ and corresponding y values $y^{(i)} = \sin(x^{(i)}) + \epsilon^{(i)}$, where $\epsilon^{(i)}$ is drawn independently from a Gaussian distribution with mean 0 and standard deviation 0.05. Figure 3.1 shows an example, with $d = 15$. We use a stopping criterion of $\eta^k \leqslant 10^{-12}$. The infinity norm of the error and the infinity norm of the update step at each iteration are plotted in Figure 3.1b. Both the training loss and the update norm are nonincreasing and have regions on which they are constant.

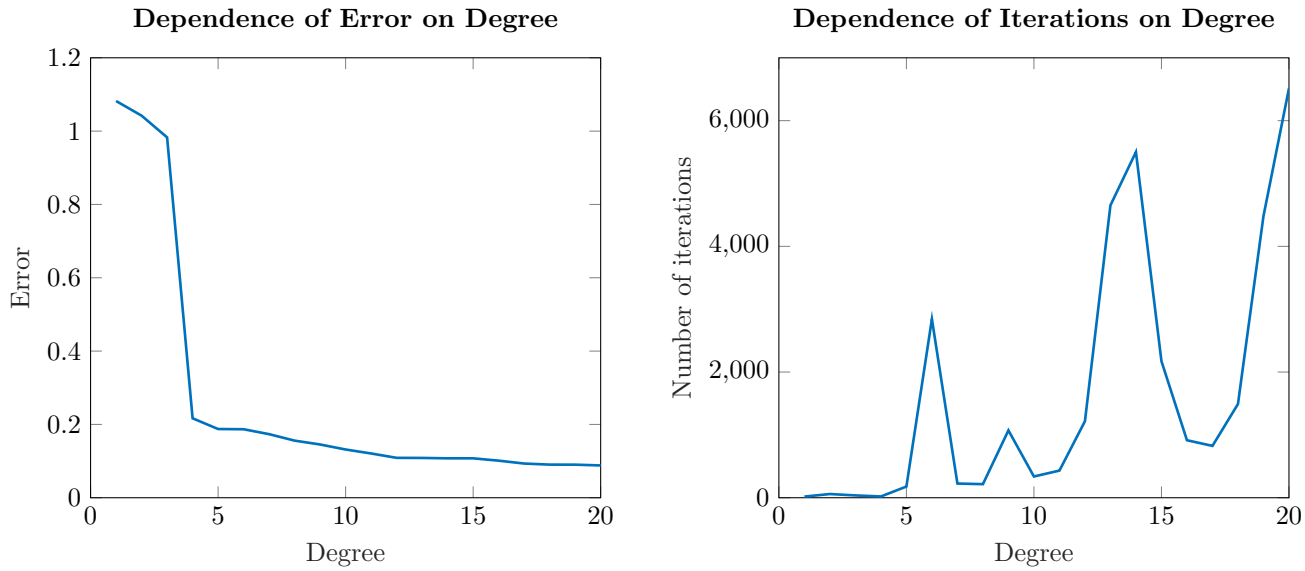


Figure 3.2: Dependence of error and number of iterations on degree of tropical rational function fit to noisy data from a sine curve. The error decreases monotonoically as a function of degree with a large drop at degree 5. The number of iterations needed to reach the stopping criterion of $\eta^k \leqslant 10^{-12}$ generally increases with the degree.

Effect of Degree Here, we investigate the relationship between the degree of tropical rational function and the error in the fit. Specifically, we generate a dataset as in the above example and use Algorithm 1 to fit a tropical rational function of degree d to the dataset for $d = 1, 2, \ldots, 20$. As a stopping criterion in Algorithm 1, we use $\eta^k \leqslant 10^{-12}$ or a maximum $k_{\max} = 10000$. Figure 3.2 shows the relationship between the degree of the rational function and the error in the fit. Note that the error decreases as a function of the degree with a large decrease in error when the degree is 5. The number of iterations needed to achieve the stopping criterion is generally increasing but is not monotonic.

3.3.2 Bivariate Data

We use the method to approximate the Matlab `peaks` dataset using degree 10 and degree 31 tropical rational functions and training until $\eta^k \leqslant 10^{-12}$. Explicitly, the `peaks` dataset consists of $2401 = 49^2$ equally spaced (x_1, x_2) pairs in $[-3, 3]^2$ and their evaluations

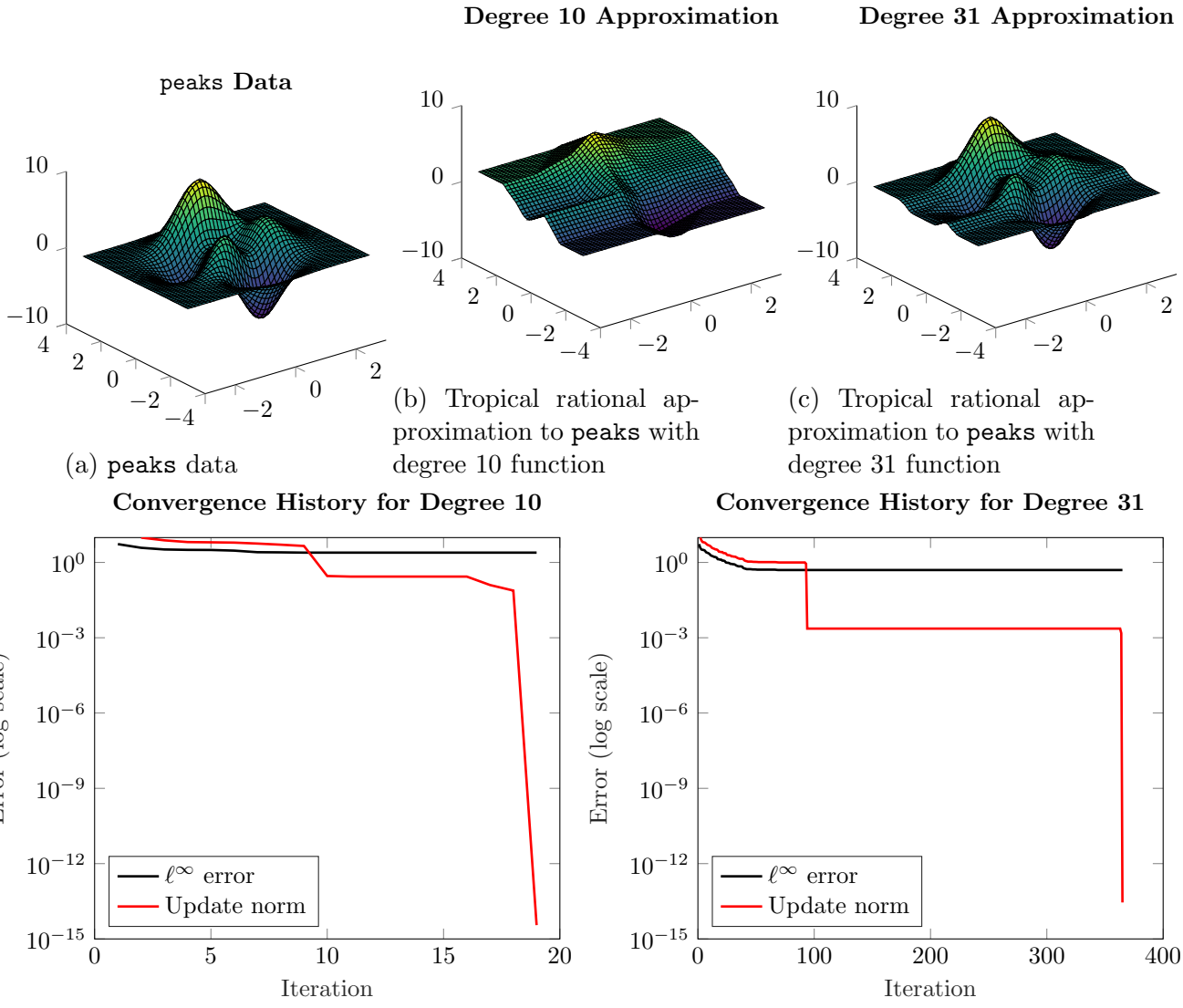
$$\texttt{peaks}(x_1, x_2) = 3(1-x_1)^2 e^{-x_1^2-(x_2+1)^2} - 10\left(\frac{x_1}{5} - x_1^3 - x_2^5\right) e^{-x_1^2-x_2^2} - \frac{1}{3}e^{-(x_1+1)^2-x_2^2}.$$

The fits and the error are shown below in Figure 3.3. Note that in both cases there is error in the regions on which the data is nearly constant despite the piecewise linear nature of the tropical rational functions. As in the univariate case, the training error and the update norm are nonincreasing and have regions where they are constant over many iterations.

Effect of Scaling Parameter In the above experiments, we directly fit a tropical rational function to the data. However, the results of [ZNL18] suggest that we should fit a function of the form $f(c\mathbf{x})$, where $c \in \mathbb{R}$ and f is a tropical rational function. To this end, we fit functions of the form $f(c\mathbf{x})$ for 21 equally spaced values of $c \in [1, 3]$ and f a tropical rational function of degree 35. For each value of c, we use a stopping criterion of $\eta^k \leqslant 10^{-12}$ or a maximum of 500 iterations of the alternating method described in Algorithm 1 to find a tropical rational function f. The dependence of the training error on c is shown in Figure 3.4 below. Note that the optimal value of c in this range is roughly 1.3. More generally, for fixed degree d, changing the value of c gives a trade-off between maximum slope and resolution between slopes. Due to this trade-off, there will, in general, be large errors for very large c because each affine piece of the tropical polynomials $p(c\mathbf{x})$ and $q(c\mathbf{x})$ will have large slopes. Conversely, there will be large errors for very small values of c because the slopes of the affine pieces of the polynomials $p(c\mathbf{x})$ and $q(c\mathbf{x})$ will be bounded.

3.3.3 Higher Dimensional Examples

We test Algorithm 1 on functions with many variables. These experiments suggest that the alternating minimization method is able to find solutions with low training loss. However, these solutions do not appear to generalize well, even on data generated from tropical rational functions.



(a) peaks data

(b) Tropical rational approximation to peaks with degree 10 function

(c) Tropical rational approximation to peaks with degree 31 function

(d) Training for degree 10 peaks regression

(e) Training for degree 31 peaks regression

Figure 3.3: Results of applying Algorithm 1 with degree 10 and 31 tropical rational functions to the peaks dataset. The resulting degree 31 function sketches the general behavior of the dataset (Figure 3.3c), while the degree 10 function fails to approximate the data (Figure 3.3b). Figures 3.3d and 3.3e display the the ℓ^∞ error $e^k = \|\mathbf{X} \boxplus \mathbf{p}^k - \mathbf{X} \boxplus \mathbf{q}^k - \mathbf{y}\|_\infty$ and the update norm $\eta^k = \| \begin{bmatrix} \mathbf{p}^{k+1} & \mathbf{q}^{k+1} \end{bmatrix}^\top - \begin{bmatrix} \mathbf{p}^k & \mathbf{q}^k \end{bmatrix}^\top \|_\infty$. For both degrees, the training loss and the update norm are each nonincreasing and contain intervals on which they are nearly constant.

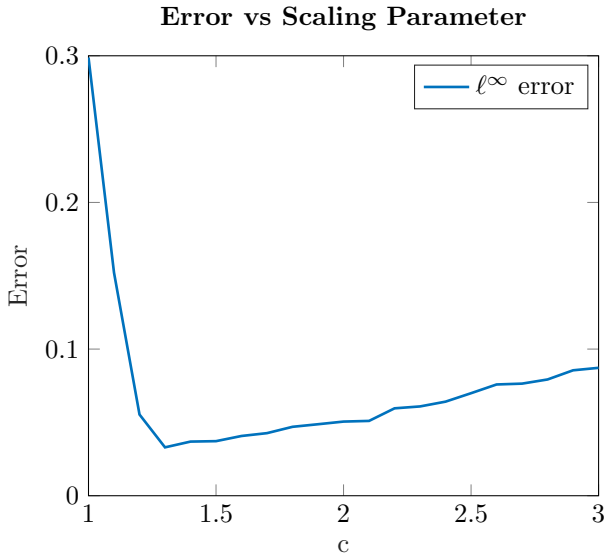


Figure 3.4: Error in approximation to the `peaks` dataset when using a degree $(35, 35)$ tropical rational function with inputs scaled by c. Here, the optimal value of c in the range $1 \leqslant c \leqslant 3$ is roughly 1.3 and gives a much lower training error than the function obtained as the output of Algorithm 1 with unscaled inputs.

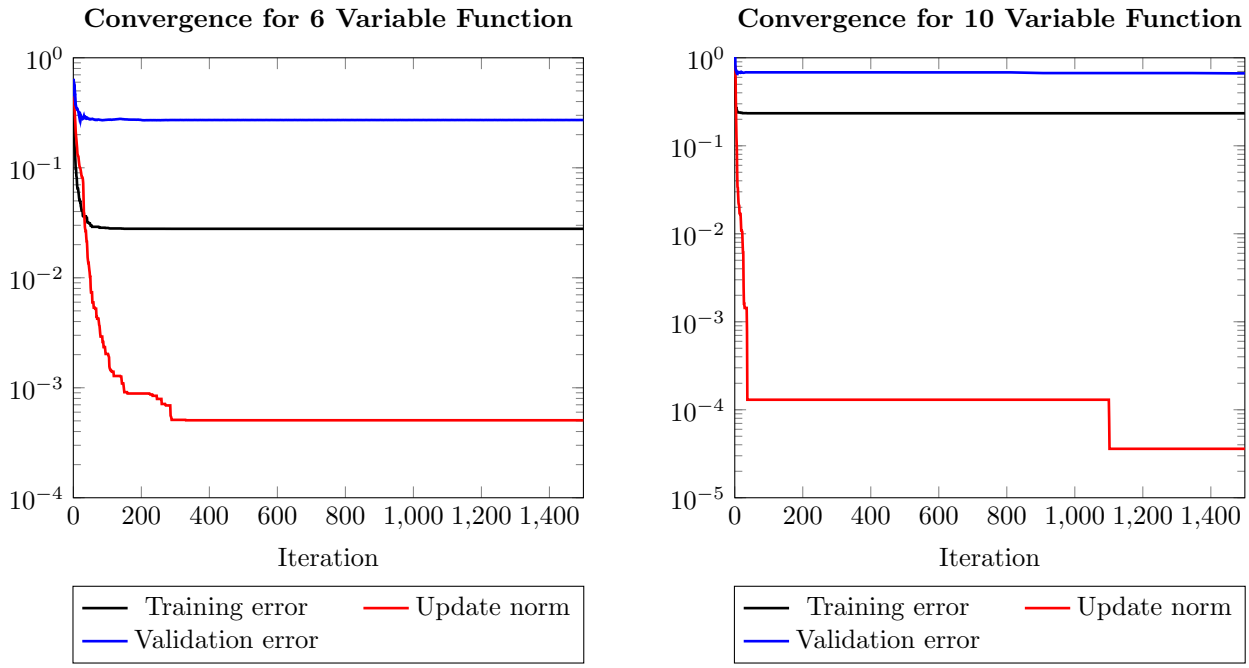
Regression on 6 Variable Function We fit a tropical rational function to the 6 variable function

$$g(\mathbf{x}) = x_1x_2x_3 + 2x_4x_5^2\sin(x_6^2)$$

on a training set consisting of $N = 10000$ points drawn uniformly at random from $[0, 1]^6$ and then test on a test set generated in the same way. Here, we fix the maximum degree of the numerator and denominator to be 3 for each variable and train until $\eta^k \leqslant 10^{-12}$ or for a maximum of 500 iterations. There are 8192 trainable parameters. The convergence behavior during training is shown in Figure 3.5a. The ℓ^∞ error on the test set is 0.2721, which is roughly 9.75 times the final training error of 0.0279.

Regression on 10 Variable Function We fit a tropical rational function to the 10 variable function

$$h(\mathbf{x}) = x_1x_2x_3 + 2x_4x_5^2\sin(x_6^2) - e^{x_7x_8x_9x_{10}}$$



(a) Degree 3 fit for 6 variable function with 10,000 data points

(b) Degree 1 fit for 10 variable function and 10,000 data points

Figure 3.5: Convergence for tropical rational approximation of 6 and 10 variable functions. The training error and update norm display similar behavior as in the low dimensional cases with regions on which they remain constant.

on a training set consisting of $N = 10000$ points drawn uniformly at random from $[0, 1]^{10}$ and then test on a test set generated in the same way. Here, we fix the maximum degree of the numerator and denominator to be 1 for each variable and train until $\eta^k \leqslant 10^{-12}$ or for a maximum of 500 iterations. There are 2048 trainable parameters. The convergence behavior during training is shown in Figure 3.5b. The ℓ^∞ error on the test set is 0.6828, which is roughly 2.9 times the final training error of 0.2342.

Data Generated from Tropical Rational Functions Here, we investigate the use of Algorithm 1 on data generated by tropical rational functions. Specifically, for $n = 6$ we investigate the use of Algorithm 1 for the recovery of a tropical rational function of degrees 1 through 5 (i.e. $W_d = \{0, 1, 2, \ldots, d\}^6$ for $1 \leqslant d \leqslant 5$). For each trial we generate a tropical rational function with coefficients sampled uniformly at random from $[-5, 5]$ as well as training and validation datasets of $N = 10000$ points sampled uniformly at random from

Degree	Relative Training Error	Relative Validation Error
1	2.372×10^{-15}	0.1271
2	5.869×10^{-15}	0.2019
3	9.108×10^{-15}	0.2869
4	1.286×10^{-14}	0.3631
5	9.373×10^{-6}	0.3598

Table 3.1: Average training and validation error on data generated from 6 variable tropical rational functions. For each degree, the training loss is low, but the validation error is high and increasing as a function of degree.

$[-5, 5]^6$. We then fit a tropical rational function $\hat{f}$ of the same degree using Algorithm 1 with a stopping criterion of $\eta^k \leqslant 10^{-8}$ or a maximum of 1000 iterations. In degrees at most 4, the method reached the stopping criterion in fewer than 1000 iterations for each trial. For degree 5, the method terminated after reaching 1000 iterations in 3 trials. In this experiment, $\mathbf{p}^0$ and $\mathbf{q}^0$ are initialized with entries drawn uniformly at random from $[-5, 5]^6$. Table 3.1 shows the average relative training and validation loss $\|\hat{f}(\mathbf{x}) - \mathbf{y}\|_\infty / \|\mathbf{y}\|_\infty$ across the five trials in each degree. Here, the training loss is low, indicating that Algorithm 1 finds a near optimal solution. However, the validation loss is high and increasing as a function of the degree. This indicates that when run to completion, Algorithm 1 solves the optimization problem (3.2) well. However, the higher validation errors suggest that the solution to (3.2) is nonunique. In particular, Algorithm 1 does not necessarily recover the coefficients of the tropical rational function used to generate the data.

3.3.4 Performance on Existing Datasets

In this section, we test the performance of our method on datasets generated from convex functions presented in [MT20] and datasets generated from nonconvex functions presented in [KL21, RK15].

Convex Functions Here, we use datasets from [MT20] and demonstrate that Algorithm 1, as a generalization of tropical polynomial regression, can be used to approximate convex

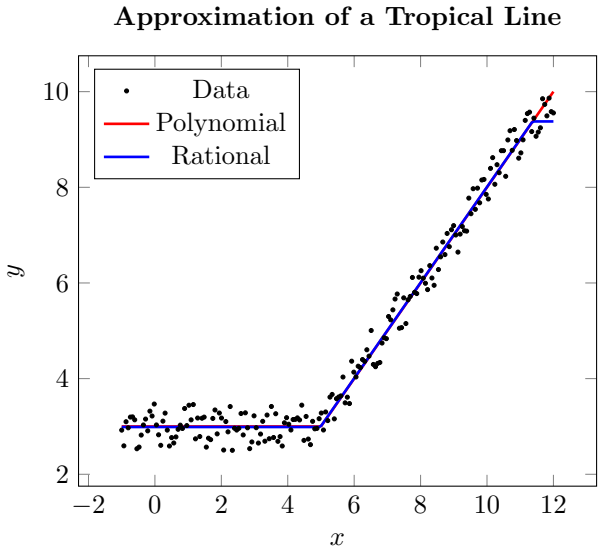


Figure 3.6: Approximation of a tropical line using tropical polynomial regression and Algorithm 1. Note that the two fits diverge near $x = 11$, where the tropical rational approximation becomes nonconvex.

functions. First, we generate data points $(x^{(i)}, g(x^{(i)}))$, where the $x^{(i)}$ are 200 equally spaced points on the interval $[-1, 12]$ and $g(x^{(i)}) = \max(3, x^{(i)} - 2) + \epsilon^{(i)}$ is a tropical line, where $\epsilon^{(i)}$ is drawn independently from a uniform distribution on $[-0.5, 0.5]$. We fit a tropical rational function and a tropical polynomial with $W = \{0, 1\}$. The results are plotted in Figure 3.6. Note that there is a deviation between the two approximations near $x = 11$, where the tropical rational approximation becomes nonconvex.

Next, we generate 500 pairs $(x_1^{(i)}, x_2^{(i)}) \in [-1, 1]^2$ uniformly at random and set $y^{(i)} = (x_1^{(i)})^2 + (x_2^{(i)})^2 + \epsilon^{(i)}$, where the $\epsilon^{(i)}$ are drawn independently from a normal distribution with mean 0 and variance 0.25^2. We then fit tropical rational functions of degrees $d = 1, 2, \ldots, 6$ to the data and record the error. Figure 3.7 shows the average and worst error across 25 such trials as well as the results reported in [MT20, Table 1]. Note that in the experimental setup of [MT20], tropical polynomials are fit to the data where the monomials are chosen via k-means clustering. In our setup, exponents for the numerator and denominator polynomials are $\{0, 1, \ldots, d_{max}\}^2$. It appears that the approach in [MT20] yields a lower error in the low-parameter setting, while the rational regression leads to lower training error in the high-

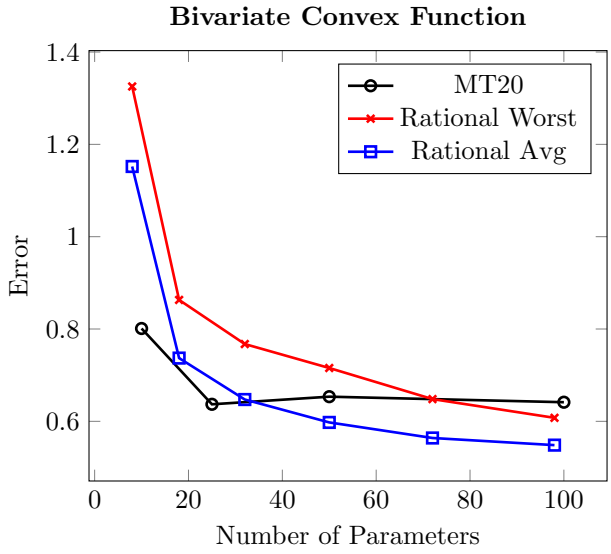


Figure 3.7: Errors in tropical approximations to a bivariate dataset generated from a convex function. Note that despite a data-blind choice of exponents, in the large parameter setting, tropical rational regression using Algorithm 1 produces approximations with lower training error than that reported in [MT20, Table 1], which used tropical polynomials with exponents chosen via k-means clustering.

parameter setting despite the data-independent monomial selection.

Nonconvex Functions Here, we test the performance of Algorithm 1 on nonconvex functions tested by [KL21, RK15]. Specifically, we consider the functions

$$\begin{aligned} g_1(x_1,x_2) &= x_1^2 - x_2^2 && \text{for } (x_1,x_2) \in [0.5,7.5]\times[0.5,3.5], \\ g_2(x_1,x_2) &= x_2^2\frac{\sin(x_1)}{x_1} && \text{for } (x_1,x_2) \in [1,3]\times[1,2], \\ \text{and } g_3(x_1,x_2) &= \exp(-10(x_1^2-x_2^2)^2) && \text{for } (x_1,x_2) \in [1,2]\times[1,2]. \end{aligned}$$

For each function, we study the effect of degree of tropical rational function and number of data points used on the error and solution time. Specifically, for $N \in \{10^2, 20^2, 50^2, 100^2\}$, we generate a dataset of N equally spaced gridpoints $(x_1^{(i)}, x_2^{(i)})$ of the function domain and their evaluations $g_j(x_1^{(i)}, x_2^{(i)})$ and use Algorithm 1 to fit tropical rational functions of degrees $1, 2, \ldots, 25$. The results are displayed in Figure 3.8. Note that the training errors

are comparable to those tested in [KL21].

3.3.5 ReLU Neural Network Initialization

Here we investigate the use of Algorithm 1 to initialize the weights of a ReLU neural network. The motivation for this approach is that the output of Algorithm 1 carries information about the training data. So, networks with weights initialized from the output of the tropical regression heuristic should start from a lower loss than those with weights drawn from a distribution that does not depend on the training data. Additionally, the computational cost of performing an iteration of Algorithm 1 is $\mathcal{O}(ND)$, which is comparable to the cost of an epoch of stochastic gradient descent for a network with D parameters. However, the networks that we are able to initialize have significantly more than D parameters. Despite the potential advantages of a tropical initialization, we find that such a scheme does not always lead to faster convergence or lower training or validation error. Moreover, the network architectures which we are able to initialize using a tropical rational function appear to have unstable training, even when initialized using well-known strategies.

In our experiments, we apply Algorithm 1 on data from the noisy sine curve and `peaks` datasets to generate approximations of the data, then use the output tropical rational function to initialize the weights of ReLU networks. The architecture of the initialized network is determined by the number of monomials in the tropical rational function f used to initialize the network. Specifically, the proof of [ZNL18, Theorem 5.4] describes one method to write a tropical rational function $f(\mathbf{x}) = p(\mathbf{x}) - q(\mathbf{x})$ as a ReLU neural network. If g and h are two tropical polynomials represented by neural networks ν and μ, respectively, then

$$(g \oplus h)(\mathbf{x}) = \sigma((\nu - \mu)(\mathbf{x})) + \sigma(\mu(\mathbf{x})) - \sigma(-\mu(\mathbf{x}))) = \begin{bmatrix} 1 & 1 & -1 \end{bmatrix} \sigma \left(\begin{bmatrix} \nu(\mathbf{x}) - \mu(\mathbf{x}) \\ \mu(\mathbf{x}) \\ -\mu(\mathbf{x}) \end{bmatrix} \right). \quad (3.4)$$

In particular, the expression (3.4) can be applied to the case in which $g(\mathbf{x}) = \mathbf{w}^\top \mathbf{x} + g_{\mathbf{w}}$ is a tropical monomial. This allows us to take the maximum of two networks by adding a layer and appropriately concatenating weight matrices in the hidden layers. In the resulting architecture, each hidden layer decreases in width. For example, a univariate degree 15 tropical rational function f can be represented via repeated applications of (3.4) as a neural network where the compositions are

$$\mathbb{R}^1 \to \mathbb{R}^{48} \to \mathbb{R}^{24} \to \mathbb{R}^{12} \to \mathbb{R}^6 \to \mathbb{R}^1.$$

For each dataset, we compare networks of the same architecture using the following initialization strategies:

- Repeated applications of (3.4) to the terms of the tropical rational function f output by Algorithm 1
- He initialization [HZRS15]
- Weights and biases drawn uniformly at random from $[-k, k]$[1], where $k = \sqrt{\frac{1}{\text{number of inputs}}}$ for each layer

All neural network parameter optimization is done in PyTorch version 1.11.0 using the Adam optimizer [KB14] to minimize the MSE loss.

Univariate Data

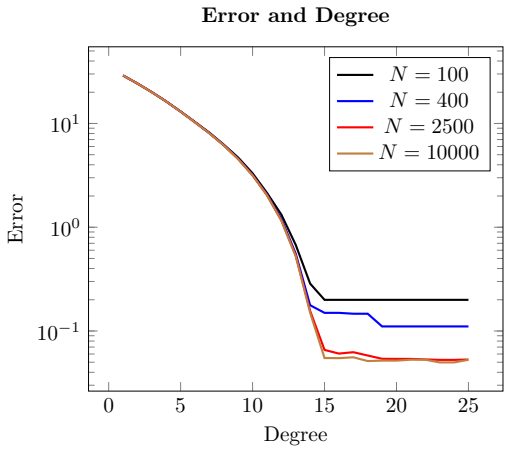
We use a degree 15 tropical rational function to initialize a neural network to fit the noisy sin curve from above. The test data consists of 200 pairs $(x^{(i)}, y^{(i)})$, where $x^{(i)}$ is randomly drawn points on the interval $[-1, 12]$ and $y^{(i)} = \sin(x^{(i)})$. The networks are trained for 1000 epochs with batches of size 64 and a learning rate of 5×10^{-6} for the tropical initialized network and 10^{-2} for the He-initialized and uniformly-initialized networks. We found choosing a

[1] This is the default initialization for linear layers in PyTorch version 1.11.0

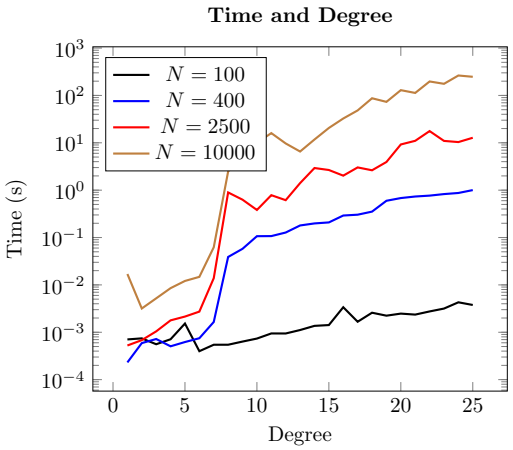
smaller learning rate for the tropical initialization important to prevent the optimization from reducing the accuracy of the model. Training and validation errors are shown in Figure 3.9. The network initialized from a tropical rational function has lower training and validation error than the network initialized using the other methods.

Bivariate Data

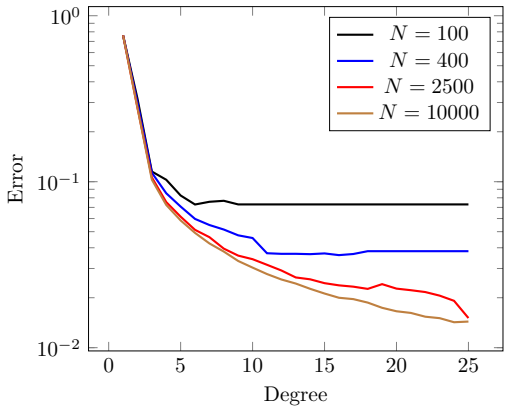
We use a degree 31 tropical rational function to initialize the `peaks` dataset using Algorithm 1 as the initialization. The networks are trained for 1000 epochs with a batch size of 64 and a learning rate of 10^{-4} for He-initialized and uniformly-initialized networks and 10^{-7} for the tropically initialized network. Results are shown in Figure 3.10. The networks initialized with He initialization and with uniform initialization reach lower training and validation errors than the tropically initialized network.



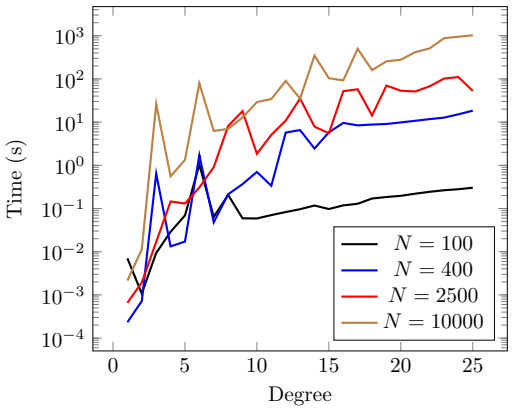
(a) $g_1(x_1, x_2) = x_1^2 - x_2^2$



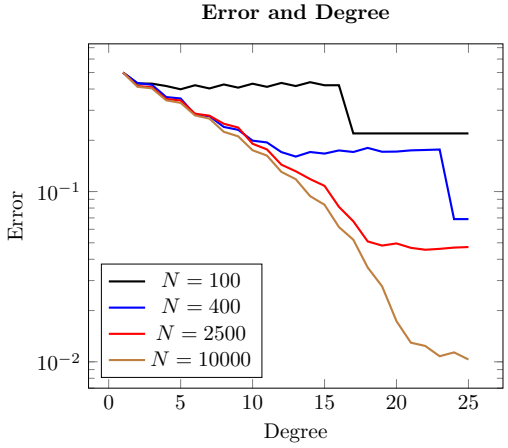
(b) $g_1(x_1, x_2) = x_1^2 - x_2^2$



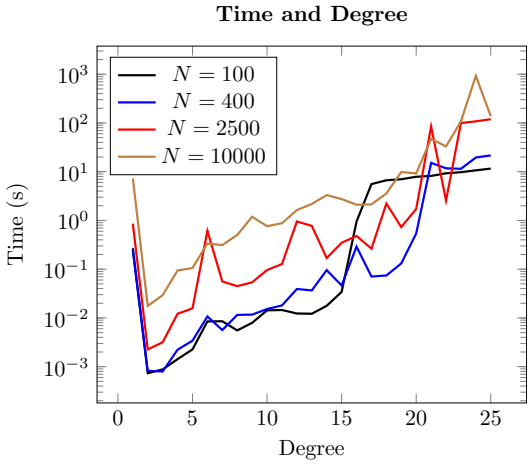
(c) $g_2(x_1, x_2) = x_2^2 \sin(x_1)/x_1$



(d) $g_2(x_1, x_2) = x_2^2 \sin(x_1)/x_1$



(e) $g_3(x_1, x_2) = \exp(-10(x_1^2 - x_2^2)^2)$



(f) $g_3(x_1, x_2) = \exp(-10(x_1^2 - x_2^2)^2)$

Figure 3.8: Performance of Algorithm 1 on datasets generated from nonconvex bivariate functions. The plots on the left display the relationship between error, degree, and number of sample points, while the figures on the right show the dependence of computation time on degree and number and sample points.

