

# Real Algebraic Geometry for Physics and Optimization

Der Fakultät für Mathematik und Informatik  
der Universität Leipzig  
angenommene

D I S S E R T A T I O N

zur Erlangung des akademischen Grades

DOCTOR RERUM NATURALIUM  
(Dr.rer.nat.)

im Fachgebiet

Mathematik

Vorgelegt

von Spec. Dmitrii Pavlov  
geboren am 05.08.1998 in Chabarowsk (Russland)

Die Annahme der Dissertation wurde empfohlen von:

1. Prof. Dr. Bernd Sturmfels (Max Planck Institute for Mathematics in the Sciences)
2. Prof. Dr. Thorsten Theobald (Goethe Universität Frankfurt am Main)
3. Dr. Matteo Parisi (Harvard University)

Die Verleihung des akademischen Grades erfolgt mit Bestehen der Verteidigung am 15.08.2024 mit dem Gesamtprädikat magna cum laude.



# Abstract

In recent years, algebraic geometry (both complex and real) has proven to be useful in numerous applications in optimization, statistics, quantum information, and physics. In this thesis, we concentrate on studying semi-algebraic sets and varieties defined over the real numbers that arise in these applied contexts.

We begin with the study of Gibbs manifolds and Gibbs varieties. Gibbs manifolds are images of affine spaces of symmetric matrices under the matrix exponential map. They appear naturally in the context of entropic regularization for semidefinite programming or entropy maximization in quantum information theory and statistical physics. The Gibbs variety is the zero locus of all polynomials that vanish on the Gibbs manifold. We compute these polynomials and show that the Gibbs variety is low-dimensional. More precisely, we give symbolic and numerical algorithms for implicitizing the Gibbs variety of a given affine space of symmetric matrices that can be implemented in computer algebra systems. We show that the dimension of the Gibbs variety is bounded by a linear function in the dimension of the given space of matrices and the size of matrices. We give an exact formula for this dimension, and an upper bound for the degree of the Gibbs variety. We apply our theory to a range of scenarios: matrix pencils, quantum optimal transport, and sparse matrices. For matrix pencils, we give a concrete description of the polynomial equations that vanish on their Gibbs variety. For the quantum optimal transport problem, we show that the corresponding Gibbs manifold is in fact a semi-algebraic set. For sparse matrices, we explore connections to graph theory and prove a formula for the dimension of Gibbs varieties of linear spaces of symmetric matrices defined by trees.

The role of Gibbs manifolds in quantum information theory leads us to consider the notion of quantum conditional independence from an algebraic perspective. We take inspiration from algebraic statistics, where graphical models encoding conditional independence relations can be described as intersections of an algebraic variety with the probability simplex, and study quantum counterparts of such models. These are families of quantum states satisfying quantum conditional independence conditions encoded by a graph. We present several ways to associate an algebraic variety to such a model. The first one is based on the notion of quantum conditional mutual information, and the corresponding variety is called the QCMI variety. The second one arises from the Petz recovery map, resulting in the Petz variety. Finally, the third one comes from Gibbs varieties of linear spaces of Hamiltonians. We study basic properties of these varieties and provide algorithms to compute their defining equations. For instance, we show that the Petz variety is irreducible, and that the construction of the Gibbs variety extends from linear spaces of symmetric matrices to unirational varieties thereof. We also study toric varieties defined by commuting Hamiltonians arising from a graph in the context of stabilizer codes. We give an efficient algorithm to compute the defining equations of such a toric variety. Moreover, we investigate a quantum analog of maximum likelihood

estimation for quantum exponential families, the so-called quantum information projection. Our main result here is a quantum analog of Birch's theorem from algebraic statistics, which allows to compute quantum information projections by using algebraic methods.

We continue with studying (semi-)algebraic geometry of minimizing dual volumes of polytopes. Similarly to Gibbs manifolds, this appears naturally in the context of regularization of convex optimization problems. The interior point of a convex polytope that leads to a polar dual of minimal volume is called the Santaló point. When translating the facet hyperplanes, the Santaló point traces out a semi-algebraic set called the Santaló patchwork. We describe and compute this set using algebraic and numerical techniques. In particular, we show that its dimension is equal to that of the original polytope and give its defining equations and inequalities. We then investigate several naturally defined algebraic varieties containing the Santaló patchwork. One of them arises from the incidence variety of a statistical model called Wachspress model. We connect the maximum likelihood degree of this model to the degree of the Zariski closure of the Santaló patchwork. We continue by treating the question of computing the Santaló points of polytopes numerically, by using homotopy continuation techniques. We also explore connections to physics, where the dual volume function gives the canonical form of a polytope. This is relevant for calculating scattering amplitudes in certain quantum field theories.

Finally, we study Grasstopes. This is yet another class of semi-algebraic sets inspired by physics. These are linear projections of the positive Grassmannian  $\text{Gr}_{\geq 0}(k, n)$  to  $\text{Gr}(k, k+m)$ . When the linear projection is given by a totally positive matrix  $Z$ , we recover the definition of the amplituhedron, a semi-algebraic set that computes scattering amplitudes in the  $\mathcal{N} = 4$  SYM quantum field theory. We divide Grasstopes into three categories (tame, wild and rational) based on the properties of the matrix  $Z$  defining the linear projection. We concentrate on the case  $m = 1$ , when the image lives in the projective space, and give a combinatorial characterization of such Grasstopes in terms of sign flips, extending the results of Karp and Williams for the amplituhedron. That is, we show that  $m = 1$  Grasstopes consist of regions of the oriented hyperplane arrangement defined by the matrix  $Z$  that are labeled by sign vectors with sufficiently many sign changes. We also suggest a notion of a Grasstope coming from an arbitrary (not necessarily realizable) oriented matroid.



## Authorship

The results presented in this thesis are based on five articles of my mine, four of which were written in collaboration with other researchers. Parts of this thesis are republished.

Chapter 1 is written by myself. It contains known definitions and results from the literature. There are no original contributions in this chapter.

Chapter 2 is based on two works. The first article [PST23] is joint work with Bernd Sturmfels and Simon Telen. All coauthors equally contributed to all aspects of the publication. I made major contributions to all results presented in Sections 2.1, 2.2, and 2.3, while my coauthors explored the connection to optimization (Sections 2.4 and 2.5). The article appeared in *Information Geometry*. Passages from this article are taken over to all sections of Chapter 2.

The second article [Pav23] is written by myself. Its results are distributed over Sections 2.1, 2.2, and 2.6. These sections contain passages from this article. This article was submitted to *Beiträge zur Algebra und Geometrie (Contributions to Algebra and Geometry)* and is under review.

Chapter 3 is based on the article [DPW23] written together with Eliana Duarte and Maximilian Wiesmann. All authors contributed equally to all aspects of the publication and mathematical results presented in the article. The article was submitted to *Advances in Applied Mathematics* and is currently under review. Passages from this article are carried over to all sections of Chapter 3.

Chapter 4 is based on joint work [PT24] with Simon Telen. My coauthor and I equally contributed to all aspects of the publication. I primarily focused on the aspects of semi-algebraic geometry (Sections 4.2, 4.3 and 4.4), while my coauthor explored the relation to algebraic statistics and numerical algorithms (Sections 4.5 and 4.6). The article was submitted to *SIAM Journal on Applied Algebra and Geometry* and is currently under review. Passages from this article are carried over to all sections of Chapter 4.

Chapter 5 is based on the article [MPP23], which is the result of a collaboration with Yelena Mandelshtam and Elizabeth Pratt. All coauthors contributed equally to all aspects of the publication and mathematical results. The article was submitted to *Combinatorial Theory* and is currently under review. Passages from this article are carried over to all sections of Chapter 5.



## Acknowledgements

Throughout my time in Leipzig I have often heard from more senior PhD students that writing the acknowledgements is the hardest part of writing a thesis. Now that I have gotten to writing one myself, I find that this is truly so. It is a very challenging task, for there are so many people to be thankful to, and just one page available to express my gratitude. I will nevertheless try to do my best here.

Firstly, I would like to thank my advisors, Bernd Sturmfels and Simon Telen. Bernd, as trite as it may sound, thank you for all the advice. And for encouragement, inspiration and ideas that have made my mathematical journey here so pleasant. Simon, thank you for your endless energy, optimism and support. Doing math can sometimes be stressful but it was never when I worked with you. A special thanks goes to my masters advisor, Gleb Pogudin. None of this would be possible without you.

This thesis is the result of my PhD studies at the Max Planck Institute for Mathematics in the Sciences in Leipzig. And I would like to thank everyone here for making this place so welcoming, lively and enjoyable. Barbara, Laura, Leonie, and Max, my dear fellow PhD students and friends, thanks for everything really, for all the fun things we did together, for your help with small things in life and for cheering me up every now and then. I could always count on you. Ben, Sachi, and Raul, thanks for the fun first year, and for reminding me about the importance of work-life balance. Sam and Sachi, thanks for the excellent Monday cake baking experiences. Mirke and Saskia, thanks a lot for all the administrative help. The famous German bureaucracy looks much less scary thanks to you. And a separate thanks to Mirke for the wonderful German classes; this was an absolute delight.

In my view, mathematics and mathematicians thrive on collaboration and sharing ideas. I would like to thank all of my collaborators, past and present, for the great research experiences I had: Bernd Sturmfels, Simon Telen, Eliana Duarte, Max Wiesmann, Yelena Mandelshtam, Lizzie Pratt, Sebastian Falkensteiner, Rafael Sendra, and Kristian Ranestad, this is for you.

Finally, a huge thank you to my family and to my friends back in Russia. Мама, Мама, Егор, Настя, Люба, Лена, Аня и Лиза, спасибо! In the realities of the world we live in, we had to meet mostly online and way less often than I would have liked. But seeing you every now and then gave me the energy and strength to finish this endeavour.





*In memory of my father.*

# Contents

<b>Introduction</b>	<b>15</b>
<b>1. Background</b>	<b>18</b>
1.1. Algebraic geometry . . . . .	18
1.1.1. Varieties and ideals . . . . .	18
1.1.2. Semi-algebraic sets . . . . .	21
1.2. Quantum information theory . . . . .	22
1.3. Discrete geometry . . . . .	23
1.3.1. Polyhedra, polytopes, and cones . . . . .	23
1.3.2. Subdivisions, triangulations, and chambers . . . . .	24
1.4. Convex optimization . . . . .	26
1.5. Algebraic combinatorics . . . . .	28
1.5.1. Hyperplane arrangements . . . . .	28
1.5.2. Oriented matroids . . . . .	29
<b>2. Gibbs Manifolds</b>	<b>32</b>
2.1. What are Gibbs manifolds and varieties? . . . . .	34
2.2. Implicitization of Gibbs varieties . . . . .	38
2.3. Pencils of quadrics . . . . .	42
2.4. Role in convex optimization . . . . .	46
2.5. Quantum optimal transport . . . . .	50
2.6. Logarithmic sparsity . . . . .	53
<b>3. Quantum Graphical Models</b>	<b>60</b>
3.1. Quantum graphical models on trees . . . . .	61
3.1.1. Quantum conditional mutual information varieties . . . . .	61
3.1.2. Petz varieties . . . . .	66
3.2. Quantum graphical models from Gibbs manifolds . . . . .	68
3.2.1. Gibbs varieties of linear systems of Hamiltonians . . . . .	68
3.2.2. Gibbs varieties of unirational varieties of Hamiltonians . . . . .	69
3.3. Toric varieties from quantum exponential families associated to graphs . . . . .	70
3.3.1. Commuting Hamiltonians from graphs . . . . .	70
3.3.2. Quantum information projections . . . . .	73

---

3.4. Stabilizer formalism . . . . .	75
<b>4. Minimizing Dual Volumes of Polytopes</b>	<b>77</b>
4.1. The Santaló point . . . . .	77
4.2. Dual volumes of polytopes . . . . .	80
4.3. The Santaló patchwork . . . . .	84
4.4. Patch varieties . . . . .	88
4.5. Wachspress models . . . . .	90
4.6. Computing Santaló points . . . . .	94
4.6.1. From likelihood equations to dual volume . . . . .	95
4.6.2. Tracking paths on Santaló patches . . . . .	97
<b>5. Grasstopes</b>	<b>99</b>
5.1. What is a Grasstope? . . . . .	100
5.2. Grasstopes for $m = 1$ : tame, wild, and rational . . . . .	103
5.3. Examples . . . . .	105
5.4. Extremal counts and oriented matroid Grasstopes . . . . .	108
<b>Bibliography</b>	<b>111</b>

# List of Figures

1.1.	Chamber complex of a pentagon in $\mathbb{R}^2$ . . . . .	25
1.2.	The ellipsope $\mathcal{E}_3$ . . . . .	27
1.3.	An oriented hyperplane arrangement and sign labels of its regions . . . . .	29
2.1.	Gibbs manifolds $\text{GM}(\mathcal{L}_\epsilon)$ and the spectrahedron from Example 2.4.2 . . . . .	47
3.1.	Left: the graph $G$ . Right: illustration of the Hamiltonian $H_1$ . . . . .	71
4.1.	Left: the pentagon $P$ from Example 4.1.1, together with its adjoint curve (red) and facet hyperplanes (blue). Right: a two-dimensional slice of the chamber complex $\mathcal{C}_A$ . . . . .	78
4.2.	Visualization of the Santaló patchwork for $A$ from (4.1.6). . . . .	79
4.3.	The Santaló patchwork (left) and chamber complex (right) from Example 4.3.4. . . . .	86
4.4.	Comparing the Santaló patchwork (green) and the Gibbs manifold (red). . . . .	87
4.5.	The Santaló patchwork (left) and chamber complex (right) from Example 4.3.5. . . . .	87
5.1.	Affine chart in which the tame Grasstope is bounded. The six lines corresponding to the rows of $Z$ are colored red, orange, yellow, green, and blue, in order, with orientations given by arrows. The shaded portion of the figure is the Grasstope, which consists exactly of the regions with at least two sign changes. . . . .	106
5.2.	The six lines corresponding to the rows of $Z$ are pictured as described, with orientations given by the arrows. The regions can then be labelled by sign patterns. The shaded portion of the figure is the Grasstope, and it consists exactly of those regions with at least two sign changes. . . . .	107
5.3.	The six lines corresponding to the rows of $Z$ are pictured as described, with orientations given by the arrows. The regions can then be labelled by sign patterns. The shaded portion of the figure is the Grasstope, and it consists exactly of those regions with at least two sign changes. In this case, the shaded region is a Möbius strip. . . . .	108
5.4.	The Grasstope of a totally positive matrix with two rows negated. The six lines are cyclically ordered with orientations indicated by arrows. Every region has at least two sign changes, so the Grasstope is all of $\mathbb{RP}^2$ . . . . .	111

# List of Tables

- 5.1. Minimal and maximal possible number of regions in a Grasstope. . . . . 110
- 5.2. Maximal number of regions from reorienting and reordering a positive matrix. 110

# Introduction

Classical algebraic geometry is devoted to the study of algebraic varieties. These are sets of solutions to systems of polynomial equations. Its methods, which originate in commutative algebra, work best when the varieties of interest are defined over the field of complex numbers (that is, we are looking for complex solutions to systems of polynomials with complex coefficients) or other algebraically closed fields. Many modern real-life problems, however, require us to study *real* solutions of polynomial systems. When passing to working with polynomials defined over the field of real numbers, one has to keep in mind that this field has a natural ordering. Therefore, it makes sense to study not just polynomial equations but also polynomial inequalities. Sets defined by systems of polynomial equations and inequalities are called semi-algebraic, and are central objects in the field of real algebraic geometry.

In recent years, algebraic geometry (both real and complex) has proven to be useful in numerous applications in optimization, statistics, quantum information and physics. In this thesis, we study semi-algebraic sets and varieties defined over the real numbers that arise in these applied contexts. This is done by using methods from real and computational algebraic geometry, as well as nonlinear algebra [MS21a]. Original research results are presented in four chapters (Chapters 2–5), and Chapter 1 gives necessary background notions in order to make this thesis as self-contained as possible. Our work relies heavily on computer algebra software, and the supplementary code we used in our research is available at <https://mathrepo.mis.mpg.de>, an online research data repository of the Max Planck Institute for Mathematics in the Sciences.

We now give a description of the results of this thesis. In Chapter 2, we introduce Gibbs manifolds and Gibbs varieties. A Gibbs manifold is the image of an affine space of real symmetric matrices under the matrix exponential map, and the corresponding Gibbs variety is its Zariski closure. The names of these objects originate in physics: if one starts from a linear space of Hamiltonians, the corresponding Gibbs manifold parametrizes Gibbs states (or thermal states) associated to the elements of this linear space [Alh23, Vig99]. Our motivation to study Gibbs manifolds from an algebraic point of view, however, comes from a different angle, namely, from regularization techniques in convex optimization. In linear programming one optimizes a linear function over a convex polytope. In the entropic regularization of linear programming [STVvR24] an important role is played by the positive part of the toric variety associated to this polytope via the moment map. A natural generalization of linear programming is semidefinite programming [BPT12], in which one optimizes a linear function over an affine section of the cone of real positive semidefinite matrices. Entropic regularization techniques are also available here, as described in Section 2.4. In this context the role of the positive part of the toric variety is played precisely by the Gibbs manifold of the linear span of symmetric matrices defining the affine section. There is also an analog of the moment map. From this point of view, Gibbs manifolds and varieties provide a natural noncommutative

extension of toric geometry. In some special cases, Gibbs manifolds are semi-algebraic sets. This is true, for instance, when the linear space we start from consists of pairwise commuting matrices (Theorem 2.1.13). In general, however, this is not the case. In particular, the dimension of the Gibbs variety can be higher than that of the corresponding Gibbs manifold. We give a simple linear bound for the dimension of the Gibbs variety in terms of the dimension of the linear space and the size of matrices we consider (Theorem 2.1.6). This bound implies that Gibbs varieties are, in an appropriate sense, low-dimensional. We also give a somewhat more intricate exact formula for the dimension of the Gibbs variety, which relies on particular eigenvalue properties of the linear space (Theorem 2.1.9), and bound the degree of this variety from above (Proposition 2.1.10). We present symbolic and numerical algorithms (Algorithms 1 and 2) to compute its defining equations. In addition, we elaborate on the connection to optimization. The theory developed in this chapter is applied to several contexts, such as pencils of quadrics (Section 2.3), quantum optimal transport (Section 2.5), and spaces of sparse matrices (Section 2.6). This leads to applications in quantum information and high-dimensional statistics.

As mentioned above, in quantum physics and quantum information theory Gibbs manifolds parametrize thermal states. These are states of maximal entropy associated to families of Hamiltonians. They can be thought of as quantum analogs of probability distributions satisfying conditional independence statements [PH11]. This observation led us to a journey of investigating algebraic approaches to the notion of quantum conditional independence, and the results are presented in Chapter 3. In algebraic statistics [Sul18], graphical models [Lau96] conveniently describe families of probability distributions satisfying conditional independence constraints given by a graph. Methods of algebraic geometry have been used with great success to understand the structure of these models, which turn out to be intersections of an algebraic variety with the probability simplex. In this spirit, we introduce quantum graphical models that parametrize quantum states satisfying certain independence constraints and give three ways to associate algebraic varieties to such a model. The first one is the already familiar Gibbs variety (Section 3.2). The second one is based on the notion of quantum conditional mutual information (Section 3.1.1). Finally, the third one comes from the Petz recovery map [Pet86] (Section 3.1.2). We study fundamental properties of these varieties and present algorithms to obtain their defining equations. Classical graphical models are obtained from these constructions by restricting to quantum states represented by diagonal matrices.

An important statistical problem of maximal likelihood estimation also has a natural quantum analog. This is given by the quantum information projection. We study quantum information projections to quantum exponential families defined by graphs that arise in the context of stabilizer codes, and prove a generalization of Birch's Theorem (Theorem 3.3.10).

Chapter 4, just like Chapter 2, is motivated by regularization problems in convex optimization. The problem we are interested in is linear programming and the regularizer is the universal barrier. This is the function giving the volume of the polar dual of the feasible polytope with respect to each of its interior points. The unique minimizer of this dual volume function on the interior of the polytope is called the Santaló point, and as the facet hyperplanes of the polytope are translated, this point traces out a set called the Santaló patchwork. We show that this set is semi-algebraic and compute its defining equations and inequalities. Moreover, we show that it is a union of basic semi-algebraic sets that are in bijection with the cells of the chamber complex of a point configuration (Corollary 4.3.3). We study Zariski closures of these sets, and provide results bounding their degree (Proposition 4.4.6, Corollary 4.5.5). We explore connections to algebraic statistics via Wachspress models [KSS20] (Section 4.5). This chapter concludes with numerical algorithms for computing Santaló points. These algorithms rely on homotopy continuation techniques. Another similarity with Chapter 2 is



that the dual volume function has a physical interpretation. It is the canonical function of the polytope as a positive geometry [AHBL17]. For special polytopes, this function computes scattering amplitudes in certain quantum field theories [AHBHY18] and string amplitudes in string theory [AHHL21].

Chapter 5, the final chapter of this thesis, is inspired by physics. One of the main recent breakthroughs in particle physics is the realization that certain scattering amplitudes (that is, functions describing the probability that a certain interaction of elementary particles occurs) can be computed as volumes of certain semi-algebraic sets defined by positivity constraints. This was first discovered for  $\mathcal{N} = 4$  Super Yang-Mills quantum field theory, and the relevant semi-algebraic set is the amplituhedron [AHT14]. The amplituhedron is the image of the positive Grassmannian  $\text{Gr}_{\geq 0}(k, n)$  under a projection to the Grassmannian  $\text{Gr}(k, k + m)$  defined by a totally positive matrix  $Z$ . If one drops the total positivity constraint on  $Z$ , one obtains the definition of a Grastope, a natural generalization of the amplituhedron. We study Grastopes for  $m = 1$  and provide a combinatorial characterization of these objects in terms of the hyperplane arrangement defined by  $Z$  and sign flips (Theorem 5.2.1 and Proposition 5.2.3). This generalizes the results of [KW19] for the amplituhedron. Finally, we suggest an abstract notion of a Grastope arising from an arbitrary oriented matroid (Section 5.4). This work is among several recently appeared *invitations to positive geometry* (cf. [Lam22]), an emerging area of mathematics on the intersection of combinatorics and algebraic geometry that studies semi-algebraic sets defined by positivity conditions. For other invitations, see [EGP<sup>+</sup>23, RST24].

# Chapter 1

## Background

In this chapter we present the notions that are necessary to develop the following chapters of this thesis. This includes basics of complex and real algebraic geometry, discrete geometry, combinatorics, convex optimization, and quantum information theory.

### 1.1. Algebraic geometry

We begin with reviewing basic notions of complex and real algebraic geometry: varieties, ideals and semi-algebraic sets. We mostly follow the exposition of [MS21a]. Since this thesis is devoted to applications of real algebraic geometry, the notions we introduce in this section will play a crucial role in every chapter. For more details on basics of complex and real algebraic geometry, we refer to [Sha13] and [BCR13] respectively.

#### 1.1.1. Varieties and ideals

We write  $\mathbb{K}[x_1, \dots, x_n]$  for the polynomial ring in the variables  $x_1, \dots, x_n$  over a field  $\mathbb{K}$ . Generally speaking, algebraic varieties are zero sets of systems of polynomials, and they are given by ideals of the relevant polynomial ring. We therefore start with defining ideals.

**Definition 1.1.1** (Ideal). Let  $R$  be a commutative ring. A nonempty subset  $I \subseteq R$  is called an *ideal* of  $R$  if for any  $a, b \in I$  we have  $a + b \in I$  and for any  $a \in I$  and  $b \in R$  we have  $ab \in I$ .

**Definition 1.1.2** (Prime ideal). An ideal  $I \subset R$  is called *prime* if  $ab \in I$  implies  $a \in I$  or  $b \in I$  for all  $a, b \in R$ .

We are now ready to define affine and projective varieties.

**Definition 1.1.3** (Affine variety). A (complex) *affine variety*  $\mathcal{V}(f_1, \dots, f_k) \subseteq \mathbb{C}^n$  is the zero set of finitely many polynomials  $f_1, \dots, f_k \in \mathbb{C}[x_1, \dots, x_n]$ .

As one could guess from the name, projective varieties live in the projective space  $\mathbb{P}^n$ . Points in this space are given by homogeneous coordinates  $(x_0 : x_1 : \dots : x_n)$ , and these are defined up to multiplication by a common scalar. This last fact means that the value of a polynomial  $f \in \mathbb{C}[x_0, \dots, x_n]$  at a point  $x \in \mathbb{P}^n$  is not well-defined. However, if  $f$  is homogeneous (that is, all of its monomials have the same degree), then we have  $f(x_0, \dots, x_n) = 0$  if and only if  $f(\lambda x_0, \dots, \lambda x_n) = 0$  for any  $\lambda \in \mathbb{C} \setminus \{0\}$ . Therefore, the notion of vanishing of a homogeneous polynomial at a point in the projective space is well-defined. This allows us to define projective varieties similarly to the affine case.

**Definition 1.1.4** (Projective variety). A (complex) projective variety  $\mathcal{V}(f_1, \dots, f_k) \subseteq \mathbb{P}^n$  is the zero set of finitely many homogeneous polynomials  $f_1, \dots, f_k \in \mathbb{C}[x_0, \dots, x_n]$ .

It is straightforward to check that the set of all polynomials that vanish on a given variety (be it affine or projective) is an ideal of the corresponding polynomial ring. We will write  $\mathcal{V}(I)$  for the variety defined by the ideal  $I$  and  $I_{\mathcal{V}}$  for the ideal defining the variety  $\mathcal{V}$ . Note that although we defined our varieties over  $\mathbb{C}$ , an analogous construction holds over any field (in particular, over  $\mathbb{R}$ ). Hilbert's basis theorem [MS21a, Corollary 1.14] states that a polynomial ring over a field is Noetherian, that is, every ideal is generated by a finite number of elements. Thus, every ideal  $I$  in  $\mathbb{K}[x_1, \dots, x_n]$  defines a variety.

**Remark 1.1.5.** Although for convenience of exposition we have chosen  $\mathcal{V}$  as the standard notation for a variety in this section, in the following chapters we typically denote varieties by capital latin letters (mainly  $X$  and  $Y$ ), as is customary in algebraic geometry.

We turn  $\mathbb{C}^n$  and  $\mathbb{P}^n$  into topological spaces by endowing them with the *Zariski topology*. The closed sets in this topology are exactly varieties. This induces a topology on every variety, in which the closed sets are its subvarieties (that is, subsets that are varieties themselves).

In algebraic geometry one often studies maps between algebraic varieties. Two important classes of these are regular and rational maps. A map between two affine varieties  $\mathcal{V}$  and  $\mathcal{W}$  is called *regular* if it is given by polynomials. A map between two (affine or projective) varieties is called *rational* if it is given by rational functions. An important property of rational maps is that they are not defined everywhere on  $\mathcal{V}$ : denominators of the rational functions giving the map define a subvariety of  $\mathcal{V}$  on which the map is not well-defined. We will denote regular maps by solid arrows ( $\mathcal{V} \rightarrow \mathcal{W}$ ) and rational maps by dashed arrows ( $\mathcal{V} \dashrightarrow \mathcal{W}$ ). Finally, we say that a map is *dominant* if its image is Zariski dense in  $\mathcal{W}$ .

We will often consider projections of affine varieties to coordinate subspaces of  $\mathbb{C}^n$ . Such a projection is a particularly simple regular map. If  $\mathcal{V}$  lives in  $\mathbb{C}^n$  with coordinates  $x_1, \dots, x_n$ , the projection  $\pi_d$  to the subspace corresponding to the first  $d$  coordinates is given by  $(x_1, \dots, x_n) \mapsto (x_1, \dots, x_d)$ . The projected variety is the Zariski closure  $\overline{\pi_d(\mathcal{V})}$  ( $\pi_d(\mathcal{V})$  itself might not be closed). In the language of ideals,  $I_{\overline{\pi_d(\mathcal{V})}}$  is obtained from  $I_{\mathcal{V}}$  by *elimination of the variables*  $x_{d+1}, \dots, x_n$ . That is,  $I_{\overline{\pi_d(\mathcal{V})}} = I_{\mathcal{V}} \cap \mathbb{C}[x_1, \dots, x_d]$  [MS21a, Theorem 4.2].

The two most fundamental invariants of a variety are its *dimension* and *degree*. Even though these two numbers have an intuitive meaning (the dimension answers the question “how big is the variety?”, and the degree tells us “how curvy it is?”), the formal definition requires some work. We shall now present this.

Our first notion on the way to dimension and degree is somewhat self-explanatory. We say that an ideal  $I \subseteq \mathbb{C}[x_1, \dots, x_n]$  is *monomial* if it is generated by monomials, i.e. products of variables. The set of all monomials in  $\mathbb{C}[x_1, \dots, x_n]$  is naturally identified with the set  $\mathbb{Z}_{\geq 0}^n$  of  $n$ -tuples of nonnegative integers:  $\mathbf{a} = (a_1, \dots, a_n) \in \mathbb{Z}_{\geq 0}^n$  corresponds to  $\mathbf{x}^{\mathbf{a}} = x_1^{a_1} \dots x_n^{a_n}$ . Now consider a total ordering  $\prec$  of  $\mathbb{Z}_{\geq 0}^n$ . We say that  $\prec$  is a *monomial order* if for all  $\mathbf{a}, \mathbf{b}$  and  $\mathbf{c} \in \mathbb{Z}_{\geq 0}^n$  firstly, we have  $(0, \dots, 0) \preceq \mathbf{a}$  and, secondly,  $\mathbf{a} \preceq \mathbf{b}$  implies  $\mathbf{a} + \mathbf{c} \preceq \mathbf{b} + \mathbf{c}$ . Here  $\mathbf{a} \preceq \mathbf{b}$  if  $\mathbf{a} \prec \mathbf{b}$  or  $\mathbf{a} = \mathbf{b}$ . Finally, we say that a monomial order  $\prec$  is *degree-compatible* if for all  $\mathbf{a}, \mathbf{b} \in \mathbb{Z}_{\geq 0}^n$  the inequality  $|\mathbf{a}| < |\mathbf{b}|$  implies  $\mathbf{a} \prec \mathbf{b}$ . Here  $|\mathbf{a}|$  is the sum of all entries of  $\mathbf{a}$ .

Now, given a monomial order  $\prec$  and a nonzero polynomial  $f \in \mathbb{C}[x_1, \dots, x_n]$ , we define the *initial monomial*  $\text{in}_{\prec}(f)$  to be the largest monomial  $\mathbf{x}^{\mathbf{a}}$  w.r.t.  $\prec$  among those that appear in  $f$  with a non-zero coefficient. The *initial ideal*  $\text{in}_{\prec}(I)$  of an ideal  $I \subseteq \mathbb{C}[x_1, \dots, x_n]$  is then naturally defined to be the monomial ideal generated by the initial monomials of all non-zero elements of  $I$ .

Our next step is to define the *Hilbert polynomial* of a monomial ideal. First, we define

the *Hilbert function* and the *Hilbert series*.

**Definition 1.1.6** (Hilbert function and series). For a monomial ideal  $I \subseteq \mathbb{C}[x_1, \dots, x_n]$  the *Hilbert function*  $h_I$  is a map  $\mathbb{Z}_{\geq 0} \rightarrow \mathbb{Z}_{\geq 0}$  defined as follows: for  $q \in \mathbb{Z}_{\geq 0}$  the value  $h_I(q)$  is the number of monomials of degree  $q$  not belonging to  $I$ . The *Hilbert series* is then the formal power series

$$\text{HS}_I(z) = \sum_{q=0}^{\infty} h_I(q)z^q.$$

The Hilbert polynomial can now be defined in the statement of the following theorem.

**Theorem 1.1.7** ([MS21a, Theorem 1.25]). The Hilbert series of a monomial ideal  $I \subseteq \mathbb{C}[x_1, \dots, x_n]$  is

$$\text{HS}_I(z) = \frac{\kappa_I(z)}{(1-z)^n},$$

where  $\kappa_I(z)$  is a polynomial with integer coefficients and  $\kappa_I(0) = 1$ . There exists a univariate polynomial  $\text{HP}(q)$  of degree  $\leq n - 1$ , known as the *Hilbert polynomial* of  $I$ , such that  $\text{HP}(q) = h_I(q)$  for all values of the integer  $q$  that are sufficiently large.

With this machinery at hand, we can finally define the dimension and the degree (first, of a monomial ideal, and then of a variety). The correctness of this definition is ensured by [MS21a, Lemma 1.33].

**Definition 1.1.8** (Dimension and degree). Let  $I \subseteq \mathbb{C}[x_1, \dots, x_n]$  be a monomial ideal and write

$$\text{HP}_I(q) = \frac{g}{(d-1)!}q^{d-1} + \text{lower-degree terms in } q.$$

If  $\text{HP}_I(q)$  is non-zero, the *dimension* of  $I$  is  $d$  and the *degree* of  $I$  is  $g$ . If  $\text{HP}_I(q)$  is identically zero, we say that  $I$  is zero-dimensional, and define the degree of  $I$  to be the dimension of a finite-dimensional vector space  $\mathbb{C}[x_1, \dots, x_n]/I$ . For a variety  $\mathcal{V}$ , its dimension and degree are those of the initial ideal  $\text{in}_{\prec}(I(\mathcal{V}))$ . Here  $\prec$  is any degree-compatible monomial order on the corresponding polynomial ring.

We now define two fundamental properties of varieties that we deal with in this thesis.

**Definition 1.1.9** (Irreducible variety). A variety  $\mathcal{V}$  is called *irreducible* if it cannot be represented as the union of two strictly smaller (w.r.t. inclusion) varieties. Equivalently,  $\mathcal{V}$  is irreducible if and only if  $I_{\mathcal{V}}$  is prime (see Definition 1.1.2).

**Definition 1.1.10** (Unirational variety). A variety  $\mathcal{V}$  is called *unirational* if it can be parametrized by rational functions, that is, if there exists a dominant rational map from  $\mathbb{C}^n$  or  $\mathbb{P}^n$  to  $\mathcal{V}$ .

A family of unirational varieties that plays a significant role in this thesis is that of *toric varieties*. From a certain point of view, toric varieties are the simplest varieties. Yet, many very general concepts of algebraic geometry can already be clearly understood when working with toric varieties. A brief introduction into the topic is presented in [MS21a, Chapter 8], a more comprehensive reference is [Tel22], and an in-depth study can be found in [CLS11]. Chapter 2 of this thesis offers a non-commutative extension of the notion of a toric variety. We now define affine and projective toric varieties.

**Definition 1.1.11** (Affine toric variety). An *affine toric variety* is the Zariski closure of the image of a map  $(\mathbb{C}^*)^n \rightarrow \mathbb{C}^p$ ,  $\mathbf{x} \mapsto (\mathbf{x}^{\mathbf{a}_1}, \dots, \mathbf{x}^{\mathbf{a}_p})$ , where  $\mathbf{a}_i \in \mathbb{Z}^n$  and  $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$ .

That is, affine toric varieties are the varieties that can be parametrized by monomials. Let  $A$  be a matrix whose columns are the vectors  $\mathbf{a}_i$  from the definition above. Such a matrix defines the toric variety and therefore its ideal. If  $A$  contains the vector  $(1, \dots, 1)$  in its row span, then the ideal of the corresponding affine toric variety is generated by homogeneous

polynomials (this follows immediately from [MS21a, Lemma 8.8]). This means that the affine toric variety is a cone over a projective variety. This motivates the following definition.

**Definition 1.1.12** (Projective toric variety). A *projective toric variety* is a variety of the form  $\mathcal{V}(I)$ , where  $I$  is a homogeneous ideal defined by a matrix  $A$  such that  $(1, \dots, 1) \in \text{rowspan}(A)$ .

We now leave the toric world and conclude this subsection with a classical example of a variety: the Grassmannian. This variety is the starting point for Chapter 5.

**Example 1.1.13** (The Grassmannian). The Grassmannian  $\text{Gr}(k, n)$  is the variety parametrizing  $k$ -dimensional subspaces of an  $n$ -dimensional complex linear space. Alternatively, it parametrizes  $(k - 1)$ -dimensional projective subspaces of an  $(n - 1)$ -dimensional projective space. A linear space, considered as an element of  $\text{Gr}(k, n)$ , can be represented by a  $k \times n$ -matrix  $A$  whose rows span the space. The Grassmannian  $\text{Gr}(k, n)$  can then be realized as a projective variety living in  $\mathbb{P}^{\binom{n}{k}-1}$  via the *Plücker embedding*, which sends a matrix  $A$  representing an element in  $\text{Gr}(k, n)$  to the vector of its maximal minors. Maximal minors of  $A$  are called *Plücker coordinates*. This embedding does not depend on the choice of matrix representatives. The Grassmannian is an irreducible unirational projective variety for any  $k$  and  $n$ . Irreducibility is proven, for instance, in [MS21a, Theorem 5.4], and unirationality follows from the construction of the Plücker embedding. This construction can be carried out over  $\mathbb{R}$ , resulting in the real Grassmannian  $\text{Gr}_{\mathbb{R}}(k, n)$  living in the real projective space.  $\blacklozenge$

## 1.1.2. Semi-algebraic sets

The reason why varieties in the previous subsection are defined over  $\mathbb{C}$  is that this is an algebraically closed field. This fact is crucial for many classical results in algebraic geometry. For instance, Bézout's theorem does not in general hold over non-closed fields [Sha13, Chapter 1, §1]. This makes studying varieties over  $\mathbb{R}$  much more difficult. One more important difference between  $\mathbb{R}$  and  $\mathbb{C}$ , however, is that the former is an ordered field. This allows to consider polynomial inequalities over the real numbers. Real algebraic geometry mainly studies sets that are defined by polynomial inequalities rather than equations. These are known as semi-algebraic sets. We now give precise definitions.

**Definition 1.1.14** (Basic semi-algebraic set). A *basic closed (resp. open) semi-algebraic set* in  $\mathbb{R}^n$  is the set of solutions to a system of finitely many polynomial inequalities of the form  $f(x_1, \dots, x_n) \geq 0$  (resp.  $f(x_1, \dots, x_n) > 0$ ), where  $f \in \mathbb{R}[x_1, \dots, x_n]$ .

**Definition 1.1.15** (Semi-algebraic set). A *semi-algebraic set* in  $\mathbb{R}^n$  is a finite boolean combination of basic semi-algebraic sets (open or closed).

One can also define semi-algebraic sets in the real projective space.

**Definition 1.1.16** (Projective semi-algebraic set). A set  $S \subseteq \mathbb{RP}^n$  is *semi-algebraic* if it is the image of a semi-algebraic set in  $\mathbb{R}^{n+1}$  under the natural map  $\mathbb{R}^{n+1} \setminus \{0\} \rightarrow \mathbb{RP}^n$ .

The class of semi-algebraic sets is closed under projections and, more generally, polynomial maps. This statement is the content of the Tarski-Seidenberg theorem [MS21a, Theorem 4.17], [BCR13, Theorem 1.4.2].

**Theorem 1.1.17** (Tarski-Seidenberg). The image of a semi-algebraic set in  $\mathbb{R}^n$  under a polynomial map to  $\mathbb{R}^m$  is again a semi-algebraic set.

Semi-algebraic sets play a significant role in convex optimization, as explained in Section 1.4. Important examples introduced in that section are spectrahedra. Here we give another example relevant for Chapter 5, continuing the story of the Grassmannian from Example 1.1.13.

**Example 1.1.18** (The totally nonnegative Grassmannian). The *totally nonnegative Grassman-*

nian  $\text{Gr}_{\geq 0}(k, n)$  is the subset of  $\text{Gr}(k, n)$  consisting of the elements whose non-zero Plücker coordinates are all real and have the same sign [Pos06, Wil21]. Each element in  $\text{Gr}_{\geq 0}(k, n)$  can be represented by a *totally nonnegative*  $k \times n$ -matrix, that is, by a matrix with nonnegative maximal minors. The *totally positive Grassmannian*  $\text{Gr}_{> 0}(k, n)$  is a subset of  $\text{Gr}(k, n)$  consisting of the elements whose Plücker coordinates are all non-zero, real, and have the same sign. Elements of  $\text{Gr}_{> 0}(k, n)$  can be represented by  $k \times n$  matrices with strictly positive maximal minors (such matrices are called *totally positive*). The condition that all Plücker coordinates have the same sign is described by inequalities of the form  $p_I p_J \geq 0$ , where  $p_I$  and  $p_J$  are Plücker coordinates, so  $\text{Gr}_{\geq 0}(k, n)$  and  $\text{Gr}_{> 0}(k, n)$  are semi-algebraic sets.  $\blacklozenge$

## 1.2. Quantum information theory

Chapter 3 is devoted to problems arising in quantum information theory. Since this is an area that is not necessarily familiar to an algebraic geometer, we briefly introduce the very basic notions of this theory. Our main references here are [NC02] and [Lan19], and we follow the exposition of [DPW23, Section 2].

**Definition 1.2.1 (Quantum state).** A *quantum state* on  $N$  *qudits* is represented by a vector  $|\psi\rangle \in \mathcal{H} = \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_N$  of unit length, where  $\mathcal{H}_i$  is the Hilbert space (with the Hermitian inner product)  $\mathcal{H}_i \cong \mathbb{C}^d$ ,  $i = 1, \dots, N$ . Here, we make use of the Dirac notation, i.e.  $|\psi\rangle$  denotes a column vector and  $\langle\psi|$  its complex conjugate transpose. In the case  $N = 1$  and  $d = 2$ ,  $|\psi\rangle$  is called a *qubit*.

**Definition 1.2.2 (Density matrix).** An *ensemble of quantum states* is a collection  $\{p_i, |\psi_i\rangle\}_i$  where  $\{p_i\}_i$  is a discrete probability distribution. The *density matrix* of an ensemble is

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|.$$

Equivalently, we can characterize density matrices as positive semidefinite endomorphisms (that is, endomorphisms with real nonnegative eigenvalues) on  $\mathcal{H}$  with unit trace. We will use the terms “quantum state” and “density matrix” interchangeably. The set of all density matrices on  $\mathcal{H}$  is denoted by  $\mathcal{D}(\mathcal{H})$ .

**Definition 1.2.3.** Let  $\rho_{AB}$  be a bipartite quantum state, i.e. a density operator on  $\mathcal{H}_A \otimes \mathcal{H}_B$ . We define the *partial trace* over the  $B$ -system on elementary tensors via

$$\text{Tr}_B(|a_i\rangle\langle a_j| \otimes |b_k\rangle\langle b_l|) := |a_i\rangle\langle a_j| \cdot \text{Tr}(|b_k\rangle\langle b_l|) = |a_i\rangle\langle a_j| \cdot \langle b_l|b_k\rangle,$$

where  $|a_i\rangle, |a_j\rangle \in \mathcal{H}_A$  and  $|b_k\rangle, |b_l\rangle \in \mathcal{H}_B$  and extend this operation linearly to  $\rho_{AB}$ . Note that  $\text{Tr}_B \rho_{AB}$  is a density operator on  $\mathcal{H}_A$  and therefore we use the notation  $\rho_A := \text{Tr}_B \rho_{AB}$ .

One can think of the partial trace operation as the quantum analog of marginalization in statistics. In physics  $\rho_A$  describes the state of the subsystem  $A$  of the composite system  $AB$ .

**Example 1.2.4.** Set  $|0\rangle := (1, 0)^T$ ,  $|1\rangle := (0, 1)^T$  as a basis of  $\mathbb{C}^2$  (this basis is called *computational basis* in the context of quantum information theory); we also adapt the notation to write  $|ij\rangle$  for  $|i\rangle \otimes |j\rangle$ ,  $i, j \in \{0, 1\}$ . Consider the *Bell state*

$$\rho_{AB} = \left(\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)\right) \left(\frac{1}{\sqrt{2}}(\langle 00| + \langle 11|)\right) = \frac{1}{2}(|00\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 00| + |11\rangle\langle 11|)$$

on  $\mathbb{C}^2 \otimes \mathbb{C}^2$ . Then the partial trace over  $B$  is computed as

$$\text{Tr}_B \rho_{AB} = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|) = \frac{1}{2}(\text{Id}_2),$$

called the *maximally mixed state*, meaning  $\rho_{AB}$  is *maximally entangled* [NC02, §2].  $\blacklozenge$

Finally, when we speak of a *Hamiltonian*  $H$ , we simply mean a real symmetric matrix. The reason we restrict to this class (and do not consider Hermitian matrices as would be natural in many contexts in quantum physics) is that symmetric matrices form a linear space, while the set of Hermitian matrices is not an algebraic variety. Likewise, we consider density matrices to be real positive semidefinite (PSD), see Definition 1.4.3. Note that  $\exp(H)$  is positive definite and thus can (up to normalization) be regarded as a quantum state.

## 1.3. Discrete geometry

The content of Chapter 4 revolves around classical discrete geometry structures that have been studied since antiquity: polytopes and cones. Several incarnations of these objects, e.g. the cone of positive semidefinite matrices, also appear in other chapters. In this section we define polytopes and cones and discuss some of their basic properties. We also introduce a number of related notions that will be of use to us in the following chapters. Our primary source here is [Zie12], while we also rely on [DLRS10] for the definition of the chamber complex. Another source containing many of the notions introduced in this section is [Jos21].

### 1.3.1. Polyhedra, polytopes, and cones

The objects we are going to define in this section are examples of convex sets. Our first definition is therefore that of convexity.

**Definition 1.3.1** (Convex set). A set  $\mathcal{K} \subseteq \mathbb{R}^d$  is called *convex* if for any two points  $x, y \in \mathcal{K}$  the line segment connecting them lies in  $\mathcal{K}$ . That is,  $\lambda x + (1 - \lambda)y \in \mathcal{K}$  for any  $\lambda \in [0, 1]$ .

We are now ready to define the main characters of this section.

**Definition 1.3.2** (Polyhedron). Let  $a_1, \dots, a_k \in \mathbb{R}^d$  and  $b_1, \dots, b_k \in \mathbb{R}$ . The set

$$\{x \in \mathbb{R}^d : \langle x, a_i \rangle \geq b_i \text{ for } i = 1, \dots, k\},$$

where  $\langle \cdot, \cdot \rangle$  is the standard scalar product on  $\mathbb{R}^d$ , is called a *polyhedron*. That is, a polyhedron is the intersection of a finite number of affine half-spaces in  $\mathbb{R}^d$ . The dimension of a polyhedron is the dimension of its affine hull (i.e. of the smallest affine-linear space containing it).

**Definition 1.3.3** (Convex hull). Let  $v_1, \dots, v_n \in \mathbb{R}^d$  be a finite collection of points. Their *convex hull* is the set

$$\text{conv}(v_1, \dots, v_n) = \{\lambda_1 v_1 + \dots + \lambda_n v_n : \lambda_1 + \dots + \lambda_n = 1, \lambda_i \geq 0 \text{ for } i = 1, \dots, n\}.$$

This is the smallest (w.r.t. inclusion) convex set containing all the points  $v_i$ .

**Definition 1.3.4** (Polytope). A *polytope* is the convex hull of finitely many points in  $\mathbb{R}^d$ .

The following statement appears as the ‘‘Main theorem for polytopes’’ in [Zie12, §1.1].

**Theorem 1.3.5.** Polytopes are exactly bounded polyhedra.

**Definition 1.3.6** (Cone). A non-empty set  $C \subseteq \mathbb{R}^d$  is a *cone* if  $\lambda x \in C$  for any  $x \in C$  and  $\lambda \in \mathbb{R}_{\geq 0}$ . If a cone is in addition a polyhedron, then we call it a *polyhedral cone*.

Note that every cone contains the origin. We now introduce the notion of the conical hull, which will be necessary to formulate an analog of Theorem 1.3.5 for cones.

**Definition 1.3.7** (Conical hull). Let  $v_1, \dots, v_n \in \mathbb{R}^d$  be a finite collection of vectors. Their *conical hull* is the set

$$\text{cone}(v_1, \dots, v_n) = \{\lambda_1 v_1 + \dots + \lambda_n v_n : \lambda_i \geq 0 \text{ for } i = 1, \dots, n\}.$$

This is the smallest (w.r.t inclusion) cone containing all the vectors  $v_i$ . If the scalars  $\lambda_i$  are only allowed to be strictly positive, one gets the definition of the *positive hull*  $\text{pos}(v_1, \dots, v_n)$ .

The following statement is the “Main theorem for cones” [Zie12, Theorem 1.3].

**Theorem 1.3.8.** A cone  $C \subseteq \mathbb{R}^d$  is polyhedral if and only if it is the conical hull of finitely many vectors.

Polyhedra, including polytopes and polyhedral cones, are objects with rich combinatorial structure. We here give just one, arguably the most basic, illustration of this claim. We say that a hyperplane  $H \subset \mathbb{R}^d$  *supports* a polyhedron  $P$  if  $H$  bounds a half-space that contains  $P$ . A *face* of  $P$  is the intersection of  $P$  with a supporting hyperplane. A face  $F$  of  $P$  is called *proper* if  $F \subsetneq P$ . Inclusion-maximal proper faces of  $P$  are called *facets*, and zero-dimensional faces are called *vertices*. The faces of  $P$  are naturally ordered by inclusion and thus form a partially ordered set called the *face lattice*  $\mathcal{F}$  of  $P$ . We say that two polyhedra have the same *combinatorial type* if their face lattices are isomorphic. We need one more definition for polyhedral cones: their one-dimensional faces are called *rays*.

We now move on to defining several important properties of polytopes and cones that play a role in this dissertation.

**Definition 1.3.9.** A full-dimensional polytope  $P \subset \mathbb{R}^d$  is called *simple* if every vertex is adjacent to exactly  $d$  facets. We call  $P$  *simplicial* if every proper face of  $P$  is a simplex (equivalently, if every facet of  $P$  contains exactly  $d$  vertices).

**Definition 1.3.10.** A cone  $C \subseteq \mathbb{R}^d$  is called *pointed* if  $C \cap -C = \{0\}$ , where  $-C = \{-x : x \in C\}$ . We call  $C$  *proper* if it is full-dimensional, closed in the Euclidean topology and pointed.

We conclude this subsection with a brief discussion on polar duality, a construction that is essential for everything that happens in Chapter 4.

**Definition 1.3.11 (Polar dual).** For any subset  $P \subseteq \mathbb{R}^d$ , the *polar dual* set is defined by

$$P^\circ := \{y \in \mathbb{R}^d : \langle x, y \rangle \leq 1 \text{ for all } x \in P\},$$

where  $\langle \cdot, \cdot \rangle$  is the standard scalar product on  $\mathbb{R}^d$ . In what follows, we will normally refer to  $P^\circ$  as simply the dual of  $P$  for the sake of brevity.

The following statement appears as Proposition A.3 in [Jos21] and demonstrates that polytopes behave well with respect to polar duality. It is proven in [Zie12, §2.3].

**Theorem 1.3.12.** Let  $P$  be a full-dimensional polytope in  $\mathbb{R}^d$  containing the origin in its interior. Let  $v_1, \dots, v_k$  be the vertices of  $P$ . The following statements hold:

- (i) The polar dual  $P^\circ$  is a full-dimensional polytope containing the origin in the interior. More precisely,

$$P^\circ = v_1^\circ \cap \dots \cap v_k^\circ.$$

- (ii) The face lattice of  $P^\circ$  is anti-isomorphic to that of  $P$ . In particular, if  $P$  is simple, then  $P^\circ$  is simplicial, and vice versa.

- (iii)  $P^{\circ\circ} = P$ .

### 1.3.2. Subdivisions, triangulations, and chambers

In discrete geometry one studies not just individual objects with interesting combinatorial properties (such as polyhedra) but also nicely arranged collections of such objects. The primary example here is the notion of a polyhedral complex.



**Definition 1.3.13** (Polyhedral complex). A *polyhedral complex* in  $\mathbb{R}^d$  is a non-empty finite collection  $\Pi$  of polyhedra that satisfies the two following properties. Firstly, any face of a polyhedron in  $\Pi$  also lies in  $\Pi$ . Secondly, any (possibly empty) intersection of polyhedra in  $\Pi$  is also in  $\Pi$ . A polyhedral complex composed of cones is called a *polyhedral fan* and its one-dimensional elements are called *rays*, in analogy with the case of individual cones. A polyhedral complex composed of polytopes is called a *polytopal complex*. We say that a polyhedral complex  $\Sigma$  *refines*  $\Pi$  if each element of  $\Sigma$  is contained in an element of  $\Pi$ . In this same case we say that  $\Pi$  *coarsens*  $\Sigma$ .

The ultimate goal of this subsection is to define the chamber complex of a polytope  $P$ . The first step to this is the notion of a polyhedral subdivision.

**Definition 1.3.14** (Subdivision). Let  $P$  be a polyhedron. A polyhedral complex  $\Pi$  is called a *subdivision* of  $P$  if the union of its elements is equal to  $P$ .

There is a particularly simple and very important class of subdivisions.

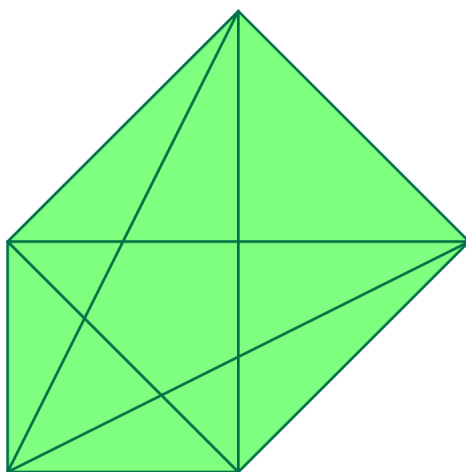
**Definition 1.3.15** (Triangulation). A subdivision  $\Pi$  of a polyhedron  $P$  is called a *triangulation* if all of its elements are simplices.

We are now ready to define the chamber complex. The definition we give here is fairly concise, and the one that is more convenient for computations is presented in Section 4.2.

**Definition 1.3.16** (Chamber complex). Let  $P$  be a full-dimensional polyhedron in  $P$ . The *chamber complex*  $\mathcal{C}(P)$  is the coarsest common refinement of all its triangulations. That is, every triangulation of  $P$  coarsens  $\mathcal{C}(P)$  and there exists no subdivision of  $P$  with this property that is strictly contained in  $\mathcal{C}(P)$ . Elements of  $\mathcal{C}(P)$  are called *chambers* of  $P$ , and full-dimensional chambers are referred to as *cells*.

In Chapter 4, we will give a slightly more general definition of the chamber complex of a point configuration (Definition 4.2.2). We conclude this subsection with a simple example.

**Example 1.3.17** (Chamber complex of a pentagon). The chamber complex of a pentagon in  $\mathbb{R}^2$  with the vertices  $(0, 0)$ ,  $(0, 1)$ ,  $(1, 0)$ ,  $(1, 2)$  and  $(2, 1)$  is shown in Figure 1.1. It has 10 zero-dimensional, 20 one-dimensional and 11 two-dimensional chambers.  $\blacklozenge$



**Figure 1.1:** Chamber complex of a pentagon in  $\mathbb{R}^2$ .

## 1.4. Convex optimization

Chapters 2 and 4 of this thesis are inspired by problems in convex optimization. A convex optimization problem consists in minimizing a convex function over a convex set. Convex sets are defined in Definition 1.3.1, and we start this section with defining convex functions.

**Definition 1.4.1** (Convex function). Let  $\mathcal{K} \subseteq \mathbb{R}^n$  be a convex set and  $f : \mathcal{K} \rightarrow \mathbb{R}$  be a real-valued function on  $\mathcal{K}$ . The function  $f$  is called *convex (on  $\mathcal{K}$ )* if for all  $0 \leq t \leq 1$  and all  $x, y \in \mathcal{K}$  we have  $f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$ .

This definition is naturally connected to that of a convex set. For  $\mathcal{K} = \mathbb{R}$ , the function is convex if and only if its epigraph (the set of points “above” its graph) is a convex set in  $\mathbb{R}^2$ .

A general problem of convex optimization can then be formulated as follows:

$$\text{minimize } f(x) \tag{1.4.1}$$

$$\text{subject to } x \in \mathcal{K}, \tag{1.4.2}$$

where  $\mathcal{K} \subseteq \mathbb{R}^n$  is a convex set and  $f$  is a convex function on  $\mathcal{K}$ . We call  $\mathcal{K}$  the *feasible set* or *feasible region* of the optimization problem.

It is straightforward to see that any linear function in  $n$  variables is convex on  $\mathbb{R}^n$ : the non-strict inequality from Definition 1.4.1 always turns into equality in this case. The two well-studied convex optimization problems that play a role in this thesis are linear and semidefinite programming. In both cases one minimizes a linear function, and the specifics of the problem is determined by a class of sets that  $\mathcal{K}$  is chosen from. In our exposition of these problems, we rely primarily on [BPT12].

**Definition 1.4.2** (Linear programming). *Linear programming* is the problem of minimizing a linear function subject to constraints given by linear equalities and inequalities. It is usually presented in *standard form*

$$\text{minimize } c^T x \tag{1.4.3}$$

$$\text{subject to } Ax = b, x \geq 0, \tag{1.4.4}$$

where  $A$  is a  $d \times n$  real matrix,  $b \in \mathbb{R}^d$ , and the minimization is performed over  $x \in \mathbb{R}^n$ . The inequality  $x \geq 0$  is interpreted component-wise, i.e.  $x_i \geq 0$  for  $i = 1, \dots, n$ .

The constraints of a linear program define a section of the positive orthant  $\mathbb{R}_{\geq 0}^n$  by an affine subspace of  $\mathbb{R}^n$ . This is a polyhedron. Polyhedra are convex sets, so linear programming is indeed a convex optimization problem.

Semidefinite programming is a generalization of linear programming, in which the decision variables  $x$  are symmetric matrices, and the constraints are given by linear matrix inequalities. Feasible regions of semidefinite programs are called spectrahedra. We now introduce all the necessary notions to formulate a semidefinite program. We write  $\mathbb{S}^n$  for the set of symmetric real  $n \times n$ -matrices. This is a real vector space of dimension  $\binom{n+1}{2}$ .