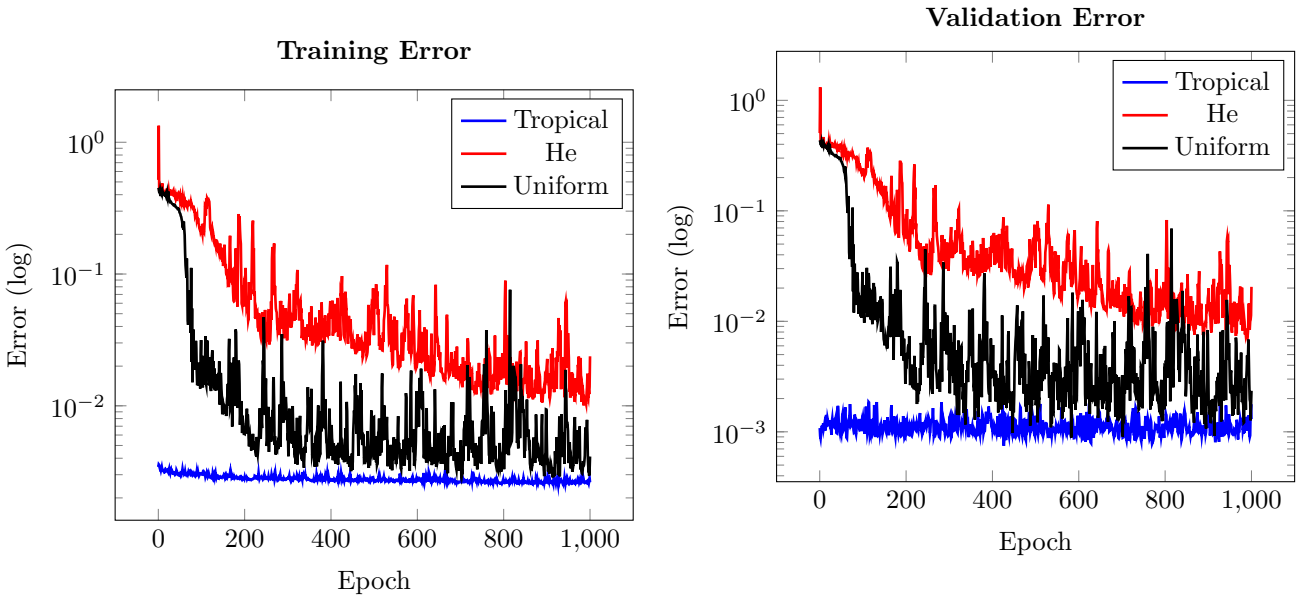


Figure 3.9: Training and validation errors for neural network fit to noisy sin data. The network initialized from a tropical rational approximation to the dataset starts and remains at lower training and validation losses than the networks initialized with the He Initialization [HZRS15]. and with uniform initialization

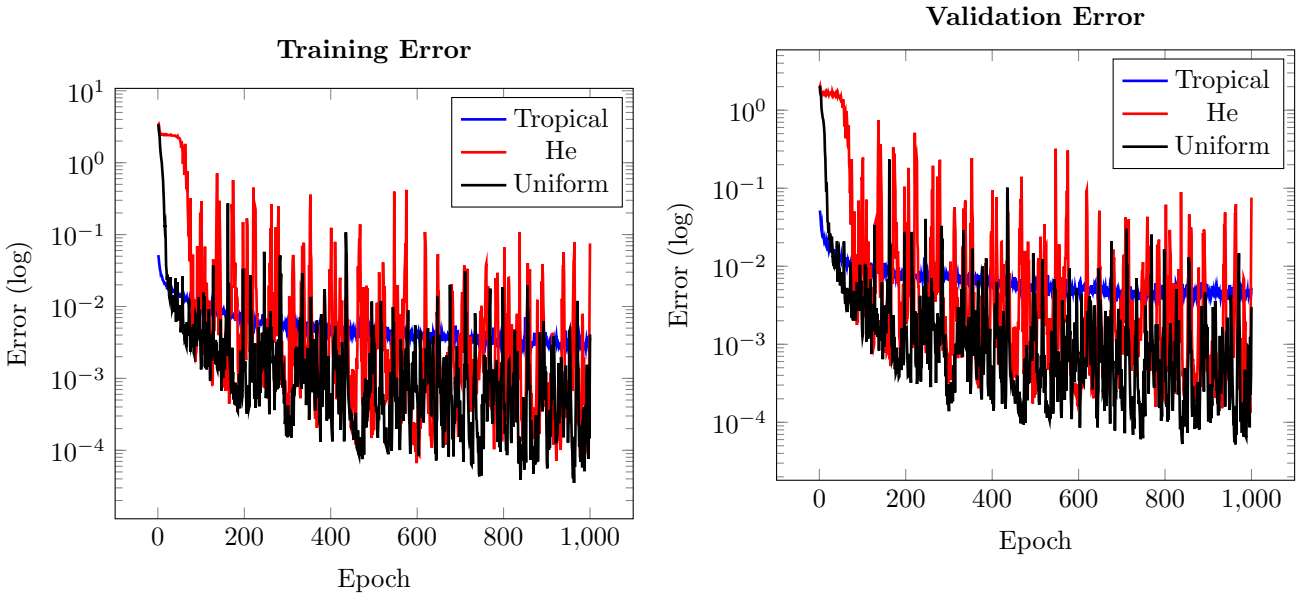


Figure 3.10: Training and validation errors for neural network fits to `peaks` data. The networks initialized the He Initialization [HZRS15] and uniform initialization reach lower training and validation errors than the tropically initialized network.

Chapter 4

Tensor-Tensor Products and Semidefinite Programs

This chapter is based on joint work with Elizabeth Newman [DN25]. The mathematical content is the same; however the presentation has been altered to align with the themes of this disseration. Additionally, [DN25] includes numerical experiments conducted by Elizabeth Newman that are omitted here.

In this chapter, we investigate a notion of positive semidefiniteness for third order tensors and define convex optimization problems over such tensors. Our point of view is derived from the $\star_M$ tensor-tensor product, a family of tensor-tensor products depending on an underlying invertible matrix M which allows for the generalization of many familiar linear algebraic properties to the tensor case [KKA15]. We show that under a reasonable notion of positive semidefiniteness for tensors with the $\star_M$ product, many desirable properties of PSD matrices carry over to the tensor case. Moreover, by leveraging the algebraic structure of the $\star_M$ product, we show that optimizing a linear functional over an affine slice of the cone of positive semidefinite tensors can be viewed as solving a block-diagonalized matrix semidefinite programing problem.

In another direction, we study the algebraic structure of the $\star_M$ product, connecting the

choice of matrix M to the representation theory of finite groups. In particular, we show that if the rows of the matrix M are chosen corresponding to a basis compatible with a decomposition into irreducible representations, then the $\star_M$ multiplication map is equivariant map for an explicit linear subspace of tensors. Combining the representation theoretic perspective on the $\star_M$ product with the observation that optimization over affine slices of the cone of positive semidefinite tensors is equivalent to a block-diagonalized semidefinite program leads to a natural connection with well-studied invariant semidefinite programs [GP04].

As applications of the tensor semidefinite programming framework, we phrase low-rank tensor completion problems as tensor semidefinite programs and provide a description of certain group invariant quadratic forms.

Notation: In this chapter, we will work with third-order tensors $\mathcal{A} \in \mathbb{K}^{n_1 \times n_2 \times n_3}$, where the field $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Tensors will be denoted with uppercase caligraphic letters $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \ldots)$. We will use `Matlab` notation for indexing, so that $\mathcal{A}_{i,j,k} \in \mathbb{K}$ is the scalar in the i^{th} row, j^{th} column of the k^{th} frontal slice of the tensor. Tensors of format $1 \times 1 \times n_3$ are called *tubes* and will be denoted using lowercase bold letters $(\mathbf{a}, \mathbf{b}, \mathbf{c}, \ldots)$. There are fixed vector space isomorphisms $\text{tube} : \mathbb{K}^{n_3} \to \mathbb{K}^{1 \times 1 \times n_3}$ and $\text{vec} : \mathbb{K}^{1 \times 1 \times n_3} \to \mathbb{K}^{n_3}$. In coordinates, we have that $\text{tube}(a)_{1,1,i} = a_i$ and $\text{vec}(\mathbf{a})_i = \mathbf{a}_{1,1,i}$. We write $a \equiv \mathbf{a}$ for $a \in \mathbb{K}^{n_3}$ and $\mathbf{a} \in \mathbb{K}^{1 \times 1 \times n_3}$ if $\text{tube}(a) = \mathbf{a}$.

4.1 The $\star_M$ Product of Tensors

The $\star_M$ product gives the structure of a commutative ring to the vector space of tubes. In this section, we recall the definition and some preliminary results on the $\star_M$ product. One important theme when working with the $\star_M$-prodcut is a dependence on coordinates. Usually, in algebra, one seeks coordinate-free definitions and properties. However, in applied settings coordinates are genearlly specified by the problem. When working with the $\star_M$ product from an algebraic perspective, we will see that there is a balancing act between the

advantages of coordinate-free and coordinate based approaches.

Informally speaking, the $\star_M$ product operates on tubes by transforming the tubes to a new basis, performing pointwise mulitplications, then pulling back to the original basis. To make this precise, we define the *mode-3* product on tensors.

Definition 4.1.1 (Mode-3 Product)**.** *Let $\mathcal{A} \in \mathbb{K}^{n_1 \times n_2 \times n_3}$ be the tensor with tubes $\mathcal{A}_{i,j,:} = \mathbf{a}_{i,j}$ and let $M \in \mathbb{K}^{p \times n_3}$. The* mode-3 product *of $\mathcal{A}$ with M, denoted $\mathcal{A} \times_3 M$, is the tensor with tubes $(\mathcal{A} \times_3 M)_{i,j} \equiv M \operatorname{vec}(\mathbf{a}_{i,j})$.*

Remark 4.1.1. In commutative algebra, one frequently understands maps between spaces of tensors in terms of the universal property of tensor products. We can cast the definition of the mode-3 product in this light as well. Specifically, the map $\mathbb{K}^{n_1} \times \mathbb{K}^{n_2} \times \mathbb{K}^{n_3} \to \mathbb{K}^{n_1} \otimes \mathbb{K}^{n_2} \otimes \mathbb{K}^{p}$ which sends $(a, b, c) \mapsto a \otimes b \otimes (Mc)$ is multilinear and the $\times_3$ product is the induced map $\mathbb{K}^{n_1} \otimes \mathbb{K}^{n_2} \otimes \mathbb{K}^{n_3} \to \mathbb{K}^{n_1} \otimes \mathbb{K}^{n_2} \otimes \mathbb{K}^{p}$.

Using the mode-3 product, we are able to define the $\star_M$ product of tubes. Note that if the matrix M is clear from context, we will denote $\mathcal{A} \times_3 M$ by $\hat{\mathcal{A}}$ and refer to $\hat{\mathcal{A}}$ as the image of $\mathcal{A}$ in the *transform domain.*

Definition 4.1.2 ($\star_M$ product on tubes [KKA15])**.** *Let $\mathbf{a}, \mathbf{b} \in \mathbb{C}^{1 \times 1 \times n_3}$ be tubes and fix $M \in GL_{n_3}(\mathbb{C})$. We define*

$$\mathbf{a} \star_M \mathbf{b} = (\mathbf{a} \times_3 M \odot \mathbf{b} \times_3 M) \times_3 M^{-1} \equiv M^{-1} \operatorname{diag}(M \operatorname{vec}(\mathbf{a}))M \operatorname{vec}(\mathbf{b}),$$

where $\odot$ is the pointwise product. The $\star_M$ product turns $\mathbb{C}^{1 \times 1 \times n_3}$ into a commutative ring with multiplicative identity $\mathbf{1}_M = \operatorname{tube}(M^{-1}\mathbb{1}_{n_3})$. The resulting ring is denoted $\mathbb{C}_M$.

Using the $\star_M$ product, we have an alternative viewpoint on tensors $\mathcal{A} \in \mathbb{C}^{n_1 \times n_2 \times n_3}$. Specifically, rather than viewing a tensor as a multiliear object over the field $\mathbb{C}$, we can view a tensor as a $\mathbb{C}_M$-linear object. That is, $\mathcal{A} \in \operatorname{Mat}_{n_1 \times n_2}(\mathbb{C}_M)$ is an $n_1 \times n_2$ matrix with entries in the ring $\mathbb{C}_M$. In particular, there is a natural notion of tensor-tensor product.

Definition 4.1.3 ($\star_M$ product of tensors [KKA15]). *Let $\mathcal{A} \in \mathbb{C}^{n_1 \times k \times n_3}$ and $\mathcal{B} \in \mathbb{C}^{k \times n_2 \times n_3}$. The product $\mathcal{A} \star_M \mathcal{B}$ is the tensor $\mathcal{C} \in \mathbb{C}^{n_1 \times n_2 \times n_3}$ with tubes*

$$\mathcal{C}_{i,j,:} = \sum_{\ell=1}^{k} \mathcal{A}_{i,\ell,:} \star_M \mathcal{B}_{\ell,j,:}.$$

Note that with this definition of tensor-tensor product, the tensor $\mathcal{A} \star_M \mathcal{B}$ represents the composition of $\mathbb{C}_M$-linear transformations of free $\mathbb{C}_M$-modules $\mathbb{C}_M^{n_2} \xrightarrow{\mathcal{B}} \mathbb{C}_M^{k} \xrightarrow{\mathcal{A}} \mathbb{C}_M^{n_1}$. Moreover, since the $\star_M$ product on tubes is defined by pointwise multiplication in the transform domain, the $\star_M$ product of two tensors can be computed completely in parallel in the transform domain. This leads to an alternate equivalent definition of $\mathcal{A} \star_M \mathcal{B}$.

Definition 4.1.4 ($\star_M$ product of tensors [KKA15]). *Let $\mathcal{A} \in \mathbb{C}^{n_1 \times k \times n_3}$ and $\mathcal{B} \in \mathbb{C}^{k \times n_2 \times n_3}$. The product $\mathcal{A} \star_M \mathcal{B}$ is the tensor $\mathcal{C} \in \mathbb{C}^{n_1 \times n_2 \times n_3}$ such that for each $k \in [n_3]$, we have*

$$\hat{\mathcal{C}}_{:,:,k} = \hat{\mathcal{A}}_{:,:,k} \hat{\mathcal{B}}_{:,:,k}.$$

The definition of the $\star_M$ product provided in Definition 4.1.4 suggests a recipe for developing tensor decompositions analogous to those found in numerical linear algebra. First, take the tensor to the transform domain, compute a matrix decomposition for each frontal slice of the transformed tensor, then pull everything back to the original basis by taking the mode-3 product with M^{-1}. Not only does this recipe work to create tensor decompositions, but such decompositions can be computed in parallel since the frontal slices are independent in the transform domain.

Following this recipe results in several important definitions which generalize familiar objects from linear algebra. We will use the following definitions in the subsequent sections.

Definition 4.1.5 (Multiplicative Identity Tensor). *The multiplicative identity tensor $\mathcal{I}_M \in \mathbb{C}^{n \times n \times n_3}$ is the unique tensor such that each frontal slice of $\widehat{\mathcal{I}_M}$ is the $n \times n$ identity matrix. Note that this implies that $\mathcal{I}_M$ has tubes $(\mathcal{I}_M)_{i,i,:} = \mathbf{1}_M$ and $(\mathcal{I}_M)_{i,j,:} = \mathbf{0}$ otherwise.*

Definition 4.1.6 (Hermitian Transpose of a Tensor)**.** *Let* $\mathcal{A} \in \mathbb{C}^{n_1 \times n_2 \times n_3}$*. The* Hermitian transpose *of* $\mathcal{A}$ *is the unique tensor* $\mathcal{A}^H \in \mathbb{C}^{n_2 \times n_1 \times n_3}$ *such that for each* $k \in n_3$*, we have*

$$\widehat{\mathcal{A}^H}_{:,:,k} = (\hat{\mathcal{A}}_{:,:,k})^H.$$

For real-valued $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ *and a real* $M \in GL_{n_3}(\mathbb{R})$*, we call* $\mathcal{A}^H$ *the* transpose *of* $\mathcal{A}$ *and denote by* $\mathcal{A}^\top$*. Moreover, we have that* $(\mathcal{A}^\top)_{i,j,:} = \mathcal{A}_{j,i,:}$*. We call the tensor* $\mathcal{A}$ symmetric *if* $\mathcal{A} = \mathcal{A}^\top$*.*

Definition 4.1.7 (Orthogonal Tensor)**.** *Let* $\mathcal{A} \in \mathbb{R}^{n \times n \times n_3}$ *and* $M \in GL_{n_3}(\mathbb{R})$ *is* $\star_M$-orthogonal *if* $\mathcal{A} \star_M \mathcal{A}^\top = \mathcal{I}_M = \mathcal{A}^\top \star_M \mathcal{A}$*. The definition of a* $\star_M$*-unitary complex tensor is analogous, with* $\mathcal{A}^H$ *replacing* $\mathcal{A}^\top$*.*

With these definitions, one can define a singular value decomposition of a given tensor. A tensor $\mathcal{S} \in \mathbb{C}^{n_1 \times n_2 \times n_3}$ is *f-diagonal* if the only nonzero tubes in $\mathcal{S}$ are $\mathcal{S}_{i,i,:}$ for $i \in [\min(n_1, n_2)]$. That is, every frontal slice of $\mathcal{S}$ is a diagonal matrix.

Theorem 4.1.8 (Existence of an SVD, t-rank [KKA15, KHAN21])**.** *Let* $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ *and* $M \in GL_{n_3}(\mathbb{R})$*. Then, there are* $\star_M$*-orthogonal tensors* $\mathcal{U}$ *and* $\mathcal{V}$ *and an f-diagonal tensor* $\mathcal{S}$ *such that*

$$\mathcal{A} = \mathcal{U} \star_M \mathcal{S} \star_M \mathcal{V}^\top = \sum_{i=1}^{r} \mathbf{s}_{i,i} \star_M \mathcal{U}_{i,:,:} \star_M (\mathcal{V}^\top)_{:,i,:}.$$

Moreover, this decomposition can be computed slice-wise in the transform domain. Specifically, for each $k \in [n_3]$*,*

$$(\mathcal{A} \times_3 M)_{:,:,k} = (\mathcal{U} \times_3 M)_{:,:,k} (\mathcal{S} \times_3 M)_{:,:,k} (\mathcal{V}^\top \times_3 M)_{:,:,k}$$

is a singular value decomposition of the matrix $(\mathcal{A} \times_3 M)_{:,:,k}$*. The number of nonzero tubes* $\mathbf{s}_{i,i}$ *in* $\mathcal{S}$ *is called the* t-rank *of* $\mathcal{A}$*.*

The formulation of the SVD is computationally tractable and gives notion of rank which

is compatible with the the $\star_M$ product structure. Finally, we note that if the matrix M is orthogonal, then the $\star_M$ product is compatible with the inner product on $\mathbb{R}^{n_1 \times n_2 \times n_3}$ which induces the Fröbenius norm.

Lemma 4.1.9. *Fix an orthogonal matrix $M \in \mathbb{R}^{n_3 \times n_3}$ and let $\mathcal{X} \in \mathbb{R}^{n \times 1 \times n_3}$*

1. *If $\mathcal{A}, \mathcal{B} \in \mathbb{R}^{1 \times n \times n_3}$, then*

$$\langle \mathcal{A} \star_M \mathcal{X}, \mathcal{B} \star_M \mathcal{X} \rangle = \langle \mathcal{X}, (\mathcal{A}^\top \star_M \mathcal{B}) \star_M \mathcal{X} \rangle.$$

2. *If $\mathcal{A} \in \mathbb{R}^{n \times n \times n_3}$, then*

$$\langle \mathcal{X}, \mathcal{A} \star_M \mathcal{X} \rangle = \langle \mathcal{A}, \mathcal{X} \star_M \mathcal{X}^\top \rangle.$$

3. *For any tensors $\mathcal{A}, \mathcal{B} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, we have that*

$$\langle \mathcal{A} \times_3 M, \mathcal{B} \times_3 M \rangle = \langle \mathcal{A}, \mathcal{B} \rangle.$$

Proof. The proofs are calculations:

1.

$$\begin{aligned}
\langle \mathcal{A} \star_M \mathcal{X}, \mathcal{B} \star_M \mathcal{X} \rangle &= \left\langle \sum_{i=1}^{n} \mathbf{a}_i \star_M \mathbf{x}_i, \sum_{j=1}^{n} \mathbf{b}_j \star_M \mathbf{x}_i \right\rangle \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} \langle \mathbf{a}_i \star_M \mathbf{x}_i, \mathbf{b}_j \star_M \mathbf{x}_j \rangle \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} \left\langle M^\top \operatorname{diag}(M \operatorname{vec}(\mathbf{a}_i)) M \operatorname{vec}(\mathbf{x}_i), M^\top \operatorname{diag}(M \operatorname{vec}(\mathbf{b}_j)) M \operatorname{vec}(\mathbf{x}_j) \right\rangle \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} \left\langle \operatorname{vec}(\mathbf{x}_i), M^\top \operatorname{diag}(M \operatorname{vec}(\mathbf{a}_i)) \operatorname{diag}(M \operatorname{vec}(\mathbf{b}_j)) M \operatorname{vec}(\mathbf{x}_j) \right\rangle \\
&= \langle \mathcal{X}, (\mathcal{A}^\top \star_M \mathcal{B}) \star_M \mathcal{X} \rangle
\end{aligned}$$

2.
$$\begin{aligned}\langle \mathcal{X}, \mathcal{A} \star_M \mathcal{X} \rangle &= \sum_{i=1}^{n} \left\langle \operatorname{vec}(\mathbf{x}_i), \left(\sum_{j=1}^{n} M^\top \operatorname{diag}(M\mathbf{a}_{i,j}) M\mathbf{x}_j \right) \right\rangle \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} \left\langle \operatorname{vec}(\mathbf{x}_i), M^\top \operatorname{diag}(M \operatorname{vec}(\mathbf{x}_j)) M \operatorname{vec}(\mathbf{a}_{i,j}) \right\rangle \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} \left\langle M^\top \operatorname{diag}(M \operatorname{vec}(\mathbf{x}_j)) M \operatorname{vec}(\mathbf{x}_i), \operatorname{vec}(\mathbf{a}_{i,j}) \right\rangle \\ &= \langle \mathcal{X} \star_M \mathcal{X}^\top, \mathcal{A} \rangle.\end{aligned}$$

3.
$$\begin{aligned}\langle \mathcal{A} \times_3 M, \mathcal{B} \times_3 M \rangle &= \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} \langle M \operatorname{vec}(\mathbf{a}_{i,j}), M \operatorname{vec}(\mathbf{b}_{i,j}) \rangle \\ &= \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} \langle \operatorname{vec}(\mathbf{a}_{i,j}), \operatorname{vec}(\mathbf{b}_{i,j}) \rangle \\ &= \langle \mathcal{A}, \mathcal{B} \rangle.\end{aligned}$$

□

In Section 4.2 below, we will show that many linear algebraic properties of positive semidefinite matrices have analogs for a class of tensors.

4.2 M-Semidefinite Tensors

In this section, we develop the theory of M-positive semidefinite (M-PSD) tensors and corresponding M-semidefinite programming problems (M-SDP). This builds on work done in [ZHW21, ZHH22, MKY24] where positive semidefinite tensors and semidefinite programming problems which respected the t-product structure were studied. In Section 4.2.1, we develop the basic properties of M-PSD tensors. In section 4.2.2, we introduce M-SDP problems. Section 4.2.3 provides an application of this framework to low-rank tensor completion problems.

4.2.1 M-Semidefinite Tensors

The notion of positive semidefinite matrices is crucial in many areas of pure and applied mathematics. It is well known that there are many equivalent conditions for a matrix to be PSD.

Theorem 4.2.1 (cf [BPT13, Appendix A.1]). *The following are equivalent for a symmetric matrix $A \in \mathbb{R}^{n\times n}$:*

1. *A is positive semidefinite.*
2. *$\langle x, Ax\rangle \geqslant 0$ for all $x \in \mathbb{R}^n$.*
3. *There is a factorization $A = BB^\top$, where $B \in \mathbb{R}^{n\times r}$ and $r = \operatorname{rk}(A)$.*
4. *All $2^n - 1$ principal minors of A are nonnegative.*
5. *Every eigenvalue of A is nonnegative.*

Following the definition of t-PSD tensors given in [ZHW21], we define M-PSD tensors through the inner product, mirroring condition 2 from Theorem 4.2.1.

Definition 4.2.2 (M-PSD Tensor). *Let $\mathcal{A} \in \mathbb{R}^{n\times n\times n_3}$ be a symmetric tensor and $M \in \mathbb{R}^{n_3\times n_3}$ an orthogonal matrix. The tensor $\mathcal{A}$ is M-positive semidefinite (M-PSD) if*

$$\langle \mathcal{X}, \mathcal{A} \star_M \mathcal{X}\rangle \geqslant 0 \text{ for all } \mathcal{X} \in \mathbb{R}^{n\times 1\times n_3}.$$

We denote the set of M-PSD tensors of format $n \times n \times n_3$ by PSD_M^n.

In the remainder of this section, we will prove results analogous to the conditions for a matrix to be PSD given by Theorem 4.2.1. We assume throughout that M is an orthogonal matrix. First, we provide a few immediate remarks about the set of M-PSD tensors.

Remark 4.2.1. 1. The set PSD_M^n is a convex cone in the real vector space of symmetric tensors. Indeed, for any $\lambda, \mu \geqslant 0$ and any $\mathcal{A}, \mathcal{B} \in \mathrm{PSD}_M^n$, it follows from the $\mathbb{R}$-linearity of the $\star_M$ product and the inner product that

$$\langle \mathcal{X}, (\lambda\mathcal{A} + \mu\mathcal{B}) \star_M \mathcal{X} \rangle = \lambda \langle \mathcal{X}, \mathcal{A} \star_M \mathcal{X} \rangle + \mu \langle \mathcal{X}, \mathcal{B} \star_M \mathcal{X} \rangle \geqslant 0 \text{ for all } \mathcal{X} \in \mathbb{R}^{n \times n \times 1}.$$

2. The dependence on M in the definition cannot be removed. That is, there exist tensors which are M-PSD but not N-PSD for $M \neq N$. As an explicit example, we have that the tensor $\mathcal{A}$ with frontal slices

$$\mathcal{A}_{:,:,1} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad \mathcal{A}_{:,:,2} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

is I-PSD but not M-PSD for $M = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$.

3. In the case $n = 1$, $\mathcal{A} \in \mathbb{R}^{1 \times 1 \times n_3}$ is simply a tube $\mathcal{A} = \mathbf{a}$. Taking inspiration from the matrix case, where positive semidefinite 1×1 matrices are simply nonnegative scalars, one would expect that the tube $\mathbf{a}$ is a "nonnegative" element of the ring $\mathbb{R}_M$. However, the ring $\mathbb{R}_M$ is not ordered, so we instead relax to the condition that $\mathbf{a}$ is a square in the ring $\mathbb{R}_M$. In fact, a tube $\mathbf{a} \in \mathbb{R}_M$ is a square if and only if every entry of $\hat{\mathbf{a}} = \mathbf{a} \times_3 M$ is nonnegative. Indeed, if $\mathbf{a} = \mathbf{b} \star_M \mathbf{b}$, then, $\hat{\mathbf{a}} = (\hat{\mathbf{b}} \vartriangle \hat{\mathbf{b}}) \equiv \begin{bmatrix} (\sum_{j=1}^{n_3} M_{1,j} b_j)^2 & (\sum_{j=1}^{n_3} M_{2,j} b_j)^2 & \dots & (\sum_{j=1}^{n_3} M_{n_3,j} b_j)^2 \end{bmatrix}^\top$ has nonnegative entries. Conversely, if $\hat{\mathbf{a}}$ has nonnegative entries, then since each nonnegative element of $\mathbb{R}$ is a square, we have that $\hat{\mathbf{a}} \equiv \begin{bmatrix} \alpha_1^2 & \alpha_2^2 & \dots & \alpha_{n_3}^2 \end{bmatrix}^\top$ and $\mathbf{a} = \mathbf{b} \star_M \mathbf{b}$, where $\mathbf{b} \equiv M^{-1} \begin{bmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_{n_3} \end{bmatrix}^\top$. This fits thematically with the $\star_M$-product framework, as nonnegativity here is equivalent to facewise nonnegativity in the transform domain.

4. Because the $\star_M$ product represents a $\mathbb{R}_M$-linear transformation $\mathbb{R}_M^n \to \mathbb{R}_M^n$, there is an underlying $\mathbb{R}$-linear transformation of the underlying $\mathbb{R}$-vector spaces $\mathbb{R}^{n\times 1\times n_3} \to \mathbb{R}^{n\times 1\times n_3}$. By composing with the isomorphisms vec and tube, we obtain an associated $\mathbb{R}$-linear transformation of $\mathbb{R}$-vector spaces $\mathbb{R}^{nn_3} \to \mathbb{R}^{nn_3}$. In a reasonable notion of M-positive semidefiniteness, one would want this $\mathbb{R}$-linear transforamtion to also be positive semidefinite, and we show below in Proposition 4.2.3 that this is indeed the case. It also follows that the tensor $\mathcal{A}$ is M-PSD if and only if the quadratic form $\mathcal{X} \mapsto \langle \mathcal{X}, \mathcal{A} \star_M \mathcal{X} \rangle$ is positive semidefinite on $\mathbb{R}^{n\times 1\times n_3}$.

Proposition 4.2.3. *Fix an* orthogonal *matrix* $M \in \mathbb{R}^{n_3\times n_3}$. *Given a symmetric tensor* $\mathcal{A} \in \mathbb{R}^{n\times n\times n_3}$, *we have that* $\mathcal{A} \in \mathrm{PSD}_M^n$ *if and only if the block matrix*

$$\begin{bmatrix} D_{1,1} & D_{1,2} & \dots & D_{1,n} \\ D_{2,1} & D_{2,2} & \dots & D_{2,n} \\ \vdots & & \ddots & \vdots \\ D_{n,1} & D_{n,2} & \dots & D_{n,n} \end{bmatrix} \qquad \textit{where} \qquad D_{i,j} = \mathrm{diag}(M \,\mathrm{vec}(\mathbf{a}_{i,j})) \tag{4.1}$$

is positive semidefinite.

Proof. Note that for $\mathcal{X} \in \mathbb{R}^{n\times 1\times n_3}$, we have that the i^{th} tube of $\mathcal{A}\star_M \mathcal{X}$ is given by $(\mathcal{A}\star_M \mathcal{X})_i = \sum_{j=1}^n \mathbf{a}_{i,j} \star_M \mathbf{x}_j$. It then follows that

$$\langle \mathcal{X}, \mathcal{A} \star_M \mathcal{X} \rangle = \sum_{i=1}^n \left(\mathrm{vec}(\mathbf{x}_i)^\top \mathrm{vec}\left(\sum_{j=1}^n \mathbf{a}_{i,j} \star_M \mathbf{x}_j \right) \right) = \sum_{i=1}^n \left(\mathrm{vec}(\mathbf{x}_i)^\top \left(\sum_{j=1}^n M^\top D_{i,j} M \,\mathrm{vec}(\mathbf{x}_j) \right) \right),$$

where $D_{i,j} = \mathrm{diag}(M \,\mathrm{vec}(\mathbf{a}_{i,j}))$. Setting $y_j = M \,\mathrm{vec}(\mathbf{x}_j) \in \mathbb{R}^{n_3}$ for each $j \in [n]$ then gives that

$$\langle \mathcal{X}, \mathcal{A} \star_M \mathcal{X} \rangle = \sum_{i=1}^n \sum_{j=1}^n y_i^\top D_{i,j} y_j.$$

So, $\langle \mathcal{X}, \mathcal{A} \star_M \mathcal{X} \rangle \geqslant 0$ for all $\mathcal{X}$ if the block matrix (4.1) is positive semidefinite. Because M is invertible, the linear transformation $\mathbb{R}^{n\times 1\times n_3} \to \mathbb{R}^{nn_3}$ which sends a tensor $\mathcal{X}$ to the vector

$$\begin{bmatrix} (M \operatorname{vec}(\mathbf{x}_1))^\top & (M \operatorname{vec}(\mathbf{x}_2))^\top & \dots & (M \operatorname{vec}(\mathbf{x}_n))^\top \end{bmatrix}^\top$$

is an isomorphism and therefore $\langle \mathcal{X}, \mathcal{A} \star_M \mathcal{X} \rangle \geqslant 0$ for all $\mathcal{X}$ only if the block matrix (4.1) is positive semidefinite. □

We now work towards an analouge of Theorem 4.2.1 for M-PSD tensors. In particular, we show that M-PSD tensors factorize, have prinicpal minors which are squares in $\mathbb{R}_M$, and have eigentubes which are squares in the ring of tubes $\mathbb{R}_M$.

Factorization:

Our first goal is to show that a symmetric tensor $\mathcal{A} \in \mathbb{R}^{n\times n\times n_3}$ is M-PSD if and only if it admits a decomposition $\mathcal{A} = \mathcal{B} \star_M \mathcal{B}^\top$, where $\mathcal{B} \in \mathbb{R}^{n\times r\times n_3}$ and $r = t\text{-rank}(\mathcal{A})$. An important consequence of the factorization will be that a tensor $\mathcal{A}$ is M-PSD if and only if each frontal slice of $\hat{\mathcal{A}}$ is a positive semidefinite matrix. Since we are relating the factorization of $\mathcal{A}$ to t-rank$(\mathcal{A})$, we will pass to the SVD of $\mathcal{A}$ in our proof. First, we note that one direction is clear:

Proposition 4.2.4. *A tensor of the form $\mathcal{A} = \mathcal{B} \star_M \mathcal{B}^\top$ for $\mathcal{B} \in \mathbb{R}^{n\times r\times n_3}$ is M-PSD.*

Proof. Note that if $\mathcal{B} = \begin{bmatrix} \mathcal{B}^{(1)} & \mathcal{B}^{(2)} & \dots & \mathcal{B}^{(r)} \end{bmatrix}$, where $\mathcal{B}^{(i)} \in \mathbb{R}^{n\times 1\times n_3}$, then $\mathcal{B} \star_M \mathcal{B}^\top = \sum_{i=1}^r \mathcal{B}^{(i)} \star_M \mathcal{B}^{(i)\top}$. Since PSD_M^n is a convex cone, it therefore suffices to show that $\mathcal{B} \star_M \mathcal{B}^\top$ is M-PSD for any $\mathcal{B} \in \mathbb{R}^{n\times 1\times n_3}$. This follows from Lemma 4.1.9 since for any $\mathcal{X} \in \mathbb{R}^{n\times 1\times n_3}$, we have that

$$\langle \mathcal{X}, (\mathcal{B} \star_M \mathcal{B}^\top) \star_M \mathcal{X} \rangle = \langle \mathcal{B}^\top \star_M \mathcal{X}, \mathcal{B}^\top \star_M \mathcal{X} \rangle = \|\mathcal{B}^\top \star_M \mathcal{X}\|_F^2 \geqslant 0.$$

□

The converse to Proposition 4.2.4 takes a bit more work. To prove that any M-PSD tensor $\mathcal{A}$ admits a decomposition $\mathcal{A} = \mathcal{B} \star_M \mathcal{B}^\top$, we pass through the SVD and reduce to the case of f-diagonal tensors. In this case, we show that every M-PSD f-diagonal tensor has diagonal tubes which are squares in the ring $\mathbb{R}_M$.

Theorem 4.2.5. *A symmetric tensor $\mathcal{A} \in \mathbb{R}^{n\times n\times n_3}$ is M-PSD if and only if $\mathcal{A} = \mathcal{B} \star_M \mathcal{B}^\top$ where $\mathcal{B} \in \mathbb{R}^{n\times r\times n_3}$, where $r = t$-rank$(\mathcal{A})$.*

Proof. The "if" direction is Proposition 4.2.4. For the converse direction, we first reduce to the f-diagonal case. Let $\mathcal{A} = \mathcal{U} \star_M \mathcal{S} \star_M \mathcal{U}^\top$ be an SVD of $\mathcal{A}$. Note that there is a decomposition of this form since the SVD is computed slicewise in the transform domain and $\mathcal{A}$ (and thus $\hat{\mathcal{A}}$) is symmetric. Next, note that by Lemma 4.1.9, we have that for any $\mathcal{X} \in \mathbb{R}^{n\times 1\times n_3}$,

$$\langle \mathcal{X}, \mathcal{A} \star_M \mathcal{X}\rangle = \langle \mathcal{X}, (\mathcal{U} \star_M \mathcal{S} \star_M \mathcal{U}^\top) \star_M \mathcal{X}\rangle = \langle \mathcal{U}^\top \star_M \mathcal{X}, \mathcal{S} \star_M (\mathcal{U}^\top \star_M \mathcal{X})\rangle.$$

Since $\mathcal{U}$ is orthogonal, it follows that $\mathcal{A}$ is PSD if and only if $\mathcal{S}$ is PSD. So, it suffices to consider the case of f-diagonal tensors. We claim that if $\mathcal{S}$ is an f-diagonal tensor which is M-PSD, then each tube along the diagonal of $\mathcal{S}$ is a square in $\mathbb{R}_M$. Fix an index $i \in [n]$ and $\ell \in [n_3]$ and set $\mathcal{X} \in \mathbb{R}^{n\times 1\times n_3}$ to be the tensor with tubes

$$\mathbf{x}_j = \begin{cases} \text{tube}(M^\top e_\ell), & j = i \\ 0, & j \neq i \end{cases}.$$

where e_ℓ is the ℓ^{th} standard basis vector of $\mathbb{R}^{n_3}$. Note that the tubes $\mathbf{x}_j$ are defined so that the only nonzero entry of $\widehat{\mathcal{X}}$ is $\widehat{\mathcal{X}}_{i,1,\ell} = 1$. Then,

$$\langle \mathcal{X}, \mathcal{S} \star_M \mathcal{X}\rangle = e_\ell^\top \operatorname{diag}(M \operatorname{vec}(\mathcal{S}_{i,i,:}))e_\ell \geqslant 0.$$

So, every entry of $M \operatorname{vec}(\mathcal{S}_{i,i,:})$ is nonnegative and therefore $\mathcal{S}_{i,i,:}$ is a square in the ring of tubes. If $\mathcal{S}$ has t-rank$(\mathcal{S}) = r$, then this implies that there is $\mathcal{Q} \in \mathbb{R}^{n\times r\times n_3}$ such that

$\mathcal{S} = \mathcal{Q} \star_M \mathcal{Q}^\top$. Taking $\mathcal{B} = \mathcal{U} \star_M \mathcal{Q}$ then yields $\mathcal{A} = \mathcal{B} \star_M \mathcal{B}^\top$. □

The decomposition given in Theorem 4.2.5 allows us to check membership in the cone PSD_M^n on frontal slices in the transform domain. This is similar to other properties with $\star_M$-prodcuts of tensors, where the tensor property is equivalent to matrix properties on frontal slices in the transform domain.

Corollary 4.2.6. *Let $M \in \mathbb{R}^{n_3 \times n_3}$ be orthogonal. A symmetric tensor $\mathcal{A} \in \mathbb{R}^{n \times n \times n_3}$ is M-PSD if and only if each frontal slice of $\hat{\mathcal{A}}$ is a PSD matrix. That is, $(\mathcal{A} \times_3 M)_{:,:,\ell} \succeq 0$ for each $\ell \in [n_3]$.*

Proof. By Theorem 4.2.5, the tensor $\mathcal{A}$ is M-PSD if and only if $\mathcal{A} = \mathcal{B} \star_M \mathcal{B}^\top$ for some $\mathcal{B} \in \mathbb{R}^{n \times r \times n_3}$. By the definition of the $\star_M$-product, this is equivalent to $\hat{\mathcal{A}}_{:,:,\ell} = \hat{\mathcal{B}}_{:,:,\ell} \hat{\mathcal{B}}^\top_{:,:,\ell}$ for each $\ell \in [n_3]$ and therefore $\mathcal{A}$ is M-PSD if and only if each frontal slice of $\hat{\mathcal{A}}$ is a PSD matrix. □

We will use Corollary 4.2.6 in our discussion of M-semidefinite programming problems below. In particular, we can view Corollary 4.2.6 as a block-diagonalization result giving n_3 independent $n \times n$ matrix variables rather than an $nn_3 \times nn_3$ matrix variables.

Principal Minors:

In this subsection, we make an analogy with statement (4) in Theorem 4.2.1, which states that every principal minor of a PSD matrix is nonnegative. To effectively develop such a result, we need a notion of determinant. For a tensor $\mathcal{A} \in \mathbb{R}^{n \times n \times n_3}$, we define the determinant

$$\det_M(\mathcal{A}) = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) \mathbf{a}_{1\sigma(1)} \star_M \mathbf{a}_{2\sigma(2)} \star_M \cdots \star_M \mathbf{a}_{n\sigma(n)}. \tag{4.2}$$

Here, the sign of a permutation $\sigma \in S_n$ is $\mathrm{sgn}(\sigma) = 1$ if σ can be written as an even number of transpositions and $\mathrm{sgn}(\sigma) = -1$ otherwise. Note that $\det_M(\mathcal{A}) \in \mathbb{R}^{1 \times 1 \times n_3}$ is a

tube and that the definition (4.2) is consistent with the point of view that $\mathcal{A}$ is an $\mathbb{R}_M$-linear transformation of $\mathbb{R}_M$-modules. If $J = \{j_1, j_2, \ldots, j_r\} \subseteq [n]$ is a subset of indices with $j_1 \leqslant j_2 \leqslant \cdots \leqslant j_r$, then the principal minor indexed by J is

$$\det_M(\mathcal{A}_J) = \sum_{\sigma \in S_r} \operatorname{sgn}(\sigma) \mathbf{a}_{j_1, j_{\sigma(1)}} \star_M \mathbf{a}_{j_2, j_{\sigma(2)}} \star_M \cdots \star_M \mathbf{a}_{j_r, j_{\sigma(r)}}.$$

Remark 4.2.2. The tube determinant of a $n \times n \times n_3$ tensor with the t-product structure was defined in [EJRR23] to be the tube $\det_t(A)$ such that

$$\widehat{\det_t(\mathcal{A})} = \operatorname{tube}\left(\begin{bmatrix} \det(\hat{\mathcal{A}}_{:,:,1}) & \det(\hat{\mathcal{A}}_{:,:,2}) & \ldots & \det(\hat{\mathcal{A}}_{:,:,n_3}) \end{bmatrix}^\top\right).$$

The definition of $\det_M$ in (4.2) agrees with this perspective. Indeed, for $M \in \mathrm{GL}_{n_3}(\mathbb{R})$ and $\mathcal{A} \in \mathbb{R}^{n \times n \times n_3}$, we have that for $D_{i,j} = \operatorname{diag}(M \operatorname{vec}(\mathbf{a}_{i,j}))$,

$$\begin{aligned} \det_M(\mathcal{A})) &= \sum_{\sigma \in S_{n_3}} \operatorname{sgn}(\sigma) \mathbf{a}_{1,\sigma(1)} \star_M \mathbf{a}_{2,\sigma(2)} \star_M \cdots \star_M \mathbf{a}_{r,\sigma(r)} \\ &\equiv M^{-1} \sum_{\sigma \in S_{n_3}} \operatorname{sgn}(\sigma) D_{1,\sigma(1)} D_{2,\sigma(2)} \cdots D_{n_3-1,\sigma(n_3-1)} M \operatorname{vec}(\mathbf{a}_{n_3,\sigma(n_3)}) \qquad (4.3) \\ &\equiv M^{-1} \begin{bmatrix} \det(\hat{\mathcal{A}}_{:,:,1}) & \det(\hat{\mathcal{A}}_{:,:,2}) & \cdots & \det(\hat{\mathcal{A}}_{:,:,n_3}) \end{bmatrix}^\top. \end{aligned}$$

Proposition 4.2.7. *A symmetric tensor $\mathcal{A} \in \mathbb{R}^{n \times n \times n_3}$ is M-PSD if and only if for each ordered subset of indices $J \subseteq [n]$, the principal minor $\det_M(\mathcal{A}_J)$ is a square in the ring $\mathbb{R}_M$.*

Proof. Let $J = \{j_1, j_2, \ldots, j_r\} \subseteq [n]$ with $j_1 \leqslant j_2 \leqslant \ldots \leqslant j_r$. By Equation (4.3), we have that

$$\det_M(\mathcal{A}_J) \equiv M^{-1} \begin{bmatrix} \det(\hat{\mathcal{A}}_{J,J,1}) & \det(\hat{\mathcal{A}}_{J,J,2}) & \ldots & \det(\hat{\mathcal{A}}_{J,J,n_3}) \end{bmatrix}.$$

Suppose that $\mathcal{A}$ is M-PSD. By Corollary 4.2.6, each frontal slice $\hat{\mathcal{A}}_{:,:,\ell}$ is positive semidefinite and therefore each minor $\det(\hat{\mathcal{A}}_{J,J,\ell})$ is nonnegative. It then follows that $\det_M(\mathcal{A}_J)$ is a square in $\mathbb{R}_M$.

Conversely, if $\det_M(\mathcal{A}_J)$ is a square in $\mathbb{R}_M$ for any ordered set of indices J, then $\det(\hat{\mathcal{A}}_{J,J,\ell}) \geqslant 0$ for each $\ell \in [n_3]$. So, every principal minor of the frontal slice $\hat{\mathcal{A}}_{:,:,\ell}$ is nonnegative and therefore $\hat{\mathcal{A}}$ has PSD frontal slices. By Corollary 4.2.6, this implies $\mathcal{A} \in \mathrm{PSD}_M^n$. □

Eigentubes:

The final characterization of PSD matrices in Theorem 4.2.1 is that a symmetric matrix A is positive semidefinite if and only if A has all nonnegative eigenvalues. We prove a weaker statement for M-PSD tensors, namely, that under a nondegeneracy condition, the tubes which act analogously to eigenvalues must be squares in $\mathbb{R}_M$. This statement is a weaker statement than Theorem 4.2.5 and Proposition 4.2.7 because the ring $\mathbb{R}_M$ has zero divisors. The spectral theory of tensors equipped with the t-product was developed in [EJRR23]. Note also that a notion of spectral theory for higher order tensors which does not depend on a tensor-tensor product has been studied in the literature; see e.g. [QL17] for an overview.

Proposition 4.2.8. *Suppose that $\mathcal{A} \in \mathbb{R}^{n\times n\times n_3}$ is M-PSD. Let $\boldsymbol{\lambda} \in \mathbb{R}_M$ and $\mathcal{X} \in \mathbb{R}^{n\times 1\times n_3}$ be such that $\mathcal{A} \star_M \mathcal{X} = \boldsymbol{\lambda} \star_M \mathcal{X}$ and that all frontal slices of $\hat{\mathcal{X}}$ are nonzero. Then $\boldsymbol{\lambda}$ is a square in $\mathbb{R}_M$.*

Proof. If $\mathcal{A} \star_M \mathcal{X} = \boldsymbol{\lambda} \star_M \mathcal{X}$, then in the transform domain, we have $\hat{\mathcal{A}} \vartriangle \hat{\mathcal{X}} = \hat{\boldsymbol{\lambda}} \vartriangle \hat{\mathcal{X}}$, where $\vartriangle$ denotes the facewise product. It then follows that for each $\ell \in [n_3]$,

$$\hat{\mathcal{A}}_{:,:,\ell}\hat{\mathcal{X}}_{:,1,\ell} = \hat{\boldsymbol{\lambda}}_{1,1,\ell}\hat{\mathcal{X}}_{:,1,\ell}.$$

By the hypothesis that all frontal slices of $\hat{\mathcal{X}}$ are nonzero, this implies that $\hat{\boldsymbol{\lambda}}_{1,1,\ell}$ is an eigenvalue of $\hat{\mathcal{A}}_{:,:,\ell}$ for each $\ell \in [n_3]$. By Corollary 4.2.6, the matrix $\hat{\mathcal{A}}_{:,:,\ell}$ is positive semidefinite and therefore $\hat{\boldsymbol{\lambda}}_{1,1,\ell} \geqslant 0$. It therefore follows that $\boldsymbol{\lambda}$ is a square in $\mathbb{R}_M$. □

Note that the requirement that all frontal slices of $\hat{\mathcal{X}}$ are nonzero is necessary for the conclusion of Proposition 4.2.8 to hold, as shown by the following example.

Example 4.2.1. Let $\mathcal{A} \in \mathbb{R}^{2\times 2\times 2}$ and $\mathcal{X} \in \mathbb{R}^{2\times 1\times 2}$ have frontal slices

$$\mathcal{A}_{:,:,1} = \mathcal{A}_{:,:,2} = I, \qquad \mathcal{X}_{:,1,1} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad \mathcal{X}_{:,1,2} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Then, if I is the 2×2 identity matrix, $\mathcal{A} \in \mathrm{PSD}_I^2$ and $\mathcal{A} \star_I \mathcal{X} = \boldsymbol{\lambda} \star_I \mathcal{X}$ for any $\boldsymbol{\lambda} \in \mathbb{R}_I$ with $\lambda_{1,1,1} = 1$. However, $\boldsymbol{\lambda}$ is a square in $\mathbb{R}_I$ if and only if $\boldsymbol{\lambda}$ has nonnegative entries. It follows that there are tubes $\boldsymbol{\lambda}$ with $\mathcal{A} \star_I \mathcal{X} = \boldsymbol{\lambda} \star_I \mathcal{X}$ which are not squares in $\mathbb{R}_I$. For example, $\boldsymbol{\lambda} = \mathrm{tube}(\begin{bmatrix} 1 & -1 \end{bmatrix}^\top)$ is such a tube. ◇

4.2.2 M-Semidefinite Programs

Using the notion of M-PSD tensors developed in the previous section, we are able to define M-semidefinite programs. Taking inspiration from regular semidefinite programming problems, these are problems of optimizing a linear functional over the cone of M-PSD tensors.

Definition 4.2.9 (M-SDP)**.** *Let $M \in \mathbb{R}^{n_3\times n_3}$ be an orthogonal matrix and fix symmetric tensors $\mathcal{C} \in \mathbb{R}^{n\times n\times n_3}$ and $\mathcal{A}^{(i)} \in \mathbb{R}^{n\times n\times n_3}$ for $i \in [k]$ and scalars $b_1, b_2, \ldots, b_k$. An* M-semidefinite programming problem *(M-SDP) is an optimization problem of the form*

$$\max \langle \mathcal{C}, \mathcal{X} \rangle \quad s.t.\ \langle \mathcal{A}^{(i)}, \mathcal{X} \rangle = b_i \ \text{ for } i \in [k], \ \ \mathcal{X} \in \mathrm{PSD}_M^n. \tag{M-SDP}$$

The problem (M-SDP) reduces to classical semidefinite programming in the case where $n_3 = 1$. On the other extreme, (M-SDP) is equivalent to linear programming in $\mathbb{R}^{n_3}$ when $n = 1$ and $M = I$. In between these extreme cases, (M-SDP) can be thought of as standard SDP with n_3 matrix variables of size $n \times n$ (or, equivalently, a single $nn_3 \times nn_3$ block diagonal matrix variable with blocks of size $n \times n$). Indeed, by Corollary 4.2.6, the condition $\mathcal{X} \in \mathrm{PSD}_M^n$ can be checked on frontal slices in the transform domain.

Proposition 4.2.10. *Fix an orthogonal matrix $M \in \mathbb{R}^{n_3\times n_3}$, and let $\mathcal{C}, \mathcal{A}^{(i)}$, and b_i be as in Definition 4.2.9. Then, a tensor $\mathcal{X}$ is feasible (optimal) to* (M-SDP) *if and only if $X^{(\ell)} = \hat{\mathcal{X}}_{:,:,\ell}$*

for $\ell \in [n_3]$ is feasible (optimal) to the following problem

$$\begin{aligned} \max \ & \sum_{\ell=1}^{n_3} \langle \hat{\mathcal{C}}_{:,:,\ell}, X^{(\ell)} \rangle \\ s.t. \ & \sum_{\ell=1}^{n_3} \langle \widehat{\mathcal{A}^{(i)}}_{:,:,\ell}, X^{(\ell)} \rangle = b_i \qquad i \in [k], \\ & X^{(\ell)} \succeq 0 \qquad\qquad\qquad\quad \ell \in [n_3]. \end{aligned}$$

Proof. By Corollary 4.2.6, we have $\mathcal{X} \in \mathrm{PSD}_M^n$ if and only if $\hat{\mathcal{X}}_{:,:,j} \succeq 0$ for each $j \in [n_3]$. Moreover, since M is orthogonal, it follows from Lemma 4.1.9 that $\langle \mathcal{C}, \mathcal{X} \rangle = \langle \hat{\mathcal{C}}, \hat{\mathcal{X}} \rangle = \sum_{j=1}^{n_3} \langle \hat{\mathcal{C}}_{:,:,j}, \hat{\mathcal{X}}_{:,:,j} \rangle$ and for each $i \in [\ell]$, we have $\langle \widehat{\mathcal{A}^{(i)}}, \mathcal{X} \rangle = \langle \widehat{\mathcal{A}^{(i)}}, \hat{\mathcal{X}} \rangle = \sum_{j=1}^{n_3} \langle \widehat{\mathcal{A}^{(i)}}_{:,:,j}, \hat{\mathcal{X}}_{:,:,j} \rangle$. □

In the special case where for each i, the tensor $\widehat{\mathcal{A}^{(i)}}$ has only one nonzero frontal slice $\widehat{\mathcal{A}^{(i)}}_{:,:,\ell_i}$, the equality constraint $\sum_{\ell=1}^{n_3} \langle \widehat{\mathcal{A}^{(i)}}_{:,:,\ell}, X^{(\ell)} \rangle = b_i$ reduces to $\langle \widehat{\mathcal{A}^{i}}_{:,:,\ell_i}, X^{(\ell_i)} \rangle = b_i$. So, in this case, (M-SDP) can be solved via n_3 independent standard SDPs with matrix variables of size $n \times n$.

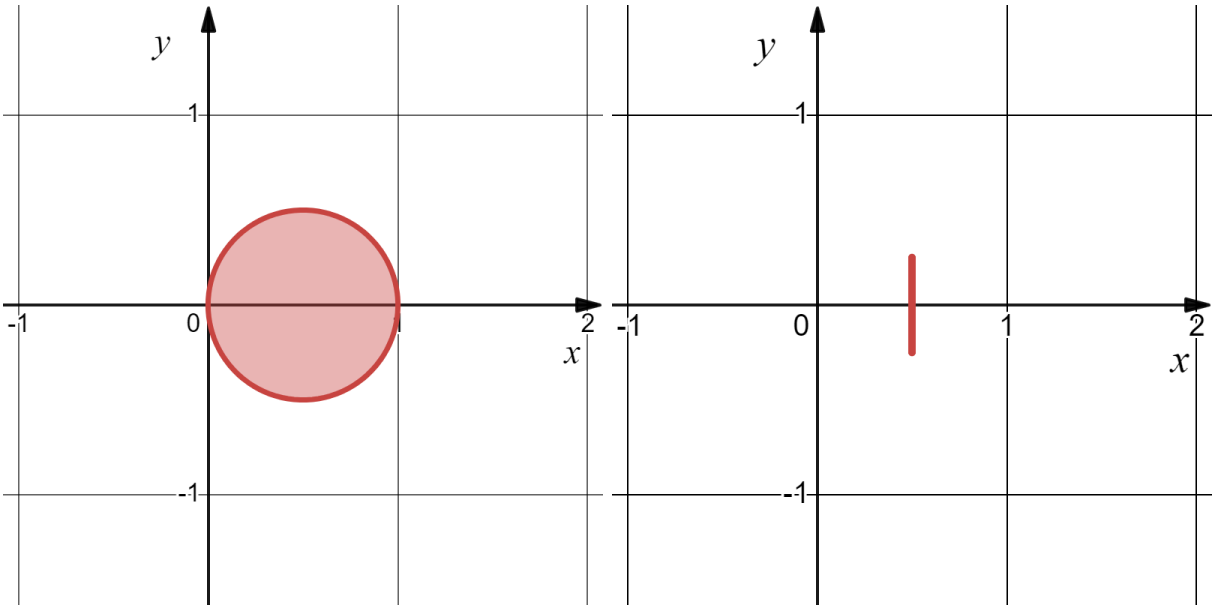
We conclude this section with an example highlighting the dependence on M of the feasible region to (M-SDP). In particular, we show that different choices of M can lead to drastically different behaviors with the feasible region, as for tensors with the same format and affine constraints, one choice of matrix leads to a feasible region with interior in $\mathbb{R}^2$ while another is a line segment in $\mathbb{R}^2$.

Example 4.2.2. Consider the $2 \times 2 \times 2$ tensor $\mathcal{X}$ with frontal slices

$$\mathcal{X}_{:,:,1} = \begin{bmatrix} x & y \\ y & (1-x) \end{bmatrix} \qquad \mathcal{X}_{:,:,2} = \begin{bmatrix} (1-x) & y \\ y & x \end{bmatrix}.$$

Now, $\mathcal{X} \succeq_I 0$ if and only if $x(1-x) - y^2 \geqslant 0$. So, the set of I-PSD tensors with this structure is given by (x, y) in the disk of radius $\frac{1}{2}$ centered at $(\frac{1}{2}, 0)$.

On the other hand, if we set $\alpha = \frac{1}{\sqrt{2}}$ and $M = \begin{bmatrix} \alpha & \alpha \\ \alpha & -\alpha \end{bmatrix}$, then the frontal slices of $\hat{\mathcal{X}} = \mathcal{X} \times_3 M$ are



(a) The set of I-PSD tensors in Example 4.2.2 (b) The set of M-PSD tensors in Example 4.2.2

Figure 4.1: The sets $\{(x, y) \mid \mathcal{X} \succeq_I 0\}$ (Figure 4.1a) and $\{(x, y) \mid \mathcal{X} \succeq_M 0\}$ (Figure 4.1b) discussed in Example 4.2.2. The region for which $\mathcal{X} \succeq_I 0$ is a disk, while the region for which $\mathcal{X} \succeq_M 0$ is a line segment.

$$\widehat{\mathcal{X}_{:,:,1}} = \begin{bmatrix} \alpha & 2\alpha y \\ 2\alpha y & \alpha \end{bmatrix} \qquad \widehat{\mathcal{X}_{:,:,2}} = \begin{bmatrix} \alpha(2x-1) & 0 \\ 0 & \alpha(1-2x) \end{bmatrix}.$$

So, $\mathcal{X} \succeq_M 0$ if and only if $x = \frac{1}{2}$ and $y^2 \leqslant \frac{1}{4}$. This is a line segment in $\mathbb{R}^2$. The two regions are shown in Figure 4.1.

◇

4.2.3 Application: Low-Rank Tensor Completion

One application of semidefinite programming in the matrix setting is in low rank matrix completion. In particular, for a matrix A with singular values $\sigma_1 \geqslant \sigma_2 \geqslant \ldots \geqslant \sigma_r \geqslant 0$, one can compute the *nuclear norm* $\|A\|_* = \sum_{j=1}^{r} \sigma_j$ via a semidefinite program through the following primal-dual pair (see e.g., [RFP10] [BPT13, Section 2.1]).

$$\begin{aligned} &\max \operatorname{trace}(A^\top Y) && \min_{W_1,W_2} \frac{1}{2}\left(\operatorname{trace} W_1 + \operatorname{trace} W_2\right) \\ &\text{s.t.} \begin{bmatrix} I & Y \\ Y^\top & I \end{bmatrix} \succeq 0 && \text{s.t.} \begin{bmatrix} W_1 & A \\ A^\top & W_2 \end{bmatrix} \succeq 0 \end{aligned} \tag{4.4}$$

In another line of research, the $\star_M$ product has been in low-rank completion of tensors (e.g., [KLL21]). Here, the authors consider a norm on $\mathbb{R}^{n_1\times n_2\times n_3}$ defined as the sum of nuclear norms of frontal slices in the transform domain:

$$\|\mathcal{X}\|_{M,*} = \sum_{\ell=1}^{n_3} \|(\mathcal{X} \times_3 M)_{:,:,\ell}\|_*. \tag{4.5}$$

We show below that an analogous formulation computes the norm (4.5) using an M-SDP and use this to formulate the tensor completion problem for fixed M as an M-SDP.

Proposition 4.2.11. *Fix an orthognal matrix $M \in \mathbb{R}^{n_3\times n_3}$. The M-nuclear norm of a tensor $\mathcal{A} \in \mathbb{R}^{n_1\times n_2\times n_3}$, defined by (4.5), can be computed by solving n_3 independent matrix SDPs or solving a single equivalent M-SDP.*

Proof. Let $\hat{\mathcal{A}} = \mathcal{A} \times_3 M$. For each $\ell \in [n_3]$, we can compute optimal solutions $W_1^{(\ell)}, W_2^{(\ell)}$ to

$$\min_{W_1,W_2} \frac{1}{2}\left(\operatorname{trace} W_1 + \operatorname{trace} W_2\right) \quad \text{s.t.} \begin{bmatrix} W_1 & \hat{\mathcal{A}}_{:,:,\ell} \\ \hat{\mathcal{A}}_{:,:,\ell}^\top & W_2 \end{bmatrix} \succeq 0 \tag{4.6}$$

where $\|\hat{\mathcal{A}}_{:,:,\ell}\|_* = \frac{1}{2}(\operatorname{trace} W_1^{(\ell)} + \operatorname{trace} W_2^{(\ell)})$. It follows that $\|\mathcal{A}\|_{M,*} = \sum_{\ell=1}^{n_3} \|\hat{\mathcal{A}}_{:,:,\ell}\|_* = \frac{1}{2}\sum_{\ell=1}^{n_3}(\operatorname{trace} W_1^{(\ell)} + \operatorname{trace} W_2^{(\ell)})$. This shows that $\|\mathcal{A}\|_{M,*}$ can be computed via n_3 independent SDPs.

We now claim that $\|\mathcal{A}\|_{M,*}$ is the optimal value of the following M-SDP:

$$\min \frac{1}{2}\left\langle \begin{bmatrix} \mathcal{I}_M & 0 \\ 0 & \mathcal{I}_M \end{bmatrix}, \begin{bmatrix} \mathcal{W}_1 & \mathcal{A} \\ \mathcal{A}^\top & \mathcal{W}_2 \end{bmatrix} \right\rangle \quad \text{s.t.} \quad \begin{bmatrix} \mathcal{W}_1 & \mathcal{A} \\ \mathcal{A}^\top & \mathcal{W}_2 \end{bmatrix} \succeq_M 0. \tag{4.7}$$

To see that the problem (4.7) provides a lower bound on $\|\mathcal{A}\|_{M,*}$, set $\widehat{\mathcal{W}}_1$ and $\widehat{\mathcal{W}}_2$ to be the

tensors with frontal slices $(\widehat{\mathcal{W}}_1)_{:,:,\ell} = W_1^{(\ell)}$ and $(\widehat{\mathcal{W}}_2)_{:,:,\ell} = W_2^{(\ell)}$, $\ell \in [n_3]$ where $W_1^{(\ell)}$ and $W_2^{(\ell)}$ are optimal solutions to (4.4) as above. Further set $\mathcal{W}_1 = \widehat{\mathcal{W}}_1 \times_3 M^{-1}$ and $\mathcal{W}_2 = \widehat{\mathcal{W}}_2 \times_3 M^{-1}$. Now, by Corollary 4.2.6, it follows that $\begin{bmatrix} \mathcal{W}_1 & \mathcal{A} \\ \mathcal{A}^\top & \mathcal{W}_2 \end{bmatrix}$ is M-PSD. Moreover, we compute that the corresponding objective value is

$$\begin{aligned}
\frac{1}{2}\left\langle \begin{bmatrix} \mathcal{I}_M & 0 \\ 0 & \mathcal{I}_M \end{bmatrix}, \begin{bmatrix} \mathcal{W}_1 & \mathcal{A} \\ \mathcal{A}^\top & \mathcal{W}_2 \end{bmatrix} \right\rangle &= \frac{1}{2}\left\langle \begin{bmatrix} \mathcal{I}_M \times_3 M & 0 \\ 0 & \mathcal{I}_M \times_3 M \end{bmatrix}, \begin{bmatrix} \mathcal{W}_1 & \mathcal{A} \\ \mathcal{A}^\top & \mathcal{W}_2 \end{bmatrix} \times_3 M \right\rangle \\
&= \frac{1}{2}\sum_{\ell=1}^{n_3}\left\langle \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}, \begin{bmatrix} W_1^{(\ell)} & \hat{\mathcal{A}}_{:,:,\ell} \\ \hat{\mathcal{A}}_{:,:,\ell}^\top & W_2^{(\ell)} \end{bmatrix} \right\rangle \\
&= \frac{1}{2}\sum_{\ell=1}^{n_3}(\operatorname{trace} W_1^{(\ell)} + \operatorname{trace} W_2^{(\ell)}) \\
&= \sum_{\ell=1}^{n_3} \|\hat{\mathcal{A}}_{:,:,\ell}\|_* \\
&= \|\mathcal{A}\|_{M,*}
\end{aligned}$$

Conversely, by Corollary 4.2.6, any M-PSD tensor with block format $\begin{bmatrix} \mathcal{W}_1 & \mathcal{A} \\ \mathcal{A}^\top & \mathcal{W}_2 \end{bmatrix}$ yields feasible solutions to (4.4) by taking frontal slices in the transform domain. Moreover, the corresponding objective value satisfies

$$\begin{aligned}
\frac{1}{2}\left\langle \begin{bmatrix} \mathcal{I}_M & 0 \\ 0 & \mathcal{I}_M \end{bmatrix}, \begin{bmatrix} \mathcal{W}_1 & \mathcal{A} \\ \mathcal{A}^\top & \mathcal{W}_2 \end{bmatrix} \right\rangle &= \frac{1}{2}\sum_{\ell=1}^{n_3}(\operatorname{trace}(\mathcal{W}_1 \times_3 M)_{:,:,\ell} + \operatorname{trace}(\mathcal{W}_2 \times_3 M)_{:,:,\ell}) \\
&\geqslant \sum_{\ell=1}^{n_3} \|(\mathcal{A} \times_3 M)_{:,:,\ell}\|_* \\
&= \|\mathcal{A}\|_{M,*}.
\end{aligned}$$

So, the problem (4.7) computes $\|\mathcal{A}\|_{M,*}$.

□

We now consider the tensor completion problem. A tensor $\mathcal{Y} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is *partially specified* if the entries $\mathcal{Y}_{i,j,k}$ are known for $(i,j,k) \in \Omega \subseteq \mathbb{Z}^3$ and unspecified $\mathcal{Y}_{i,j,k}$ if $(i,j,k) \notin \Omega$. We seek a low M-rank solution to the tensor completion problem, using the M-nuclear norm (4.5) as a proxy for rank in the objective function:

$$\min_{\mathcal{A}} \|\mathcal{A}\|_{M,*} \text{ s.t. } \mathcal{A}_{i,j,k} = \mathcal{Y}_{i,j,k} \text{ for all } (i,j,k) \in \Omega. \tag{4.8}$$

We rewrite (4.8) as an M-SDP:

$$\min_{\mathcal{A},\mathcal{W}_1,\mathcal{W}_2} \frac{1}{2} \left\langle \begin{bmatrix} \mathcal{I}_M & 0 \\ 0 & \mathcal{I}_M \end{bmatrix}, \begin{bmatrix} \mathcal{W}_1 & \mathcal{A} \\ \mathcal{A}^\top & \mathcal{W}_2 \end{bmatrix} \right\rangle \quad \text{s.t.} \quad \mathcal{A}_{i,j,k} = \mathcal{Y}_{i,j,k} \text{ for all } (i,j,k) \in \Omega, \begin{bmatrix} \mathcal{X} & \mathcal{A} \\ \mathcal{A}^\top & \mathcal{Z} \end{bmatrix} \succeq_M 0. \tag{4.9}$$

Note that if $\mathcal{Y} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, then the tensors involved in (4.9) have sizes $\mathcal{W}_1 \in \mathbb{R}^{n_1 \times n_1 \times n_3}$, $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, and $\mathcal{W}_2 \in \mathbb{R}^{n_2 \times n_2 \times n_3}$.

4.3 Group Representations and $\star_M$-Products

In this section, we investigate the algebraic structure of the $\star_M$-product thorugh the lens of representation theory of finite groups. This will allow us to connect the M-SDP framework develooped in the previous section to invariant SDPs in Section 4.3.2. Representation theoretic methods have become increasingly popular in data-centric mathematics. One particularly active area which takes this perpective is the area of "geometric deep learning"; see, e.g., [BBCV21, LN23]

First, we note that from a purely commutative algebraic point of view, the rings $\mathbb{C}_M$ are all isomorphic. So, an algebraic interpretation of the $\star_M$-product needs to involve more than just the ring structure on $\mathbb{C}_M$.

Proposition 4.3.1. *Fix n_3. Then, for any $M \in GL_{n_3}(\mathbb{C})$, we have that $\mathbb{C}_M \simeq \mathbb{C}_I$ in the category of $\mathbb{C}$-algebras.*

Proof. The map $\phi : \mathbb{C}_M \to \mathbb{C}_I$ defined by $\phi(\mathbf{a}) = \mathbf{a} \times_3 M$ is bijective and $\mathbb{C}$-linear since M is invertible. It is a $\mathbb{C}$-algebra homomorphism because

$$\begin{aligned} \phi(\mathbf{a} \star_M \mathbf{b}) &= \text{tube}(M^{-1} \operatorname{diag}(M \operatorname{vec}(\mathbf{a}))M \operatorname{vec}(\mathbf{b})) \times_3 M \\ &= \text{tube}(\operatorname{diag}(M \operatorname{vec}(\mathbf{a}))M \operatorname{vec}(\mathbf{b})) \\ &= \phi(\mathbf{a}) \star_I \phi(\mathbf{b}). \end{aligned} \tag{4.10}$$

□

As one motivation of our point of view, notice that the map ϕ used in the proof of Proposition 4.3.1 is simply a change of ($\mathbb{C}$-vector space) basis on the rings. So, it is natural to ask "What is a *good* choice of basis?" To answer this question, we assume that our problem has some underlying symmetry and show that a "good" choice of basis is one which is compatible with the decomposition of $\mathbb{C}^{n_3}$ into irreducible representations of the underlying group action.

4.3.1 Equivariance Properties of the $\star_M$-product

In this subsection, we develop an interpretation of the $\star_M$ product based on the representation theory of finite groups. Specifically, we derive conditions on a finite group G, a representation $\rho : G \to GL_{n_3}(\mathbb{C})$, the matrix $M \in \mathbb{C}^{n_3 \times n_3}$, and tubes $\mathbf{a}_{i,j} \in \mathbb{C}^{1\times 1\times n_3}$ which ensure that multiplication $\mathcal{A} \star_M \mathcal{B}$ is ρ-equivariant, that is, $\mathcal{A} \star_M (g \cdot \mathcal{B}) = g \cdot (\mathcal{A} \star_M \mathcal{B})$ for any $\mathcal{B}$ of compatible size. These conditions are concrete and can be interpreted in terms of linear equations involving the rows of the matrix M. The representation theoretic results may be of independent interest and have potential applications to other problems involving $\star_M$-products of tensors. So, we work in the more general setting of complex-valued tensors.

We first reduce the problem to determining the equivariance of the multiplication map $T_{\mathbf{a}} : \mathbb{C}^{1\times 1\times n_3} \to \mathbb{C}^{1\times 1\times n_3}$.

For notational convenience, if $\rho : G \to GL_{n_3}(\mathbb{C})$ is a representation of the group G, then for a tensor $\mathcal{A} \in \mathbb{C}^{n_1\times n_2\times n_3}$, we denote $g \cdot \mathcal{A} = \mathcal{A} \times_3 \rho(g)$.