**Definition 1.4.3** (PSD matrix). A matrix  $A \in \mathbb{S}^n$  is called *positive semidefinite (PSD)* if  $x^T A x \geq 0$  for all  $x \in \mathbb{R}^n$ , and *positive definite* if  $x^T A x > 0$  for all nonzero  $x \in \mathbb{R}^n$ . Equivalently,  $A$  is positive semidefinite if all of its eigenvalues are nonnegative, and positive definite if all of its eigenvalues are strictly positive.

We denote the set of positive semidefinite  $n \times n$ -matrices by  $\mathbb{S}_+^n$ . If  $A, B \in \mathbb{S}_+^n$  and  $x, y \geq 0$ , one sees from the first definition that  $xA + yB \in \mathbb{S}_+^n$ . Thus,  $\mathbb{S}_+^n$  is a closed convex cone in  $\mathbb{S}^n \cong \mathbb{R}^{\binom{n+1}{2}}$ . This cone is proper [BPT12, page 20] and, in particular, full-dimensional. Its interior in the Euclidean topology is the set of positive definite matrices, which is an open

cone. We therefore denote this open cone by  $\text{int}(\mathbb{S}_+^n)$ . One more piece of notation: instead of  $A \in \mathbb{S}_+^n$ , we sometimes write  $A \succcurlyeq 0$ .

**Definition 1.4.4** (Linear matrix inequality). A *linear matrix inequality* has the form

$$A_0 + \sum_{i=1}^d A_i x_i \succcurlyeq 0 \text{ for } A_i \in \mathbb{S}^n.$$

**Definition 1.4.5** (Spectrahedron). A *spectrahedron* is a set of the form

$$S = \left\{ (x_1, \dots, x_d) \in \mathbb{R}^d : A_0 + \sum_{i=1}^d A_i x_i \succcurlyeq 0 \right\}.$$

Note that when the matrices  $A_i$  are diagonal, one recovers the definition of a polyhedron.

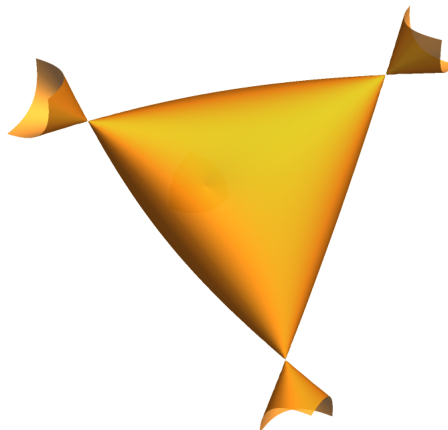
As the coefficients  $x_i \in \mathbb{R}$  vary, the matrices  $A_0 + \sum_{i=1}^d A_i x_i$  trace out an  $d$ -dimensional affine subspace of  $\mathbb{S}^n$ . We can therefore interpret the corresponding spectrahedron as the intersection of this affine space with the PSD cone  $\mathbb{S}_+^n$ . In this interpretation spectrahedra live in  $\mathbb{S}^n$  rather than in  $\mathbb{R}^d$ . We will make use of this when defining semidefinite programming in standard primal form.

Affine subspaces are convex sets, and so is the PSD cone. The class of convex sets is closed under taking intersections. Therefore, spectrahedra are convex sets. Moreover, the PSD cone is a basic semi-algebraic set [BPT12, Proposition A.1]. Thus, any spectrahedron is a basic semi-algebraic set as well. This clarifies our claim from Subsection 1.1.2.

**Example 1.4.6** (The elliptope). One of the most classical examples of a spectrahedron is the three-dimensional elliptope  $\mathcal{E}_3$ . This set is defined as follows:

$$\mathcal{E}_3 = \left\{ (x, y, z) \in \mathbb{R}^3 : \begin{bmatrix} 1 & x & y \\ x & 1 & z \\ y & z & 1 \end{bmatrix} \succcurlyeq 0 \right\}.$$

It is the bounded component in Figure 1.2. ◆



**Figure 1.2:** The elliptope  $\mathcal{E}_3$ .

One final note before we define semidefinite programs is that the space of symmetric matrices  $\mathbb{S}^n$  has the scalar product  $\langle X, Y \rangle = \text{Tr}(XY)$ , where  $\text{Tr}$  denotes the matrix trace.

**Definition 1.4.7** (Semifinite programming (SDP)). *Semidefinite programming (SDP)* is the problem of minimizing a linear function over a spectrahedron. It usually is presented in *standard primal form*

$$\text{minimize } \langle C, X \rangle \tag{1.4.5}$$

$$\text{subject to } \langle A_i, X \rangle = b_i, \quad i = 1, \dots, m, \tag{1.4.6}$$

$$X \succcurlyeq 0, \tag{1.4.7}$$

where  $C, A_i \in \mathbb{S}^n$ . The optimization is performed over  $X \in \mathbb{S}_+^n$ .

When the matrices  $C$  and  $A_i$  are diagonal, one recovers the setup of linear programming.

We conclude this section with an observation that allows for a generalization of SDP which we will however not consider in this thesis. The constraints of a linear program are  $x \geq 0$  and  $Ax = b$ . The first one defines the positive orthant  $\mathbb{R}_{\geq 0}^n$ , which is a proper cone. The second one defines its affine section. The setup of semidefinite programming is very similar: the place of the positive orthant is taken by the PSD cone  $\mathbb{S}_+^n$ , and we still consider affine sections thereof. This is generalized by *conic programming*, in which one optimizes a linear function over an affine section of an arbitrary proper cone. Powerful algebraic methods are available for solving many conic programs (see, for instance, [Nem06]).

## 1.5. Algebraic combinatorics

In Chapter 5 we study Grasstopes from the point of view of algebraic combinatorics. The main tools we use are hyperplane arrangements and oriented matroids. In this section we define these objects, based on [BLVS<sup>+</sup>99] and following the exposition of [MPP23, Section 5].

### 1.5.1. Hyperplane arrangements

**Definition 1.5.1** (Hyperplane arrangement). A (*real*) *projective hyperplane arrangement*  $\mathbf{P} = \{H_1, \dots, H_n\}$  is a finite set of hyperplanes in the real projective space  $\mathbb{R}\mathbb{P}^k$ . Analogously, an *affine hyperplane arrangement*  $\mathbf{A}$  is a finite set of hyperplanes in  $\mathbb{R}^k$ .

**Definition 1.5.2** (Simple arrangement). A hyperplane arrangement is called *simple* if the intersection of any  $i \leq k$  hyperplanes in the arrangement has codimension exactly  $i$  in the ambient  $k$ -dimensional space, and the intersection of any  $i > k$  hyperplanes is empty.

**Definition 1.5.3** (Region). A *region* of a hyperplane arrangement is a connected component of the complement of the union of the hyperplanes in the arrangement.

In [Zas75], Zaslavsky gives formulae for the total number of regions  $t(\mathbf{A})$  and the number of bounded regions  $b(\mathbf{A})$  of a simple affine arrangement  $\mathbf{A}$  of  $n$  hyperplanes in  $\mathbb{R}^k$ :

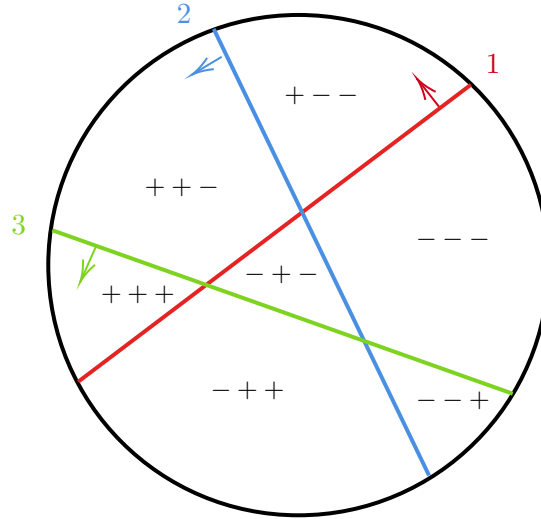
$$t(\mathbf{A}) = 1 + n + \binom{n}{2} + \dots + \binom{n}{k},$$

$$b(\mathbf{A}) = \binom{n-1}{k}.$$

Note that since intersections of two hyperplanes in any projective arrangement  $\mathbf{P}$  have codimension two in  $\mathbb{P}^k$  and there are a finite number of them, we can always find some hyperplane avoiding them, and hence an affine chart which contains all of them (see Figure

1.3). A projective arrangement  $\mathbf{P}$  naturally induces an affine arrangement  $\mathbf{A}$  in this chart. If a region of  $\mathbf{P}$  intersects the chosen hyperplane at infinity, it induces two unbounded regions of  $\mathbf{A}$ . Otherwise it induces a single bounded region. Thus, the total number of regions in  $\mathbf{P}$  is

$$r(\mathbf{P}) = b(\mathbf{A}) + \frac{t(\mathbf{A}) - b(\mathbf{A})}{2}. \quad (1.5.1)$$



**Figure 1.3:** An oriented hyperplane arrangement and sign labels of its regions

**Definition 1.5.4** (Oriented arrangement). One can order the hyperplanes in a given affine arrangement  $\mathbf{A}$ . Each of the hyperplanes  $H_i \in \mathbf{A}$  splits  $\mathbb{R}^k$  into two half-spaces. One can declare one of them to be positive with respect to  $H_i$  and the other one to be negative. We will call  $\mathbf{A}$  together with the ordering of hyperplanes and the choice of the positive half-space for each of them an *oriented hyperplane arrangement*. Regions of an oriented hyperplane arrangement are naturally labeled by ordered sequences of + and - signs, reflecting in which half-space with respect to each hyperplane a given region lies. An affine hyperplane arrangement  $\mathbf{A}$  in  $\mathbb{R}^k$  naturally induces a projective arrangement  $\mathbf{P}$  in  $\mathbb{RP}^{k-1}$ . As explained above, two unbounded regions of  $\mathbf{A}$  give rise to one region of  $\mathbf{P}$ . On the level of sign labels in  $\mathbf{P}$ , we identify the label  $\sigma$  and the label  $-\sigma$  obtained from  $\sigma$  by flipping all the signs.

An example of an oriented arrangement with sign labels is shown in Figure 1.3. As we will see in Section 1.5.2, oriented hyperplane arrangements define realizable oriented matroids.

## 1.5.2. Oriented matroids

Much of the material on hyperplane arrangements can be naturally generalized to the setting of oriented matroids. In this section, we review the basics of oriented matroid theory and give a dictionary between the language of hyperplane arrangements and that of oriented matroids. We begin by discussing signed circuits, which we will often abbreviate as circuits if the meaning is clear from context.

**Definition 1.5.5** (Signed circuit axioms). An *oriented matroid* consists of a ground set  $\mathcal{E}$  and a collection  $\mathcal{C}$  of tuples of the form  $X = (X^+, X^-)$  called *circuits*, where  $X^+, X^-$  are disjoint subsets of  $\mathcal{E}$  satisfying

- (i)  $(\emptyset, \emptyset)$  is not a circuit.

- (ii) If  $X$  is a circuit, then so is  $-X = (X^-, X^+)$ .
- (iii) No proper subset of a circuit is a circuit.
- (iv) (Elimination). If  $X_0$  and  $X_1$  are two signed circuits with  $X_0 \neq X_1$  and  $e \in X_0^+ \cap X_1^-$ , then there is a third circuit  $X \in \mathcal{C}$  with  $X^+ \subset (X_0^+ \cup X_1^+) \setminus \{e\}$  and  $X^- \subset (X_0^- \cup X_1^-) \setminus \{e\}$ .

We can obtain the oriented matroid of a matrix as follows.

**Definition 1.5.6** (Oriented matroid of a matrix). Fix a matrix  $A$  and let  $\sum_i \lambda_i v_i = 0$  be a minimal linear dependency among its rows. Associate to this dependency the signed set  $X = (X^-, X^+)$ , where

$$\begin{aligned} X^- &= \{i : \lambda_i < 0\}, \\ X^+ &= \{i : \lambda_i > 0\}. \end{aligned}$$

Then the *oriented matroid*  $\mathcal{M}_A$  associated to  $A$  has as its signed circuits the signed sets coming from minimal linear dependencies, that is, from left kernel vectors with minimal support.

One can check that oriented matroids of matrices satisfy the signed circuit axioms.

**Example 1.5.7** (Oriented matroid of a matrix  $A$ ). Consider the matrix

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & -3 & 4 \end{bmatrix}.$$

Then the only linear dependency up to scaling is  $v_4 - 4v_3 + 3v_2 - v_1 = 0$ . Thus the only circuit (when only one of  $X, -X$  is considered) is  $24\bar{1}3$ , where we use a more compact notation in which the bar indicates being in the negative part of the signed set.  $\blacklozenge$

**Remark 1.5.8.** One can also define  $\mathcal{M}_A$  by signed bases; the matroid is then the map which assigns to each size  $k$  subset  $I \subset [n]$  the sign of the determinant of  $A_I$ . This definition is called the *chirotope definition* [BLVS<sup>+</sup>99, page 6] and satisfies the *chirotope axioms*, which we do not describe here. In particular, the database [Fin] that we use in Chapter 5 indexes matroids by their chirotope.

We recall a few definitions we will need to explain how a hyperplane arrangement can be viewed as an oriented matroid.

**Definition 1.5.9** (Composition). Let  $X = (X^+, X^-)$  and  $Y = (Y^+, Y^-)$  be signed sets. Then their *composition*  $X \circ Y$  is  $(X^+ \cup (Y^+ \setminus X^-), X^- \cup (Y^- \setminus X^+))$ .

**Definition 1.5.10** (Orthogonality). Let  $X = (X^+, X^-)$  and  $Y = (Y^+, Y^-)$  be signed sets. Define  $S(X, Y) = (X^+ \cap Y^-) \cup (X^- \cap Y^+)$ . We say  $X$  and  $Y$  are *orthogonal* if  $S(X, Y)$  and  $S(X, -Y)$  are both empty or both non-empty.

We define *vectors* as compositions of circuits, and *cocircuits* and *covectors* as the circuits and vectors of the dual matroid [BLVS<sup>+</sup>99, page 4], respectively. An equivalent definition which is easier for computing is that the covectors of  $\mathcal{M}$  are the signed sets which are orthogonal to all vectors of  $\mathcal{M}$ . For more detail, see [BLVS<sup>+</sup>99, Chapter 1]. There is yet another definition in terms of vector configurations.

**Definition 1.5.11** (Oriented matroid of a vector configuration). One can also view the rows  $v_i$  of the matrix  $A$  as vectors in  $\mathbb{R}^k$ . For such a vector configuration, the covectors of  $\mathcal{M}_A$  can be defined as the set of tuples  $Y_H = (Y_H^+, Y_H^-)$  as  $H$  runs over oriented hyperplanes, where  $Y_H^+$  is

the set of vectors in the positive half-space defined by  $H$ , and  $Y_H^-$  is the set of vectors in the negative half-space. The *cocircuits* are the minimal covectors. They arise from hyperplanes that are spanned by subsets of  $\{v_1, \dots, v_n\}$ .

**Definition 1.5.12** (Oriented matroid of a hyperplane arrangement). Let  $H_i$  be the hyperplane given by the vanishing of  $l_i(x) = a_i \cdot x$ . Then the oriented matroid of the arrangement  $\{H_i\}_{i=1}^n$  is the matroid of the vector configuration given by  $a_1, \dots, a_n$ , or equivalently, the oriented matroid of the matrix with rows  $a_1, \dots, a_n$ .

**Remark 1.5.13.** Faces of a hyperplane arrangement correspond to covectors of its oriented matroid, and regions correspond to maximal covectors. The *rank* (denoted by  $r$ ) of the matroid is  $k + 1$ , where  $k$  is the dimension of the ambient space [BLVS<sup>+</sup>99, Chapter 1].

**Example 1.5.14** (Oriented matroid of a matrix  $A$ ). Let  $A$  be the matrix

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & -3 & 4 \end{bmatrix}$$

as in Example 1.5.7. The cocircuits are given by considering hyperplanes  $H$  spanned by pairs of rows. For example, if  $H = \text{Span}\{a_1, a_3\} = \{y = 0\}$ , then  $a_2$  is in the positive half-space (since the  $y$ -coordinate is 1) and  $a_4$  is in the negative half-space (since the  $y$  coordinate is  $-3$ ). Thus we obtain the cocircuit  $2\bar{4}$ .

The total set of cocircuits (not including negations) is  $\{2\bar{4}, 3\bar{4}, 2\bar{3}, 1\bar{4}, 3\bar{1}, 1\bar{2}\}$ . In the previous sections we have used sign vector notation; that is, to each signed set we associate a length  $n$  vector with  $\pm$  at index  $i$  if  $i$  is in  $X^\pm$  and 0 otherwise. Applying this convention, one can check that the covectors are exactly the sign vectors with fewer than 3 sign changes. For example,  $2\bar{4} \circ 3\bar{1} = 2\bar{3}\bar{1}\bar{4}$ , which corresponds to  $(- + + -)$  with two sign changes.  $\blacklozenge$

**Definition 1.5.15.** An oriented matroid is *realizable* if it arises from a hyperplane arrangement over  $\mathbb{R}$ .

In Section 5.4 we will use the correspondence between hyperplane arrangements and matroids to suggest an abstract definition of a Grasstopo. This correspondence will also prove helpful for our computations, presented in the same section.

## Chapter 2

# Gibbs Manifolds

Toric varieties provide the geometric foundations for many successes in the mathematical sciences. In statistics they appear as discrete exponential families [Sul18, p. 2], and their ideals reveal Markov bases for sampling from conditional distributions [DS98]. In optimization, they furnish nonnegativity certificates [FdW22] and govern the entropic regularization of linear programming [STVvR24]. Sightings in phylogenetics, stochastic analysis, Gaussian inference and chemical reaction networks motivate the slogan *the world is toric* [MS21a, Section 8.3].

In all of these applications, the key player is the positive part of the toric variety. That real manifold is identified with a convex polytope by the moment map [MS21a, Theorem 8.24]. The fibers of the underlying linear map are polytopes of complementary dimension, and each fiber intersects the toric variety uniquely, in the *Birch point*. This is the unique maximizer of the entropy over the polytope [PS05, Theorem 1.10]. In statistical physics and computer science [Vig99], the Birch point is called the *Gibbs distribution*. This name refers to the maximum entropy state in a quantum system, and is the reason behind the title of this chapter.

This chapter offers a non-commutative extension of applied toric geometry. In that extension, points in  $\mathbb{R}^n$  are replaced by real symmetric  $n \times n$ -matrices, and linear programming is replaced by semidefinite programming. There is a moment map which takes the cone of positive semidefinite matrices onto a spectrahedral shadow, and whose fibers are spectrahedra of complementary dimension. The Gibbs manifold plays the role of the positive toric variety. Each spectrahedron intersects the Gibbs manifold uniquely, in the *Gibbs point*. Just like in the toric case, we study these objects algebraically by passing to the Zariski closure of our positive manifold. The resulting analogues of toric varieties are called *Gibbs varieties*.

We illustrate these concepts for the following linear space of symmetric  $3 \times 3$ -matrices:

$$\mathcal{L} = \left\{ \begin{bmatrix} y_1 + y_2 + y_3 & y_1 & y_2 \\ y_1 & y_1 + y_2 + y_3 & y_3 \\ y_2 & y_3 & y_1 + y_2 + y_3 \end{bmatrix} : y_1, y_2, y_3 \in \mathbb{R} \right\}. \quad (2.0.1)$$

The Gibbs manifold  $\text{GM}(\mathcal{L})$  is obtained by applying the exponential function to each matrix in  $\mathcal{L}$ . Since the matrix logarithm is the inverse to the matrix exponential, it is a 3-dimensional manifold, contained in the 6-dimensional cone  $\text{int}(\mathbb{S}_+^3)$  of positive definite  $3 \times 3$ -matrices.

Consider the quotient map from the matrix space  $\mathbb{S}^3 \simeq \mathbb{R}^6$  onto  $\mathbb{S}^3/\mathcal{L} \simeq \mathbb{R}^3$ . This takes a positive semidefinite matrix  $X = [x_{ij}]$  to its inner products with the matrices in a basis of  $\mathcal{L}$ :

$$\pi : \mathbb{S}_+^3 \rightarrow \mathbb{R}^3 : X \mapsto (\text{Tr}(X) + 2x_{12}, \text{Tr}(X) + 2x_{13}, \text{Tr}(X) + 2x_{23}).$$

Precisely this map appeared in the statistical study of Gaussian models in [SU10, Example 1.1]. The fibers  $\pi^{-1}(b)$  are three-dimensional spectrahedra, and these serve as feasible regions in optimization, both for semidefinite programming and for maximum likelihood estimation.



We here consider yet another convex optimization problem over the spectrahedron  $\pi^{-1}(b)$ , namely maximizing the *von Neumann entropy*  $h(X) = \text{Tr}(X - X \cdot \log(X))$ . This problem has a unique local and global maximum, at the intersection  $\pi^{-1}(b) \cap \text{GM}(\mathcal{L})$ . See Theorem 2.4.1. This *Gibbs point* is the maximizer of the entropy over the spectrahedron. Therefore, the Gibbs manifold  $\text{GM}(\mathcal{L})$  is the set of Gibbs points in all fibers  $\pi^{-1}(b)$ , as  $b$  ranges over  $\mathbb{R}^3$ .

To study these objects algebraically, we ask for the polynomials that vanish on  $\text{GM}(\mathcal{L})$ . The zeros of these polynomials form the *Gibbs variety*  $\text{GV}(\mathcal{L})$ . Thus, the Gibbs variety is the Zariski closure of the Gibbs manifold. In our example, the Gibbs manifold has dimension 3, whereas the Gibbs variety has dimension 5. The latter is the cubic hypersurface  $\text{GV}(\mathcal{L}) =$

$$\{X \in \mathbb{S}^3 : (x_{11} - x_{22})(x_{11} - x_{33})(x_{22} - x_{33}) = x_{33}(x_{13}^2 - x_{23}^2) + x_{22}(x_{23}^2 - x_{12}^2) + x_{11}(x_{12}^2 - x_{13}^2)\}.$$

As promised, the study of Gibbs manifolds and Gibbs varieties is a non-commutative extension of applied toric geometry. Indeed, every toric variety is a Gibbs variety arising from diagonal matrices. For instance, the toric surface  $\{x \in \mathbb{R}^3 : x_1 x_3 = x_2^2\}$  is realized as

$$\text{GV}(\mathcal{L}') = \{X \in \mathbb{S}^3 : x_{11}x_{33} - x_{22}^2 = x_{12} = x_{13} = x_{23} = 0\}$$

for the diagonal matrix pencil

$$\mathcal{L}' = \left\{ \begin{bmatrix} 2y_1 & 0 & 0 \\ 0 & y_1 + y_2 & 0 \\ 0 & 0 & 2y_2 \end{bmatrix} : y_1, y_2 \in \mathbb{R} \right\}. \quad (2.0.2)$$

However, even for diagonal matrices, the dimension of the Gibbs variety can exceed that of the Gibbs manifold. To see this, replace the matrix entry  $2y_1$  by  $\sqrt{2}y_1$  in the definition of  $\mathcal{L}'$ . This explains why transcendental number theory will make an appearance in our study.

Our presentation in this chapter is organized as follows. Section 2.1 gives a more thorough introduction to Gibbs manifolds and Gibbs varieties. Theorem 2.1.6 states that the dimension of the Gibbs variety is usually quite small and Theorem 2.1.9 gives a formula for this dimension. The proofs of these results are presented in Section 2.2. In that section we present algorithms for computing the prime ideal of the Gibbs variety. This is an implicitization problem, where the parametrization uses transcendental functions. We compare exact symbolic methods for solving that problem with a numerical approach. A key ingredient is the Galois group for the eigenvalues of a linear space of symmetric matrices. We implemented our algorithms in *Julia*, making use of the computer algebra package `Oscar.jl` [OSC24]. Code and data for this chapter are available at <https://mathrepo.mis.mpg.de/GibbsManifolds>.

In Section 2.3 we study the Gibbs varieties given by two-dimensional spaces of symmetric matrices. This rests on the classical Segre-Kronecker classification of matrix pencils [FMS21].

In Section 2.4 we turn to the application that led us to study Gibbs manifolds, namely *entropic regularization* in convex optimization. That section develops the natural generalization of the geometric results in [STVvR24] from linear programming to semidefinite programming. In Section 2.5 we study *quantum optimal transport* [CEFZ22]. This is the semidefinite programming analogue to the classical optimal transport problem [STVvR24, Section 3]. We show that its Gibbs manifold is the positive part of a Segre variety in matrix space. We conclude in Section 2.6 with a discussion of logarithmic sparsity. A matrix is called logarithmically sparse if its matrix logarithm has many zeros, and once a sparsity pattern is fixed, such matrices form a Gibbs manifold. We study these special Gibbs manifolds and explore a connection to graph theory.

## 2.1. What are Gibbs manifolds and varieties?

We write  $\mathbb{S}^n$  for the space of symmetric  $n \times n$ -matrices. This is a real vector space of dimension  $\binom{n+1}{2}$ . The subset of positive semidefinite matrices is denoted  $\mathbb{S}_+^n$ . This is a full-dimensional closed semi-algebraic convex cone in  $\mathbb{S}^n$ , known as the *PSD cone*. The PSD cone is self-dual with respect to the trace inner product, given by  $\langle X, Y \rangle := \text{Tr}(XY)$  for  $X, Y \in \mathbb{S}^n$ .

The matrix exponential function is defined by the usual power series, which converges for all real and complex  $n \times n$ -matrices. It maps symmetric matrices to positive definite symmetric matrices. The zero matrix  $0_n$  is mapped to the identity matrix  $\text{id}_n$ . We write

$$\exp : \mathbb{S}^n \rightarrow \text{int}(\mathbb{S}_+^n), \quad X \mapsto \sum_{i=0}^{\infty} \frac{1}{i!} X^i.$$

This map is invertible, with the inverse given by the familiar series for the logarithm:

$$\log : \text{int}(\mathbb{S}_+^n) \rightarrow \mathbb{S}^n, \quad Y \mapsto \sum_{j=1}^{\infty} \frac{(-1)^{j-1}}{j} (Y - \text{id}_n)^j.$$

We next introduce the geometric objects studied in this chapter. We fix any matrix  $A_0 \in \mathbb{S}^n$  and  $d$  linearly independent matrices  $A_1, A_2, \dots, A_d$ , also in  $\mathbb{S}^n$ . We write  $\mathcal{L}$  for the affine subspace  $A_0 + \text{span}_{\mathbb{R}}(A_1, \dots, A_d)$  of the vector space  $\mathbb{S}^n \simeq \mathbb{R}^{\binom{n+1}{2}}$ . Thus,  $\mathcal{L}$  is an *affine space of symmetric matrices* (ASSM) of dimension  $d$ . If  $A_0 = 0$ , then  $\mathcal{L}$  is a *linear space of symmetric matrices* (LSSM). We are interested in the image of  $\mathcal{L}$  under the exponential map:

**Definition 2.1.1** (Gibbs manifold). The *Gibbs manifold*  $\text{GM}(\mathcal{L})$  of  $\mathcal{L}$  is the  $d$ -dimensional manifold  $\exp(\mathcal{L}) \subset \mathbb{S}_+^n$ .

This is indeed a  $d$ -dimensional manifold inside the convex cone  $\mathbb{S}_+^n$ . It is diffeomorphic to  $\mathbb{R}^d$ , with the identification given by the exponential map and the logarithm map.

In notable special cases (e.g. that in Section 2.5), the Gibbs manifold is semi-algebraic, namely it is the intersection of an algebraic variety with the PSD cone. However, this fails in general, as seen above. It is still interesting to ask which polynomial relations hold between the entries of any matrix in  $\text{GM}(\mathcal{L})$ . This motivates the following definition.

**Definition 2.1.2** (Gibbs variety). The *Gibbs variety*  $\text{GV}(\mathcal{L})$  of  $\mathcal{L}$  is the Zariski closure of  $\text{GM}(\mathcal{L})$  in  $\mathbb{C}^{\binom{n+1}{2}}$ .

**Example 2.1.3** ( $n = 4, d = 2$ ). Consider the 2-dimensional linear space of symmetric matrices

$$\mathcal{L} = \left\{ \begin{bmatrix} 0 & 0 & 0 & y_1 \\ 0 & 0 & y_1 & y_2 \\ 0 & y_1 & y_2 & 0 \\ y_1 & y_2 & 0 & 0 \end{bmatrix} : y_1, y_2 \in \mathbb{R} \right\} \subset \mathbb{S}^4.$$

Its Gibbs manifold  $\text{GM}(\mathcal{L})$  is a surface in  $\mathbb{S}^4 \simeq \mathbb{R}^{10}$ . The Gibbs variety  $\text{GV}(\mathcal{L})$  has dimension five and degree three. It consists of all symmetric matrices  $X = (x_{ij})$  whose entries satisfy

$$\begin{aligned} x_{13} - x_{22} + x_{44} &= x_{14} - x_{23} + x_{34} = x_{24} - x_{33} + x_{44} = 0, \\ \text{and} \quad \text{rank} \begin{bmatrix} x_{11} - x_{44} & x_{12} - x_{34} & x_{22} - x_{33} \\ x_{12} & x_{22} - x_{44} & x_{23} - x_{34} \end{bmatrix} &\leq 1. \end{aligned} \tag{2.1.1}$$

This follows from the general result on matrix pencils in Theorem 2.3.4. ◆

We now investigate various properties of Gibbs varieties. We begin with an amusing fact that if  $\mathcal{L}$  is an LSSM with some additional symmetries, the Gibbs variety respects them.

We consider the tuple of variables  $\mathbf{x} = \{x_{ij} : 1 \leq i \leq j \leq n\}$ . An element  $\sigma$  of the symmetric group  $S_n$  acts on the polynomial ring  $\mathbb{C}[\mathbf{x}]$  by sending  $x_{ij}$  to  $x_{\sigma(i)\sigma(j)}$  for  $1 \leq i \leq j \leq n$  (we identify the variables  $x_{ij}$  and  $x_{ji}$ ). We will also consider the action of  $S_n$  on  $\mathbb{S}^n$  by simultaneously permuting rows and columns of a matrix.

**Proposition 2.1.4.** Let  $\mathcal{L}$  be an LSSM of  $n \times n$  matrices that is invariant under the action of  $\sigma \in S_n$  as a set, i.e.  $\sigma(\mathcal{L}) = \mathcal{L}$ . Then the ideal  $I(\text{GV}(\mathcal{L}))$  of the corresponding Gibbs variety is also invariant under the action of  $\sigma$ .

*Proof.* To prove the Proposition, it suffices to show that if  $B \in \mathcal{L}$  is obtained from  $A \in \mathcal{L}$  by simultaneously permuting rows and columns, then  $\exp(B)$  is obtained from  $\exp(A)$  in the same way. Since  $\exp(B)$  is a formal power series in  $B$ , it suffices to show that  $B^k$  is obtained from  $A^k$  by simultaneously permuting rows and columns for any non-negative integer  $k$ . The latter fact immediately follows from the matrix multiplication formula.  $\square$

Note that Proposition 2.1.4 is equivalent to the well-known fact that  $\exp(P^{-1}AP) = P^{-1}\exp(A)P$  for a permutation (and, more generally, orthogonal) matrix  $P$ . In Section 2.6, we use Proposition 2.1.4 to show that ideals of Gibbs varieties of sparse LSSMs defined by graphs are invariant under permutations of variables induced by graph automorphisms.

**Example 2.1.5.** Consider the LSSM

$$\mathcal{L} = \left\{ \begin{bmatrix} y_1 + y_2 + y_3 & y_1 & y_2 \\ y_1 & y_1 + y_2 + y_3 & y_3 \\ y_2 & y_3 & y_1 + y_2 + y_3 \end{bmatrix} : y_1, y_2, y_3 \in \mathbb{R} \right\}.$$

This linear space is spanned by the matrices

$$A_1 = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

The transposition  $\sigma = (12) \in S_3$  acts on  $\mathbb{S}^3$  in the following way:

$$\begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{12} & x_{22} & x_{23} \\ x_{13} & x_{23} & x_{33} \end{bmatrix} \mapsto \begin{bmatrix} x_{22} & x_{12} & x_{23} \\ x_{12} & x_{11} & x_{13} \\ x_{23} & x_{13} & x_{33} \end{bmatrix}.$$

This action restricts to a linear automorphism of  $\mathcal{L}$  defined by sending  $A_2$  to  $A_3$  and  $A_3$  to  $A_2$ , while leaving  $A_1$  intact. The Gibbs variety of  $\mathcal{L}$  is a hypersurface in  $\mathbb{C}^6$  whose prime ideal is generated by a single polynomial

$$p(x_{11}, x_{12}, x_{13}, x_{22}, x_{23}, x_{33}) = (x_{11} - x_{22})(x_{11} - x_{33})(x_{22} - x_{33}) - x_{33}(x_{13}^2 - x_{23}^2) - x_{22}(x_{23}^2 - x_{12}^2) - x_{11}(x_{12}^2 - x_{13}^2).$$

The action of  $\sigma$  on  $\mathbb{C}[x_{11}, x_{12}, x_{13}, x_{22}, x_{23}, x_{33}]$  sends  $p$  to  $-p$  and preserves the ideal.  $\blacklozenge$

The following dimension bounds constitute one of our main results on Gibbs varieties.

**Theorem 2.1.6.** Let  $\mathcal{L} \subset \mathbb{S}^n$  be an ASSM of dimension  $d$ . The dimension of the Gibbs variety  $\text{GV}(\mathcal{L})$  is at most  $n + d$ . If  $A_0 = 0$ , i.e.  $\mathcal{L}$  is an LSSM, then  $\dim \text{GV}(\mathcal{L})$  is at most  $n + d - 1$ .

These bounds are often attained, e.g. in Example 2.1.3. Our proof of Theorem 2.1.6 appears in Section 2.2, in the context of algorithms for computing the ideal of  $\text{GV}(\mathcal{L})$ .

It turns out that when  $\mathcal{L}$  is an LSSM, Theorem 2.1.6, extremely simple in its statement, can be refined into an exact formula for the dimension of the Gibbs variety. This refinement rests on the notions of eigenvalues of an ASSM  $\mathcal{L}$  and of the  $\mathcal{L}$ -centralizer of  $A \in \mathcal{L}$ .

**Definition 2.1.7 (Centralizer).** Let  $A$  be an  $n \times n$ -matrix, and  $\mathcal{L}$  be an LSSM of  $n \times n$ -matrices. The *centralizer*  $C(A)$  of  $A$  is the set of all matrices that commute with  $A$ . The  $\mathcal{L}$ -*centralizer*  $C_{\mathcal{L}}(A)$  of  $A$  is  $C(A) \cap \mathcal{L}$ .

Note that  $C_{\mathcal{L}}(A)$  is a linear subspace of  $\mathcal{L}$ . Its dimension is independent of  $A$  for a generic choice of  $A \in \mathcal{L}$  (see Section 2.6 for an explanation). This generic dimension will be denoted by  $k$  in Theorem 2.1.9.

Since any ASSM can be written in the form  $\mathcal{L} = \{A_0 + y_1 A_1 + \cdots + y_d A_d : y_i \in \mathbb{R}\}$ , it can be identified with a single matrix with entries in the field  $\mathbb{R}(y_1, \dots, y_d)$  of rational functions in the variables  $y_1, \dots, y_d$ . The eigenvalues of this matrix are algebraic functions in  $y_1, \dots, y_d$  (that is, elements of the algebraic closure  $\overline{\mathbb{R}(y_1, \dots, y_d)}$ ) and will be referred to as the eigenvalues of the corresponding ASSM  $\mathcal{L}$ .

**Example 2.1.8.** Consider the LSSM spanned by  $A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$  and  $A_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ . It is identified with the matrix  $\begin{bmatrix} y_1 & y_2 \\ y_2 & y_1 \end{bmatrix}$  over  $\mathbb{R}(y_1, y_2)$ . Its eigenvalues are  $y_1 - y_2$  and  $y_1 + y_2$ .  $\blacklozenge$

**Theorem 2.1.9.** Let  $\mathcal{L}$  be an LSSM of  $n \times n$ -matrices of dimension  $d$ . Assume that  $\mathcal{L}$  has distinct eigenvalues. Let  $k$  be the dimension of the  $\mathcal{L}$ -centralizer of a generic element in  $\mathcal{L}$  and  $m$  the dimension of the  $\mathbb{Q}$ -linear space spanned by the eigenvalues of  $\mathcal{L}$ . Then  $\dim \text{GV}(\mathcal{L}) = m + d - k$ .

The proof of Theorem 2.1.9 is also presented in Section 2.2. We now give a degree bound for the Gibbs variety of an LSSM  $\mathcal{L}$ , whose proof is contained in Section 2.2 as well. In what follows,  $\mathcal{V}(I)$  denotes the affine variety (over  $\mathbb{C}$ ) defined by the polynomial ideal  $I$ .

**Proposition 2.1.10.** Let  $\mathcal{L}$  be an LSSM of  $n \times n$ -matrices with  $\mathbb{Q}$ -linearly independent eigenvalues. Then  $\deg \text{GV}(\mathcal{L}) \leq n^{\binom{n+1}{2} + 2n}$ .

Example 2.1.5 shows that this degree bound can be quite pessimistic: the actual degree of the Gibbs variety is 3, while Proposition 2.1.10 gives  $3^{12}$ .

While it might be difficult to find all polynomials that vanish on the Gibbs manifold, finding *linear* relations is sometimes easier. Such relations are useful for semidefinite optimization, see Remark 2.4.3. This brings us to the final geometric object of this chapter.

**Definition 2.1.11 (Gibbs plane).** The *Gibbs plane*  $\text{GP}(\mathcal{L})$  is the smallest affine space containing  $\text{GV}(\mathcal{L})$ .

Clearly, we have the chain of inclusions  $\text{GM}(\mathcal{L}) \subseteq \text{GV}(\mathcal{L}) \subseteq \text{GP}(\mathcal{L}) \subseteq \mathbb{C}^{\binom{n+1}{2}}$ .

**Example 2.1.12.** The Gibbs plane of the LSSM  $\mathcal{L}$  from Example 2.1.3 is the 7-dimensional linear space in  $\mathbb{C}^{10}$  that is defined by the linear relations listed in the first row of (2.1.1).  $\blacklozenge$

We claimed that this chapter offers a generalization of toric varieties. In what follows, we make that claim precise, by discussing the case when  $\mathcal{L}$  is a commuting family. This means that the symmetric matrices  $A_0, A_1, \dots, A_d$  commute pairwise, i.e.  $A_i A_j = A_j A_i$  for all  $i, j$ . We now assume that this holds. Then the ASSM  $\mathcal{L}$  can be diagonalized, i.e. there is an orthogonal matrix  $V$  such that  $\Lambda_i = V^T A_i V$  is a diagonal matrix, for all  $i$ . The vector  $\lambda_i \in \mathbb{R}^n$  of diagonal entries in  $\Lambda_i = \text{diag}(\lambda_i)$  contains the eigenvalues of  $A_i$ .

The matrix exponential of any element in  $\mathcal{L}$  can be computed as follows:

$$\exp(A_0 + y_1 A_1 + \cdots + y_d A_d) = V \cdot \exp(\Lambda_0 + y_1 \Lambda_1 + \cdots + y_d \Lambda_d) \cdot V^T. \quad (2.1.2)$$

Let  $\mathcal{D}$  denote this ASSM of diagonal matrices, i.e.  $\mathcal{D} = \{\Lambda_0 + y_1 \Lambda_1 + \cdots + y_d \Lambda_d : y \in \mathbb{R}^d\}$ . Then the linear change of coordinates given by  $V$  identifies the respective Gibbs manifolds:

$$\text{GM}(\mathcal{L}) = V \cdot \text{GM}(\mathcal{D}) \cdot V^T. \quad (2.1.3)$$

The same statement holds for the Gibbs varieties and the Gibbs planes:

$$\text{GV}(\mathcal{L}) = V \cdot \text{GV}(\mathcal{D}) \cdot V^T \quad \text{and} \quad \text{GP}(\mathcal{L}) = V \cdot \text{GP}(\mathcal{D}) \cdot V^T. \quad (2.1.4)$$

The dimensions of these objects are determined by arithmetic properties of the eigenvalues.

Recall that  $\Lambda_i = \text{diag}(\lambda_i)$  where  $\lambda_i$  is a vector in  $\mathbb{R}^n$ . Let  $\Lambda$  denote the linear subspace of  $\mathbb{R}^n$  that is spanned by the  $d$  vectors  $\lambda_1, \dots, \lambda_d$ . We have  $\mathcal{D} = \lambda_0 + \Lambda$ , and therefore

$$\text{GM}(\mathcal{D}) = \exp(\lambda_0) \star \exp(\Lambda) = \{(e^{\lambda_{01}} w_1, \dots, e^{\lambda_{0n}} w_n) : w \in \exp(\Lambda)\} \subset \mathbb{R}^n.$$

Here  $\star$  denotes coordinate-wise multiplication in  $\mathbb{R}^n$ . Let  $\Lambda_{\mathbb{Q}}$  be the smallest vector subspace of  $\mathbb{R}^n$  spanned by elements of  $\mathbb{Q}^n$  which contains  $\Lambda$ . Its dimension  $d_{\mathbb{Q}} = \dim \Lambda_{\mathbb{Q}}$  satisfies  $d \leq d_{\mathbb{Q}} \leq n$ . Fix lattice vectors  $a_1, a_2, \dots, a_{d_{\mathbb{Q}}}$  in  $\mathbb{Z}^n$  that form a basis of  $\Lambda_{\mathbb{Q}}$ . Then, inside the  $n$ -dimensional linear space defined by the diagonality condition, we have

$$\text{GV}(\mathcal{D}) = \overline{\left\{ \left( e^{\lambda_{01}} \prod_{i=1}^{d_{\mathbb{Q}}} z_i^{a_{i1}}, e^{\lambda_{02}} \prod_{i=1}^{d_{\mathbb{Q}}} z_i^{a_{i2}}, \dots, e^{\lambda_{0n}} \prod_{i=1}^{d_{\mathbb{Q}}} z_i^{a_{in}} \right) : z \in (\mathbb{C}^*)^{d_{\mathbb{Q}}} \right\}}. \quad (2.1.5)$$

This is a toric variety of dimension  $d_{\mathbb{Q}}$ . The Zariski closure is taken in  $\mathbb{C}^n$ . The Gibbs manifold  $\text{GM}(\mathcal{D})$  is a  $d$ -dimensional subset of the real points in  $\text{GV}(\mathcal{D})$  for which  $z$  has strictly positive coordinates. We summarize our discussion in the following theorem.

**Theorem 2.1.13.** Let  $\mathcal{L}$  be an affine space of pairwise commuting symmetric matrices. Then the Gibbs variety  $\text{GV}(\mathcal{L})$  is a toric variety of dimension  $d_{\mathbb{Q}}$ , given by (2.1.4) and (2.1.5).

For an illustration, consider the seemingly simple case  $d = 1$  and  $A_0 = 0$ . Here,  $\text{GM}(\mathcal{L})$  is the curve formed by all powers of  $\exp(A_1)$ , and  $\text{GV}(\mathcal{L})$  is a toric variety of generally higher dimension. This scenario is reminiscent of that studied in [GS21].

**Example 2.1.14.** Let  $n = 3$  and consider the LSSM  $\mathcal{L}$  spanned by  $A_1 = \begin{bmatrix} 4 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{bmatrix}$ . We have

$$A_1 = V \cdot \text{diag}(\lambda) \cdot V^T, \quad \text{where} \quad \lambda = (2, 4 + \sqrt{2}, 4 - \sqrt{2}) \quad \text{and} \quad V = \frac{1}{2} \begin{bmatrix} 0 & \sqrt{2} & -\sqrt{2} \\ -\sqrt{2} & 1 & 1 \\ \sqrt{2} & 1 & 1 \end{bmatrix}.$$

Here,  $\mathcal{D} = \Lambda = \mathbb{R}\lambda$ ,  $d_{\mathbb{Q}} = 2$ , and  $\Lambda_{\mathbb{Q}} = \mathbb{R}\{(1, 2, 2), (0, 1, -1)\} = \{p \in \mathbb{R}^3 : 4p_1 = p_2 + p_3\}$ . Hence  $\text{GV}(\mathcal{D})$  is the toric surface  $\{q_{11}^4 = q_{22}q_{33}\}$  in  $\text{GP}(\mathcal{D}) = \{Q \in \mathbb{S}^3 : q_{12} = q_{13} = q_{23} = 0\}$ . We transform that surface into the original coordinates via (2.1.4). The computation reveals

$$\text{GV}(\mathcal{L}) = \{X \in \text{GP}(\mathcal{L}) : x_{23}^4 - 4x_{23}^3 x_{33} + 6x_{23}^2 x_{33}^2 - 4x_{23} x_{33}^3 + x_{33}^4 + 2x_{13}^2 - x_{23}^2 - 2x_{23} x_{33} - x_{33}^2 = 0\}.$$

The ambient 3-space is  $\text{GP}(\mathcal{L}) = \{X \in \mathbb{S}^3 : x_{11} - x_{23} - x_{33} = x_{12} - x_{13} = x_{22} - x_{33} = 0\}$ .  $\blacklozenge$

This concludes our discussion of the toric Gibbs varieties arising from pairwise commuting matrices. In the next section we turn to the general case, which requires new ideas.

## 2.2. Implicitization of Gibbs varieties

Implicitization is the computational problem of finding implicit equations for an object that comes in the form of a parametrization. When the parametrizing functions are rational functions, these equations are polynomials and can be found using resultants or Gröbner bases [MS21a, Section 4.2]. A different approach rests on polynomial interpolation and numerical nonlinear algebra. This section studies implicitization of Gibbs varieties. The difficulty here is in the fact that Gibbs manifolds are transcendental, since their parametrizations involve the exponential function. We start out by presenting our proof of Theorem 2.1.6.

As in Section 2.1,  $\mathcal{L} = A_0 + \text{span}_{\mathbb{R}}(A_1, \dots, A_d)$  is a  $d$ -dimensional affine space of symmetric  $n \times n$ -matrices. Its elements are  $A_0 + y_1 A_1 + \dots + y_d A_d$ . We shall parametrize the Gibbs manifold  $\text{GM}(\mathcal{L})$  in terms of the coordinates  $y_1, \dots, y_d$  on  $\mathcal{L}$ . This uses the following formula.

**Theorem 2.2.1** (Sylvester [Sy183]). Let  $f : D \rightarrow \mathbb{R}$  be an analytic function on an open set  $D \subset \mathbb{R}$  and  $M \in \mathbb{R}^{n \times n}$  a matrix that has  $n$  distinct eigenvalues  $\lambda_1, \dots, \lambda_n$  in  $D$ . Then

$$f(M) = \sum_{i=1}^n f(\lambda_i) M_i, \quad \text{with} \quad M_i = \prod_{j \neq i} \frac{1}{\lambda_i - \lambda_j} (M - \lambda_j \cdot \text{id}_n).$$

Note that the product on the right hand side is taken in the commutative matrix ring  $\mathbb{R}[M]$ .

*Proof of Theorem 2.1.6.* The characteristic polynomial of  $A(y) = A_0 + y_1 A_1 + \dots + y_d A_d$  is

$$P_{\mathcal{L}}(\lambda; y) = \det(A(y) - \lambda \cdot \text{id}_n) = c_0(y) + c_1(y) \lambda + \dots + c_{n-1}(y) \lambda^{n-1} + (-1)^n \lambda^n.$$

Its zeros  $\lambda$  are algebraic functions of the coordinates  $y = (y_1, \dots, y_d)$  on  $\mathcal{L}$ .

We first assume that  $\mathcal{L}$  has distinct eigenvalues, i.e. there is a Zariski open subset  $U \subset \mathbb{R}^d$  such that  $P_{\mathcal{L}}(\lambda; y^*)$  has  $n$  distinct real roots  $\lambda$  for all  $y^* \in U$ . Sylvester's formula writes the entries of  $\exp(A(y))$  as rational functions of  $y$ ,  $\lambda_i(y)$  and  $e^{\lambda_i(y)}$  for  $y \in U$ . These functions are multisymmetric in the pairs  $(\lambda_i, e^{\lambda_i})$ . They evaluate to convergent power series on  $\mathbb{R}^d$ .

Let  $V$  be the subvariety of  $U \times \mathbb{R}^n$  that is defined by the equations

$$c_i(y) = (-1)^i \sigma_{n-i}(\lambda) \quad \text{for} \quad i = 0, \dots, n-1, \quad (2.2.1)$$

where  $\sigma_i(\lambda)$  is the  $i^{\text{th}}$  elementary symmetric polynomial evaluated at  $(\lambda_1, \dots, \lambda_n)$ . We have  $\dim V = d$ . Define a map  $\phi : V \times \mathbb{R}^n \rightarrow \mathbb{S}^n$ , using coordinates  $z_1, \dots, z_n$  on  $\mathbb{R}^n$ , as follows:

$$(y_1, \dots, y_d, \lambda_1, \dots, \lambda_n, z_1, \dots, z_n) \mapsto \sum_{i=1}^n z_i \prod_{j \neq i} \frac{1}{\lambda_i - \lambda_j} (A(y) - \lambda_j \cdot \text{id}_n). \quad (2.2.2)$$

The closure  $\overline{\phi(V \times \mathbb{R}^n)}$  of the image of this map is a variety. It contains the Gibbs variety: setting  $z_i = e^{\lambda_i}$  parametrizes a dense subset of the Gibbs manifold, by Theorem 2.2.1.

The Gibbs variety of the LSSM  $\mathbb{R}\mathcal{L}$  spanned by the ASSM  $\mathcal{L}$  also lies in  $\overline{\phi(V \times \mathbb{R}^n)}$ , because  $\exp(y_0 A(y)) = \phi(y_0 \cdot y, y_0 \cdot \lambda, e^{y_0 \cdot \lambda})$  for any  $y \in U$  and  $y_0 \in \mathbb{R} \setminus \{0\}$ . We thus have

$$\dim \text{GV}(\mathcal{L}) \leq \dim \overline{\phi(V \times \mathbb{R}^n)} \leq d + n \quad \text{and} \quad \dim \text{GV}(\mathbb{R}\mathcal{L}) \leq d + n.$$

Finally, suppose that  $\mathcal{L}$  is an LSSM, i.e.  $A_0 = 0$ . Then  $\mathcal{L}$  is the linear span of an ASSM of dimension  $d-1$  in  $\mathbb{S}^n$ . The second inequality therefore gives  $\dim \text{GV}(\mathcal{L}) \leq d + n - 1$ .

We finally consider the case when  $\mathcal{L}$  has  $m < n$  distinct eigenvalues. Since symmetric matrices are diagonalizable, Sylvester's formula can easily be adapted to this case: it suffices to sum over the distinct eigenvalues of  $M$ , and to adjust the parametrization (2.2.2) accordingly. That is, we replace  $n$  by  $m$ . See [HJ94, Chapter 6.1, Problem 14] for details.  $\square$

**Remark 2.2.2.** If the points  $\exp(\lambda(y)) = (e^{\lambda_1(y)}, \dots, e^{\lambda_n(y)})$ ,  $y \in U$ , lie in a lower-dimensional subvariety  $W \subset \mathbb{R}^n$ , then the proof of Theorem 2.1.6 gives the better bound  $\dim \text{GV}(\mathcal{L}) \leq d + \dim W$ . We saw this in Example 2.1.14. In general, no such subvariety  $W$  exists, i.e. one expects  $W = \mathbb{R}^n$ . This is an issue of Galois theory, to be discussed at the end of this section.

For ease of exposition, we work only with LSSMs in the rest of this section. That is, we set  $A_0 = 0$ . We comment on the generalization to ASSMs in Remark 2.2.7. Our discussion and the proof of Theorem 2.1.6 suggest Algorithm 1, for computing the ideal of the Gibbs variety of an LSSM  $\mathcal{L}$ . That ideal lives in a polynomial ring  $\mathbb{R}[\mathbf{x}]$  whose variables are the

---

**Algorithm 1** Implicitization of the Gibbs variety of an LSSM  $\mathcal{L}$ , defined over  $\mathbb{Q}$

---

**Input:** Linearly independent matrices  $A_1, \dots, A_d \in \mathbb{S}^n$  with rational entries

**Output:** Polynomials that define  $\text{GV}(\mathcal{L})$ , where  $\mathcal{L} = \text{span}_{\mathbb{R}}(A_1, \dots, A_d)$

1: Compute the characteristic polynomial  $P_{\mathcal{L}}(\lambda; y) = c_0(y) + c_1(y)\lambda + \dots + c_n(y)\lambda^n$

**Require:**  $P_{\mathcal{L}}(\lambda; y)$  has  $n$  distinct roots in  $\overline{\mathbb{R}(y)}$

2:  $E'_1 \leftarrow \{\text{the } n \text{ polynomials } (-1)^i \sigma_{n-i}(\lambda) - c_i(y) \text{ in (2.2.1)}\}$

3:  $E_1 \leftarrow \{\text{generators of any associated prime over } \mathbb{Q} \text{ of } \langle E'_1 \rangle\}$

4:  $E_2 \leftarrow \{\text{the entries of } \phi(y, \lambda, z) - X\}$ , with  $X = (x_{ij})$  a symmetric matrix of variables

5:  $E_2, D \leftarrow$  clear denominators in  $E_2$  and record the least common denominator  $D$

6: **if** the roots  $\lambda_1, \dots, \lambda_n$  of  $P_{\mathcal{L}}(\lambda; y)$  are  $\mathbb{Q}$ -linearly dependent **then**

7:  $E_3 \leftarrow \{z^\alpha - z^\beta : \sum \alpha_i \lambda_i = \sum \beta_j \lambda_j, \alpha, \beta \in \mathbb{Z}_{\geq 0}^n\}$

8: **else**

9:  $E_3 \leftarrow \emptyset$

10:  $I \leftarrow$  the ideal generated by  $E_1, E_2, E_3$  in the polynomial ring  $\mathbb{R}[y, \lambda, z, \mathbf{x}]$

11:  $I \leftarrow I : D^\infty$

12:  $J \leftarrow$  elimination ideal obtained by eliminating  $y, \lambda, z$  from  $I$

13: **return** a set of generators of  $J$

---

entries of a symmetric  $n \times n$ -matrix  $X$ . The algorithm builds three subsets  $E_1, E_2, E_3$  of the larger polynomial ring  $\mathbb{R}[y, \lambda, z, \mathbf{x}]$ . After the saturation (step 11), the auxiliary variables  $y, \lambda, z$  are eliminated. The equations  $E'_1$  come from (2.2.1). They constrain  $(y, \lambda)$  to lie in  $V$ . The set  $E_1$  generates an associated prime of  $\langle E'_1 \rangle$  (step 3), see the discussion preceding Theorem 2.2.6. The equations  $E_2$  come from the parametrization (2.2.2). Note that, if  $\mathcal{L}$  has  $m < n$  distinct eigenvalues, this formula can be adjusted as in the end of the proof of Theorem 2.1.6, and the requirement after step 1 can be dropped. Later in the algorithm, one replaces  $n$  with  $m$ . It is necessary to clear denominators in order to obtain polynomials (step 5). The saturation by  $D$  avoids spurious components arising from this step. Finally,  $E_3$  accounts for toric relations between the  $z_i$  arising from  $\mathbb{Q}$ -linear relations among the  $\lambda_i$ . If no such relations exist, Theorem 2.2.5 ensures that the assignment  $E_3 \leftarrow \emptyset$  in step 9 is correct.

Steps 6 and 7 in Algorithm 1 require a detailed discussion. Further below we shall explain the  $\mathbb{Q}$ -linear independence of eigenvalues, how to check this, and how to compute  $E_3$ . Ignoring this for now, one can also run Algorithm 1 with  $E_3 = \emptyset$ . Then step 13 still returns polynomials that vanish on the Gibbs variety  $\text{GV}(\mathcal{L})$  but these may cut out a larger variety.

We implemented Algorithm 1 in Julia (v1.8.3), using `Oscar.jl` [OSC24], and tested it on many examples. The code is available at <https://mathrepo.mis.mpg.de/GibbsManifolds>.

**Example 2.2.3.** The Gibbs variety  $\text{GV}(\mathcal{L})$  for the LSSM  $\mathcal{L}$  in (2.0.1) has the parametrization

$$\phi = \sum_{i=1}^3 \frac{z_i}{q(\lambda_i, y_1, y_2, y_3)} \begin{bmatrix} p_{11}(\lambda_i, y_1, y_2, y_3) & p_{12}(\lambda_i, y_1, y_2, y_3) & p_{13}(\lambda_i, y_1, y_2, y_3) \\ p_{12}(\lambda_i, y_1, y_2, y_3) & p_{22}(\lambda_i, y_1, y_2, y_3) & p_{23}(\lambda_i, y_1, y_2, y_3) \\ p_{13}(\lambda_i, y_1, y_2, y_3) & p_{23}(\lambda_i, y_1, y_2, y_3) & p_{33}(\lambda_i, y_1, y_2, y_3) \end{bmatrix}, \text{ where}$$

$$\begin{aligned}
q &= 2y_1^2 + 6y_1y_2 + 2y_2^2 + 6y_1y_3 + 6y_2y_3 + 2y_3^2 - 6y_1\lambda - 6y_2\lambda - 6y_3\lambda + 3\lambda^2, \\
p_{11} &= y_1^2 + 2y_1y_2 + y_2^2 + 2y_1y_3 + 2y_2y_3 - 2y_1\lambda - 2y_2\lambda - 2y_3\lambda + \lambda^2, \\
p_{12} &= -y_1^2 - y_1y_2 - y_1y_3 + y_2y_3 + y_1\lambda, \\
p_{13} &= -y_1y_2 - y_2^2 + y_1y_3 - y_2y_3 + y_2\lambda, \\
p_{22} &= y_1^2 + 2y_1y_2 + 2y_1y_3 + 2y_2y_3 + y_3^2 - 2y_1\lambda - 2y_2\lambda - 2y_3\lambda + \lambda^2, \\
p_{23} &= y_1y_2 - y_1y_3 - y_2y_3 - y_3^2 + y_3\lambda, \\
p_{33} &= 2y_1y_2 + y_2^2 + 2y_1y_3 + 2y_2y_3 + y_3^2 - 2y_1\lambda - 2y_2\lambda - 2y_3\lambda + \lambda^2.
\end{aligned}$$

Our `Julia` code for Algorithm 1 easily finds the cubic polynomial defining  $\text{GV}(\mathcal{L})$ .  $\blacklozenge$

In spite of such successes, symbolic implicitization is limited to small  $n$  and  $d$ . Numerical computations can help, in some cases, to find equations for more challenging Gibbs varieties.