Lemma 4.3.2. *Let $\mathcal{A} \in \mathbb{C}^{n_1\times n_2\times n_3}$ and fix a group G and representation $\rho : G \to GL_{n_3}(\mathbb{C})$. The map $\mathcal{A} \star_M \bullet : \mathbb{C}^{n_2\times k\times n_3} \to \mathbb{C}^{n_1\times k\times n_3}$ is ρ-equivariant for any $k \in \mathbb{N}$ if and only if the maps $T_{\mathbf{a}_{i,j}} : \mathbb{C}^{1\times 1\times n_3} \to \mathbb{C}^{1\times 1\times n_3}$ are ρ-equivariant for all $i \in [n_1]$ and $j \in [n_2]$.*

Proof. If the maps $T_{\mathbf{a}_{i,j}} : \mathbb{C}^{1\times 1\times n_3} \to \mathbb{C}^{1\times 1\times n_3}$ are ρ-equivariant for all $i \in [n_1]$ and $j \in [n_2]$ then for any $g \in G$, and any $\mathcal{B} \in \mathbb{C}^{n_2\times k\times n_3}$, we compute that

$$(\mathcal{A} \star_M (g \cdot \mathcal{B}))_{i,j,:} = \sum_{\ell=1}^{n_2} \mathbf{a}_{i,\ell} \star_M (g \cdot \mathbf{b}_{\ell,j}) = \sum_{\ell=1}^{n_2} g \cdot (\mathbf{a}_{i,\ell} \star_M \mathbf{b}_{\ell,k}) = g \cdot (\mathcal{A} \star_M \mathcal{B})_{i,j,:}.$$

Conversely, suppose that there is some $\mathbf{b} \in \mathbb{C}^{1\times 1\times n_3}$, $g \in G$, and $i \in [n_1], j \in [n_2]$ such that $\mathbf{a}_{i,j} \star_M (g \cdot \mathbf{b}) \neq g \cdot (\mathbf{a}_{i,j} \star_M \mathbf{b})$. Then, if $\mathcal{B} \in \mathbb{C}^{n_2\times 1\times n_3}$ is the tensor with $\mathcal{B}_{j,1} = \mathbf{b}$ we have $\mathcal{A} \star_M (g \cdot \mathcal{B}) \neq g \cdot (\mathcal{A} \star_M \mathcal{B})$. □

To make the ideas of equivariance concrete, we take the t-product as our motivating example. Specifically, we show that the multiplication maps $T_{\mathbf{a}} : \mathbb{C}^{n_3} \to \mathbb{C}^{n_3}$ are all equivariant under a representation of the cyclic group of order n_3.

Example 4.3.1 (Cyclic Equivariance of the tensor t-product)**.** The t-product is the $\star_M$-product with M taken to be the (unnormalized) discrete Fourier matrix $F \in \mathbb{C}^{n_3\times n_3}$ where each entry is a power of a complex root of unity; that is, $F_{j+1,k+1} = e^{2\pi \mathbb{i} jk/n_3}$ for $j, k = 0, \ldots, n_3 - 1$ [KHAN21]. The matrix representative of the linear transformation $T_{\mathbf{a}} : \mathbb{C}^{n_3} \to$

$\mathbb{C}^{n_3}$ is the circulant matrix

$$F^{-1}\operatorname{diag}(F\operatorname{vec}(\mathbf{a}))F = \begin{bmatrix} a_1 & a_{n_3} & \dots & a_2 \\ a_2 & a_1 & \dots & a_3 \\ \vdots & & \ddots & \vdots \\ a_{n_3} & a_{n_3-1} & \dots & a_1 \end{bmatrix} = \sum_{i=1}^{n_3} a_i Z^{i-1} \text{ where } Z = \begin{bmatrix} 0 & 0 & \dots & 1 \\ 1 & 0 & \dots & 0 \\ \vdots & \ddots & & \vdots \\ 0 & \dots & 1 & 0 \end{bmatrix} \tag{4.11}$$

Note that the matrix Z arises in the regular representation of the cyclic group C_{n_3} of order n_3, defined on a generator g of G by $\rho(g) = Z$.

Given some $\mathbf{a} \in \mathbb{C}^{1\times 1\times n_3}$, we show that the linear map $T_{\mathbf{a}} : \mathbb{C}^{1\times 1\times n_3} \to \mathbb{C}^{1\times 1\times n_3}$ is ρ-equivariant. Specifically, if $g \in C_{n_3}$ be the generator of the cyclic group of order n_3. Then, for any $\mathbf{b} \in \mathbb{C}^{1\times 1\times n_3}$, we have

$$\mathbf{a} \star_F (g \cdot \mathbf{b}) \equiv \left(\sum_{i=1}^{n_3} a_i Z^{i-1}\right)(Z\operatorname{vec}(\mathbf{b})) = Z\left(\sum_{i=1}^{n_3} a_i Z^{i-1}\right)\operatorname{vec}(\mathbf{b}) \equiv g \cdot (\mathbf{a} \star_F \mathbf{b}). \tag{4.12}$$

$\diamond$

In general, the multiplication maps $T_{\mathbf{a}}$ will only be equivariant for some subset of tubes. We want to characterize this subset.

Recall from Section 2.5 that if G is a finite group and ρ is a faithful representation, then the decomposition of $\mathbb{C}^{n_3}$ into irreducible representations of ρ takes the form

$$\mathbb{C}^{n_3} \simeq V_1 \oplus V_2 \oplus \cdots \oplus V_m,$$

where $\dim V_j = d_j$. This gives a corresponding (non-canonical) basis of $\mathbb{C}^{n_3}$

$$B = \{v_{1,1}, v_{1,2}, \dots, v_{1,d_1}, v_{2,1}, \dots, v_{m,d_m}\},$$

where $V_j = \operatorname{span}\{v_{j,1}, v_{j,2}, \dots, v_{j,d_j}\}$. If the matrix M changes basis from the standard basis

to B, then for each $g \in G$, the matrix $M\rho(g)M^{-1}$ will be block diagonal, with blocks of size $d_1, d_2, \ldots, d_m$. Schur's Lemma (Lemma 2.5.4) characterizes ρ-equivariance of linear transformations $\mathbb{C}^{n_3} \to \mathbb{C}^{n_3}$. We connect this characterization to the matrix M.

Theorem 4.3.3. *Fix a finite group G and a representation $\rho : G \to GL_{n_3}(\mathbb{C})$. Suppose that the decomposition of $\mathbb{C}^{n_3}$ into irreducible representations of ρ is given by*

$$\mathbb{C}^{n_3} \simeq V_1 \oplus V_2 \oplus \cdots \oplus V_m,$$

where $\dim_{\mathbb{C}}(V_j) = d_j$. Let $M \in GL_{n_3}(\mathbb{C})$ be a change of basis compatible with this decomposition. There is a vector subspace $W_\rho \subseteq \mathbb{C}^{1\times 1\times n_3}$ of tubes $\mathbf{a}$ for which the multiplication map $T_{\mathbf{a}}(\mathbf{b}) = \mathbf{a} *_M \mathbf{b}$ is ρ-equivariant given by

$$W_\rho = \left\{ \text{tube}(x) \;\middle|\; \text{diag}(Mx) = \begin{bmatrix} c_1 I_{d_1} & & & \\ & c_2 I_{d_2} & & \\ & & \ddots & \\ & & & c_m I_{d_m} \end{bmatrix}, \; c_1, c_2, \ldots, c_m \in \mathbb{C} \right\}.$$

Proof. Note that by the choice of M, we have that the structured matrix for multiplication by $\mathbf{a}$ factors as in the diagram:

$$\begin{array}{ccc} \mathbb{C}^{n_3} & \xrightarrow{M} & V_1 \oplus V_2 \oplus \cdots \oplus V_m \\ \downarrow{\scriptstyle T_{\mathbf{a}}} & & \downarrow{\scriptstyle \text{diag}(M\,\text{vec}(\mathbf{a}))} \\ \mathbb{C}^{n_3} & \xleftarrow[M^{-1}]{} & V_1 \oplus V_2 \oplus \cdots \oplus V_m \end{array}$$

In particular, the map $T_{\mathbf{a}}$ is ρ-equivariant if and only if the map given by $\text{diag}(M\,\text{vec}(\mathbf{a}))$ is ρ-equivariant. By Schur's Lemma (Lemma 2.5.4), this happens if and only if the restrictions to each V_i are given by multiplication by a constant. This happens if and only if $\text{diag}(M\,\text{vec}(\mathbf{a}))$ has the desired block structure. □

Theorem 4.3.3 gives explicit conditions linear conditions on the space of tubes for the

multiplication by **a** map to be equivariant in terms of the rows of the matrix M. To do so, we form an auxillary matrix $V \in \mathbb{C}^{n_3 \times m}$, with block structure

$$V = \begin{bmatrix} \mathbb{1}_{d_1} & 0 & \dots & 0 \\ 0 & \mathbb{1}_{d_2} & & 0 \\ \vdots & & \ddots & \\ 0 & 0 & \dots & \mathbb{1}_{d_m} \end{bmatrix},$$

where $\mathbb{1}_{d_i} \in \mathbb{C}^{d_i \times 1}$ is the constant vector of all ones. Then, for any $a \in \mathbb{C}^{n_3}$, $\text{tube}(a) \in W_\rho$ if and only if we can find coefficients $c \in \mathbb{C}^m$ such that $Ma = Vc$.

Example 4.3.2 (The Symmetric Group S_3)**.** Consider the symmetric group

$$S_3 = \langle \sigma, \tau \mid \sigma^2 = \tau^3 = 1, \sigma\tau = \tau^2\sigma \rangle.$$

Let $\rho : S_3 \to GL_3(\mathbb{C})$ be the permutation representation so that

$$\rho(\sigma) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \rho(\tau) = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

The decomposition of $\mathbb{C}^3$ into irreducibles is then $\mathbb{C}^3 \simeq V_1 \oplus V_2$ for $V_1 = \text{span}_{\mathbb{C}}\{(1,1,1)\}$ and $V_2 = \text{span}_{\mathbb{C}}\{(1,-1,0),(1,0,-1)\}$. We construct M so that the rows correspond to these bases of V_1 and V_2.

$$M = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix} \quad \text{and} \quad M^{-1} = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix}.$$

To determine the ρ-equivariant transformations, $T_{\mathbf{a}}$, we solve for **a** using the relation

$Ma = Vc$. In particular, we compute the kernel of the augmented matrix $\left[\begin{array}{c|c} M & V \end{array}\right]$; that is,

$$\ker\left(\left[\begin{array}{ccc|cc} 1 & 1 & 1 & 1 & 0 \\ 1 & -1 & 0 & 0 & 1 \\ 1 & 0 & -1 & 0 & 1 \end{array}\right]\right) = \left\{ a \in \mathbb{C}^3, c_1, c_2 \in \mathbb{C} \;\middle|\; \begin{array}{rcl} a_1 + a_2 + a_3 &=& c_1 \\ a_1 - a_2 &=& c_2 \\ a_1 - a_3 &=& c_2 \end{array} \right\}.$$

Theorem 4.3.3 then implies that $T_{\mathbf{a}}$ is ρ-equivariant if and only if $a_1 - a_2 = a_1 - a_3 = c_2$ for some constant $c_2 \in \mathbb{C}$. Therefore, the set of tubes which yield ρ-equivariant transformations are

$$W_\rho = \{\mathbf{a} \in \mathbb{C}^{1\times 1\times 3} \mid a_1 - a_2 = a_1 - a_3\} = \operatorname{span}_{\mathbb{C}}\{\operatorname{tube}(1,0,0), \operatorname{tube}(0,1,1)\}$$

Let $\mathbf{a}_1 = \operatorname{tube}(1,0,0)$ and let $\mathbf{a}_2 = \operatorname{tube}(0,1,1)$ be the basis tubes of the W_ρ. Let $\mathbf{b} \in \mathbb{C}^{1\times 1\times 3}$ be arbitrary and consider $T_{\mathbf{a}}(\mathbf{b})$ for each basis vector; that is,

$$T_{\mathbf{a}_1}(\mathbf{b}) = \mathbf{a}_1 \star_M \mathbf{b} \equiv \underbrace{M^{-1}\operatorname{diag}(M\operatorname{vec}(\mathbf{a}_1))M}_{I_3}\operatorname{vec}(\mathbf{b}) = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}, \quad \text{and} \tag{4.13a}$$

$$T_{\mathbf{a}_2}(\mathbf{b}) = \mathbf{a}_2 \star_M \mathbf{b} \equiv \underbrace{M^{-1}\operatorname{diag}(M\operatorname{vec}(\mathbf{a}_2))M}_{Z_3}\operatorname{vec}(\mathbf{b}) = \begin{bmatrix} b_2 + b_3 \\ b_1 + b_3 \\ b_1 + b_2 \end{bmatrix}. \tag{4.13b}$$

We see that $T_{\mathbf{a}_i}(g \cdot \mathbf{b}) = g \cdot \mathbf{b}$ for any $g \in S_3$, since S_3 acts by permuting indices. ◇

There are several immediate corollaries to Theorem 4.3.3.

Corollary 4.3.4. *Fix a finite group G, a representation $\rho : G \to GL_{n_3}(\mathbb{C})$, and $M \in GL_{n_3}(\mathbb{C})$. The multiplication map $T_{\mathbf{a}}(\mathbf{b}) = \mathbf{a} *_M \mathbf{b}$ is ρ-equivariant for all $\mathbf{a} \in \mathbb{C}^{1\times 1\times n_3}$ if and only if $M\rho(g)M^{-1}$ is diagonal for each $g \in G$.*

Proof. By Theorem 4.3.3 if $T_{\mathbf{a}}$ is ρ-equivariant for all $\mathbf{a} \in \mathbb{C}^{1\times 1\times n_3}$, then $W_\rho = \mathbb{C}^{1\times 1\times n_3}$. By the description of W_ρ, any diagonal matrix D arises as $\operatorname{diag}(M \operatorname{vec}(\mathbf{a}))$ for some $\mathbf{a} \in W_\rho$ and since the multiplication map $T_{\mathbf{a}}$ is ρ-equivariant, we have that for any $g \in G$,

$$DM\rho(g)M^{-1} = M\rho(g)M^{-1}D.$$

This implies that $M\rho(g)M^{-1}$ is diagonal by taking D to be the matrices with $D_{i,i} = 1$ and $D_{j,j} = 0$ for i ranging over $[n_3]$ and $j \neq i$. The other converse follows since diagonal matrices commute. □

Corollary 4.3.5. *If G is a nonabelian group and ρ a faithful representation of G, then there exist tubes $\mathbf{a} \in \mathbb{C}^{1\times 1\times n_3}$ such that the multiplication map $T_{\mathbf{a}}(\mathbf{b}) = \mathbf{a} \star_M \mathbf{b}$ is* not *ρ-equivariant.*

Proof. If ρ is a faithful representation of G and $g, h \in G$ have $gh \neq hg$, then $\rho(g)\rho(h) \neq \rho(h)\rho(g)$. Because the matrices $\rho(g)$ and $\rho(h)$ do not commute, they are not simultaneously diagonalizable and the result follows by Corollary 4.3.4. □

So, while the t-product provides an example for which every multiplication map $T_{\mathbf{a}}$ is ρ-equivariant for the cyclic group, it is an anomoly in the sense that W_ρ will be a proper subspace in most cases, for example symmetric groups.

4.3.2 Connection to Invariant SDPs

We now use the representation theoretic interpretation of the $\star_M$ product to study invariant semidefinite programs. The study of invariant SDPs was initiated by Gaterman and Parillo in [GP04]. We overview the main ideas here.

An SDP is said to be invariant with respect to the group representation ρ the cost function and feasible set are invariant under the action of G on symmetric matrices. Note that if $\rho : G \to GL_{n_3}(\mathbb{C})$ is a representation, then G acts on symmetric $n_3 \times n_3$ matrices by $g \cdot X = \rho(g)^\top X \rho(g)$. So, if $\rho(g)$ is orthogonal for all g and the feasible set is invariant under the action of G, then $\rho(g)X = X\rho(g)$, i.e., the linear transformation defined by X is

ρ-equivariant. So, if $\mathbb{C}^3 \simeq V_1 \oplus V_2 \oplus \cdots \oplus V_m$ is the decomposition of $\mathbb{C}^3$ into irreducibles with $\dim V_i = d_i$, then we can change basis so that $A = P^{-1}XP$ has block format

$$A = \begin{bmatrix} A_{1,1} & A_{1,2} & \dots & A_{1,m} \\ A_{2,1} & A_{2,2} & \dots & A_{2,m} \\ \vdots & & \ddots & \vdots \\ A_{m,1} & A_{m,2} & \dots & A_{m,m} \end{bmatrix},$$

where $A_{i,j} : V_i \to V_j$. Furthermore, by Schur's Lemma, we know that $A_{i,j} = 0$ if $V_i \neq V_j$ and $A_{i,j} = c_{i,j} I_{d_i}$ if $V_i = V_j$. Once again changing basis allows us to block diagonalize A, where the size of the blocks depend on d_i and the multiplicity of each irreducible V_i (that is, the number of copies it has in the decomposition of $\mathbb{C}^{n_3}$).

So, the approach to invariant SDPs involves a change of basis related to a group action and results in a block diagonalization. In the previous sections, we have seen that M-PSD tensors correspond to block diagonal PSD matrices and that the choice of change of basis M can be related to a group action. So, it is natural to connect M-SDP problems with invariant SDP problems.

The equality constraints in (M-SDP) allow one to enforce that multiplication by the tensor $\mathcal{X}$ satisfies group equivariance properties. Indeed, if $\rho : G \to GL_{n_3}(\mathbb{R})$ and M are as in Theorem 4.3.3, then the condition that multiplication is ρ-equivariant is a linear condition on the space of tubes and can therefore be written using equalities of the form $\langle \mathcal{A}^{(i)}, \mathcal{X} \rangle = 0$ for some symmetric tensors $\mathcal{A}^{(i)}$. In particular, group equivariance can be encoded in the constraints of (M-SDP), and such cases yield invariant semidefinite programs.

Theorem 4.3.6. *Let G be a finite group and $\rho : G \to GL_{n_3}(\mathbb{R})$ a representation with $\rho(g)$ orthogonal for each $g \in G$. Let $M \in \mathbb{R}^{n_3 \times n_3}$ and W_ρ be as in the statement of Theorem 4.3.3. Set $\mathcal{L} \subseteq \mathbb{R}^{n \times n \times n_3}$ to be the vector space of symmetric tensors whose tubes are elements of W_ρ. Then, for any symmetric tensors $\mathcal{C}, \mathcal{A}^{(i)}$ of size $n \times n \times n_3$ and any scalars b_i, $i \in [k]$, the M-SDP*

$$\max\langle \mathcal{C}, \mathcal{X} \rangle \ \ s.t.\ \langle \mathcal{A}_i, \mathcal{X} \rangle = b_i \ for\ i \in [k], \ \ \mathcal{X} \in \mathrm{PSD}_M^n \cap \mathcal{L}$$

is an invariant SDP with respect to the representation $\hat{\rho} : G \to GL_{nn_3}(\mathbb{R})$ *given by* $\hat{\rho}(g) = I_n \otimes \rho(g)$*, where* I_n *is the* $n \times n$ *identity matrix and* $\otimes$ *is the matrix kronecker product.*

Proof. Note that for any tensor $\mathcal{X} \in \mathbb{R}^{n \times n \times n_3}$, the multiplication map $\mathcal{X} \star_M \bullet : \mathbb{R}^{n \times 1 \times n_3} \to \mathbb{R}^{n \times 1 \times n_3}$ gives a linear transformation $\mathbb{R}^{nn_3} \to \mathbb{R}^{nn_3}$ with matrix representative

$$\hat{X} = \begin{bmatrix} M^{-1}D_{1,1}M & M^{-1}D_{1,2}M & \dots & M^{-1}D_{1,n}M \\ M^{-1}D_{2,1}M & M^{-1}D_{2,2}M & \dots & M^{-1}D_{2,n}M \\ \vdots & \vdots & \ddots & \vdots \\ M^{-1}D_{n,1}M & M^{-1}D_{n,2}M & \dots & M^{-1}D_{n,n}M \end{bmatrix}, \qquad D_{i,j} = \mathrm{diag}(M\,\mathrm{vec}(\mathbf{x}_{i,j}))$$

Now, if $\mathcal{X} \in \mathcal{L}$, then $\mathbf{x}_{i,j} \in W_\rho$ for each i, j and therefore $M^{-1}D_{i,j}M$ commutes with $\rho(g)$ for any $g \in G$. Since $\rho(g)^\top = (\rho(g))^{-1} = \rho(g^{-1})$,

$$\rho(g)^\top M^{-1} D_{i,j} M = M^{-1} D_{i,j} M \rho(g)^\top.$$

It then follows that for any $g \in G$,

$$\hat{\rho}(g)^\top \hat{X} \hat{\rho}(g) = \hat{X} \hat{\rho}(g)^\top \hat{\rho}(g) = \hat{X}.$$

To conclude, we need to show that if $\mathcal{B} \in \mathbb{R}^{n \times n \times n_3}$ is a symmetric tensor and $\hat{B}$ is its matrix representative, then $\langle \mathcal{B}, \mathcal{X} \rangle = \mathrm{trace}(\hat{B}\hat{X})$. Note that if $\mathbf{a}, \mathbf{b} \in \mathbb{R}^{1 \times 1 \times n_3}$ are tubes, then because M is orthogonal,

$$\begin{aligned}\langle \mathbf{a}, \mathbf{b} \rangle &= \langle M \operatorname{vec}(\mathbf{a}), M \operatorname{vec}(\mathbf{b}) \rangle \\ &= \operatorname{trace}(\operatorname{diag}(M \operatorname{vec}(\mathbf{a})) \operatorname{diag}(M \operatorname{vec}(\mathbf{b}))) \\ &= \operatorname{trace}(M^{-1} \operatorname{diag}(M \operatorname{vec}(\mathbf{a})) M M^{-1} \operatorname{diag}(M \operatorname{vec}(\mathbf{b})) M)\end{aligned}$$

It then follows that

$$\begin{aligned}\langle \mathcal{B}, \mathcal{X} \rangle &= \sum_{i=1}^{n} \left(\sum_{k=1}^{n} \langle \mathbf{b}_{i,k}, \mathbf{x}_{k,i} \rangle \right) \\ &= \sum_{i=1}^{n} \left(\sum_{k=1}^{n} \operatorname{trace}(M^{-1} \operatorname{diag}(M \operatorname{vec}(\mathbf{b}_{i,k})) M M^{-1} \operatorname{diag}(M \operatorname{vec}(\mathbf{x}_{k,i})) M) \right) \\ &= \operatorname{trace}(\hat{B} \hat{X}).\end{aligned}$$

□

We conclude this section with a simple example which highlights the connection between M-SDP and invariant SDP.

Example 4.3.3. Consider the symmetric group $S_2 = \langle id, \sigma \mid \sigma^2 = id \rangle$ and let $\rho : S_2 \to GL_2(\mathbb{R})$ be the representation with

$$\rho(\sigma) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}. \tag{4.14}$$

The corresponding decomposition of $\mathbb{R}^2$ into irreducibles is $\mathbb{R}^2 \simeq \operatorname{span}_{\mathbb{R}}\{(1,1)\} \oplus \operatorname{span}_{\mathbb{R}}\{(1,-1)\}$. Note that each irreducible in this decomposition has multiplicity 1.

Let $M = \begin{bmatrix} \alpha & \alpha \\ \alpha & -\alpha \end{bmatrix}$, where $\alpha = \frac{1}{\sqrt{2}}$. Since $M\rho(\sigma)M^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ is diagonal, it follows from Corollary 4.3.4 that the multiplication map $T_{\mathbf{a}}$ is ρ-equivariant for all tubes $\mathbf{a} \in \mathbb{R}^{1\times 1\times 2}$. That is, with notation as in the statement of Theorem 4.3.6, we have $\mathcal{L} = \mathbb{R}^{2\times 2\times 2}$ and therefore $\mathrm{PSD}_M^2 \cap \mathcal{L} = \mathrm{PSD}_M^2$.

Consider the following tensor $\mathcal{X} \in \mathbb{R}^{2\times 2\times 2}$ and its transformed $\widehat{\mathcal{X}} \in \mathbb{R}^{2\times 2\times 2}$ given by

$$\mathcal{X}_{:,:,1} = \begin{bmatrix} x_1 & 1 \\ 1 & y_1 \end{bmatrix} \qquad \mathcal{X}_{:,:,2} = \begin{bmatrix} x_2 & 1 \\ 1 & y_2 \end{bmatrix} \tag{4.15a}$$

$$\widehat{\mathcal{X}}_{:,:,1} = \begin{bmatrix} \alpha(x_1+x_2) & 2\alpha \\ 2\alpha & \alpha(y_1+y_2) \end{bmatrix}, \qquad \widehat{\mathcal{X}}_{:,:,2} = \begin{bmatrix} \alpha(x_1-x_2) & 0 \\ 0 & \alpha(y_1-y_2) \end{bmatrix}. \tag{4.15b}$$

We will show that the corresponding M-SDP

$$\min_{(x_1,x_2),(y_1,y_2)} \langle \mathcal{I}^*, \mathcal{X} \rangle \qquad \text{s.t.} \qquad \mathcal{X} \succeq_M 0 \tag{4.16}$$

is invariant under the group action of S_2.

By Proposition 4.2.10, we can rewrite (4.16) using block matrices with blocks from the transform domain resulting in the equivalent M-SDP

$$\min_{(x_1,x_2),(y_1,y_2)} \langle \mathcal{I}^*, \mathcal{X} \rangle \qquad \text{s.t.} \qquad \underbrace{\begin{bmatrix} \widehat{\mathcal{X}}_{:,:,1} & \\ & \widehat{\mathcal{X}}_{:,:,2} \end{bmatrix}}_{\texttt{bdiag}(\widehat{\mathcal{X}})} \succeq 0. \tag{4.17}$$

Because the frontal slices in the transform domain are symmetric matrices, the block diagonal matrix $\texttt{bdiag}(\widehat{\mathcal{X}})$ is a symmetric matrix that is invariant under the induced action $\hat{\rho}(\sigma) = I_2 \otimes \rho(\sigma)$. Moreover, the objective function $\langle \mathcal{I}^*, \mathcal{X} \rangle = x_1 + x_2 + y_1 + y_2$ is invariant under $\hat{\rho}$, which acts as permutations $(x_i, y_i) \mapsto (x_{g(i)}, y_{g(i)})$ for $g \in S_2$. Since the objective function and feasible region of (4.17) are both invariant under $\hat{\rho}$, the problem (4.17) is an invariant SDP.

Note that the decomposition of $\mathbb{R}^4$ into irreducible representations of S_2 consists of the trivial representation and the sign representation, each with multiplicity 2, and the corresponding blocks in (4.17) are 2×2.

$\diamond$

4.3.3 Application: Invariant Quadratic Forms

A well-known application of semidefinite programming is as a relaxation of polynomial optimization problems. The central idea is to search for nonnegativity certificates of a polynomial p in the form of a decomposition of p into a sum of squares; see e.g., [Las01, Par03]. The maximum value of γ such that $p(x) - \gamma$ is a sum of squares (SOS) is then a (possibly strict) lower bound on $\inf_{x \in \mathbb{R}^n} p(x)$. Additionally, a polynomial $p(x) - \gamma$ is a sum of squares if and only if there is a positive semidefinite Gram matrix Q such that $p(x) - \gamma = \xi^\top Q \xi$, where ξ is a vector of all monomials of degree at most $d = \deg(p)$. Searching for the maximum γ such that $p(x) - \gamma$ is a sum of squares is therefore a semidefinite programming problem. Here, we use M-PSD tensors to study a subset of SOS polynomials. As a starting point, we mirror the definition of block-circulant SOS polynomial from [ZHH22] to define M-SOS polynomials.

It will be convenient to use the notation of fold to turn a vector of length mk into a tensor of format $m \times 1 \times k$. More precisely, if $v = \begin{bmatrix} v_1^\top & v_2^\top & \dots & v_m^\top \end{bmatrix}^\top$, where $v_i \in \mathbb{R}^k$ for each $i \in [m]$, then we set

$$\operatorname*{fold}_k(v) = \begin{bmatrix} \operatorname{tube}(v_1) & \operatorname{tube}(v_2) & \dots & \operatorname{tube}(v_m) \end{bmatrix}^\top \in \mathbb{R}^{m \times 1 \times k}.$$

Definition 4.3.7 (M-SOS Polynomial)**.** *Suppose that* $f \in \mathbb{R}[x_1, x_2, \dots, x_k]_{\leqslant 2d}$ *and that* n_3 *divides* $N = \binom{k+d}{d}$. *Set* $[x]_d = \begin{bmatrix} 1 & x_1 & x_2 & \dots & x_k & x_1^2 & x_1 x_2 & \dots & x_k^d \end{bmatrix}^\top$ *and* $\mathcal{X}_d^{(n_3)} = \operatorname{fold}_{n_3}([x]_d) \in \mathbb{R}^{(N/n_3) \times 1 \times n_3}$. *We say that* f *is an* M-SOS *polynomial if there are* $\mathcal{Q}_1, \dots, \mathcal{Q}_r \in \mathbb{R}^{1 \times (N/n_3) \times n_3}$ *such that*

$$f(x) = \sum_{j=1}^{r} \operatorname{vec}(\mathcal{Q}_j *_M \mathcal{X}_d^{(n_3)})^\top \operatorname{vec}(\mathcal{Q}_j *_M \mathcal{X}_d^{(n_3)}). \tag{4.18}$$

Remark 4.3.1. Note that if $n_3 = 1$, then M is a nonzero real number and a polynomial is M-SOS if and only if it is SOS.

The definition (4.18) highlights that the polynomial f is a sum of squares, since for any tensor $\mathcal{A}$, the inner product expands as $\langle \mathcal{A}, \mathcal{A} \rangle = \|\mathcal{A}\|_F^2 = \sum a_{i,j,k}^2$. As discussed above, a

sum of squares polynomial has a positive semidefinite Gram matrix. The additional structure implied by the $\star_M$ product in the decomposition (4.18) yields an M-PSD tensor.

Proposition 4.3.8. *f is an M-SOS polynomial if and only if there is $\mathcal{Q} \in \mathrm{PSD}_M^n$ such that*

$$f(x) = \langle \mathcal{X}_d^{(n_3)}, \mathcal{Q} *_M \mathcal{X}_d^{(n_3)} \rangle.$$

Proof. Let $\mathcal{Q}_1, \ldots, \mathcal{Q}_r \in \mathbb{R}^{1 \times (N/n_3) \times n_3}$ such that

$$f(x) = \sum_{j=1}^{r} \mathrm{vec}(\mathcal{Q}_j *_M \mathcal{X}_d^{(n_3)})^\top \mathrm{vec}(\mathcal{Q}_j *_M \mathcal{X}_d^{(n_3)}).$$

Then, because $\mathrm{vec}(\mathcal{Q}_j *_M \mathcal{X}_d^{(n_3)})^\top \mathrm{vec}(\mathcal{Q}_j *_M \mathcal{X}_d^{(n_3)}) = \langle \mathcal{Q}_j *_M \mathcal{X}_d^{(n_3)}, \mathcal{Q}_j *_M \mathcal{X}_d^{(n_3)} \rangle$ for each $j \in [r]$ it follows from Lemma 4.1.9 and Proposition 4.2.4 that

$$f(x) = \sum_{j=1}^{r} \langle \mathcal{Q}_j *_M \mathcal{X}_d^{(n_3)}, \mathcal{Q}_j *_M \mathcal{X}_d^{(n_3)} \rangle = \left\langle \mathcal{X}_d^{(n_3)}, \left(\sum_{j=1}^{r} \mathcal{Q}_j^\top *_M \mathcal{Q}_j \right) *_M \mathcal{X}_d^{(n_3)} \right\rangle$$

and $\mathcal{Q} = \left(\sum_{j=1}^{r} \mathcal{Q}_j^\top *_M \mathcal{Q}_j \right)$ is M-PSD.

Conversely, let $\mathcal{Q} \in \mathrm{PSD}_M^n$ be such that $f(x) = \langle \mathcal{X}_d^{(n_3)}, \mathcal{Q} *_M \mathcal{X}_d^{(n_3)} \rangle$. By Theorem 4.2.5, there is a decomposition $\mathcal{Q} = \mathcal{B}^\top *_M \mathcal{B}$ with $\mathcal{B} \in \mathbb{R}^{r \times n \times n_3}$. This yields that f is M-SOS by taking $\mathcal{Q}_i = \mathcal{B}_{i,:,:}$ for each $i \in [r]$. □

There has been recent interest in sums of squares polynomials which are invariant under a group action on the variables, see e.g., [GP04, RTAL13, HHS21]. In this setting, the theory of invariant SDPs is leveraged to better understand the SDPs which certify that an invariant sos polynomial is a sum of squares. Since the results of this section relate the $\star_M$-product to invariant SDPs, we work towards an analogy in the tensor case. While every M-SDP corresponding to an invariant SOS program results in an invariant polynomial, the converse is not true without an additional assumption on the multiplicity of the irreducible representations appearing in the decomposition of $\mathbb{R}^{n_3}$.

Theorem 4.3.9. *Fix an orthogonal matrix $M \in \mathbb{R}^{n_3 \times n_3}$. Let G be a finite group and $\rho : G \to GL_{n_3}(\mathbb{R})$ be a representation with each $\rho(g)$ orthogonal, and set $W_\rho \subseteq \mathbb{R}^{1\times 1\times n_3}$ to be the vector subspace of tubes for which multiplication is ρ-equivariant. Consider a quadratic form $f \in \mathbb{R}[x_{i,j} \mid i \in [m], j \in [n_3]]_2$ and the action of G defined on variables by $g \cdot \begin{bmatrix} x_{i,1}, x_{i,2}, \ldots, x_{i,n_3} \end{bmatrix}^\top = \rho(g) \begin{bmatrix} x_{i,1}, x_{i,2}, \ldots, x_{i,n_3} \end{bmatrix}^\top$ for each $i \in [m]$.*

If $f = \langle \mathcal{X}_1^{(n_3)}, \mathcal{Q} *_M \mathcal{X}_1^{(n_3)} \rangle$ for some M-PSD tensor $\mathcal{Q}$ with tubes in W_ρ then f is SOS and invariant under the action of G. The converse holds if each irreducible representation in the decomposition of $\mathbb{R}^{n_3}$ appears with multiplicity one.

Proof. Suppose that $f = \langle \mathcal{X}_1^{(n_3)}, \mathcal{Q} *_M \mathcal{X}_1^{(n_3)} \rangle$ where $\mathcal{Q}$ is M-PSD and has tubes in W_ρ. Let

$$\xi = \begin{bmatrix} x_{1,1} & x_{1,2} & \ldots & x_{1,n_3} & x_{2,1} & \ldots & x_{m,n_3} \end{bmatrix}^\top$$

be the length mn_3 vector consisting of all variables, and set

$$\hat{Q} = \begin{bmatrix} M^{-1}D_{1,1}M & M^{-1}D_{1,2}M & \ldots & M^{-1}D_{1,m}M \\ M^{-1}D_{2,1}M & M^{-1}D_{2,2}M & \ldots & M^{-1}D_{2,m}M \\ \vdots & \vdots & \ddots & \vdots \\ M^{-1}D_{m,1}M & M^{-1}D_{m,2}M & \ldots & M^{-1}D_{m,m}M \end{bmatrix} \qquad D_{i,j} = \operatorname{diag}(M \operatorname{vec}(\mathbf{q}_{i,j})).$$

Then, $f = \xi^\top \hat{Q} \xi$. The action of $g \in G$ on f is given by $g \cdot f = ((I_m \otimes \rho(g))\xi)^\top \hat{Q}((I_m \otimes \rho(g))\xi)$. Now, since $\mathbf{q}_{i,j} \in W_\rho$ for each $i, j \in [m]$, it follows that $\rho(g)$ commutes with $M^{-1}D_{i,j}M$ for each $g \in G$. Since $\rho(g)$ is orthogonal for each $g \in G$, it then follows that

$$g \cdot f = ((I_m \otimes \rho(g))\xi)^\top \hat{Q}((I_m \otimes \rho(g))\xi) = \xi^\top \hat{Q} \xi = f$$

Conversely, suppose that f is SOS and invariant under the action of G. Then, there is a $(mn_3) \times (mn_3)$ positive semidefinite matrix Q such that $f = \xi^\top Q \xi$ and $(I_m \otimes \rho(g))^\top Q (I_m \otimes \rho(g)) = Q$. Let

$$Q = \begin{bmatrix} Q_{1,1} & Q_{1,2} & \dots & Q_{1,m} \\ Q_{2,1} & Q_{2,2} & \dots & Q_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ Q_{m,1} & Q_{m,2} & \dots & Q_{m,m} \end{bmatrix},$$

where each $Q_{i,j}$ is $n_3 \times n_3$. Now, for each i, j and each $g \in G$, $\rho(g)^\top Q_{i,j} \rho(g) = Q_{i,j}$ since $(I_m \otimes \rho(g))^\top Q (I_m \otimes \rho(g)) = Q$. By hypothesis, the decomposition of $\mathbb{R}^{n_3}$ into irreducibles is given by

$$\mathbb{R}^{n_3} \simeq V_1 \oplus V_2 \oplus \cdots \oplus V_s,$$

where each V_i is unique. Since M corresponds to a symmetry adapted basis of $\mathbb{R}^{n_3}$ it follows from Schur's Lemma (Lemma 2.5.4) that

$$M Q_{i,j} M^{-1} = \begin{bmatrix} c_1 I_{d_1} & & & \\ & c_2 I_{d_2} & & \\ & & \ddots & \\ & & & c_s I_{d_s} \end{bmatrix}$$

where $d_t = \dim V_t$. By Theorem 4.3.3, there exists a tube $\mathbf{q}_{i,j} \in W_\rho$ such that $M Q_{i,j} M^{-1} = \operatorname{diag}(M \operatorname{vec}(\mathbf{q}_{i,j}))$. This in turn implies that

$$Q = \begin{bmatrix} M^{-1} D_{1,1} M & M^{-1} D_{1,2} M & \dots & M^{-1} D_{1,m} M \\ M^{-1} D_{2,1} M & M^{-1} D_{2,2} M & \dots & M^{-1} D_{2,m} M \\ \vdots & \vdots & \ddots & \vdots \\ M^{-1} D_{m,1} M & M^{-1} D_{m,2} M & \dots & M^{-1} D_{m,m} M \end{bmatrix} \qquad D_{i,j} = \operatorname{diag}(M \operatorname{vec}(\mathbf{q}_{i,j})).$$

and therefore if $\mathcal{Q} \in \mathbb{R}^{m \times m \times n_3}$ is the tensor with tubes $\mathcal{Q}_{i,j,:} = \mathbf{q}_{i,j}$, then $f = \langle \mathcal{X}_1^{(n_3)}, \mathcal{Q} \star_M \mathcal{X}_1^{(n_3)} \rangle$. □

We conclude this section with examples demonstrating Theorem 4.3.9 and its limitations.