**Example 2.2.4.** We consider the LSSM of  $4 \times 4$  Hankel matrices with upper left entry zero:

$$\mathcal{L} = \left\{ \begin{bmatrix} 0 & y_2 & y_3 & y_4 \\ y_2 & y_3 & y_4 & y_5 \\ y_3 & y_4 & y_5 & y_6 \\ y_4 & y_5 & y_6 & y_7 \end{bmatrix} : (y_2, \dots, y_7) \in \mathbb{R}^6 \right\}.$$

Algorithm 1 failed to compute its Gibbs variety. We proceed using numerics as follows. Fix a degree  $D > 0$  and let  $N = \binom{9+D}{D}$  be the number of monomials in the 10 coordinates  $x_{11}, \dots, x_{44}$  on  $\mathbb{S}^4$ . We create  $M \geq N$  samples on  $\text{GM}(\mathcal{L})$  by plugging in random values for the six parameters  $y_i$  and applying the matrix exponential. Finding all vanishing equations of degree  $D$  on these samples amounts to computing the kernel of an  $M \times N$  Vandermonde matrix. If this matrix has full rank, then there are no relations of degree  $D$ . We implemented this procedure in `Julia`. In our example, Theorem 2.1.6 says that  $\text{GV}(\mathcal{L})$  is contained in a hypersurface. Using our numerical method, we find one defining equation of degree  $D = 6$ . We used  $M = 5205 \geq N = 5005$  samples. Our sextic has 853 terms with integer coefficients:

$$x_{11}^3 x_{22} x_{24} x_{34} - x_{11}^3 x_{23}^2 x_{34} - x_{11}^3 x_{23} x_{24}^2 + \dots + 3x_{23} x_{24}^2 x_{33} x_{34}^2 + x_{24}^4 x_{33} x_{34} - x_{24}^3 x_{33}^2 x_{34}.$$

Its Newton polytope has the f-vector  $(456, 5538, 21560, 41172, 44707, 29088, 11236, 2370, 211)$ . In fact, the Gibbs variety in this example is precisely the hypersurface defined by this sextic. This follows from Theorem 2.1.9.

Note that the package `Oscar.jl` conveniently allows to perform symbolic and numerical implicitization and polyhedral computations in the same programming environment.

We emphasize that our numerical `Julia` code is set up to find *exact* integer coefficients. For this, we first normalize the numerical approximation of the coefficient vector by setting its first (numerically) nonzero entry to one. Then we rationalize the coefficients using the built in command `rationalize` in `Julia`, with error tolerance `tol = 1e-7`. Correctness of the result is proved by checking that the resulting polynomial vanishes on the parametrization.  $\blacklozenge$

We now turn to  $\mathbb{Q}$ -linear relations among eigenvalues of  $\mathcal{L}$ . Our arithmetic discussion begins with a version of [Ax71, (SP)], which is well-known in transcendental number theory.

**Theorem 2.2.5 (Ax-Schanuel).** If the eigenvalues  $\lambda_1, \dots, \lambda_n$  of the LSSM  $\mathcal{L}$  are  $\mathbb{Q}$ -linearly independent, then  $e^{\lambda_1}, \dots, e^{\lambda_n}$  are algebraically independent over the field  $\mathbb{C}(y_1, \dots, y_d)$ .

On the other hand, suppose that the eigenvalues of  $\mathcal{L}$  satisfy some non-trivial linear relation over  $\mathbb{Q}$ . We can then find nonnegative integers  $\alpha_i$  and  $\beta_j$ , not all zero, such that

$$\sum_{i=1}^n \alpha_i \lambda_i = \sum_{j=1}^n \beta_j \lambda_j. \tag{2.2.3}$$

This implies that the exponentials of the eigenvalues satisfy the toric relations

$$\prod_{i=1}^n z_i^{\alpha_i} = \prod_{j=1}^n z_j^{\beta_j}. \tag{2.2.4}$$



The linear relations (2.2.3) can be found from the ideal  $\langle E'_1 \rangle$  in step 2 which specifies that the  $\lambda_i$  are the eigenvalues of  $A(y)$ . This ideal is radical if we assume that  $\mathcal{L}$  has distinct eigenvalues. We compute the prime decomposition of the ideal over  $\mathbb{Q}$ . All prime components are equivalent under permuting the  $\lambda_i$ , so we replace  $\langle E'_1 \rangle$  by any of these prime ideals in step 3. We compute (2.2.3) as the linear forms in that prime ideal. Using (2.2.4), we compute the toric ideal  $\langle E_3 \rangle$  in step 7, which is also prime. This ideal defines a toric variety  $W'$ , whose  $S_n$ -orbit is the variety  $W$  in Remark 2.2.2. We arrive at the following result.

**Theorem 2.2.6.** Let  $\mathcal{L} \subset \mathbb{S}^n$  be an LSSM with distinct eigenvalues. The Gibbs variety  $\text{GV}(\mathcal{L})$  is irreducible and unirational, and the ideal  $J$  found in Algorithm 1 is its prime ideal.

*Proof.* Sylvester's formula yields a rational parametrization  $\psi$  of  $\text{GV}(\mathcal{L})$  with parameters  $y_1, \dots, y_d, z_1, \dots, z_n$ . The parameters  $\lambda_i$  in (2.2.2) can be omitted: the entries in the image are multisymmetric in  $(\lambda_i, z_i)$ , so that they can be expressed in terms of elementary symmetric polynomials of the  $\lambda_i$  [Bri04, Theorem 1]. The point  $(z_1, \dots, z_n)$  lies on the toric variety  $W'$  defined above. The domain  $\mathbb{C}^d \times W'$  of  $\psi$  is an irreducible variety, and it is also rational. The image of  $\psi$  is the Gibbs variety  $\text{GV}(\mathcal{L})$ , which is therefore unirational and irreducible. The ideals given by  $E_1$  and  $E_2$  in Algorithm 1 are prime, after saturation, and elimination in step 12 preserves primality. Hence the output in  $J$  in step 13 is the desired prime ideal.  $\square$

We are now ready to prove Theorem 2.1.9 and Proposition 2.1.10.

*Proof of Theorem 2.1.9.* The dimension of a generic fiber of the map  $\phi$  from the proof of Theorem 2.1.6 is equal to the dimension of the centralizer of a generic element in this fiber, i.e. to  $k$ . This is explained in Section 2.3, in the proof of Theorem 2.3.6. Note that  $\phi : V \times \mathbb{R}^n \rightarrow \mathbb{S}^n$  restricts to a dominant map from  $V \times W'$  to  $\text{GV}(\mathcal{L})$ , where  $W'$  is the variety from Theorem 2.2.6. This domain is irreducible, being the product of irreducible varieties, and has dimension  $m + d$ .  $\text{GV}(\mathcal{L})$  is also irreducible by Theorem 2.2.6. Thus, by fiber dimension theorem [Har13, Exercise II.3.22],  $\dim \text{GV}(\mathcal{L}) = m + d - k$ .  $\square$

*Proof of Proposition 2.1.10.* By Algorithm 1, the prime ideal  $J$  of  $\text{GV}(\mathcal{L})$  is obtained by elimination from the ideal  $I \subseteq \mathbb{C}[x_{ij}, \lambda_i, z_i, y_i]$  of a polynomial ring in  $\binom{n+1}{2} + 2n + d$  variables generated by polynomials of degree at most  $n$ . By [MS21a, Theorem 4.2], the variety  $\mathcal{V}(J)$  is a projection of  $\mathcal{V}(I)$  and therefore  $\deg \mathcal{V}(J) \leq \deg \mathcal{V}(I)$ . Therefore,  $\deg \text{GV}(\mathcal{L}) = \deg \mathcal{V}(J) \leq \deg \mathcal{V}(I)$ . The variety  $\mathcal{V}(I)$  lives in the affine space of dimension  $\binom{n+1}{2} + 2n + d$ , where  $d = \dim \mathcal{L}$ . Note that  $\dim \mathcal{V}(I) \geq \dim \mathcal{L}$  and thus  $\text{codim } \mathcal{V}(I) \leq \binom{n+1}{2} + 2n$ . Therefore, by Bézout's theorem [Ful98, Theorem 12.3], we have  $\deg \mathcal{V}(I) \leq n^{\binom{n+1}{2} + 2n}$ .  $\square$

Once the degree of the Gibbs variety is known, one can use numerical techniques to find its defining equations. In general, this allows to compute defining equations (in a set-theoretic sense) of Gibbs varieties that are infeasible for symbolic algorithms, as seen in Example 2.2.4. This is due to the fact that any variety of degree  $d$  can be set-theoretically defined by equations of degree at most  $d$ . Such a description is given by the *Chow equations* of a variety [DS95, Proposition 3.1].

We now present Algorithm 2 for finding the equations of the Gibbs variety numerically. This is based on [BKS18, Chapter 5]. We write  $\langle P \rangle$  for the ideal generated by  $P \subseteq \mathbb{C}[\mathbf{x}]$ .

Correctness of Algorithm 2 is ensured by the genericity condition imposed on the samples picked in Step 3. Unfortunately, the degree upper bound in Proposition 2.1.10 restricts the practical applicability of this algorithm to  $n \leq 3$ . However, if the Gibbs variety is a hypersurface, then the algorithm can terminate immediately after finding a single algebraic equation. The degree of this equation is usually much lower than the degree bound in

---

**Algorithm 2** Numerical implicitization of Gibbs varieties of known degree

---

**Input:** An LSSM  $\mathcal{L}$  given as an  $\mathbb{R}$ -span of  $d$  linearly independent matrices  $A_1, \dots, A_d$ , degree  $k$  of  $\text{GV}(\mathcal{L})$ ;

**Output:** A set of equations that define  $\text{GV}(\mathcal{L})$  set-theoretically.

**Require:**  $\mathcal{L}$  has  $\mathbb{Q}$ -linearly independent eigenvalues.

- 1:  $I \leftarrow \{0\}$ ,  $l \leftarrow 1$ ,  $N \leftarrow \binom{n+1}{2}$
  - 2: **for**  $l = 1$  to  $k$  **do**
  - 3:     Pick  $M > \binom{N+l-1}{l}$  random samples in  $\mathcal{L}$  such that they are not contained in a lower-dimensional subspace of  $\mathcal{L}$
  - 4:     Let  $E$  be the set of matrix exponentials of the  $M$  picked samples
  - 5:     Construct a Vandermonde matrix  $A$  by evaluating all monomials of degree  $l$  on the elements of  $E$
  - 6:     Let  $I_l$  be the basis of  $\ker A$
  - 7:      $I \leftarrow \langle I \cup I_l \rangle$
- return** a set of generators of  $I$ .
- 

Proposition 2.1.10 (as seen in Example 2.1.5) and therefore the defining equation can be found with this algorithm for larger  $n$ . This was the strategy used in Example 2.2.4.

As promised, we conclude this section with explaining the role of Galois theory in symbolic implicitization of Gibbs varieties. We define the *Galois group*  $G_{\mathcal{L}}$  of an LSSM  $\mathcal{L}$  to be the Galois group of the characteristic polynomial  $P_{\mathcal{L}}(\lambda, y)$  over the field  $\mathbb{Q}(y_1, \dots, y_d)$ . Note that  $G_{\mathcal{L}}$  is the subgroup of the symmetric group  $S_n$  whose elements are permutations that fix each associated prime of  $\langle E'_1 \rangle$ . Hence the index of the Galois group  $G_{\mathcal{L}}$  in  $S_n$  is the number of associated primes. In particular, the Galois group equals  $S_n$  if and only if the ideal  $\langle E'_1 \rangle$  formed in step 2 of Algorithm 1 is prime.

The existence of linear relations (2.2.3) depends on the Galois group  $G_{\mathcal{L}}$ . If the Galois group is small then the primes of  $\langle E_1 \rangle$  are large, and more likely to contain linear forms. There is a substantial literature in number theory on this topic. See [Gir82, Gir99] and the references therein. For instance, by Kitaoka [Kit17, Proposition 2], there are no linear relations if  $n$  is prime, or if  $n \geq 6$  and the Galois group is  $S_n$  or  $A_n$ . If this holds,  $E_3 = \emptyset$  in step 9 of Algorithm 1.

The computation of Galois groups is a well-studied topic in symbolic computation and number theory. Especially promising are methods based on numerical algebraic geometry (e.g. in [HRS18]). These fit well with the approach to implicitization in Example 2.2.4. A very interesting research question is to classify LSSMs by their Galois groups.

**Remark 2.2.7.** We briefly comment on how to adjust Algorithm 1 to compute the Gibbs variety of an ASSM  $\mathcal{L}$  with  $A_0 \neq 0$ . In this case, algebraic relations between  $e^{\lambda_1}, \dots, e^{\lambda_n}$  come from  $\mathbb{Q}$ -linear relations between the eigenvalues of  $\mathcal{L}$ , but this time modulo  $\mathbb{C}$ : an affine relation  $\sum \alpha_i \lambda_i = \sum \beta_j \lambda_j + \gamma$  gives  $z^\alpha - e^\gamma \cdot z^\beta = 0$ , where  $z_i = e^{\lambda_i}$ ,  $\alpha_i, \beta_j \in \mathbb{Z}_{\geq 0}$ ,  $\gamma \in \mathbb{C}$ . Here  $\gamma$  is a  $\mathbb{Q}$ -linear combination of eigenvalues of  $A_0$ . Theorem 2.2.6 holds for ASSMs as well, provided that these  $\mathbb{Q}$ -linear relations modulo  $\mathbb{C}$  can be computed in practice. This can usually not be done over  $\mathbb{Q}$ . We leave this algorithmic challenge for future research.

## 2.3. Pencils of quadrics

In this section we study the Gibbs variety  $\text{GV}(\mathcal{L})$  where  $\mathcal{L} \subset \mathbb{S}^n$  is a pencil of quadrics, i.e. an LSSM of dimension  $d = 2$ . We follow the exposition in [FMS21], where pencils  $\mathcal{L}$  are classified

by Segre symbols. The *Segre symbol*  $\sigma = \sigma(\mathcal{L})$  is a multiset of partitions that sum up to  $n$ . It is computed as follows. Pick a basis  $\{A_1, A_2\}$  of  $\mathcal{L}$ , where  $A_2$  is invertible, and find the Jordan canonical form of  $A_1 A_2^{-1}$ . Each eigenvalue determines a partition, according to the sizes of the corresponding Jordan blocks. The multiset of these partitions is the Segre symbol  $\sigma$ .

We use the canonical form in [FMS21, Section 2]. Suppose the Segre symbol is  $\sigma = [\sigma_1, \dots, \sigma_r]$ , where the  $i$ th partition  $\sigma_i$  equals  $(\sigma_{i,1} \geq \sigma_{i,2} \geq \dots \geq \sigma_{i,n} \geq 0)$ . There are  $r$  groups of blocks, one for each eigenvalue  $\alpha_i$  of  $A_1 A_2^{-1}$ . The  $j$ -th matrix in the  $i$ -th group is the  $\sigma_{i,j} \times \sigma_{i,j}$ -matrix

$$y_1 \cdot \begin{bmatrix} 0 & 0 & \dots & 0 & \alpha_i \\ 0 & 0 & \dots & \alpha_i & 1 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \alpha_i & 1 & \ddots & 0 \\ \alpha_i & 1 & \dots & 0 & 0 \end{bmatrix} + y_2 \cdot \begin{bmatrix} 0 & \dots & 0 & 0 & 1 \\ 0 & \dots & 0 & 1 & 0 \\ 0 & \dots & 1 & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 1 & \dots & 0 & 0 & 0 \end{bmatrix}.$$

There are 13 Segre symbols for  $n = 4$ ; see [FMS21, Example 3.1]. It is instructive to compute their Gibbs varieties. All possible dimensions, 2, 3, 4 and 5, are attained. Dimension 2 arises for the diagonal pencil  $\mathcal{L}_\sigma = \text{diag}(\alpha_1 y_1 + y_2, \alpha_2 y_1 + y_2, \alpha_3 y_1 + y_2, \alpha_4 y_1 + y_2)$ , with Segre symbol  $\sigma = [1, 1, 1, 1]$ . When the  $\alpha_i$  are distinct integers,  $\text{GV}(\mathcal{L}_\sigma) = \text{GM}(\mathcal{L}_\sigma)$  is a toric surface. This is similar to (2.0.2). Dimension 5 arises for  $\sigma = [4]$ , see Example 2.1.3.

The following examples, also computed with Algorithm 1, exhibit the dimensions 5, 4, 3.

**Example 2.3.1.** Consider the Segre symbol  $\sigma = [3, 1]$ . The canonical pencil  $\mathcal{L}_{[3,1]}$  is spanned by

$$\begin{bmatrix} 0 & 0 & \alpha_1 & 0 \\ 0 & \alpha_1 & 1 & 0 \\ \alpha_1 & 1 & 0 & 0 \\ 0 & 0 & 0 & \alpha_2 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \text{for } \alpha_1, \alpha_2 \in \mathbb{R} \text{ distinct.}$$

Here,  $\dim \text{GV}(\mathcal{L}_{[3,1]}) = 5$ , the upper bound in Theorem 2.1.6. Algorithm 1 produces the ideal

$$J = \langle x_{14}, x_{24}, x_{34}, x_{13} - x_{22} + x_{33}, x_{12}^2 - x_{11}x_{22} - x_{12}x_{23} + x_{11}x_{33} + x_{22}x_{33} - x_{33}^2 \rangle.$$

If  $\alpha_1 = \alpha_2$ , then the Segre symbol changes to  $\sigma = [(3, 1)]$ . We now find the additional cubic

$$x_{11}x_{22}x_{33} + 2x_{12}x_{13}x_{23} - x_{13}^2x_{22} - x_{11}x_{23}^2 - x_{12}^2x_{33} - x_{44} \in J. \quad (2.3.1)$$

This cuts down the dimension by one, and we now have  $\dim \text{GV}(\mathcal{L}_{[(3,1)]}) = 4$ .  $\blacklozenge$

**Example 2.3.2.** Consider the Segre symbol  $\sigma = [(2, 2)]$ . The pencil  $\mathcal{L}_{[(2,2)]}$  is spanned by

$$\begin{bmatrix} 0 & \alpha & 0 & 0 \\ \alpha & 1 & 0 & 0 \\ 0 & 0 & 0 & \alpha \\ 0 & 0 & \alpha & 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad \text{for some } \alpha \in \mathbb{R}.$$

A version of Algorithm 1 for LSSMs with multiple eigenvalues produces the ideal

$$J = \langle x_{11} - x_{33}, x_{12} - x_{34}, x_{22} - x_{44}, x_{13}, x_{14}, x_{23}, x_{24} \rangle.$$

The Gibbs variety  $\text{GV}(\mathcal{L}_{[(2,2)]})$  is 3-dimensional and equals the Gibbs plane  $\text{GP}(\mathcal{L}_{[(2,2)]})$ .  $\blacklozenge$

The cubic (2.3.1) which distinguishes the Segre symbols  $[3, 1]$  and  $[(3, 1)]$  is explained by the following result. This applies not just to pencils but to all ASSMs with block structure.

**Proposition 2.3.3.** Let  $\mathcal{L}$  be a block-diagonal ASSM with  $r$  blocks  $X_i(y)$  of size  $\tau_i$ , where  $\tau_1 + \dots + \tau_r = n$ . The Gibbs plane  $\text{GP}(\mathcal{L})$  is contained in  $\mathbb{S}^{\tau_1} \times \dots \times \mathbb{S}^{\tau_r} \subset \mathbb{S}^n$ . Moreover, with the notation  $\mathcal{J} = \{\{i, j\} \in \binom{[r]}{2} : \text{Tr}(X_i(y)) = \text{Tr}(X_j(y))\}$ , we have

$$\text{GV}(\mathcal{L}) \subseteq \{(X_1, \dots, X_r) \in \text{GP}(\mathcal{L}) : \det(X_i) = \det(X_j) \text{ for all } \{i, j\} \in \mathcal{J}\}.$$

*Proof.* Block-diagonal matrices are exponentiated block-wise. The entries outside the diagonal blocks are zero. The statement follows from  $\det(\exp(X_i(y))) = \exp(\text{Tr}(X_i(y)))$ .  $\square$

Proposition 2.3.3 holds for the canonical pencil  $\mathcal{L}_\sigma$  of any Segre symbol  $\sigma$ . First of all, for all indices  $(i, j)$  outside the diagonal blocks, we have  $x_{ij} = 0$  on the Gibbs plane  $\text{GP}(\mathcal{L}_\sigma)$ . Next, one has equations for the exponential of a single block, like those in Theorem 2.3.4 below. Finally, there are equations that link the blocks corresponding to entries  $\sigma_{ij}$  of the same partition  $\sigma_i$ . Some of these come from trace equalities between blocks of  $\mathcal{L}_\sigma$ , and this is the scope of Proposition 2.3.3. In particular, blocks  $ij$  and  $ik$  for which  $\sigma_{ij} = \sigma_{ik} \pmod{2}$  exponentiate to  $X_{ij} \in \mathbb{S}_+^{\sigma_{ij}}$  and  $X_{ik} \in \mathbb{S}_+^{\sigma_{ik}}$  with equal determinant. We saw this in (2.3.1). In all examples we computed, the three classes of equations above determine the Gibbs variety.

We now derive the equations that hold for the exponential of a single block. To this end, we fix  $\sigma = [n]$  with  $\alpha_1 = 0$ . The canonical LSSM  $\mathcal{L}_{[n]}$  consists of the symmetric matrices

$$Y = \begin{bmatrix} 0 & 0 & \dots & 0 & y_1 \\ 0 & 0 & \dots & y_1 & y_2 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & y_1 & y_2 & \vdots & 0 \\ y_1 & y_2 & \vdots & 0 & 0 \end{bmatrix}.$$

The case  $n = 4$  was featured in Example 2.1.3. In what follows we generalize that example.

**Theorem 2.3.4.** The following linear equations hold on the Gibbs plane  $\text{GP}(\mathcal{L}_{[n]})$ :

$$x_{i-1,j} + x_{i+1,j} = x_{i,j-1} + x_{i,j+1} \quad \text{for } 2 \leq i < j \leq n. \quad (2.3.2)$$

The  $2 \times 2$ -minors of the following  $2 \times (n-1)$ -matrix vanish on the Gibbs variety  $\text{GV}(\mathcal{L}_{[n]})$ :

$$D(X) = \begin{bmatrix} x_{11} & x_{12} & x_{22} & \dots \\ x_{12} & x_{22} & x_{23} & \dots \end{bmatrix} - \begin{bmatrix} x_{n,n} & x_{n-1,n} & x_{n-1,n-1} & \dots \\ 0 & x_{n,n} & x_{n-1,n} & \dots \end{bmatrix}. \quad (2.3.3)$$

If the Galois group  $G_{\mathcal{L}_{[n]}}$  is the symmetric group  $S_n$ , then the prime ideal of  $\text{GV}(\mathcal{L}_{[n]})$  is generated by (2.3.2) and (2.3.3), and we have  $\dim \text{GP}(\mathcal{L}_n) = 2n - 1$  and  $\dim \text{GV}(\mathcal{L}_{[n]}) = n + 1$ .

**Remark 2.3.5.** We conjecture that  $G_{\mathcal{L}_{[n]}} = S_n$ . This was verified computationally for many values of  $n$ , but we currently do not have a proof that works for all  $n$ . This gap underscores the need, pointed out at the end of Section 2.2, for a study of the Galois groups of LSSMs.

*Proof.* We claim that the linear equations (2.3.2) hold for every non-negative integer power of  $Y$ . This implies that they hold for  $\exp(Y)$ . We will show this by induction. The equations clearly hold for  $Y^0 = \text{id}_n$ . Suppose they hold for  $(m_{ij}) = M = Y^k$ . Write  $(b_{ij}) = B := Y^{k+1} = MY$ .

The two-banded structure of  $Y$  implies  $b_{i,j} = y_1 \cdot m_{i,n-j+1} + y_2 \cdot m_{i,n-j+2}$  for  $1 \leq i < j$ . The following identity holds for  $2 \leq i < j$ , and it shows that  $\exp(Y)$  satisfies the equations (2.3.2):

$$\begin{aligned}
b_{i-1,j} - b_{i,j-1} - b_{i,j+1} + b_{i+1,j} &= y_1 \cdot m_{i-1,n-j+1} + y_2 \cdot m_{i-1,n-j+2} - y_1 \cdot m_{i,n-j+2} - \\
&\quad y_2 \cdot m_{i,n-j+3} - y_1 \cdot m_{i,n-j} - y_2 \cdot m_{i,n-j+1} + y_1 \cdot m_{i+1,n-j+1} + y_2 \cdot m_{i+1,n-j+2} \\
&= y_1 \cdot (m_{i-1,n-j+1} - m_{i,n-j+2} - m_{i,n-j} + m_{i+1,n-j+1}) + \\
&\quad y_2 \cdot (m_{i-1,n-j+2} - m_{i,n-j+3} - m_{i,n-j+1} + m_{i+1,n-j+2}) = 0.
\end{aligned}$$

We next consider the matrix  $D(X)$  in (2.3.3). We must show that  $D(X)$  has rank  $\leq 1$  for  $X \in \text{GV}(\mathcal{L}_{[n]})$ . We claim that the rows of  $D(Y^k)$  are proportional with the same coefficient for all  $k \in \mathbb{Z}_{\geq 0}$ . This will imply that the rows of  $D(\exp(Y))$  are proportional. For the proof, let  $\vec{v}_1$  and  $\vec{v}_2$  be the rows of  $D(B)$ , where  $B = Y^k$ . We will show that  $y_1 \vec{v}_1 + y_2 \vec{v}_2 = 0$ .

First note that  $D(\text{id}_n) = 0$ . Also note that each column of  $D(B)$  has the form

$$\begin{bmatrix} b_{i,i} - b_{n+1-i,n+1-i} \\ b_{i,i+1} - b_{n+1-i,n+2-i} \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} b_{i,i+1} - b_{n-i,n+1-i} \\ b_{i+1,i+1} - b_{n-i+1,n-i+1} \end{bmatrix}.$$

We start with the left case. We must show  $y_1(b_{i,i} - b_{n+1-i,n+1-i}) + y_2(b_{i,i+1} - b_{n+1-i,n+2-i}) = 0$ .

Recall from above that  $b_{i,j} = y_1 \cdot m_{i,n-j+1} + y_2 \cdot m_{i,n-j+2}$ , where  $(m_{i,j}) = M = Y^{k-1}$  for  $i < j$ . Using this and the fact that the powers of  $Y$  are symmetric, we write

$$\begin{aligned}
y_1(b_{i,i} - b_{n+1-i,n+1-i}) + y_2(b_{i,i+1} - b_{n+1-i,n+2-i}) &= \\
&= y_1((y_1 \cdot m_{i,n-i+1} + y_2 \cdot m_{i,n-i+2}) - (y_1 \cdot m_{n+1-i,i} + y_2 \cdot m_{n+1-i,i+1})) + \\
&\quad y_2((y_1 \cdot m_{i,n-i} + y_2 \cdot m_{i,n-i+1}) - (y_1 \cdot m_{n+1-i,i-1} + y_2 \cdot m_{n+1-i,i})) \\
&= y_1 y_2 (m_{i,n-i+2} - m_{i+1,n+1-i} + m_{i,n-i} - m_{n+1-i,i-1}) = 0,
\end{aligned}$$

where the last equality follows from (2.3.2). Now, for the second case we have

$$\begin{aligned}
y_1(b_{i,i+1} - b_{n-i,n+1-i}) + y_2(b_{i+1,i+1} - b_{n-i+1,n-i+1}) &= \\
&= y_1(y_1 \cdot m_{i,n-i} + y_2 \cdot m_{i,n-i+1} - y_1 \cdot m_{n-i,i} - y_2 \cdot m_{n-i,i+1}) + \\
&\quad y_2(y_1 \cdot m_{i+1,n-i} + y_2 \cdot m_{i+1,n-i+1} - y_1 \cdot m_{n-i+1,i} - y_2 \cdot m_{n-i+1,i+1}) = 0.
\end{aligned}$$

This proves that the  $2 \times 2$ -minors of  $D(X)$  vanish on the Gibbs variety  $\text{GV}(\mathcal{L}_{[n]})$ .

Suppose now that the eigenvalues of  $Y$  are  $\mathbb{Q}$ -linearly independent. We can check this directly for  $n \leq 5$ . For  $n \geq 6$  it follows from our hypothesis  $G_{\mathcal{L}_{[n]}} = S_n$ , by [Kit17, Proposition 2]. That hypothesis implies  $\dim \text{GV}(\mathcal{L}_{[n]}) = n + 1$ , by Theorems 2.1.6 and 2.2.5.

For the primality statement, we note that the matrix  $D(X)$  is 1-generic in the sense of Eisenbud [Eis87, Eis88]. By [Eis87, Theorem 1], the  $2 \times 2$ -minors of  $D(X)$  generate a prime ideal of codimension  $n - 2$  in the coordinates of the  $(2n - 1)$ -dimensional space given by (2.3.2). The equality of dimensions yields  $\dim \text{GP}(\mathcal{L}_{[n]}) = 2n - 1$ , and we conclude that our linear and quadratic constraints generate the prime ideal of  $\text{GV}(\mathcal{L}_{[n]})$ .  $\square$

Theorem 2.1.6 and its refinement in Remark 2.2.2 furnish an upper bound on the dimension of any Gibbs variety. This raises the question when this bound is attained. In general, this question is answered by Theorem 2.1.9. In what follows, we offer a more concrete complete answer for  $d = 2$ . Let  $\mathcal{L}$  be a pencil with eigenvalues  $\lambda_i(y)$ , and let  $W$  denote the Zariski closure in  $\mathbb{R}^n$  of the set of points  $\exp(\lambda(y)) = (e^{\lambda_1(y)}, \dots, e^{\lambda_n(y)})$ ,  $y \in \mathbb{R}^2$ .

**Theorem 2.3.6.** Let  $\mathcal{L} = \text{span}_{\mathbb{R}}(A_1, A_2)$ , where  $A_1 A_2 \neq A_2 A_1$ . Then  $\dim \text{GV}(\mathcal{L}) = \dim(W) + 1$ . In particular, if the Galois group  $G_{\mathcal{L}}$  is the symmetric group  $S_n$  then  $\dim \text{GV}(\mathcal{L}) = n + 1$ .

*Proof.* We claim that the fibers of the map  $\phi : V \times W \rightarrow \mathbb{S}^n$  defined by (2.2.2) are one-dimensional. Let  $B \in \phi(V \times W)$  and consider any point  $p = (y_1, y_2, \lambda_1, \dots, \lambda_n, z_1, \dots, z_n) \in \phi^{-1}(B)$ . The condition that  $p$  lies in the fiber  $\phi^{-1}(B)$  is equivalent to

- (1)  $z_1, \dots, z_n$  are the eigenvalues of  $B$ , and
- (2)  $X = y_1 A_1 + y_2 A_2$  and  $B$  have the same eigenvectors, and
- (3)  $\lambda_1, \dots, \lambda_n$  are the eigenvalues of  $X$ .

Condition (1) follows from Theorem 2.2.1 for  $f = \exp$ . It implies that there are only finitely many possibilities for the  $z$ -coordinates of the point  $p$  in the fiber: up to permutations, they are the eigenvalues of  $B$ . Condition (3) follows from  $(y_1, y_2, \lambda_1, \dots, \lambda_n) \in V$ . It says that the  $\lambda$ -coordinates are determined, up to permutation, by  $y_1, y_2$ . Therefore, it suffices to show that the matrices in  $\mathcal{L}$  whose eigenvectors are those of  $B$  form a one-dimensional subvariety.

Symmetric matrices have common eigenvectors if and only if they commute. Define  $S = \{X = y_1 A_1 + y_2 A_2 \in \mathcal{L} : X \cdot B = B \cdot X\} \subset \mathcal{L}$ . This is a pairwise commuting linear subspace. Note that  $S$  contains a nonzero matrix  $X$ , since there is a point in  $\phi^{-1}(B)$  whose  $y$ -coordinates define a nonzero matrix in  $\mathcal{L}$ . Therefore  $\dim S \geq 1$ . Since  $A_1 A_2 \neq A_2 A_1$ , we also have  $\dim S \leq 1$ . Hence  $\dim S = \dim \phi^{-1}(B) = 1$  and the upper bound  $\dim W + 1$  for the dimension of  $\text{GV}(\mathcal{L})$ , which is given by Remark 2.2.2, is attained in our situation.  $\square$

## 2.4. Role in convex optimization

In this section we show how Gibbs manifolds arise from entropic regularization in optimization (cf. [STVvR24]). We fix an arbitrary linear map  $\pi : \mathbb{S}^n \rightarrow \mathbb{R}^d$ . This can be written in the form

$$\pi(X) = (\langle A_1, X \rangle, \langle A_2, X \rangle, \dots, \langle A_d, X \rangle).$$

Here the  $A_i \in \mathbb{S}^n$ , and  $\langle A_i, X \rangle := \text{Tr}(A_i X)$ . The image  $\pi(\mathbb{S}_+^n)$  of the PSD cone  $\mathbb{S}_+^n$  under our linear map  $\pi$  is a *spectrahedral shadow*. Hence it is a full-dimensional semi-algebraic convex cone in  $\mathbb{R}^d$ . Interestingly,  $\pi(\mathbb{S}_+^n)$  can fail to be closed, as explained in [JS21].

*Semidefinite programming (SDP)* is the following convex optimization problem:

$$\text{Minimize } \langle C, X \rangle \quad \text{subject to } X \in \mathbb{S}_+^n \text{ and } \pi(X) = b. \quad (2.4.1)$$

See Section 1.4 or [MS21a, Chapter 12] for more details. The instance (2.4.1) is specified by the cost matrix  $C \in \mathbb{S}^n$  and the right hand side vector  $b \in \mathbb{R}^d$ . The feasible region  $\mathbb{S}_+^n \cap \pi^{-1}(b)$  is a *spectrahedron*. The SDP problem (2.4.1) is feasible, i.e. the spectrahedron is non-empty, if and only if  $b$  is in  $\pi(\mathbb{S}_+^n)$ .

Consider the LSSM  $\mathcal{L} = \text{span}_{\mathbb{R}}(A_1, \dots, A_d)$ . We usually assume that  $\mathcal{L}$  contains a positive definite matrix. This hypothesis ensures that each spectrahedron  $\pi^{-1}(b)$  is compact.

As an extension of [STVvR24, eqn (2)], we now define the entropic regularization of SDP:

$$\text{Minimize } \langle C, X \rangle - \epsilon \cdot h(X) \quad \text{subject to } X \in \mathbb{S}_+^n \text{ and } \pi(X) = b. \quad (2.4.2)$$

Here  $\epsilon > 0$  is a small parameter, and  $h$  denotes the *von Neumann entropy*, here defined as

$$h : \mathbb{S}_+^n \rightarrow \mathbb{R}, \quad X \mapsto \text{Tr}(X - X \cdot \log(X)).$$

We note that  $h$  is invariant under the action of the orthogonal group on  $\mathbb{S}_+^n$ . This implies that  $h(X) = \sum_{i=1}^n (\lambda_i - \lambda_i \log(\lambda_i))$ , where  $\lambda_1, \dots, \lambda_n$  are the eigenvalues of  $X$ . Hence the von Neumann entropy  $h$  is the matrix version of the entropy function on  $\mathbb{R}_+^n$  used in [STVvR24].

Our next result makes the role of Gibbs manifolds in semidefinite programming explicit. The following ASSM is obtained by incorporating  $\epsilon$  and the cost matrix  $C$  into the LSSM:

$$\mathcal{L}_\epsilon := \mathcal{L} - \frac{1}{\epsilon}C \quad \text{for any } \epsilon > 0.$$

Here we allow the case  $\epsilon = \infty$ , where the dependency on  $C$  disappears and the ASSM is simply the LSSM, i.e.  $\mathcal{L}_\infty = \mathcal{L}$ . The following theorem is the main result in this section.

**Theorem 2.4.1.** For  $b \in \pi(\mathbb{S}_+^n)$ , the intersection of  $\pi^{-1}(b)$  with the Gibbs manifold  $\text{GM}(\mathcal{L}_\epsilon)$  consists of a single point  $X_\epsilon^*$ . This point is the optimal solution to the regularized SDP (2.4.2). For  $\epsilon = \infty$ , it is the unique maximizer of von Neumann entropy on the spectrahedron  $\pi^{-1}(b)$ .

The importance of this result for SDP lies in taking the limit as  $\epsilon$  tends to zero. This limit  $\lim_{\epsilon \rightarrow 0} X_\epsilon^*$  exists and it is an optimal solution to (2.4.1). The optimal solution is unique for generic  $C$ . Entropic regularization is about approximating that limit.

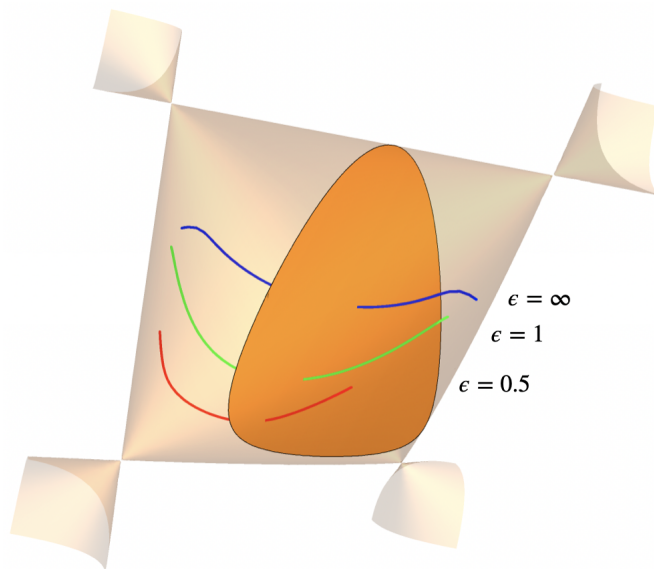
**Example 2.4.2.** Consider the four matrices

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad A_4 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

and the linear constraints  $\langle A_1, X \rangle = \langle A_2, X \rangle = \langle A_3, X \rangle = 1$ ,  $\langle A_4, X \rangle = 0$ . These cut out a two-dimensional spectrahedron inside the six-dimensional cone  $\mathbb{S}_+^3$ . Note that the first three constraints define the familiar ellipsope from Example 1.4.6. The Gibbs manifolds  $\text{GM}(\mathcal{L}_\epsilon)$  for the SDP problem with the cost matrix

$$C = \frac{1}{2} \begin{bmatrix} 0 & -1 & 1 \\ -1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

are four-dimensional. When intersected with the affine space cut out by the first three constraints, they become curves. These curves, along with the corresponding spectrahedron, are shown in Figure 2.1. ◆



**Figure 2.1:** Gibbs manifolds  $\text{GM}(\mathcal{L}_\epsilon)$  and the spectrahedron from Example 2.4.2

**Remark 2.4.3.** Theorem 2.4.1 implies that adding the condition  $X \in \text{GV}(\mathcal{L}_\epsilon)$  to (2.4.2) leaves the optimizer unchanged. Hence, if we know equations for the Gibbs variety, we may shrink the feasible region by adding polynomial constraints. Most practical are the affine-linear equations: imposing  $X \in \text{GP}(\mathcal{L}_\epsilon)$  allows to solve (2.4.2) on a spectrahedron of lower dimension.

To prove Theorem 2.4.1, we derive two key properties of the von Neumann entropy:

**Proposition 2.4.4.** The function  $h$  satisfies:

- (a)  $h$  is strictly concave on the PSD cone  $\mathbb{S}_+^n$ , and
- (b) the gradient of  $h$  is the negative matrix logarithm:  $\nabla(h)(X) = -\log(X)$ .

*Proof.* For (a), we use a classical result by Davis [Dav57]. The function  $h$  is invariant in the sense that its value  $h(X)$  depends on the eigenvalues of  $X$ . In fact, it is a symmetric function of the  $n$  eigenvalues  $\lambda_1, \lambda_2, \dots, \lambda_n$ . This function equals  $h(\lambda_1, \lambda_2, \dots, \lambda_n) = \sum_{i=1}^n (\lambda_i - \lambda_i \log(\lambda_i))$ , and this is a concave function on  $\mathbb{R}_+^n$ . The assertion hence follows from the theorem in [Dav57].

For (b) we prove a more general result. For convenience, we change variables  $Y = X - \text{id}_n$  so that  $f(Y) = h(Y + \text{id}_n)$  is analytic at  $Y = 0$ . Fix any function  $f : \mathbb{R} \rightarrow \mathbb{R}$  that is analytic in a neighborhood of the origin. Then  $Y \mapsto \text{Tr}(f(Y))$  is a well-defined real-valued analytic function of  $n \times n$ -matrices  $Y = (y_{ij})$  that are close to zero. The gradient of this function is the  $n \times n$ -matrix whose entries are the partial derivatives  $\partial \text{trace}(f(Y)) / \partial y_{ij}$ . We claim that

$$\nabla \text{Tr}(f(Y)) = f'(Y^T). \quad (2.4.3)$$

Both sides are linear in  $f$ , and  $f$  is analytic, so it suffices to prove this for monomials, i.e.

$$\nabla \text{Tr}(Y^k) = k \cdot (Y^T)^{k-1} \quad \text{for all integers } k \geq 1. \quad (2.4.4)$$

Note that  $\text{trace}(Y^k)$  is a homogeneous polynomial of degree  $k$  in the matrix entries  $y_{ij}$ , namely it is the sum over all products  $y_{i_1 i_2} y_{i_2 i_3} \cdots y_{i_{k-2} i_{k-1}} y_{i_{k-1} i_1}$  that represent closed walks in the complete graph on  $k$  nodes. When taking the derivative  $\partial / \partial y_{ij}$  of that sum, we obtain  $k$  times the sum over all walks that start at node  $j$  and end at node  $i$ . Here each walk occurs with the factor  $k$  because  $y_{ij}$  can be inserted in  $k$  different ways to create one of the closed walks above. This polynomial of degree  $k-1$  is the entry of the matrix power  $Y^{k-1}$  in row  $j$  and column  $i$ , so it is the entry of its transpose  $(Y^T)^{k-1}$  in row  $i$  and column  $j$ . To prove the proposition, we now apply (2.4.3) to the function  $f(y) = (y+1) - (y+1) \cdot \log(y+1)$ .  $\square$

If  $\mathcal{L} = \mathcal{D}$  consists of diagonal matrices then the Gibbs manifold  $\text{GM}(\mathcal{D})$  is a discrete exponential family [Sul18, §6.2], and  $\pi(\text{GM}(\mathcal{D}))$  is the associated convex polytope. This uses the toric moment map [MS21a, Theorem 8.24]. In particular, if the linear space  $\mathcal{D}$  is defined over  $\mathbb{Q}$  then the polytope is rational and the Zariski closure of  $\text{GM}(\mathcal{D})$  is the toric variety of that polytope. If the space  $\mathcal{D}$  is not defined over  $\mathbb{Q}$  then  $\text{GM}(\mathcal{D})$  is an analytic toric manifold, whose Zariski closure is the larger toric variety  $\text{GV}(\mathcal{D}) = \text{GM}(\mathcal{D}_{\mathbb{Q}})$  seen in (2.1.5).

The key step to proving Theorem 2.4.1 is a non-abelian version of the toric moment map.

**Theorem 2.4.5.** The restriction of the linear map  $\pi : \mathbb{S}_+^n \rightarrow \mathbb{R}^d$  to the Gibbs manifold  $\text{GM}(\mathcal{L})$  defines a bijection between  $\text{GM}(\mathcal{L})$  and the open spectrahedral shadow  $\text{int}(\pi(\mathbb{S}_+^n))$  in  $\mathbb{R}^d$ .

*Proof.* Fix an arbitrary positive definite matrix  $X \in \text{int}(\mathbb{S}_+^n)$  and set  $b = \pi(X)$ . We must show that the spectrahedron  $\pi^{-1}(b)$  contains precisely one point that lies in  $\text{GM}(\mathcal{L})$ .

Consider the restriction of the von Neumann entropy  $h$  to the spectrahedron  $\pi^{-1}(b)$ . This restriction is strictly concave on the convex body  $\pi^{-1}(b)$  by Proposition 2.4.4. Therefore  $h$



attains a unique maximum  $X^*$  in the relative interior of  $\pi^{-1}(b)$ . The first order condition at this maximum tells us that  $\nabla(h)(X^*) = -\log(X^*)$  lies in  $\mathcal{L}$ , which is the span of the gradients of the constraints  $\langle A_i, X \rangle = b_i$ . Hence, the optimal matrix  $X^*$  lies in the Gibbs manifold

$$\text{GM}(\mathcal{L}) = \{X \in \mathbb{S}_+^n : \log(X) \in \mathcal{L}\}.$$

The assignment  $b \mapsto X^* = X^*(b)$  is well defined and continuous on the interior of the cone  $\pi(\mathbb{S}_+^n)$ . We have shown that it is a section of the linear map  $\pi$ , which means  $\pi(X^*(b)) = b$ . It is also surjective onto  $\text{GM}(\mathcal{L})$ , because  $X^*(\pi(X)) = X$ , for  $X \in \text{GM}(\mathcal{L})$ . We conclude that  $\pi$  defines a homeomorphism between  $\text{GM}(\mathcal{L})$  and  $\text{int}(\pi(\mathbb{S}_+^n))$ .  $\square$

*Proof of Theorem 2.4.1.* For any fixed  $\epsilon > 0$ , the minimizer  $X^* = X_\epsilon^*$  of the regularized problem (2.4.2) lies in the interior of the spectrahedron  $\pi^{-1}(b)$ . This is because the gradient of the entropy function diverges at the boundary (Proposition 2.4.4). By the same convexity argument as in the proof of Theorem 2.4.5, the objective function in (2.4.2) has only one critical point  $X^*$  in the spectrahedron  $\pi^{-1}(b)$ . It satisfies the first order optimality conditions, which impose  $C + \epsilon \cdot \log(X^*) \in \mathcal{L}$ . Therefore  $X^* \in \text{GM}(\mathcal{L}_\epsilon)$ , and  $\pi^{-1}(b) \cap \text{GM}(\mathcal{L}_\epsilon) = \{X_\epsilon^*\}$ .  $\square$

We can now turn the discussion around and offer a definition of Gibbs manifolds and Gibbs varieties purely in terms of convex optimization. Fix any LSSM  $\mathcal{L}$  of dimension  $d$  in  $\mathbb{S}^n$ . This defines a canonical linear map  $\pi : \mathbb{S}_+^n \rightarrow \mathbb{S}^n/\mathcal{L}^\perp \simeq \mathbb{R}^d$ . Each fiber  $\pi^{-1}(b)$  is a spectrahedron. If this is non-empty then the entropy  $h(X)$  has a unique maximizer  $X^*(b)$  in  $\pi^{-1}(b)$ . The Gibbs manifold  $\text{GM}(\mathcal{L})$  is the set of these entropy maximizers  $X^*(b)$  for  $b \in \mathbb{R}^d$ . The Gibbs variety  $\text{GV}(\mathcal{L})$  is defined by all polynomial constraints satisfied by these  $X^*(b)$ .

This extends naturally to any ASSM  $A_0 + \mathcal{L}$ . Here we maximize the concave function  $h(X) + \langle A_0, X \rangle$  over the spectrahedra  $\pi^{-1}(b)$ . The Gibbs manifold  $\text{GM}(A_0 + \mathcal{L})$  collects all maximizers, and the Gibbs variety  $\text{GV}(A_0 + \mathcal{L})$  is defined by their polynomial constraints.

**Example 2.4.6.** Let  $\mathcal{L}$  denote the space of all Hankel matrices  $[y_{i+j-1}]_{1 \leq i, j \leq n}$  in  $\mathbb{S}^n$ . This LSSM has dimension  $d = 2n - 1$ . The linear map  $\pi : \mathbb{S}_+^n \rightarrow \mathbb{R}^d$  takes any positive definite matrix  $X$  to a nonnegative polynomial  $b = b(t)$  in one variable  $t$  of degree  $2n - 2$ . We have  $b(t) = (1, t, \dots, t^{n-1})X(1, t, \dots, t^{n-1})^T$ , so the matrix  $X$  gives a sum-of-squares (SOS) representation of  $b(t)$ . The fiber  $\pi^{-1}(b)$  is the *Gram spectrahedron* [Sch22] of the polynomial  $b$ . The entropy maximizer  $X^*(b)$  in the Gram spectrahedron is a favorite SOS representation of  $b$ . The Gibbs manifold  $\text{GM}(\mathcal{L})$  gathers the favorite SOS representations for all non-negative polynomials  $b$ . The Gibbs variety  $\text{GV}(\mathcal{L})$ , which has dimension  $\leq 3n - 2$ , is the tightest outer approximation of  $\text{GM}(\mathcal{L})$  that is definable by polynomials in the matrix entries.

In Example 2.2.4 we saw a variant of  $\mathcal{L}$ , namely the sub-LSSM where the upper left entry of the Hankel matrix was fixed to be zero. If  $C = -E_{11}$  is the corresponding negated matrix unit, then (2.4.1) is the problem of minimizing  $b(t)$  over  $t \in \mathbb{R}$ . See [MS21a, Section 12.3] for a first introduction to polynomial optimization via SOS representations. It would be interesting to explore the potential of entropic regularization for polynomial optimization.  $\blacklozenge$

One of the topics of [STVvR24] was a scaling algorithm for solving the optimization problem (2.4.2) for linear programming (LP), i.e. the case when  $A_1, \dots, A_d$  are diagonal matrices. This algorithm extends the Darroch-Ratcliff algorithm for Iterative Proportional Fitting in statistics. Combining this with a method for driving  $\epsilon$  to zero leads to a numerical algorithm for large-scale LP problems, such as the optimal transport problems in [STVvR24, Section 3].

We are hopeful that the scaling algorithm can be extended to the problem (2.4.2) in full generality. By combining this with a method for driving  $\epsilon$  to zero, one obtains a numerical framework for solving SDP problems such as quantum optimal transport in Section 2.5.

One important geometric object for SDP is the limiting Gibbs manifold,  $\lim_{\epsilon \rightarrow 0} \text{GM}(\mathcal{L}_\epsilon)$ . This is the set of optimal solutions, as  $b$  ranges over  $\mathbb{R}^d$ . In the case of LP, with  $C$  generic, it is the simplicial complex which forms the regular triangulation given by  $C$ . This reveals the combinatorial essence of entropic regularization of LP, see [STVvR24, Theorem 7]. From the perspective of *positive geometry*, it would be worthwhile to study  $\lim_{\epsilon \rightarrow 0} \text{GM}(\mathcal{L}_\epsilon)$  for SDP. This set is semi-algebraic, and it defines a nonlinear subdivision of the spectrahedral shadow  $\pi(\mathbb{S}_+^n)$ . If we vary the cost matrix  $C$ , the theory of *fiber bodies* in [MM23] becomes relevant.

## 2.5. Quantum optimal transport

In this section we examine a semidefinite programming analog of the classical optimal transport problem, known as *quantum optimal transport* (QOT). We follow the presentation by Cole, Eckstein, Friedland, and Życzkowski in [CEFŻ22]. We consider the space  $\mathbb{S}^{d_1 d_2}$  of real symmetric matrices  $X$  of size  $d_1 d_2 \times d_1 d_2$ . Rows and columns are indexed by  $[d_1] \times [d_2]$ . Thus, we write  $X = (x_{ijkl})$ , where  $(i, j)$  and  $(k, l)$  are in  $[d_1] \times [d_2]$ . The matrix being symmetric means that  $x_{ijkl} = x_{klji}$  for all indices. Each such matrix is mapped to a pair of two *partial traces* by the following linear map:

$$\mathbb{S}^{d_1 d_2} \rightarrow \mathbb{S}^{d_1} \times \mathbb{S}^{d_2}, \quad X \mapsto (Y, Z),$$

where the  $d_1 \times d_1$ -matrix  $Y = (y_{ik})$  satisfies  $y_{ik} = \sum_{j=1}^{d_2} x_{ijkj}$ , and the  $d_2 \times d_2$ -matrix  $Z = (z_{jl})$  satisfies  $z_{jl} = \sum_{i=1}^{d_1} x_{ijil}$ . This is done in accordance with Definition 1.2.3. If  $X$  is positive semidefinite then so are its partial traces  $Y$  and  $Z$ . Hence our *marginalization map* restricts to a linear projection of closed convex cones, denoted

$$\mu : \mathbb{S}_+^{d_1 d_2} \rightarrow \mathbb{S}_+^{d_1} \times \mathbb{S}_+^{d_2}, \quad X \mapsto (Y, Z). \quad (2.5.1)$$

Diagonal matrices in  $\mathbb{S}_+^{d_1 d_2}$  can be identified with rectangular matrices of format  $d_1 \times d_2$  whose entries are nonnegative. The map  $\mu$  takes such a rectangular matrix to its row sums and column sums. Hence the restriction of  $\mu$  to diagonal matrices in  $\mathbb{S}_+^{d_1 d_2}$  is precisely the linear map that defines classical optimal transport in the discrete setting of [STVvR24, Section 3.1].

The quantum optimal transportation problem (QOT) is the task of minimizing a linear function  $X \mapsto \langle C, X \rangle$  over any transportation spectrahedron  $\mu^{-1}(Y, Z)$ . This is an SDP. Our main theorem in this section states that the Gibbs manifold of  $\mu$  is semi-algebraic.

**Theorem 2.5.1.** The Gibbs manifold  $\text{GM}(\mathcal{L})$  for QOT is a semi-algebraic subset of  $\mathbb{S}_+^{d_1 d_2}$ . It consists of all symmetric matrices  $Y \otimes Z$ , where  $Y \in \mathbb{S}_+^{d_1}$  and  $Z \in \mathbb{S}_+^{d_2}$ . The Gibbs variety  $\text{GV}(\mathcal{L}) \subset \mathbb{S}^{d_1 d_2}$  is linearly isomorphic to the cone over the Segre variety  $\mathbb{P}^{\binom{d_1+1}{2}-1} \times \mathbb{P}^{\binom{d_2+1}{2}-1}$ .

The image of the marginalization map  $\mu$  generalizes the polytope  $\Delta_{d_1-1} \times \Delta_{d_2-1}$ , and the fibers of  $\mu$  are quantum versions of transportation polytopes. These shapes are now nonlinear.

**Lemma 2.5.2.** The image of the map  $\mu$  is a convex cone of dimension  $\binom{d_1+1}{2} + \binom{d_2+1}{2} - 1$ :

$$\text{image}(\mu) = \{(Y, Z) \in \mathbb{S}_+^{d_1} \times \mathbb{S}_+^{d_2} : \text{Tr}(Y) = \text{Tr}(Z)\}. \quad (2.5.2)$$

For any point  $(Y, Z)$  in the relative interior of this cone, the *transportation spectrahedron*  $\mu^{-1}(Y, Z)$  is a compact convex body of dimension  $\frac{1}{2}(d_1 - 1)(d_2 - 1)(d_1 d_2 + d_1 + d_2 + 2)$ .

*Proof of Lemma 2.5.2.* The partial trace map  $\mu$  restricts to tensor products as follows:

$$\mu(Y \otimes Z) = (\text{Tr}(Z) \cdot Y, \text{Tr}(Y) \cdot Z). \quad (2.5.3)$$

Hence, if  $Y \in \mathbb{S}_+^{d_1}$  and  $Z \in \mathbb{S}_+^{d_2}$  satisfy  $t = \text{Tr}(Y) = \text{Tr}(Z)$  then  $\frac{1}{t}Y \otimes Z$  is a positive semidefinite matrix in the fiber  $\mu^{-1}(Y, Z)$ . This shows that the image is as claimed on the right hand side of (2.5.2). The image is a spectrahedral cone of dimension  $\binom{d_1+1}{2} + \binom{d_2+1}{2} - 1$ . Subtracting this from  $\dim \mathbb{S}_+^{d_1 d_2} = \binom{d_1 d_2 + 1}{2}$  yields the dimension of the interior fibers.  $\square$

**Example 2.5.3** ( $d_1=d_2=2$ ). The map  $\mu$  projects positive semidefinite symmetric  $4 \times 4$ -matrices

$$X = \begin{bmatrix} x_{1111} & x_{1112} & x_{1121} & x_{1122} \\ x_{1112} & x_{1212} & x_{1221} & x_{1222} \\ x_{1121} & x_{1221} & x_{2121} & x_{2122} \\ x_{1122} & x_{1222} & x_{2122} & x_{2222} \end{bmatrix}$$

onto a 5-dimensional convex cone, given by the direct product of two disks. The formula is

$$Y = \begin{bmatrix} x_{1111} + x_{1212} & x_{1121} + x_{1222} \\ x_{1121} + x_{1222} & x_{2121} + x_{2222} \end{bmatrix} \quad \text{and} \quad Z = \begin{bmatrix} x_{1111} + x_{2121} & x_{1112} + x_{2122} \\ x_{1112} + x_{2122} & x_{1212} + x_{2222} \end{bmatrix}.$$

The fibers of this map  $\mu$  are the 5-dimensional transportation spectrahedra  $\mu^{-1}(Y, Z)$ .