Example 4.3.4. Let S_3 be the symmetric group on three elements and $\rho : S_3 \to GL_3(\mathbb{R})$ be the permutation representation. Set

$$M = \begin{bmatrix} -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{bmatrix}, \quad \mathbf{q}_1 = (-\sqrt{3}) \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \text{ and } \mathbf{q}_2 = (\sqrt{18}) \begin{bmatrix} -1 \\ \sqrt{3} \\ 1 \end{bmatrix}$$

The map $T_{\mathbf{a}} : \mathbb{R}_M \to \mathbb{R}_M$ is ρ-equivariant if and only if $\mathbf{a} \in \text{span}_{\mathbb{R}}\{\mathbf{q}_1, \mathbf{q}_2\}$. We construct an M-SOS quadratic form $f \in \mathbb{R}[x_1, x_2, x_3, y_1, y_2, y_3]$ as

$$\begin{aligned} f(x,y) &= \left\langle \begin{bmatrix} x \\ y \end{bmatrix}, \left(\begin{bmatrix} \mathbf{q}_1 + \mathbf{q}_2 \\ \mathbf{q}_2 \end{bmatrix} *_M \begin{bmatrix} \mathbf{q}_1 + \mathbf{q}_2 & \mathbf{q}_2 \end{bmatrix}^\top \right) *_M \begin{bmatrix} x \\ y \end{bmatrix} \right\rangle \\ &= \left((-3-\sqrt{2})x_1 + (3-\sqrt{2})x_2 + (3-\sqrt{2})x_3 - (4+\sqrt{2})y_1 + (2-\sqrt{2})y_2 + (2-\sqrt{2})y_3 \right)^2 \\ &+ \left((3-\sqrt{2})x_1 + (-3-\sqrt{2})x_2 + (3-\sqrt{2})x_3 + (2-\sqrt{2})y_1 - (4+\sqrt{2})y_2 + (2-\sqrt{2})y_3 \right)^2 \\ &+ \left((3-\sqrt{2})x_1 + (3-\sqrt{2})x_2 + (-3-\sqrt{2})x_3 + (2-\sqrt{2})y_1 + (2-\sqrt{2})y_2 - (4+\sqrt{2})y_3 \right)^2 \end{aligned}$$

Note that f is invariant under

$$(x_1, x_2, x_3, y_1, y_2, y_3) \mapsto (x_{g(1)}, x_{g(2)}, x_{g(3)}, y_{g(1)}, y_{g(2)}, y_{g(3)}) \text{ for } g \in S_3.$$

$\diamond$

We also show that the condition that each irreducible representation appearing in the decomposition of $\mathbb{R}^{n_3}$ appears with multiplicity is necessary via an example.

Example 4.3.5. Let $S_2 = \{id, \sigma\}$, and $\rho : S_2 \to GL_3(\mathbb{R})$ be the representation with

$\rho(\sigma) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$. That is, σ acts on vectors by swapping the second and third coordinates. Note that the decomposition of $\mathbb{R}^3$ into irreducible representations is given by

$$\mathbb{R}^3 \simeq \operatorname{span}(1,0,0) \oplus \operatorname{span}(0,1,1) \oplus \operatorname{span}(0,1,-1),$$

and that the trivial representation appears with multiplicity 2, given by $\operatorname{span}(1,0,0)$ and $\operatorname{span}(0,1,1)$.

Let $\alpha = \frac{1}{\sqrt{2}}$ and $M = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \alpha & \alpha \\ 0 & \alpha & -\alpha \end{bmatrix}$. Note that an SOS quadratic form $f \in \mathbb{R}[x_1, x_2, x_3]_2$ is M-SOS if and only if a Gram matrix Q for f satisfies the condition that MQM^{-1} is diagonal.

Let

$$f(x) = (x_1 + x_2 + x_3)^2 + (x_2 - x_3)^2 = x_1^2 + 2x_2^2 + 2x_3^2 + 2x_1x_2 + 2x_1x_3.$$

The first square $(x_1 + x_2 + x_3)^2$ is the square of the sum of elements lying in each copy of the trivial representation of S_2. The unique Gram matrix for f is $Q = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix}$. Now,

$$MQM^{-1} = \begin{bmatrix} 1 & \sqrt{2} & 0 \\ \sqrt{2} & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

is not diagonal. So, there is no tube $\mathbf{q} \in \mathbb{R}^{1\times 1\times 3}$ such that $Q = M^{-1} \operatorname{diag}(M \operatorname{vec}(\mathbf{q}))M$ and therefore f cannot be M-SOS. $\diamond$

Chapter 5

Topological Approach to Aggregations of Quadratic Inequalities

The content of this chapter is based on joint work with Greg Blekherman and appears in [BD25]. This chapter presents a condensed version of the results with fewer examples than the journal version. Additionally, the introductory sections of this chapter have been changed to better position this work in the context of the dissertation.

In this chapter, we study the properties of semialgebraic sets defined by three quadratic inequalities from the point of view of algebraic topology. In doing so, we address open problems in real algebraic geometry as well as optimization from a unified viewpoint.

In real algebraic geometry, one is frequently interested in certificates of emptiness for real varieties in terms of sums of squares of polynomials. Consider a variety defined by a complete intersection of three quadrics Q_1, Q_2, Q_3. One certificate that $\mathcal{V}_{\mathbb{R}}(Q_1, Q_2, Q_3) \subseteq \mathbb{RP}^n$ is empty is a positive definite linear combination of the Q_i :

$$\sum_{i=1}^{3} \lambda_i Q_i = A \succ 0.$$

In this case, we have that the quadratic form $f(x) = x^\top A x$ is strictly positive (and therefore nonvanishing) on $\mathbb{RP}^n$ but $f \in \langle Q_1, Q_2, Q_3 \rangle$. Note that (up to relabeling the Q_i) this gives a

sum of squares certificate that Q_3 is positive on the variety $\mathcal{V}_{\mathbb{R}}(Q_1, Q_2)$ by rearranging. That is

$$Q_3 = f + \langle Q_1, Q_2 \rangle.$$

However, because the variety $\mathcal{V}_{\mathbb{R}}(Q_1, Q_2)$ has degree 4 and codimension 2, it follows from Theorem 2.3.8 that there are choices for Q_3 such that $\mathcal{V}_{\mathbb{R}}(Q_1, Q_2, Q_3) = \varnothing$ but Q_3 is not a sum of squares mod $\langle Q_1, Q_2 \rangle$. So, a natural question is "Under what conditions is $\mathcal{V}_{\mathbb{R}}(Q_1, Q_2, Q_3) = \varnothing$?"

The other problem comes from optimization. The optimization of a linear functional over a set S defined by quadratic inequalities is a computationally difficult problem. However, one can equivalently optimize the functional over the convex hull of the set S. For this approach to be tractable, one needs an efficient description of $\text{conv}(S)$. One approach to finding an efficient description is by *aggregations* (nonnegative linear combinations) of the defining inequalities [Yil09, DMnS22, BDS24]. In the two quadratics case, $\text{conv}(S)$ can always be represented by aggregations [Yil09]. In the three quadratics case, one is guaranteed such a representation if there is a positive definite linear combination of the defining quadratics. However, this condition is not necessary.

We address both problems by studying the algebraic topology of the associated *spectral curve*–the real algebraic plane curve cut out by $\det\left(\sum_{i=1}^{3} \lambda_i Q_i\right)$. Specifically, spectral sequences derived in [AL12] relate the homology of semialgebraic sets defined by quadratic inequalities to the cohomology of linear combinations of the defining quadratics with specified index and coefficients in a specified polyhedral cone. These sets are precisely the intersections of components of the complement of the spectral curve with a polyhedral cone.

For both problems, we show that the positive definite linear combination (PDLC) condition relaxes to a condition on the hyperbolicity of the spectral curve. This is a strict generalization of the PDLC condition, as the spectral curve is hyperbolic when there is a positive definite linear combination, as in Example 2.3.1. Moreover, these results highlight

the interplay of algebra, convex geometry, and optimization prevalent in this dissertation.

To make precise statements of the main results of this chapter, we will first fix notation.

Notation In this chapter, we will fix the following notation. We fix three symmetric $(n+1) \times (n+1)$ matrices Q_1, Q_2, Q_3. To these three matrices, we associate the following:

- The functions $f_i(x) = \begin{bmatrix} x^\top & 1 \end{bmatrix} Q_i \begin{bmatrix} x^\top & 1 \end{bmatrix}^\top$ for $i \in [3]$.
- The homogenizations $f^h(x, x_{n+1}) = \begin{bmatrix} x^\top & x_{n+1} \end{bmatrix} Q_i \begin{bmatrix} x^\top & x_{n+1} \end{bmatrix}^\top$ for $i \in [3]$.
- The semialgebraic set

$$S = \{x \in \mathbb{R}^n \mid f_i(x) \leqslant 0 \text{ for all } i \in [3]\}.$$

- The homogenized version of S:

$$S^h = \left\{(x, x_{n+1}) \in \mathbb{R}^n \times \mathbb{R} \mid f_i^h(x) \leqslant 0 \text{ for all } i \in [3]\right\}.$$

Given an element $\lambda \in \mathbb{R}^3$

- We consider linear combinations of the matrices $Q_\lambda = \sum_{i=1}^3 \lambda_i Q_i$, quadratics $f_\lambda = \sum_{i=1}^3 \lambda_i f_i$, and homogeneous quadratics $f_\lambda^h = \sum_{i=1}^3 \lambda_i f_i^h$.
- We consider an associated (affine) semialgebraic set defined by f_λ

$$S_\lambda = \{x \in \mathbb{R}^n \mid f_\lambda(x) \leqslant 0\}.$$

- The homogenization of S_λ:

$$S_\lambda^h = \{(x, x_{n+1}) \in \mathbb{R}^n \times \mathbb{R} \mid f_\lambda^h(x, x_{n+1}) \leqslant 0\}.$$

With this setup, we define the *spectral curve* to be the curve in $\mathbb{RP}^2$ cut out by the polynomial

$$g(\lambda) = \det\left(\lambda_1 Q_1 + \lambda_2 Q_2 + \lambda_3 Q_3\right).$$

Finally, we will need notation for the restriction of many of these objects to a hyperplane $H \subseteq \mathbb{R}^{n+1}$. We denote by $Q_\lambda|_H$ the restriction of the matrix Q_λ to the n-dimensional space H. Note that the signature and singularity of this restriction is well-defined.

Statement of Main Results and Relationship to Prior Work We first address the real algebraic geometry problem of finding a certificate of emptiness for the real variety $\mathcal{V}_\mathbb{R}(f_1^h, f_2^h, f_3^h) \subseteq \mathbb{RP}^n$. Under the assumption that the spectral curve is smooth, our theorem relates the emptiness of $\mathcal{V}_\mathbb{R}(f_1^h, f_2^h, f_3^h)$ to the hyperbolicity of the spectral curve and a condition on eigenvalues of matrices in $\text{span}_\mathbb{R}\{Q_1, Q_2, Q_3\}$.

Theorem 5.0.1. *Suppose that the spectral curve is smooth and that $n \neq 2$. Then, $\mathcal{V}_\mathbb{R}(f_1^h, f_2^h, f_3^h)$ is empty if and only if one of the following holds:*

1. *The spectral curve is hyperbolic and there is $\mu \in \mathbb{R}^3$ such that Q_μ has n positive eigenvalues*

2. *$n = 3$, and the spectral curve has no real points.*

In the case $n = 2$, the spectral curve is not smooth if the projective variety $\mathcal{V}_\mathbb{R}(f_1^h, f_2^h, f_3^h)$ is nonempty.

The result of Theorem 5.0.1 in the case $n = 3$ can be found in [PSV11, Theorem 7.8]. Additionally, the "only if" direction for $n \geqslant 4$ is in [Agr88]. Our statement of Theorem 5.0.1 unifies these existing results and provides a converse to the statement in [Agr88].

Additionally, Theorem 5.0.1 can be interpreted in light of the Helton-Vinnikov Theorem (Theorem 2.3.13 [HV07]), which states that any hyperbolic plane curve possesses a definite

determinantal representation. In this case, the variety defined by the associated quadratic forms is necessarily empty. Theorem 5.0.1, on the other hand, says that there is only one other possibility for a determinantal representation of a hyperbolic curve where the variety defined by the corresponding quadratic forms is empty–there must be a linear combination of the defining matrices which obtains n positive eigenvalues and therefore there must be a hyperbolicity cone $\mathcal{P}$ of g whose interior consists of $\mu \in \mathbb{R}^3$ such that Q_μ has $n-1$ positive and two negative eigenvalues.

As an extension of Theorem 5.0.1, we study the emptiness of solution sets of systems defined by three quadratic *inequalities* under the assumption that the variety $\mathcal{V}_{\mathbb{R}}(f_1^h, f_2^h, f_3^h)$ is empty.

Proposition 5.0.2. *Suppose that the spectral curve is smooth and nonempty and that the variety $\mathcal{V}_{\mathbb{R}}(f_1^h, f_2^h, f_3^h)$ is empty. Then, for a nonzero polyhedral cone $K \subseteq \mathbb{R}^3$, we have that the set*

$$\{X(K) = \{[x] \in \mathbb{RP}^n \mid f_i^h([x]) \in K \text{ for all } i \in [3]\}\}$$

is empty if and only if either

1. *there is $\mu \in K^\circ$ such that $Q_\mu \succ 0$, or*
2. *the set $\Omega^n = \{\lambda \in K^\circ \cap \mathbb{S}^2 \mid Q_\lambda$ has at least n positive eigenvalues$\}$ has $\dim_{\mathbb{Z}_2} H^1(\Omega^n) = 1$.*

While the statement and proof of 5.0.2 requires the language of cohomology, we note that the condition can be checked via convex geometry. Specifically, by Theorem 5.0.1, we have that when the variety $\mathcal{V}_{\mathbb{R}}(f_1^h, f_2^h, f_3^h)$ is empty, the spectral curve must be hyperbolic with hyperbolicity cone $\mathcal{P}$ whose interior contains either positive definite combinations or matrices with exactly two negative eigenvalues. In the PDLC case, the statement of Proposition 5.0.2 asserts that there is a nonempty intersection of the cones K° and $\mathcal{P}$. In the non-PDLC case, the condition on the first cohomology group of Ω^n is equivalent to $\mathcal{P} \subset K^\circ$.

We use these tools to study aggregations of quadratic inequalities, and in particular, the existence of representations

$$\overline{\text{conv}}(S) = \bigcap_{\lambda \in \Lambda} S_\lambda. \tag{5.1}$$

To relate aggregations to convexity, we need to restrict our attention to subsets of aggregations. We say that an aggregation $\lambda \in \mathbb{R}^3_+$ is *permissible* if Q_λ has at most one negative eigenvalue. Note that if λ is permissible, then $\text{int}(S_\lambda)$ is either empty, convex, or the disjoint union of two convex sets. We further stratify permissible aggregations into *good aggregations*–λ such that $\text{conv}(S) \subseteq S_\lambda$ and *bad aggregations*– permissible aggregations which are not good. Using this language, we are able to develop a sufficient condition for a representation of the form (5.1).

Theorem 5.0.3. *Assume that* $\text{int}(S) \neq \varnothing$ *and that* S *has* no points at infinity*, i.e.*

$$\left\{(x,0) \in \mathbb{R}^n \times \mathbb{R} \mid f_i^h(x,0) \leqslant 0 \text{ for all } i \in [3]\right\} = \varnothing.$$

Suppose further that the variety $\mathcal{V}_{\mathbb{R}}(f_1^h, f_2^h, f_3^h) = \varnothing$ *and the spectral curve is smooth and nonempty. Then,*

1. *If* $n \geqslant 3$*, the spectral curve* g *is hyperbolic. Let* $\mathcal{P} \subseteq \mathbb{R}^3$ *be the hyperbolicity cone of* g *such that* $\text{int}(\mathcal{P})$ *consists of either positive definite matrices or matrices with exactly two negative eigenvalues. If no nonzero aggregation lies in* $\mathcal{P}$*, then there are* $k \leqslant 6$ *good aggregations* $\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)} \in \mathbb{R}^3_+$ *such that*

$$\overline{\text{conv}}(S) = \bigcap_{i=1}^{k} S_{\lambda^{(i)}}.$$

2. *If* $n = 2$ *and* g *is hyperbolic, then* (i) *still applies. If* g *is not hyperbolic, then there is a (possibly infinite) subset* $\Lambda_1 \subseteq \mathbb{R}^3_+$ *of good aggregations such that*

$$\overline{\text{conv}}(S) = \bigcap_{\lambda \in \Lambda_1} S_\lambda.$$

The condition in Theorem 5.0.3 is related to the main results of [DMnS22, BDS24]. In these settings, it was determined that PDLC is sufficient for a description via aggregations of the convex hull of sets defined three *strict* inequalities. The results in [DMnS22, BDS24] were derived using hidden convexity properties of quadratic maps. However, using our topological approach, we are able to provide a sufficient condition which does not rely on PDLC and deal with the technically more subtle case of closed inequalities.

We further study the set of permissible aggregations, providing finiteness results.

Theorem 5.0.4. *Suppose that $\mathcal{V}_{\mathbb{R}}(f_1^h, f_2^h, f_3^h)$ is empty and that the spectral curve is smooth and hyperbolic. Then, there is a finite subset Λ_1 of permissible aggregations such that*

$$\bigcap_{\lambda \text{ permissible}} S_\lambda = \bigcap_{\lambda \in \Lambda_1} S_\lambda.$$

Finally, when restricting to the PDLC case, we are able to improve the bound on the number of necessary aggregations:

Theorem 5.0.5. *If Q_1, Q_2, Q_3 satisfy PDLC, $S = \text{cl}(\text{int}(S))$, and $\text{int}(S) \neq \varnothing$, then there is a subset $\{\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(r)}\}$ of good aggregations with $r \leqslant 4$ such that*

$$\overline{\text{conv}}(S) = \bigcap_{i=1}^{r} S_{\lambda^{(i)}}.$$

Theorem 5.0.5 is related to [BDS24, Conjecture 3.2], where the authors conjecture the existence of a set T defined by the intersection of three strict quadratic inequalities satisfying PDLC such that $\text{conv}(T)$ cannot be described using fewer than six good aggregations. Theorem 5.0.5 says that if the set defined by the non-strict inequalities is sufficiently regular, then four aggregations is sufficient.

5.1 Homogeneous Quadratic Maps

In this section we briefly review the necessary background on homogeneous quadratic maps and establish notation. The main goal of this section is to state [AL12, Theorem A], which provides a spectral sequence relating the topology of the solution set of a system of quadratic inequalities with the topology of sets of linear combinations of the defining matrices. This theorem can be thought of as a topological duality theorem for quadratics.

A *homogeneous quadratic map* $p : \mathbb{R}^{n+1} \to \mathbb{R}^m$ is a map

$$p(x) = (p_1(x), p_2(x), \ldots, p_m(x)),$$

where each p_i is a quadratic form on $\mathbb{R}^{n+1}$. Note that because each p_i is a quadratic form, we have that $p(\lambda x) = \lambda^2 p(x)$ for any $\lambda \in \mathbb{R}$. In particular, for a point $[x] \in \mathbb{RP}^n$, the evaluation $p([x])$ is well-defined up to nonnegative scaling. So, for $K \subseteq \mathbb{R}^m$ a polyhedral cone and $p : \mathbb{R}^{n+1} \to \mathbb{R}^m$ a homogeneous quadratic map, we set

$$X(K, p) = \{[x] \in \mathbb{RP}^n \mid p([x]) \in K\}.$$

When the polyhedral cone K and the homogeneous quadratic map p are clear from context, we drop them from the notation and simply write $X = X(K, p)$. As an example, if we set $f^h = (f_1^h, f_2^h, f_3^h)$ and take the polyhedral cone $K = -\mathbb{R}^3_+$, then

$$X(-\mathbb{R}^3_+, f^h) = \left\{[(x, x_{n+1})] \in \mathbb{RP}^n \mid (x, x_{n+1}) \in S^h \backslash \{0\}\right\}.$$

That is, $X(-\mathbb{R}^3_+, f)$ is the image of S^h in projective space. The sets $X(K, p)$ will play the role of "primal" objects in the results of this chapter. We also define dual objects coming from linear combinations of the defining quadratic forms with coefficients in a specified cone. Specifically, if $p = (p_1, p_2, \ldots, p_m)$ is a homogeneous quadratic map and P_i is the symmetric matrix representative of p_i, then we set

$$\Omega^j(K,p) = \left\{ \lambda \in K^\circ \cap \mathbb{S}^{m-1} \;\middle|\; \sum_{i=1}^m \lambda_i P_i \text{ has at least } j \text{ positive eigenvalues} \right\}.$$

We also define a relative version of $\Omega^j(K,p)$ for the restriction of the quadratic map p to a hyperplane $H \subseteq \mathbb{R}^{n+1}$. Specifically,

$$\Omega^j_H(K,p) = \left\{ \lambda \in K^\circ \cap \mathbb{S}^{m-1} \;\middle|\; \sum_{i=1}^m \lambda_i P_i|_H \text{ has at least } j \text{ positive eigenvalues} \right\}.$$

We will frequently work with the setting where the components of a homogeneous quadratic map are defined by some $\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(m)} \in \mathbb{R}^3$. In this case, if $A \in \mathbb{R}^{m\times 3}$ is the matrix whose rows are $(\lambda^{(j)})^\top$, then we set

$$f^h_A = (f^h_{\lambda^{(1)}}, f^h_{\lambda^{(2)}}, \ldots, f^h_{\lambda^{(m)}}).$$

Since $f^h_A(x) = Af^h(x)$ (i.e. f^h_A factors through $\mathbb{R}^3$), we can change cones to pull back the sets Ω^j to $\mathbb{R}^3$ and only work with the homogeneous quadratic map f^h. Specifically, we see that for a fixed polyhedral cone $K \subseteq \mathbb{R}^m$, we have

$$\begin{aligned} X(K, f^h_A) &= \{[x] \in \mathbb{RP}^n \mid f^h_A([x]) \in K\} \\ &= \{[x] \in \mathbb{RP}^n \mid f^h([x]) \in A^{-1}(K)\} \\ &= X(A^{-1}(K), f^h), \end{aligned}$$

where $A^{-1}(K)$ is the preimage of K under the map $A : \mathbb{R}^3 \to \mathbb{R}^m$. Moreover, we compute that

$$\begin{aligned}(A^\top K^\circ)^\circ &= \{v \in \mathbb{R}^3 \mid \langle A^\top y, v\rangle \leqslant 0 \text{ for all } y \in K^\circ\}\\ &= \{v \in \mathbb{R}^3 \mid \langle y, Av\rangle \leqslant 0 \text{ for all } y \in K^\circ\}\\ &= \{v \in \mathbb{R}^3 \mid Av \in (K^\circ)^\circ\}\\ &= A^{-1}(K).\end{aligned}$$

So, $A^{-1}(K)$ has polar dual $A^\top K^\circ \subseteq \mathbb{R}^3$, which we will leverage for convenient computations of the sets Ω^j. Specifically, we will frequently encounter the case where $K = -\mathbb{R}^m_+$, in which case we see that

$$\Omega^j(A^{-1}(-\mathbb{R}^m_+), f^h) = \{\lambda \in \text{cone}(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^m) \cap \mathbb{S}^2 \mid Q_\lambda \text{ has at least } j \text{ positive eigenvalues.}\}.$$

We can now state the spectral sequence of [AL12, Theorem A], which will provide the basis for many of our computations in this chapter.

Theorem 5.1.1 ([AL12])**.** *Let $p : \mathbb{R}^{n+1} \to \mathbb{R}^m$ be a homogeneous quadratic map and let $K \subseteq \mathbb{R}^m$ be a polyhedral cone. Then, there is a first quadrant, cohomology spectral sequence (E_r, d_r) converging to $H_{n-*}(X(K,p))$ such that $E_2^{i,j} = H^i(K^\circ \cap B^m, \Omega^j)$. Moreover, there is an explicit formula for the differential d_2.*

It follows from the long exact sequence in cohomology of the pair $(K^\circ \cap B^m, \Omega^j)$ that

$$E_2^{ij} \cong \begin{cases} H^{i-1}(\Omega^{j+1}), & i \geqslant 2,\ \Omega^{j+1} \neq \varnothing \\ H^0(\Omega^{j+1})/\mathbb{Z}_2, & i = 1,\ \Omega^{j+1} \neq \varnothing \\ \mathbb{Z}_2, & i = 0,\ \Omega^{j+1} = \varnothing \\ 0, & \text{otherwise} \end{cases}. \tag{5.2}$$

5.2 Certificates of Emptiness for Systems of Quadratics

In this section we use Theorem 5.1.1 to prove Theorem 5.0.1 and Proposition 5.0.2. We begin with a preparatory lemma about determinantal curves.

Lemma 5.2.1. *If $\lambda \in \Omega^n(\{0\})$ (i.e. Q_λ has n positive eigenvalues) and the spectral curve is smooth, then $[\lambda] \in \mathbb{RP}^2$ lies in the interior of an oval of the spectral curve of depth at least $\lfloor \frac{n+1}{2} \rfloor - 1$.*

Proof. Let $x \in \mathbb{S}^2$ be such that $[x] \in \mathbb{RP}^2$ is on the exterior of every oval of the spectral curve. By [Vin93], it follows that if $n+1$ is even, then Q_x has $\frac{n+1}{2}$ positive and negative eigenvalues, and if $n+1$ is odd, then one of Q_x or $-Q_x$ has $\lfloor \frac{n+1}{2} \rfloor + 1$ positive eigenvalues and $\lfloor \frac{n+1}{2} \rfloor$ negative eigenvalues. Since g is smooth, it follows that if $[\lambda]$ is on the interior of an oval of depth k but on the exterior of all ovals of depth at least $k+1$, then Q_λ has at most $\lfloor \frac{n+1}{2} \rfloor + k$ positive eigenvalues for $n+1$ even and $\lfloor \frac{n+1}{2} + k + 1$ positive eigenvalues for $n+1$ odd. In either case, if Q_λ has at least n positive eigenvalues, it therefore follows that λ is in the interior of an oval of depth at least $\lfloor \frac{n}{2} \rfloor - 1$. □

We will start with the case where $X = X(\{0\}, f^h) = \mathcal{V}_\mathbb{R}(f_1^h, f_2^h, f_3^h)$. The hyperbolicity of the spectral curve is an immediate computation using Lemma 5.2.1 and the spectral sequence of Theorem 5.1.1.

Proposition 5.2.2 (cf [Agr88])**.** *Suppose that $n \geqslant 4$ and that the spectral curve is smooth and nonempty. If $\mathcal{V}_\mathbb{R}(f_1^h, f_2^h, f_3^h)$ is empty, then the spectral curve is hyperbolic and $\Omega^n(\{0\})$ is nonempty, i.e., there is $\mu \in \mathbb{R}^3$ such that Q_μ has n positive eigenvalues.*

Proof. Set $X = \mathcal{V}_\mathbb{R}(f_1^h, f_2^h, f_3^h)$ and $\Omega^j = \Omega^j(\{0\}, f^h)$ for notational convenience. If PDLC holds then we are done.

We now show that $\Omega^n \neq \varnothing$. Suppose for the sake of a contradiction that $\Omega^n = \varnothing$ and let (E_r, d_r) be the spectral sequence of Theorem 5.1.1. By the isomorphisms (5.2), we have that

$E_2^{0,n} \cong \mathbb{Z}_2$. On the other hand, we have that $H_0(X) = 0$ since X is empty. Therefore, since $H^i(\Omega^j) = 0$ for all $i \geqslant 2$ for all j, it must be the case that at least one of $d_2^{0,n} : \mathbb{Z}_2 \to H^1(\Omega^n)$ or $d_3^{0,n} : \mathbb{Z}_2 \to H^2(\Omega^{n-1})$ is nonzero. However, since $\Omega^n = \varnothing$, it follows that $H^1(\Omega^n) = 0$. Similarly, since the spectral curve is nonempty, $H^2(\Omega^{n-1}) = 0$. So, both differentials must be zero, the desired contradiction.

The remaining case is that $\Omega^n \neq \varnothing$ and PDLC does not hold. We want to show that in this case, the spectral curve is hyperbolic. As before, let (E_r, d_r) be the spectral sequence of Theorem 5.1.1. We will use the fact that $H_0(X) = 0$ to compute the cohomology groups $H^i(\Omega^j)$. Note that since $\Omega^n \neq \varnothing$, it follows that there is Q_μ with at most one positive eigenvalue and therefore $\Omega^{j+1} \subseteq \mathbb{S}^2$ is a proper subset for each $j \geqslant 1$. In particular, this implies that $E_2^{3,j} = H^2(\Omega^{j+1}) = 0$ for all $j \geqslant 1$. Since $n \geqslant 4$, this implies that $E_2^{3,j} = 0$ for $j \geqslant n-3$. So, the E_2 page has the following form:

E_2

n	$\mathbb{Z}_2$	0	0	0
$n-1$	0	$H^0(\Omega^n)/\mathbb{Z}_2$	$H^1(\Omega^n)$	0
$n-2$	0	$H^0(\Omega^{n-1})/\mathbb{Z}_2$	$H^1(\Omega^{n-1})$	0
$n-3$	0	$H^0(\Omega^{n-2})/\mathbb{Z}_2$	$H^1(\Omega^{n-2})$	0
	0	1	2	3

Since $0 = H_0(X) = \bigoplus_{i+j=n} E_\infty^{i,j}$, it follows that $d_r^{0,n}$ is injective for some $r \geqslant 2$. Because $E^{i,j} = 0$ for all j when $i \geqslant 4$, we in fact have that $d_2^{0,n}$ must be injective. So $H^1(\Omega^n) \neq 0$.

Since $H^1(\Omega^n) \neq 0$, and since $\lambda \in \Omega^n$ implies that $[\lambda]$ lies in the interior of an oval of the spectral curve of depth at least $\lfloor \frac{n+1}{2} \rfloor - 1$, the spectral curve must have a nest of ovals of depth $\lfloor \frac{n+1}{2} \rfloor$. Therefore the spectral curve is hyperbolic. □

The converse is more difficult. In particular, we need to compute the differential d_2 of

the spectral sequence. In [AL12], the authors relate the value of the differential d_2 to sets of matrices with repeated eigenvalues. To make the precise statement, we set $\mathcal{Q}$ to be the vector space of quadratic forms on $\mathbb{R}^{n+1}$. For $q \in \mathcal{Q}$, we set $\rho_1(q) \geqslant \rho_2(q) \geqslant \ldots \geqslant \rho_{n+1}(q)$ to be the eigenvalues of q and $\mathcal{D}_j = \{q \in \mathcal{Q} \mid \rho_j(q) > \rho_{j+1}(q)\}$. With this notation, we have the following theorem.

Theorem 5.2.3 ([AL12, Theorem B])**.** *There is a formula for the differential d_2 in terms of matrices with repeated eigenvalues. Explicitly,*

$$d_2(x) = (x \smile \bar{f}^* \gamma_{1,j})|_{(K^\circ \cap B^3, \Omega^j(K))} \text{ for } x \in H^*(K^\circ \cap B^3, \Omega^{j+1}).$$

Here, $\bar{f} : \mathbb{R}^3 \to \mathcal{Q}$ is given by $\bar{f}(\lambda) = f^h_\lambda$ and $\bar{f}^*$ is the induced map on cohomology. The value of the class $\gamma_{1,j} \in H^2(\mathcal{Q}, \mathcal{D}_j)$ on the image of $\sigma : B^2 \to \mathcal{Q}$ with $\sigma(\partial B^2) \subseteq \mathcal{D}_j$ is equal to the intersection number of $\sigma(B^2)$ and $\mathcal{Q} \backslash \mathcal{D}_j \mod 2$.

We will be interested in applying Theorem 5.2.3 to compute $d_2^{0,n}(1)$ in the case where g is hyperbolic with a hyperbolicity cone $\mathcal{P}$ such that the interior of $\mathcal{P}$ consists of λ such that Q_λ has exactly $n-1$ positive and two negative eigenvalues. In particular, we will want to show that there is exactly one $\lambda \in \text{int}(\mathcal{P}) \cap \mathbb{S}^2$ such that Q_λ has a repeated negative eigenvalue. In this case, with notation as in Theorem 5.2.3, we have that $\gamma_{1,n} \pitchfork 0$ so that $d_2^{0,n}(1) \neq 0$.

Lemma 5.2.4. *Suppose that $\Omega^n(\{0\})$ is nonempty and that g is smooth and hyperbolic with hyperbolicity cone $\mathcal{P}$ which contains matrices with $n-1$ positive and two negative eigenvalues. Then, if $\lambda^{(1)}, \lambda^{(2)} \in \text{int}(\mathcal{P})$ are such that $Q_{\lambda^{(1)}}$ and $Q_{\lambda^{(2)}}$ each have repeated eigenvalue -1, then for each $t \in \mathbb{R}$, we have that $tQ_{\lambda^{(1)}} + (1-t)Q_{\lambda^{(2)}}$ has at least two negative eigenvalues.*

Proof. Note that the statement holds for all $t \in [0,1]$ by the convexity of $\mathcal{P}$. Let V be the two dimensional subspace of $\mathbb{R}^{n+1}$ corresponding to the eigenvalue -1 of $Q_{\lambda^{(1)}}$ and let U be the two dimensional subspace of $\mathbb{R}^{n+1}$ corresponding to the eigenvalue -1 of $Q_{\lambda^{(2)}}$.

Then, if $t \geqslant 1$ and $v \in V$ has $v^\top v = 1$,

$$v^\top(tQ_{\lambda^{(1)}} + (1-t)Q_{\lambda^{(2)}})v = -t + (1-t)v^\top Q_{\lambda^{(2)}}v \leqslant -1.$$

Similarly, if $u \in U$ has $u^\top u = 1$ and $t \leqslant 0$, then

$$u^\top(tQ_{\lambda^1} + (1-t)Q_{\lambda^{(2)}})u = t(u^\top Q_{\lambda^{(1)}}u + 1) - 1 \leqslant -1.$$

So, the claim follows from the variational characterization of eigenvalues. □

Using Lemma 5.2.4, we show that there is exactly one $\lambda \in \text{int}(\mathcal{P}) \cap \mathbb{S}^2$ such that Q_λ has a repeated negative eigenvalue. This will allow us to compute $d_2^{0,n}(1) \neq 0$.

Lemma 5.2.5. *Suppose that $\Omega^n(\{0\})$ is nonempty and that g is smooth and hyperbolic with hyperbolicity cone $\mathcal{P}$ which contains matrices with $n-1$ positive and two negative eigenvalues. Then, there is a unique $\lambda \in \mathcal{P} \cap \mathbb{S}^2$ such that Q_λ has a repeated negative eigenvalue.*

Proof. First, there can be at most one such λ. Suppose for the sake of a contradiction that $\lambda^{(1)}, \lambda^{(2)} \in \mathcal{P} \cap \mathbb{S}^2$ had repeated negative eigenvalue. Then, for some rescaling $\hat{\lambda^{(1)}}$ and $\hat{\lambda^{(2)}}$ of $\lambda^{(1)}$ and $\lambda^{(2)}$, respectively, we would have $Q_{\hat{\lambda^{(1)}}}$ and $Q_{\hat{\lambda^{(2)}}}$ with repeated negative eigenvalue -1. By Lemma 5.2.4, we have that for all $t \in \mathbb{R}$, the matrix $Q_{t\hat{\lambda^{(1)}}+(1-t)\hat{\lambda^{(2)}}}$ must have at least two negative eigenvalues. Since g is hyperbolic with respect to both $\lambda^{(1)}$ and $\lambda^{(2)}$, it must be the case that the univariate polynomial $\det(Q_{\hat{\lambda^{(2)}}} + t(Q_{\hat{\lambda^{(1)}}-\hat{\lambda^{(2)}}}))$ has all zeros at infinity, but this contradicts the smoothness of g.

We now show that there is at least one $\lambda \in \mathcal{P} \cap \mathbb{S}^2$ such that Q_λ has a repeated negative eigenvalue. Let $\lambda \in \text{int}(\mathcal{P}) \cap \mathbb{S}^2$ and let $v^{(1)}, v^{(2)}, \ldots, v^{(n+1)}$ be an orthonormal basis of eigenvectors of Q_λ corresponding to the eigenvalues $\rho_1(Q_\lambda) \geqslant \rho_2(Q_\lambda) \geqslant \ldots \geqslant \rho_{n+1}(Q_\lambda)$. Set

$$B = \begin{bmatrix} \frac{1}{\sqrt{|\rho_1(Q_\lambda)|}}v^{(1)} & \frac{1}{\sqrt{|\rho_2(Q_\lambda)|}}v^{(2)} & \frac{1}{\sqrt{|\rho_{n+1}(Q_\lambda)|}}v^{(n+1)} \end{bmatrix}.$$

Then, $x \mapsto Bx$ defines a real change of coordinates on $\mathbb{RP}^n$. In these coordinates, the quadratic form f_λ^h is represented by the diagonal matrix

$$B^\top Q_\lambda B = \operatorname{diag}(1, 1, \ldots, 1, -1, -1).$$

By the above discussion, λ is the unique element of $\mathcal{P} \cap \mathbb{S}^2$ such that $B^\top Q_\lambda B$ has a repeated negative eigenvalue. Using Theorem 5.2.3, we therefore obtain that

$$\mathcal{V}_{\mathbb{R}}(B^\top Q_1 B, B^\top Q_2 B, B^\top Q_3 B) = \varnothing.$$

Since nonexistence of real points on a variety is preserved under a real change of coordinates on $\mathbb{RP}^n$, this implies that $\mathcal{V}_{\mathbb{R}}(f_1^h, f_2^h, f_3^h)$ is empty. By Theorem 5.2.3, this implies that the differential $d_2^{0,n}$ is nontrivial and therefore there is an odd number of $\lambda \in \mathcal{P} \cap \mathbb{S}^2$ such that Q_λ has a repeated negative eigenvalue. By the preceding discussion, this λ must be unique. □

We are now able to prove the $n \geqslant 4$ case of Theorem 5.0.1.

Proposition 5.2.6. *Suppose that $n \geqslant 4$ and that the spectral curve is smooth. If g is hyperbolic with hyperbolicity cone $\mathcal{P}$ such that* $\operatorname{int}(\mathcal{P})$ *has either positive definite matrices or matrices with exactly two negative eigenvalues, then $\mathcal{V}_{\mathbb{R}}(f_1^h, f_2^h, f_3^h) = \varnothing$.*

Proof. If $\mathcal{P}$ contains positive definite matrices, then we are done. Otherwise, $\operatorname{int}(\mathcal{P})$ contains matrices with $n-1$ positive and two negative eigenvalues. Set $X = X(\{0\}, f^h)$ and $\Omega^j = \Omega^j(\{0\})$. If (E_r, d_r) is the spectral sequence of Theorem 5.1.1, then the E_2 page has the form

E_2

n	$\mathbb{Z}_2$	0	0
$n-1$	0	0	$H^1(\Omega^n) \cong \mathbb{Z}_2$
$n-2$	0	0	0
$n-3$	0	0	0
	0	1	2

So, we have that

$$H_0(X) \cong \bigoplus_{i+j=n} E_\infty^{i,j} \cong \ker(d_2^{0,n}).$$

To show that $X = \varnothing$, it therefore suffices to show that $d_2^{0,n}$ is injective, i.e, $d_2^{0,n}(1) \neq 0$. Let $\sigma : B^2 \to \mathbb{S}^2$ be a representative of the nontrivial class in $H^2(\mathbb{S}^2, \Omega^n)$. Then, with notation as in the statement of Theorem 5.2.3, we have that $\gamma_{1,n}(\bar{f}(\sigma))$ is equal to the number (mod 2) of $\lambda \in \mathcal{P} \cap \mathbb{S}^2$ such that Q_λ has a repeated negative eigenvalue. By Lemma 5.2.5, there is a unique such λ. Therefore,

$$d_2^{0,n}(1)(\sigma) = 1 \smile \bar{f}^* \gamma_{1,n}(\sigma) = \gamma_{1,n}(\bar{f}(\sigma)) = 1.$$

So $d_2^{0,n}(1)$ is not the zero map and therefore $d_2^{0,n}$ is injective. So,

$$H_0(X) \cong \ker(d_2^{0,n}) \cong 0$$

and $X = \varnothing$. □

We are now prepared to prove the full statement of Theorem 5.0.1. The small n ($n \leqslant 3$) cases are treated separately.

Proof of Theorem 5.0.1 . We separate the proof by cases for the value of n.

$n \geqslant 4$: This is given by Propositions 5.2.2 and 5.2.6.

$n = 3$: See [PSV11, Theorem 7.8].

$n = 1$: If Q_1, Q_2, Q_3 are linearly independent, then they span the entire space of quadrics and therefore satisfy PDLC.

$n = 2$: Recall here that the statement of Theorem 5.0.1 asserts that the spectral curve is not smooth when $\mathcal{V}_{\mathbb{R}}(f_1^h, f_2^h, f_3^h)$ is nonempty.

Suppose that $v \in \mathbb{R}^3$ is a nonzero point with $[v] \in \mathcal{V}_{\mathbb{R}}(f_1^h, f_2^h, f_3^h)$. Then, we have that the vectors $Q_1 v, Q_2 v$, and $Q_3 v$ are all orthogonal to v and therefore $\dim(\operatorname{span}_{\mathbb{R}}(Q_1, Q_2, Q_3)) \leqslant 2$. Let $\alpha \in \mathbb{R}^3$ be nonzero such that $0 = \sum_{i=1}^3 \alpha_i Q_i v = Q_\alpha v$. If Q_α has rank 1, then $[\alpha]$ is necessarily a singular point of $\mathcal{V}_{\mathbb{R}}(g)$.

Suppose that α has rank 2. Let $\mathcal{D} \subseteq \operatorname{Sym}_3(\mathbb{R})$ be the hypersurface in the space of real symmetric 3×3 matrices defined by the vanishing of the determinant. Note that $Q_\alpha \in \mathcal{D}$ since $\operatorname{rk}(Q_\alpha) \neq 3$. By Jacobi's formula, we have that $\nabla \det(Q_\alpha) = \operatorname{adj} Q_\alpha = cvv^\top$ for some constant c. This implies that the tangent space to $\mathcal{D}$ at Q_α, $T_{Q_\alpha}\mathcal{D}$ is the space of quadratic forms which vanish at v since $\langle \operatorname{adj} Q_\alpha, Q \rangle = 0$ implies that $\operatorname{trace}(vv^\top Q) = v^\top Q v = 0$. But then $Q_i \in T_{Q_\alpha}\mathcal{D}$ for each $i \in [3]$ and $\operatorname{span}_{\mathbb{R}}(\{Q_i \mid i \in [3]\})$ intersects $\mathcal{D}$ non-transversely at Q_α. So, $[\alpha]$ is a singular point of the spectral curve. □

Example 5.2.1 ([PSV12, Example 5.2])**.** As stated in Example 2.3.2, the polynomial

$$p(x,y,z) = \det\left(\begin{bmatrix} 25x & 0 & 12y - 32x & -60z \\ 0 & 25x & 10z & 24x + 16y \\ 12y - 32x & 10z & 6x + 4y & 0 \\ -60z & 24x + 16y & 0 & 6x + 4y \end{bmatrix}\right)$$

is hyperbolic with respect to $(1,0,0)$ and there are no values of (x,y,z) which result in a positive definite matrix. We interpret this example in terms of Theorem 5.0.1. Set

$$Q_1 = \begin{bmatrix} 25 & 0 & -32 & 0 \\ 0 & 25 & 0 & 24 \\ -32 & 0 & 6 & 0 \\ 0 & 24 & 0 & 6 \end{bmatrix}, \quad Q_2 = \begin{bmatrix} 0 & 0 & 12 & 0 \\ 0 & 0 & 0 & 16 \\ 12 & 0 & 4 & 0 \\ 0 & 16 & 0 & 4 \end{bmatrix}, \text{ and } Q_3 = \begin{bmatrix} 0 & 0 & 0 & -60 \\ 0 & 0 & 10 & 0 \\ 0 & 10 & 0 & 0 \\ -60 & 0 & 0 & 0 \end{bmatrix}.$$

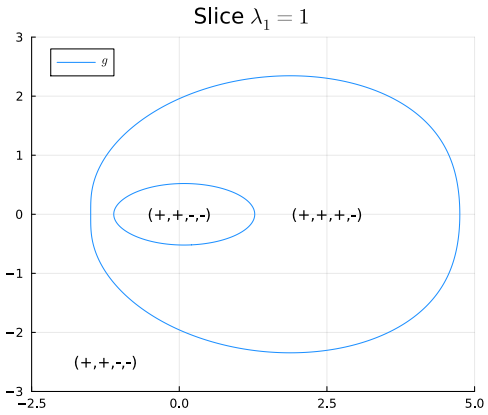


Figure 5.1: An affine slice of the spectral curve for Example 5.2.1

Note that we have that the spectral curve $g(\lambda) = g(\det(\sum_{i=1}^{3} \lambda_i Q_i)) = p(\lambda_1, \lambda_2, \lambda_3)$. Figure 5.1 shows the spectral curve in the affine slice $\lambda_1 = 1$ and the signature of matrices in each connected component of the complement of the curve. Here, for example, the label $(+, +, -, -)$ means that matrices in this region have two positive and two negative eigenvalues. As predicted by Theorem 5.0.1, we see that the spectral curve is hyperbolic and has a hyperbolicity cone whose interior consists of matrices with exactly two negative eigenvalues.

$\diamond$

We now turn our attention to systems of quadratic inequalities. This will be useful for the study of quadratically constrained optimization problems in the later sections of this chapter. In particular, we prove the dichotomy of Proposition 5.0.2: a set defined by quadratic inequalities is empty if and only if combinations of the defining inequalities achieve a positive definite matrix or the set with n positive eigenvalues has nontrivial first cohomology.

Proof of Proposition 5.0.2. For notational convenience, we set $X = X(K, f^h)$ and $\Omega^j = \Omega^j(K, f^h)$. Suppose first that $X = \varnothing$ and $\Omega^{n+1} = \varnothing$. We want to show that $H^1(\Omega^n)$ is nontrivial. As in the proof of Proposition 5.2.2, it must be the case that Ω^n is nonempty since otherwise $\dim_{\mathbb{Z}_2}(H_0(X)) \geqslant 1$. So, if (E_r, d_r) is the spectral sequence of Theorem 5.1.1, the E_2 page must have the following form:

E_2

n	$\mathbb{Z}_2$	0	0
$n-1$	0	$H^0(\Omega^n)/\mathbb{Z}_2$	$H^1(\Omega^n)$
$n-2$	0	$H^0(\Omega^{n-1})/\mathbb{Z}_2$	$H^1(\Omega^{n-1})$
	0	1	2

Note that we have used the fact that $H^2(\Omega^j) = \varnothing$ for all j since K° is a proper subset of $\mathbb{R}^3$. Since $X = \varnothing$, it must be the case that $d_2^{0,n} : \mathbb{Z}_2 \to H^1(\Omega^n)$ is injective, as $E_\infty^{0,n} \cong \ker(d_2^{0,n})$ is a direct summand of $H_0(X) \cong 0$. So, $H^1(\Omega^n) \neq 0$.

We now show that the conditions on Ω^n and Ω^{n+1} are sufficient for the emptiness of X.

If $\mu \in K^\circ \cap \mathbb{S}^2$ has $Q_\mu \succ 0$, then $X = \varnothing$. Indeed, if $f^h(x) = (f_1^h(x), f_2^h(x), f_3^h(x)) \in K$, then $f_\mu^h(x) \leqslant 0$ since $\mu \in K^\circ$. On the other hand $f_\mu^h(x) > 0$ for all $[x] \in \mathbb{RP}^n$ since $Q_\mu \succ 0$.

Finally, if Ω^{n+1} is empty and Ω^n has nontrivial first cohomology, then it suffices to show that the differential $d_2^{0,n} : \mathbb{Z}_2 \to H^1(\Omega^n)$ is nonzero. By Theorem 5.2.3, this happens if there is a continuous map $\sigma : B^2 \to \operatorname{span}_{\mathbb{R}}\{Q_1, Q_2, Q_3\}$ such such that $\sigma(\partial B^2) \subseteq \Omega^n$ and $\operatorname{int}(\sigma(B^2))$ contains a unique matrix with repeated negative eigenvalue. Since the variety $\mathcal{V}_{\mathbb{R}}(f_1^h, f_2^h, f_3^h)$ is empty, we know that such a map σ exists, as in the proof of Proposition 5.2.6. So, the differential $d_2^{0,n} \neq 0$ and therefore $E_3^{0,n} \cong E_\infty^{0,n} \cong 0$. Moreover, we see that $E_2^{1,n-1} \cong E_2^{2,n-2} \cong 0$ from the hyperbolicity and smoothness of the spectral curve. So,

$$H_0(X) \cong E_\infty^{0,n} \oplus E_\infty^{1,n-1} \oplus E_\infty^{2,n-2} \cong 0$$

and X is therefore empty. □

The result of Proposition 5.0.2 is that when the real variety $\mathcal{V}_{\mathbb{R}}(f_1^h, f_2^h, f_3^h)$ is empty and the spectral curve is smooth and hyperbolic, the emptiness of a system of quadratic inequalities can be determined by convex geometry. Specifically, the system of inequalities

defined by a polyhedral cone K is empty if and only if K° intersects the cone of positive definite linear combinations of the Q_i or if the cone K° contains the hyperbolicity cone $\mathcal{P}$ of g which certifies the emptiness of the real variety.

5.3 Reduction to Finite Subsets of Aggregations

In this section, we use the topological techniques of Proposition 5.0.2 to determine upper bounds on the number of necessary aggregations in a description

$$\overline{\text{conv}}(S) = \bigcap_{\lambda \in \Lambda_1} S_\lambda.$$

First, we recall some known results about aggregations of quadratic inequalities before proving our upper bounds in the general empty variety setting and improved upper bounds in the PDLC setting.

5.3.1 Some Preliminaries on Aggregations

Recall that we have set

$$\Lambda = \{\lambda \in \mathbb{R}^3_+ \mid Q_\lambda \text{ has exactly one negative eigenvalue}\}.$$

An aggregation $\lambda \in \Lambda$ is a *good aggregation* if $\text{conv}(S) \subseteq S_\lambda$ and a *bad aggregation* otherwise.

Lemma 5.3.1 ([BDS24])**.** *If $\lambda \in \Lambda$, then $\text{int}(S_\lambda)$ is either a convex set or a union of two disjoint convex sets. Moreover, λ is a bad aggregation if and only if $\text{int}(S_\lambda)$ is a union of two disjoint convex sets and $\text{int}(S)$ has nonempty intersection with both components.*

In [BDS24], the authors study bounds on the number of necessary aggregations in the PDLC case. The strategy is to improve a given aggregation by translating in the direction of a positive semidefinite matrix. By repeatedly applying this strategy, the authors show that

the set defined by good aggregations can be defined by good aggregations with at most two nonzero entries.

Proposition 5.3.2 ([BDS24]). *Suppose that PDLC holds and set* $\Theta = \{\theta \in \mathbb{R}^3 \mid Q_\theta \succeq 0\}$ *and* $\Lambda_1 = \{\lambda \in \Lambda \mid \lambda$ *is a good aggregation and* $S_\lambda \neq \mathbb{R}^n\}$. *If* $\lambda \in \Lambda_1$ *and* $\theta \in \Theta$ *are such that* $\lambda' = \lambda + \theta \in \mathbb{R}^3_+ \backslash \{0\}$, *then* $\lambda' \in \Lambda$ *and* $S_{\lambda'} \subseteq S_\lambda$.

Proposition 5.3.3 ([BDS24]). *Suppose that PDLC holds. Set* Λ_1 *as in Proposition 5.3.2 and* $\Lambda_2 = \{\lambda \in \Lambda_1 \mid |\operatorname{supp}(\lambda)| \leqslant 2\}$. *Then,* $\bigcap_{\lambda\in\Lambda_1} S_\lambda = \bigcap_{\lambda\in\Lambda_2} S_\lambda$.

5.3.2 Upper Bounds when $\mathcal{V}_{\mathbb{R}}(f_1^h, f_2^h, f_3^h) = \varnothing$

To determine such bounds, we interpret the redundancy of a given aggregation in terms of a system of quadratic inequalities. The idea is to ensure that the zero set of an new aggregation does not intersect the set defined by the previous aggregations, resulting in a strictly smaller subset.

Proposition 5.3.4. *Let* $\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k+1)} \in \Lambda$. *Set* $f_k^h = (f^h_{\lambda^{(1)}}, f^h_{\lambda^{(2)}}, \ldots, f^h_{\lambda^{(k)}})$ *and* $f_k^h = (f^h_{\lambda^{(1)}}, f^h_{\lambda^{(2)}}, \ldots, f^h_{\lambda^{(k+1)}})$. *Suppose that* $\mathcal{V}_{\mathbb{R}}(f^h_{\lambda^{(k+1)}}) \cap X(-\mathbb{R}^{k+1}_+, f^h_{k+1}) = \varnothing$ *and that* $X(-\mathbb{R}^k_+, f^h_k)$ *and* $X(-\mathbb{R}^{k+1}_+, f^h_{k+1})$ *have the same number of connected components. Then,*

$$X(-\mathbb{R}^{k+1}, f^h_{k+1}) = X(-\mathbb{R}^k, f^h_k).$$

Proof. Note that $X(-\mathbb{R}^{k+1}_+, f^h_{k+1}) \subseteq X(-\mathbb{R}^k_+, f^h_k)$. For the reverse inclusion, note that if $f^h_{\lambda^{(k+1)}}(x) \neq 0$ for all $[x] \in X(-\mathbb{R}^{k+1}_+, f^h_{k+1})$, then $f^h_{\lambda^{(k+1)}}$ has constant nonzero sign on each connected component of $X(-\mathbb{R}^k, f^h_k)$. Since $X(-\mathbb{R}^k_+, f^h_k)$ and $X(-\mathbb{R}^{k+1}_+, f^h_{k+1})$ have the same number of connected components, this implies that $f_{\lambda^{(k+1)}}([x]) < 0$ for all $[x] \in X(-\mathbb{R}^k, f^h_k)$. Therefore $X(-\mathbb{R}^{k+1}, f^h_{k+1}) = X(-\mathbb{R}^k, f^h_k)$. □

Note that the statement of the conclusion of Proposition 5.3.4 implies that $\bigcap_{i=1}^k S_{\lambda^{(i)}} = \bigcap_{i=1}^{k+1} S_{\lambda^{(i)}}$. Indeed, a point in $x \in \left(\bigcap_{i=1}^k S_{\lambda^{(i)}}\right) \setminus \left(\bigcap_{i=1}^{k+1} S_{\lambda^{(i)}}\right)$ would yield a point $[(x,1)] \in X(-\mathbb{R}^k_+, f^h_k) \backslash X(-\mathbb{R}^{k+1}, f^h_{k+1})$.

The advantage of Proposition 5.3.4 is that we will be able to conclude that given a finite list of aggregations, some new aggregation is unnecssary by applying Proposition 5.0.2 with a cone of the form $-\mathbb{R}^k \times \{0\}$. When $-\mathbb{R}^k \times \{0\}$ is pulled back to $\mathbb{R}^3$, we will therefore need to study sets of the form $\Omega^j(\text{cone}(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)})^\circ)$.

As a starting point, taking inspiration from Proposition 5.3.3, we consider a subset Λ' of Λ which consists of aggregations which generate extreme rays of Λ and have at least one zero entry:

$$\Lambda' = \{\lambda \in \text{ex}(\Lambda) \mid |\text{supp}(\lambda)| \leqslant 2\}.$$

Note that $|\Lambda'| \leqslant 6$ since each facet of $\mathbb{R}^3_+$ can contain at most two extreme rays of Λ. We will work towards applying Proposition 5.3.4 in the case that $\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)}$ is an enumeration of Λ' and $\lambda^{(k+1)} \in \Lambda \backslash \text{cone}(\Lambda')$. So, we need to understand the relevant cone.

Lemma 5.3.5. *If $\lambda \in \Lambda'$, then either λ is a standard basis vector of $\mathbb{R}^3$ or $g(\lambda) = 0$.*

Proof. If λ is not a standard basis vector, then $\lambda = \lambda_i e_i + \lambda_j e_j$ for soe $i \neq j \in [3]$ and $\lambda_i, \lambda_j > 0$. Since $\lambda \in \Lambda'$, if $g(\lambda) \neq 0$, then Q_λ has exactly n positive and one negative eigenvalue. So, for any $\epsilon > 0$ sufficiently small, we have that $\lambda - \epsilon(e_i + e_j)$ and $\lambda + \epsilon(e_i + e_j)$ are both elements of $\mathbb{R}^3_+$. Moreover, $Q_{\lambda \pm \epsilon(e_i + e_j)}$ also has exactly n positive and one negative eigenvalue. But then, λ does not span an extreme ray of Λ, a contradiction. □

Lemma 5.3.6. *Supose that $\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)}$ is an enumeration of Λ' and that $\lambda^{(k+1)} \in \Lambda \backslash \text{cone}(\Lambda')$. Let $A \in \mathbb{R}^{(k+1)\times 3}$ have rows $(\lambda^{(i)})^\top$. If, for fixed $i \in [k]$, the set*

$$F = \{t\lambda^{(k+1)} + s\lambda^{(i)} \mid t \in \mathbb{R}, s \in \mathbb{R}_+\}$$

is a face of $A^\top(\mathbb{R}^k_+ \times \mathbb{R})$, then $g(\lambda^{(i)}) = 0$.

Proof. Note that if $\lambda^{(k+1)} \in \Lambda \backslash \text{cone}(\Lambda')$, then $\lambda^{(k+1)}$ has strictly positive entries. Let $v \in \mathbb{R}^3$ be a supporting vector to the face F so that $v^\top \lambda^{(k+1)} = v^\top \lambda^{(i)} = 0$ and $v^\top \lambda^{(j)} > 0$ for

$j \in [k]\backslash\{i\}$. Suppose for the sake of a contradiction that $g(\lambda^{(i)}) \neq 0$. By Lemma 5.3.5, this implies that $\lambda^{(i)}$ is a standard basis vector of $\mathbb{R}^3$, without loss of generality $\lambda^{(i)} = e_1$.

Since $\lambda^{(i)} = e_1$ and $g(\lambda^{(i)}) \neq 0$, this implies that $Q_{\lambda^{(i)}} = Q_1$ has exactly n positive and one negative eigenvalue. So, either $e_2 \in \Lambda'$ or there is $t \in [0, 1)$ such that $te_1 + (1-t)e_2 \in \Lambda'$. But then, $v^\top e_2 > 0$. Similarly, $v^\top e_3 > 0$. Since $\lambda^{(k+1)}$ has strictly positive entries, this implies that $v^\top \lambda^{(k+1)} > 0$, a contradiction with the construction of v. □

Lemma 5.3.6 says that the elements of Λ' involved with faces of the cone $A^\top(-\mathbb{R}^k \times \mathbb{R})$ lie on the spectral curve. In the setting where the spectral curve is smooth and hyperbolic, they will necessarily lie on the oval of depth $\lfloor \frac{n+1}{2} \rfloor$ or $\lfloor \frac{n+1}{2} \rfloor - 1$. We will be particularly interested in the latter case (since the innermost oval bounds a convex region.) Fix C to be the part of the affine cone over the oval $\lfloor \frac{n+1}{2} \rfloor - 1$ of the spectral curve such that $\lambda \in C$ implies that Q_λ has $n-1$ positive eigenvalues, one negative eigenvalue, and 0 as an eigenvalue of multiplicity one. Additionally, fix $\mathcal{P}$ to be the hyperbolicity cone of the spectral curve consisting of either positive definite matrices or matrices with exactly two negative eigenvalues.

Lemma 5.3.7. *Let $\lambda \in \Lambda \cap C$. Then, if $e \in \mathcal{P}$, there is exactly one root t^* of $g(te + (1-t)\lambda)$ for $t \in (0, 1]$ and $t^*e + (1-t^*)\lambda$ lies on $\partial\mathcal{P}$.*

Proof. First, we can reduce to the case where g is a definite representation of the spectral curve since any hyperbolic plane curve has a definite representation and the number of intersection points of the spectral curve and a fixed line $L = \{te + (1-t)\lambda \mid t \in \mathbb{R}\}$ is invariant to a change in the representation of the curve.

Now, there is at least one such root since $e \in \mathcal{P}$. On the disjoint union $(-\infty, 0) \cup (1, \infty)$, there are at least $n-1$ roots since for sufficiently large values of t, the matrices $Q_{\lambda+t(e+\lambda)}$ and $Q_{\lambda-t(e-\lambda)}$ have opposite signature. Since 0 is a root and there is a root in $(0, 1]$, there can be no other roots, as $g(te + (1-t)\lambda)$ has degree $n+1$ as a univariate polynomial in the variable t. □

Lemma 5.3.8. *Suppose that $C \cap \mathbb{R}^3_+ \neq \varnothing$ and that $S \neq \varnothing$. Then, each connected component of $C \cap \mathbb{R}^3_+$ intersects a proper face of $\mathbb{R}^3_+$. Moreover, if a connected component of $C \cap \mathbb{R}^3_+$ intersects $\operatorname{int}(\mathbb{R}^3_+)$, then there exist $\lambda^{(1)} \neq \lambda^{(2)}$ in $C \cap \mathbb{R}^3_+$ with $|\operatorname{supp}(\lambda^{(i)})| \leqslant 2$.*

Proof. Note that the interior of C consists of λ such that Q_λ has n positive and one negative eigenvalue. If a connected component C_1 of $C \cap \mathbb{R}^3_+$ does not intersect a proper face of $\mathbb{R}^3_+$, then C_1 is contained entirely in $\operatorname{int}(\mathbb{R}^3_+)$. This implies that the image of C_1 in $\mathbb{RP}^2$ is an oval of the spectral curve of depth $\lfloor \frac{n+1}{2} \rfloor - 1$. Since there is only one such component, it must be the case that $C_1 = C$. but then the hyperbolicity cone $\mathcal{P}$ is contained in $\operatorname{int}(\mathbb{R}^3_+)$, which implies that $S = \varnothing$ by Proposition 5.0.2. So, it must be the case that C_1 intersects a proper face of $\mathbb{R}^3_+$.

The preceding paragraph shows that $\mathbb{R}^3_+$ cannot contain the entirety of the region bounded by C. So, if a connected component C_1 of $C \cap \mathbb{R}^3_+$ contains an element with strictly positive entries and there is a unique element λ with $|\operatorname{supp}(\lambda)| \leqslant 2$, then the entirety f the region bounded by C_1 is contained in $\mathbb{R}^3_+$. This implies that $C_1 = C$, a contradiction as above. □

The result of Lemma 5.3.8 is that we have control of the elements on the oval of the spectral curve of submaximal depth. We will leverage this below to obtain certificates of redundancy for aggregations in terms of the intersection of polyhedral cones with hyperbolicity cones.

Theorem 5.3.9. *Suppose that $\mathcal{V}_{\mathbb{R}}(f_1^h, f_2^h, f_3^h)$ is empty but $S \neq \varnothing$, the spectral curve is smooth and hyperbolic, and that $\Omega^n(\operatorname{cone}(\Lambda')^\circ)$ is contractible. Then, $\bigcap_{\lambda \in \Lambda'} S_\lambda = \bigcap_{\lambda \in \Lambda} S_\lambda$.*

Proof. Enumerate $\Lambda' = \{\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)}\}$ and let $\lambda^{(k+1)} \in \Lambda \backslash \operatorname{cone}(\Lambda')$. Note that if $g(\lambda^{(k+1)}) \neq 0$, then we can write $\lambda^{(k+1)}$ as a conical combination of an element of $\operatorname{cone}(\Lambda')$ and an element of C, where C is as above. In particular, it suffices to take $\lambda^{(k+1)} \in C$. Moreover, by the hypothesis that $S \neq \varnothing$, we have that $\mathcal{P} \nsubseteq \mathbb{R}^3_+$. This implies that there is $e \in \mathcal{P}$ and $\lambda^{(i)}, \lambda^{(j)} \in \Lambda$ such that $t\lambda^{(k+1)} + (1-t)e = \lambda$ for some $\lambda \in \operatorname{cone}(\lambda^{(i)}, \lambda^{(j)})$ at $0 \leqslant t \leqslant 1$. Since $\lambda^{(k+1)} \notin C$, we can take $\lambda^{(i)}, \lambda^{(j)} \in$C.

Set $K = \text{cone}(\lambda^{(i)}, \lambda^{(j)}, \lambda^{(k+1)}, -\lambda^{(k+1)})$. We will show that $\mathcal{P} \subseteq \text{int}(K^\circ)$. For notational convenience, denote by A the matrix such that $A^\top = \begin{bmatrix} \lambda^{(i)} & \lambda^{(j)} & \lambda^{(k+1)} \end{bmatrix}$, so that $K^\circ = A^\top(\mathbb{R}^2_+ \times \mathbb{R})$.

Suppose for the sake of a contradiction that $A^\top(\mathbb{R}^2_+ \times \mathbb{R}) \cap \mathcal{P} = \{0\}$. Let v be normal to a separating hyperplane oriented so that $v^\top y > 0$ for all $y \in \mathcal{P}$ and $v^\top w \leqslant 0$ for all $w \in A^\top(\mathbb{R}^2_+ \times \mathbb{R})$. Then, $v^\top \lambda^{(k+1)} < 0$ since $t\lambda^{(k+1)} + (1-t)e = \lambda$. But then, for $(c_1, c_2, d) \in \mathbb{R}^2_+, \mathbb{R}$, we have that $v^\top(d\lambda^{(k+1)} + c_1\lambda^{(i)} + c_2\lambda^{(j)})$ can take both positive and negative values since d is unconstrained. This contradicts the assumption that v was normal to a separating hyperplane. So, $\mathcal{P}$ and $A^\top(\mathbb{R}^2_+ \times \mathbb{R})$ intersect nontrivially. In fact, it must be the case that $\mathcal{P} \subseteq \text{int}(A^\top(\mathbb{R}^2_+ \times \mathbb{R}))$. Indeed, the face $\{t\lambda^{(k)} + s\lambda^{(i)} \mid t \in \mathbb{R}, s \in \mathbb{R}_+\}$ cannot intersect $\mathcal{P}$ (and similarly for the face defined by $\lambda^{(k)}$ and $\lambda^{(j)}$). This is a consequence of Lemma 5.3.7 since if there were an element e in this intersection, then there would be two roots of g when restricted to the line segment between $\lambda^{(k+1)}$ and e.

Since $\mathcal{P} \subseteq \text{int}(K^\circ)$, we have that either $H^1(\Omega^n(K^\circ)) \neq 0$ or $\Omega^{n+1}(K^\circ) \neq \varnothing$. By Proposition 5.0.2, this implies that $X(K^\circ, f^h) = \varnothing$. That is,

$$\mathcal{V}_\mathbb{R}(f^h_{\lambda^{(k+1)}}) \cap X(-\mathbb{R}^2, (f^h_{\lambda^{(1)}}, f^h_{\lambda^{(2)}})) = \varnothing.$$

This in turn implies that $\mathcal{V}_\mathbb{R}(f^h_{\lambda^{(k+1)}}) \cap X(\text{cone}(\Lambda')^\circ, f^h) = \varnothing$. By the hypothesis that $\Omega^n(\text{cone}(\Lambda')^\circ)$ is contractible and Proposition 5.3.4, this implies that $\lambda^{(k+1)}$ is unecessary, i.e., $\bigcap_{i=1}^k S_{\lambda^{(i)}} = \bigcap_{i=1}^{k+1} S_{\lambda^{(i)}}$. □

The following example demonstrates the proof of Theorem 5.3.9.

Example 5.3.1 (Certifying the Redundancy of an Aggregation)**.** Figure 5.2 illustrates the argument of the proof of Theorem 5.3.9. Again, we take the system of quadratics from Example 5.2.1 so that the curve $\mathcal{V}_\mathbb{R}(g)$ is hyperbolic but g is not a definite representation. The aggregation μ which lies in the interior of $\mathbb{R}^3_+$ does not contribute to the intersection of all aggregations of correct signature, which is given by $\bigcap_{\lambda \in \Lambda'} S_\lambda$. This is certified by the fact

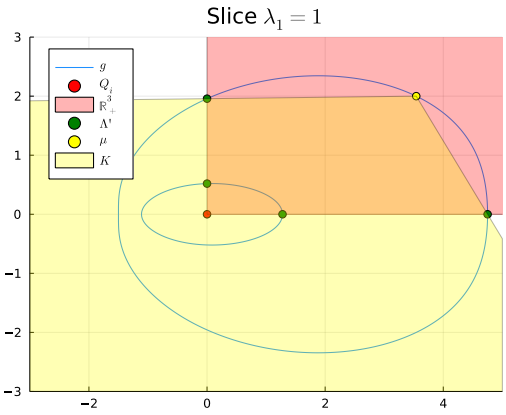


Figure 5.2: An illustration of Theorem 5.3.9. The aggregation μ is unnecessary for the description of $\bigcap_{\lambda\in\Lambda'} S_\lambda$ since $\mathcal{P} \subseteq K$.

that the cone $K = A^\top(\mathbb{R}^4_+ \times \mathbb{R})$ contains the hyperbolicity cone of g in its interior. Here, $A^\top = \begin{bmatrix} \lambda^{(1)} & \lambda^{(2)} & \lambda^{(3)} & \lambda^{(4)} & \mu \end{bmatrix}$ for an enumeration $\Lambda' = \{\lambda^{(i)} \mid i \in [4]\}$. In cohomological terms, $H^1(\Omega^n(K^\circ)) = \mathbb{Z}_2$.