To illustrate the QOT problem, we fix the margins and the cost matrix as follows:

$$Y = \begin{bmatrix} 5 & 1 \\ 1 & 6 \end{bmatrix} \quad \text{and} \quad Z = \begin{bmatrix} 7 & 2 \\ 2 & 4 \end{bmatrix} \quad \text{and} \quad C = \begin{bmatrix} 2 & 3 & 5 & 7 \\ 3 & 11 & 13 & 17 \\ 5 & 13 & 23 & 29 \\ 7 & 17 & 29 & 31 \end{bmatrix}. \quad (2.5.4)$$

We wish to minimize  $\langle C, X \rangle$  subject to  $\mu(X) = (Y, Z)$ . The optimal solution  $X^*$  is equal to

$$\begin{bmatrix} 3.579128995196972555885181314 & 2.148103387337332721011731020 & 2.671254991031789281229265149 & -2.07566204542024789990696017 \\ 2.148103387337332721011731020 & 1.420871004803027444114818686 & 1.16978382139276763200240537 & -1.671254991031789281229265149 \\ 2.671254991031789281229265149 & 1.16978382139276763200240537 & 3.420871004803027444114818686 & -0.14810338733733272101173102 \\ -2.07566204542024789990696017 & -1.671254991031789281229265149 & -0.14810338733733272101173102 & 2.579128995196972555885181314 \end{bmatrix}.$$

This matrix has rank 2. The optimal value equals  $v = 156.964485798827271035367539305\dots$ . This is an algebraic number of degree 12. Its exact representation is the minimal polynomial

$$\begin{aligned} &125v^{12} - 465480v^{11} + 770321646v^{10} - 744236670798v^9 + 463560077206539v^8 - 193865445786866004v^7 \\ &+ 54901023652716544539v^6 - 10330064181552258647604v^5 + 1219620644420527588643307v^4 \\ &- 77994100149206862070472310v^3 + 1395374211380010273312826701v^2 \\ &+ 83502957914204004050312708316v - 2047417613706778627978564647804 = 0. \end{aligned}$$

This was derived from the KKT equations in [NRS10, Theorem 3]. We conclude that the algebraic degree of QOT for  $d_1 = d_2 = 2$  is equal to 12. This is smaller than the algebraic degree of semidefinite programming, which is 42. That is the entry for  $m=5$  and  $n=4$  in [NRS10, Table 2].

This drop arises because QOT is a very special SDP. The LSSM for our QOT problem is

$$\mathcal{L} = \left\{ \begin{bmatrix} y_1 + y_3 & y_5 & y_4 & 0 \\ y_5 & y_1 & 0 & y_4 \\ y_4 & 0 & y_2 + y_3 & y_5 \\ 0 & y_4 & y_5 & y_2 \end{bmatrix} : y_1, y_2, y_3, y_4, y_5 \in \mathbb{R} \right\}. \quad (2.5.5)$$

This defines our 5-dimensional Gibbs manifold  $\text{GM}(\mathcal{L})$  in the 10-dimensional cone  $\mathbb{S}_+^4$ . Theorem 2.5.1 states that it equals the positive part of the Gibbs variety, i.e.  $\text{GM}(\mathcal{L}) = \text{GV}(\mathcal{L}) \cap \mathbb{S}_+^4$ .

We compute the entropy maximizer inside the 5-dimensional transportation spectrahedron  $\mu^{-1}(Y, Z)$  for the marginal matrices  $Y$  and  $Z$  in (2.5.4). Notably, its entries are rational:

$$\mu^{-1}(Y, Z) \cap \text{GV}(\mathcal{L}) = \mu^{-1}(Y, Z) \cap \text{GM}(\mathcal{L}) = \left\{ \frac{1}{11} \begin{bmatrix} 35 & 10 & 7 & 2 \\ 10 & 20 & 2 & 4 \\ 7 & 2 & 42 & 12 \\ 2 & 4 & 12 & 24 \end{bmatrix} \right\}. \quad \blacklozenge$$

*Proof of Theorem 2.5.1.* By linear extension, the equation (2.5.3) serves as a definition of the marginalization map  $\mu$  on  $\mathbb{S}^{d_1 d_2}$ . We observe the following for the inner product on  $\mathbb{S}^{d_1 d_2}$ :

$$\begin{aligned} \text{Tr}((A \otimes \text{id}_{d_2})(Y \otimes Z)) &= \text{Tr}(Z) \cdot \text{Tr}(AY) \quad \text{for all } A \in \mathbb{S}^{d_1} \\ \text{and } \text{Tr}((\text{id}_{d_1} \otimes B)(Y \otimes Z)) &= \text{Tr}(Y) \cdot \text{Tr}(BZ) \quad \text{for all } B \in \mathbb{S}^{d_2}. \end{aligned}$$

Therefore, the  $(i, j)$  entry of  $\text{Tr}(Z) \cdot Y$  is obtained as  $\frac{1}{2} \langle (E_{ij} + E_{ji}) \otimes \text{id}_{d_2}, Y \otimes Z \rangle$ , where  $E_{ij}$  is the  $(i, j)$ -th matrix unit. A similar observation holds for the entries of  $\text{Tr}(Y) \cdot Z$ . This means that  $\mu(X)$  is computed by evaluating  $\text{Tr}((A \otimes \text{id}_{d_2})X)$  and  $\text{Tr}((\text{id}_{d_1} \otimes B)X)$ , where  $A$  ranges over a basis of  $\mathbb{S}^{d_1}$  and  $B$  ranges over a basis of  $\mathbb{S}^{d_2}$ . Therefore, we have

$$\mathcal{L} = \{ A \otimes \text{id}_{d_2} + \text{id}_{d_1} \otimes B : A \in \mathbb{S}^{d_1} \text{ and } B \in \mathbb{S}^{d_2} \}. \quad (2.5.6)$$

Now, the key step in the proof consists of the following formula for the matrix logarithm

$$\log(Y \otimes Z) = \log(Y) \otimes \text{id}_{d_2} + \text{id}_{d_1} \otimes \log(Z).$$

This holds for positive semidefinite matrices  $Y$  and  $Z$ , and it is verified by diagonalizing these matrices. By setting  $Y = \exp(A)$  and  $Z = \exp(B)$ , we now conclude that the Gibbs manifold  $\text{GM}(\mathcal{L})$  consists of all tensor products  $Y \otimes Z$  where  $Y \in \mathbb{S}_+^{d_1}$  and  $Z \in \mathbb{S}_+^{d_2}$ .

We have shown that  $\text{GM}(\mathcal{L})$  is the intersection of a variety with  $\mathbb{S}_+^{d_1 d_2}$ . This variety must be the Gibbs variety  $\text{GV}(\mathcal{L})$ . More precisely,  $\text{GV}(\mathcal{L})$  consists of all tensor products  $Y \otimes Z$  where  $Y, Z$  are complex symmetric. This is the cone over the Segre variety, which is the projective variety in  $\mathbb{P}^{\binom{d_1 d_2 + 1}{2} - 1}$  whose points are the tensor products  $X = Y \otimes Z$ .  $\square$

We have the following immediate consequence of the proof of Theorem 2.5.1. The entropy maximizers have rational entries. This explains the matrix at the end of Example 2.5.3

**Corollary 2.5.4.** The Gibbs point for QOT is given by  $\frac{Y \otimes Z}{\text{Tr}(Y)}$ , with  $Y, Z$  the given margins.

At this point, we revisit Section 2.2 and study its thread for the LSSM in (2.5.6).

**Example 2.5.5.** We apply Algorithm 1 to the LSSM  $\mathcal{L}$  in (2.5.5). The eigenvalues of  $\mathcal{L}$  are distinct, and the ideal  $\langle E'_1 \rangle$  in step 2 is the intersection of six prime ideals. One of them is

$$\begin{aligned} &\langle \lambda_1 + \lambda_2 - y_1 - y_2 - y_3, \lambda_3 + \lambda_4 - y_1 - y_2 - y_3, \\ &2\lambda_2\lambda_4 - \lambda_2y_1 - \lambda_4y_1 - \lambda_2y_2 - \lambda_4y_2 + 2y_1y_2 - \lambda_2y_3 - \lambda_4y_3 + y_1y_3 + y_2y_3 + y_3^2 - 2y_4^2 + 2y_5^2, \\ &\lambda_2^2 + \lambda_4^2 - \lambda_2y_1 - \lambda_4y_1 - \lambda_2y_2 - \lambda_4y_2 + 2y_1y_2 - \lambda_2y_3 - \lambda_4y_3 + y_1y_3 + y_2y_3 - 2y_4^2 - 2y_5^2 \rangle. \end{aligned}$$

The other five associated primes are found by permuting indices of  $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ . Hence, the Galois group  $G_{\mathcal{L}}$  is the Klein four-group  $S_2 \times S_2$  in  $S_4$ , and we infer the linear relation  $\lambda_1 + \lambda_2 - \lambda_3 - \lambda_4$ . The set  $E_3$  in step 7 is the singleton  $\{z_1 z_2 - z_3 z_4\}$ . The elimination in step 12 reveals the prime ideal in  $\mathbb{R}[\mathbf{x}]$  that is shown for arbitrary  $d_1, d_2$  in Corollary 2.5.6.  $\blacklozenge$

Our final result is derived from Theorem 2.5.1 using toric algebra [MS21a, Chapter 8].

**Corollary 2.5.6.** The Gibbs variety for QOT is parametrized by monomials  $x_{ijkl} = y_{ik} z_{jl}$  that are not all distinct. Its prime ideal in  $\mathbb{R}[X]$  is minimally generated by the  $2 \times 2$  minors of a matrix of format  $\binom{d_1+1}{2} \times \binom{d_2+1}{2}$ , together with  $\binom{d_1}{2} \binom{d_2}{2}$  linear forms in the entries of  $X$ .

In Chapter 3, we extend QOT to quantum graphical models [WG23]. In statistics, every undirected graph  $G$  on  $s$  vertices defines such a model [Sul18, Section 13.2]. The graphical model lives in the probability simplex  $\Delta_{d_1 d_2 \dots d_s - 1}$ . Its points are nonnegative tensors of format  $d_1 \times d_2 \times \dots \times d_s$  whose entries sum to 1. The quantum graphical model lives in the high-dimensional PSD cone  $\mathbb{S}_+^{d_1 d_2 \dots d_s}$ , where the marginalization records the partial trace for every clique in  $G$ . We will study the Gibbs manifold and the Gibbs varieties for these models. We will see that even for decomposable graphs  $G$  the dimension of the Gibbs variety is higher than that of the Gibbs manifold. Theorem 2.5.1 therefore shows that QOT is a very special example of a quantum graphical model. The graph  $G$  here has two nodes and no edges.

## 2.6. Logarithmic sparsity

We conclude this chapter with studying the phenomenon of logarithmic sparsity. Logarithmically sparse symmetric matrices are positive definite matrices for which the matrix logarithm is sparse. Such matrices arise in high-dimensional statistics [Bat17], where structural assumptions about covariance matrices are necessary for giving consistent estimators, and sparsity assumptions are natural to make. We will elaborate more on the statistical relevance in the end of this section. Once the sparsity pattern is fixed, the corresponding set of logarithmically sparse matrices forms a Gibbs manifold. In this section, we study these special Gibbs manifolds. We show that the dimension of the corresponding Gibbs variety is generally close to the upper bound given by Theorem 2.1.6 and can be computed by using simple linear algebra. We also present one more implicitization algorithm for Gibbs varieties defined by logarithmic sparsity.

Every simple undirected graph  $G$  on  $n$  nodes with edge set  $E(G) \subseteq \{(i, j) : 1 \leq i < j \leq n\}$  defines a sparsity pattern on  $n \times n$  symmetric matrices in the following way.

**Definition 2.6.1** (Sparsity from graphs). We say that  $A = (a_{ij}) \in \mathbb{S}^n$  satisfies the sparsity condition given by  $G$  if  $a_{ij} = 0$  whenever  $i \neq j$  and  $(i, j) \notin E(G)$ . Note that the diagonal entries of a matrix are never constrained to be zero. The set of all symmetric matrices satisfying the sparsity condition given by  $G$  forms an LSSM, which we will denote by  $\mathcal{L}_G$ .

Note that if  $G$  has  $n$  nodes and  $e$  edges, then  $\dim \mathcal{L}_G = n + e$ .

**Example 2.6.2.** Let  $n = 4$  and  $E(G) = \{(1, 2), (1, 3), (2, 4)\}$ . The corresponding LSSM

$$\mathcal{L}_G = \begin{bmatrix} y_{11} & 0 & 0 & y_{14} \\ 0 & y_{22} & y_{23} & 0 \\ 0 & y_{23} & y_{33} & y_{34} \\ y_{14} & 0 & y_{34} & y_{44} \end{bmatrix}$$

is cut out by the equations  $y_{12} = y_{13} = y_{24} = 0$ . ◆

**Definition 2.6.3** (Logarithmic sparsity). We say that  $A \in \text{int}(\mathbb{S}_+^n)$  satisfies the logarithmic sparsity condition given by  $G$  if  $\log A \in \mathcal{L}_G$ .

We are interested in an algebraic description of the set of matrices that satisfy a logarithmic sparsity pattern given by  $G$ . This set of matrices is precisely the Gibbs manifold of  $\mathcal{L}_G$ . Since disconnected graphs correspond to LSSMs with block-diagonal structure and block-diagonal matrices are exponentiated block-wise, we will only consider the case of connected  $G$ .

We note that an automorphism  $\sigma$  of a graph  $G$  does not change the associated linear space but induces a permutation of variables  $\mathbf{x}$  (namely,  $x_{ij}$  is sent to  $x_{\sigma(i)\sigma(j)}$ ). Therefore, by Proposition 2.1.4, the ideal of the Gibbs variety of  $\mathcal{L}_G$  in the polynomial ring  $\mathbb{C}[\mathbf{x}]$  is invariant under permutations of variables induced by automorphisms of  $G$ .

LSSMs given by graphs are nice in the sense that finding the dimension of their Gibbs varieties can be reduced to a simple linear algebra procedure of computing matrix centralizers. This is justified by the following result, which implies that for an LSSM  $\mathcal{L}_G$  arising from a graph on  $n$  nodes one always has  $m = n$  in Theorem 2.1.9.

**Proposition 2.6.4.** Let  $\mathcal{L}_G$  be an LSSM given by a simple connected graph  $G$  on  $n$  nodes. Then its eigenvalues are  $\mathbb{Q}$ -linearly independent.

*Proof.* By setting the variables  $y_{ij}$  to zero for  $i \neq j$  and the variables  $y_{ii}$  to  $n$   $\mathbb{Q}$ -linearly independent algebraic numbers, we obtain a diagonal element of  $\mathcal{L}_G$  whose eigenvalues are linearly independent over  $\mathbb{Q}$ . This immediately implies  $\mathbb{Q}$ -linear independence of the eigenvalues of  $\mathcal{L}_G$ . □

We now address the question of computing the  $\mathcal{L}_G$ -centralizer of a generic element  $A \in \mathcal{L}_G$ . One way to do this is by straightforwardly solving the system of  $\binom{n}{2}$  equations  $XA - AX = 0$  in the variables  $x_{ij}$  over the field  $\mathbb{Q}(a_{ij})$  (the minimal field over which the coefficients of the system are defined), where  $x_{ij}$  are the entries of  $X \in \mathcal{L}_G$  and  $a_{ij}$  are the entries of  $A$ . However, there is a way to give a more explicit description of the  $\mathcal{L}_G$ -centralizer.

Note that by Proposition 2.6.4 the eigenvalues of  $\mathcal{L}_G$  are  $\mathbb{Q}$ -linearly independent. In particular, this implies that the eigenvalues of  $A \in \mathcal{L}_G$  are generically distinct and that  $A$  is generically non-derogatory [HJ85, Definition 1.4.4]. Therefore, by [HJ94, Theorem 4.4.17, Corollary 4.4.18], we have  $C(A) = \text{span}_{\mathbb{R}}(\text{id}_n, A, \dots, A^{n-1})$ , where  $\text{id}_n$  is the  $n \times n$  identity matrix. Hence, finding  $C_{\mathcal{L}_G}(A)$  reduces to intersecting  $\text{span}_{\mathbb{R}}(\text{id}_n, A, \dots, A^{n-1})$  with  $\mathcal{L}_G$ . Such an intersection can be found by solving a system of linear equations  $p_0 \text{id}_n + p_1 A + \dots + p_{n-1} A^{n-1} = \sum_{(i,j) \notin S_G} c_{ij} E_{ij}$  in the variables  $p_0, \dots, p_{n-1}, c_{ij}$ . Since  $\text{id}_n$  and  $A$  are both in  $\mathcal{L}_G$ , the intersection is at least two-dimensional and we arrive at the following proposition.

**Proposition 2.6.5.** Let  $G$  be a simple connected graph on  $n$  nodes with  $e$  edges. Then  $\dim \text{GV}(\mathcal{L}_G) \leq 2n + e - 2$ .

Note that by the same argument, the upper bound  $n + d - 1$  from Theorem 2.1.6 for the dimension of the Gibbs variety of an arbitrary LSSM can be improved to  $n + d - 2$  for any LSSM containing the identity matrix.

We conjecture that  $\dim \text{GV}(\mathcal{L}_G) = \min\left(2n + e - 2, \binom{n+1}{2}\right)$ . When  $2n + e - 2 \leq \binom{n+1}{2}$ , the conjecture is equivalent to the statement that  $\{A^2, \dots, A^{n-1}\} \cup \{E_{ij} | (i, j) \in E(G)\} \cup \{E_{ii} | i = 1, \dots, n\}$  is a linearly independent set. Here  $E(G)$  denotes the set of edges of  $G$ . This conjecture is true when  $G$  is a tree (that is, a graph with no loops), as we will see below.

We continue by characterizing Gibbs varieties for LSSMs of simple connected graphs on  $n \leq 4$  vertices. Direct computation shows that for  $n \leq 3$  we always have  $\dim \text{GV}(\mathcal{L}_G) = \binom{n+1}{2}$  and therefore  $\text{GV}(\mathcal{L}_G)$  is the entire ambient space  $\mathbb{C}^{\binom{n+1}{2}}$ . For  $n = 4$  there are 6 non-isomorphic simple connected graphs, 2 of which are trees. If  $G$  is not a tree, we once again have  $\dim \text{GV}(\mathcal{L}_G) = 10 = \binom{n+1}{2}$  and  $\text{GV}(\mathcal{L}_G) = \mathbb{C}^{\binom{n+1}{2}}$ . If  $G$  is a tree, then  $\text{GV}(\mathcal{L}_G)$  is a hypersurface. We discuss the defining equations of these hypersurfaces in Example 2.6.7.

We now concentrate on the case when  $G$  is a tree. Trees are an important class of graphs that give rise to LSSMs with the smallest possible dimension for a given number of nodes. It is remarkable that for such LSSMs the dimension of the Gibbs variety only depends on the number of nodes in the graph (or, equivalently, the size of the matrices), and the dependence is linear. In what follows, we write  $\mathbb{Q}(A)$  for the field of rational functions in the entries  $a_{ij}$  of the matrix  $A$  over  $\mathbb{Q}$ .

**Theorem 2.6.6.** Let  $\mathcal{L}_G$  be an LSSM given by a tree  $G$  on  $n$  nodes. Then  $\dim \text{GV}(\mathcal{L}_G) = 3n - 3$ .

*Proof.* By Proposition 2.6.4 the dimension of the  $\mathbb{Q}$ -linear space spanned by the eigenvalues of  $\mathcal{L}_G$  is equal to  $n$ . The dimension of  $\mathcal{L}_G$  is equal to  $2n - 1$ , since  $G$  is a tree and therefore has  $n - 1$  edges. It remains to compute the dimension of the  $\mathcal{L}_G$ -centralizer of a generic element in  $\mathcal{L}_G$ . Suppose  $A \in \mathcal{L}_G$ . We are looking for solutions of the equation  $AY - YA = 0$ ,  $Y \in \mathcal{L}_G$ . This is a system of homogeneous linear equations in the unknowns  $y_{ij}$ . We have  $(AY - YA)_{ik} = \sum a_{ij} y_{jk} - \sum y_{ij} a_{jk}$ . Note that since  $Y \in \mathcal{L}_G$ ,  $y_{ij}$  is generically non-zero if and only if  $(i, j)$  is an edge of  $G$  or  $i = j$ . The same is generically true for  $a_{ij}$ . Thus,  $(AY - YA)_{ik}$  is not identically zero if and only if there exists  $j$  such that  $(i, j)$  and  $(j, k)$  are edges of  $G$  or if  $(i, k)$  is itself an edge of  $G$ . In terms of the graph  $G$ , this means that  $(AY - YA)_{ik}$  is not identically zero if and only if there is a path of edge length at most 2 from  $i$  to  $k$ . Since  $G$  is a tree, there is at most one such path. Therefore, if  $i$  and  $k$  are connected

by a path of edge length 2 via the node  $j$ , the corresponding entry of  $AY - YA$  is equal to  $a_{ij}y_{jk} - a_{jk}y_{ij}$ . It is equal to zero if  $y_{jk}$  is proportional to  $y_{ij}$  with the coefficient  $a_{ij}/a_{jk}$  (note that  $a_{jk}$  is generically non-zero). Since  $G$  is connected, we conclude that all the  $y_{ij}$  with  $i \neq j$  are proportional to each other with coefficients prescribed by  $A$ . If  $i$  and  $k$  are connected by an edge, the corresponding entry of  $AY - YA$  is equal to  $y_{ii}a_{ik} - y_{kk}a_{ik} - (a_{ii} - a_{kk})y_{ik}$ . If it is equal to zero, then  $y_{kk} = y_{ii} - \frac{a_{ii} - a_{kk}}{a_{ik}}y_{ik}$ . We conclude that, since  $G$  is connected and all the  $y_{ik}$  are proportional to each other over  $\mathbb{Q}(A)$ , all the  $y_{ii}$  can be expressed as  $\mathbb{Q}(A)$ -linear combinations of  $y_{11}$  and just one  $y_{jk}$  with  $j \neq k$ . Therefore, the centralizer, which is the solution space of the considered linear system, is at most 2-dimensional. Since it contains  $\text{id}_n$  and  $A$ , it is exactly two-dimensional. The statement of the theorem now follows from Theorem 2.1.9 for  $m = n$ ,  $d = 2n - 1$  and  $k = 2$ .  $\square$

**Example 2.6.7.** For  $n = 4$  there are exactly two non-isomorphic trees, shown below. By Theorem 2.6.6, the dimension of their Gibbs varieties is equal to 9. Therefore, these Gibbs varieties are hypersurfaces in  $\mathbb{C}^{\binom{n+1}{2}} = \mathbb{C}^{10}$ .



The corresponding LSSMs are

$$\begin{bmatrix} y_{11} & y_{12} & 0 & 0 \\ y_{12} & y_{22} & y_{23} & 0 \\ 0 & y_{23} & y_{33} & y_{34} \\ 0 & 0 & y_{34} & y_{44} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} y_{11} & y_{12} & y_{13} & y_{14} \\ y_{12} & y_{22} & 0 & 0 \\ y_{13} & 0 & y_{33} & 0 \\ y_{14} & 0 & 0 & y_{44} \end{bmatrix},$$

respectively.

For the 4-chain, the graph on the left, the Gibbs variety is defined by a single homogeneous equation of degree 6 that has 96 terms. For the graph on the right the defining equation is also homogeneous of degree 6. It has 60 terms. These equations were found with Algorithm 2.  $\blacklozenge$

We now add a bit more structure to our LSSMs by considering graphs that come with a coloring. Sparse LSSMs defined by colored graphs appear in the study of colored Gaussian graphical models in algebraic statistics [HL08], [SU10]. In this section, we study the properties of Gibbs varieties of such LSSMs.

Consider the graph  $G$  and suppose its vertices are labeled by  $p$  colors and the edges are labeled by  $q$  colors. The corresponding LSSM  $\mathcal{L}$  is defined by the following sets of equations:

- (i)  $x_{ij} = 0$  if  $(i, j)$  is not an edge of  $G$
- (ii)  $x_{ii} = x_{jj}$  if the vertices  $i$  and  $j$  have the same color.
- (iii)  $x_{ij} = x_{kl}$  if  $(i, j)$  and  $(k, l)$  are edges of  $G$  that have the same color.

We have  $\dim \mathcal{L} = p + q$ .

We will denote colored graphs by  $\mathcal{G}$  and the corresponding LSSMs by  $\mathcal{L}_{\mathcal{G}}$ . The corresponding uncolored graph will be denoted by  $G$ , as usual. Note that since  $\mathcal{L}_{\mathcal{G}} \subseteq \mathcal{L}_G$ , the inclusion of the Gibbs varieties also holds:  $\text{GV}(\mathcal{L}_{\mathcal{G}}) \subseteq \text{GV}(\mathcal{L}_G)$ . Since the identity matrix is in  $\mathcal{L}_{\mathcal{G}}$  for any  $\mathcal{G}$ , the dimension bound from Proposition 2.6.5 holds for colored graphs as well. This is reformulated in terms of numbers of colors in the following result.

**Proposition 2.6.8.** Let  $\mathcal{G}$  be a colored graph on  $n$  nodes in which vertices are labeled by  $p$  colors and edges are labeled by  $q$  colors. Then  $\dim \text{GV}(\mathcal{L}_{\mathcal{G}}) \leq n + p + q - 2$ .

**Definition 2.6.9** (Colored sparsity). We say that  $X \in \mathbb{S}_+^n$  satisfies the *colored sparsity pattern* given by  $\mathcal{G}$  if  $X \in \mathcal{L}_{\mathcal{G}}$ .

Note that if  $\mathcal{G}$  is a colored graph, the eigenvalues of  $\mathcal{L}_{\mathcal{G}}$  are not necessarily  $\mathbb{Q}$ -linearly independent. Therefore, the upper bound from Proposition 2.6.8 is not always attained.

**Example 2.6.10.** Consider the graph . The corresponding LSSM is

$$\begin{bmatrix} y_1 & y_2 & 0 \\ y_2 & y_1 & y_3 \\ 0 & y_3 & y_1 \end{bmatrix}.$$

The eigenvalues of this LSSM are  $\mathbb{Q}$ -linearly dependent: they satisfy the equation  $2\lambda_1 = \lambda_2 + \lambda_3$ . We have  $\dim \text{GV}(\mathcal{L}) = 3 < n + p + q - 2 = 3 + 1 + 2 - 2 = 4$ , which can be verified using Algorithm 1. Note that in this case  $\dim \text{GV}(\mathcal{L}_{\mathcal{G}}) = \dim \text{GM}(\mathcal{L}_{\mathcal{G}})$ .  $\blacklozenge$

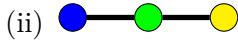
In order to illustrate how different colorings of the same graph affect the Gibbs variety, we conclude our discussion of colored graphs with analyzing those of them for which the underlying graph is the 3-chain. This is done using Algorithm 1.



The corresponding LSSM is

$$\mathcal{L}_{\mathcal{G}} = \begin{bmatrix} y_1 & y_4 & 0 \\ y_4 & y_2 & y_5 \\ 0 & y_5 & y_3 \end{bmatrix}.$$

$\dim \text{GV}(\mathcal{L}_{\mathcal{G}}) = 6$  and there are no polynomial equations that hold on the Gibbs variety.



The corresponding LSSM is

$$\mathcal{L}_{\mathcal{G}} = \begin{bmatrix} y_1 & y_4 & 0 \\ y_4 & y_2 & y_4 \\ 0 & y_4 & y_3 \end{bmatrix}.$$

$\dim \text{GV}(\mathcal{L}_{\mathcal{G}}) = 5$  and the Gibbs variety is a cubic hypersurface whose prime ideal is generated by the polynomial

$$\begin{aligned} x_{11}x_{13}x_{23} - x_{12}^2x_{23} + x_{12}x_{22}x_{13} - x_{12}x_{13}^2 - \\ - x_{12}x_{13}x_{33} + x_{12}x_{23}^2 - x_{22}x_{13}x_{23} + x_{13}^2x_{23}. \end{aligned}$$



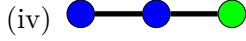
The corresponding LSSM is

$$\mathcal{L}_{\mathcal{G}} = \begin{bmatrix} y_1 & y_3 & 0 \\ y_3 & y_1 & y_4 \\ 0 & y_4 & y_2 \end{bmatrix}.$$

$\dim \text{GV}(\mathcal{L}_{\mathcal{G}}) = 5$  and the Gibbs variety is a cubic hypersurface. Its prime ideal is generated by the polynomial

$$\begin{aligned} -x_{11}x_{12}x_{23} + x_{11}x_{22}x_{13} - x_{11}x_{13}x_{33} + x_{12}x_{22}x_{23} - \\ - x_{22}^2x_{13} + x_{22}x_{13}x_{33} + x_{13}^3 - x_{13}x_{23}^2. \end{aligned}$$





The corresponding LSSM is

$$\mathcal{L}_{\mathcal{G}} = \begin{bmatrix} y_1 & y_3 & 0 \\ y_3 & y_1 & y_3 \\ 0 & y_3 & y_2 \end{bmatrix}.$$

$\dim \text{GV}(\mathcal{L}_{\mathcal{G}}) = 4$ . The Gibbs variety is a complete intersection, its prime ideal is generated by the polynomials

$$\begin{aligned} & x_{11} - x_{22} + x_{33}, \\ & -x_{12}x_{23} + x_{22}x_{13} - x_{13}^2 - x_{13}x_{33} + x_{23}^2. \end{aligned}$$



The corresponding LSSM is

$$\mathcal{L}_{\mathcal{G}} = \begin{bmatrix} y_1 & y_3 & 0 \\ y_3 & y_2 & y_4 \\ 0 & y_4 & y_1 \end{bmatrix}.$$

$\dim \text{GV}(\mathcal{L}_{\mathcal{G}}) = 5$  and the Gibbs variety is a cubic hypersurface. Its prime ideal is generated by the polynomial

$$-x_{11}x_{12}x_{23} + x_{12}^2x_{13} + x_{12}x_{23}x_{33} - x_{13}x_{23}^2.$$



The corresponding LSSM is

$$\mathcal{L}_{\mathcal{G}} = \begin{bmatrix} y_1 & y_3 & 0 \\ y_3 & y_2 & y_3 \\ 0 & y_3 & y_1 \end{bmatrix}.$$

$\dim \text{GV}(\mathcal{L}_{\mathcal{G}}) = 4$  and the Gibbs variety is an affine subspace with the prime ideal generated by  $x_{12} - x_{23}$  and  $x_{11} - x_{33}$ .




The corresponding LSSM

$$\mathcal{L}_{\mathcal{G}} = \begin{bmatrix} y_1 & y_2 & 0 \\ y_2 & y_1 & y_3 \\ 0 & y_3 & y_1 \end{bmatrix}$$

appeared in Example 2.6.10. We have  $\dim \text{GV}(\mathcal{L}_{\mathcal{G}}) = 3$ . The prime ideal of the Gibbs variety is generated by 7 polynomials:

$$\begin{aligned} & x_{12}x_{13} - x_{22}x_{23} + x_{23}x_{33}, \\ & x_{11}x_{13} - x_{12}x_{23} + x_{13}x_{33}, \\ & x_{11}x_{22} - x_{11}x_{33} - x_{22}^2 + x_{22}x_{33} + x_{13}^2, \\ & x_{12}^2 - x_{22}^2 + x_{13}^2 + x_{33}^2, \\ & x_{11}x_{12} - x_{12}x_{22} + x_{13}x_{23}, \\ & x_{11}^2 - x_{22}^2 + x_{13}^2 + x_{23}^2, \\ & -x_{12}x_{22}x_{23} + x_{12}x_{23}x_{33} + x_{22}^2x_{13} - x_{13}^3 - x_{13}x_{33}^2. \end{aligned}$$

(viii) 

The corresponding LSSM is

$$\mathcal{L}_G = \begin{bmatrix} y_1 & y_2 & 0 \\ y_2 & y_1 & y_2 \\ 0 & y_2 & y_1 \end{bmatrix}.$$

This is a commuting family and therefore, by Theorem 2.1.13,  $\dim \text{GV}(\mathcal{L}_G) = 2$  and  $\text{GM}(\mathcal{L}_G) = \text{GV}(\mathcal{L}_G) \cap \text{int}(\mathbb{S}_+^3)$ . The prime ideal of the Gibbs variety is generated by 3 linear forms and 1 quadric:  $x_{22} - x_{13} - x_{33}$ ,  $x_{12} - x_{23}$ ,  $x_{11} - x_{33}$  and  $-2x_{13}x_{33} + x_{23}^2$ .

Since the logarithm is an analytic function on  $\mathbb{R}_{>0}$ , the set of matrices satisfying the logarithmic sparsity pattern given by a graph  $G$  can be defined via formal power series equations. One way to write these equations is by using Sylvester's formula (Theorem 2.2.1).

By setting  $f$  in Sylvester's formula to be the logarithm function, we obtain a parametrization of  $\log X$  with rational functions in the entries  $x_{ij}$  of  $X$ , the eigenvalues  $\lambda_i$  of  $X$  and their logarithms  $\log \lambda_i$ . The logarithmic sparsity condition induced on  $X$  requires that some components of this parametrization are zero and therefore gives a system of polynomial equations in  $x_{ij}$ ,  $\lambda_i$  and  $\log \lambda_i$ . By eliminating the variables  $\lambda_i$  and  $\log \lambda_i$  from this system while taking into account the polynomial relations between  $\lambda_i$  and  $x_{ij}$  given by the coefficients of the characteristic polynomial, we obtain a set of defining equations of  $\text{GV}(\mathcal{L}_G)$ . This procedure is described by Algorithm 3. The notation  $I : a^\infty$  stands for saturation of an ideal  $I$  of a ring  $R$  by an element  $a \in R$ , that is,  $I : a^\infty := \{b \in R : \exists N \in \mathbb{Z}_{\geq 0} : a^N b \in I\}$ . The quantities  $x_{ij}$ ,  $\lambda_i$  and  $\log \lambda_i$  are treated by the algorithm as variables in a polynomial ring, without a priori algebraic dependencies between them. We note that the strategy of Algorithm 3 is quite similar to that Algorithm 1. The difference is that now we set  $f$  in Sylvester's formula to be the logarithm rather than the exponential function. In addition, due to the additional structure of LSSMs defined by graphs, we avoid introducing the variables  $z$ , so our computations are performed in a polynomial ring with fewer variables, and fewer variables are eliminated.

---

**Algorithm 3** Implicitization of the Gibbs variety of  $\mathcal{L}_G$  given by a graph  $G$

---

**Input:** A simple undirected connected graph  $G$ ;

**Output:** A set of defining equations of  $\text{GV}(\mathcal{L}_G)$ .

- 1:  $S \leftarrow \{(i, j) : 1 \leq i \leq j \leq n \text{ and } (i, j) \notin E(G)\}$ .
  - 2:  $\{a_{ij}\} = A \leftarrow \sum_{i=1}^n \log(\lambda_i) X_i$ , with  $X_i = \prod_{j \neq i} \frac{1}{\lambda_i - \lambda_j} (X - \lambda_j \cdot \text{id}_n)$ , where  $X = (x_{ij})$  is a symmetric matrix of variables.
  - 3:  $E_1 \leftarrow \{a_{ij} : (i, j) \in S\}$ .
  - 4: Clear the denominators in  $E_1$  and record the least common denominator  $D$ .
  - 5: Compute the characteristic polynomial  $P_X(x_{ij}; \lambda) = \det(X - \lambda \text{id}_n) = c_0(x_{ij}) + c_1(x_{ij})\lambda + \dots + c_n(x_{ij})\lambda^n$ .
  - 6:  $E_2 \leftarrow \{\text{the } n \text{ polynomials } (-1)^i \sigma_{n-i}(\lambda) - c_i(x)\}$ , where  $\sigma_{n-i}(\lambda)$  is the  $(n-i)$ -th elementary symmetric polynomial in the variables  $\lambda_1, \dots, \lambda_n$ .
  - 7: Let  $I$  be the ideal in  $\mathbb{Q}[x_{ij}, \lambda, \log \lambda]$  generated by  $E_1$  and  $E_2$ .
  - 8:  $I \leftarrow I : D^\infty$ .
  - 9:  $J \leftarrow I \cap \mathbb{Q}[x_{ij}]$ .
  - 10: **return** a set of generators of  $J$ .
- 

**Theorem 2.6.11.** Algorithm 3 is correct. The ideal  $J$  in step 9 is the prime ideal of  $\text{GV}(\mathcal{L}_G)$ .

*Proof.* Since the eigenvalues of  $\mathcal{L}_G$  are  $\mathbb{Q}$ -linearly independent, the ideal generated by  $E_2$  is prime. Moreover, there is no  $\mathbb{C}$ -algebraic relation between the eigenvalues of  $X$  and their logarithms that holds for any positive definite  $X$ . This is a consequence of Theorem 2.2.5. Any  $\mathbb{Q}$ -linear relation imposed on the logarithms of eigenvalues defines a set of positive codimension in the set of positive definite matrices. Thus, one can assume that the logarithms of eigenvalues, considered as functions in the entries of an indeterminate positive definite matrix  $X$ , are  $\mathbb{Q}$ -linear independent and apply Theorem 2.2.5. These two facts ensure that all the algebraic relations between  $X$ ,  $\lambda$  and  $\log \lambda$  are accounted for, and that the algorithm is thus correct. The ideal generated by  $E_1$  and  $E_2$  is therefore also prime, after saturation, and elimination in step 9 preserves primality.  $\square$

Note that since  $J$  is prime,  $\text{GV}(\mathcal{L}_G)$  is irreducible, as stated in Theorem 2.2.6.

We now briefly comment on the statistical relevance of logarithmic sparsity. A typical problem in high-dimensional statistics is estimating the covariance matrix of a random vector of length  $n$  from  $l \ll n$  samples. It is known that no consistent estimator can be derived in such setup without making additional assumptions on the structure of the covariance matrix. This problem can in some cases be solved by assuming that the covariance matrix has a fixed logarithmic sparsity pattern [Bat17], [Bat23]. An advantage of this assumption is that once a logarithmic sparsity pattern is induced on the covariance matrix  $C$ , it is also automatically induced on the concentration matrix  $K = C^{-1}$ , since  $(\exp L)^{-1} = \exp(-L)$ . In principle, one could relax the structural assumption of logarithmic sparsity and replace it by the assumption that the covariance matrix is an element of the Gibbs variety. The advantage of such relaxation is that checking whether a given set of polynomial equations is satisfied by the matrix is generally simpler than computing the matrix logarithm and then checking whether it satisfies the sparsity condition.

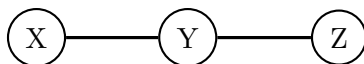
We conclude this chapter with a discussion of possible future research directions. One problem already mentioned above is to classify LSSMs by their Galois groups. A very natural question to ask is when is the Gibbs manifold semi-algebraic. It is clear that a necessary condition is  $\dim \text{GV}(\mathcal{L}) = \dim \text{GM}(\mathcal{L})$ . Is this condition also sufficient? We saw that this is the case, for instance, in Theorem 2.1.13 but we do not currently have a proof for the general case. Finally, in the context of regularization of LP and SDP, what are other regularizers giving rise to smooth manifolds via first-order optimality conditions? Is there a common framework to compute the defining equations of their Zariski closures? As we will see in Chapter 4, even for fairly well-studied and widely used functions such as the universal barrier, the analog of the Gibbs manifold is not everywhere smooth.

## Chapter 3

# Quantum Graphical Models

The goal of this chapter is to consider *quantum graphical models* [LP08] from the point of view of algebraic geometry, with the aim of offering a new perspective on open problems in quantum information theory. Roughly speaking, when passing from the classical to the quantum setting, we replace probability distributions with *density matrices*, with the classical case being recovered when these matrices are diagonal. The graph describes a physical quantum system with nodes representing subsystems. Such models have also been coined *quantum Markov networks* in the quantum information theory literature [BP12, DGMM20, PH11] and have applications to quantum many-body systems, quantum error correction and the study of entanglement. In this chapter we describe different approaches to obtain an algebraic variety associated to a quantum graphical model.

In algebraic statistics, *graphical models* [Lau96] play a prominent role, with applications to, among others, phylogenetics, causal inference and medical diagnosis [KF09, MDLW18]. Such a model arises from a graph imposing certain conditional independence statements on random variables represented by nodes in the graph. As an example [PS05, Example 1.29], consider the chain graph  $G$  on three vertices



with binary random variables  $X, Y$  and  $Z$ . This graph encodes the conditional independence statement  $X \perp\!\!\!\perp Z \mid Y$  (“ $X$  is independent of  $Z$  given  $Y$ ”). This gives rise to a statistical model described by the algebraic variety

$$\mathcal{M}_G = \mathcal{V}(p_{001}p_{100} - p_{000}p_{101}, p_{011}p_{110} - p_{010}p_{111}) \subseteq \mathbb{P}^7 = \text{Proj}(\mathbb{C}[p_{000}, \dots, p_{111}]).$$

More generally, graphical models for discrete and Gaussian random variables are algebraic varieties. This algebraic perspective advances both the theoretical foundations for these statistical models and the development of new computational methods for practical use. At the core of these advances lies the understanding of the implicit and parametric descriptions of the model and its likelihood geometry. This proved to be useful in model selection, causal discovery, and maximum likelihood estimation [Eva20, GMS06, LUSB14, URB13].

Motivated by this, we find ways to associate algebraic varieties to quantum graphical models and make progress towards understanding their implicit and parametric descriptions. This leads to a number of interesting varieties and new computational challenges. The role of the maximum likelihood estimator is taken by the *quantum information projection*. We study its geometry for *quantum exponential families* of commuting Hamiltonians, e.g. Hamiltonians arising in the context of *graph states*.

This chapter is organized as follows. In Section 3.1, we introduce the *quantum conditional mutual information (QCMI) variety* and the *Petz variety*. The former is obtained from studying the structure of quantum states satisfying *strong subadditivity* with equality [HJPW04]. The latter is related to the Petz recovery map [Pet86] and the solution of the Quantum Marginal Problem for the 3-chain [TV15]. In both instances, the graph imposes quantum conditional independence statements, in direct analogy to the classical case. In Section 3.2, we suggest a notion of a quantum graphical model as the Gibbs manifold of a certain family of Hamiltonians and we consider the corresponding Gibbs variety. Here, the graph encodes a locality structure imposed on the Hamiltonians [BP12]. This section also includes more results on Gibbs varieties. In Section 3.3 we present results on quantum exponential families coming from *stabilizer codes* [NC02, §10]. One particular example are families of Hamiltonians associated to graph states [HEB04]. We study the quantum information projection [NGKG13] and relate it to the classical theory of maximum likelihood estimation, proving a generalization of Birch's Theorem. We conclude with a short section on the stabilizer formalism, which allows us to strengthen the results of Section 3.3.

Throughout the sections we provide algorithms to compute the varieties appearing in our study and present computational examples. Those are implemented in `Julia` making use of the computer algebra package `Oscar.jl` [OSC24] and the numerical algebraic geometry tool `HomotopyContinuation.jl` [BT18]. The code is available at <https://mathrepo.mis.mpg.de/QuantumGraphicalModels>.

For basic notions of quantum information theory and the notational conventions used in this chapter we refer to Section 1.2. The ambient space of the algebraic varieties we consider in this chapter is  $\mathbb{S}^n \cong \mathbb{C}^{n(n+1)/2}$ , each point being a complex symmetric matrix. To recover a specific quantum model, we intersect the variety with the PSD cone and the hyperplane of trace one matrices. The space of real symmetric matrices will be denoted  $\mathbb{S}_{\mathbb{R}}^n \cong \mathbb{R}^{n(n+1)/2}$ . Note a slight difference in notation compared to Chapter 2.

## 3.1. Quantum graphical models on trees

It is an important and mostly open problem in quantum information theory to describe the set of compatible density matrices on subsystems of a composite system. This is known as the Quantum Marginal Problem.

**Problem 3.1.1** (Quantum Marginal Problem). Let  $S = \{1, \dots, N\}$  be a composite system on  $N$  qudits and suppose we are given density matrices  $\rho_{S_1}, \dots, \rho_{S_n}$  of  $n$  subsystems  $S_1, \dots, S_n \subseteq S$ . What conditions do  $\rho_{S_1}, \dots, \rho_{S_n}$  have to satisfy to arise from  $\rho_S$  as  $\rho_{S_i} = \text{Tr}_{S \setminus S_i} \rho_S$ ?

So far, for general graphs, this problem has only been solved in the case of disjoint subsystems  $S_i$ . See [TV15] for a survey. However, for trees it is possible to reconstruct a quantum state from its two-body marginals [DGM21, DGMM20]. This can be done using algebraic methods and motivates the algebro-geometric notions of quantum graphical models we introduce in this section. Associating algebraic varieties to quantum graphical models and studying their defining equations might provide a new way of attacking this problem.

### 3.1.1. Quantum conditional mutual information varieties

In this section we give the definition of quantum conditional mutual information (QCMI) and collect some of its properties. The vanishing of QCMI can be thought of as a quantum analog of conditional independence and gives rise to an algebraic variety called the *QCMI variety*.

The *von Neumann entropy*  $S(\rho)$  of a quantum state  $\rho$  is  $S(\rho) := -\text{Tr}(\rho \log \rho)$ . Note a slight difference with the definition in Section 2.4: we now omit the term  $\text{Tr}(\rho)$  since quantum states have unit trace. The von Neumann entropy is a straightforward generalization of the classical Shannon entropy; here, the logarithm has base two. Let  $\rho_{ABC}$  be a tripartite state. The *quantum conditional mutual information* between  $A$  and  $C$  given  $B$  is then

$$I(A : C | B) := S(AB) + S(BC) - S(ABC) - S(B),$$

where  $S(ABC) = S(\rho_{ABC})$ ,  $S(AB) = S(\text{Tr}_C \rho_{ABC})$  etc. Note that if one replaces the von Neumann entropy with the Shannon entropy in the above definition, one obtains the classical conditional mutual information  $I_{\text{cl}}(A : C | B)$  between random variables  $A$  and  $C$  given  $B$ . The identity  $I_{\text{cl}}(A : C | B) = 0$  is well-known to be equivalent to the conditional independence  $A \perp\!\!\!\perp C | B$  and leads to two possible different factorizations of the joint probability distribution  $p(a, b, c)$  [HJPW04]. The vanishing of QCMDI of a tripartite system behaves similarly, implying a more involved factorization of the density matrix of the tripartite system (see Construction 3.1.4).

The following constitutes a quantum analog of the conditional independence axioms for probability distributions, see [LP08, Theorem 4.5].

**Proposition 3.1.2.** Let  $S$  be a composite quantum system with disjoint subsystems  $A, B, C, D \subseteq S$ . Then the following implications hold:

- (i)  $I(A : C | B) = 0 \Rightarrow I(C : A | B) = 0$  (Symmetry),
- (ii)  $I(A : CD | B) = 0 \Rightarrow I(A : C | B) = 0$  (Decomposition),
- (iii)  $I(A : CD | B) = 0 \Rightarrow I(A : C | BD) = 0$  (Weak Union),
- (iv)  $I(A : B | CD) = 0$  and  $I(A : D | C) = 0 \Rightarrow I(A : BD | C) = 0$  (Contraction).

The QCMDI is closely related to the celebrated *strong subadditivity* (SSA) inequality [LR73]

$$S(ABC) + S(B) \leq S(AB) + S(BC).$$

The case of equality in SSA, i.e.  $I(A : C | B) = 0$ , has been intensively studied; the main result is the following theorem from [HJPW04].

**Theorem 3.1.3.** A quantum state  $\rho_{ABC}$  on  $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$  satisfies SSA with equality if and only if there exists a decomposition of  $\mathcal{H}_B$  as

$$\mathcal{H}_B = \bigoplus_j \mathcal{H}_{b_j^L} \otimes \mathcal{H}_{b_j^R}$$

such that  $\rho_{ABC}$  decomposes as

$$\rho_{ABC} = \bigoplus_j q_j \rho_{Ab_j^L} \otimes \rho_{b_j^R C},$$

where  $\{q_j\}_j$  is a probability distribution and  $\rho_{Ab_j^L}, \rho_{b_j^R C}$  are states on  $\mathcal{H}_A \otimes \mathcal{H}_{b_j^L}$  and  $\mathcal{H}_{b_j^R} \otimes \mathcal{H}_C$ , respectively.

**Construction 3.1.4** (QCMDI variety of the 3-chain graph). The following reformulation of Theorem 3.1.3 plays a central role in the construction of the QCMDI variety. Setting  $\Lambda_{AB} := \bigoplus_j q_j \rho_{Ab_j^L} \otimes \text{Id}_{b_j^R C}$  and  $\Lambda_{BC} := \bigoplus_j \text{Id}_{Ab_j^L} \otimes \rho_{b_j^R C}$ , we arrive at

$$I(A : C | B) = 0 \text{ if and only if } \rho_{ABC} = \Lambda_{AB} \Lambda_{BC} \text{ with } [\Lambda_{AB}, \Lambda_{BC}] = 0, \quad (3.1.1)$$

where  $\Lambda_{AB}, \Lambda_{BC}$  are symmetric matrices acting on  $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$  and  $\Lambda_{AB}$  acts as identity on  $\mathcal{H}_C$ , and, likewise,  $\Lambda_{BC}$  acts as identity on  $\mathcal{H}_A$  [BP12]. Then the right hand side of (3.1.1) gives rise to a parametrization of a variety we denote by  $\mathcal{Q}_{I(A:C|B)}$ .

**Example 3.1.5** ( $\mathcal{Q}_{I(A:C|B)}$  in the qubit case). Let  $\mathcal{H}_A \cong \mathcal{H}_B \cong \mathcal{H}_C \cong \mathbb{C}^2$  and write  $\Lambda_{AB} = M \otimes \text{Id}_2$ ,  $\Lambda_{BC} = \text{Id}_2 \otimes N$  for  $M, N \in \mathbb{S}^4$ . In this case, the parametrization of the variety  $\mathcal{Q}_{I(A:C|B)}$  induced by the right hand side of (3.1.1) sends

$$M = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \\ x_2 & x_5 & x_6 & x_7 \\ x_3 & x_6 & x_8 & x_9 \\ x_4 & x_7 & x_9 & x_{10} \end{bmatrix} \quad \text{and} \quad N = \begin{bmatrix} y_1 & y_2 & y_3 & y_4 \\ y_2 & y_5 & y_6 & y_7 \\ y_3 & y_6 & y_8 & y_9 \\ y_4 & y_7 & y_9 & y_{10} \end{bmatrix}$$

to the matrix

$$\begin{bmatrix} x_1y_1 + x_2y_4 & x_1y_2 + x_2y_5 & x_1y_4 + x_2y_6 & x_1y_7 + x_2y_9 & x_4y_1 + x_7y_4 & x_4y_2 + x_7y_5 & x_4y_4 + x_7y_6 & x_4y_7 + x_7y_9 \\ x_1y_2 + x_2y_7 & x_1y_3 + x_2y_8 & x_1y_5 + x_2y_9 & x_1y_8 + x_2y_{10} & x_4y_2 + x_7y_7 & x_4y_3 + x_7y_8 & x_4y_5 + x_7y_9 & x_4y_8 + x_7y_{10} \\ x_2y_1 + x_3y_4 & x_2y_2 + x_3y_5 & x_2y_4 + x_3y_6 & x_2y_7 + x_3y_9 & x_5y_1 + x_8y_4 & x_5y_2 + x_8y_5 & x_5y_4 + x_8y_6 & x_5y_7 + x_8y_9 \\ x_2y_2 + x_3y_7 & x_2y_3 + x_3y_8 & x_2y_5 + x_3y_9 & x_2y_8 + x_3y_{10} & x_5y_2 + x_8y_7 & x_5y_3 + x_8y_8 & x_5y_5 + x_8y_9 & x_5y_8 + x_8y_{10} \\ x_4y_1 + x_5y_4 & x_4y_2 + x_5y_5 & x_4y_4 + x_5y_6 & x_4y_7 + x_5y_9 & x_6y_1 + x_9y_4 & x_6y_2 + x_9y_5 & x_6y_4 + x_9y_6 & x_6y_7 + x_9y_9 \\ x_4y_2 + x_5y_7 & x_4y_3 + x_5y_8 & x_4y_5 + x_5y_9 & x_4y_8 + x_5y_{10} & x_6y_2 + x_9y_7 & x_6y_3 + x_9y_8 & x_6y_5 + x_9y_9 & x_6y_8 + x_9y_{10} \\ x_7y_1 + x_8y_4 & x_7y_2 + x_8y_5 & x_7y_4 + x_8y_6 & x_7y_7 + x_8y_9 & x_9y_1 + x_{10}y_4 & x_9y_2 + x_{10}y_5 & x_9y_4 + x_{10}y_6 & x_9y_7 + x_{10}y_9 \\ x_7y_2 + x_8y_7 & x_7y_3 + x_8y_8 & x_7y_5 + x_8y_9 & x_7y_8 + x_8y_{10} & x_9y_2 + x_{10}y_7 & x_9y_3 + x_{10}y_8 & x_9y_5 + x_{10}y_9 & x_9y_8 + x_{10}y_{10} \end{bmatrix}.$$

This results in a twelve-dimensional variety inside  $\mathbb{S}^8$  cut out by 735 equations in degrees one to five; the degree of  $\mathcal{Q}_{I(A:C|B)}$  is 110. As all of these equations are homogeneous,  $\mathcal{Q}_{I(A:C|B)}$  can be considered as a subvariety of  $\mathbb{P}^{35} = \text{Proj}(\mathbb{C}[z_1, \dots, z_{36}])$ . Among these equations only two are linear:

$$z_{14} - z_{18} + z_{23} - z_{29} = 0, \quad z_{12} - z_{16} - z_{25} + z_{31} = 0,$$

and just one has degree five:

$$\begin{aligned} & -z_{13}z_{22}z_{29}z_{31}z_{33} - z_{13}z_{22}z_{31}^2z_{35} + z_{13}z_{24}z_{29}^2z_{33} + z_{13}z_{24}z_{29}z_{31}z_{35} + z_{22}^2z_{29}z_{31}z_{33} \\ & + z_{22}^2z_{31}^2z_{35} + z_{22}z_{24}z_{25}z_{29}z_{35} - z_{22}z_{24}z_{25}z_{31}z_{33} - z_{22}z_{24}z_{29}^2z_{33} - 2z_{22}z_{24}z_{29}z_{31}z_{35} \\ & + z_{22}z_{24}z_{31}^2z_{33} - z_{23}z_{24}^2z_{29}z_{35} + z_{23}z_{24}^2z_{31}z_{33} + z_{24}^2z_{29}^2z_{35} - z_{24}^2z_{29}z_{31}z_{33} = 0. \end{aligned}$$

Here the variables  $z_1, \dots, z_{36}$  denote the entries of a symmetric  $8 \times 8$ -matrix written in order starting from left to right and continuing from top to bottom. Note that if you set all non-diagonal entries in  $M$  and  $N$  to zero, this results in a monomial parametrization of the classical graphical model of the 3-chain as presented in the beginning of this chapter.  $\blacklozenge$

**Proposition 3.1.6.** The variety  $\mathcal{Q}_{I(A:C|B)}$  is irreducible.

*Proof.* Under the composition of morphisms

$$\text{U}(8) \times \mathbb{R}^4 \times \mathbb{R}^4 \rightarrow \mathbb{S}^8 \times \mathbb{S}^8 \xrightarrow{\text{mult.}} \mathbb{S}^8$$

given by

$$\begin{aligned} (U, \lambda, \mu) & \mapsto (U \text{diag}(\lambda_1, \lambda_1, \dots, \lambda_4, \lambda_4)U^{-1}, U \text{diag}(\lambda_1, \dots, \lambda_4, \lambda_1, \dots, \lambda_4)U^{-1}) \\ (M, N) & \mapsto M \cdot N, \end{aligned}$$

where  $\lambda = (\lambda_1, \dots, \lambda_4)$  and  $\mu = (\mu_1, \dots, \mu_4)$ ,  $\mathcal{Q}_{I(A:C|B)}$  is the image of an irreducible variety.  $\square$

In analogy to the classical theory of graphical models, we associate QCMDI statements to separations in an undirected graph. A *separator* between two sets of nodes  $A$  and  $C$  in a graph  $G$  is a set  $B$  of nodes such that every path from a node in  $A$  to a node in  $C$  contains a node in  $B$ . For classical graphical models on undirected graphs, the Hammersley–Clifford Theorem [Lau96, Theorem 3.9] (see also [CH71]) states that a positive probability distribution satisfies the conditional independence statements associated to separations in a graph if and only if it factorizes according to the graph. One might attempt to achieve a similar factorization theorem for quantum graphical models. However, such a description is not available for arbitrary graphs. Still, there is a “quantum Hammersley–Clifford Theorem” for trees.

**Theorem 3.1.7** ([PH11, Theorem 1]). Let  $G = (V, E)$  with  $V = \{v_1, \dots, v_N\}$  be a tree and let  $\rho$  be a positive definite quantum state on a Hilbert space  $\mathcal{H} = \mathcal{H}_1 \otimes \dots \otimes \mathcal{H}_N$  satisfying all QCMDI statements imposed by  $G$ . Then  $\rho$  can be written as the exponential of a sum of local commuting Hamiltonians, i.e.  $\rho = \exp(H)$  with

$$H = \sum_{C \in \mathcal{C}(G)} h_C, \quad [h_C, h_{C'}] = 0 \text{ for all } C, C' \in \mathcal{C}(G),$$

where  $\mathcal{C}(G)$  is the set of cliques of  $G$  and  $h_C$  is only nontrivial on the clique  $C$ , i.e.  $h_C$  is an endomorphism on  $\mathcal{H}$  acting as identity on each  $\mathcal{H}_i$  where  $v_i \notin C$ .

This quantum Hammersley–Clifford Theorem is a generalization of Equation (3.1.1) to trees. Along with Example 3.1.5, this suggests the following construction of the QCMDI variety of a tree and an associated quantum graphical model.

**Construction 3.1.8** (QCMDI variety of a tree). Let  $G = (V, E)$  be an undirected tree with vertices labelled  $S_1, \dots, S_N$ . Let  $\rho_V = \rho_{S_1 \dots S_N}$  be a quantum state on  $\mathcal{H}_1 \otimes \dots \otimes \mathcal{H}_N$ . For each triple of vertices  $S_i, S_j, S_k$  such that  $S_j$  separates  $S_i$  from  $S_k$  in  $G$ , we impose the QCMDI statement  $I(S_i : S_k | S_j) = 0$ , i.e. we require

$$\text{Tr}_{V \setminus \{S_i, S_j, S_k\}} \rho_V = \Lambda_{S_i S_j} \Lambda_{S_j S_k} \text{ with } [\Lambda_{S_i S_j}, \Lambda_{S_j S_k}] = 0 \quad (3.1.2)$$

as in (3.1.1). Moreover, for any two QCMDI statements  $I(S_i : S_k | S_j) = 0 = I(S_{i'} : S_{k'} | S_{j'})$  we impose the *compatibility constraints*

$$\text{Tr}_{\mathcal{T} \setminus (\mathcal{T} \cap \mathcal{T}')} \rho_{\mathcal{T}} = \text{Tr}_{\mathcal{T}' \setminus (\mathcal{T} \cap \mathcal{T}')} \rho_{\mathcal{T}'}, \text{ where } \mathcal{T} = (S_i, S_j, S_k), \mathcal{T}' = (S_{i'}, S_{j'}, S_{k'}). \quad (3.1.3)$$

In the qubit case, this construction gives rise to Algorithm 4 whose output is a variety inside  $\mathbb{S}^{2^N}$ . We call this variety the *QCMDI variety* associated to  $G$  and denote it by  $\mathcal{Q}_G$ . Algorithm 4 constructs the QCMDI variety by considering the conditions (3.1.2), (3.1.3) as polynomial constraints in the entries of an arbitrary density matrix  $\rho$  and of the matrices  $\Lambda_{S_i S_j}, \Lambda_{S_j S_k}$ , then it eliminates the  $\Lambda$  parameters. The QCMDI variety  $\mathcal{Q}_G$  defines a *quantum graphical model*  $\mathcal{M}_G$  by restricting to PSD matrices with trace one inside  $\mathcal{Q}_G$ . Note that Algorithm 4 and the notion of the QCMDI variety generalize straightforwardly to arbitrary qudit systems.

**Remark 3.1.9.** Let  $G$  be the path graph on three vertices with ordered vertex labels  $A, B$  and  $C$  and consider the qubit case  $\mathcal{H}_A \cong \mathcal{H}_B \cong \mathcal{H}_C \cong \mathbb{C}^2$ . Then  $\mathcal{Q}_G$  is the variety  $\mathcal{Q}_{I(A:C|B)}$  from Example 3.1.5. The computations in this example were carried out using Algorithm 4.

**Remark 3.1.10.** Note that as we consider trees, it is equivalent to impose QCMDI statements on triples of vertices as in Construction 3.1.8 or to impose a *global quantum Markov property*, in the sense that one imposes the QCMDI statement  $I(A : C | B)$  for any triple of sets of vertices  $A, B, C \subseteq V$  such that  $B$  separates  $A$  from  $C$ , as one can derive the latter from the former by using the Weak Union and Contraction axioms from Proposition 3.1.2.



**Algorithm 4** Computing the QCMI variety  $\mathcal{Q}_G$ 


---

**Input:** A graph  $G = (V, E)$   
**Output:** Polynomials defining  $\mathcal{Q}_G \subseteq \mathbb{S}^{2^N}$

- 1:  $N \leftarrow \#V$
- 2:  $\rho_V \leftarrow$  symmetric  $2^N \times 2^N$ -matrix consisting of variables  $\rho_{11}, \rho_{12}, \dots, \rho_{2^N 2^N}$
- 3:  $\mathcal{E} \leftarrow \emptyset$  initialize list of equations
- 4: **for** every triple of vertices  $\mathcal{T} = (S_i, S_j, S_k)$  such that  $S_j$  separates  $S_i$  from  $S_k$  in  $G$  **do**
- 5:    $\Lambda_{S_i S_j} \leftarrow (\lambda_{lm}^{\mathcal{T}}) \otimes \text{Id}_2$  where  $(\lambda_{lm}^{\mathcal{T}})$  is a symmetric  $4 \times 4$ -matrix of variables
- 6:    $\Lambda_{S_j S_k} \leftarrow \text{Id}_2 \otimes (\mu_{lm}^{\mathcal{T}})$  where  $(\mu_{lm}^{\mathcal{T}})$  is a symmetric  $4 \times 4$ -matrix of variables
- 7:    $\mathcal{E}' \leftarrow$  entries of  $\text{Tr}_{V \setminus \mathcal{T}} \rho_V - \Lambda_{S_i S_j} \Lambda_{S_j S_k}$
- 8:    $\mathcal{E}'' \leftarrow$  entries of  $[\Lambda_{S_i S_j}, \Lambda_{S_j S_k}]$
- 9:    $\mathcal{E} \leftarrow \mathcal{E} \cup \mathcal{E}' \cup \mathcal{E}''$
- 10: **for** every pair of triples of vertices  $\mathcal{T} = (S_i, S_j, S_k)$  and  $\mathcal{T}' = (S_{i'}, S_{j'}, S_{k'})$  **do**
- 11:    $\mathcal{E}''' \leftarrow$  entries of  $\text{Tr}_{\mathcal{T} \setminus (\mathcal{T} \cap \mathcal{T}')} \rho_{\mathcal{T}} - \text{Tr}_{\mathcal{T}' \setminus (\mathcal{T} \cap \mathcal{T}')} \rho_{\mathcal{T}'}$
- 12:    $\mathcal{E} \leftarrow \mathcal{E} \cup \mathcal{E}'''$
- 13:  $I \leftarrow$  ideal generated by  $\mathcal{E}$  in  $\mathbb{C}[\rho, \lambda, \mu]$
- 14:  $J \leftarrow$  elimination ideal  $I \cap \mathbb{C}[\rho]$
- 15: **return** a set of generators of  $J$

---

**Example 3.1.11.** Consider the claw tree  $G$  on four vertices with labels  $A, B, C, D$ , the set of edges  $\{\{A, D\}, \{B, D\}, \{C, D\}\}$ , and the corresponding Hilbert spaces  $\mathcal{H}_A \cong \mathcal{H}_B \cong \mathcal{H}_C \cong \mathcal{H}_D \cong \mathbb{C}^2$ . The Hilbert space of the full system is  $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C \otimes \mathcal{H}_D \cong \mathbb{C}^{16}$ . Every path in  $G$  with three vertices imposes a QCMI statement and any such path contains the node  $D$ . The model  $\mathcal{M}_G$  consists of all density matrices  $\rho$  that satisfy the three QCMI statements

$$I(A : B | D) = I(A : C | D) = I(B : C | D) = 0.$$

These lead to factorizations of the marginal subsystems as

$$\begin{aligned} \text{Tr}_A \rho_{ABCD} &= \rho_{BCD} = \Lambda_{BD} \Lambda_{CD} \text{ with } [\Lambda_{BD}, \Lambda_{CD}] = 0, \\ \text{Tr}_B \rho_{ABCD} &= \rho_{ACD} = \Lambda_{AD} \Lambda_{CD} \text{ with } [\Lambda_{AD}, \Lambda_{CD}] = 0, \\ \text{Tr}_C \rho_{ABCD} &= \rho_{ABD} = \Lambda_{AD} \Lambda_{BD} \text{ with } [\Lambda_{AD}, \Lambda_{BD}] = 0. \end{aligned}$$

In addition, there are compatibility conditions on the marginals which lead to

$$\text{Tr}_B \rho_{ABD} = \text{Tr}_C \rho_{ACD}, \quad \text{Tr}_A \rho_{ABD} = \text{Tr}_C \rho_{BCD}, \quad \text{Tr}_A \rho_{ACD} = \text{Tr}_B \rho_{BCD}.$$

It is computationally very challenging to obtain defining equations of  $\mathcal{Q}_G$  as Algorithm 4 would involve eliminating 60 variables in a polynomial ring in 196 variables, which is infeasible with current computational resources.  $\blacklozenge$

**Question 3.1.12.** Is the variety  $\mathcal{Q}_G$  irreducible for any tree  $G$ ?

It would be desirable to find a parametrization of  $\mathcal{Q}_G$ . Quantum information theory provides a map that recovers a unique quantum state compatible with given two-body marginals on a tree, called the *Petz recovery map* [HJPW04, Pet86]. However, our algebraic version of this map does not parametrize the QCMI variety  $\mathcal{Q}_G$  due to the fact that we are working with complex symmetric matrices instead of Hermitian matrices. The Petz recovery map therefore gives rise to a different variety, which we introduce in the next subsection.

### 3.1.2. Petz varieties

The Quantum Marginal Problem asks about how to reconstruct a quantum state of a composite system from the states of its subsystems. In the case of the 3-chain graph with ordered vertices  $A, B$  and  $C$ , one can ask for a quantum state  $\rho_{ABC}$  with given two-body marginals  $\rho_{AB}$  and  $\rho_{BC}$  and satisfying the quantum Markov condition  $I(A : C | B) = 0$ . The answer to this particular problem is given by the Petz recovery map. This map is of algebraic nature and gives rise to an algebraic variety, the *Petz variety*, which we study in this subsection.

We start by introducing the Petz recovery map for the 3-chain graph with the associated Hilbert space  $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$ . Let  $\mathcal{C}$  be the set of pairs of compatible invertible density operators on  $\mathcal{H}_A \otimes \mathcal{H}_B$  and  $\mathcal{H}_B \otimes \mathcal{H}_C$ , respectively, i.e. an element in  $\mathcal{C}$  is of the form  $(\rho_{AB}, \rho_{BC})$ , where  $\rho_{AB}$  and  $\rho_{BC}$  are invertible density operators satisfying the compatibility condition  $\text{Tr}_A \rho_{AB} = \text{Tr}_C \rho_{BC}$ . The Petz recovery map  $\mathcal{R}_G$  for the 3-chain graph  $G$  is

$$\mathcal{R}_G: \mathcal{C} \rightarrow \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$$

$$\mathcal{R}_G(\rho_{AB}, \rho_{BC}) = (\rho_{AB}^{1/2} \otimes \text{Id}_C)(\text{Id}_A \otimes \rho_B^{-1/2} \otimes \text{Id}_C)(\text{Id}_A \otimes \rho_{BC})(\text{Id}_A \otimes \rho_B^{-1/2} \otimes \text{Id}_C)(\rho_{AB}^{1/2} \otimes \text{Id}_C), \quad (3.1.4)$$

where  $\text{Id}_A, \text{Id}_B$  and  $\text{Id}_C$  are the identity operators on  $\mathcal{H}_A, \mathcal{H}_B$  and  $\mathcal{H}_C$ , respectively. The recovered state is compatible with the marginals and satisfies  $I(A : C | B) = 0$ . Moreover, it is the unique maximizer of the von Neumann entropy among all states on  $\mathcal{H}$  [DGMM20, Theorem 1]. This, in particular, shows that the map  $\mathcal{R}'_G$  defined by

$$\mathcal{R}'_G(\rho_{AB}, \rho_{BC}) = (\text{Id}_A \otimes \rho_{BC}^{1/2})(\text{Id}_A \otimes \rho_B^{-1/2} \otimes \text{Id}_C)(\rho_{AB} \otimes \text{Id}_C)(\text{Id}_A \otimes \rho_B^{-1/2} \otimes \text{Id}_C)(\text{Id}_A \otimes \rho_{BC}^{1/2})$$

recovers the same state [DGMM20, Remark 2].

The Petz recovery map (3.1.4) gives rise to a rational map  $R_G$ . From this point forward we restrict to the qubit case  $\mathcal{H}_A \cong \mathcal{H}_B \cong \mathcal{H}_C \cong \mathbb{C}^2$ ; the general case is a straightforward generalization. Let  $\rho_{AB}^{1/2} = X = (x_{ij})$  be a  $4 \times 4$ -symmetric matrix of variables  $x_{11}, x_{12}, \dots, x_{44}$ . In the same way, let  $\rho_{BC}^{1/2} = Y = (y_{ij})$ . Finally, let  $\rho_B^{1/2} = Z = (z_{ij})$  be a  $2 \times 2$ -symmetric matrix of variables. To reflect the required marginal compatibilities between  $\rho_{AB}, \rho_{BC}$  and  $\rho_B$ , we impose the conditions  $\text{Tr}_C(Y^2) = \text{Tr}_A(X^2) = Z^2$ . These conditions cut out a variety  $V$  in  $\mathbb{S}_{\mathbb{R}}^4 \times \mathbb{S}_{\mathbb{R}}^4 \times \mathbb{S}_{\mathbb{R}}^2$ . Analogously to (3.1.4), the map  $R_G: V \dashrightarrow \mathbb{S}_{\mathbb{R}}^8$  sends a point  $(x, y, z) \in V$  to

$$(x \otimes \text{Id}_C)(\text{Id}_A \otimes z^{-1} \otimes \text{Id}_C)(\text{Id}_A \otimes y)(\text{Id}_A \otimes z^{-1} \otimes \text{Id}_C)(x \otimes \text{Id}_C). \quad (3.1.5)$$

We call the Zariski closure of  $R_G(V)$  inside  $\mathbb{S}^8$ , the space of *complex* symmetric  $8 \times 8$ -matrices, the *Petz variety* of  $G$  and denote it  $\mathcal{P}_G$ .

**Remark 3.1.13.** The expression (3.1.5) for  $R_G$  gives a concrete polynomial parametrization of  $\mathcal{P}_G$ . The polynomials appearing in  $R_G$  have degree five and have a minimum of 20 and maximum of 32 terms. The number of parameter variables is 23 while the variety  $V$  has dimension 17. Algorithm 5 provides a symbolic routine to compute the ideal of the Petz variety for arbitrary trees. When restricting to the case where  $x, y$  and  $z$  are diagonal, (3.1.5) gives yet another parametrization of the classical graphical model of the 3-chain from the beginning of this chapter.

**Proposition 3.1.14.** Let  $G$  be the 3-chain graph. The Petz variety  $\mathcal{P}_G$  is irreducible.

*Proof.* Consider the subset  $S \subset \mathbb{S}_{\mathbb{R}}^8 \times \mathbb{S}_{\mathbb{R}}^8$  consisting of pairs of invertible matrices whose partial traces agree. The condition that their partial traces agree defines a linear subspace of  $\mathbb{S}_{\mathbb{R}}^8 \times \mathbb{S}_{\mathbb{R}}^8$ . Linear spaces are irreducible and taking out the locus of positive codimension where the matrices become singular preserves irreducibility. Therefore,  $S$  is irreducible. Note

that  $\mathcal{R}_G$  can be considered as a map on  $S$  and the Zariski closure of its image coincides with  $\mathcal{P}_G$  since square roots and inverses of symmetric matrices are again symmetric. Therefore, since  $\mathcal{R}_G$  is continuous,  $\mathcal{P}_G$  is irreducible.  $\square$

The Petz map can be generalized to arbitrary trees by iteratively applying the procedure for 3-chains of a tree  $G$  [DGM21, DGMM20]. This is done by taking two leaves  $v_1$  and  $v_2$  of a tree  $G$ , and replacing  $G \setminus \{v_1, v_2\}$  by a single vertex representing a joint state on this subgraph. The joint state on  $G$  is then expressed in terms of states on  $G \setminus \{v_1\}$  and  $G \setminus \{v_2\}$  via (3.1.4); by applying this procedure iteratively to  $G \setminus \{v_1\}$  and  $G \setminus \{v_2\}$ , we reduce to the level of two-body marginals. This process leads to a map as in (3.1.4) involving only one- and two-body marginals; again, we denote the resulting map by  $\mathcal{R}_G$ . Note that the expression for  $\mathcal{R}_G$  depends on the choice of  $v_1$  and  $v_2$  in each iteration.

We now generalize the construction of the Petz variety to arbitrary trees.

**Construction 3.1.15** (Petz variety). Let  $G$  be a tree with  $N$  vertices and let us fix an expression for  $\mathcal{R}_G$  as obtained in the previous paragraph. Let  $\varrho_1$  and  $\varrho_2$  be the sets of one- and two-body marginals occurring in  $\mathcal{R}_G$ . Moreover, let  $V$  be the variety inside  $S = (\mathbb{S}_{\mathbb{R}}^2)^{\#\varrho_1} \times (\mathbb{S}_{\mathbb{R}}^4)^{\#\varrho_2}$  consisting of tuples of symmetric matrices satisfying compatibility constraints according to  $G$ . In analogy to (3.1.4),  $\mathcal{R}_G$  gives rise to a rational map  $R_G: V \dashrightarrow \mathbb{S}_{\mathbb{R}}^{2^N}$ . The *Petz variety*  $\mathcal{P}_G$  of  $G$  is defined as  $\overline{R_G(V)} \subseteq \mathbb{S}^{2^N}$ .

Algorithm 5 makes this construction explicit and computes the ideal of  $\mathcal{P}_G$ .

---

**Algorithm 5** Computing the Petz variety  $\mathcal{P}_G$

---

**Input:** A graph  $G$  with  $N$  vertices

**Output:** An ideal defining the Petz variety  $\mathcal{P}_G \subseteq \mathbb{S}^{2^N}$

- 1:  $R_G \leftarrow$  expression for  $\mathcal{R}_G$  in terms of one- and two-body marginals
  - 2:  $\varrho_1, \varrho_2 \leftarrow$  sets of one- and two-body marginals, respectively, involved in  $R_G$
  - 3: **for** all  $\rho_v \in \varrho_1$  **do**
  - 4:      $Z_v \leftarrow (z_{ij}^v)$  symmetric  $2 \times 2$ -matrix of variables
  - 5: **for** all  $\rho_{v_1 v_2} \in \varrho_2$  **do**
  - 6:      $X_{v_1 v_2} \leftarrow (x_{ij}^{v_1 v_2})$  symmetric  $4 \times 4$ -matrix of variables
  - 7:  $S \leftarrow (\mathbb{S}_{\mathbb{R}}^2)^{\#\varrho_1}_{\{Z_{v_i}\}} \times (\mathbb{S}_{\mathbb{R}}^4)^{\#\varrho_2}_{\{X_{v_i v_j}\}}$
  - 8:  $\mathcal{E} \leftarrow \emptyset$
  - 9: **for** all pairs  $(\rho_{v_1 v_2}, \rho_{w_1 w_2}) \in \varrho_2^2$  such that  $v_2 = w_1$  **do**
  - 10:      $\mathcal{E} \leftarrow \mathcal{E} \cup \{\text{entries of } \text{Tr}_{v_1}(X_{v_1 v_2}^2) - \text{Tr}_{w_2}(X_{w_1 w_2}^2)\} \cup \{\text{entries of } \text{Tr}_{v_1}(X_{v_1 v_2}^2) - Z_{v_2}^2\}$
  - 11:  $V \leftarrow$  variety defined by  $\mathcal{E}$  inside  $S$
  - 12:  $R_G \leftarrow R_G$  with every  $\rho_{v_1 v_2}$  replaced by  $X_{v_1 v_2}$  and every  $\rho_v$  replaced by  $Z_v$
  - 13: **return**  $\ker(R_G: \mathbb{C}[\mathbb{S}^{2^N}] \rightarrow \mathbb{C}[V])$
- 

**Proposition 3.1.16.** The Petz variety  $\mathcal{P}_G$  does not depend on the choice of expression for the recovery map  $\mathcal{R}_G$ .

*Proof.* By the same argument as for the 3-chain graph above, the map  $\mathcal{R}_G$  does not depend on the chosen expression. Let  $\mathcal{C}$  be the domain of  $\mathcal{R}_G$ ; each element of  $\mathcal{C}$  is a tuple consisting of  $\#E$ -many compatible invertible two-body marginals, where  $E$  is the set of edges of  $G$ . Consider the set  $(\mathbb{S}_{\mathbb{R}}^4)^{\#E}$  of  $\#E$ -tuples of real symmetric  $4 \times 4$ -matrices; it is Zariski dense in the set of complex symmetric matrices. We have  $\mathcal{R}_G(\mathcal{C} \cap (\mathbb{S}_{\mathbb{R}}^4)^{\#E}) = R_G(V \cap U)$  where  $U \subseteq \mathbb{C}^{2^N \times 2^N}$  is the Zariski dense open set of invertible matrices. The set  $R_G(V \cap U)$

is Zariski dense in the Petz variety  $\mathcal{P}_G$ . Let  $\mathcal{R}'_G$  be another expression for the recovery map, then  $\mathcal{R}_G(\mathcal{C} \cap (\mathbb{S}_{\mathbb{R}}^4)^{\#E}) = R'_G(V \cap U)$  and denote by  $\mathcal{P}'_G$  the variety defined by  $R'_G$ . It follows that  $R'_G(V \cap U) = R_G(V \cap U)$  so  $\mathcal{P}_G$  and  $\mathcal{P}'_G$  agree on a dense open set, thus  $\mathcal{P}_G = \overline{R_G(V)} = \overline{R'_G(V)} = \mathcal{P}'_G$ .  $\square$

**Proposition 3.1.17.** For any tree  $G$ , the Petz variety  $\mathcal{P}_G$  is irreducible.

*Proof.* The proof is analogous to that of Proposition 3.1.14:  $\mathcal{P}_G$  can be represented as the Zariski closure of the image of a linear space under a continuous map.  $\square$

**Remark 3.1.18.** Computing the defining equations of the Petz variety is very challenging. Even in the case of the 3-chain graph, Algorithm 5 does not terminate as it involves symbolic computations in a polynomial ring with 59 variables. Applying numerical implicitization techniques is also not straightforward for the same reason.

**Remark 3.1.19.** If we considered Hermitian matrices, the set of states recovered by the Petz map would coincide with the set of states satisfying SSA with equality [HJPW04]. However, since the ambient space of our varieties is that of complex symmetric matrices, the QCM variety  $\mathcal{Q}_G$  and the Petz variety  $\mathcal{P}_G$  are not the same.

## 3.2. Quantum graphical models from Gibbs manifolds

In this section, we revisit the concluding remark of Section 2.5 and consider a new class of quantum graphical models, which arise as Gibbs manifolds of families of Hamiltonians. These serve as examples of *quantum exponential families* [Zho08].

In physics, the Gibbs manifold parametrizes thermal states of a family of Hamiltonians. Those states are crucial to quantum many-body systems theory and computation [Alh23]. When the LSSM  $\mathcal{L}$  consists of diagonal matrices with rational entries, Theorem 2.1.13 ensures that the corresponding Gibbs manifold is semi-algebraic. For such  $\mathcal{L}$  the resulting Gibbs variety is toric and recovers the classical notion of exponential families [Efr22]. Moreover, the Gibbs manifold in this case is the intersection of the Gibbs variety with the PSD cone.

### 3.2.1. Gibbs varieties of linear systems of Hamiltonians

The quantum Hammersley–Clifford Theorem (Theorem 3.1.7) suggests to consider exponentials of *local* Hamiltonians, i.e. those that act non-trivially only on a small subsystem. However, we do not consider the class of local *commuting* Hamiltonians as they neither form an LSSM nor a unirational variety (see Subsection 3.2.2).

To a simple, undirected graph  $G = (V, E)$  we associate an LSSM of Hamiltonians as follows. For each clique  $C$  in  $G$ , let  $\mathcal{L}_C$  be the LSSM given by all Hamiltonians *supported* on  $C$ , i.e. those that act nontrivially only on the tensor factors  $\mathcal{H}_i$  such that  $v_i \in C$  and act as identity on all other subsystems. More precisely,  $\mathcal{L}_C = \otimes_i L_C^i$  where  $L_C^i = \mathbb{S}^{d_i}$  for  $v_i \in C$  and  $L_C^i = \text{Id}_{d_i}$  otherwise. The family of Hamiltonians associated to  $G$  is then  $\mathcal{L}_G = \sum_{C \in \mathcal{C}(G)} \mathcal{L}_C$  where the sum runs over all cliques of  $G$  [WG23, eqn (17)]. The quantum graphical model is  $\text{GM}(\mathcal{L}_G)$  intersected with the space of trace one matrices. The Gibbs variety  $\text{GV}(\mathcal{L}_G)$  gives an algebraic description of this model.

**Example 3.2.1.** Consider the 3-chain graph  $G$  and assume we are in the qubit case, i.e.  $\mathcal{H}_A \cong \mathcal{H}_B \cong \mathcal{H}_C \cong \mathbb{C}^2$ . Then we have  $\mathcal{L}_G = \mathbb{S}^2 \otimes \mathbb{S}^2 \otimes \text{Id}_2 + \text{Id}_2 \otimes \mathbb{S}^2 \otimes \mathbb{S}^2$ . Using numerical algebraic geometry techniques, we verify that no linear or quadratic equations vanish on

$\text{GV}(\mathcal{L}_G)$ . The higher degree equations are not amenable to our computational techniques. The dimension of  $\text{GM}(\mathcal{L}_G)$  is 15 while the dimension of  $\text{GV}(\mathcal{L}_G)$  is at most 22.  $\blacklozenge$

As we have seen in Example 3.2.1 and in several examples in Chapter 2, the defining ideals of Gibbs varieties are often difficult to compute. In view of this, one might hope to simplify the defining equations of the Gibbs variety by restricting the family of Hamiltonians to a subset inside  $\mathcal{L}_G$ . This approach is pursued in the next subsection.

### 3.2.2. Gibbs varieties of unirational varieties of Hamiltonians

A natural subset to consider inside  $\mathcal{L}_G$  is  $X_G := \sum_{C \in \mathcal{C}(G)} X_C$ , where  $X_C$  is the set of decomposable tensors supported on  $C$ . Note that  $X_G$  is not a linear space. However, it is still a unirational variety. This motivates the following extension of the notion of Gibbs varieties.

**Definition 3.2.2.** Let  $X$  be a unirational variety of symmetric matrices of size  $n \times n$ . The *Gibbs variety*  $\text{GV}(X)$  of  $X$  is the Zariski closure of  $\exp(X) \subseteq \mathbb{S}^n$ .

A number of concepts related to Gibbs varieties of linear spaces generalize to the case of smooth unirational varieties of symmetric matrices. If  $X$  is unirational and has dimension  $d$ , then it can be parametrized by rational functions in  $d$  variables  $y_1, \dots, y_d$  [Oja90, Proposition 1.1]. Therefore, one can think of  $X$  as a single matrix with entries in  $\mathbb{C}(y_1, \dots, y_d)$ . The eigenvalues of this matrix are elements of  $\overline{\mathbb{C}(y_1, \dots, y_d)}$  and will be referred to as the eigenvalues of  $X$ . If  $A \in \mathbb{S}^n$ , then the  $X$ -centralizer of  $A$  is the set  $\mathcal{Z}_X(A) = \{B \in X \mid AB - BA = 0\}$ , in full analogy with Definition 2.1.7. We collect properties of Gibbs varieties of unirational varieties in the following statement.

**Proposition 3.2.3.** Let  $X$  be a smooth unirational variety of  $n \times n$ -symmetric matrices of dimension  $d$ . Let  $m$  be the dimension of the  $\mathbb{Q}$ -linear space spanned by the eigenvalues of  $X$  and let  $k$  be the dimension of the  $X$ -centralizer of a generic element of  $X$ . Then  $\dim(\text{GV}(X)) = m + d - k$ . In particular,  $\dim(\text{GV}(X)) \leq n + d$ . Moreover, if  $X$  has distinct eigenvalues, then  $\text{GV}(X)$  is irreducible and unirational.

*Proof.* This proposition generalizes Theorem 2.1.9 and Theorem 2.2.6. Proofs of these statements carry over to the case of unirational varieties, since they only use the fact that an LSSM can be parametrized by rational functions in  $y_1, \dots, y_d$  and do not depend on these functions being linear.  $\square$

Note that symbolic (Algorithm 1) and numerical (Algorithm 2) implicitization methods for Gibbs varieties generalize accordingly.

**Example 3.2.4.** Again, consider the 3-chain graph  $G$  in the qubit case. The associated unirational variety is  $X_G = \{K \otimes L \otimes \text{Id}_2 + \text{Id}_2 \otimes M \otimes N \mid K, L, M, N \in \mathbb{S}^2\} \subseteq \mathbb{S}^8$ . The dimension of  $X$  is equal to 10 inside the 36-dimensional space of symmetric  $8 \times 8$ -matrices. The Gibbs variety  $\text{GV}(X)$  is a 14-dimensional irreducible variety cut out by nine linear forms and 66 quadratic equations in  $\mathbb{S}^8$ . These results were obtained by using techniques of numerical algebraic geometry. More precisely, we create a sample of points on  $\text{GM}(X)$  and then interpolate with polynomials of a fixed degree by setting up a Vandermonde matrix and computing its kernel via QR-decomposition to obtain a sparse representation, see [BKSW18]. This procedure yields polynomials of degree one and two. As these equations cut out an irreducible variety of the correct dimension, we obtain generators of the prime ideal of  $\text{GV}(X)$ .

The equations we obtained exhibit a remarkably simple structure. For instance, all polynomials have coefficients  $\pm 1$  and consist of at most eight terms. It would be very interesting to obtain a theoretical explanation of this phenomenon.  $\blacklozenge$

### 3.3. Toric varieties from quantum exponential families associated to graphs

In this section we explore a completely different family of Hamiltonians  $\mathfrak{H}_G$  associated to a graph  $G$ . This results in quantum exponential families that have a richer structure than the ones considered in Section 3.2. However, it should be emphasized that, unlike the previous constructions, this is not a generalization of classical graphical models. On the other hand, all the results in this section hold for general undirected graphs, not only trees.

#### 3.3.1. Commuting Hamiltonians from graphs

In quantum physics, to an undirected graph  $G$  one associates an LSSM  $\mathfrak{H}_G$  which gives rise to the definition of *graph states*, used in the study of entanglement, e.g. [HEB04]. More precisely, the graph state associated to  $G$  is the *stabilizer state* of  $\mathfrak{H}_G$ . Stabilizer states appear in the framework of the stabilizer formalism used in quantum error correction [NC02, §10.5]. In fact, all results in this section generalize to stabilizers. For more details, we provide an introduction to the stabilizer formalism in Section 3.4. However, instead of studying the graph states associated to  $\mathfrak{H}_G$ , here we focus on its Gibbs variety. This latter perspective gives yet another example of quantum exponential families.

Let  $G = (V, E)$  be a graph with vertices  $V = \{v_1, \dots, v_N\}$ . To each vertex  $v_i$ , we associate a Hamiltonian  $H_i = \bigotimes_{j=1}^N H_{i,j}$  with

$$H_{i,j} = \begin{cases} \sigma_X & \text{if } i = j, \\ \sigma_Z & \text{if } (i, j) \in E, \\ \text{Id}_2 & \text{else.} \end{cases}$$

Here,  $\sigma_X$  and  $\sigma_Z$  are the *Pauli-X* and *Pauli-Z* matrices

$$\sigma_X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Denote the linear span of this set of Hamiltonians by  $\mathfrak{H}_G := \langle H_i \mid i = 1, \dots, N \rangle$ . The Hamiltonians  $H_i$  are elements of the *Pauli group*  $\mathcal{P}_N$  where

$$\mathcal{P}_1 := \{\pm \text{Id}_2, \pm i \text{Id}_2, \pm \sigma_X, \pm i \sigma_X, \pm \sigma_Y, \pm i \sigma_Y, \pm \sigma_Z, \pm i \sigma_Z\}$$

and  $\mathcal{P}_N$  is the set of all  $N$ -fold tensor products of elements of  $\mathcal{P}_1$  equipped with multiplication as the group operation. Here,  $\sigma_Y$  denotes the Pauli-Y matrix

$$\sigma_Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}.$$

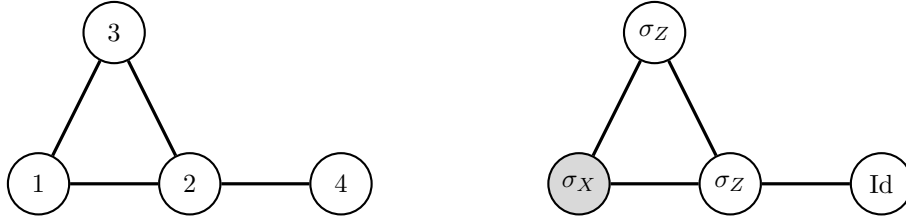
**Example 3.3.1.** Consider the graph  $G$  on four vertices depicted in Figure 3.1. The Hamiltonian  $H_1$  is given by

$$H_1 = \sigma_X \otimes \sigma_Z \otimes \sigma_Z \otimes \text{Id}_2.$$

The linear space  $\mathfrak{H}_G$  is spanned by the four Hamiltonians

$$\sigma_X \otimes \sigma_Z \otimes \sigma_Z \otimes \text{Id}_2, \sigma_Z \otimes \sigma_X \otimes \sigma_Z \otimes \sigma_Z, \sigma_Z \otimes \sigma_Z \otimes \sigma_X \otimes \text{Id}_2, \text{Id}_2 \otimes \sigma_Z \otimes \text{Id}_2 \otimes \sigma_X.$$

◆



**Figure 3.1:** Left: the graph  $G$ . Right: illustration of the Hamiltonian  $H_1$ .

In the following we consider the Gibbs variety of  $\mathfrak{H}_G$ . We start by showing that  $\mathfrak{H}_G$  is a commuting family, implying that  $\text{GV}(\mathfrak{H}_G)$  is *toric* after a linear change of coordinates and the Gibbs manifold  $\text{GM}(\mathfrak{H}_G)$  is semi-algebraic (Theorem 2.1.13).

**Lemma 3.3.2.** Any two matrices  $H, H' \in \mathfrak{H}_G$  commute.

*Proof.* W.l.o.g. assume  $H$  and  $H'$  are generators  $H_m$  and  $H_n$  of  $\mathfrak{H}_G$ . Note that the Pauli matrices satisfy the commutation relation

$$[\sigma_j, \sigma_k] = 2i\epsilon_{jkl}\sigma_l,$$

where  $\epsilon_{jkl}$  is the Levi-Civita symbol (it is zero if  $j, k$  and  $l$  are not pairwise distinct and is the sign of the permutation  $(jkl)$  otherwise) and we denote  $\sigma_1 = \sigma_X, \sigma_2 = \sigma_Y$  and  $\sigma_3 = \sigma_Z$ . For  $P \in \mathcal{P}_N$ , let  $\text{Supp}_X(P) := \{j \in [N] \mid \sigma^{(j)} = \sigma_X\}$  and  $\text{Supp}_Z(P) := \{j \in [N] \mid \sigma^{(j)} = \sigma_Z\}$  denote the supports of  $\sigma_X$  and  $\sigma_Z$ , respectively. Then two Pauli product matrices  $P, Q \in \mathcal{P}_N$  containing only  $\text{Id}_2, \sigma_X$  or  $\sigma_Z$  as tensor factors commute if and only if

$$\#(\text{Supp}_X(P) \cap \text{Supp}_Z(Q)) + \#(\text{Supp}_Z(P) \cap \text{Supp}_X(Q)) \equiv 0 \pmod{2}. \quad (3.3.1)$$

Let  $\mathbb{N}(v)$  denote the set of neighbouring vertices of  $v$  in  $G$ . Assume  $v_n \in \mathbb{N}(v_m)$ ; then the left-hand side of (3.3.1) for  $P = H_m$  and  $Q = H_n$  becomes  $\#\{m\} + \#\{n\} = 2$ . Finally, if  $v_n \notin \mathbb{N}(v_m) \cup \{v_m\}$  the left-hand side of (3.3.1) is just zero.  $\square$

Let us briefly recall from Section 2.1 how to obtain the toric variety and the coordinate change from  $\text{GV}(\mathfrak{H}_G)$ . The symmetric matrices  $H_1, \dots, H_N \in \mathfrak{H}_G$  are simultaneously diagonalizable, i.e. there exist an orthogonal matrix  $U$  and diagonal matrices  $D_1, \dots, D_N$  such that  $U^{-1}H_iU = D_i$  for  $i = 1, \dots, N$ . The exponential of an element in  $\mathfrak{H}_G$  can then be written as

$$\exp(x_1H_1 + \dots + x_NH_N) = U \exp(x_1D_1 + \dots + x_ND_N)U^{-1}$$

and thus  $\text{GV}(\mathfrak{H}_G) = U \cdot \text{GV}(\mathfrak{D}) \cdot U^{-1}$  where  $\mathfrak{D} = \langle D_1, \dots, D_N \rangle$ . Let  $D_i = \text{diag}(d_i), i = 1, \dots, N$  for  $d_i \in \mathbb{R}^{2^N}$  and let  $\mathcal{D} = \langle d_1, \dots, d_N \rangle \subseteq \mathbb{R}^{2^N}$  be the  $\mathbb{R}$ -vector space spanned by the diagonals. Consider the smallest vector subspace  $\mathcal{D}_{\mathbb{Q}} \subseteq \mathbb{R}^{2^N}$  containing  $\mathcal{D}$  that is spanned by elements of  $\mathbb{Q}^{2^N}$  and choose an integral basis  $a_1, \dots, a_N \in \mathbb{Z}^{2^N}$  of  $\mathcal{D}_{\mathbb{Q}}$ . If  $A$  denotes the  $N \times 2^N$ -matrix with rows  $a_1, \dots, a_N$  then  $\text{GV}(\mathfrak{D})$  is the toric variety  $X_A$  associated to  $A$ .

We will need the following standard fact from the theory of quantum stabilizer codes. A proof is provided in Section 3.4.

**Lemma 3.3.3** ([NC02, Prop. 10.5]). Let  $S = \langle P_1, \dots, P_{N-k} \rangle$  be a subgroup of  $\mathcal{P}_N$  generated by  $N-k$  independent<sup>1</sup> and commuting Pauli product matrices such that  $-\text{Id}_{2^N} \notin S$ . Then the vector space  $V_S := \{v \in \mathbb{R}^{2^N} \mid P_i v = v \forall i = 1, \dots, N-k\}$  of simultaneous  $(+1)$ -eigenvectors has dimension  $2^k$ .

<sup>1</sup>i.e.  $\forall i = 1, \dots, N-k: \langle P_1, \dots, \widehat{P}_i, \dots, P_{N-k} \rangle \neq S$  where the hat denotes that this generator is omitted.

**Remark 3.3.4.** Note that any Pauli product matrix  $P \in \mathcal{P}_N$  has eigenvalues  $\pm 1$ , both with multiplicity  $2^{N-1}$  each. Lemma 3.3.3 can then be rephrased as follows: the  $(\pm 1)$ -eigenspaces of  $P_i$  intersect the eigenspaces of all  $P_1, \dots, P_{i-1}$  in half their dimension. This fact is essential to the next theorem establishing a strong connection between quantum exponential families and classical algebraic statistics.

**Theorem 3.3.5.** For any graph  $G$  with  $N$  vertices,  $\text{GV}(\mathfrak{H}_G)$  is an independence model on  $N$  binary random variables after a linear change of coordinates.

*Proof.* As shown above,  $\text{GV}(\mathfrak{H}_G) = U \cdot X_A \cdot U^{-1}$  where the rows of  $A$  are the diagonal entries of  $D_i = U^{-1}H_iU$  for  $i = 1, \dots, N$  and  $X_A$  is the (affine) toric variety associated to  $A$ . By Remark 3.3.4, we can assume  $A$  to be of the form

$$A = \begin{bmatrix} -1 & -1 & -1 & -1 & \dots & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & \dots & 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 & \dots & 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & \dots & 1 & 1 & 1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ -1 & -1 & 1 & 1 & \dots & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & \dots & -1 & -1 & 1 & 1 \\ -1 & 1 & -1 & 1 & \dots & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & \dots & -1 & 1 & -1 & 1 \end{bmatrix} \quad (3.3.2)$$

i.e. the columns of  $A$  are the vertices of the  $N$ -dimensional hypercube  $[-1, 1]^N$ . Thus,  $X_A$  is an independence model on  $N$  binary random variables.  $\square$

**Remark 3.3.6.** The variety  $X_A$  above is not the independence model in its standard description. For example, for  $N = 3$  we have

$$A = \begin{bmatrix} -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 \\ -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 \end{bmatrix};$$

the prime ideal of  $X_A$  is

$$I_A = \langle x_1x_8 - 1, x_6x_7 - x_5x_8, x_4x_7 - x_3x_8, x_2x_7 - 1, x_4x_6 - x_2x_8, \\ x_3x_6 - 1, x_4x_5 - 1, x_3x_5 - x_1x_7, x_2x_5 - x_1x_6, x_2x_3 - x_1x_4 \rangle \subseteq \mathbb{C}[x_1, \dots, x_8].$$

The matrix  $A'$  of the independence model has as columns the vertices of the hypercube  $[0, 1]^N$ . Adding a row of ones to  $A$  and to  $A'$  yields the same variety. Thus,  $X_A$  is an affine patch of the independence model, i.e. of the Segre variety  $\sigma(\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1) \subseteq \mathbb{P}^7$ .

A priori, it is not obvious how to obtain defining equations for  $\text{GV}(\mathfrak{H}_G)$  computationally in an efficient manner. However, Theorem 3.3.5 gives rise to Algorithm 6 making computations of defining ideals for graphs with four or more vertices feasible.

Note that Step 3 in Algorithm 6 can be pre-computed. This allows to reduce finding the equations of  $\text{GV}(\mathfrak{H}_G)$  to a linear algebra problem, therefore reducing the computational complexity. An implementation of this algorithm is available at <https://mathrepo.mis.mpg.de/QuantumGraphicalModels>.

**Example 3.3.7.** Let  $G$  be the graph from Example 3.3.1. Using Algorithm 6, we compute that  $\text{GV}(\mathfrak{H}_G) \subseteq \mathbb{S}^{16}$  is defined by 296 quadratic equations in 136 variables. The average number of terms of each generator is about 1982. This highlights the fact that the equations defining the Gibbs variety can be quite involved. It would be impossible to compute these equations without using the additional structure of  $\mathfrak{H}_G$  being diagonalizable. Both Algorithm 1 and Algorithm 2 failed to compute this example.  $\blacklozenge$



**Algorithm 6** Computing defining equations of  $\text{GV}(\mathfrak{H}_G)$ **Input:** A graph  $G$ **Output:** Polynomials defining  $\text{GV}(\mathfrak{H}_G)$ 

- 1: Compute  $\mathfrak{H}_G = \langle H_1, \dots, H_N \rangle$
- 2:  $U \leftarrow$  matrix simultaneously diagonalizing  $H_1, \dots, H_N$
- 3:  $I \leftarrow$  ideal of the independence model defined by  $A$  as in (3.3.2) in variables  $p_{i,i}$  for  $i = 1, \dots, 2^N$
- 4:  $Y = (y_{ij}) \leftarrow$  linear coordinate change according to  $UPU^{-1}$  where  $P = (p_{ij})$
- 5: **for** all generators  $g_k$  of  $I$  **do**
- 6:      $h_k \leftarrow g_k$  changed to  $Y$ -coordinates
- 7:  $J \leftarrow$  ideal generated by all  $h_k$  and  $y_{ij} = 0$  for all  $i \neq j \in \{1, \dots, 2^N\}$
- 8: **return** a set of generators of  $J$

**3.3.2. Quantum information projections**

Given an arbitrary quantum state  $\rho$  we can ask for the member  $\tilde{\rho} \in \mathcal{Q}$  of some quantum exponential family  $\mathcal{Q} = \text{GM}(\mathcal{L})$  “closest” to  $\rho$ . Here, “closest” means minimizing the *quantum relative entropy*.

**Definition 3.3.8.** The *quantum relative entropy*  $D(\rho||\sigma)$  between a state  $\rho$  and a positive semidefinite operator  $\sigma$  is

$$D(\rho||\sigma) := \begin{cases} \text{Tr}(\rho(\log(\rho) - \log(\sigma))) & \text{if } \text{Supp}(\rho) \subseteq \text{Supp}(\sigma) \\ +\infty & \text{otherwise.} \end{cases}$$

Here, the support of a linear operator is the subspace orthogonal to the kernel with respect to the standard inner product on  $\mathbb{C}^n$ , and all logarithms are taken to have base two.

This is a quantum generalization of the Kullback–Leibler divergence in classical information theory. Note that, similarly to the Kullback–Leibler divergence, the quantum relative entropy is *not* an actual metric as it is not symmetric and does not satisfy the triangle inequality. However, it does satisfy non-negativity (quantum Gibbs’ inequality). More precisely, if  $\text{Tr}(\sigma) \leq 1$  we have  $D(\rho||\sigma) \geq 0$  with equality if and only if  $\rho = \sigma$ . See [Will13, §11.8] for an extensive reference.

**Definition 3.3.9.** The *quantum information projection*  $\tilde{\rho}$  of a quantum state  $\rho$  to a quantum exponential family  $\mathcal{Q}$  is the element of  $\mathcal{Q}$  which is the closest to  $\rho$  with respect to the quantum relative entropy

$$\tilde{\rho} = \underset{\rho' \in \mathcal{Q}}{\text{argmin}} D(\rho||\rho').$$

The quantum information projection is unique and has been characterized in the case where  $\mathcal{Q}$  consists of exponentials of local Hamiltonians [NGKG13, Lemma 2]. Since the Gibbs manifold considered in the previous subsection is semi-algebraic, we can use algebraic techniques to find the quantum information projection in this case. The following theorem gives an algebraic characterization of the quantum information projection for a quantum exponential family  $\mathcal{Q}$  of commuting Hamiltonians, in particular for  $\mathcal{Q} = \text{GM}(\mathfrak{H}_G)$ .

**Theorem 3.3.10.** Let  $\mathfrak{H} = \langle H_1, \dots, H_k \rangle$  be a linear span of commuting Hamiltonians in  $\mathbb{S}_{\mathbb{R}}^d$ , fix a positive definite matrix  $\rho \in \mathbb{S}_{\mathbb{R}}^d$  and let  $b_i := \langle H_i, \rho \rangle = \text{Tr}(H_i \rho)$  for  $i = 1, \dots, k$ . Let  $M_\rho$  be the affine linear space defined by

$$M_\rho := \{A \in \mathbb{S}_{\mathbb{R}}^d \mid \langle H_i, A \rangle = b_i \text{ for } i = 1, \dots, k\}.$$

Then  $M_\rho \cap \text{GM}(\mathfrak{H})$  consists of a unique point  $\rho^*$ . It is the maximiser of the von Neumann entropy inside  $M_\rho$  and the quantum information projection of  $\rho$  to  $\text{GM}(\mathfrak{H})$ .

**Remark 3.3.11.** This result generalizes Birch's Theorem [DSS08, Proposition 2.1.5] to quantum exponential families that become toric after a linear change of coordinates.

*Proof.* The fact that  $M_\rho \cap \text{GM}(\mathfrak{H}) = \{\rho^*\}$  and  $\rho^*$  is the unique point maximising the von Neumann entropy is a direct consequence of Theorem 2.4.1. It remains to show that this point is the quantum information projection of  $\rho$  to  $\text{GM}(\mathfrak{H})$ .

Let  $\tilde{\rho}$  be the quantum information projection of  $\rho$  to  $\text{GM}(\mathfrak{H})$ . As in the discussion preceding Lemma 3.3.3, let  $U$  be the matrix diagonalizing  $\mathfrak{H}$  into  $\mathfrak{D} = \langle D_1, \dots, D_k \rangle$ , i.e.  $H_i = UD_iU^{-1}$  for  $i = 1, \dots, k$ , so  $\tilde{\rho} \in U \text{GM}(\mathfrak{D})U^{-1}$ . Minimizing the quantum relative entropy between  $\rho$  and  $\text{GM}(\mathfrak{H})$  is then equivalent to minimizing the quantum relative entropy between  $U^{-1}\rho U$  and  $\text{GM}(\mathfrak{D})$ . Let  $\sigma := U^{-1}\rho U$ ; then we want to maximize  $\text{Tr}(\sigma \log(\Delta))$  over diagonal matrices  $\Delta \in \text{GM}(\mathfrak{D})$ , i.e. find  $\tilde{\rho}' = \text{diag}(\hat{\delta})$  such that

$$\tilde{\rho}' = \underset{\Delta \in \text{GM}(\mathfrak{D})}{\text{argmax}} \sum_j \sigma_{jj} \log \Delta_{jj}.$$

This is the same problem as finding the maximum likelihood estimator on the exponential family  $\text{GM}(\mathfrak{D})$  given data  $u = (\sigma_{11}, \sigma_{22}, \dots, \sigma_{dd})$ . Note that every coordinate of  $u$  is nonzero. By Birch's Theorem,  $\hat{\delta}$  is the unique point on  $\text{GM}(\mathfrak{D})$  satisfying  $A\hat{\delta} = Au$  where  $A$  is the matrix whose  $i$ th row is the diagonal of  $D_i$  as in (3.3.2). Observe that

$$(Au)_i = \sum_j (D_i)_{jj} \sigma_{jj} = \text{Tr}(D_i \sigma) = \text{Tr}(D_i U^{-1} \rho U) = \text{Tr}(UD_i U^{-1} \rho) = \text{Tr}(H_i \rho) = b_i;$$

analogously,  $(A\hat{\delta})_i = \text{Tr}(H_i \tilde{\rho})$ . This shows  $\tilde{\rho} \in M_\rho \cap \text{GM}(\mathfrak{H})$  and thus  $\tilde{\rho} = \rho^*$ .  $\square$

Theorem 3.3.10 provides a way to compute the quantum information projection to  $\text{GM}(\mathfrak{H}_G)$  algorithmically by using numerical algebraic geometry; concretely, one can first compute  $M_\rho \cap \text{GV}(\mathfrak{H}_G)$  and then choose the unique point lying in the PSD cone.

**Example 3.3.12.** Consider the positive definite matrix

$$\rho = \begin{bmatrix} 84 & -22 & 11 & -51 & -15 & -8 & -26 & 4 \\ -22 & 51 & -5 & -7 & 23 & -13 & 17 & 40 \\ 11 & -5 & 51 & 25 & -16 & -3 & 9 & 28 \\ -51 & -7 & 25 & 70 & -19 & 17 & 18 & -26 \\ -15 & 23 & -16 & -19 & 92 & 32 & 23 & 24 \\ -8 & -13 & -3 & 17 & 32 & 62 & 2 & -36 \\ -26 & 17 & 9 & 18 & 23 & 2 & 94 & 10 \\ 4 & 40 & 28 & -26 & 24 & -36 & 10 & 109 \end{bmatrix}$$

and the 3-chain graph  $G$ . The intersection  $M_\rho \cap \text{GV}(\mathfrak{H}_G)$  consists of six real matrices. Only one of them is positive semidefinite, namely the matrix

$$\tilde{\rho} = \begin{bmatrix} 20.5417 & -12.5 & -20.5 & -12.4746 & -5.5 & 3.34685 & -5.48884 & -3.34006 \\ -12.5 & 20.5417 & 12.4746 & 20.5 & 3.34685 & -5.5 & 3.34006 & 5.48884 \\ -20.5 & 12.4746 & 20.5417 & 12.5 & 5.48884 & -3.34006 & 5.5 & 3.34685 \\ -12.4746 & 20.5 & 12.5 & 20.5417 & 3.34006 & -5.48884 & 3.34685 & 5.5 \\ -5.5 & 3.34685 & 5.48884 & 3.34006 & 20.5417 & -12.5 & 20.5 & 12.4746 \\ 3.34685 & -5.5 & -3.34006 & -5.48884 & -12.5 & 20.5417 & -12.4746 & -20.5 \\ -5.48884 & 3.34006 & 5.5 & 3.34685 & 20.5 & -12.4746 & 20.5417 & 12.5 \\ -3.34006 & 5.48884 & 3.34685 & 5.5 & 12.4746 & -20.5 & 12.5 & 20.5417 \end{bmatrix}.$$

This matrix is the quantum information projection of  $\rho$  to  $\text{GM}(\mathfrak{H}_G)$ .  $\blacklozenge$

### 3.4. Stabilizer formalism

The purpose of this last section is to provide a proof of Lemma 3.3.3 to make this chapter self-contained, and to generalize Theorem 3.3.5 by introducing the stabilizer formalism. This framework is commonly used in quantum error correction for a very convenient description of quantum code spaces. In our exposition we follow [NC02, §10.5.1].

Any subgroup  $S \leq \mathcal{P}_N$  of the Pauli group acts on the vector space of  $N$  qubit states by multiplication. The *vector space stabilized by  $S$*  is denoted  $V_S$  and we call  $S$  the *stabilizer* of  $V_S$ . In quantum error correction,  $V_S$  is the code space.

Let  $S$  be generated by  $S = \langle p_1, \dots, p_l \rangle$ ; the generators  $p_1, \dots, p_l$  are called *independent* if for all  $i = 1, \dots, l$ :  $\langle p_1, \dots, \hat{p}_i, \dots, p_l \rangle \not\subseteq S$ , where the hat means that the element is omitted.

**Lemma 3.4.1** ([NC02, Proposition 10.5]). Let  $S = \langle p_1, \dots, p_{N-k} \rangle \leq \mathcal{P}_N$  be generated by  $N - k$  independent and commuting Pauli product matrices such that  $-\text{Id}_{2^N} \notin S$ . Then  $V_S$  has dimension  $2^k$ .

*Proof.* First note that any Pauli matrix  $\sigma \in \{\sigma_X, \sigma_Y, \sigma_Z\}$  has eigenvalues  $\pm 1$ , and the projector on the  $\pm 1$ -eigenspace of  $\sigma$  is  $\frac{\text{Id}_{2^N} \pm \sigma}{2}$ . For any  $\mathbf{x} = (x_1, \dots, x_{N-k}) \in (\mathbb{Z}/2\mathbb{Z})^{N-k}$ , define

$$P_S^{\mathbf{x}} := \frac{1}{2^{N-k}} \prod_{j=1}^{N-k} (\text{Id}_{2^N} + (-1)^{x_j} p_j);$$

$P_S^{\mathbf{0}}$  is the projector onto  $V_S$  as the composition of projectors onto the eigenspaces of the  $p_i$ 's.

**Claim 3.4.2.** For any  $\mathbf{x} \in (\mathbb{Z}/2\mathbb{Z})^{N-k}$ ,  $\dim(\text{Im}(P_S^{\mathbf{x}})) = \dim(\text{Im}(P_S^{\mathbf{0}}))$ .

We represent a Pauli product matrix  $p \in \mathcal{P}_N$  as a (row) vector  $v_p \in (\mathbb{Z}/2\mathbb{Z})^{2N}$  as follows:

$$(v_p)_i = \begin{cases} 1 & \text{if } i \leq N \text{ and the } i\text{-th tensor factor of } v_p \text{ is either } \sigma_X \text{ or } \sigma_Y, \\ 1 & \text{if } i > N \text{ and the } (i - N)\text{-th tensor factor of } v_p \text{ is either } \sigma_Z \text{ or } \sigma_Y, \\ 0 & \text{else.} \end{cases}$$

Then two Pauli product matrices  $p, p' \in \mathcal{P}_N$  commute if and only if  $v_p \Lambda v_{p'}^T = 0$ , where  $\Lambda$  is the  $2N \times 2N$ -matrix

$$\Lambda = \begin{bmatrix} 0 & \text{Id}_N \\ \text{Id}_N & 0 \end{bmatrix}.$$

Let  $p_1, \dots, p_{N-k}$  be the independent generators of  $S$ . For any  $i = 1, \dots, N - k$  there exists a  $p \in \mathcal{P}_N$  such that  $pp_i p^\dagger = -p_i$  and  $pp_j p^\dagger = p_j$  for all  $j \neq i$ . Indeed, consider the  $(N - k) \times 2N$ -matrix  $P$  with rows  $v_{p_1}, \dots, v_{p_{N-k}}$ ; as the generators are independent, one can check that the rows of  $P$  are linearly independent. Therefore, the linear system  $P\Lambda x = e_i$ , where  $e_i$  is the  $i$ th standard basis vector, has a solution, say  $s \in (\mathbb{Z}/2\mathbb{Z})^{2N}$ . Then we define  $p \in \mathcal{P}_N$  by  $v_p = s^T$ . Thus, for any  $j \neq i$  we have  $v_{p_j} \Lambda v_p = 0$ , so  $p$  and  $p_j$  commute and  $pp_j p^\dagger = p_j$ . Moreover,  $v_{p_i} \Lambda v_p = 1$ , hence  $pp_i p^\dagger = -p_i$ .

The argument above shows that for any  $\mathbf{x} \in (\mathbb{Z}/2\mathbb{Z})^{N-k}$ , there exists  $p_{\mathbf{x}} \in \mathcal{P}_N$  such that  $P_S^{\mathbf{x}} = p_{\mathbf{x}} P_S^{\mathbf{0}} p_{\mathbf{x}}^\dagger$ , proving the claim.

Let  $\mathbf{x}, \mathbf{x}' \in (\mathbb{Z}/2\mathbb{Z})^{N-k}$  be two distinct vectors, i.e. there exists an  $i \in \{1, \dots, N - k\}$  such that  $\mathbf{x}_i \neq \mathbf{x}'_i$ . Then  $\text{Im}(P_S^{\mathbf{x}})$  and  $\text{Im}(P_S^{\mathbf{x}'})$  are orthogonal. Indeed, the Hilbert–Schmidt inner product between  $P_S^{\mathbf{x}}$  and  $P_S^{\mathbf{x}'}$  evaluates to

$$\langle P_S^{\mathbf{x}}, P_S^{\mathbf{x}'} \rangle = \frac{1}{2^{2(N-k)}} \text{Tr} \left( (\text{Id} + p_i)(\text{Id} - p_i) \prod_{j \neq i} (\text{Id} + (-1)^{x_j} p_j)(\text{Id} + (-1)^{x'_j} p_j) \right) = 0$$

as  $(\text{Id} + p_i)/2$  and  $(\text{Id} - p_i)/2$  are projectors on complementary eigenspaces of  $p_i$ .