◇

We also show via example that the condition that $\Omega^n(\text{cone}(\Lambda')^\circ)$ is contractible is a nontrivial hypothesis.

Example 5.3.2 (A Necessary Aggregation with Positive Entries)**.** We construct a system of quadratics such that $\bigcap_{\lambda\in\Lambda'} S_\lambda$ has three connected components, but for some μ with strictly positive entries, $S_\mu \cap (\bigcap_{\lambda\in\Lambda'} S_\lambda)$ has only two components. As in Example 5.2.1, we start with the matrices

$$M_1 = \begin{bmatrix} 25 & 0 & -32 & 0 \\ 0 & 25 & 0 & 24 \\ -32 & 0 & 6 & 0 \\ 0 & 24 & 0 & 6 \end{bmatrix}, \quad M_2 = \begin{bmatrix} 0 & 0 & 12 & 0 \\ 0 & 0 & 0 & 16 \\ 12 & 0 & 4 & 0 \\ 0 & 16 & 0 & 4 \end{bmatrix}, \text{ and } M_3 = \begin{bmatrix} 0 & 0 & 0 & -60 \\ 0 & 0 & 10 & 0 \\ 0 & 10 & 0 & 0 \\ -60 & 0 & 0 & 0 \end{bmatrix}.$$

The matrices do not satisfy PDLC and $\det(xM_1 + yM_2 + zM_3)$ is hyperbolic with respect

to $(1,0,0)$. We construct the defining quadratics $f_i(x,y,z) = \begin{bmatrix} x & y & z & 1 \end{bmatrix} Q_i \begin{bmatrix} x & y & z & 1 \end{bmatrix}^\top$ with Q_i defined as follows:

$$Q_1 = M_1 + 1.5M_2 = \begin{bmatrix} 25 & 0 & -14 & 0 \\ 0 & 25 & 0 & 48 \\ -14 & 0 & 12 & 0 \\ 0 & 48 & 0 & 12 \end{bmatrix}$$

$$Q_2 = M_1 - 2M_2 + 2M_3 = \begin{bmatrix} 25 & 0 & -56 & -120 \\ 0 & 25 & 20 & -8 \\ -56 & 20 & -2 & 0 \\ -120 & -8 & 0 & -2 \end{bmatrix}$$

$$Q_3 = M_1 - 2M_2 - 2M_3 = \begin{bmatrix} 25 & 0 & -56 & 120 \\ 0 & 25 & -20 & -8 \\ -56 & -20 & -2 & 0 \\ 120 & -8 & 0 & -2 \end{bmatrix}$$

We numerically compute that the elements of Λ' are appropriately normalized vectors in the directions of the following vectors $\lambda^{(1)}, \lambda^{(2)}, \lambda^{(3)}$. Since the semialgebraic set S_λ is invariant under positive scalings of λ, we work directly with the $\lambda^{(i)}$. We also take an aggregation μ which lies on the oval of depth $\lfloor \frac{n+1}{2} \rfloor - 1$ with strictly positive entries.

$$\lambda^{(1)} = e_1, \quad \lambda^{(2)} = (1, 0.68725, 0), \quad \lambda^{(3)} = (1, 0, 0.68725), \quad \mu = (0.1429, 0.4286, 0.4286)$$

The intersection of the cone $K = \text{cone}(\lambda^{(1)}, \lambda^{(2)}, \lambda^{(3)})$ with $\Omega^n(-\mathbb{R}^3_+)$ has three connected components and the intersection of the cone $\tilde{K} = \text{cone}(\lambda^{(1)}, \lambda^{(2)}, \lambda^{(3)}, \mu)$ with $\Omega^n(-\mathbb{R}^3_+)$ has two connected components. The corresponding semialgebraic sets $\bigcap_{i=1}^3 S_{\lambda^{(i)}}$ and $S_\mu \cap \left(\bigcap_{i=1}^3 S_{\lambda^{(i)}}\right)$ have three and two connected components, respectively. However, as in the

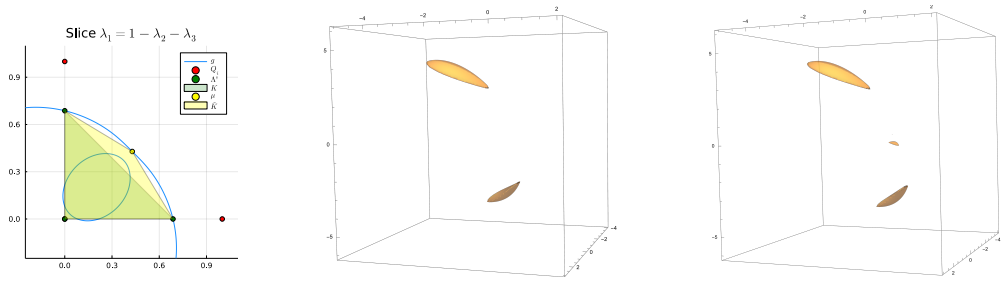


Figure 5.3: Plots for Example 5.3.2. The left figure displays the spectral curve and relevant aggregations and cones. The center figure shows the set $S_\mu \cap (\bigcap_{\lambda\in\Lambda'} S_\lambda)$ which has two connected components. The right figure shows $\bigcap_{\lambda\in\Lambda'} S_\lambda$, which has three connected components.

proof of Theorem 5.3.9, $\mathcal{V}_{\mathbb{R}}(f_\mu^h) \cap \left(\bigcap_{i=1}^3 S_{\lambda^{(i)}}\right) = \varnothing$. So intersecting with S_μ has the effect of cutting off a component of $\left(\bigcap_{i=1}^3 S_{\lambda^{(i)}}\right)$. In particular, we see that $\bigcap_{\lambda\in\Lambda'} S_\lambda \neq \bigcap_{\lambda\in\Lambda} S_\lambda$. The spectral curve and the relevant semialgebraic sets are shown in Figure 5.3.

◇

As shown in Example 5.3.2, the set defined by the intersection of all permissible aggregations is not necessarily defined by the intersection of aggregations in Λ' when $\Omega^n(\circ(\Lambda)')$ is not contractible. However, we are able to determine a bound on the number of aggregations with strictly positive entries which are needed in such a description in the case where PDLC is not satisfied and the set defined by all permissible aggregations is not connected. As a starting point, we observe that adding an aggregation $\lambda \in \operatorname{int}(\mathbb{R}^3_+)$ to Λ' cannot make the number of connected components of the set defined by the intersection of aggregations increase.

Lemma 5.3.10. *Set $K_0 = \operatorname{cone}(\Lambda')$ and let $\mu^{(1)}, \mu^{(2)}, \ldots, \mu^{(N)} \in \Lambda$ have strictly positive entries. For each $j \in [N]$, set $K_j = \operatorname{cone}(\Lambda' \cup \bigcup_{i=1}^j \{\mu^{(i)}\})$. Then, every connected component of $\Omega^n(K_j^\circ)$ intersects a component of $\Omega^n(K_0^\circ)$. Moreover, we have that*

$$1 \leqslant \dim H^0(\Omega^n(K_N^\circ)) \leqslant \dim H^0(\Omega^n(K_{N-1}^\circ)) \leqslant \cdots \leqslant \dim H^0(\Omega^n(K_1^\circ)) \leqslant \dim H^0(\Omega^n(K_0^\circ)).$$

Proof. First, we show that there are no connected components of $\Omega^n(K_j^\circ)$ which consist of points with all positive entries. Suppose for the sake of a contradiction that there was such a component. Then, its boundary would form an oval of $\mathcal{V}_\mathbb{R}(g)$ whose interior contains matrices with n positivev eigenvalues. If $\mathcal{V}_\mathbb{R}(f_1^h, f_2^h, f_3^h) = \varnothing$, then $\mathcal{V}_\mathbb{R}(g)$ is hyperbolic and the only oval which contains matrices with n positive eigenvalues in its interior is the oval of depth $\lfloor \frac{n+1}{2} \rfloor - 1$. If K_j contains this oval, then it also contains the oval of maximal depth, and thus the hyperbolicity cone of g which certifies the emptiness of the variety. By Proposition 5.0.2, this would implies that $X(K_j^\circ, f^h) = \varnothing$, a contradiction since $X(-\mathbb{R}^3_+, f^h) \subseteq X(K_j^\circ, f^h)$.

So, every connected component of $\Omega^n(K_j^\circ)$ has nonempty intersection with $\Omega^n(K_0^\circ)$ since there is λ in each component with some coordinate $\lambda_i = 0$. Therefore, either $\lambda \in \Lambda'$ or λ is a conical combination of two elements of Λ'. Since $\Omega^n(K_0^\circ) \subseteq \Omega^n(K_{j-1}^\circ) \subseteq \Omega^n(K_j^\circ)$ for all j, this implies that every connected component of $\Omega^n(K_j^\circ)$ intersects $\Omega^n(K_{j-1}^\circ)$.

The last claim then follows since there are at least as many connected components of $\Omega^n(K_{j-1}^\circ)$ as connected components of $\Omega^n(K_j^\circ)$. Finally, since $X(-\mathbb{R}^3_+, f^h) \neq \varnothing$, we have that $\dim H^0(\Omega^n(K_N^\circ)) \geqslant 1$. □

We are now prepared to prove Theorem 5.0.4, which states that the intersection over all permissible aggregations is given by the intersection over a finite subset Λ_1 of permissible aggregations. The essential idea is that the number of aggregations with strictly positive components which can be necessary is bounded above by the number of connected components of $\dim H^0(\Omega^n(\text{cone}(\Lambda')^\circ))$, or equivalently the number of connected components of $X(\text{cone}(\Lambda')^\circ, f^h)$ by an application of Theorem 5.1.1.

Proof of Theorem 5.0.4. If $\Omega^n(\text{cone}(\Lambda')^\circ)$ is contractible, then $\Lambda_1 = \Lambda'$ by Theorem 5.3.9. Otherwise, if $\Omega^n(\text{cone}(\Lambda')^\circ)$ is not contractible, we can apply Lemma 5.3.8. Specifically, fix an enumeration $\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)}$ of Λ'. A list $\mu^{(1)}, \mu^{(2)}, \ldots, \mu^{(N)}$ of aggregations with strictly positive entries only contains necessary elements only if $\dim H^0(\Omega^n(K^circ_j)) < \dim H^0(\Omega^n(K_{j-1}^\circ))$. In particular, the maximal length N of a sequence of aggregations with positive entries which can be necessary is at most $\dim H^0(\Omega^n(\text{cone}(\Lambda')^\circ))$. □

5.3.3 Improved Upper Bounds in the PDLC Case

In this section, we study the topology of the set of good aggregations as a subset of the set of all permissible aggregations. We pay particular attention to the PDLC case, where, under assumptions on the regularity of the set S, we obtain an improved bound on the number of necessary aggregations to describe $\overline{\text{conv}}(S)$. As a starting point, we show that when the set S has no low-dimensional components, i.e. $S = \overline{\text{int}(S)}$, then sets of good and bad aggregations are connected in the set of permissible aggregations.

Proposition 5.3.11. *Suppose that $S = \overline{\text{int}(S)}$. Then, each connected component of Λ consists of only good aggregations or only bad aggregations.*

Proof. We first show that the set of good aggregations is closed in Λ. Let $\lambda^{(i)} \to \lambda$ be a sequence of aggregations converging to $\lambda \in \Lambda$. Recall that the matrix Q_λ has the block structure $Q_\lambda = \begin{bmatrix} A_\lambda & b_\lambda \\ b_\lambda^\top & c_\lambda \end{bmatrix}$, where the quadratic form associated to $A_\lambda \in \mathbb{R}^{n\times n}$ is the homogeneous part of f_λ. If $\text{int}(S_{\lambda^{(i)}})$ is convex for sufficiently large i, then $A_{\lambda^{(i)}} \succeq 0$ for sufficiently large i. This in turn implies that $\text{int}(S_\lambda)$ is convex since the cone of positive semidefinite matrices is closed and therefore $A_\lambda \succeq 0$. Otherwise, suppose that $\text{int}(S_{\lambda^{(i)}})$ has two convex components for large i. Set $(\alpha^{(i)}, \beta^{(i)}) \in \mathbb{R}^n \times \mathbb{R}$ to be the unit length eigenvector corresponding to the negative eigenvalue of $Q_{\lambda^{(i)}}$ for each i, with sign chosen such that $(\alpha^{(i)})^\top x \leqslant \beta^{(i)}$ for all $x \in S$. This is possible since each $\lambda^{(i)}$ is a good aggregation and the affine hyperplane defined by $(\alpha^{(i)})^\top x = \beta^{(i)}$ separates the two components of $\text{int}(S_{\lambda^{(i)}})$. Now, $\alpha^{(i)} \to \alpha$ and $\beta^{(i)} \to \beta$ where (α, β) is an eigenvector of Q_λ corresponding to the negative eigenvalue and such that $\alpha^\top x \leqslant \beta$ for all $x \in S$. So, λ is a good aggregation.

Next, we show that the set of bad aggregations is also closed in Λ. Let $\lambda^{(i)} \to \lambda$ be a sequence of bad aggregations which converges to $\lambda \in \Lambda$. Let $(\alpha^{(i)}, \beta^{(i)})$ be the unit length eigenvector of $Q_{\lambda^{(i)}}$ corresponding to the negative eigenvalue of $Q_{\lambda^{(i)}}$ chosen with orientation such that $(\alpha^{(i)})^\top \to \alpha^\top$ and $\beta^{(i)} \to \beta$ where (α, β) is an eigenvector of Q_λ corresponding to the negative eigenvalue. Since each $\lambda^{(i)}$ is a bad aggregation, there are $x, y \in \text{int}(S)$ such

that $(\alpha^{(i)})^\top x < \beta^{(i)}$ and $(\alpha^{(i)})^\top y > \beta^{(i)}$ for all i sufficiently large. So, $\alpha^\top x \leqslant \beta$ and $\alpha^\top y \geqslant \beta$. Since $x, y \in \operatorname{int}(S)$, this in turn implies that there are $\hat{x}, \hat{y} \in \operatorname{int}(S)$ such that $\alpha^\top \hat{x} < \beta$ and $\alpha^\top \hat{y} > \beta$. So, λ is a bad aggregation.

To conclude, note that if Λ^* is a connected component of Λ, then Λ^* can be written as a disjoint union of good aggregations in Λ^* and bad aggregations in Λ^*. Since Λ^* is connected this implies that Λ^* consists entirely of good aggregations or entirely of bad aggregations. □

We will now restrict our attention to the PDLC case. Here, unlike in the previous sections, we do not assume that the spectral curve $\mathcal{V}_{\mathbb{R}}(g)$ is smooth. First, we show that the PDLC hypothesis implies that exactly one connected component of aggregations of correct signature consists of good aggregations.

Proposition 5.3.12. *Suppose that PDLC holds,* $\operatorname{int}(S) \neq \varnothing$*, and* S *has no points at infinity. Then, exactly one connected component of* $\Omega^n(-\mathbb{R}^3_+)$ *consists of good aggregations.*

Proof. Note that $\overline{\operatorname{conv}}(S)$ can be described as the intersection of finitely many good aggregations by [BDS24] since the $Q_\lambda \neq 0$ for all $\lambda \neq 0$ by the hypothesis that the Q_i are linearly independent and the map f^h satisfies hidden hyperplane convexity since PDLC holds. By the results of [BDS24], there are good aggregations $\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(r)}$ with $r \leqslant 6$ such that $\overline{\operatorname{conv}}(S) = \bigcap_{i=1}^r S_{\lambda^{(i)}}$. Set $K = \operatorname{cone}(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(r)})$.

Suppose for the sake of a contradiction that $\Omega^n(K^\circ)$ has at least two components. From the spectral sequence of Theorem 5.1.1, it follows that $\dim H_0(X(K^\circ, f^h)) \geqslant 2$. It therefore follows that the set T defined by $T = \bigcap_{\lambda \text{ a good aggregation}} S_\lambda$ has at least two components. On the other hand, if S has no points at infinity then every connected component of T must intersect S since otherwise there would be an affine hyperplane H with $H \cap T \neq \varnothing$ but $H \cap \operatorname{conv}(S) = \varnothing$. This can't happen when PDLC holds by [BDS24]. So, S intersects multiple connected components of T, which in turn implies that $\operatorname{conv}(S)$ intersects multiple connected components of T, which is a contradiction since $\operatorname{conv}(S)$ is connected. So, it must be the case that $\Omega^n(K^\circ)$ has only one component and therefore exactly one connected

component of $\Omega^n(-\mathbb{R}^3_+)$ contains good aggregations. □

We will show that the set defined by the intersection of all good aggregations can be defined by the intersection of at most four good aggregations by following a similar strategy to that of Proposition 5.3.2. Note that if PDLC holds and $S \neq \varnothing$, then the hyperbolicity cone $\mathcal{P}$ of g which contains positive semidefinite matrices does not intersect $\mathbb{R}^3_+$ away from 0. However, it can be the case that the hyperbolicity cone of g which contains negative definite matrices intersects $\mathbb{R}^3_+$ and moreover is completely contained in $\mathbb{R}^3_+$. When the hyperbolicity cone of g which contains negative definite matrices is not completely contained in $\mathbb{R}^3_+$, the strategy from Proposition 5.3.2 applies directly.

Proposition 5.3.13. *Suppose that PDLC holds,* $\operatorname{int}(S) \neq \varnothing$*, and S has no points at infinity. Assume that the hyperbolicity cone $\mathcal{P}$ of g which contains negative definite matrices is not contained in* $\mathbb{R}^3_+$*. Then, there are* $\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(r)}$ *such that* $r \leqslant 4$ *and* $\overline{\operatorname{conv}}(S) = \bigcap_{i=1}^{r} S_{\lambda^{(i)}}$.

Proof. If the hyperbolicity cone $\mathcal{P}$ of g which contains negative definite matrices is not conttained in $\mathbb{R}^3_+$, then there is $\theta \in \mathbb{R}^3$ such that $Q_\theta \succeq 0$ and at least one component of θ is strictly positive, without loss generality $\theta_3 > 0$. Note that since $\operatorname{int}(S) \neq \varnothing$, it follows that $\theta \notin \mathbb{R}^3_+$ and at least one of $\theta_1, \theta_2 < 0$.

If $\lambda = \lambda_1 e_1 + \lambda_2 e_2 \in \Lambda$ is an aggregation for $\lambda_1, \lambda_2 > 0$, we have that $\lambda + \epsilon\theta \in \mathbb{R}^3_+ \backslash \{0\}$, where $\epsilon = \min\left\{ \frac{-\lambda_i}{\theta_i} \mid \theta_i < 0 \right\} > 0$. But then, by Proposition 5.3.2, $\lambda + \epsilon\theta$ is a good aggregation with $S_{\lambda+\epsilon\theta} \subseteq S_\lambda$. By the choice of ϵ, we additionally have that $\lambda + \epsilon\theta \in \operatorname{cone}(e_1, e_3) \cup \operatorname{cone}(e_2, e_3)$. In particular, by improving aggregations in this way, no aggregation in $\operatorname{cone}(e_1, e_2)$ is necessary. By Proposition 5.3.3 and the fact that all aggregations in $\operatorname{cone}(e_i, e_j)$ can be described using at most two aggregations, it follows that $\overline{\operatorname{conv}}(S)$ can be described as the intersection of at most four aggregations. □

The case where the cone of negative definite matrices is completely contained in $\mathbb{R}^3_+$ is more subtle. To deal with this case, we fix the following notation stratifying subsets of aggregations.

- $\Lambda_1 = \{\lambda \in \Lambda \mid \lambda \text{ is a good aggregation and } S_\lambda \neq \mathbb{R}^n\}$.
- $\Lambda_2 = \{\lambda \in \Lambda_1 \mid |\operatorname{supp}(\lambda)| \leqslant 2\}$.
- $\Lambda_3 = \{\lambda \in \Lambda_2 \mid \|\lambda\| = 1 \text{ and } \lambda \text{ generates an extreme ray of } \Lambda_2\}$.

We will also need the following technical lemmas regarding the relationship between the negative definite cone and $\operatorname{cone}(\Lambda_3)$.

Lemma 5.3.14. *Suppose that the hyperbolicity cone of g which contains negative definite matrices is strictly contained in $\mathbb{R}^3_+$. Suppose further that $\lambda \in \Lambda_3$ has $g(\lambda) = 0$ and that $\omega \in \mathbb{R}^3$ has $Q_\omega \prec 0$. Then, for any $t \in (0,1)$, the matrix $Q_{t\lambda+(1-t)\omega}$ has at most $n-1$ positive eigenvalues.*

Proof. Consider the univariate polynomial $g(t\lambda + (1-t)\omega)$. Suppose for the sake of a contradiction that there is $t^* \in (0,1)$ such that $Q_{t^*\lambda+(1-t^*)\omega}$ has n positive eigenvalues. Since $g(\lambda) = 0$ and since Q_ω has $n+1$ negative eigenvalues, there are $n+1$ roots of $g(t\lambda+(1-t)\omega)$ on the interval $(0,1]$ when counted with multiplicity. There must additionally be a root t^* with $t^* < 0$ since the line connecting λ and ω must have two points of intersection with the boundary of the hyperbolicity cone. But then, the nonconstant degree $n+1$ polynomial $g(t\lambda + (1-t)\omega)$ has at least $n+2$ roots, the desired contradiction. □

Lemma 5.3.15. *Suppose that $n \geqslant 3$ and that the hyperbolicity cone of g which contains negative definite matrices is strictly contained in $\mathbb{R}^3_+$. Let $\omega \in \operatorname{cone}(\Lambda_3)$ be nonzero. Then, Q_ω has at least one nonnegative eigenvalue.*

Proof. By Carathéodory's Theorem, ω is a conical combination of some $\lambda^{(1)}, \lambda^{(2)}, \lambda^{(3)} \in \Lambda_3$. Let V_1, V_2, V_3 be the n-dimensional subspaces of $\mathbb{R}^{n+1}$ such that $v^\top Q_{\lambda^{(i)}} v \geqslant 0$ for $v \in V_1$. Since $n \geqslant 3$, we have that the vector subspace $V = V_1 \cap V_2 \cap V_3$ has codimension at most 3 in $\mathbb{R}^{n+1}$ and therefore $\dim(V) \geqslant 1$. Now, for a nonzero $v \in V$ and $t_1, t_2, t_3 \geqslant 0$, we compute that

$$v^\top \left(\sum_{i=1}^{3} t_i Q_{\lambda^{(i)}} \right) v \geqslant 0$$

and therefore Q_ω has at least one nonnegative eigenvalue for each $\omega \in \operatorname{cone}(\lambda^{(1)}, \lambda^{(2)}, \lambda^{(3)})$. □

We can now prove the improved bound on the number of needed aggregations in the PDLC case.

Proof of Theorem 5.0.5. The case for which the hyperbolicity cone of g which contains negative definite matrices is *not* contained in $\mathbb{R}^3_+$ is proven in Proposition 5.3.13. For the case where the hyperbolicity cone is contained in $\mathbb{R}^3_+$, we separate by the cases $n = 1, n = 2, n \geqslant 3$.

Note that it follows from [BDS24] that $\overline{\operatorname{conv}}(S) = \bigcap_{\lambda \in \Lambda_2} S_\lambda$ and that since a convex combination of good aggregations is a good aggregation if it is permissible, $\overline{\operatorname{conv}}(S) = \bigcap_{\lambda \in |\Lambda_3|} S_\lambda$. So, it suffices to bound $|\Lambda_3|$. Recall that if $\lambda \in \Lambda_3$, then Q_λ has exactly one negative eigenvalue, as otherwise, $\operatorname{int}(S) = \varnothing$. Finally recall that the set of good aggregations is connected by Propositions 5.3.11 and 5.3.12.

In the cases $n = 1$ and $n = 2$, we are able to use the structure of the spectral curve to explicitly bound $|\Lambda_3|$.

$n = 1$: If $\lambda \in \mathbb{R}^3_+$, then either $Q_\lambda \prec 0$ or $\lambda \in \Lambda$. Since the cone of negative definite matrices is contained in $\mathbb{R}^3_+$, it follows that every nonzero $\lambda \in \mathbb{R}^3_+$ with $|\operatorname{supp}(\lambda)| \leqslant 2$ has exactly one negative eigenvalue. So, Λ is connected and therefore the set of good aggregations is connected and $\Lambda_3 = \{e_1, e_2, e_3\}$, so $|\Lambda_3| = 3 \leqslant 4$.

$n = 2$: In this case, if $\lambda \in \Lambda_3$ is not a standard basis vector, then λ lies on the non-oval component of the spectral curve. Note that the line through two standard basis vectors e_j and e_k can only intersect this component once. So, if a face $\operatorname{cone}(e_j, e_k)$ of $\mathbb{R}^3_+$ contains two elements of Λ_3, it must be the case that either these two components are e_j and e_k or that exactly one of e_j and e_k is an element of Λ_3. So, up to relabeling the Q_i, there are three possibilities for the set Λ_3:

$$\Lambda_3 \in \left\{\{e_1, e_2, e_3\}, \{e_1, \lambda^{(2)}, \lambda^{(3)}\}, \{e_2, e_3, \lambda^{(2)}, \lambda^{(3)}\}\right\},$$

where $\lambda^{(i)} \in \text{cone}(e_1, e_i)$. In all cases, $|\Lambda_3| \leqslant 4$.

$n \geqslant 3$: The case $n \geqslant 3$ is more technical. Let $\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(r)}$ be such that $\overline{\text{conv}}(S) = \bigcap_{i=1}^{r} S_{\lambda^{(i)}}$. Set $K = \text{cone}(\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(r)})$, and note that at most two of the $\lambda^{(i)}$ can lie on a given facet $\text{cone}(e_j, e_k)$ of $\mathbb{R}^3_+$. Note that $\Omega^n := \Omega^n(K^\circ)$ is nonempty and connected. Indeed, if Ω^n were empty, then S would intersect nontirivially with every affine hyperplane in $\mathbb{R}^n$ by Proposition 5.0.2 and therefore $\overline{\text{conv}}(S) = \mathbb{R}^n$. Connectedness follows from Proposition 5.3.12.

We now show that the connected component of Λ which contains Ω^n contains exactly two aggregations $\lambda^{(1)}$ and $\lambda^{(2)}$ (up to relabeling) and that this implies that $\overline{\text{conv}}(S) = S_{\lambda^{(1)}} \cap S_{\lambda^{(2)}}$. Note that we know that there are $\lambda^{(1)}, \lambda^{(2)}$ in the connected component of Λ which contains Ω^n coming from the intersection of the spectral curve with the boundary of $\mathbb{R}^3_+$.

Suppose for the sake of a contradiction that there is $\lambda^{(3)}$ in the same connected component of Λ as $\lambda^{(1)}$ and $\lambda^{(2)}$. Given ω with $Q_\omega \prec 0$, we consider the lines L_i connecting ω and $\lambda^{(i)}$ in $\mathbb{R}^3$ and their restrictions to the sphere $\hat{L}_i = \{\frac{\ell}{\|\ell\|} \mid \ell \in L_i\}$. For any $i \in [3]$ and $\mu \in \hat{L}_i$, we have that Q_μ has at most $n-1$ positive eigenvalues. In particular, we see that $\hat{L}_i \cap \Omega^n = \varnothing$ for each $i \in [3]$ and therefore Ω^n has at least two connected components, a contradiction.

To show that $\overline{\text{conv}}(S) = S_{\lambda^{(1)}} \cap S_{\lambda^{(2)}}$, we note that it suffice to show that for any affine hyperplane $H \subseteq \mathbb{R}^n$,

$$\left(\bigcap_{i=1}^{r} S_{\lambda^{(i)}}\right) \cap H = \varnothing \iff S_{\lambda^{(1)}} \cap S_{\lambda^{(2)}} \cap H = \varnothing.$$

Let H be an affine hyperplane such that $(\bigcap_{i=1}^{r} S_{\lambda^{(i)}}) \cap H = \varnothing$. By Proposition 5.0.2, we know that $\Omega^n_H(K^\circ) \neq \varnothing$, since PDLC is satisfied and therefore $H^1(\Omega^{n-1}_H(K^\circ)) = 0$. By the Cauchy Interlacing theorem, it must be the case that $\Omega^n_H(K^\circ) \subseteq \Omega^n$. Moreover, g_H

interlaces g so that the hyperbolicity cone of g_H containing positive definite matrices cannot be completely contained in $\mathbb{R}^3_+$ and therefore $\Omega^n_H(K^\circ) \cap \text{cone}(\lambda^{(1)}, \lambda^{(2)}) \neq \varnothing$. But then, since $\Omega^n_H(K^\circ) \cap \text{cone}(\lambda^{(1)}, \lambda^{(2)}) \neq \varnothing$, it must be the case that $S_{\lambda^{(1)}} \cap S_{\lambda^{(2)}} \cap H = \varnothing$ by Proposition 5.0.2. □

5.4 Computing the Convex Hull

In this section, we derive a sufficient condition for the expression

$$\overline{\text{conv}}(S) = \bigcap_{\lambda \in \Lambda_1} S_\lambda,$$

proving Theorem 5.0.3. Our strategy mirrors the strategy in [DMnS22, BDS24]. Loosely speaking, given a valid inequality on S, and H the hyperplane defining this valid inequality, we want to show that there is an aggregation λ such that the restriction $Q_\lambda|_H$ is positive definite.

Recall that we say that the set S has *no points at infinity* if

$$S^h \cap \{(x, 0) \in \mathbb{R}^{n+1}\} = \{0\}$$

Lemma 5.4.1. *Suppose that S has no points at infinity. If $\alpha \in \mathbb{R}^n$ and $\beta \in \mathbb{R}$ are such that $\alpha^\top x < \beta$ for all $x \in S$, and if $H = \{(u, u_{n+1}) \mid \alpha^\top u = \beta u_{n+1}\}$, is the associated hyperplane in $\mathbb{R}^{n+1}$, then $S^h \cap H = \{0\}$.*

Proof. If $0 \neq (\hat{u}, \hat{u}_{n+1}) \in S^h \cap H$, then $\frac{\hat{u}}{\hat{u}_{n+1}} \in S$ and $\alpha^\top \left(\frac{\hat{u}}{\hat{u}_{n+1}}\right) = \beta$. □

We now record a strategy for obtaining sufficient conditions for the expression of $\overline{\text{conv}}(S)$ in terms of good aggregations. Variants of this strategy were used in [DMnS22, BDS24].

Proposition 5.4.2. *Suppose that* $\text{int}(S) \neq \varnothing$ *and that the matrices Q_i satisfy the following property:*

If $\alpha^\top x < \beta$ is a valid inequality on S, then for the hyperplane
$H = \{(x, x_{n+1}) \mid \alpha^\top x = \beta x_{n+1}\} \subseteq \mathbb{R}^{n+1}$, *there exists* $\lambda \in \mathbb{R}^3_+$ *such that* $Q_\lambda|_H \succ 0$. (*)

Then, there is a set $\Lambda_1 \subseteq \mathbb{R}^3_+$ *of good aggregations such that* $\overline{\text{conv}}(S) = \bigcap_{\lambda\in\Lambda_1} S_\lambda$.

Proof. Let $y \notin \overline{\text{conv}}(S)$. Then, there is $\alpha \in \mathbb{R}^n$ such that $\alpha^\top x < \alpha^\top y$ for all $x \in S$. Set $\beta = \alpha^\top y$ and $H = \{(x, x_{n+1}) \mid \alpha^\top x = \beta x_{n+1}\}$. If property (*) holds, then there is an aggregation $\lambda \in \mathbb{R}^3_+$ such that $Q_\lambda|_H \succ 0$. By the Cauchy Interlacing Theorem and the hypothesis that $\text{int}(S) \neq \varnothing$, it follows that Q_λ has exactly one negative eigenvalue. Since $Q_\lambda|_H \succ 0$, this in turn implies that either S_λ is convex or S_λ is the disjoint union of two convex components separated by the affine hyperplane $\{x \in \mathbb{R}^n \mid \alpha^\top x = \beta\}$. Since $\alpha^\top x < \beta$ is a valid inequality on S, this implies that $\overline{\text{conv}}(S)$ is contained in a single convex connected component of S_λ and therefore λ is a good aggregation which certifies that $y \notin \overline{\text{conv}}(S)$. □

So, in order to develop a sufficient condition for $\overline{\text{conv}}(S)$ to be given by aggregations, we can search for properties which ensure that $S^h \cap H = \{0\}$ is certified by positive definite aggregations $Q_\lambda|_H$ for any hyperplane H. We will restrict our attention to the setting where $\mathcal{V}_\mathbb{R}(f_1^h, f_2^h, f_3^h) = \varnothing$ and the spectral curve is smooth. By Theorem 5.0.1, we know that the spectral curve must be hyperbolic with a hyperbolicity cone $\mathcal{P}$ such that $\text{int}(\mathcal{P})$ has either positive definite matrices or matrices with exactly two negative eigenvalues. On the other hand, if $\mathcal{V}_\mathbb{R}(f_1^h, f_2^h, f_3^h) = \varnothing$, and $H \subseteq \mathbb{RP}^n$ is a hyperplane, then $\mathcal{V}_\mathbb{R}(f_1^h, f_2^h, f_3^h) \cap H$ must be empty. So, the restricted spectral curve $g_H(\lambda) = \det(Q_\lambda|_H)$ must also be hyperbolic if it is smooth and have hyperbolicity cone $\mathcal{P}_H$ such that $\text{int}(\mathcal{P}_H)$ either has positive definite combinations or combinations with exactly two negative eigenvalues.

So, we want to understand the way the ovals of $\mathcal{V}_\mathbb{R}(g)$ and $\mathcal{V}_\mathbb{R}(g_H)$ interact. Note that if Q_1, Q_2, Q_3 satisfy PDLC then so do $Q_1|_H, Q_2|_H, Q_3|_H$ and the ovals of $\mathcal{V}_\mathbb{R}(g_H)$ interlace those of $\mathcal{V}_\mathbb{R}(g)$ [KPV15]. When $\mathcal{P}$ contains matrices with two negative eigenvalues, the realtionship

is more subtle. Restricting our attention to a pencil gives the following restriction on the interactions between the ovals [Tho76, Tho91].

Lemma 5.4.3. *Suppose that g is smooth and that $\lambda^{(1)}, \lambda^{(2)} \in \mathbb{R}^3$ are such that $g(\lambda^{(i)}) = 0$ for $i = 1, 2$ and that $g(t\lambda^{(1)} + (1-t)\lambda^{(2)}) \neq 0$ for $t \in (0,1)$. Suppose further that $g_H(\lambda^{(i)}) \neq 0$ for $i = 1, 2$. The number of zeros of $g_H(t\lambda^{(1)} + (1-t)\lambda^{(2)})$ for $t \in (0,1)$ is even if $Q_{\lambda^{(1)}}$ and $Q_{\lambda^{(2)}}$ have the same signature and odd otherwise.*

Proof. Let $r \in \mathbb{Z}$ be such that $Q_\lambda^{(1)}$ has r positive eigenvalues, $n-r$ negative eigenvalues, and 0 as an eigenvalue of multiplicity one. Then, $Q_{\lambda^{(1)}}|_H$ has r positive eigenvalues and $n-r$ negative eigenvalues as well by the Cauchy Interlacing Theorem.