Finally, observe that

$$\sum_{\mathbf{x} \in (\mathbb{Z}/2\mathbb{Z})^{N-k}} P_S^{\mathbf{x}} = \text{Id}_{2^N},$$

so the  $2^{N-k}$  many vector spaces  $\text{Im}(P_S^{\mathbf{x}})$  form an equidimensional partition of  $\mathbb{C}^{2^N}$ , hence  $\dim(V_S) = \dim(\text{Im}(P_S^{\mathbf{0}})) = 2^k$ .  $\square$

We will now see that Theorem 3.3.5 does not rely on the structure of the graph but generalizes to stabilizers.

**Theorem 3.4.3.** Let  $S = \langle p_1, \dots, p_N \rangle \leq \mathcal{P}_N$  be generated by  $N$  independent and commuting Pauli product matrices such that  $-\text{Id}_{2^N} \notin S$ . Then  $\text{GV}(S)$  is an independence model on  $N$  binary random variables after a linear change of coordinates.

*Proof.* The proof of Theorem 3.3.5 immediately extends to this setup with Lemma 3.4.1.  $\square$

We now summarize open research questions that are motivated by this chapter. Are QCMI varieties irreducible for all trees? As seen in Example 3.2.4, Gibbs varieties of unirational varieties of Hamiltonians defined by graphs exhibit a remarkably simple structure in their defining equations. Is there a theoretical explanation for this fact? We have seen that computing the defining equations of QCMI varieties, Petz varieties and Gibbs varieties is computationally very challenging. Are there ways to exploit the structure and theoretical properties of these varieties to facilitate implicitization, in the spirit of Example 3.3.7?

## Chapter 4

# Minimizing Dual Volumes of Polytopes

This chapter studies the (semi-)algebraic geometry of minimizing volumes of dual polytopes. At this point it is instructive to revisit basics of discrete geometry in Section 1.3. Motivations for our study include optimization, statistics and particle physics.

### 4.1. The Santaló point

We start with some terminology. A polytope  $P \subset \mathbb{R}^m$  is the convex hull of finitely many points. If  $P$  has dimension  $m$ , then each point  $y$  in its interior defines a dual polytope

$$(P - y)^\circ = \{z \in \mathbb{R}^m : \langle y' - y, z \rangle \leq 1, \text{ for all } y' \in P\}.$$

The function  $y \mapsto \text{vol}_m(P - y)^\circ$  is strictly convex on the interior of  $P$ . In fact, this is true when  $P$  is replaced by any convex body, see the proof of Proposition 1(i) in [MW98]. It follows that there is a unique minimizer  $y^* \in \text{int}(P)$ . This is called the *Santaló point* of  $P$ :

$$y^* = \underset{y \in \text{int}(P)}{\text{argmin}} \text{vol}_m(P - y)^\circ = \underset{y \in \text{int}(P)}{\text{argmin}} \int_{(P-y)^\circ} dz_1 \cdots dz_m. \quad (4.1.1)$$

A special property of polytopes, compared to general convex bodies, is that our volume function is *rational*. It follows from Theorems 3.1 and 3.2 in [Gae20] that

$$\text{vol}_m(P - y)^\circ = \gamma \cdot \frac{\alpha_P(y)}{\ell_1(y) \cdots \ell_k(y)}, \quad (4.1.2)$$

where  $\gamma$  is a nonzero real constant,  $\ell_i(y) = 0$  is an affine-linear equation defining the  $i$ -th facet hyperplane of  $P$ , and  $\alpha_P(y)$  is the adjoint polynomial. We will recall a formula for  $\alpha_P$  in Section 4.2. Having established the identity (4.1.2), computing the Santaló point of  $P$  comes down to minimizing a convex rational function or, equivalently, its logarithm.

**Example 4.1.1** ( $m = 2, k = 5$ ). We consider the pentagon  $P$  in  $\mathbb{R}^2$  given by the inequalities

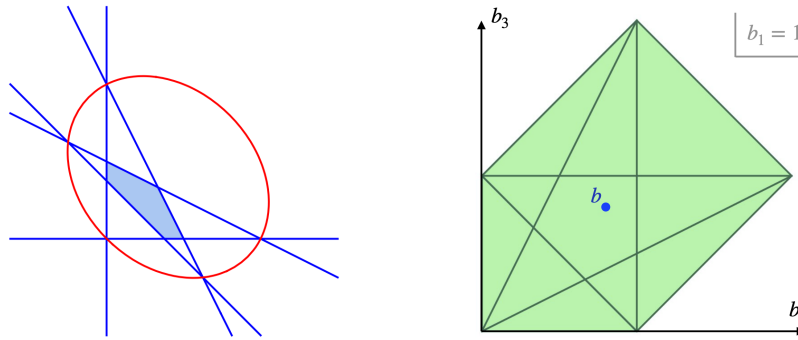
$$y_1 + \frac{1}{5} \geq 0, \quad y_2 + \frac{1}{5} \geq 0, \quad 2y_1 + 2y_2 + \frac{1}{5} \geq 0, \quad -2y_1 - y_2 + \frac{1}{5} \geq 0, \quad -y_1 - 2y_2 + \frac{1}{5} \geq 0.$$

It is shown, together with the poles and zeros of  $\text{vol}_2(P - y)^\circ$ , in Figure 4.1 (left). We have

$$\text{vol}_2(P - y)^\circ = \frac{1}{125} \frac{-50y_1^2 - 25y_1y_2 + 15y_1 - 50y_2^2 + 15y_2 + 11}{(y_1 + \frac{1}{5})(y_2 + \frac{1}{5})(2y_1 + 2y_2 + \frac{1}{5})(-2y_1 - y_2 + \frac{1}{5})(-y_1 - 2y_2 + \frac{1}{5})}. \quad (4.1.3)$$

The Santaló point minimizes this function on  $\text{int}(P)$ :  $y^* = (-0.00311069, -0.00311069)$ .





**Figure 4.1:** Left: the pentagon  $P$  from Example 4.1.1, together with its adjoint curve (red) and facet hyperplanes (blue). Right: a two-dimensional slice of the chamber complex  $\mathcal{C}_A$ .

The first motivation for computing Santaló points comes from convex optimization [NN94]. In that context,  $P$  is the feasible region of a linear program, whose optimal solution is typically a vertex of  $P$ . Interior point methods approximate that vertex by first optimizing a strictly convex (barrier) function. The resulting interior optimizer is then tracked to the optimal vertex by varying a regularization parameter, in the spirit of Section 2.4. For more details, see [Gül96, NN94], where (4.1.2) is called the universal barrier. For a summary, see the introduction of [STVvR24].

We are interested in how the Santaló point varies when the facet hyperplanes of  $P$  are translated. More precisely, we fix a nonnegative  $d \times n$ -matrix  $A \in \mathbb{R}_{\geq 0}^{d \times n}$  of rank  $d$ , none of whose columns is the zero vector, and consider the fibers of the projection  $A : \mathbb{R}_{\geq 0}^n \rightarrow \mathbb{R}^d$ :

$$P_b = \{x \in \mathbb{R}_{\geq 0}^n : Ax = b\}, \quad b \in \text{pos}(A).$$

Here  $\text{pos}(A)$  is the image of  $A : \mathbb{R}_{\geq 0}^n \rightarrow \mathbb{R}^d$ . If  $b$  lies in  $\text{pos}(A)$ , then  $P_b$  is a polytope of dimension  $m = n - d$ . A point  $x$  in its relative interior defines a full-dimensional polytope  $P_b - x$  in the  $(n - d)$ -dimensional vector space  $\ker A \simeq \mathbb{R}^{n-d}$ . We define

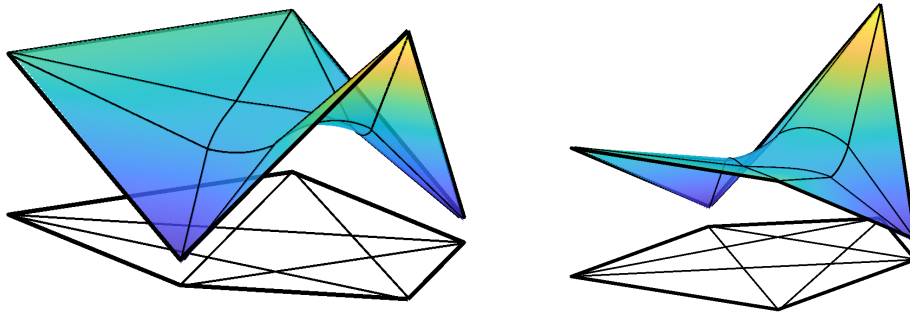
$$V : \mathbb{R}_{> 0}^n \longrightarrow \mathbb{R}_{\geq 0}, \quad x \longmapsto \text{vol}_{n-d}(P_{Ax} - x)^\circ. \quad (4.1.4)$$

This is defined up to a scaling factor, which depends on the choice of basis for  $\ker A$ . We prove that this global volume function is piecewise rational, meaning that it is a rational function when restricted to certain  $n$ -dimensional subcones of  $\mathbb{R}_{> 0}^n$  (Proposition 4.2.5). These subcones correspond to the cells of the chamber complex  $\mathcal{C}_A$  associated to  $A$ , see e.g. [BGS93] for more details. Moreover, on each of these subcones,  $V$  is homogeneous of degree  $d - n$  (Proposition 4.2.5).

Each fiber  $P_b$  has a unique Santaló point. This defines a natural section of  $A : \mathbb{R}_{> 0}^n \rightarrow \mathbb{R}^d$ :

$$x^*(b) = \underset{x \in \text{int}(P_b)}{\text{argmin}} V(x). \quad (4.1.5)$$

The image  $x^*(b)$  of a point  $b$  can be efficiently computed using numerical methods, see Example 4.6.1. The map  $x^* : \text{pos}(A) \rightarrow \mathbb{R}_{> 0}^n$  is piecewise algebraic. Its image is called the Santaló patchwork. We show that the Santaló patchwork is a union of  $d$ -dimensional basic semi-algebraic sets, one for each  $d$ -dimensional cell in the chamber complex  $\mathcal{C}_A$  (Corollary 4.3.3). We give inequalities for each of its pieces (called Santaló patches), and bound the degree of their Zariski closures.



**Figure 4.2:** Visualization of the Santaló patchwork for  $A$  from (4.1.6).

**Example 4.1.2** ( $d = 3, n = 5$ ). The pentagon in Example 4.1.1 is the fiber  $P_b - x$  for the data

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 2 & 1 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 & 1 \end{bmatrix}, \quad b = \frac{1}{5} \begin{bmatrix} 5 \\ 4 \\ 4 \end{bmatrix}, \quad x = \frac{1}{5} [1 \ 1 \ 1 \ 1 \ 1]^T. \quad (4.1.6)$$

The coordinates  $y_1$  and  $y_2$  in Example 4.1.1 are with respect to the following basis of  $\ker A$ :

$$B = \frac{1}{18} \begin{bmatrix} 5 & -4 & 2 & -6 & 3 \\ -4 & 5 & 2 & 3 & -6 \end{bmatrix}^T.$$

The columns of  $A$  are the vertices of a pentagon in  $\mathbb{R}^3$ . They define the polyhedral complex shown in Figure 4.1 (right). The chamber complex  $\mathcal{C}_A$  is the polyhedral fan over that complex. There are 11 3-dimensional cells. Our  $b$  lies in the central pentagonal cell. For any  $x \in \mathbb{R}_{\geq 0}^5$  such that  $Ax$  lies in this cell, we have the following formula for the volume function  $V(x)$ :

$$V(x) = \frac{3x_1x_2x_3 + 2x_1x_3x_5 + 2x_1x_4x_5 + 2x_2x_3x_4 + 2x_2x_4x_5}{x_1x_2x_3x_4x_5}. \quad (4.1.7)$$

This is given by Corollary 4.2.6. To match this with (4.1.3), use  $Ax = b$  and  $B^T x = y$  to switch from  $x$ - to  $(b, y)$ -coordinates and substitute  $b = (1, 4/5, 4/5)$ . A different rational function is needed when  $b$  belongs to a different cell, because the combinatorial type of  $P_b$  changes. For instance, one checks that for  $b = (1, 6/5, 4/5)$ ,  $P_b$  is a quadrilateral. Each cell in  $\mathcal{C}_A$  gives a patch of the Santaló patchwork, which is a 3-dimensional semi-algebraic set in  $\mathbb{R}_{\geq 0}^5$ . Intersecting this with the 4-dimensional simplex  $\{\sum_{i=1}^5 x_i = 1\}$  and projecting to  $\mathbb{R}^3$ , we obtain Figure 4.2.  $\blacklozenge$

Understanding the degrees of Santaló patches relies on insights from algebraic statistics [DSS08]. Minimizing the logarithm of the dual volume has the interpretation of maximum likelihood estimation for a particular discrete statistical model, called Wachspress model [KSS20, Section 2]. Every righthand side vector  $b \in \text{pos}(A)$  defines a Wachspress model. The maximum likelihood degree (ML degree) [CHKS06] of this model is constant for generic  $b$  in the interior of a cell in the chamber complex. We conjecture that, under mild genericity assumptions, it gives a lower bound for the degree of the corresponding Santaló patch, see Conjecture 4.4.7 and Proposition 4.5.4. Example 4.5.8 gives evidence that this lower bound is close to the actual degree of the Santaló patch. We show how to compute the ML degree numerically, and Conjecture 4.5.7 gives a formula for polygons. A sketch of proof is included.

The outline of this chapter is as follows. Section 4.2 studies the volume function (4.1.4). Sections 4.3 and 4.4 describe the Santaló patchwork and its Zariski closure. Section 4.5

makes the link to Wachspress models. Finally, in Section 4.6 we discuss homotopy based methods for computing Santaló points. First, we use monodromy to compute the Santaló point of some fiber  $P_{b_0}$ . Next, we compute the Santaló point of a new fiber  $P_{b_1}$  from that of  $P_{b_0}$ , such that  $b_0$  and  $b_1$  belong to the same chamber of  $\mathcal{C}_A$ . We use numerical homotopy continuation [SW<sup>+</sup>05] to track  $x^*(b_0)$  to  $x^*(b_1)$  along a smooth path on the Santaló patchwork. Our algorithms are implemented in a Julia package `Santaló.jl`, which is available at <https://mathrepo.mis.mpg.de/Santaló>.

The content of this chapter fits nicely into a broader story of semi-algebraic sets in convex optimization, algebraic statistics and particle physics. Different strictly convex objective functions used in interior point methods give rise to other interesting geometric objects, see [DLSV12,STVvR24]. For the log-barrier function  $V(x) = -\sum_{i=1}^n \log x_i$ , the role of the Santaló patchwork is played by the positive reciprocal linear space associated to the row span of the matrix  $A$ . The Santaló point is replaced by the analytic center. Entropic regularization uses  $V(x) = \sum_{i=1}^n x_i \log x_i - x_i$  and leads naturally to consider the positive toric  $d$ -fold associated to  $A$ , with the Birch point being its unique intersection with  $P_b$ . From a statistical point of view, these scenarios correspond to maximum likelihood estimation for linear models and exponential families respectively. Next to optimization and statistics, the dual volume function (4.1.2) shows up in particle physics as the canonical function of  $P$ , viewed as a positive geometry [AHBL17]. This enters in the proof of Propostion 4.2.5. For some specific polytopes,  $V(x)$  is a scattering amplitude [AHBHY18]. Recently, dual volumes have been used in the study of toric singularities [MS21b].

All of these connections motivate our effort to study the Santaló geometry of polytopes. This chapter provides new theoretical insights into Santaló points, and practical tools for computing them. It leads to several new possible research directions, see Section 4.6.

## 4.2. Dual volumes of polytopes

To avoid confusion, below we write  $Q \subset \mathbb{R}^m$  for a full-dimensional polytope (where, usually,  $m = n - d$ ), and  $P_b \subset \mathbb{R}^n$  for the  $(n - d)$ -dimensional fibers of  $A : \mathbb{R}_{\geq 0}^n \rightarrow \text{pos}(A)$ .

This section describes the dual volume function (4.1.2) of a full-dimensional polytope  $Q \subset \mathbb{R}^m$ . We start with the numerator of this rational function, called the *adjoint polynomial*  $\alpha_Q(y)$ . Recall that  $Q$  is *simple* if each vertex is adjacent to exactly  $m$  facets.

Suppose  $Q$  is simple and has minimal facet representation

$$Q = \{y' \in \mathbb{R}^m : \langle w_i, y' \rangle + c_i \geq 0, i = 1, \dots, k\}. \quad (4.2.1)$$

Here  $w_i \in \mathbb{R}^m$  and  $c_i \in \mathbb{R}$ . The adjoint polynomial of  $Q$ , introduced by Warren [War96], is

$$\alpha_Q(y) = \text{vol}_m(Q - y)^\circ \cdot \prod_{i=1}^k (\langle w_i, y \rangle + c_i). \quad (4.2.2)$$

For completeness, we include a proof of a convenient formula for  $\alpha_Q(y)$ . We collect the vectors  $w_i$  in an  $m \times k$  matrix  $W$  and write  $W_I$  for the submatrix of columns indexed by  $I \subset \{1, \dots, k\}$ . Let  $\mathcal{V}(Q)$  be the set of vertices of  $Q$ . For each  $v \in \mathcal{V}(Q)$ , we let  $I(v) = \{i : \langle w_i, v \rangle + c_i = 0\} \subset \{1, \dots, k\}$  be the  $m$ -element index set of the facets containing  $v$ .

**Proposition 4.2.1.** For a simple full-dimensional polytope  $Q \subset \mathbb{R}^m$  with minimal facet representation (4.2.1) the adjoint polynomial  $\alpha_Q(y)$  is given by

$$\alpha_Q(y) = \sum_{v \in \mathcal{V}(Q)} |\det W_{I(v)}| \cdot \prod_{i \notin I(v)} (\langle w_i, y \rangle + c_i). \quad (4.2.3)$$



*Proof.* For  $y \in \text{int}(Q)$  the translated polytope  $Q - y$  has the following facet representation:

$$Q - y = \{y' \in \mathbb{R}^m : \langle w_i, y' \rangle + (\langle w_i, y \rangle + c_i) \geq 0, i = 1, \dots, k\}.$$

The dual polytope is then simplicial and can be described as

$$(Q - y)^\circ = \text{conv} \left( \left\{ \frac{w_i}{\langle w_i, y \rangle + c_i} : i = 1, \dots, k \right\} \right).$$

We compute its volume as the sum of volumes over the pieces of its triangulation:

$$\begin{aligned} \text{vol}_m(Q - y)^\circ &= \sum_{v \in \mathcal{V}(Q)} \text{vol}_m \left( \text{conv} \left( \{0\} \cup \bigcup_{i \in I(v)} \left\{ \frac{w_i}{\langle w_i, y \rangle + c_i} \right\} \right) \right) = \\ &= \sum_{v \in \mathcal{V}(Q)} |\det W_{I(v)}| \prod_{i \in I(v)} (\langle w_i, y \rangle + c_i)^{-1}. \end{aligned}$$

Since by definition  $\alpha_Q(y) = \text{vol}_m(Q - y)^\circ \cdot \prod_{i=1}^k (\langle w_i, y \rangle + c_i)$ , we get the formula in (4.2.3).  $\square$

To avoid confusion, we point out that what we call the adjoint of  $Q$  is the adjoint of the dual polytope  $Q^\circ$  in some of the literature [KR20, War96]. The variety inside  $\mathbb{R}^m$  defined by  $\alpha_Q$  is the *adjoint hypersurface* associated to  $Q$ , see [KR20]. When the facet hyperplanes of  $Q$  form a simple arrangement (that is, the intersection of any  $i$  hyperplanes has codimension  $i$ ), the adjoint hypersurface is the unique hypersurface of minimal degree interpolating the *residual arrangement* of  $Q$ . This arrangement is the union of all affine spaces that are contained in the intersections of facet hyperplanes but do not contain a face of  $Q$  [KR20, Theorem 6]. In Figure 4.1 (left), the residual arrangement consists of 5 points defining a unique adjoint conic.

We now switch back to the setting where  $m = n - d$  and the polytope  $Q$  arises as a fiber  $P_b$  of the linear projection  $A : \mathbb{R}_{\geq 0}^n \rightarrow \mathbb{R}^d$  for some  $A \in \mathbb{R}_{\geq 0}^{d \times n}$ . If  $x$  is an interior point of  $P_b$ , then the translate  $P_b - x$  is a full-dimensional polytope inside  $\ker A \cong \mathbb{R}^{n-d}$ . We are interested in minimizing its dual volume  $\text{vol}_{n-d}(P_b - x)^\circ$  with respect to  $x$ . In order to treat this problem algebraically, we will first project  $P_b$  to  $\ker A$ . To do so, fix an  $n \times (n - d)$ -matrix  $B$  whose columns span  $\ker A$ . The projection of  $P_b$  is denoted by  $Q_b = B^T \cdot P_b$  and the coordinates  $y$  on  $\ker A$  are induced from  $y = B^T x$ .

By construction, the matrix obtained by concatenating  $A$  and  $B^T$  vertically is an  $n \times n$ -matrix of full rank. It therefore defines an invertible coordinate change

$$\begin{bmatrix} b \\ y \end{bmatrix} = \begin{bmatrix} A \\ B^T \end{bmatrix} x. \quad (4.2.4)$$

This means that in order to compute the Santaló point  $x^*(b)$  of  $P_b$ , it is sufficient to compute the Santaló point  $y^*(b)$  of  $Q_b$  and then apply the inverse coordinate change:

$$x^*(b) = \begin{bmatrix} A \\ B^T \end{bmatrix}^{-1} \begin{bmatrix} b \\ y^*(b) \end{bmatrix}. \quad (4.2.5)$$

We will now study the dual volume function  $\text{vol}_{n-d}(Q_b - y)^\circ$  for the polytope  $Q_b$ . Our aim is to show that this is a piecewise rational function of  $y$  and  $b$ . A key role will be played by the *chamber complex*  $\mathcal{C}_A$  of  $\text{cone}(A) = \text{pos}(A)$ , the conical hull of the columns of  $A$ . We now give a definition that is slightly more general than Definition 1.3.16 and is more convenient for computations.

Let  $a_i$  denote the  $i$ -th column of  $A$ . For a nonempty subset  $\sigma \subset [n] = \{1, \dots, n\}$  we define  $A_\sigma = \{a_i : i \in \sigma\}$  to be the submatrix with columns indexed by  $\sigma$ .

**Definition 4.2.2.** For  $b \in \text{cone}(A)$ , define the *chamber*  $C_b := \bigcap_{\text{cone}(A_\sigma) \ni b} \text{cone}(A_\sigma)$ . The *chamber complex* of  $A$  is the collection of all such chambers:

$$\mathcal{C}_A := \{C_b : b \in \text{cone}(A)\}.$$

In the rest of this chapter, full-dimensional chambers are called *cells* of  $\mathcal{C}_A$ .

For more details on the chamber complex, see [BGS93] and [DLRS10, Chapter 5].

**Proposition 4.2.3.** For each  $b$  in the interior of a cell  $C \in \mathcal{C}_A$ , the  $(n-d)$ -dimensional polytopes  $P_b$  and  $Q_b$  are simple, and so are their facet hyperplane arrangements. As  $b$  varies over  $\text{int}(C)$ , the combinatorial types of  $P_b$  and  $Q_b$  are equal and constant.

*Proof.* Since  $b$  is in  $\text{pos}(A)$ , the interior of  $\text{cone}(A)$ ,  $P_b$  has dimension  $n-d$ . Since every vertex  $v$  of  $P_b$  is a solution of  $Av = b$  with  $v_i = 0$  for  $n-d$  entries of  $v$  [BT97, Theorem 2.4], it is on exactly  $n-d$  facet hyperplanes, and the polytope  $P_b$  is simple. For essentially the same reason, the facet hyperplane arrangement of  $P_b$  is simple for any  $b \in \text{int}(C)$ .

The affine span of  $P_b$  is parallel to  $\ker A$ . The matrix  $B$  whose columns span  $\ker A$  defines a projection to  $\ker A$ , and the projected polytope  $Q_b = B^T \cdot P_b$  has the same dimension and combinatorial type as  $P_b$ . The fact that the combinatorial type of  $P_b$  stays the same as  $b$  varies over a given chamber  $C \in \mathcal{C}_A$  appears as Theorem 18 in [AH23].  $\square$

**Example 4.2.4.** The columns of the matrix  $A$  from Example 4.1.2 define the vertices of a pentagon shown in Figure 4.1 (right). The positive hull  $\text{pos}(A)$  is a cone over this pentagon, and the chamber complex  $\mathcal{C}_A$  is the fan over the polyhedral complex obtained by taking the common refinement of all triangulations of this pentagon. The chamber complex has 11 cells: one cone over a pentagon and 10 cones over triangles. When  $b$  is in the central cell, the polytope  $P_b$  is itself a pentagon. When  $b$  is in one of the five cells that share a facet with the central one,  $P_b$  is a quadrilateral. Finally, when  $b$  is one of the five remaining cells,  $P_b$  is a triangle. The following code snippet computes the chamber complex in Macaulay2 [GS].

```
A = matrix{{1,1,1,1,1},{2,1,0,1,0},{1,2,0,0,1}}
B = {{5,-4},{-4,5},{2,2},{-6,3},{3,-6}}
F = gfanSecondaryFan B
all_fulldim_cones = cones(n,F)
all_rays = rays(F)
matrices = apply(all_fulldim_cones, s -> A*submatrix(all_rays,s))
cells_CA = apply(matrices,i->posHull(i))
```

The list `cells_CA` contains all cells of  $\mathcal{C}_A$ . Our computation follows [AH23, Remark 21].  $\blacklozenge$

**Proposition 4.2.5.** Let  $C \in \mathcal{C}_A$  be a cell. Let  $n_C$  be the number of facets of  $P_b$  for  $b \in \text{int}(C)$  and let  $Q_b = B^T \cdot P_b$ , for some kernel matrix  $B \in \mathbb{R}^{n \times (n-d)}$  of  $A$ . The function  $f(b, y) = \text{vol}_{n-d}(Q_b - y)^\circ$  is a homogeneous rational function on

$$\{(b, y) : b \in C \cap \text{pos}(A), y \in \text{int}(Q_b)\},$$

of degree  $d-n$ . Its numerator has degree  $d-n+n_C$  and the denominator has degree  $n_C$ .

*Proof.* We prove the statement for  $b \in \text{int}(C)$ . The result extends to  $b \in C \cap \text{pos}(A)$  by continuity. The dual volume function can be expressed as follows:

$$f(b, y) = \text{vol}_{n-d}(Q_b - y)^\circ = \gamma(b) \cdot \frac{\alpha(b, y)}{\ell_1(b, y) \cdots \ell_{n_C}(b, y)}.$$

Here  $\gamma$  is a nonzero function of  $b$ ,  $\ell_i(b, y) = 0$  is a linear equation defining the  $i$ -th facet hyperplane of  $Q_b$ , and  $\alpha(b, y)$  is the adjoint polynomial of  $Q_b$ , see (4.1.2). The proposition will follow from analyzing these functions. By construction, the  $\ell_i(b, y)$  can be chosen as  $n_C$  of the (homogeneous) linear entries of the following vector:

$$\begin{bmatrix} A \\ B^T \end{bmatrix}^{-1} \begin{bmatrix} b \\ y \end{bmatrix}. \quad (4.2.6)$$

We denote these by  $\ell_i(b, y) = c_i(b) + \langle w_i, y \rangle$ , where  $w_i \in \mathbb{R}^{n-d}$  and  $c_i(b)$  are homogeneous linear forms in  $b$ . By Proposition 4.2.3,  $Q_b$  is a simple polytope. Hence, we can apply (4.2.3) to compute the adjoint polynomial  $\alpha(b, y)$ :

$$\alpha(b, y) = \sum_{v \in \mathcal{V}(Q_b)} \left( |\det W_{I(v)}| \cdot \prod_{i \notin I(v)} (c_i(b) + \langle w_i, y \rangle) \right).$$

Since  $Q_b$  is simple, each vertex is adjacent to exactly  $n - d$  facets. This means that, up to the prefactor,  $\alpha(b, y)$  is a nonzero sum of homogeneous polynomials of degree  $n_C - (n - d)$ . We have now determined the function  $\text{vol}_{n-d}(Q_b - y)^\circ$  up to an overall scaling by  $\gamma(b)$ . The proposition is proved once we show that  $\gamma(b) \in \mathbb{R} \setminus \{0\}$  is a constant. For this, we rely on the theory of *positive geometries* [AHBL17, Lam22]. Since the dual volume is the canonical function of  $Q_b$  as a positive geometry [Lam22, Theorem 3], the residues of this function at the vertices of  $Q_b$  must be equal to  $\pm 1$  for any  $b \in C$ . Taking the residue at  $u \in \mathcal{V}(Q_b)$  results in

$$\text{res}_u \text{vol}_{n-d}(Q_b - y)^\circ = \gamma(b) \kappa_u \frac{\alpha(b, u)}{\prod_{i \notin I(u)} (c_i(b) + \langle w_i, u \rangle)} = \pm 1,$$

where  $\kappa_u = (\det W_{I(u)})^{-1} \in \mathbb{R} \setminus \{0\}$ . Using the fact that  $\alpha(b, u)$  equals

$$\alpha(b, u) = |\det W_{I(u)}| \cdot \prod_{i \notin I(u)} (c_i(b) + \langle w_i, u \rangle),$$

we see that  $\gamma(b) = \pm(\det W_{I(u)} / |\det W_{I(u)}|) = \pm 1$  is indeed a nonzero constant.  $\square$

In  $x$ -coordinates, the proof of Proposition 4.2.5 leads to nice expressions like (4.1.7) for the dual volume  $V(x)$  from (4.1.4). For any  $b \in \text{int}(C)$ , let  $\mathcal{F}_C \subset [n]$  be the indices of the entries of (4.2.6) which correspond to facets of  $Q_b$  and, for each vertex of  $Q_b$ , let  $I(v) \subset \mathcal{F}_C$  be the set of indices of facets containing  $v$ . These sets are independent of the choice of  $b \in \text{int}(C)$ . The set of all index sets  $I(v)$  records the vertices of  $Q_b$  for  $b \in \text{int}(C)$ . We denote it by  $\mathcal{V}_C$ . For an index set  $I \subset [n]$ , we write  $x_I = \prod_{i \in I} x_i$  for the corresponding product of  $x$ -variables. Since  $A \cdot W = 0$ , we have  $\det W_{I(v)} = \pm \gamma \det A_{[n] \setminus I(v)}$  for some  $\gamma \in \mathbb{R}$ , which shows the following.

**Corollary 4.2.6.** Let  $C \in \mathcal{C}_A$  be a cell. The restriction of the dual volume function  $V(x) = \text{vol}_{n-d}(B^T \cdot P_{Ax} - B^T \cdot x)$  to the cone  $\{x \in \mathbb{R}_{>0}^n : Ax \in C\}$  is given by

$$V_C(x) = \gamma \cdot \frac{\sum_{I(v) \in \mathcal{V}_C} |\det A_{[n] \setminus I(v)}| \cdot x_{\mathcal{F}_C \setminus I(v)}}{x_{\mathcal{F}_C}}$$

for some positive constant  $\gamma$  which depends on the choice of  $B$ .

We conclude this section by using Proposition 4.2.5 to derive the degree bound for the algebraic boundary of an important class of objects in convex geometry, the so called *Santaló regions*. These are defined in [MW98] for an arbitrary convex body  $K$  and any  $a \in \mathbb{R}_{>0}$ :

$$K_a := \{x \in \text{int}(K) : \text{vol}(K - x)^\circ - \text{vol}(K - x^*)^\circ \leq a\},$$

where  $x^*$  is the Santaló point of  $K$ . When  $K$  is a polytope, the dual volume function is rational, and  $K_a$  is a semi-algebraic set. When  $K$  is simple, Proposition 4.2.5 says that the algebraic boundary of each Santaló region has degree  $\leq n_C$ , the number of facets of  $K$ .

### 4.3. The Santaló patchwork

As shown in the previous section, the dual volume function  $f(b, y) = \text{vol}_{n-d}(Q_b - y)^\circ$  is a piecewise rational function in  $b$  and  $y$ , with one piece  $f_C(b, y)$  per chamber  $C \in \mathcal{C}_A$ . As noted in the beginning of this chapter, for a fixed  $b$  this function is strictly convex with respect to  $y$  on the interior of  $Q_b$ , and therefore attains a unique minimum at  $y^*(b)$ , which is the Santaló point of  $Q_b = B^T \cdot P_b$ . The Santaló point  $x^*(b)$  of  $P_b$  is then recovered via the linear change of coordinates given in (4.2.5). In this section we introduce the *Santaló patchwork*, a semi-algebraic set keeping track of the Santaló points  $x^*(b)$  for all  $b \in \text{pos}(A)$ .

**Definition 4.3.1.** The *Santaló patchwork*  $\text{SP}(A)$  of  $A \in \mathbb{R}_{\geq 0}^{d \times n}$  is the image of the map  $\phi : \text{pos}(A) \rightarrow \mathbb{R}_{>0}^n$ , which sends  $b$  to the Santaló point  $x^*(b) = \arg \min_{x \in P_b} \text{vol}_{n-d}(Q_b - B^T x)^\circ$ .

**Proposition 4.3.2.** The map  $\phi$  from Definition 4.3.1 is a homeomorphism onto  $\text{SP}(A)$ .

*Proof.* It is convenient to work in  $(b, y)$  coordinates first. Let  $\Sigma(B)$  be the open cone

$$\Sigma(B) = \left\{ (b, y) \in \mathbb{R}^n : \begin{bmatrix} A \\ B^T \end{bmatrix}^{-1} \begin{bmatrix} b \\ y \end{bmatrix} > 0 \right\}.$$

It is clear that  $\Sigma(B) \simeq \mathbb{R}_{>0}^n$  via the linear coordinate change  $\begin{pmatrix} A \\ B^T \end{pmatrix}$ . The map  $\phi$  factors as  $\phi = \begin{pmatrix} A \\ B^T \end{pmatrix}^{-1} \circ \psi$ , where  $\psi(b) = (b, y^*(b)) \in \Sigma(b)$ . It suffices to show that  $\psi$  is a homeomorphism onto its image. First, we note that the restriction of  $\psi$  to the interior of any cell  $C \in \mathcal{C}_A$  is given by algebraic functions and is therefore continuous. Indeed, for a fixed  $b \in \text{int}(C)$ ,  $y^*(b)$  minimizes the rational function  $f_C(b, y) = \text{vol}_{n-d}(Q_b - y)^\circ$ . Let  $b_0$  be a point in the Euclidean boundary  $\partial C \cap \text{pos}(A)$ . By continuity of the dual volume,  $f_C(b_0, y)$  is the dual volume of  $Q_{b_0} - y$  for any  $y \in \text{int}(Q_{b_0})$ . The Santaló point  $y^*(b_0)$  is the unique minimizer of this function on  $\text{int}(Q_{b_0})$ . Since the dual volume is strictly convex on  $\text{int}(Q_{b_0})$  [MW98, Proof of Proposition 1(i)],  $y^*(b_0)$  is a non-degenerate solution to the system of algebraic equations

$$\frac{\partial_{y_i} f_C(b_0, y)}{f_C(b_0, y)} = 0, \quad \text{for } i = 1, \dots, n - d. \quad (4.3.1)$$

By the Implicit Function Theorem, there exist a neighborhood  $\Omega(b_0, C) \subset \text{pos}(A)$  of  $b_0$  and a unique algebraic function  $y_C^*(b)$  such that  $y_C^*(b_0) = y^*(b_0)$  and

$$\frac{\partial_{y_i} f_C(b, y_C^*(b))}{f_C(b, y_C^*(b))} = 0, \quad \text{for } i = 1, \dots, n - d \text{ and } b \in \Omega(b_0, C). \quad (4.3.2)$$

Moreover, being a solution of (4.3.2),  $y_C^*(b)$  minimizes the dual volume  $\text{vol}_{n-d}(Q_b - y)$  for  $b \in \Omega(b_0, C) \cap C$ , that is,  $y_C^*(b) = y^*(b)$  for  $b \in \Omega(b_0, C) \cap C$ . Note that by construction, for

two cells  $C, C' \in \mathcal{C}_A$  and for  $b_0 \in C \cap C' \cap \text{pos}(A)$ , we have  $y_C^*(b_0) = y_{C'}^*(b_0) = y^*(b_0)$ . Since  $\text{pos}(A)$  is covered by  $C \cap \text{pos}(A)$  for cells  $C \in \mathcal{C}_A$ , we get that  $y^*(b)$  is continuous on  $\text{pos}(A)$ . We conclude that  $\psi$  is injective and continuous, so it is a homeomorphism between  $\text{pos}(A)$  and its image, the graph of  $y^*(b)$ . See Figure 4.2 for an illustration of such a graph.  $\square$

We now find a description of  $\text{SP}(A)$  as a finite union of basic semi-algebraic sets. This will imply that  $\text{SP}(A)$  is a semi-algebraic set. For  $b \in \text{int}(C)$ , the Santaló point  $x^*(b)$  is the unique positive point among the critical points of the following (equality) constrained optimization problem:

$$\text{minimize } \log V_C(x), \quad \text{subject to } Ax = b. \quad (4.3.3)$$

Here  $V_C(x)$  is the rational function in Corollary 4.2.6. We simplify the notation by setting

$$\gamma = 1, \quad \alpha_C(x) = \sum_{I(v) \in \mathcal{V}_C} |\det A_{[n] \setminus I(v)}| \cdot x_{\mathcal{F}_C \setminus I(v)}, \quad \text{and} \quad V_C(x) = \frac{\alpha_C(x)}{x_{\mathcal{F}_C}}. \quad (4.3.4)$$

Recall that  $x_{\mathcal{F}_C} = \prod_{i \in \mathcal{F}_C} x_i$  is the product of all variables  $x_i$  which contribute a facet in the cell  $C$ . Note that  $x_i$  contributes a facet if and only if every  $b \in \text{int}(C)$  is in the interior of the convex hull of all but the  $i$ -th column of  $A$ . Furthermore,  $\alpha_C(x)$  depends only on  $x_i, i \in \mathcal{F}_C$ . The partial derivatives of  $\log V_C$  with respect to the variables  $x$  are given by

$$\partial_{x_i}(\log V_C) = \begin{cases} \frac{\partial_{x_i} \alpha_C}{\alpha_C} - \frac{1}{x_i} & i \in \mathcal{F}_C, \\ 0 & i \in [n] \setminus \mathcal{F}_C. \end{cases}$$

Here we write  $\partial_{x_i}$  for  $\frac{\partial}{\partial x_i}$ . Applying the method of Lagrange multipliers to (4.3.3) we obtain the following set of rational function equations in the variables  $x, \lambda = (\lambda_1, \dots, \lambda_d)$ :

$$(\partial_{x_1}(\log V_C), \dots, \partial_{x_n}(\log V_C))^T = A^T \cdot \lambda \quad \text{and} \quad Ax = b.$$

To eliminate the multipliers  $\lambda$ , we apply  $B^T$  to the left- and righthand side of the first set of equations. Writing  $B_C$  for the submatrix of  $B$  whose rows are indexed by  $\mathcal{F}_C$ , we obtain

$$B_C^T \cdot \left( \frac{\partial_{x_i} \alpha_C}{\alpha_C} - \frac{1}{x_i} \right)_{i \in \mathcal{F}_C} = 0 \quad \text{and} \quad Ax = b.$$

These equations make sense for minimizing the dual volume of  $P_b$  only when  $Ax = b \in C \cap \text{pos}(A)$ , and the minimizer is the unique solution in that cone. We define the *Santaló patch* of the cell  $C \in \mathcal{C}_A$  to be the following basic semi-algebraic set:

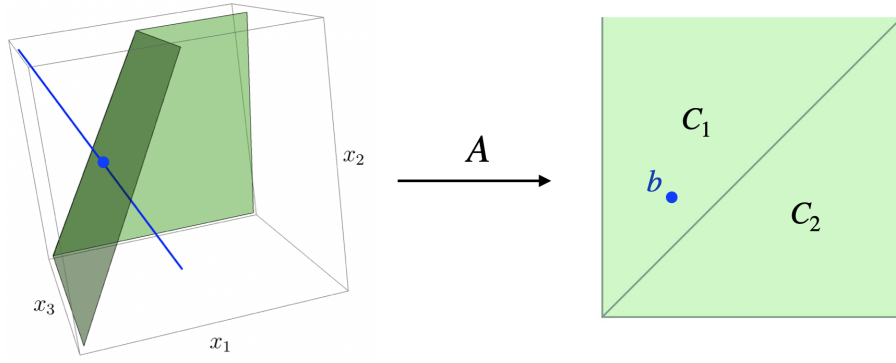
$$S_C = \left\{ x \in \mathbb{R}_{>0}^n : Ax \in C \cap \text{pos}(A) \quad \text{and} \quad B_C^T \cdot \left( \frac{\partial_{x_i} \alpha_C}{\alpha_C} - \frac{1}{x_i} \right)_{i \in \mathcal{F}_C} = 0 \right\}. \quad (4.3.5)$$

Notice that the rational equations in this definition make sense, since  $\alpha_C$  and the coordinate functions  $x_i$  are positive on  $\mathbb{R}_{>0}^n$ . We now state a consequence of the proof of Proposition 4.3.2.

**Corollary 4.3.3.** For a cell  $C \in \mathcal{C}_A$ ,  $\phi|_{C \cap \text{pos}(A)} : C \cap \text{pos}(A) \rightarrow S_C$  is a homeomorphism. In particular, the Santaló patchwork  $\text{SP}(A)$  is the union of the Santaló patches:

$$\text{SP}(A) = \bigcup_{C \in \mathcal{C}_A} S_C,$$

where the union is taken over the cells of  $\mathcal{C}_A$ .



**Figure 4.3:** The Santaló patchwork (left) and chamber complex (right) from Example 4.3.4.

**Example 4.3.4** ( $d = 2, n = 3$ ). Consider the matrix  $A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & 2 \end{pmatrix}$ . The open cone  $\text{pos}(A)$  is  $\mathbb{R}_{>0}^2$  and the polytope  $P_b$ , for  $b \in \text{pos}(A)$ , is a line segment. The complex  $\mathcal{C}_A$  has two cells:

$$C_1 = \{(b_1, b_2) \in \mathbb{R}_{\geq 0}^2 : b_1 \leq b_2\}, \quad C_2 = \{(b_1, b_2) \in \mathbb{R}_{\geq 0}^2 : b_1 \geq b_2\}.$$

For  $C_1$ , we have  $\mathcal{F}_{C_1} = \{1, 2\}$  and  $\mathcal{V}_{C_1} = \{\{1\}, \{2\}\}$ . The dual volume function is

$$V_{C_1}(x_1, x_2, x_3) = \frac{\begin{vmatrix} 1 & 0 \\ 1 & 2 \end{vmatrix} \cdot x_2 + \begin{vmatrix} 2 & 0 \\ 0 & 2 \end{vmatrix} \cdot x_1}{x_1 x_2} = \frac{2x_2 + 4x_1}{x_1 x_2}.$$

Notice that  $V_{C_1}$  does not depend on  $x_3$ , because  $x_3 = 0$  does not contribute a facet to the line segment  $P_b$ ,  $b \in \text{int}(C_1)$ . Setting  $B = \begin{pmatrix} 1 & -2 & 1 \end{pmatrix}^T$  gives  $B_{C_1}^T = \begin{pmatrix} 1 & -2 \end{pmatrix}$ . We find the following inequality description of the Santaló patch  $S_{C_1}$ :

$$\begin{aligned} S_{C_1} &= \left\{ x \in \mathbb{R}_{>0}^3 : 2x_1 + x_2 \leq x_2 + 2x_3, \left( \frac{4}{2x_2 + 4x_1} - \frac{1}{x_1} \right) - 2 \left( \frac{2}{2x_2 + 4x_1} - \frac{1}{x_2} \right) = 0 \right\} \\ &= \left\{ x \in \mathbb{R}_{>0}^3 : x_1 \leq x_3, 2x_1 - x_2 = 0 \right\}. \end{aligned}$$

With an analogous computation we find the following data for the cell  $C_2$ :

$$V_{C_2}(x_1, x_2, x_3) = \frac{4x_3 + 2x_2}{x_2 x_3}, \quad S_{C_2} = \left\{ x \in \mathbb{R}_{>0}^3 : x_1 \geq x_3, 2x_3 - x_2 = 0 \right\}.$$

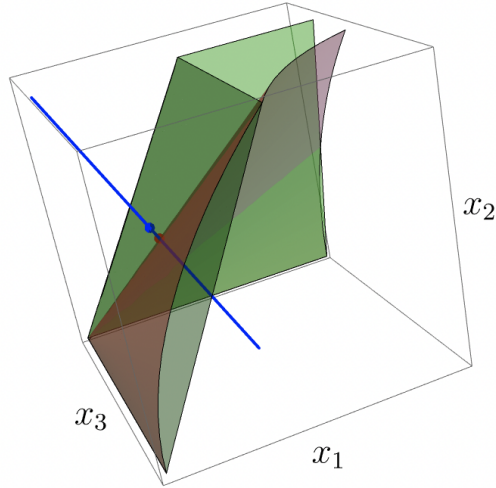
We conclude that the Santaló patchwork  $\text{SP}(A)$  is the union of two 2-dimensional cones in  $\mathbb{R}^3$ . The projection  $A : \text{SP}(A) \rightarrow \text{pos}(A)$  is a homeomorphism, see Figure 4.3. To underscore the analogy with Chapter 2, in Figure 4.4 we show the Santaló patchwork of  $A$  together with the Gibbs manifold of  $A$  (i.e. with the positive part of the toric surface  $x_1 x_3 = x_2^2$  of  $A$ ).  $\blacklozenge$

**Example 4.3.5** ( $d = 2, n = 4$ ). The chamber complex  $\mathcal{C}_A$  for  $A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \end{pmatrix}$  has three cells:

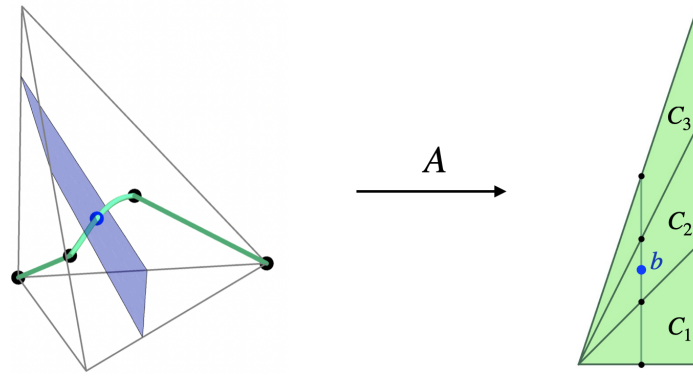
$$C_1 = \{b_2 \geq 0, b_1 \geq b_2\}, \quad C_2 = \{b_1 \leq b_2, 2b_1 \geq b_2\}, \quad C_3 = \{2b_1 \leq b_2, 3b_1 \geq b_2\}.$$

For  $b \in \text{int}(C_1)$  and  $b \in \text{int}(C_3)$ ,  $P_b$  is a triangle, and for  $b \in \text{int}(C_2)$ , it is a quadrilateral:

$$\begin{aligned} \mathcal{F}_{C_1} &= \{2, 3, 4\}, & \mathcal{V}_{C_1} &= \{\{2, 3\}, \{2, 4\}, \{3, 4\}\}, \\ \mathcal{F}_{C_2} &= \{1, 2, 3, 4\}, & \mathcal{V}_{C_2} &= \{\{1, 3\}, \{1, 4\}, \{2, 3\}, \{2, 4\}\}, \\ \mathcal{F}_{C_3} &= \{1, 2, 3\}, & \mathcal{V}_{C_3} &= \{\{1, 2\}, \{1, 3\}, \{2, 3\}\}. \end{aligned}$$



**Figure 4.4:** Comparing the Santaló patchwork (green) and the Gibbs manifold (red).



**Figure 4.5:** The Santaló patchwork (left) and chamber complex (right) from Example 4.3.5.

With these data, it is straightforward to write down the dual volume functions:

$$V_{C_1} = \frac{3x_4 + 2x_3 + x_2}{x_2x_3x_4}, \quad V_{C_2} = \frac{2x_2x_4 + x_2x_3 + 3x_1x_4 + 2x_1x_3}{x_1x_2x_3x_4}, \quad V_{C_3} = \frac{x_3 + 2x_2 + 3x_1}{x_1x_2x_3}.$$

The Santaló patches are 2-dimensional semi-algebraic subsets of  $\mathbb{R}^4$ . They are given by

$$\begin{aligned} S_{C_1} &= \{x > 0, Ax \in C_1, 2x_3 - x_2 = 3x_4 - 2x_3 = 0\}, \\ S_{C_2} &= \{x > 0, Ax \in C_2, x_1x_2 - 6x_1x_4 + 2x_2x_3 + x_3x_4 = x_1x_2 - 4x_1x_3 + 4x_2x_4 - x_3x_4 = 0\}, \\ S_{C_3} &= \{x > 0, Ax \in C_3, -x_3 + 2x_2 = -2x_2 + 3x_1 = 0\}. \end{aligned}$$

To visualize the Santaló patchwork, we restrict  $A : \mathbb{R}_{>0}^4 \rightarrow \text{pos}(A)$  to the probability simplex  $\Delta_3 = \{x > 0, x_1 + x_2 + x_3 + x_4 = 1\}$ . The image of this restriction is the interior of the line segment that is the convex hull of the columns of  $A$ . The set  $\text{SP}(A) \cap \Delta_3$  is a piece-wise algebraic curve, homeomorphic to this line segment, see Figure 4.5. Note the similarity between Figure 4.5 and [STVvR24, Figure 2], where dual volume is replaced by entropy.  $\blacklozenge$

**Example 4.3.6.** The Santaló patchwork for the matrix  $A$  in our running example (Example 4.1.2) consists of 11 patches, one for each cell in the chamber complex shown in Example 4.1.1. These 11 patches are separated by the black curves on the surfaces in Figure 4.2.  $\blacklozenge$

The following statement is a tautology. It emphasizes the role of  $\text{SP}(A)$  in solving (4.1.5).

**Proposition 4.3.7.** The Santaló point of  $P_b$  is given by  $x^*(b) = \phi(b) = \text{SP}(A) \cap P_b$ .

**Example 4.3.8.** For  $A$  from Example 4.3.4, the polytope  $P_b$  for  $b = (1, 2)$  is the blue line segment in Figure 4.3 (left). Its intersection point with  $\text{SP}(A)$  is the center of that line segment, which is its Santaló point  $x^*(b)$ . For  $A$  from Example 4.3.5,  $P_b$  for  $b = (1, 3/2)$  is the blue quadrilateral in Figure 4.5. Again,  $\text{SP}(A) \cap P_b$  is the Santaló point.  $\blacklozenge$

## 4.4. Patch varieties

Section 4.3 describes the set of all solutions to the optimization problem (4.1.5) as a semi-algebraic set called the Santaló patchwork. For computations, it is often convenient to work with algebraic sets instead. This section studies algebraic varieties containing the Santaló patches  $S_C$  defined in (4.3.5). A natural thing to do is take the Zariski closure. We define

$$X_C = \overline{S_C} \subset \mathbb{C}^n.$$

We call  $X_C$  the *patch variety* of the cell  $C$ . A simple way to find equations vanishing on  $X_C$  is by dropping the inequalities in (4.3.5). Let  $\mathcal{X}_C \subset \mathbb{C}^n$  be the Zariski closure of the set

$$\mathcal{X}_C^\circ = \left\{ x \in \mathbb{C}^n : \alpha_C(x) \prod_{i \in \mathcal{F}_C} x_i \neq 0 \quad \text{and} \quad B_C^T \cdot \left( \frac{\partial_{x_i} \alpha_C}{\alpha_C} - \frac{1}{x_i} \right)_{i \in \mathcal{F}_C} = 0 \right\}.$$

**Theorem 4.4.1.** The patch variety  $X_C$  is a  $d$ -dimensional irreducible component of  $\mathcal{X}_C$ .

*Proof.* We switch to  $(b, y)$ -coordinates using the transformation from (4.2.4). We view

$$B_C^T \cdot \left( \frac{\partial_{x_i} \alpha_C}{\alpha_C} - \frac{1}{x_i} \right)_{i \in \mathcal{F}_C} = 0 \tag{4.4.1}$$

as equations in  $y$ , parametrized by  $b$ . For  $b_0 \in C \cap \text{pos}(A)$ , by strict convexity of  $V_C(x)$ , the Santaló point  $x^*(b_0) \sim (b_0, y^*(b_0))$  corresponds to an isolated solution. By [SW<sup>+</sup>05, Theorem A.14.1], it follows that  $x^*(b_0)$  lies on a  $d$ -dimensional irreducible component  $\mathcal{Z}^\circ$  of  $\mathcal{X}_C^\circ$ , and hence on an irreducible component  $\mathcal{Z} = \overline{\mathcal{Z}^\circ}$  of  $\mathcal{X}_C$ . This is true for every  $b_0 \in C \cap \text{pos}(A)$ , so that  $S_C$  is contained in  $\mathcal{Z}$ . Hence,  $X_C \subset \mathcal{Z}$  and it has dimension at most  $d$ . By Corollary 4.3.3,  $S_C$  is  $d$ -dimensional, so  $X_C$  has dimension at least  $d$ . We conclude that  $X_C = \mathcal{Z} \subset \mathcal{X}_C$ .  $\square$

**Remark 4.4.2.** Notice that, by construction, the Santaló patch  $S_C$  is stable under simultaneous scaling of the coordinates:  $x^*(t \cdot b) = t \cdot x^*(b)$  for any  $t \in \mathbb{R}_{>0}$ . It follows that the ideal of  $X_C$  can be generated by homogeneous equations. Furthermore, since the equations defining  $\mathcal{X}_C^\circ$  are homogeneous (of degree  $-1$ ),  $I(\mathcal{X}_C)$  is a homogeneous ideal as well.

If  $A \in \mathbb{Q}_{\geq 0}^{d \times n}$  has rational entries, the vanishing ideal  $I(\mathcal{X}_C)$  of  $\mathcal{X}_C$  can be computed using computer algebra software such as `Macaulay2` [GS] or `Oscar.jl` [OSC24] as follows. Consider the ideal of the ring  $\mathbb{Q}[(x_i)_{i \in \mathcal{F}_C}, (\sigma_i)_{i \in \mathcal{F}_C}, z]$  generated by the  $n - d + n_C + 1$  equations

$$B_C^T \cdot (z \partial_{x_i} \alpha_C - \sigma_i)_{i \in \mathcal{F}_C}, \quad x_i \sigma_i - 1, i \in \mathcal{F}_C, \quad \text{and} \quad \alpha_C(x) z - 1.$$

From this ideal, eliminate the variables  $\sigma_i, i \in \mathcal{F}_C$  and  $z$ . The result is  $I(\mathcal{X}_C)$ .



**Example 4.4.3.** We perform the computation explained above for our running Example 4.1.2, for the 3-dimensional cell  $C \in \mathcal{C}_A$  with five facets containing  $b = (1, 4/5, 4/5)$ . The adjoint is

$$\alpha_C(x) = 3x_1x_2x_3 + 2x_1x_3x_5 + 2x_1x_4x_5 + 2x_2x_3x_4 + 2x_2x_4x_5,$$

i.e., the numerator of (4.1.4). The elimination takes place in a polynomial ring with 11 variables. The ideal  $I(\mathcal{X}_C)$  is prime, homogeneous, and of degree 14. It is generated by five quintics. Here is how to compute  $\alpha_C$  and  $I(\mathcal{X}_C)$  using our Julia package `Santalo.jl`, available at the online repository <https://mathrepo.mis.mpg.de/Santalo>:

---

```
using Santalo # load the package 1
A = [1 1 1 1 1; 2 1 0 1 0; 1 2 0 0 1]; b = 1//5*[5; 4; 4]; 2
R, alpha = adjoint_x(A,b) 3
T, J = ideal_XC(A,b) 4
```

---

The outputs in line 3 are the adjoint `alpha` =  $\alpha_C$  and a polynomial ring `R` containing it. In line 4, we compute the ideal `J` =  $I(\mathcal{X}_C)$  and a polynomial ring `T` containing it.  $\blacklozenge$

Next, we ask whether  $\mathcal{X}_C$  may fail to be equidimensional. I.e., can it have components of dimension  $> d$ ? We do not know the answer, but we expect that for general matrices we even have  $X_C = \mathcal{X}_C$  (see Conjecture 4.4.7). We show that the answer is *no* if we perturb the objective function  $V_C(x)$  slightly. More precisely, we consider the new objective function

$$V_{C,u}(x) = \frac{\alpha_C(x)^{u_0}}{\prod_{i \in \mathcal{F}_C} x_i^{u_i}}. \quad (4.4.2)$$

Here  $u_0, u_i, i \in \mathcal{F}_C$  are new parameters. Setting  $u = \mathbf{1} = (1, \dots, 1)$  recovers our original objective function  $V_C$ . We will see in Section 4.4 that these new parameters have a natural statistical interpretation. The critical point equations of  $\log V_{C,u}$  define the incidence

$$\mathcal{X}_C^\circ = \left\{ (x, u) \in \mathbb{C}^n \times \mathbb{C}^{n_C+1} : \alpha_C(x) \prod_{i \in \mathcal{F}_C} x_i \neq 0 \quad \text{and} \quad B_C^T \cdot \left( \frac{u_0 \partial_{x_i} \alpha_C}{\alpha_C} - \frac{u_i}{x_i} \right)_{i \in \mathcal{F}_C} = 0 \right\}.$$

We write  $\pi_u : \mathcal{X}_C^\circ \rightarrow \mathbb{C}^{n_C+1}$  for the projection  $(x, u) \mapsto u$ , and denote its fiber  $\pi_u^{-1}(u)$  by  $\mathcal{X}_{C,u}^\circ$ . The variety  $\mathcal{X}_C^\circ$  is  $\mathcal{X}_{C,1}^\circ$ . The Zariski closure of  $\mathcal{X}_C^\circ$  is  $\mathcal{X}_C \subset \mathbb{C}^n \times \mathbb{C}^{n_C+1}$ . Fibers of  $\pi_u : \mathcal{X}_C \rightarrow \mathbb{C}^{n_C+1}$  are denoted by  $\mathcal{X}_{C,u}$ . We have  $\overline{\mathcal{X}_{C,u}^\circ} \subset \mathcal{X}_{C,u}$ , and in particular  $\mathcal{X}_C \subset \mathcal{X}_{C,1}$ .

**Proposition 4.4.4.** The varieties  $\mathcal{X}_C^\circ, \mathcal{X}_C$  are irreducible of dimension  $n_C + 1 + d$ . There is a dense open subset  $U \subset \mathbb{C}^{n_C+1}$  such that, for  $u \in U$ ,  $\mathcal{X}_{C,u}$  is pure dimensional of dimension  $d$ .

*Proof.* We consider the other projection  $\pi_x : \mathcal{X}_C^\circ \rightarrow \mathbb{C}^n$  which sends  $(x, u)$  to  $x$ . A fiber  $\pi_x^{-1}(x)$  is defined by linear equations in  $u_0, u_i, i \in \mathcal{F}_C$ . These equations are linearly independent, because  $B_C$  has rank  $n - d$ . This last claim follows from the fact that the rows of  $B$  giving rise to  $B_C$  are indexed by  $\mathcal{F}_C$ , which means that they contain the rays of the normal fan to a full-dimensional polytope  $Q_b$  for  $b \in \text{int}(C)$ . Hence, all fibers of  $\pi_x$  are linear, and hence irreducible, of dimension  $n_C + 1 - n + d$ . By [Sha13, Chapter 1, §6, Theorem 8],  $\mathcal{X}_C^\circ$  is irreducible of dimension  $n_C + 1 + d$ . The same holds for  $\mathcal{X}_C$ . Since the map  $\pi_u : \mathcal{X}_C \rightarrow \mathbb{C}^{n_C+1}$  is dominant, the proposition now follows from [Sha13, Chapter 1, §6, Theorem 7].  $\square$

The following statement summarizes the role of these varieties in the study of the Santaló point of  $P_b$ : they provide useful semi-algebraic descriptions.

**Theorem 4.4.5.** Let  $b \in C \cap \text{pos}(A)$  for some cell  $C \in \mathcal{C}_A$  and let  $P_b^\circ = \text{relint}(P_b)$ . The Santaló point  $x^*(b)$  is given by

$$x^*(b) = \text{SP}(A) \cap P_b^\circ = S_C \cap P_b^\circ = X_C \cap P_b^\circ = \mathcal{X}_C^\circ \cap P_b^\circ = \mathcal{X}_C \cap P_b^\circ = \mathcal{X}_{C,1}^\circ \cap P_b^\circ = \mathcal{X}_{C,1} \cap P_b^\circ.$$

*Proof.* The first two equalities are essentially Proposition 4.3.7. The equality  $x^*(b) = \mathcal{X}_C^\circ \cap P_b^\circ$  follows from strict convexity of the dual volume function on  $P_b$ : there is only one critical point of  $\log V_C$  on  $P_b^\circ$ . Since  $(\mathcal{X}_C \setminus \mathcal{X}_C^\circ) \cap \mathbb{R}_{>0}^n = \emptyset$ , replacing  $\mathcal{X}_C^\circ$  with  $\mathcal{X}_C$  does not change the intersection with  $P_b^\circ$ . The equality  $x^*(b) = X_C \cap P_b^\circ$  now follows from  $x^*(b) \in X_C \subset \mathcal{X}_C$ . The last two equalities follow from  $\mathcal{X}_{C,1}^\circ = \mathcal{X}_C^\circ$  and  $(\mathcal{X}_{C,1} \setminus \mathcal{X}_{C,1}^\circ) \cap \mathbb{R}_{>0}^n = \emptyset$ .  $\square$

Next, we state a naive degree bound for the varieties defined in this section.

**Proposition 4.4.6.** For  $\bullet = X_C, \mathcal{X}_C$  or  $\mathcal{X}_{C,u}$ , for generic  $u$ , we have the inequality

$$\deg(\bullet) \leq (2n_C - n + d - 1)^{n-d}.$$

*Proof.* For  $\mathcal{X}_C$ , this follows from clearing denominators in (4.4.1) and applying Bézout's bound [Har13, Chapter I, Theorem 7.7]. For  $X_C$ , we use Theorem 4.4.1. Finally, for  $\mathcal{X}_{C,u}$ , note that for generic  $u \in \mathbb{C}^{n_C+1}$  we have  $\mathcal{X}_{C,u} = \overline{\mathcal{X}_{C,u}^\circ}$ . Adding the parameters  $u$  to the equations (4.4.1) does not change the Bézout number.  $\square$

The bound from Proposition 4.4.6 is pessimistic. E.g., for Example 4.4.3 it reads  $14 \leq 49$ . In particular, the varieties  $X_C = \mathcal{X}_C$  and  $\mathcal{X}_{C,u}$  have the same degree in that example. In the next section, we use insights from algebraic statistics to prove a lower bound on  $\deg \mathcal{X}_{C,u}$  for generic  $u$  (Corollary 4.5.5). That bound is relevant for our homotopy method for computing Santaló points in Section 4.6. Also, in experiments, we find that it approximates the actual degree more closely (Example 4.5.8). As motivated by the next conjecture, which is suggested by the examples we computed, we here mean both the degree of  $X_C$  and  $\mathcal{X}_{C,u}$ .

**Conjecture 4.4.7.** For generic matrices  $A \in \mathbb{R}_{\geq 0}^{d \times n}$  and for each cell  $C \in \mathcal{C}_A$ , there exists a dense open subset  $U \subset \mathbb{C}^{n_C+1}$  such that the variety  $\mathcal{X}_{C,u}$  is irreducible of dimension  $d$  for  $u \in U$ . Moreover,  $\mathbf{1} \in U$  and we have  $\mathcal{X}_{C,1} = \mathcal{X}_C = X_C$ .

## 4.5. Wachspress models

In algebraic statistics [DSS08, Sul18], a statistical model for a discrete random variable with  $N+1$  states is the intersection of a complex algebraic variety  $M_{\mathbb{C}}$  with the probability simplex

$$\Delta_N = \{(p_0, \dots, p_N) \in \mathbb{R}_{>0}^{N+1} : p_0 + \dots + p_N = 1\}.$$

We denote this model by  $M = M_{\mathbb{C}} \cap \Delta_N$ , and require that this intersection is non-empty. For our purposes, it suffices to consider parametric models, i.e., models that come with a rational parametrization. This is true for many commonly used models, including exponential families and (conditional) independence models. Let  $p_i(y) = p_i(y_1, \dots, y_m)$ ,  $i = 0, \dots, N$  be rational functions of  $m < N$  variables such that  $\sum_{i=0}^N p_i = 1$ . The variety  $M_{\mathbb{C}}$  is the closure of the image of the rational map  $\mathbb{C}^m \dashrightarrow \mathbb{C}^{N+1}$  given by  $y \mapsto (p_0(y), \dots, p_N(y))$ . *Maximum likelihood estimation* for the model  $M$  means finding the probability distribution  $p \in M$  which makes an experimental observation  $u \in \mathbb{N}^{N+1}$  most likely. More precisely, suppose that state  $i$  was observed  $u_i$  times in an experiment. One maximizes the *log-likelihood function*

$$L_u = \log p_0^{u_0} p_1^{u_1} \cdots p_N^{u_N}$$

subject to the constraint  $p \in M$ . To study this problem algebraically, one often relaxes it to finding all complex critical points of  $L_u$  on an open subset of  $M_{\mathbb{C}}$ . In our parametric setting, we solve the system of rational function equations

$$\partial_{y_i} L_u(y) = \sum_{j=0}^N u_j \frac{\partial_{y_i} p_j(y)}{p_j(y)} = 0, \quad i = 1, \dots, m \quad \text{for } y \in \mathbb{C}^m \setminus D. \quad (4.5.1)$$

Here  $D \subset \mathbb{C}^m$  is the union of the supports of the divisors  $\text{div}(p_j)$ ,  $j = 0, \dots, N$ . That is, it is the union of all hypersurfaces in  $\mathbb{C}^m$  along which one of the  $p_j$  has a zero or a pole. We refer to these equations as the *likelihood equations* for the model  $M$ . The number of complex solutions  $y \in \mathbb{C}^m \setminus D$  for generic, complex data  $u \in \mathbb{C}^{N+1}$  is an invariant called the *maximum likelihood degree* (ML degree) of  $M_{\mathbb{C}}$  [CHKS06], which we denote by  $\text{MLdeg}(M_{\mathbb{C}})$ . This assumes that the map  $\mathbb{C}^m \setminus D \rightarrow M_{\mathbb{C}}$  given by  $y \mapsto (p_0(y), \dots, p_N(y))$  is birational.

The models that are relevant to our story are called *Wachspress models*. These are associated to simple polytopes  $Q \subset \mathbb{R}^m$ , and the number of states equals the number of vertices  $|\mathcal{V}(Q)|$ . We use the notation (4.2.1) for the face description of  $Q$ . The parametrizing functions of our model are naturally obtained from the formula (4.2.3) for the adjoint:

$$p_v(y) = \frac{|\det W_{I(v)}| \cdot \prod_{i \notin I(v)} (c_i + \langle w_i, y \rangle)}{\alpha_Q(y)}, \quad v \in \mathcal{V}(Q). \quad (4.5.2)$$

This gives a rational map  $\mathbb{C}^m \dashrightarrow \mathbb{P}^{|\mathcal{V}(Q)|-1}$ , whose image closure  $M_{\mathbb{C}}(Q)$  is the *Wachspress variety* of  $Q$ . Note that the coordinates  $p_v$  for  $v \in \mathcal{V}(Q)$  sum to 1 by construction. These varieties appear in the context of geometric modelling [GPS10], and Wachspress surfaces were studied in [IS14]. To the best of our knowledge, the interpretation as a statistical model first appeared in [KSS20, Section 2]. Bayesian integrals for these models were studied in [BSST23]. The divisor  $D$  from (4.5.1) for the Wachspress model  $M_{\mathbb{C}}(Q)$  is the union of the adjoint hypersurface  $\alpha_Q(y) = 0$  and the facet hyperplanes  $c_i + \langle w_i, y \rangle = 0$ . We denote this by  $D(Q)$ .

**Lemma 4.5.1.** Let  $Q \subset \mathbb{R}^m$  be a simple polytope with Wachspress model  $M_{\mathbb{C}}(Q)$ . Let  $\mathcal{H} \subset \mathbb{P}^N$  be the divisor  $(\sum_{i=0}^N p_i) \prod_{i=0}^N p_i = 0$  for  $N = |\mathcal{V}(Q)| - 1$ . The map  $\psi: \mathbb{C}^m \setminus D(Q) \rightarrow M_{\mathbb{C}}(Q) \setminus \mathcal{H}$  given by  $y \mapsto (p_v(y))_{v \in \mathcal{V}(Q)}$ , with  $p_v$  from (4.5.2), is an isomorphism.

*Proof.* First note that the morphism  $\psi$  is well-defined. The functions  $p_v$  are regular on  $\mathbb{C}^m \setminus D(Q)$ , and the image of  $\mathbb{C}^m \setminus D(Q)$  is contained in the complement of  $\mathcal{H}$ .

It remains to show that  $\psi$  is invertible. Consider the automorphism  $\varphi$  of  $\mathbb{P}^N$  defined by  $p_v \mapsto \frac{p_v}{|\det W_{I(v)}|}$ . The map  $\varphi \circ \psi$  is a restriction of the *Wachspress map* defined in [KR20, Equation (5)], which is invertible by [KR20, Theorem 4]. Thus,  $\psi$  is invertible too.  $\square$

**Corollary 4.5.2.** The maximum likelihood degree  $\text{MLdeg}(M_{\mathbb{C}}(Q))$  of the Wachspress model of  $Q$  equals the absolute value of the Euler characteristic  $\chi(\mathbb{C}^m \setminus D(Q)) = \chi(M_{\mathbb{C}}(Q) \setminus \mathcal{H})$ .

*Proof.*  $M_{\mathbb{C}}(Q) \setminus \mathcal{H}$  is smooth (Lemma 4.5.1), so [HS14, Theorem 1.7] implies the statement.  $\square$

Solving the likelihood equations of  $M_{\mathbb{C}}(Q)$  with data  $u$  is equivalent to computing the intersection of the fiber  $\mathcal{X}_{C,w}^{\circ}$ , defined in Section 4.3, with a linear space. The parameters  $u' \in \mathbb{C}^{n_C+1}$  are obtained from  $u$  via a linear map. This is the content of our next theorem.

**Theorem 4.5.3.** Let  $Q_{b_0} = B^T \cdot P_{b_0}$  for  $b_0 \in C \cap \text{pos}(A)$ , where  $C \in \mathcal{C}_A$  is a cell in the chamber complex of  $A$ . The complex critical points of the log-likelihood function  $L_u(y)$  for

the Wachspress model  $M_{\mathbb{C}}(Q_{b_0})$  with data  $u = (u_v)_{v \in \mathcal{V}(Q_{b_0})}$  are in one-to-one correspondence with the complex critical points of  $V_{C,u'}(x)$  from (4.4.2) on  $\{Ax = b_0\}$ , where  $u'$  has entries

$$u'_0 = - \sum_{v \in \mathcal{V}(Q_b)} u_v, \quad u'_i = - \sum_{v: i \notin I(v)} u_v, \quad i \in \mathcal{F}_C. \quad (4.5.3)$$

More precisely, the critical points  $y_{\text{crit}} \in \mathbb{C}^{n-d} \setminus D(Q_{b_0})$  of  $L_u(y)$  are  $B^T \cdot x_{\text{crit}}$ , where  $x_{\text{crit}}$  ranges over the points in the intersection  $\mathcal{X}_{C,u'}^{\circ} \cap \{Ax = b_0\}$ ,

*Proof.* The likelihood function  $L_u(y)$  for the data  $u = (u_v)_{v \in \mathcal{V}(Q_{b_0})}$  is given by

$$\exp L_u(y) = \prod_{v \in \mathcal{V}(Q_{b_0})} \left( \frac{|\det W_{I(v)}| \cdot \prod_{i \in \mathcal{F}_C \setminus I(v)} (c_i(b_0) + \langle w_i, y \rangle)}{\alpha_{Q_{b_0}}(y)} \right)^{u_v} = V_{C,u'} \left( \begin{bmatrix} A \\ B^T \end{bmatrix}^{-1} \begin{bmatrix} b_0 \\ y \end{bmatrix} \right).$$

This uses the change of coordinates (4.2.4):  $x_i = c_i(b_0) + \langle w_i, y \rangle$ . The chain rule gives

$$\partial_{y_j} \log V_{C,u'} = \sum_{i \in \mathcal{F}_C} \partial_{x_i} \log V_{C,u'} \cdot \frac{dx_i}{dy_j}.$$

It follows that the likelihood equations for the Wachspress model  $M_{\mathbb{C}}(Q_{b_0})$  are equivalent to

$$W_C^T \cdot \left( \frac{u'_0 \partial_{x_i} \alpha_C}{\alpha_C} - \frac{u'_i}{x_i} \right)_{i \in \mathcal{F}_C} = 0 \quad \text{and} \quad Ax = b_0.$$

Here  $Ax = b_0$  is  $b = b_0$  in  $x$ -coordinates, and  $W_C$  is the submatrix of the matrix  $W$  of facet normals whose rows are indexed by  $\mathcal{F}_C$ . The column span of  $W_C$  equals that of  $B_C$  by construction, so these are precisely the equations for  $\mathcal{X}_{C,u'}^{\circ} \cap \{Ax = b_0\}$ .  $\square$

Our next statement uses the following definition. An isolated solution  $x$  to the  $n$  equations

$$B_C^T \cdot \left( \frac{\tilde{u}_0 \partial_{x_i} \alpha_C}{\alpha_C} - \frac{\tilde{u}_i}{x_i} \right)_{i \in \mathcal{F}_C} = 0 \quad \text{and} \quad Ax = b_0 \quad (4.5.4)$$

for fixed  $\tilde{u} \in \mathbb{C}^{n_c+1}$  is *regular* if the rank of the Jacobian matrix at  $x$  is  $n$ .

**Proposition 4.5.4.** Let  $b_0 \in C \cap \text{pos}(A)$  for some cell  $C \in \mathcal{C}_A$ . There is a dense open subset  $U \subset \mathbb{C}^{n_c+1}$  such that, for  $\tilde{u} \in U$ , the set  $I(\tilde{u}, b_0) := \mathcal{X}_{C,\tilde{u}}^{\circ} \cap \{Ax = b_0\}$  consists of  $\text{MLdeg}(M_{\mathbb{C}}(Q_{b_0}))$  regular points. Moreover, the number of regular isolated points in  $I(\tilde{u}, b_0)$  for any  $\tilde{u}$  is at most  $\text{MLdeg}(M_{\mathbb{C}}(Q_{b_0}))$ .

*Proof.* The data points  $u = (u_v)_{v \in \mathcal{V}(Q_{b_0})}$  for the Wachspress model  $M_{\mathbb{C}}(Q_{b_0})$  parametrize a linear subspace  $H$  of  $\mathbb{C}^{n_c+1}$  via (4.5.3). By Theorem 4.5.3 and the definition of the ML degree, the number of points in  $I(u', b_0)$  is  $\text{MLdeg}(M_{\mathbb{C}}(Q_{b_0}))$  for generic  $u' \in H$ . By Corollary 4.5.2, this number equals the signed Euler characteristic of  $\mathbb{C}^m \setminus D(Q_{b_0})$ . By [Huh13, Theorem 1], that Euler characteristic is the number of regular complex critical points of

$$V_{C,\tilde{u}} \left( \begin{bmatrix} A \\ B^T \end{bmatrix}^{-1} \begin{bmatrix} b_0 \\ y \end{bmatrix} \right) = \frac{\alpha_{Q_{b_0}}(y)^{\tilde{u}_0}}{\prod_{i \in \mathcal{F}_C} (c_i(b_0) + \langle w_i, y \rangle)^{\tilde{u}_i}}.$$

for generic  $\tilde{u} \in \mathbb{C}^{n_c+1}$ . The final statement about the upper bound follows from the fact that the generic number of regular isolated solutions to the system of equations (4.5.4) equals the maximal number of regular isolated solutions, see for instance [SW<sup>+</sup>05, Theorem 7.1.1].  $\square$

**Corollary 4.5.5.** For any cell  $C \in \mathcal{C}_A$  and generic  $\tilde{u} \in \mathbb{C}^{n_C+1}$  the degree of the variety  $\mathcal{X}_{C,\tilde{u}} \subset \mathbb{C}^n$  is at least  $\text{MLdeg}(M_{\mathbb{C}}(Q_{b_0}))$ , where  $b_0$  is a generic point in  $\text{int}(C)$ .

*Proof.* By Proposition 4.4.4,  $\mathcal{X}_{C,\tilde{u}}$  is pure dimensional of dimension  $d$  for generic  $u$ . Its degree is the maximal number of regular intersection points with a linear space of codimension  $d$ . This is at least the cardinality of  $I(\tilde{u}, b_0)$ . The statement is a consequence of Proposition 4.5.4.  $\square$

Though the Santaló point of  $Q_{b_0}$  is one of the regular intersection points in  $I(\mathbf{1}, b_0) = \mathcal{X}_{C,\mathbf{1}} \cap \{Ax = b_0\}$  (Theorem 4.4.5), the usefulness of the results in this section for our original problem may seem somewhat mysterious. It will become clear in Section 4.6 that Proposition 4.5.4 is crucial for our homotopy continuation based algorithm for computing Santaló points.

**Remark 4.5.6.** Dual volume minimization is not the only convex optimization problem on  $P_b$  that has the interpretation of a maximum likelihood estimation problem. Other commonly used objective functions lead to maximum likelihood estimation for different models. We briefly discuss the cases  $V(x) = -\sum_{i=1}^n \log x_i$  (log-barrier) and  $V(x) = \sum_{i=1}^n x_i \log x_i - x_i$  (entropic regularization) mentioned in the beginning of this chapter. In each case, there are  $N + 1 = n$  states. For ease of exposition, we make some additional assumptions on  $A$ .

First, for  $V(x) = -\sum_{i=1}^n \log x_i$ , assume that the entries of each column of  $A$  sum to the same number  $c$ . The statistical model  $M$  in this context is the linear model obtained by intersecting the row span  $M_{\mathbb{C}}$  of  $A$  with  $\Delta_{n-1}$ . It is parametrized by  $p_i(y) = (y^T a_i)/(y^T A \mathbf{1})$ , where  $a_i$  is the  $i$ -th column of  $A$  and  $\mathbf{1} \in \mathbb{R}^n$  is the all-ones vector. One checks that the maximum likelihood estimate for the data  $u = (1, \dots, 1)$  is the unique positive minimizer of the log-barrier function  $V(x)$  on the affine-linear space  $\{Ax = b\}$ , where  $b = c^{-1} A \mathbf{1}$ .

For  $V(x) = \sum_{i=1}^n x_i \log x_i - x_i$ , the model comes from a toric variety. We assume that the first row of  $A$  is the all-ones vector  $\mathbf{1}$  and write  $a_i \in \mathbb{R}_{\geq 0}^d$  for the  $i$ -th column. These columns define a monomial map, whose image is  $M_{\mathbb{C}}$ . Concretely, let  $f(y) = y^{a_1} + y^{a_2} + \dots + y^{a_n}$  and consider the rational parametrization functions  $p_i(y) = y^{a_i}/f(y)$ , parametrizing  $M_{\mathbb{C}}$ . For any data vector  $u = (u_1, \dots, u_n) \in \mathbb{N}^n$ , let  $\tilde{u} = (\sum_{i=1}^n u_i)^{-1} \cdot u$  be the *empirical distribution*. As a consequence of Birch's theorem [DSS08, Proposition 2.1.5], the maximum likelihood estimate for the model  $M$  is the unique positive minimizer of the entropy  $V(x)$  on  $\{Ax = A\tilde{u}\}$ .

There is no explicit formula yet for the maximum likelihood degree of the Wachspress model  $M_{\mathbb{C}}(Q)$ . We end the section with conjectures for polygons in the plane. We represent a *generic*  $n$ -gon by a fiber  $P_b$  of  $A : \mathbb{R}_{\geq 0}^n \rightarrow \text{cone}(A)$ , where  $A \in \mathbb{R}_{\geq 0}^{(n-2) \times n}$  is generic among those matrices for which there is a cell in  $\mathcal{C}_A$  whose fibers are  $n$ -gons. Concretely, let

$$C_{\max} = \text{cone}(A_{[n] \setminus \{1\}}) \cap \text{cone}(A_{[n] \setminus \{2\}}) \cap \dots \cap \text{cone}(A_{[n] \setminus \{n\}}) \neq \emptyset \quad \text{and} \quad \dim(C_{\max}) = n - 2.$$

This uses the notation introduced before Definition 4.2.2. In general,  $C_{\max}$  is a union of cells in  $\mathcal{C}_A$ . We pick  $b \in \text{int}(C)$  for any cell  $C \subset C_{\max}$ .

**Conjecture 4.5.7.** Let  $Q = B^T \cdot P_b$  be a generic  $n$ -gon. The maximum likelihood degree of the corresponding Wachspress model is  $\text{MLdeg}(M_{\mathbb{C}}(Q)) = (n-1)(n-2) + (n-3)(n-5) - 1$ .

*Sketch of proof.* By Corollary 4.5.2, we have  $\text{MLdeg}(M_{\mathbb{C}}(Q)) = \chi(\mathbb{C}^2 \setminus D(Q))$ , where  $D_Q$  is the curve  $\{\alpha_Q(y) \prod_{i=1}^n l_i(y) = 0\}$ . Here we write  $l_i(y) = c_i + \langle w_i, y \rangle$  for the equations of the lines defining the edges of  $Q$ . The excision property of the Euler characteristic gives

$$\chi(\mathbb{C}^2 \setminus D(Q)) = \chi(\mathbb{C}^2 \setminus \{\prod_{i=1}^n l_i(y) = 0\}) - \chi(\{\alpha_Q(y) = 0\} \setminus \{\prod_{i=1}^n l_i(y) = 0\}).$$

Since the line arrangement of  $l_1, \dots, l_n$  is generic, the first term is  $\binom{n-1}{2}$  [HKS05, Equation (8)]. On the second term, we use excision once more:

$$\chi(\{\alpha_Q(y) = 0\} \setminus \{\prod_{i=1}^n l_i(y) = 0\}) = \chi(\{\alpha_Q(y) = 0\}) - \chi(\{\alpha_Q = 0\} \cap \{\prod_{i=1}^n l_i(y) = 0\}).$$

Here the second term is  $-\binom{n-1}{2} - 1$ , the number of *residual points* of  $Q$  [KPR<sup>+</sup>21, Section 2.1]. What's missing is the Euler characteristic of the affine curve  $\chi(\{\alpha_Q(y) = 0\})$ . We conjecture that, for generic  $Q$ , this curve is generic in the sense of [Huh13, Theorem 3], with Newton polytope equal to that of  $(1 + y_1 + y_2)^{n-3}$ . That would imply that its Euler characteristic equals  $-(n-3)^2 + 2(n-3)$ . Summing all this up gives the formula in the conjecture.  $\square$

In the spirit of Corollary 4.5.5, we can compare the number  $(n-1)(n-2) + (n-3)(n-5) - 1$  to the degree of the variety  $\mathcal{X}_{C,u}$ , and hence that of  $\mathcal{X}_C$  and  $X_C$  (Conjecture 4.4.7).

**Example 4.5.8.** For  $n = 3, 4, \dots, 11$  we generate a totally positive matrix  $A \in (\mathbb{R})_{\geq 0}^{(n-2) \times n}$  (meaning that all  $(n-2)$ -minors are positive) and we pick a cell  $C \subset C_{\max}$ . Using the numerical homotopy continuation techniques discussed in the next section, we compute that

$n$	3	4	5	6	7	8	9	10	11
$(n-1)(n-2) + (n-3)(n-5) - 1$	1	4	11	22	37	56	79	106	137
$\deg(\mathcal{X}_{C,u}) = \deg(\mathcal{X}_C) = \deg(X_C)$	1	4	14	27	44	65	90	119	152

For instance, for  $n = 5$ , a generic linear space  $\{\tilde{A}x = \tilde{b}\}$  of dimension 2 intersects  $\mathcal{X}_{C,u}$  in 14 points. By Proposition 4.5.4, the *special* linear space  $\{Ax = b\}$  leads to only 11 points. Hence, the lower bound in 4.5.5 may be strict. The table leads us to conjecture that for  $n \geq 5$ ,

$$\deg(X_C) = (n-1)(n-2) + (n-3)(n-5) - 1 + 2(n-3) - 1.$$

Code for computing these degrees is found at <https://mathrepo.mis.mpg.de/Santalo>.



## 4.6. Computing Santaló points

We discuss how to compute Santaló points numerically. We consider two different situations. First, the input is a polytope  $Q \subset \mathbb{R}^m$ , and the output is its Santaló point  $y^*$  from (4.1.1). Our continuation algorithm exploits the likelihood geometry from Section 4.5. The second scenario computes the Santaló point  $x^*(b_1)$  from  $x^*(b_0)$ , assuming  $b_1$  lies in the same cell  $C \in \mathcal{C}_A$  as  $b_0$ . The strategy here is to track a real path on the Santaló patch  $S_C$ . These algorithms are implemented in Julia (v1.9.1) using `Oscar.jl` (v0.14.0) [OSC24] and `HomotopyContinuation.jl` (v2.9.3) [BT18]. All code is available at <https://mathrepo.mis.mpg.de/Santalo>.

The computational paradigm behind our algorithms is that of *homotopy continuation*. We briefly recall the main idea and refer to the standard textbook [SW<sup>+</sup>05] for more details. Let  $F : (\mathbb{C}^n \setminus D) \times \mathbb{C}^m \rightarrow \mathbb{C}^n$  be a map whose coordinates are rational functions in  $x = (x_1, \dots, x_n)$ , depending polynomially on  $m$  parameters  $q = (q_1, \dots, q_m)$ . We assume that the denominators of these functions do not depend on  $q$ , and their vanishing locus is contained in the hypersurface  $D \subset \mathbb{C}^n$ , so that  $F$  is a regular map. We consider the incidence variety

$$Y = F^{-1}(0) = \{(x, q) \in (\mathbb{C}^n \setminus D) \times \mathbb{C}^m : F(x, q) = 0\}.$$

A fiber of the natural projection  $\pi_q : Y \rightarrow \mathbb{C}^m$  is denoted by  $Y_{q_0} = \pi_q^{-1}(q_0)$ . It consists of all solutions to the  $n$  equations in  $n$  variables  $F(x, q_0) = 0$ . A solution  $(x_0, q_0)$  in  $Y_{q_0}$  is

called *isolated* and *regular* if the Jacobian of  $F$  (with respect to  $x$ ) evaluated at  $(x_0, q_0)$  is an invertible  $n \times n$ -matrix. Typically, one has computed all isolated regular solutions in  $Y_{q_0}$  and is interested in computing those in  $Y_{q_1}$ , for some parameters  $q_0 \neq q_1 \in \mathbb{C}^m$ . Homotopy continuation rests on the *parameter continuation theorem* [SW<sup>+</sup>05, Theorem 7.1.4]. First, this states that the number of isolated regular solutions in  $Y_{q_0}$  is constant for  $q_0 \in \mathbb{C}^m \setminus \nabla$ , where  $\nabla \subset \mathbb{C}^m$  is a proper subvariety. Second, let  $\gamma : [0, 1] \rightarrow \mathbb{C}^m$  be a continuous path such that  $\gamma(0) = q_0$ ,  $\gamma(1) = q_1$  and  $\gamma([0, 1]) \cap \nabla = \emptyset$ . Since  $q_0 \notin \nabla$ , each isolated regular solution  $(x_0, q_0) \in Y_{q_0}$  defines a unique smooth solution path  $(t, x(t))$  satisfying

$$F(x(t), \gamma(t)) = 0, \quad t \in [0, 1], \quad x(0) = x_0.$$

Moreover, the limits of all these solution paths as  $t \rightarrow 1$  contain all isolated regular solutions in  $Y_{q_1}$ . Numerical path trackers, such as that implemented in `HomotopyContinuation.jl`, track these solution paths numerically for  $t$  going from 0 to 1. For obvious reasons, the system of equations  $F(x, q_0) = 0$  is called the *start system*, and  $F(x, q_1) = 0$  is the *target system*.

A useful algorithm for finding all isolated regular solutions in  $Y_{q_0}$ , i.e., the solutions to the start system, is itself based on homotopy continuation. It uses *monodromy loops* [DHJ<sup>+</sup>19]. The method needs the assumption that one solution  $(x_0, q_0) \in Y_{q_0}$  is known. One chooses  $\gamma$  to be a closed path, i.e.,  $\gamma(0) = \gamma(1) = q_0$ . If this path encircles a ramification point of the branched cover  $\pi_q : Y \rightarrow \mathbb{C}^m$ , then the corresponding solution path  $(t, x(t))$  may provide a new regular isolated solution in  $Y_{q_0}$ :  $x(1) \neq x(0)$ . If  $Y$  is irreducible, then all isolated regular solutions can be found by repeating this process [DHJ<sup>+</sup>19, Remark 2.2]. To know when enough loops are tracked, it is very useful to compute the maximal number of solutions from a theoretical argument. This is one of the purposes of Proposition 4.5.4 and Conjecture 4.5.7. The monodromy method, and in particular its implementation in the command `monodromy_solve` in `HomotopyContinuation.jl`, is very efficient and reliable in practice.

### 4.6.1. From likelihood equations to dual volume

Let  $Q \subset \mathbb{R}^m$  be a full-dimensional simple polytope with minimal facet representation

$$Q = \{y \in \mathbb{R}^m : W y + c \geq 0\}, \quad \text{for } W \in \mathbb{R}^{n \times m}, c \in \mathbb{R}^n.$$

Let  $A \in \mathbb{R}^{d \times n}$  be a cokernel matrix of  $W$  ( $A \cdot W = 0$ ). Here  $d = n - m$ , and  $A$  can be chosen so that its entries are nonnegative. Setting  $x = W y + c$ , we see that  $Q$  is a projection of  $P_b = \{x \in \mathbb{R}_{\geq 0}^n : Ax = b\}$ , with  $b = Ac$ . More precisely,  $Q$  is given by  $W^\dagger \cdot (P_b - c)$ , where  $W^\dagger \in \mathbb{R}^{m \times n}$  is the pseudo-inverse of  $W$ . Though we assumed nonnegative entries to guarantee compact fibers of  $A : \mathbb{R}_{\geq 0}^n \rightarrow \text{pos}(A)$ , it is not necessary to find a nonnegative representation for computing the Santaló point. We think of the likelihood equations (4.5.4) as a system of equations with variables  $x_1, \dots, x_n$  and parameters  $q = (u_0, (u_i)_{i \in \mathcal{F}_C})$ :

$$F(x; q) = \begin{pmatrix} B_C^T \cdot \left( \frac{u_0 \partial_{x_i} \alpha_C}{\alpha_C} - \frac{u_i}{x_i} \right)_{i \in \mathcal{F}_C} \\ Ax - b \end{pmatrix} = 0. \quad (4.6.1)$$

In order to solve this for generic parameters  $q_0 \in \mathbb{C}^{n_C+1}$  using monodromy loops, we need to compute one regular solution in  $Y_{q_0}$ . This is done as follows. Select a random point  $x_0 \in \mathbb{C}^n$  so that  $Ax_0 = b$  and solve the linear system  $F(x_0; q)$  for  $q$ . We can pick any solution to these linear equations as the start parameters  $q_0$ . Since  $Y = \mathcal{X}_C^\circ$  is irreducible, see Proposition

4.4.4, all other solutions to  $F(x, q_0)$  can be found using monodromy loops. By Proposition 4.5.4, the number of solutions is the ML degree of the Wachspress model  $M_{\mathbb{C}}(Q)$ .

Once we have computed  $Y_{q_0}$ , we set  $\gamma(t) = (1-t) \cdot q_0 + t \cdot \mathbf{1}$  and track the  $\text{MLdeg}(M_{\mathbb{C}}(Q))$ -many solution paths for  $t \in [0, 1]$ . Precisely one of the end points is positive. Indeed, the regular isolated solutions for  $q_1 = \mathbf{1}$  are critical points of the logarithm of the dual volume function on  $Q$ . Among them, the Santaló point is the unique positive point, by convexity.

**Example 4.6.1.** We illustrate the code on our running example using the data in (4.1.6):

```
using Santalo # load the package 1
A = [1 1 1 1 1; 2 1 0 1 0; 1 2 0 0 1] 2
B = transpose(1//18*[5 -4 2 -6 3; -4 5 2 3 -6]) 3
b = 1//5*[5; 4; 4] 4
Q = compute_Q(A,b,B) # Q = B~T*Pb 5
ystar = get_santaló_point(Q) # Santaló point in y-coordinates 6
xstar = get_santaló_point(A,b) # Santaló point in x-coordinates 7
```

The result  $y^*$  is as reported in Example 4.1.1, and  $x^* \approx (0.197, 0.197, 0.188, 0.210, 0.210)$ .  $\blacklozenge$

**Example 4.6.2.** The user can also construct a polytope  $Q$  using the functionalities of `Oscar.jl` and use it as input for the function `get_santaló_point`. As a 3D example, we consider the permutahedron; a simple polytope with  $f$ -vector  $(24, 36, 14)$ .

```
using Oscar # load the package Oscar to construct polytopes 1
Q = project_full(permutahedron(3)) 2
ystar = get_santaló_point(Q) # output: (2.5, 2.5, 2.5) 3
```

Here  $Q$  is the convex hull of all points  $(j, k, l)$ , where  $(i, j, k, l) \in S_4$  is a permutation of  $(1, 2, 3, 4)$ . This permutahedron is represented by the following values for  $A$  and  $b$ :

$$A = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 3 \\ 7 \\ 4 \\ 5 \\ 5 \\ 5 \\ 3 \\ 5 \\ 5 \\ 5 \\ 5 \\ 7 \end{bmatrix}.$$

We note that  $b$  does not lie in the interior of a full dimensional cell of  $\mathcal{C}_A$ : the facet hyperplane arrangement of the permutahedron is not simple (see Proposition 4.2.3). Still, because  $Q$  is a simple polytope, the adjoint polynomial  $\alpha_C(x)$  can be computed using the formula in (4.3.4). It has degree 11, and all its coefficients are equal:

$$\begin{aligned} & x_1x_2x_3x_4x_5x_6x_7x_8x_9x_{12}x_{14} + x_1x_2x_3x_4x_5x_6x_8x_9x_{12}x_{13}x_{14} + x_1x_2x_3x_4x_5x_6x_8x_{10}x_{11}x_{12}x_{13} + \\ & x_1x_2x_3x_4x_5x_8x_9x_{11}x_{12}x_{13}x_{14} + x_1x_2x_3x_4x_5x_8x_{10}x_{11}x_{12}x_{13}x_{14} + x_1x_2x_3x_4x_6x_7x_8x_9x_{10}x_{11}x_{14} + \\ & x_1x_2x_3x_4x_6x_7x_8x_9x_{11}x_{12}x_{14} + x_1x_2x_3x_5x_6x_7x_9x_{10}x_{11}x_{13}x_{14} + x_1x_2x_3x_5x_6x_7x_{10}x_{11}x_{12}x_{13}x_{14} + \\ & x_1x_2x_3x_5x_7x_8x_9x_{11}x_{12}x_{13}x_{14} + x_1x_2x_3x_5x_7x_8x_{10}x_{11}x_{12}x_{13}x_{14} + x_1x_2x_3x_6x_7x_8x_9x_{10}x_{11}x_{13}x_{14} + \\ & x_1x_2x_3x_6x_7x_8x_9x_{11}x_{12}x_{13}x_{14} + x_1x_2x_4x_5x_6x_7x_8x_9x_{10}x_{11}x_{14} + x_1x_3x_4x_5x_6x_7x_8x_{10}x_{11}x_{12}x_{13} + \\ & x_1x_3x_4x_5x_6x_7x_9x_{10}x_{11}x_{13}x_{14} + x_1x_3x_4x_5x_6x_7x_{10}x_{11}x_{12}x_{13}x_{14} + x_1x_4x_5x_6x_7x_8x_9x_{10}x_{11}x_{13}x_{14} + \\ & x_2x_3x_4x_5x_6x_7x_8x_9x_{10}x_{12}x_{14} + x_2x_3x_4x_5x_6x_8x_9x_{10}x_{11}x_{12}x_{13} + x_2x_3x_4x_5x_6x_8x_9x_{10}x_{12}x_{13}x_{14} + \\ & x_2x_4x_5x_6x_7x_8x_9x_{10}x_{11}x_{12}x_{14} + x_3x_4x_5x_6x_7x_8x_9x_{10}x_{11}x_{12}x_{13} + x_4x_5x_6x_7x_8x_9x_{10}x_{11}x_{12}x_{13}x_{14}. \end{aligned}$$



This is found using `adjoint_x(A, b)`, as in Example 4.4.3. The command `get_santalo_point` computes the Santaló point by first solving the likelihood equations for random parameters:

```
A, b, W, c = free_representation(Q) 1
solve_likelihood_startsystem(A, b) 2
```

The first line computes the representations  $P_b = \{x \geq 0, Ax = b\}$  and  $Q = \{Wy + c \geq 0\}$ . The result of line 2 shows that the ML degree of the Wachspress model of the permutahedron is 569. Interestingly, we find that the Santaló point of the permutahedra of dimensions 2, 3, 4 and 5 is  $A^\dagger \cdot b$ . That is, it is the closest point to the origin satisfying  $Ax = b$ .  $\blacklozenge$

### 4.6.2. Tracking paths on Santaló patches

Suppose the Santaló point  $x^*(b_0)$  of  $P_{b_0}$  was computed for some  $b_0 \in \text{int}(C)$ , where  $C \in \mathcal{C}_A$  is a cell. We are interested in computing  $x^*(b_1)$  for some  $b_1 \in C$  contained in the same cell. Note that  $b_1$  is not necessarily contained in the interior of  $C$ . In particular  $P_{b_1}$  is not necessarily simple. This time, the parametric equations depend only on  $b$ :

$$F(x; b) = \begin{pmatrix} B_C^T \cdot \left( \frac{\partial_{x_i} \alpha_C}{\alpha_C} - \frac{1}{x_i} \right)_{i \in \mathcal{F}_C} \\ Ax - b \end{pmatrix} = 0. \tag{4.6.2}$$

The path  $\gamma$  is  $\gamma(t) = (1 - t) \cdot b_0 + t \cdot b_1$ . At every  $t \in [0, 1]$  the solution path  $(t, x(t))$  described by the Santaló point is smooth: it is a regular solution to the equations  $F(x; \gamma(t))$  by convexity of the dual volume. In this homotopy, we track only one path, and all computations can be done over the real numbers. This feature of our problem makes the procedure extra efficient.

**Example 4.6.3.** In our running Example 4.1.1, we can set  $b_0 = (1, 4/5, 4/5)$ ,  $b_1 = (1, 1, 4/5)$ , see Figure 4.1. The fiber  $P_{b_1}$  is a quadrilateral:  $b_1$  lies on the boundary of the pentagonal cell in  $\mathcal{C}_A$ . As  $t$  moves from 0 to 1, the Santaló point  $x^*(\gamma(t))$  of  $P_{\gamma(t)}$  describes a path on the Santaló patchwork from Figure 4.2. In the  $(y_1, y_2)$ -plane, this is a path in the interior of pentagon  $Q_{\gamma(t)}$  which degenerates continuously to a quadrilateral. The Santaló point  $x^*(b_0)$  was computed in Example 4.6.1. The command `santalo_path` computes  $x^*(b_1)$  from  $x^*(b_0)$ :

```
A = [1 1 1 1 1; 2 1 0 1 0; 1 2 0 0 1] 1
b0 = 1//5*[5; 4; 4]; b1 = 1//5*[5; 5; 4]; 2
x0 = get_santalo_point(A, b0) 3
x1 = santalo_path(A, b0, b1, x0) 4
```

The result is  $x^*(b_1) = (0.291, 0.181, 0.145, 0.237, 0.146)$ .  $\blacklozenge$

We conclude with a summary of ideas for future research. Two challenges are provided by Conjectures 4.4.7 and 4.5.7. More generally, it is interesting to find formulas for the maximum likelihood degree of Wachspress models in terms of the combinatorics of the polytope.

In the context of linear programming, it is relevant to study the strictly convex objective function  $c^T x + \log V_{C, \epsilon}(x)$ , with  $V_{C, u}$  as in (4.4.2), for varying values of  $\epsilon \in \mathbb{R}_{\geq 0}$ . For  $\epsilon \rightarrow \infty$ , we recover the dual volume objective. For  $\epsilon \rightarrow 0$ , we are solving a linear program. We propose to study the degeneration of the Santaló patchwork as  $\epsilon$  moves from  $\infty$  to 0.

Next to their important role in convex optimization, we believe that *generalized Santaló points*, meaning critical points of  $\log V_{C, u}(x)$  from (4.4.2), can be used for the numerical evaluation of Euler integrals via the saddle point method [MHMT23, Section 5, problem 1].

Another next step is to go *beyond convex polytopes*. The Santaló point is well-defined for any full-dimensional convex body. One could start with spectrahedra, which is natural in the context of semidefinite programming. The Santaló patchwork of a spectrahedron replaces the Gibbs manifold for entropic regularization (Section 2.4) when the volumetric barrier is used.

Finally, we propose to study the broader context of Remark 4.5.6: which strictly convex functions give rise to interesting semi-algebraic subsets of  $\mathbb{R}_{>0}^n$ ? Furthermore, when and how are these semi-algebraic sets naturally connected to maximum likelihood estimation?

## Chapter 5

# Grasstopes

The (tree) amplituhedron, introduced by Arkani-Hamed and Trnka in [AHT14], is a geometric object playing an important role in calculations of scattering amplitudes in planar  $N = 4$  Super-Yang-Mills theory. It is defined as the image of the totally nonnegative Grassmannian  $\text{Gr}_{\geq 0}(k, n)$  under a totally positive linear map  $\tilde{Z} : \text{Gr}(k, n) \dashrightarrow \text{Gr}(k, k + m)$  given by an  $n \times (k + m)$ -matrix  $Z$ . While immediate physical relevance of the amplituhedron becomes manifest at  $m = 4$ , it is an object of independent mathematical interest for any  $m$ . It is known that when  $k = 1$ , it is a cyclic polytope [Stu88] and when  $k + m = n$ , it is isomorphic to the totally nonnegative Grassmannian  $\text{Gr}_{\geq 0}(k, n)$ .

In recent years, the amplituhedron has been studied extensively from the point of view of algebraic combinatorics for  $m = 1, 2, 4$  (see [EZLT21, GL20, KM23, KW19, LPW23, PSBW23]). The structure of the amplituhedron in the  $m = 1$  case is particularly simple: Karp and Williams [KW19] show that it is linearly homeomorphic to the complex of bounded cells of an affine hyperplane arrangement and therefore is homeomorphic to a closed ball.

One reason that the amplituhedron is so amenable to combinatorial study is that the totally nonnegative Grassmannian itself has a rich combinatorial structure. In particular,  $\text{Gr}_{\geq 0}(k, n)$  admits a stratification by positroid cells, which are all homeomorphic to open balls [Pos06]. However, amplituhedra are images of very special linear maps, just as cyclic polytopes are very special polytopes. From this point of view, it makes sense to consider images of the totally nonnegative Grassmannian under arbitrary linear maps. In [Lam16], Lam considers the images of positroid cells under arbitrary linear maps, and calls them *Grassmann polytopes*. Images of the whole totally nonnegative Grassmannian are referred to as *full Grassmann polytopes*.

While amplituhedra have attracted significant attention from the mathematical community, many results in this area rely on the total positivity of the map  $\tilde{Z}$ , and much less is known about more general Grassmann polytopes. Most of the known results are assembled in [Lam16, Part 2]. Many of them are limited to the case when  $\tilde{Z}$  does not have base points on the totally nonnegative Grassmannian, i.e. is regular (well-defined) on  $\text{Gr}_{\geq 0}(k, n)$ .

In this final chapter, we initiate the study of general full Grassmann polytopes (*Grasstopes*) and focus on the case of  $m = 1$ , when the ambient Grassmannian is the dual projective space  $(\mathbb{RP}^k)^\vee$ . We extend the results of [KW19], showing that when  $\tilde{Z}$  is regular on  $\text{Gr}_{\geq 0}(k, n)$ , the resulting Grasstopes is a union of cells of a projective hyperplane arrangement satisfying a certain sign variation condition. This, in particular, implies that Grasstopes arising from such  $\tilde{Z}$  are closed and connected, in accordance with [Lam16, Proposition 15.2]. When there are no additional restrictions on  $\tilde{Z}$ , we show that the image of the totally positive Grassmannian  $\text{Gr}_{> 0}(k, n)$  can still be characterized in terms of sign changes, although the image of  $\text{Gr}_{\geq 0}(k, n)$

might have irregular boundary. We also show that, unlike amplituhedra, general  $m = 1$  Grasstopes are not necessarily homeomorphic to closed balls or even contractible.

This chapter is organized as follows. In Section 5.1, we define Grasstopes. We divide Grasstopes into three categories (tame, wild, and rational) based on the properties of the map  $\tilde{Z}$ . We introduce the concepts necessary for the sign variation characterization of  $m = 1$  Grasstopes and prove some auxiliary results about general Grasstopes. In Section 5.2, we study the combinatorics and geometry of Grasstopes for  $m = 1$  and prove the sign variation characterization results for tame and wild Grasstopes, as well as for open rational Grasstopes. Section 5.3 is devoted to examples. Finally, in Section 5.4 we investigate how many regions of a hyperplane arrangement can be in an  $m = 1$  Grasstopes, and, based on the sign variation characterization, suggest a definition of the Grasstopes of an arbitrary oriented matroid.

## 5.1. What is a Grasstopes?

Our mathematical story in this section starts with the notion of the totally nonnegative Grassmannian  $\text{Gr}_{\geq 0}(k, n)$ . It is instructive to revisit Examples 1.1.13 and 1.1.18 for a definition. Let  $Z$  be a real  $n \times (k + m)$ -matrix of full rank, where  $k + m \leq n$ . The matrix  $Z$  defines a rational map  $\tilde{Z} : \text{Gr}(k, n) \dashrightarrow \text{Gr}(k, k + m)$  by  $[A] \mapsto [AZ]$ , where  $[A]$  is the class in  $\text{Gr}(k, n)$  of a matrix  $A$ .

**Definition 5.1.1** (Grasstopes). The image  $\tilde{Z}(\text{Gr}_{\geq 0}(k, n)) \subseteq \text{Gr}(k, k + m)$  is called the  $(n, k, m)$ -Grasstopes of  $Z$  and is denoted by  $\mathcal{G}_{n,k,m}(Z)$ .

The totally nonnegative Grassmannian  $\text{Gr}_{\geq 0}(k, n)$  is a semi-algebraic set (that is, it can be described by polynomial equations and inequalities) and  $\mathcal{G}_{n,k,m}(Z)$  is its image under a polynomial map. Thus, it follows from the Tarski-Seidenberg theorem (Theorem 1.1.17) that  $\mathcal{G}_{n,k,m}(Z)$  is also semi-algebraic.

In [Lam16]  $\mathcal{G}_{n,k,m}(Z)$  are called *full Grassmann polytopes*. When  $Z$  is a totally positive matrix, one recovers the definition of the *(tree) amplituhedron*  $\mathcal{A}_{n,k,m}(Z)$  [AHT14], a geometric object of fundamental importance to calculating scattering amplitudes in particle physics.

Note that the matrix  $AZ$  defines an element in  $\text{Gr}(k, k + m)$  if and only if  $AZ$  has full rank. It is a priori not guaranteed that the map  $\tilde{Z}$  is *well-defined* on  $\text{Gr}_{\geq 0}(k, n)$ , that is, that  $AZ$  has full rank for any totally nonnegative  $k \times n$  matrix  $A$ . In general,  $\tilde{Z}$  has base locus

$$\mathcal{B}(\tilde{Z}) := \{V : \dim(V \cap \ker Z^T) \geq 1\} \subset \text{Gr}(k, n),$$

which may or may not intersect  $\text{Gr}_{\geq 0}(k, n)$ . Note that  $\mathcal{B}(\tilde{Z})$  is a Schubert variety, so in particular, it is closed in  $\text{Gr}(k, n)$  [Lam16, Section 17]. We will often view  $\text{Gr}_{\geq 0}(k, n)$  as a parameter space of  $(k - 1)$ -dimensional subspaces of  $\mathbb{RP}^{n-1}$ , in which case the base locus is all projective subspaces which are not disjoint from  $\mathbb{P}(\ker Z^T)$ .

Finding combinatorial conditions for  $\tilde{Z}$  to be well-defined on  $\text{Gr}_{\geq 0}(k, n)$  had been an active area of research for several years. In [Lam16, Proposition 15.2] Lam proved that if  $\tilde{Z}$  is well-defined on  $\text{Gr}_{\geq 0}(k, n)$ , then  $\mathcal{G}_{n,k,m}(Z)$  is closed and connected. In the same proposition he showed that the following condition is sufficient for  $\tilde{Z}$  to be well-defined.

There exists a  $(k + m) \times k$ -matrix  $M$  such that all  $k \times k$ -minors of  $ZM$  are positive. (5.1.1)

Geometrically, condition (5.1.1) means that the element of  $\text{Gr}(k + m, n)$  represented by  $Z$  contains a totally positive  $k$ -dimensional subspace, that is, an element of  $\text{Gr}_{> 0}(k, n)$ . Lam also conjectured that this condition is necessary for  $\tilde{Z}$  to be well-defined on  $\text{Gr}_{\geq 0}(k, n)$ . A combinatorial criterion was given in [Kar17, Theorem 4.2]. Based on this criterion, Galashin

gave a counterexample to Lam's conjecture (see [KW19, Remark 9.3] and Example 5.3.2). This conjecture gives rise to the following definition.

**Definition 5.1.2** (Tame Grasstope). The Grasstope  $\mathcal{G}_{n,k,m}(Z)$  is called *tame* if  $Z$  satisfies (5.1.1).

Let  $\phi : V \rightarrow W$  be a map of vector spaces. We define its  $k^{\text{th}}$  exterior power  $\wedge^k \phi : \wedge^k V \rightarrow \wedge^k W$  by  $v_1 \wedge \dots \wedge v_k \mapsto \phi(v_1) \wedge \dots \wedge \phi(v_k)$ . If  $M$  is a matrix representing  $\phi$  in bases  $\{e_i\}$  of  $V$  and  $\{\tilde{e}_j\}$  of  $W$ , then we denote the matrix representing  $\wedge^k \phi$  in the induced bases of  $\wedge_k V$  and  $\wedge_k W$  by  $\wedge_k M$ . Then the matrix of  $\tilde{Z}$  is  $\wedge^k Z$ , where we use the standard embedding  $\text{Gr}(k, n) \rightarrow \wedge^k(\mathbb{R}^n)$ ,  $\text{Span}\{v_1, \dots, v_k\} \mapsto v_1 \wedge \dots \wedge v_k$ . Concretely, the entries of  $\wedge^k M$  are the  $k \times k$ -minors of  $M$ . Those minors are ordered by multi-indices in reverse lexicographic order (according to Macaulay2 convention [GS]).

We now give a simple geometric criterion of tameness. In what follows, for the sake of simplicity, we slightly abuse notation and write  $\mathcal{G}_{n,k,m}(Z)$  both for the Grasstope as a subset of  $\text{Gr}(k, k+m)$  and its image under the Plücker embedding.

**Proposition 5.1.3.** An  $n \times (k+m)$ -matrix  $Z$  satisfies (5.1.1) (i.e.  $\mathcal{G}_{n,k,m}(Z)$  is tame) if and only if there exists a hyperplane in  $\mathbb{RP}^{\binom{k+m}{k}-1}$ , corresponding to a point in  $\text{Gr}(k, k+m)$  which does not intersect  $\mathcal{G}_{n,k,m}(Z)$ .

*Proof.* Any  $(k+m) \times k$ -matrix  $M$  defines a hyperplane in  $\mathbb{RP}^{\binom{k+m}{k}-1}$  by its Plücker coordinates: a point  $p \in \mathbb{RP}^{\binom{k+m}{k}-1}$  lies on the hyperplane defined by  $M$  if and only if  $(\wedge_k M)^T p = 0$ . Suppose that  $Z$  satisfies (5.1.1) and  $M$  is a  $(k+m) \times k$ -matrix such that  $ZM$  is totally positive. Suppose that a point in  $\mathcal{G}_{n,k,m}(Z)$  lies on the hyperplane defined by  $M$ . Then, for some  $A \in \text{Gr}_{\geq 0}(k, n)$ , it holds that  $(\wedge_k M)^T (\wedge_k AZ)^T = 0$ , which implies that  $(\wedge_k(ZM))^T (\wedge_k A)^T = 0$ . Since all  $k \times k$ -minors of  $A$  are non-negative, and at least one is nonzero, it is not possible for all  $k \times k$ -minors of  $ZM$  to have the same sign. Thus, if there exists  $M$  such that  $ZM$  has all positive (or negative)  $k \times k$ -minors, then the image  $\mathcal{G}_{n,k,m}(Z)$  does not intersect the hyperplane defined by  $M$ .

Now suppose that  $Z$  does not satisfy (5.1.1), that is, for any  $M$  the matrix  $ZM$  has either a zero  $k \times k$ -minor or at least one positive and one negative  $k \times k$ -minor. We will show that there exists a matrix  $A \in \text{Gr}_{\geq 0}(k, n)$  such that  $(\wedge_k(ZM))^T (\wedge_k A)^T = 0$ , so that the hyperplane defined by  $M$  intersects  $\tilde{Z}(\text{Gr}_{\geq 0}(k, n))$ .

In the first case, when  $ZM$  has a zero minor in the  $i^{\text{th}}$  Plücker coordinate, one can find an element  $A \in \text{Gr}_{\geq 0}(k, n)$  which has all Plücker coordinates equal to zero except for the  $i^{\text{th}}$  one. In this case,  $(\wedge_k(ZM))^T (\wedge_k A)^T = 0$ .

Now consider the second case, in which  $ZM$  has at least one positive and one negative  $k \times k$ -minor. By the pigeonhole principle, there exists a set of column indices  $I = \{i_1, \dots, i_{k-1}\}$  such that two Plücker coordinates involving  $I$  (which we label  $p_{I \cup \{i\}}$  and  $p_{I \cup \{j\}}$ ) have different signs. Then, take  $(q_1, \dots, q_k : \dots : q_{n-k+1}, \dots, q_n) \in \mathbb{RP}^{\binom{n}{k}-1}$  such that all coordinates except for  $q_{I \cup \{i\}}, q_{I \cup \{j\}}$  are zero, and  $q_{I \cup \{i\}} = |p_{I \cup \{j\}}|$  and  $q_{I \cup \{j\}} = |p_{I \cup \{i\}}|$ . Then, all Plücker relations are satisfied so this point represents an element  $A \in \text{Gr}_{\geq 0}(k, n)$ . We have  $(\wedge_k(ZM))^T (\wedge_k A)^T = 0$ , so the hyperplane given by  $M$  intersects the Grasstope  $\mathcal{G}_{n,k,m}(Z)$ .  $\square$

Note that when  $m = 1$ , every hyperplane in  $\mathbb{RP}^{\binom{k+m}{k}-1}$  corresponds to some point in  $\text{Gr}(k, k+m)$ . Then by choosing a hyperplane disjoint from  $\mathcal{G}_{n,k,1}(Z)$  to be the hyperplane at infinity, we arrive at the following result.

**Corollary 5.1.4.** The Grasstope  $\mathcal{G}_{n,k,1}(Z)$  is tame if and only if its image under the Plücker embedding is contained in some affine chart of  $\mathbb{RP}^{\binom{k+m}{k}-1}$ .