If $Q_{\lambda^{(1)}}$ and $Q_{\lambda^{(2)}}$ have the same signature, then $Q_\lambda^{(2)}|_H$ also has r positive and $n-r$ negative eigenvalues. So, $g_H(\lambda^{(1)})$ and $g_H(\lambda^{(2)})$ have the same sign and therefore $g_H(t\lambda^{(1)} + (1-t)\lambda^{(2)})$ has an even number of zeros for $t \in (0,1)$.

If $Q_{\lambda^{(2)}}$ has $r+1$ positive and $n-r-1$ negative eigenvalues and 0 as an eigenvalues of multiplicity one, then $Q_{\lambda^{(2)}}|_H$ has $r+1$ positive and $n-r-1$ negative eigenvalues by the Cauchy Interlacing Theorem. In particular, $g_H(\lambda^{(1)})$ and $g_H(\lambda^{(2)})$ have opposite sign so that $g_H(t\lambda^{(1)} + (1-t)\lambda^{(2)})$ must have an odd number of zeros on $t \in (0,1)$. The case where $Q_{\lambda^{(2)}}$ has $r-1$ positive eigenvalues is similar. □

In the case where g is hyperbolic and $\mathcal{P}$ contains matrices with two negative eigenvalues, Lemma 5.4.3 says that there are only three possibilities for the ovals of $\mathcal{V}_\mathbb{R}(g_H)$. The ovals of $\mathcal{V}_\mathbb{R}(g_H)$ interlace the ovals of $\mathcal{V}_\mathbb{R}(g)$ where the signature changes, and the hyperbolicity cone $\mathcal{P}_H$ of g_H either is contained in the hyperolicity cone $\mathcal{P}$, is between the ovals of $\mathcal{V}_\mathbb{R}(g)$ of depth $\lfloor \frac{n+1}{2} \rfloor - 1$ and $\lfloor \frac{n+1}{2} \rfloor$ (and therefore contains $Q_\lambda|_H \succ 0$), or is between the ovals of $\mathcal{V}_\mathbb{R}(g)$ of depth $\lfloor \frac{n+1}{2} \rfloor - 2$ and $\lfloor \frac{n+1}{2} \rfloor - 1$. We will need to refine the information computed by Lemma 5.4.3 to eliminate the last possibility. To do so, we use the following spectral sequence from [AL12] which computes the relative homology of hyperplane sections of sets $X(K, f)$.

Theorem 5.4.4 ([AL12, Theorem D]). *Fix a polyhedral cone $K \subseteq \mathbb{R}^m$, a homogeneous quadratic map $p : \mathbb{R}^{n+1} \to \mathbb{R}^m$, and a hyperplane $\bar{H} \subseteq \mathbb{RP}^n$. There is a first quadrant cohomology spectral sequence (G_r, d_r) converging to $H_{n-*}(X(K,p), X(K,p) \cap \bar{H})$ with*

$$G_2^{i,j} = H^i(\Omega_H^j(K), \Omega^{j+1}(K)) \text{ for } j > 0, \;\; G_2^{i,0} = H^i(K^\circ \cap B^n, \Omega^1(K)).$$

We will apply Theorem 5.4.4 to hyperplane sections of an empty variety to understand the relationship between the spectral curves $g(\lambda)$ and $g_H(\lambda)$. In particular, we will see that it must be the case that every noncontractible loop in $\Omega_H^{n-1}(\{0\})$ can be deformed to a noncontractible loop in $\Omega^n(\{0\})$.

Corollary 5.4.5. *Fix a hyperplane $H \subseteq \mathbb{R}^{n+1}$. Suppose that $\mathcal{V}_\mathbb{R}(f_1^h, f_2^h, f_3^h) = \varnothing$ and that the polynomial g is smooth and hyperbolic. Suppose further that the hyperbolicity cone $\mathcal{P}$ of g does not contain a positive definite matrix. Then, $H^1(\Omega_H^{n-1}(\{0\}), \Omega^n(\{0\})) = 0$.*

Proof. Let (G_r, d_r) be the spectral sequence of Theorem 5.4.4. For notational convenience, set $X = \mathcal{V}_\mathbb{R}(f_1^h, f_2^h, f_3^h)$, $\Omega^j = \Omega^j(\{0\})$, and $\Omega_H^j = \Omega^j(\{0\})$.

If $n = 2$, then $G_\infty^{1,1} \cong \ker(d_2^{1,1} : G_2^{1,1} \to G_2^{3,0})$. Since X and $X \cap \bar{H}$ are both empty, it follows that $0 = H_0(X, X \cap \bar{H}) \cong G_\infty^{2,0} \oplus G_\infty^{1,1} \oplus G_\infty^{0,2}$. It then follows that $G_\infty^{1,1}$ must be zero and that $d_2^{1,1}$ is injective. Since $G_2^{3,0} = H^3(K^\circ \cap B^3, \Omega^1) = H^3(B^3, \mathbb{S}^2) = 0$, it therefore follows that $G_2^{1,1} = H^1(\Omega_H^1, \Omega^2) = 0$.

Now suppose that $n \geqslant 3$. We start by showing that $G_2^{i,j} = 0$ when $i \geqslant 3$ and $j \geqslant 1$. From the long exact sequence of the pair $(\Omega_H^j, \Omega^j + 1)$, there is an exact sequence

$$\cdots \to H^{i-1}(\Omega^{j+1}) \to H^i(\Omega_H^j, \Omega^{j+1}) \to H^i(\Omega_H^j) = 0 \to \ldots,$$

where $H^i(\Omega_H^j) = 0$ for $i \geqslant 3$ since $\Omega_H^j \subseteq \mathbb{S}^2$. So, it suffices to show $H^{i-1}(\Omega^{j+1}) = 0$ for $i \geqslant 3$ and $j \geqslant 1$. By Theorem 5.0.1 and the hypothesis that PDLC does not hold, we see that for $n \geqslant 4$, Ω^{j+1} is homotopy equivalent to the union of two points when $j = 1$, homotopy equivalent to a point when $2 \leqslant j \leqslant n-2$, homotopy equivalent to $\mathbb{S}^1$ when $j = n-1$, and

empty when $j = n$. When $n = 3$, we see that Ω^{1+1} is homotopy equivalent to a disjoint union of two points, Ω^{2+1} is homotopy equialent to $\mathbb{S}^1$, and Ω^{3+1} is empty. In all cases, this implies that $H^{i-1}(\Omega^{j+1}) = 0$ for $i \geqslant 3$ and $j \geqslant 1$. This in turn implies that $H^i(\Omega_H^j, \Omega^{j+1}) = 0$ since $H^{i-1}(\Omega^{j+1}) \cong H^i(\Omega_H^j, \Omega^{j+1})$.

Next, we note that since g is smooth and PDLC does not hold, it must be the case that $\Omega^1 = \mathbb{S}^2$ and therefore $G^{i,0} = H^i(B^3, \Omega^1) = H^i(B^3, \mathbb{S}^2) = 0$ for $i \geqslant 4$.

So, we have shown that $G_2^{i,j} = 0$ for $i \geqslant 3, j \geqslant 1$ and for $i \geqslant 4$ with no assumption on j. Since $n \geqslant 3$, it therefore follows that $H^1(\Omega_H^{n-1}, \Omega^n) \cong G_2^{1,n-1}$ has stabilized so that $H^1(\Omega_H^{n-1}, \Omega^n) \cong G_2^{1,n-1}$. Since X and $X \cap \bar{H}$ are empty, it follows that

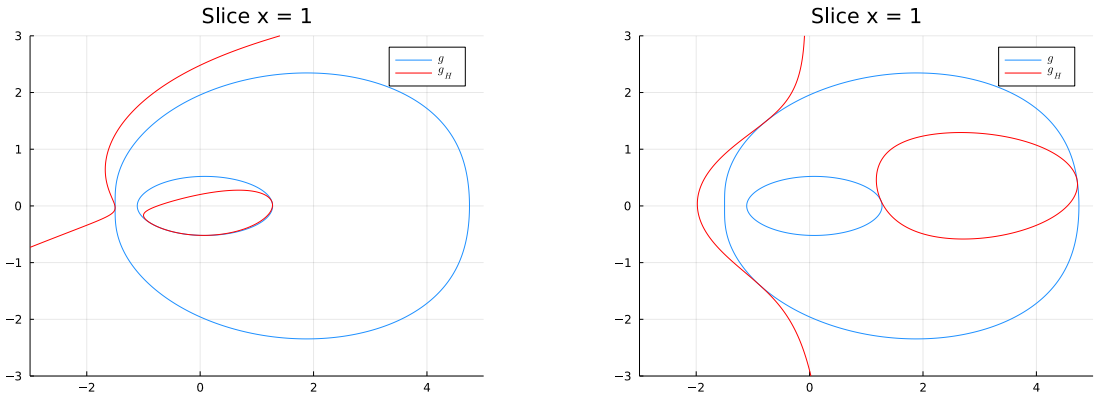
$$0 = H_0(X, X \cap \bar{H}) = \bigoplus_{j=0}^{n} G_\infty^{n-j,j}$$

and therefore $H^1(\Omega_H^{n-1}, \Omega^n) = 0$. □

Corollary 5.4.5 gives conditions on the relative cohomology groups of the Ω^j. We apply this to the setting where g and g_H are both smooth and hyperbolic and PDLC does not hold in order to determine the relationship between the hyperbolicity cones $\mathcal{P}$ and $\mathcal{P}_H$.

Theorem 5.4.6. *Suppose that $\mathcal{V}_\mathbb{R}(f_1^h, f_2^h, f_3^h) = \varnothing$ and that g is smooth and hyperbolic with hyperbolicity cone $\mathcal{P}$ such that $\operatorname{int}(\mathcal{P})$ contains matrices with exactly two negative eigenvalues. Then, if g_H is smooth and hyperbolic, either $\mathcal{P}_H$ contains positive definite matrices or $\mathcal{P}_H \subseteq \mathcal{P}$.*

Proof. Suppose that $\mathcal{P}_H$ does not contain positive definite matrices. By Theorem 5.0.1, it follwos that if $\lambda \in \operatorname{int}\mathcal{P}_H$, then $Q_\lambda|_H$ has exactly two negative eigenvalues and $n-2$ positive eigenvalues. Suppose for the sake of a contradiction that $\mathcal{P}_H \nsubseteq \mathcal{P}$. By Lemma 5.4.3 and the Cauchy Interlacing Theorem, this implies that the image of $\mathcal{P}_H \cap \mathbb{S}^2$ in $\mathbb{RP}^2$ is contained between the ovals of $\mathcal{V}_\mathbb{R}(g)$ of depth $\lfloor \frac{n+1}{2} \rfloor - 2$ and $\lfloor \frac{n+1}{2} \rfloor - 1$. So the images of $\mathcal{P}_H \cap \mathbb{S}^2$ and $\mathcal{P} \cap \mathbb{S}^2$ in $\mathbb{RP}^2$ are bounded by two disjoint ovals, neither of which contains the other. Since the image of Ω^n in $\mathbb{RP}^2$ is the region on the interior of the oval of $\mathcal{V}_\mathbb{R}(g)$ of



(a) The hyperbolicity cone $\mathcal{P}_H$ is contained in $\mathcal{P}$

(b) The hyperbolicity cone $\mathcal{P}_H$ lies between the ovals of g of maximal and submaximal depth and the restrictions $Q_1|_H$, $Q_2|_H$, $Q_3|_H$ satisfy PDLC.

Figure 5.4: Examples of the possible containment patterns for the hyperbolic curves g and g_H as given by Theorem 5.4.6.

depth $\lfloor\frac{n+1}{2}\rfloor - 1$ and the exterior of the oval of $\mathcal{V}_{\mathbb{R}}(g)$ of depth $\lfloor\frac{n+1}{2}\rfloor$, we see that there is a representative σ of the nontrivial class in $H_1(\Omega_H^n) \cong H^1(\Omega_H^n)$ which has nontrival intersection with the image of Ω^n. In particular, this implies that σ gives rise to a nontrivial class in $H_1(\Omega_H^{n-1}, \Omega^n) \cong H^1(\Omega_H^{n-1}, \Omega^n)$, contradicting Corollary 5.4.5. □

Examples of the two containment patterns are shown in Figure 5.4. Using the ideas of Theorem 5.4.6, we are able to prove Theorem 5.0.3. In particular, we see that in order to ensure that all certificates of emptiness for sets $S^h \cap H$ are obtained from positive definite aggregations $Q_\lambda|_H \succ 0$, it suffices to separate the cone $\mathcal{P}$ from the cone $\mathbb{R}^3_+$ of aggregations.

Proof of Theorem 5.4.6. We first show that condition (*) is satisfied before turning to the finiteness statement when the spectral curve is hyperbolic. Note that the PDLC case is settled by [BDS24], so assume that $\mathcal{P}$ does not contain a positive definite matrix.

Let H be a hyperplane such that $S^h \cap H = \{0\}$. This implies that the set $X(K, f^h) = \varnothing$ for $K = -\mathbb{R}^3_+$ the nonpositive orthant. By Proposition 5.0.2, either there is $\lambda \in \mathbb{R}^3_+$ such that $Q_\lambda|_H \succ 0$ or we have $H^1(\Omega_H^{n-1}(K)) \neq 0$. In either case, g_H is hyperbolic if it is smooth. Since $\Omega^{n-1}(K) \subseteq K^\circ \cap \mathbb{S}^2$, a nontrivial $H^1(\Omega_H^{n-1})$ would imply that $\mathcal{P}_H \subseteq K^\circ = \mathbb{R}^3_+$. This

cannot happen by Theorem 5.4.6 since no nontrivial aggregation lies in $\mathcal{P}$, i.e., $\mathcal{P}_H \cap \mathbb{R}^3_+ \subseteq \mathcal{P} \cap \mathbb{R}^3_+ = \{0\}$. If g_H is not smooth and $H^1(\Omega_H^{n-1}(K)) \neq 0$, then by the Cauchy Interlacing Theorem and the fact that $\mathcal{P} \cap \mathbb{R}^3_+ = 0$, there is $\lambda \in K^\circ$ such that $[\lambda] \in \mathbb{RP}^2$ lies on the outside of the oval $\mathcal{V}_\mathbb{R}(g)$ of depth $\lfloor \frac{n+1}{2} \rfloor - 1$ and such that $Q_\lambda|_H$ has at most $n-2$ positive eigenvalues. But then a nontrivial class in $H^1(\Omega_H^{n-1}(K))$ would give a nontrival class of $H^1(\Omega_H^{n-1}(K), \Omega^n(K))$, contradicting Corollary 5.4.5. So, there is $\lambda \in \mathbb{R}^3_+$ such that $Q_\lambda|_H \succ 0$ and thereofre (*) is satisfied.

In the case $n = 2$, if g is not hyperbolic, then (*) holds. Indeed, a nontrivial class in $H^1(\Omega^1_H(K))$ would give a nontrivial class in $H^1(\Omega^1_H(K), \Omega^2(K))$. Since this contradicts Corollary 5.4.5, this implies that there is $\lambda \in \mathbb{R}^3_+$ such that $Q_\lambda|_H \succ 0$.

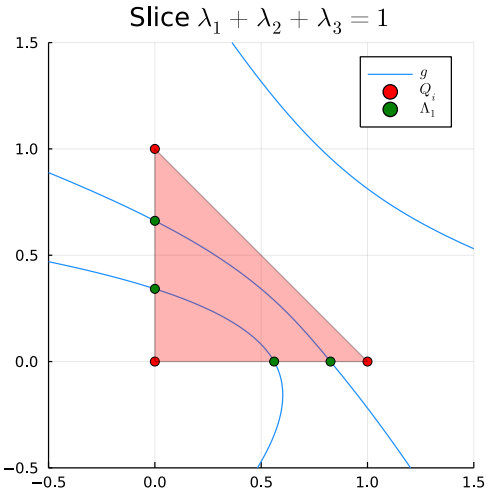
Finally, we show that when the spectral curve is hyperbolic, a finite number of good aggregations recovers $\overline{\text{conv}}(S)$. For each $y \in \mathbb{R}^n \backslash \overline{\text{conv}}(S)$, let $\lambda^{(y)} \in \mathbb{R}^3_+$ be a good aggregations certifying that $y \notin \overline{\text{conv}}(S)$. Setting $\Lambda^* = \{\lambda^{(y)} \mid y \in \mathbb{R}^n \backslash \overline{\text{conv}}(S)\}$ and $\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(k)}$ to be the unit length generators of the extreme rays of Λ^* which have $|\operatorname{supp}(\lambda^{(i)})| \leqslant 2$ gives that $\overline{\text{conv}}(S) = \bigcap_{i=1}^k S_{\lambda^{(i)}}$ by arguing as in the proof of Theorem 5.3.9. Note that $k \leqslant 6$ since each $\lambda^{(i)}$ generates an extreme ray of Λ^* and has $|\operatorname{supp}(\lambda^{(i)})| \leqslant 2$. □

We conclude with an example demonstrating the result of Theorem 5.0.3.

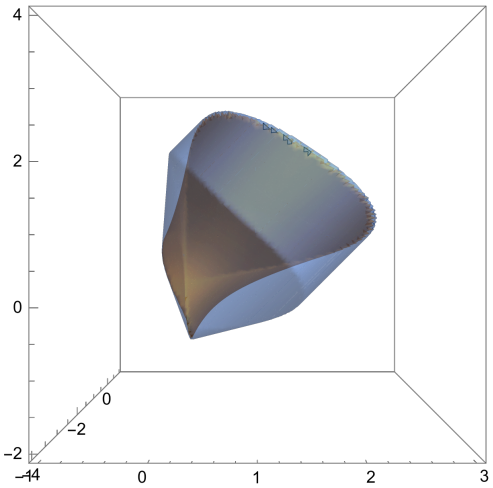
Example 5.4.1. We continue with the system of three quadratics as in Example 5.2.1. By construction, the nontrivial aggregation $(1, 0, 0)$ lies in the hyperbolicity cone of g and we see that $\overline{\text{conv}}(S)$ is a strict subset of the set defined by the intersection of good aggregations S_λ. This is demonstrated in Figure 5.5. However, if we modify the system of quadratics to be given by

$$\tilde{Q}_1 = (0.3Q_1 + 0.4Q_2 + 0.3Q_3), \quad \tilde{Q}_2 = Q_2, \quad \tilde{Q}_3 = Q_3,$$

then no nontrivial aggregation lies in the hyperbolicity cone of g and we are able to describe $\overline{\text{conv}}(S)$ as an intersection $\bigcap_{i=1}^3 S_{\lambda^{(i)}}$. This is shown in Figure 5.6 below.

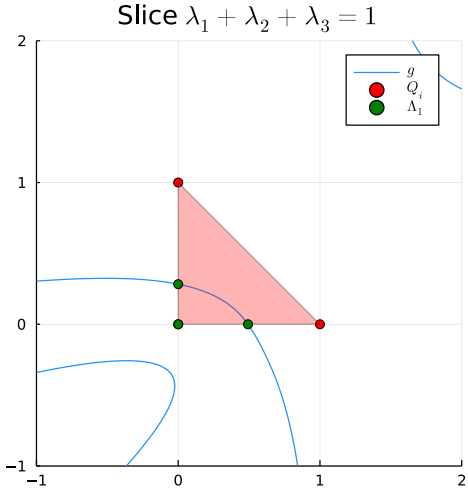


(a) Spectral Curve

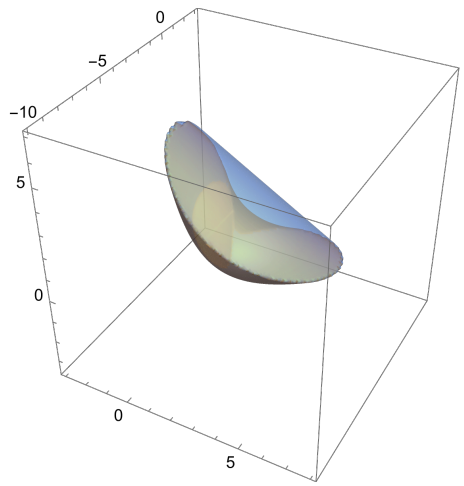


(b) Approximation

Figure 5.5: Plots corresponding to the system (Q_1, Q_2, Q_3) in Example 5.4.1. The set S in orange and approximation of $\overline{\text{conv}}(S)$ via aggregations in blue (Figure 5.5b) and the spectral curve and cone $\mathbb{R}^3_+$ (Figure 5.5a)



(a) Spectral Curve



(b) Approximation

Figure 5.6: Plots corresponding to the modified system $(\tilde{Q}_1, \tilde{Q}_2, \tilde{Q}_3)$ in Example 5.4.1. The set S in orange and approximation of $\overline{\text{conv}}(S)$ via aggregations in blue (Figure 5.6b) and the spectral curve and cone $\mathbb{R}^3_+$ (Figure 5.6a)

◇

Chapter 6

Conclusions and Future Directions

In this dissertation, we examined three problems at the intersection of algebraic and convex geometry and optimziation, using a broad range of tools from each. In Chapter 3, we presented a heuristic for regression problems using tropical algebraic structure and showed that this heuristic connected to underlying convex geometry. In Chapter 4, we studied semidefinite programs which were additionally compatible with a tensor structure. In doing so, we saw that there were additional connections to group theory. Finally, in Chapter 5, we provided a unified veiwpoint on a problem in real algebraic geometry and a problem in quadratically constrained optimization, connecting both problems to convex structures in "dual" objects.

Future Directions

There are many directions for extending the work presented in this dissertation. The area of tropical geometry for machine learning has been rapidly expanding in recent years. Future work could help to develop a better theoretical understanding of the convergence behavior of Algorithm 1. Specifically, one may be able to leverage the geometric structure of the loss to better understand the behavior of the iterates. Additionally, future work could augment the polynomial regression steps using the ideas in [Hoo19, TM19, TTM22] to develop variants of

Algorithm 1 for use with different norms or which enforce sparsity patterns or a regularization term. More generally, the development of a procedure for monomial selection remains open.

One potential application domain is in ReLU network initialization. In this work, we successfully initialized a ReLU network using a tropical rational function for a univariate regression task, while the tropical initialization was outperformed by existing initialization strategies for a bivariate regression task. This indicates the potential for future work to develop a better understanding of network initialization. In particular, the network architectures used in our experiments are limited, and a full understanding of correspondences between network architectures and tropical functions is currently an open problem; however some progress has been made recently, for example in [BLM24].

In the realm of tensor-tensor products and semidefinite programs, there is great potential for developing a bilevel optimization scheme in tensor completion. That is, optimize over the choice of M as well as the tensor completion. Such bilevel optimzation schemes have been developed in [NK24]. Since the semidefinite formuation presented in Chapter 4 is a parametrized convex problem, it may be possible to exploit this structure while optimizing over the choice of M.

In the study of systems of quadratics, there are multiple directions for future research. First, our analysis relied heavily on the fact that we were examining systems of three quadratic inequalities, so that the corresponding spectral object was an algebraic curve. An obvious next step is to study the shape of the spectral hypersurface in $\mathbb{RP}^{m-1}$ for systems of m quadratics. Here, we no longer expect the hypersurface to be hyperbolic, and PDLC is no longer a sufficient condition for the convex hull to be given by aggregations, so the problem is significantly more challenging. In an alternate direction, it is reasonable to expect to be able to leverage the dichotomy presented in Theorem 5.0.1 to construct upper bounds on the degree of sums of squares multipliers for positivstellensatz certificates for a quadratic forms on varieties defined by the complete intersection of two quadrics. Finally, the computational implementation and implications of the results in Chapter 5 remain open.

Bibliography

[ABMM18] Raman Arora, Amitabh Basu, Poorya Mianjy, and Anirbit Mukherjee. Understanding deep neural networks with rectified linear units. In *International Conference on Learning Representations (ICLR)*, 2018.

[Agr88] A. A. Agrachëv. Homology of the intersections of real quadrics. *Dokl. Akad. Nauk SSSR*, 299(5):1033–1036, 1988.

[AL12] A. Agrachev and A. Lerario. Systems of quadratic inequalities. *Proc. Lond. Math. Soc. (3)*, 105(3):622–660, 2012.

[Bar02] Alexander Barvinok. *A course in convexity*, volume 54 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2002.

[BBCV21] Michael M Bronstein, Joan Bruna, Taco Cohen, and Petar Veličković. Geometric deep learning: Grids, groups, graphs, geodesics, and gauges. *arXiv preprint arXiv:2104.13478*, 2021.

[BD25] Grigoriy Blekherman and Alex Dunbar. A topological approach to simple descriptions of convex hulls of sets defined by three quadrics. *SIAM J. Appl. Algebra Geom.*, 9(2):310–342, 2025.

[BDS24] Grigoriy Blekherman, Santanu S. Dey, and Shengding Sun. Aggregations of Quadratic Inequalities and Hidden Hyperplane Convexity. *SIAM J. Optim.*, 34(1):98–126, 2024.

[BLM24] Marie-Charlotte Brandenburg, Georg Loho, and Guido Montúfar. The real tropical geometry of neural networks. *arXiv preprint arXiv:2403.11871*, 2024.

[BPT13] Grigoriy Blekherman, Pablo A. Parrilo, and Rekha R. Thomas, editors. *Semidefinite optimization and convex algebraic geometry*, volume 13 of *MOS-SIAM Series on Optimization*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Optimization Society, Philadelphia, PA, 2013.

[BSV16] Grigoriy Blekherman, Gregory G. Smith, and Mauricio Velasco. Sums of squares and varieties of minimal degree. *J. Amer. Math. Soc.*, 29(3):893–913, 2016.

[BV04] Stephen P Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[CG79] Raymond Cuninghame-Green. *Minimax Algebra*, volume 166 of *Lecture Notes in Economics and Mathematical Systems*. Springer Berlin Heidelberg, Berlin, Heidelberg, 1979.

[CM19] Vasileios Charisopoulos and Petros Maragos. A tropical approach to neural networks with piecewise linear activations, 2019.

[DIK12] Alex Degtyarev, Ilia Itenberg, and Viatcheslav Kharlamov. On the number of components of a complete intersection of real quadrics. In *Perspectives in analysis, geometry, and topology*, volume 296 of *Progr. Math.*, pages 81–107. Birkhäuser/Springer, New York, 2012.

[DMnS22] Santanu S. Dey, Gonzalo Muñoz, and Felipe Serrano. On obtaining the convex hull of quadratic inequalities via aggregations. *SIAM J. Optim.*, 32(2):659–686, 2022.

[DN25] Alex Dunbar and Elizabeth Newman. Tensor-tensor products, group representations, and semidefinite programming. *In Preparation*, 2025.

[DR24] Alex Dunbar and Lars Ruthotto. Alternating minimization for regression with tropical rational functions. *Algebr. Stat.*, 15(1):85–111, 2024.

[EJRR23] Anas El Hachimi, Khalide Jbilou, Ahmed Ratnani, and Lothar Reichel. Spectral computation with third-order tensors using the t-product. *Applied Numerical Mathematics*, 193:1–21, 2023.

[FH91] William Fulton and Joe Harris. *Representation theory*, volume 129 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1991. A first course, Readings in Mathematics.

[Går59] Lars Gårding. An inequality for hyperbolic polynomials. *J. Math. Mech.*, 8:957–965, 1959.

[GP04] Karin Gatermann and Pablo A. Parrilo. Symmetry groups, semidefinite programs, and sums of squares. *J. Pure Appl. Algebra*, 192(1-3):95–128, 2004.

[Har59] Philip Hartman. On functions representable as a difference of convex functions. *Pacific J. Math.*, 9:707–713, 1959.

[Hat02] Allen Hatcher. *Algebraic topology*. Cambridge University Press, Cambridge, 2002.

[HHS21] Alexander Heaton, Serkan Hoşten, and Isabelle Shankar. Symmetry adapted Gram spectrahedra. *SIAM J. Appl. Algebra Geom.*, 5(1):140–164, 2021.

[Hil88] David Hilbert. Ueber die Darstellung definiter Formen als Summe von Formenquadraten. *Math. Ann.*, 32(3):342–350, 1888.

[Hoo19] James Hook. Max-plus linear inverse problems: 2-norm regression and system identification of max-plus linear dynamical systems with gaussian noise. *Linear Algebra and its Applications*, 579:1–31, 2019.

[HV07] J. William Helton and Victor Vinnikov. Linear matrix inequality representation of sets. *Communications on Pure and Applied Mathematics*, 60(5):654–674, 2007.

[HZRS15] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, December 2015.

[KB14] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization, 2014.

[KHAN21] Misha E. Kilmer, Lior Horesh, Haim Avron, and Elizabeth Newman. Tensor-tensor algebra for optimal representation and compression of multiway data. *Proc. Natl. Acad. Sci. USA*, 118(28):Paper No. e2015851118, 12, 2021.

[KKA15] Eric Kernfeld, Misha Kilmer, and Shuchin Aeron. Tensor-tensor products with invertible linear transforms. *Linear Algebra Appl.*, 485:545–570, 2015.

[KL21] Kody Kazda and Xiang Li. Nonconvex multivariate piecewise-linear fitting using the difference-of-convex representation. *Computers & Chemical Engineering*, 150:107310, July 2021.

[KLL21] Hao Kong, Canyi Lu, and Zhouchen Lin. Tensor Q-rank: new data dependent definition of tensor rank. *Mach. Learn.*, 110(7):1867–1900, 2021.

[KPV15] Mario Kummer, Daniel Plaumann, and Cynthia Vinzant. Hyperbolic polynomials, interlacers, and sums of squares. *Math. Program.*, 153(1):223–245, 2015.

[Las01] Jean B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM J. Optim.*, 11(3):796–817, 2000/01.

[LN23] Lek-Heng Lim and Bradley J. Nelson. What is... an equivariant neural network? *Notices Amer. Math. Soc.*, 70(4):619–625, 2023.

[Mar08] Murray Marshall. *Positive polynomials and sums of squares*, volume 146 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2008.

[MB09] Alessandro Magnani and Stephen P. Boyd. Convex piecewise-linear fitting. *Optimization and Engineering*, 10(1):1–17, March 2009.

[McC01] J. McCleary. *A User's Guide to Spectral Sequences*. A User's Guide to Spectral Sequences. Cambridge University Press, 2001.

[MCT21] Petros Maragos, Vasileios Charisopoulos, and Emmanouil Theodosis. Tropical geometry and machine learning. *Proceedings of the IEEE*, 109(5):728–755, 2021.

[MKY24] Hiroki Marumo, Sunyoung Kim, and Makoto Yamashita. T-semidefinite programming relaxation with third-order tensors for constrained polynomial optimization, 2024.

[MRZ22] Guido Montúfar, Yue Ren, and Leon Zhang. Sharp bounds for the number of regions of maxout networks and vertices of Minkowski sums. *SIAM J. Appl. Algebra Geom.*, 6(4):618–649, 2022.

[MS15] Diane Maclagan and Bernd Sturmfels. *Introduction to Tropical Geometry*, volume 161 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2015.

[MT19] Petros Maragos and Emmanouil Theodosis. Tropical geometry and piecewise-linear approximation of curves and surfaces on weighted lattices, 2019.

[MT20] Petros Maragos and Emmanouil Theodosis. Multivariate tropical regression and piecewise-linear surface fitting. In *ICASSP 2020-2020 IEEE International*

Conference on Acoustics, Speech and Signal Processing (ICASSP), pages 3822–3826. IEEE, 2020.

[NK24] Elizabeth Newman and Katherine Keegan. Optimal matrix-mimetic tensor algebras via variable projection. *arXiv preprint arXiv:2406.06942*, 2024.

[Par03] Pablo A. Parrilo. Semidefinite programming relaxations for semialgebraic problems. volume 96, pages 293–320. 2003. Algebraic and geometric methods in discrete optimization.

[Pow21] Victoria Powers. *Certificates of positivity for real polynomials—theory, practice, and applications*, volume 69 of *Developments in Mathematics*. Springer, Cham, [2021] ©2021.

[PSV11] Daniel Plaumann, Bernd Sturmfels, and Cynthia Vinzant. Quartic curves and their bitangents. *J. Symbolic Comput.*, 46(6):712–733, 2011.

[PSV12] Daniel Plaumann, Bernd Sturmfels, and Cynthia Vinzant. Computing linear matrix representations of Helton-Vinnikov curves. In *Mathematical methods in systems, optimization, and control*, volume 222 of *Oper. Theory Adv. Appl.*, pages 259–277. Birkhäuser/Springer Basel AG, Basel, 2012.

[PT07] Imre Pólik and Tamás Terlaky. A survey of the s-lemma. *SIAM Review*, 49(3):371–418, 2007.

[QL17] Liqun Qi and Ziyan Luo. *Tensor Analysis*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 2017.

[RFP10] Benjamin Recht, Maryam Fazel, and Pablo A. Parrilo. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM Rev.*, 52(3):471–501, 2010.

[RK15] Steffen Rebennack and Josef Kallrath. Continuous piecewise linear delta-approximations for bivariate and multivariate functions. *Journal of Optimization Theory and Applications*, 167(1):102–117, 2015.

[RTAL13] Cordian Riener, Thorsten Theobald, Lina Jansson Andrén, and Jean B. Lasserre. Exploiting symmetries in SDP-relaxations for polynomial optimization. *Math. Oper. Res.*, 38(1):122–141, 2013.

[SM19] Georgios Smyrnis and Petros Maragos. Tropical polynomial division and neural networks. *CoRR*, abs/1911.12922, 2019.

[SM20] Georgios Smyrnis and Petros Maragos. Multiclass neural network minimization via tropical newton polytope approximation. In *International Conference on Machine Learning*, pages 9068–9077. PMLR, 2020.

[Tho76] Robert C Thompson. The characteristic polynomial of a principal subpencil of a hermitian matrix pencil. *Linear Algebra and its Applications*, 14(2):135–177, 1976.

[Tho91] Robert C Thompson. Pencils of complex and real symmetric and skew matrices. *Linear Algebra and its Applications*, 147:323–371, 1991.

[TM19] Anastasios Tsiamis and Petros Maragos. Sparsity in max-plus algebra and systems. *Discrete Event Dynamic Systems*, 29(2):163–189, 2019.

[TPS21] Martin Trimmel, Henning Petzka, and Cristian Sminchisescu. Tropex: An algorithm for extracting linear terms in deep neural networks. In *International Conference on Learning Representations*, 2021.

[TTM22] Nikolaos Tsilivis, Anastasios Tsiamis, and Petros Maragos. Toward a sparsity theory on weighted lattices. *Journal of Mathematical Imaging and Vision*, pages 1–13, 2022.

[TV12] Alejandro Toriello and Juan Pablo Vielma. Fitting piecewise linear continuous functions. *European Journal of Operational Research*, 219(1):86–95, May 2012.

[Vin93] Victor Vinnikov. Selfadjoint determinantal representations of real plane curves. *Math. Ann.*, 296(3):453–479, 1993.

[Yil09] Uğur Yildiran. Convex hull of two quadratic constraints is an LMI set. *IMA J. Math. Control Inform.*, 26(4):417–450, 2009.

[ZHH22] Meng-Meng Zheng, Zheng-Hai Huang, and Sheng-Long Hu. Unconstrained minimization of block-circulant polynomials via semidefinite program in third-order tensor space. *J. Global Optim.*, 84(2):415–440, 2022.

[ZHW21] Meng-Meng Zheng, Zheng-Hai Huang, and Yong Wang. T-positive semidefiniteness of third-order symmetric tensors and T-semidefinite programming. *Comput. Optim. Appl.*, 78(1):239–272, 2021.

[ZNL18] Liwen Zhang, Gregory Naitzat, and Lek-Heng Lim. Tropical geometry of deep neural networks. In Jennifer Dy and Andreas Krause, editors, *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pages 5824–5832. PMLR, 10–15 Jul 2018.