Tame Grasstopes share many nice properties with amplituhedra. In particular, for  $m = 1$  they are homeomorphic to closed balls and can be described as complexes of bounded cells of affine hyperplane arrangements [KW19, Section 9]. The focus of this chapter, however, is to take a step away from the tame case and study Grasstopes that behave somewhat less regularly. We note that the terms “tame” and “wild” for Grasstopes are due to Thomas Lam.

**Definition 5.1.5** (Wild Grasstope). If the map  $\tilde{Z}$  is well-defined on  $\text{Gr}(k, n)_{\geq 0}$  but  $Z$  does not satisfy (5.1.1), then the Grasstope  $\mathcal{G}_{n,k,m}(Z)$  is called *wild*.

Even though  $\tilde{Z}$  might not be well-defined on  $\text{Gr}_{\geq 0}(k, n)$ , it still makes sense to consider the image of  $\text{Gr}_{\geq 0}(k, n) \setminus (\mathcal{B}(\tilde{Z}) \cap \text{Gr}_{\geq 0}(k, n))$ , where  $\mathcal{B}(\tilde{Z})$  is the base locus of  $\tilde{Z}$ . In this case we will still write  $\tilde{Z}(\text{Gr}_{\geq 0}(k, n))$  for this image.

**Definition 5.1.6** (Rational Grasstope). Suppose the map  $\tilde{Z}$  is not well-defined on  $\text{Gr}(k, n)_{\geq 0}$ . Then the image  $\mathcal{G}_{n,k,m}(Z) = \tilde{Z}(\text{Gr}(k, n)_{\geq 0})$  is called a *rational* Grasstope. The image  $\mathcal{G}_{n,k,m}^{\circ}(Z) = \tilde{Z}(\text{Gr}(k, n)_{> 0})$  of the totally positive Grassmannian is called an *open rational* Grasstope. This is indeed an open set, as shown in Proposition 5.2.4.

We conclude this section with technical results that will prove useful in characterization of  $m = 1$  Grasstopes in Section 5.2.

Given a point  $u$  in  $\mathbb{RP}^n$ , we associate a sign pattern  $\sigma = (\sigma_0, \dots, \sigma_n) \in \{+, -, 0\}^{n+1}$  to  $u$  in the following way. Pick  $i$  such that  $u_i \neq 0$  and set  $\sigma_i := +$ . Then  $\sigma_j := \text{sign}(u_i u_j)$ , which is a well-defined function of homogeneous coordinates of  $u$ . Since we associate sign labels to points in projective space, we will identify sign labels  $\sigma$  and  $-\sigma$ . Each orthant of the  $\mathbb{RP}^n$  consists of the points with the same sign pattern. For instance, the sign pattern  $(+ : + : - : +)$  for a point in  $\mathbb{RP}^3$  represents the orthant defined by

$$\{u_0 u_1 > 0, u_0 u_2 < 0, u_0 u_3 > 0, u_1 u_2 < 0, u_1 u_3 > 0, u_2 u_3 < 0\}.$$

Given a point  $x \in \mathbb{RP}^k$  being the image of a hyperplane  $X = \text{Span}\{w_1, \dots, w_k\}$  under the Plücker embedding of  $\text{Gr}(k, k+1)$ , and any point  $v \in \mathbb{RP}^k$ , one may consider the map

$$T : \mathbb{RP}^k \times \mathbb{RP}^k \rightarrow \mathbb{R}$$

$$x, v \mapsto \det \begin{bmatrix} | & & | & | \\ w_1 & \dots & w_k & v \\ | & & | & | \end{bmatrix} = \sum_{j=1}^k (-1)^j p_{1 \dots \hat{j} \dots k} v_j.$$

Then  $T(x, v) = 0$  if and only if  $v$  is contained in  $X$ . If  $v = Z_i$  for some row  $Z_i$  of  $Z$ , then  $T(x, Z_i)$  is called a *twistor coordinate of  $X$  with respect to  $Z$*  [Wil21, Definition 4.5].

**Remark 5.1.7.** One may also fix either argument  $x$  or  $v$  of  $T$  to get a linear form on  $\mathbb{RP}^k$ . In this chapter we will consider the twistor coordinates as functions of  $x$  by setting  $v = Z_i$  and denote the resulting forms by  $l_i(x)$ . We will see in Theorem 5.2.1 that the vanishing loci of the  $l_i$ 's are exactly the hyperplanes which contain the boundaries of  $\mathcal{G}_{n,k,m}(Z)$ .

We now recall the definition of sign variation from [KW19].

**Definition 5.1.8** (Sign variation). Given a sequence  $v$  of  $n$  real numbers, let  $\text{var}(v)$  be the number of sign changes in  $v$  (zeros are ignored). Let  $\overline{\text{var}}(v) := \max\{\text{var}(w) : w \in \mathbb{R}^n \text{ such that } w_i = v_i \text{ for all } i \in [n] \text{ with } v_i \neq 0\}$ , i.e.  $\overline{\text{var}}(v)$  is the maximum number of sign changes in  $v$  after a sign for each zero component is chosen. Note that both  $\text{var}$  and  $\overline{\text{var}}$  are well-defined functions of homogeneous coordinates of a point in  $\mathbb{RP}^{n-1}$ .

We call a  $(k-1)$ -dimensional subspace of  $\mathbb{RP}^{n-1}$  *positive* if it is a point in  $\text{Gr}_{\geq 0}(k, n)$  and *totally positive* if it is a point in  $\text{Gr}_{> 0}(k, n)$ . Given a point  $u \in \mathbb{RP}^{n-1}$ , we define the *hyperplane  $H_u$  given by  $u$*  to be the hyperplane orthogonal to  $u$  with respect to the standard dot product, i.e.  $H_u := \{v \in \mathbb{RP}^{n-1} : u \cdot v = 0\}$ . Then we have the following proposition.

**Proposition 5.1.9.** Let  $H_u$  be the hyperplane given by  $u \in \mathbb{RP}^{n-1}$ .

- (i)  $H_u$  contains a positive subspace of dimension  $k - 1$  if and only if  $\overline{\text{var}}(u) \geq k$ .
- (ii)  $H_u$  contains a totally positive subspace of dimension  $k - 1$  if and only if  $\text{var}(u) \geq k$ .

*Proof.* We start by proving the first statement. A hyperplane  $H_u$  contains a positive subspace of dimension  $k$  if and only if  $u$  is in the kernel of some matrix of the form  $\begin{pmatrix} A \\ B \end{pmatrix}$  where  $A \in \text{Gr}_{\geq 0}(k, n)$ , and  $B \in \text{Gr}(n - k - 1, n)$ . Here  $A$  represents the positive subspace, and  $B$  the additional points to define a hyperplane. Note that  $u \in \ker \begin{pmatrix} A \\ B \end{pmatrix}$  if and only if  $u \in \ker(A)$  and  $u \in \ker(B)$ . Then, the “only if” direction follows directly from [GK50, Theorems V1 and V6] (also [KW19, Theorem 3.4(i)]). The “if” direction also follows, with the note that one can always find  $n - k - 1$  additional points to form the matrix  $B$  such that  $u \in \ker(B)$ . The second statement is proved analogously, with replacing  $\text{Gr}_{\geq 0}(k, n)$  by  $\text{Gr}_{> 0}(k, n)$  and using [KW19, Theorem 3.4(ii)].  $\square$

## 5.2. Grasstopes for $m = 1$ : tame, wild, and rational

We begin by stating our main theorem which describes any  $m = 1$  Grasstope  $\mathcal{G}_{n,k,1}(Z)$  arising from a well-defined map  $\tilde{Z}$  as a subset of  $\mathbb{RP}^k \cong \text{Gr}(k, k + 1)$ . This theorem recovers and generalizes many of the results of Karp and Williams describing  $m = 1$  amplituhedra and tame Grasstopes [KW19, Section 6].

**Theorem 5.2.1.** Suppose  $\tilde{Z} : \text{Gr}(k, n) \dashrightarrow \text{Gr}(k, k + 1)$  is well-defined on  $\text{Gr}_{\geq 0}(k, n)$ . Then the Grasstope  $\mathcal{G}_{n,k,1}(Z)$  consists of the points  $\mathbf{x} \in \mathbb{RP}^k$  such that  $\overline{\text{var}}(L(\mathbf{x})) \geq k$ , where  $L(\mathbf{x})$  is the vector of twistor coordinates of  $\mathbf{x}$  with respect to  $Z$ .

*Proof.* Let  $L(\mathbf{x}) = (l_1(\mathbf{x}) : \dots : l_n(\mathbf{x})) \in \mathbb{RP}^{n-1}$  (see Remark 5.1.7). By Proposition 5.1.9, it suffices to show that  $H_{L(\mathbf{x})}$  contains a positive subspace if and only if  $\mathbf{x} \in \mathcal{G}_{n,k,1}(Z)$ .

For the “if” direction, suppose that  $A \in \text{Gr}_{\geq 0}(k, n)$  and  $\mathbf{x}$  is the vector of Plücker coordinates of  $[AZ]$ . Then for each row  $A_i$  of  $A$ , we have

$$L(\mathbf{x}) \cdot A_i = \sum_{j=1}^n T(\mathbf{x}, Z_j) A_{ij} = T(\mathbf{x}, \sum_{j=1}^n A_{ij} Z_j) = 0,$$

since  $\sum A_{ij} Z_j$  is a row of  $AZ$ . Thus  $H_{L(\mathbf{x})}$  contains  $A$ .

For the “only if” direction, suppose that  $H_{L(\mathbf{x})}$  contains a positive subspace  $A$ , and let  $v \in \ker Z^T$ . Then

$$L(\mathbf{x}) \cdot v = \sum_{i=1}^n T(\mathbf{x}, Z_i) v_i = T(\mathbf{x}, \sum_{i=1}^n Z_i v_i) = T(\mathbf{x}, 0) = 0.$$

So  $H_{L(\mathbf{x})}$  contains  $\mathbb{P}(\ker Z^T)$ . Since  $\mathbb{P}(\ker Z^T) \cap A = \emptyset$  by regularity of  $\tilde{Z}$  on  $\text{Gr}_{\geq 0}(k, n)$ , we have by dimension considerations that any two hyperplanes containing  $A$  and  $\mathbb{P}(\ker Z^T)$  must be equal. However, by the “if” direction, if  $\mathbf{a}$  is the vector of Plücker coordinates of  $[AZ]$ , then  $H_{L(\mathbf{a})}$  contains  $A$ . In addition,  $H_{L(\mathbf{a})}$  contains  $\mathbb{P}(\ker Z^T)$ . Therefore  $H_{L(\mathbf{x})} = H_{L(\mathbf{a})}$ . Since  $u \mapsto H_u$  is injective, we obtain  $\mathbf{x} = \mathbf{a}$ .  $\square$

Note that in the proof of Theorem 5.2.1, the well-defined condition is necessary, since otherwise there might be positive subspaces which intersect  $\mathbb{P}(\ker Z^T)$ . In this case, there is no guarantee that a hyperplane containing a positive subspace also contains a positive subspace disjoint from  $\mathbb{P}(\ker Z^T)$ , which might cause problems on the boundary of the Grasstope.

However, as we show in the following results, we can still describe the open rational Grasstope in the case that the map  $\tilde{Z}$  is not well-defined on  $\text{Gr}_{\geq 0}(k, n)$ . We first need the following lemma about totally positive subspaces to show that intersection with  $\mathbb{P}(\ker Z^T)$  does not cause any issues.

**Lemma 5.2.2.** Let  $H$  be a hyperplane in  $\mathbb{R}\mathbb{P}^{n-1}$  containing an  $(n - k - 1)$ -dimensional subspace  $P$ . If  $H$  contains a totally positive  $(k - 1)$ -dimensional subspace, then it contains a totally positive  $(k - 1)$ -dimensional subspace disjoint from  $P$ .

*Proof.* Let  $V \subset H$  be a totally positive  $(k - 1)$ -dimensional subspace. Let  $l = \dim(V \cap P)$ . Then  $\dim(V + P) = n - l - 2$  and  $V + P \subset H$ . In particular, the codimension of  $V + P$  inside  $H$  is equal to  $l$ . Pick  $l + 1$  points  $q_1, \dots, q_{l+1}$  in general position in  $H \setminus (V + P)$ . Let  $M$  be a matrix representing  $V + P \in \text{Gr}(n - l - 1, n)$  such that the first  $k + 1$  rows of  $M$  represent  $V \in \text{Gr}_{>0}(k, n)$  and there is a subset of rows  $r_1, \dots, r_{l+1}$  of  $M$  representing  $V \cap P \in \text{Gr}(l, n)$ . Denote by  $B$  the submatrix of  $M$  given by its first  $k + 1$  rows. Note that the  $k - l$  rows  $s_1, \dots, s_{k-l}$  of  $B$  that are not  $r_1, \dots, r_{l+1}$  span a subspace  $W$  disjoint from  $P$ . Consider a matrix  $M'$  obtained from  $M$  by replacing  $r_i$  with  $r_i + \varepsilon_i \tilde{q}_i$  for  $i = 1, \dots, l + 1$ , where  $\varepsilon_i > 0$  and  $\tilde{q}_i$  is a vector of homogeneous coordinates of the point  $q_i$ . Since the points  $q_1, \dots, q_{l+1}$  are in general position in  $H \setminus (V + P)$ , the matrix  $M'$  has full rank and represents an  $(n - l - 2)$ -dimensional subspace of  $H$ . Denote the first  $k + 1$  rows of  $M'$  by  $B'$ . Then for  $\varepsilon_1, \dots, \varepsilon_{l+1}$  small enough  $B'$  represents a subspace  $V' \in \text{Gr}_{>0}(k, n)$ . We claim that  $V'$  is disjoint from  $P$ . To show this, consider the matrix  $D$  obtained by stacking  $B'$  below a matrix  $C$  representing  $P$ . Note that  $D$  represents the projective subspace  $V' + P$ . Since the points in  $\mathbb{R}\mathbb{P}^{n-1}$  defined by rows  $r_1, \dots, r_{l+1}$  are in  $P$ , by performing row operations on  $D$ , one can reduce it to the form  $D' = \begin{pmatrix} C \\ Q \end{pmatrix}$ , where  $Q$  consists of the rows  $s_1, \dots, s_{k-l}, \tilde{q}_1, \dots, \tilde{q}_{l+1}$ . Due to the choice of  $q_1, \dots, q_{l+1}$  and the fact that  $W$  is disjoint from  $P$ ,  $D'$  has full rank. This means that  $\dim(V' + P) = \dim(V') + \dim(P)$ , so  $V'$  and  $P$  are disjoint.  $\square$

Following Lemma 5.2.2, we are ready to describe the open rational Grasstope  $\tilde{Z}(\text{Gr}_{>0}(k, n))$ . Recall that  $\mathcal{B}(\tilde{Z})$  denotes the set of base points of  $\tilde{Z}$ .

**Proposition 5.2.3.** For a map  $\tilde{Z} : \text{Gr}_{\geq 0}(k, n) \dashrightarrow \text{Gr}(k, k + 1)$ , given by a matrix  $Z$ , the open Grasstope  $\tilde{Z}(\text{Gr}_{>0}(k, n))$  consists of the points  $\mathbf{x} \in \mathbb{R}\mathbb{P}^k$  such that  $\text{var}(L(\mathbf{x})) \geq k$ , where  $L(\mathbf{x})$  is the vector of twistor coordinates of  $\mathbf{x}$  with respect to  $Z$ .

*Proof.* We will prove this statement by slightly modifying the proof of Theorem 5.2.1. Fix a totally positive matrix  $A \in \text{Gr}_{>0}(k, n) \setminus \mathcal{B}(\tilde{Z})$  and let  $\mathbf{x}$  be the vector of Plücker coordinates of  $[AZ]$ . Then the hyperplane  $H_{L(\mathbf{x})}$  contains  $\mathbb{P}(\ker Z^T)$  and the totally positive subspace  $A$ . Conversely, if a hyperplane  $H$  contains  $\mathbb{P}(\ker Z^T)$  and a totally positive subspace  $A$ , then, by Lemma 5.2.2,  $H$  contains a totally positive subspace  $A'$  disjoint from  $\mathbb{P}(\ker Z^T)$  (thus,  $A' \notin \mathcal{B}(\tilde{Z})$ ). The hyperplane is then uniquely determined by its containment of  $\mathbb{P}(\ker Z^T)$  and  $A'$ , so it must be  $H_{L(\mathbf{x})}$ , where  $\mathbf{x}$  is the vector of Plücker coordinates of  $[A'Z]$ . We now use the second part of Proposition 5.1.9 to conclude that the open Grasstope  $\tilde{Z}(\text{Gr}_{>0}(k, n))$  consists of the points in  $\mathbf{x} \in \mathbb{R}\mathbb{P}^k$  such that  $\text{var}(L(\mathbf{x})) \geq k$ .  $\square$

**Proposition 5.2.4.** The open rational Grasstope of  $Z$  is open and, if  $Z$  has no zero rows, the rational Grasstope of  $Z$  is contained in the closure of the open rational Grasstope of  $Z$ .



*Proof.* Let  $Z$  define a rational Grasstope and consider a point  $\mathbf{x}$  with  $\text{var}(L(\mathbf{x})) \geq k$ , where  $L(\mathbf{x})$  is the vector of twistor coordinates of  $\mathbf{x}$  with respect to  $Z$ . Then for all points  $\mathbf{x}'$  in a sufficiently small neighborhood around it,  $L(\mathbf{x}')$  has the same signs as  $L(\mathbf{x})$  in all the indices of the nonzero entries of  $L(\mathbf{x})$ . Since changing the zero entries cannot decrease the sign variation, the neighborhood is contained in the open rational Grasstope.

Similarly to the reasoning of Theorem 5.2.1, it follows from the second part of Proposition 5.1.9 that the rational Grasstope of  $Z$  must be contained in the set  $C := \{\mathbf{x} : \overline{\text{var}}(L(\mathbf{x})) \geq k\}$ . We show that when  $Z$  has no zero rows,  $C$  is the closure of the open rational Grasstope of  $Z$ . First, we show that the complement of  $C$ , the set  $\{\mathbf{x} : \overline{\text{var}}(L(\mathbf{x})) < k\}$ , is open. Let  $\mathbf{x}$  be such that  $\overline{\text{var}}(L(\mathbf{x})) < k$ . Then for all points  $\mathbf{x}'$  in a sufficiently small neighborhood around  $\mathbf{x}$ ,  $L(\mathbf{x}')$  has the same signs as  $L(\mathbf{x})$  in all the indices of the nonzero entries of  $L(\mathbf{x})$ . Since changing the values of the zero entries cannot increase  $\overline{\text{var}}$ , we know  $\overline{\text{var}}(L(\mathbf{x}')) < k$ . Therefore  $C$  is closed.

Now consider a point  $\mathbf{x} \in C$  and an open neighborhood  $N$  of points around it. We will show that there is some  $\mathbf{x}' \in N$  with  $\text{var}(L(\mathbf{x}')) \geq k$ , which is sufficient to conclude that  $C$  is the closure of the open rational Grasstope of  $Z$ . Since  $Z$  has no zero rows, each zero entry of  $L(\mathbf{x})$  corresponds to containment of  $\mathbf{x}$  in a hyperplane. Any open neighborhood around  $\mathbf{x}$  contains points on either side of the hyperplane. Similarly, if  $\mathbf{x}$  lies in the intersection of several hyperplanes, any open neighborhood of  $\mathbf{x}$  contains points in each orthant defined by these hyperplanes, so any sign pattern can be achieved in the entries which are zero in  $L(\mathbf{x})$ . In particular, this means that there is a point  $\mathbf{x}'$  with signs in the nonzero entries of  $L(\mathbf{x})$  equal to the signs of the corresponding entries of  $L(\mathbf{x}')$  (since we can restrict  $N$  to be small enough such that the signs in the nonzero entries of  $L(\mathbf{x})$  are unchanged), and signs of the zero entries replaced by the signs which ensure that  $\text{var}(L(\mathbf{x}')) \geq k$ .  $\square$

In Section 5.3, we show that the rational Grasstope may be equal to the closure of the open rational Grasstope. However, we do not know if this holds in general, since it might be possible for points in the boundary to fall out. Thus the question of fully describing which parts of the boundary are contained in  $m = 1$  rational Grasstopes remains open.

## 5.3. Examples

In this section we provide examples of the families of Grasstopes we considered. We begin with an example of a tame Grasstope.

**Example 5.3.1** (A tame Grasstope). Let

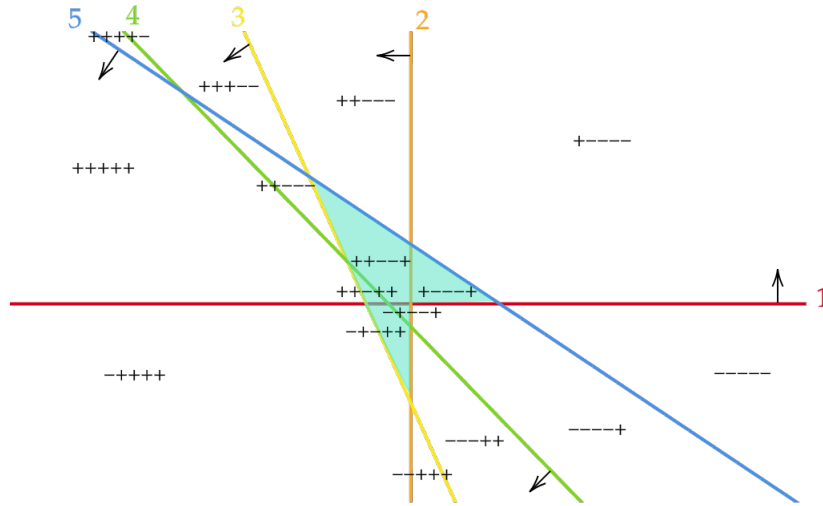
$$Z = \begin{bmatrix} 1 & 0 & -1 & -3 & -2 \\ 0 & 1 & 1 & 2 & 1 \\ 0 & 0 & 1 & -1 & -2 \end{bmatrix}^T.$$

This matrix is not totally positive, since  $p_{123} = 1$  and  $p_{124} = -1$ . However, the first two rows of  $Z^T$  span a totally positive line, so  $Z$  satisfies (5.1.1) with the matrix  $M$  being

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}^T.$$

Therefore the resulting Grasstope is tame but is not an amplituhedron. The rows of  $Z$  define 5 linear forms, as noted in Remark 5.1.7:

$$l_1 = z, \quad l_2 = -y, \quad l_3 = x - y - z, \quad l_4 = -x - 2y - 3z, \quad l_5 = -2x - y - 2z.$$



**Figure 5.1:** Affine chart in which the tame Grasstope is bounded. The six lines corresponding to the rows of  $Z$  are colored red, orange, yellow, green, and blue, in order, with orientations given by arrows. The shaded portion of the figure is the Grasstope, which consists exactly of the regions with at least two sign changes.

Not every affine chart that we choose results in a bounded picture. For instance, if we map  $(x : y : z) \mapsto (x + y : y : z)$  and dehomogenize with respect to the first coordinate, the resulting picture is unbounded. However, as predicted by Corollary 5.1.4, there are lines disjoint from the Grasstope. One of them is  $\{-4x + z = 0\}$  in  $\mathbb{RP}^2$ . By picking it to be the line at infinity (that is, by mapping  $(x : y : z) \mapsto (-4x + z : y : z)$  and dehomogenizing with respect to the first coordinate), we obtain affine lines given by the linear forms

$$\tilde{l}_1 = \tilde{y}, \quad \tilde{l}_2 = -\tilde{x}, \quad \tilde{l}_3 = \frac{-1 - 4\tilde{x} - 3\tilde{y}}{4}, \quad \tilde{l}_4 = \frac{-1 - 8\tilde{x} - 13\tilde{y}}{4}, \quad \tilde{l}_5 = \frac{1 - 2\tilde{x} - 5\tilde{y}}{2}.$$

Each line has an orientation, with the positive half-space given by the points  $(\tilde{x}, \tilde{y})$  for which  $\tilde{l}_i(\tilde{x}, \tilde{y}) \geq 0$ . The lines divide the affine plane into regions, each of which has a corresponding sign vector with  $i^{\text{th}}$  coordinate being + if  $\tilde{l}_i(\tilde{x}, \tilde{y}) > 0$  for all  $(\tilde{x}, \tilde{y})$  in the region, and - if  $\tilde{l}_i(\tilde{x}, \tilde{y}) < 0$ . The Grasstope of  $Z$  consists exactly of those points in the regions for which  $\overline{\text{var}}(u) \geq 2$ , as can be seen in Figure 5.1. ◆

**Example 5.3.2** (A wild Grasstope). Let

$$Z = \begin{bmatrix} 2 & 2 & 0 & -1 & 1 & 0 \\ 2 & 3 & 1 & 0 & 2 & 0 \\ 2 & 2 & 0 & 0 & 2 & 1 \end{bmatrix}^T.$$

This is the example found by Galashin (and communicated to us by Lam [Lam23]) to show that wild Grasstopes exist. Indeed, suppose there exists a  $3 \times 2$  matrix  $M$  such that  $ZM$  has positive  $2 \times 2$ -minors. The  $2 \times 2$ -minors of  $ZM$  are the entries of  $\wedge_2(ZM) = \wedge_2(Z) \times \wedge_2(M) =$

$$= \begin{bmatrix} 2 & 2 & 2 & 2 & 3 & 1 & 2 & 1 & -1 & -2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 2 & 0 & 2 & 2 & 0 & -2 & 2 & 2 & 0 & -1 & 1 \\ -2 & -2 & -2 & 0 & 0 & 0 & 0 & 2 & 2 & 0 & 2 & 3 & 1 & 0 & 2 \end{bmatrix}^T \begin{bmatrix} p_{12} \\ p_{13} \\ p_{23} \end{bmatrix},$$

where  $p_{12}, p_{13}, p_{23}$  are the minors of  $M$ . Then, column 4 of  $(\wedge_2 Z)^T$  tells us that  $p_{12} + p_{13} > 0$  but column 10 tells us that  $p_{12} + p_{13} < 0$ , so no such matrix  $M$  can exist.

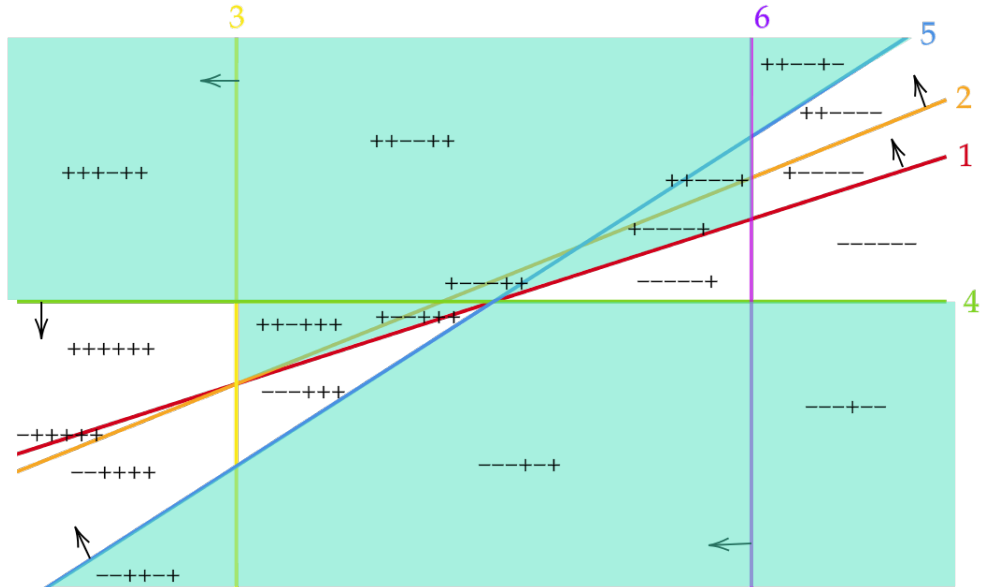
The six rows of  $Z$  correspond to the six linear forms

$$\begin{aligned} l_1 &= 2x - 2y + 2z, & l_2 &= 2x - 3y + 2z, & l_3 &= -y \\ l_4 &= -z, & l_5 &= 2x - 2y + z, & l_6 &= x. \end{aligned}$$

Mapping  $(x : y : z) \mapsto (x + y : y : z)$  and dehomogenizing with respect to the first coordinate, we obtain affine lines given by the linear forms

$$\begin{aligned} \tilde{l}_1 &= 2\tilde{y} - 4\tilde{x} + 2, & \tilde{l}_2 &= 2\tilde{y} - 5\tilde{x} + 2, & \tilde{l}_3 &= -\tilde{x}, \\ \tilde{l}_4 &= -\tilde{y}, & \tilde{l}_5 &= \tilde{y} - 4\tilde{x} + 2, & \tilde{l}_6 &= -\tilde{x} + 1. \end{aligned}$$

We draw these in the affine plane and color them (in order) red, orange, yellow, green, blue, and purple. We also give the lines orientations with the positive half-space given by the points  $(\tilde{x}, \tilde{y})$  for which  $\tilde{l}(\tilde{x}, \tilde{y}) > 0$ . Then the Grasstopes of  $Z$  consists exactly of those points in the regions between the lines for which  $\overline{\text{var}}(u) \geq 2$ , as can be seen in Figure 5.2.  $\blacklozenge$



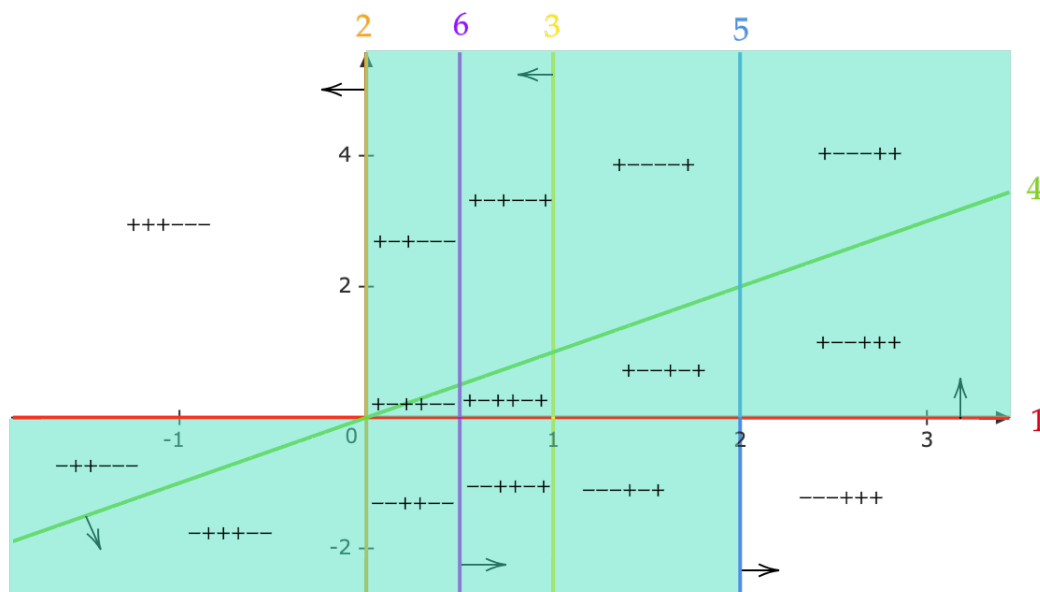
**Figure 5.2:** The six lines corresponding to the rows of  $Z$  are pictured as described, with orientations given by the arrows. The regions can then be labelled by sign patterns. The shaded portion of the figure is the Grasstopes, and it consists exactly of those regions with at least two sign changes.

**Example 5.3.3** (A rational Grasstopes with closed boundary and Möbius strip topology). Let

$$Z = \begin{bmatrix} 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 1 & -1 \\ 0 & 0 & 3 & 0 & -2 & -1 \end{bmatrix}^T.$$

Note that  $(1, 1, 1, 1, 1, 1) \in \ker(Z^T)$ , so the map  $\tilde{Z}$  has base points on  $\text{Gr}_{\geq 0}(2, 6)$ . Following Proposition 5.2.3 we can still describe its open Grasstopes. As in the previous examples, we find 6 dehomogenized linear forms corresponding to six affine lines

$$\begin{aligned} \tilde{l}_1 &= \tilde{y}, & \tilde{l}_2 &= -\tilde{x}, & \tilde{l}_3 &= -3\tilde{x} + 3, \\ \tilde{l}_4 &= \tilde{x} - \tilde{y}, & \tilde{l}_5 &= \tilde{x} - 2, & \tilde{l}_6 &= 2\tilde{x} - 1. \end{aligned}$$



**Figure 5.3:** The six lines corresponding to the rows of  $Z$  are pictured as described, with orientations given by the arrows. The regions can then be labelled by sign patterns. The shaded portion of the figure is the Grasstopes, and it consists exactly of those regions with at least two sign changes. In this case, the shaded region is a Möbius strip.

Then we can find the open rational Grasstopes of  $Z$  as in Figure 5.3.

We claim that the rational Grasstopes of  $Z$  is the closure of the open rational Grasstopes. One can check this by directly finding Plücker coordinates for some  $M \in \text{Gr}_{\geq 0}(2, 6)$  which map to the points that lie on the boundary. For example, to find such  $M$  for any point of the form  $(0, a)$  with  $a \geq 0$ , we solve  $\wedge_2(Z) \times \wedge_2(M) = (1 : 0 : a)$ . Then  $\wedge_2(M)$  must have nonnegative entries which satisfy the Plücker relations. This process is made easier if one recalls that the Plücker relations are trivially satisfied if all entries are zero except for ones which correspond to pairwise overlapping submatrices, that is any two nonzero minors come from submatrices which share a column. In this case, since

$$\wedge_2(Z) = \begin{bmatrix} 1 & 0 & 0 & -1 & 1 & 0 & 1 & 0 & 0 & -1 & -1 & 0 & 0 & 1 & 0 \\ 0 & 3 & 0 & 0 & 0 & 3 & -2 & 0 & 0 & 2 & -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 3 & 0 & 0 & 3 & 0 & -2 & -3 & 2 & 0 & -1 & 3 & 1 & -3 \end{bmatrix}^T,$$

we find that the point  $(1 : 0 : a)$  is given by  $\wedge_2(M) = (0 : 0 : a/3 : 0 : 1 : 0 : 0 : 0 : 0 : 0 : 0 : 0 : 0 : 0 : 0)$ . Since the only nonzero entries correspond to the minors  $p_{23}$  and  $p_{24}$ , the Plücker relations are satisfied. Thus the portion of the boundary line  $x = 0$  with  $y \geq 0$  is part of the rational Grasstopes of  $Z$ . One can similarly check that all other parts of all six lines are included. Therefore the rational Grasstopes of  $Z$  is closed. Furthermore, topologically, as one can see from Figure 5.3, it is a Möbius strip. ♦

### 5.4. Extremal counts and oriented matroid Grasstopes

This section relies heavily on the machinery of oriented matroids. An introduction is given in Section 1.5, and [BLVS<sup>+</sup>99] is a comprehensive reference.

Given an  $n \times (k + 1)$ -matrix  $Z$ , one may ask questions about the topology of the resulting  $m = 1$  Grasstopes. For instance, is it closed, connected, contractible? How many regions

of the hyperplane arrangement does it contain? For the  $m = 1$  amplituhedron, the answer to the first three questions is “yes” [KW19, Corollary 6.18]. As for the latter, the  $m = 1$  amplituhedron contains as few regions as possible, that is, all possible sign vectors with  $\text{var} < k$  appear as labels of regions in the corresponding arrangement [KW19, Proposition 6.14]. In this section, we investigate this last question for more general  $m = 1$  Grasstopes.

We start by recalling Equation (1.5.1) for the number of regions in a projective hyperplane arrangement  $\mathbf{P}$  given the total number of regions  $t(\mathbf{A})$  and the number of bounded regions  $b(\mathbf{A})$  in the corresponding affine arrangement  $\mathbf{A}$ :

$$r(\mathbf{P}) = b(\mathbf{A}) + \frac{t(\mathbf{A}) - b(\mathbf{A})}{2}. \quad (5.4.1)$$

We write  $\beta(k, n)$  for the number of possible sign patterns of length  $n$  with sign variation less than  $k$  and  $\gamma(k, n)$  for the number of sign patterns with variation greater or equal than  $k$  (we identify sign patterns  $\sigma$  and  $-\sigma$ ). Note that  $\beta(k, n) = 1 + \binom{n-1}{2} + \dots + \binom{n-1}{k}$ , and  $\gamma(k, n) = 2^{n-1} - \beta(k, n)$ . Theorem 5.2.1 and Proposition 5.2.4 then give the lower bound  $r(\mathbf{P}) - \beta(k, n)$  for the number of regions in  $\mathcal{G}_{n,k,1}(Z)$ , where  $\mathbf{P}$  is the hyperplane arrangement defined by  $Z$ . An upper bound is given by the minimum of  $\gamma(k, n)$  and  $r(\mathbf{P})$ .

The database [Fin] contains a catalog of isomorphism classes of oriented matroids [Fin01, Section 6]. Each matroid is indexed by a vector of signs of its bases and each hyperplane arrangement corresponds to a realizable matroid, as explained in the previous section. An arrangement is simple if the vector of signs of bases of its matroid does not contain zeros, that is, the matroid is uniform.

We iterate over all uniform oriented matroid isomorphism classes in this catalog for small values of  $k$  and  $n$ . Note that, for all the values of  $k$  and  $n$  which we consider, all uniform matroids are realizable [FMM13], that is, arise from hyperplane arrangements. Within each isomorphism class, we iterate over all possible reorderings of the ground set and reorientations. If the matroid is realizable, at the level of the matrix  $Z$  defining the arrangement as described in Section 5.1, reorderings correspond to permuting the rows and reorientations to negating certain rows. For each isomorphism class, ordering of the hyperplanes, and a choice of orientation, we compute the number of regions in the corresponding Grasstope (that is, the number of maximal covectors with sign variation greater or equal than  $k$ ; see Remark 1.5.13).

It turns out that for many values of  $k$  and  $n$  the minimal and maximal number of regions in the Grasstope when iterating over reorderings and reorientations does not depend on the oriented matroid isomorphism class. The minimal and maximal number of regions in the Grasstope for these values of  $k$  and  $n$  are presented in Table 5.1. The Python code used to extract this data is available at <https://mathrepo.mis.mpg.de/Grasstopes>. We therefore have a computational proof of the following statement.

**Proposition 5.4.1.** For each pair of values of  $k$  and  $n$  in Table 5.1 the minimal and maximal possible number of regions in a Grasstope arising from a simple arrangement of  $n$  hyperplanes in  $\mathbb{P}^k$  do not depend on the choice of arrangement.

Out of the entries in Table 5.1, note that for  $k = 2$  and the pairs  $(3, 5)$  and  $(4, 6)$ , there is only one oriented matroid up to isomorphism [Fin]. For  $k = 2$ , this is because any matrix may be turned into a totally positive matrix by permuting the rows. This can be done by viewing rows as vectors in the plane and arranging them in counterclockwise position.

The pair  $(3, 7)$  does not appear in Table 5.1. This is the first time we see variation depending on which simple arrangement we choose, with the maximal number of regions ranging from 38 to 42. The reorientations and reorderings of a totally positive matrix (i.e. the amplituhedron case) give at most 42 regions, while all other oriented matroid classes

$k, n$	Minimal	Maximal	$r(\mathbf{P})$	$\beta(k, n)$	$\gamma(k, n)$
2, 6	10	16	16	6	26
2, 7	15	22	22	7	57
3, 5	4	5	15	11	5
3, 6	10	16	26	16	16
4, 6	5	6	31	26	6
4, 7	15	22	57	42	22

**Table 5.1:** Minimal and maximal possible number of regions in a Grasstope.

achieve fewer. We can see the maximal numbers of regions for other small  $k, n$  in Table 5.2. It would be interesting to see whether the maximal number of regions attained by reorienting and reordering a totally positive matrix attains the upper bound in general.

$k, n$	Maximal	$r(\mathbf{P})$	$\gamma(k, n)$
3, 7	42	42	42
3, 8	64	64	99
4, 8	64	99	64
5, 8	29	120	29
2, 9	37	37	247
3, 9	93	93	219
4, 9	163	163	163

**Table 5.2:** Maximal number of regions from reorienting and reordering a positive matrix.

**Example 5.4.2.** Any totally positive  $6 \times 3$ -matrix with the second and fourth rows negated yields a Grasstope which includes all 16 regions counted by Equation (1.5.1). The resulting hyperplane arrangement is cyclic, with just two orientations flipped. See Figure 5.4 to see all of the regions labelled with sign patterns.  $\blacklozenge$

**Example 5.4.3.** An example of a  $6 \times 3$ -matrix whose Grasstope has 16 regions is any totally positive matrix with the 2nd and 4th rows swapped. For examples of totally positive matrices, one can take the Vandermonde matrix

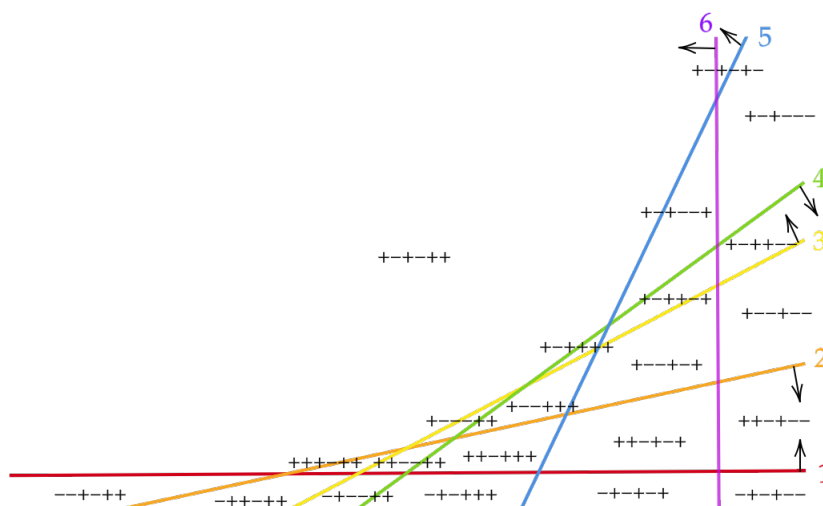
$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ d_1 & d_2 & d_3 & d_4 & d_5 & d_6 \\ d_1^2 & d_2^2 & d_3^2 & d_4^2 & d_5^2 & d_6^2 \end{bmatrix}^T$$

with  $0 < d_1 < \dots < d_6$ .  $\blacklozenge$

Note that the lower bound  $r(\mathbf{P}) - \beta(k, n)$  for the number of regions in a Grasstope is actually attained for all  $k$  and  $n$ , by the  $m = 1$  amplituhedron [KW19, Proposition 6.14, Theorem 6.16]. The upper bound is also attained by the  $m = 1$  amplituhedron for the values in Tables 5.1 and 5.2. One interesting question to study is to determine whether this holds in general, and to describe all oriented matroids achieving this upper bound.

The dictionary between hyperplane arrangements and oriented matroids (Remark 1.5.13) guides us to generalize our definition of a Grasstope to oriented matroids that are not necessarily realizable, such that the definitions agree when  $\mathcal{M} = \mathcal{M}_Z$ .

**Definition 5.4.4** (Grasstope of an oriented matroid). Let  $\mathcal{M}$  be an oriented matroid of rank  $r$ , and  $<$  be a total order on the ground set  $\mathcal{E}$  of  $\mathcal{M}$ . Then we define the Grasstope  $\mathcal{G}(\mathcal{M}, <)$  to be the subset of covectors  $\{v : \overline{var}(v) \geq r - 1\}$ , where  $r$  is the rank of  $\mathcal{M}$  and the variation is with respect to  $<$ .



**Figure 5.4:** The Grasstopes of a totally positive matrix with two rows negated. The six lines are cyclically ordered with orientations indicated by arrows. Every region has at least two sign changes, so the Grasstopes is all of  $\mathbb{RP}^2$ .

Note that by Topological Representation Theorem [FL78, Theorem 20], every oriented matroid arises from a pseudoline arrangement, with covectors labelling the cells of this arrangement. In particular, the Grasstopes  $\mathcal{G}(\mathcal{M}, <)$  can be identified with the union of cells of a pseudoline arrangement that satisfy the sign variation condition from Definition 5.4.4. Therefore, Grasstopes of oriented matroids are meaningful geometric objects, and topological concepts such as connectedness and contractibility generalize naturally to them. Studying their topological properties is an interesting topic for future research.

Finally, we use our code to analyze a non-realizable example, which does not attain the upper bound.

**Example 5.4.5.** Consider the non-realizable matroid  $\text{FMR}(8)$  of rank 4 on 8 elements, whose signed cocircuits are given in [RS88, Table 1]. Reorientations and reorderings give at least 34 and at most 63 regions. Thus, unlike the amplituhedron,  $\text{FMR}(8)$  does not achieve the upper bound of 64. ♦

The material presented in this chapter is, to the best of our knowledge, the first attempt to study Grasstopes systematically. It therefore offers a plethora of ideas and questions for future research. The most obvious way to proceed is to study Grasstopes for higher  $m$ , in particular for  $m = 2$  and  $m = 4$ , where abundant combinatorial results are already available for the amplituhedron. However, even at  $m = 1$ , there are still many interesting questions. What is the topology of  $m = 1$  Grasstopes, how bad can it be? How to characterize the boundary of open Grasstopes? What are the combinatorial and geometric properties of oriented matroid Grasstopes when the matroid is realized by a pseudoline arrangement?

# Bibliography

- [AH23] Yulia Alexandr and Serkan Hoşten. Maximum Information Divergence from Linear and Toric Models. *arXiv:2308.15598*, 2023.
- [Alh23] Álvaro M Alhambra. Quantum Many-Body Systems in Thermal Equilibrium. *PRX Quantum*, 4:040201, 2023.
- [AHBHY18] Nima Arkani-Hamed, Yuntao Bai, Song He, and Gongwang Yan. Scattering Forms and the Positive Geometry of Kinematics, Color and the Worldsheet. *Journal of High Energy Physics*, 2018(5):1–78, 2018.
- [AHBL17] Nima Arkani-Hamed, Yuntao Bai, and Thomas Lam. Positive Geometries and Canonical Forms. *Journal of High Energy Physics*, 2017(11):1–124, 2017.
- [AHHL21] Nima Arkani-Hamed, Song He, and Thomas Lam. Stringy Canonical Forms. *Journal of High Energy Physics*, 2021(2):1–62, 2021.
- [AHT14] Nima Arkani-Hamed and Jaroslav Trnka. The Amplituhedron. *Journal of High Energy Physics*, 2014(10):1–33, 2014.
- [Ax71] James Ax. On Schanuel’s Conjectures. *Annals of Mathematics*, 93(2):252–268, 1971.
- [Bat17] Heather S Battey. Eigen Structure of a New Class of Structured Covariance and Inverse Covariance Matrices. *Bernoulli*, 23:3166–3177, 2017.
- [Bat23] Heather S Battey. Inducement of Population Sparsity. *Canadian Journal of Statistics*, 51(3):760–768, 2023.
- [BT97] Dimitris Bertsimas and John N Tsitsiklis. *Introduction to Linear Optimization*. Athena Scientific, 1997.
- [BGS93] Louis J Billera, Israel M Gelfand, and Bernd Sturmfels. Duality and Minors of Secondary Polyhedra. *Journal of Combinatorial Theory, Series B*, 57(2):258–268, 1993.
- [BLVS<sup>+</sup>99] Anders Björner, Michel Las Vergnas, Bernd Sturmfels, Neil White, and Günter M Ziegler. *Oriented Matroids*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2 edition, 1999.
- [BPT12] Grigoriy Blekherman, Pablo A Parrilo, and Rekha R Thomas. *Semidefinite Optimization and Convex Algebraic Geometry*, volume 13 of *MOS-SIAM Series on Optimization*. SIAM, 2012.



- [BCR13] Jacek Bochnak, Michel Coste, and Marie-Françoise Roy. *Real Algebraic Geometry*, volume 36. Springer Science & Business Media, 2013.
- [BSST23] Michael Borinsky, Anna-Laura Sattelberger, Bernd Sturmfels, and Simon Telen. Bayesian Integrals on Toric Varieties. *SIAM Journal on Applied Algebra and Geometry*, 7(1):77–103, 2023.
- [BKSU18] Paul Breiding, Sara Kališnik, Bernd Sturmfels, and Madeleine Weinstein. Learning Algebraic Varieties from Samples. *Revista Matemática Complutense*, 31:545–593, 2018.
- [BT18] Paul Breiding and Sascha Timme. HomotopyContinuation.jl: A package for homotopy continuation in Julia. In *Mathematical Software–ICMS 2018: 6th International Conference, South Bend, IN, USA, July 24–27, 2018, Proceedings 6*, pages 458–465. Springer, 2018.
- [Bri04] Emmanuel Briand. When is the Algebra of Multisymmetric Polynomials Generated by the Elementary Multisymmetric Polynomials? *Beiträge zur Algebra und Geometrie: Contributions to Algebra and Geometry*, 45 (2), 353–368., 2004.
- [BP12] Winton Brown and David Poulin. Quantum Markov Networks and Commuting Hamiltonians. *arXiv:1206.0755*, 2012.
- [CHKS06] Fabrizio Catanese, Serkan Hoşten, Amit Khetan, and Bernd Sturmfels. The Maximum Likelihood Degree. *American Journal of Mathematics*, 128(3):671–697, 2006.
- [CH71] Peter Clifford and John M Hammersley. Markov Fields on Finite Graphs and Lattices. 1971.
- [CEĹ22] Sam Cole, Michał Eckstein, Shmuel Friedland, and Karol Życzkowski. Quantum Monge-Kantorovich Problem and Transport Distance Between Density Matrices. *Physical Review Letters*, 129(11):110402, 2022.
- [CLS11] David A Cox, John B Little, and Henry K Schenck. *Toric Varieties*, volume 124. American Mathematical Soc., 2011.
- [DS95] John Dalbec and Bernd Sturmfels. Introduction to Chow Forms. In Neil L White, editor, *Invariant Methods in Discrete and Computational Geometry*, pages 37–58. Springer Netherlands, Dordrecht, 1995.
- [Dav57] Chandler Davis. All Convex Invariant Functions of Hermitian Matrices. *Archiv der Mathematik*, 8(4):276–278, 1957.
- [DLSV12] Jesús A De Loera, Bernd Sturmfels, and Cynthia Vinzant. The Central Curve in Linear Programming. *Foundations of Computational Mathematics*, 12:509–540, 2012.
- [DLRS10] Jesús A De Loera, Jörg Rambau, and Francisco Santos. *Triangulations*. Springer Berlin, Heidelberg, 2010.
- [DGM21] Serena Di Giorgio and Paulo Mateus. On the Complexity of Finding the Maximum Entropy Compatible Quantum State. *Mathematics*, 9(2), 2021.

- [DGMM20] Serena Di Giorgio, Paulo Mateus, and Bruno Mera. Recoverability from Direct Quantum Correlations. *Journal of Physics A: Mathematical and Theoretical*, 53(18):185301, 2020.
- [DS98] Persi Diaconis and Bernd Sturmfels. Algebraic Algorithms for Sampling from Conditional Distributions. *The Annals of Statistics*, 26(1):363–397, 1998.
- [DSS08] Mathias Drton, Bernd Sturmfels, and Seth Sullivant. *Lectures on Algebraic Statistics*, volume 39. Springer Science & Business Media, 2008.
- [DPW23] Eliana Duarte, Dmitrii Pavlov, and Maximilian Wiesmann. Algebraic Geometry of Quantum Graphical Models. *arXiv:2308.11538*, 2023.
- [DHJ<sup>+</sup>19] Timothy Duff, Cvetelina Hill, Anders Jensen, Kisun Lee, Anton Leykin, and Jeff Sommars. Solving Polynomial Systems via Homotopy Continuation and Monodromy. *IMA Journal of Numerical Analysis*, 39(3):1421–1446, 2019.
- [EGP<sup>+</sup>23] Nick Early, Alheydis Geiger, Marta Panizzut, Bernd Sturmfels, and Claudia He Yun. Positive del Pezzo Geometry. *arXiv:2306.13604*, 2023.
- [Efr22] Bradley Efron. *Exponential Families in Theory and Practice*. Institute of Mathematical Statistics Textbooks. Cambridge University Press, 2022.
- [Eis87] David Eisenbud. On the Resiliency of Determinantal Ideals. In *Commutative Algebra and Combinatorics*, volume 11, pages 29–39. Mathematical Society of Japan, 1987.
- [Eis88] David Eisenbud. Linear Sections of Determinantal Varieties. *American Journal of Mathematics*, 110(3):541–575, 1988.
- [Eva20] Robin J Evans. Model Selection and Local Geometry. *The Annals of Statistics*, 48(6):3513 – 3544, 2020.
- [EZLT21] Chaim Even-Zohar, Tsviqa Lakrec, and Ran J Tessler. The Amplituhedron BCFW Triangulation. *arXiv:2112.02703*, 2021.
- [FMS21] Claudia Fevola, Yelena Mandelshtam, and Bernd Sturmfels. Pencils of Quadrics: Old and New. *Le Matematiche*, 76(2), 2021.
- [Fin] Lukas Finschi. Homepage of Oriented Matroids. <https://finschi.com/math/om/?p=home>.
- [Fin01] Lukas Finschi. *A Graph Theoretical Approach for Reconstruction and Generation of Oriented Matroids*. PhD Dissertation, Swiss Federal Institute of Technology Zürich, 2001.
- [FL78] Jon Folkman and Jim Lawrence. Oriented Matroids. *Journal of Combinatorial Theory, Series B*, 25(2):199–236, 1978.
- [FdW22] Jens Forsgård and Timo de Wolff. The Algebraic Boundary of the Sonc-cone. *SIAM Journal on Applied Algebra and Geometry*, 6(3):468–502, 2022.
- [FMM13] Komei Fukuda, Hiroyuki Miyata, and Sonoko Moriyama. Complete Enumeration of Small Realizable Oriented Matroids. *Discrete & Computational Geometry*, 49:359–381, 2013.

- [Ful98] William Fulton. *Intersection Theory*. Springer Science+Business Media, New York, 1998.
- [Gae20] Christian Gaetz. Positive Geometries Learning Seminar, Canonical Forms of Polytopes from Adjoints. *Unpublished lecture notes, available at <https://sites.google.com/view/crgaetz/research>*, 2020.
- [GL20] Pavel Galashin and Thomas Lam. Parity Duality for the Amplituhedron. *Compositio Mathematica*, 156(11):2207–2262, 2020.
- [GS21] Francesco Galuppi and Mima Stanojkovski. Toric Varieties from Cyclic Matrix Semigroups. *Rendiconti dell’Istituto di Matematica dell’Università di Trieste: an International Journal of Mathematics*, 53(17), 2021.
- [GK50] Feliks R Gantmaher and Mark G Krein. Oscillyacionye Matricy i Yadra i Malye Kolebaniya Mehaniceskih Sistem (in Russian). *Gosudarstv. Isdat. Tehn.-Teor. Lit.*, 1950.
- [GPS10] Luis David Garcia-Puente and Frank Sottile. Linear Precision for Parametric Patches. *Advances in Computational Mathematics*, 33:191–214, 2010.
- [GMS06] Dan Geiger, Christopher Meek, and Bernd Sturmfels. On the Toric Algebra of Graphical Models. *The Annals of Statistics*, 34(3):1463 – 1492, 2006.
- [Gir82] Kurt Girstmair. Linear Dependence of Zeros of Polynomials and Construction of Primitive Elements. *Manuscripta mathematica*, 39(1):81–97, 1982.
- [Gir99] Kurt Girstmair. Linear Relations between Roots of Polynomials. *Acta Arithmetica*, 89(1):53–96, 1999.
- [GS] Daniel R Grayson and Michael E Stillman. Macaulay2, a Software System for Research in Algebraic Geometry. Available at <http://www2.macaulay2.com>.
- [Gül96] Osman Güler. Barrier Functions in Interior Point Methods. *Mathematics of Operations Research*, 21(4):860–885, 1996.
- [Har13] Robin Hartshorne. *Algebraic Geometry*, volume 52. Springer Science & Business Media, 2013.
- [HRS18] Jonathan D Hauenstein, Jose Israel Rodriguez, and Frank Sottile. Numerical Computation of Galois Groups. *Foundations of Computational Mathematics*, 18(4):867–890, 2018.
- [HJPW04] Patrick Hayden, Richard Jozsa, Dénes Petz, and Andreas Winter. Structure of States which Satisfy Strong Subadditivity of Quantum Entropy with Equality. *Communications in Mathematical Physics*, 246:359–374, 2004.
- [HEB04] Marc Hein, Jens Eisert, and Hans J Briegel. Multiparty Entanglement in Graph States. *Physical Review A*, 69(6):062311, 2004.
- [HL08] Søren Højsgaard and Steffen L Lauritzen. Graphical Gaussian Models with Edge and Vertex Symmetries. *Journal of the Royal Statistical Society*, 70:1005–1027, 2008.
- [HJ85] Roger Horn and Charles R Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1985.

- [HJ94] Roger Horn and Charles R Johnson. *Topics in Matrix Analysis*. Cambridge University Press Cambridge, UK, 1994.
- [HKS05] Serkan Hosten, Amit Khetan, and Bernd Sturmfels. Solving the Likelihood Equations. *Foundations of Computational Mathematics*, 5:389–407, 2005.
- [Huh13] June Huh. The Maximum Likelihood Degree of a Very Affine Variety. *Compositio Mathematica*, 149(8):1245–1266, 2013.
- [HS14] June Huh and Bernd Sturmfels. Likelihood Geometry. *Combinatorial algebraic geometry*, 2108:63–117, 2014.
- [IS14] Corey Irving and Hal Schenck. Geometry of Wachspress Surfaces. *Algebra & Number Theory*, 8(2):369–396, 2014.
- [JS21] Yuhan Jiang and Bernd Sturmfels. Bad Projections of the PSD Cone. *Collectanea mathematica*, 72(2):261–280, 2021.
- [Jos21] Michael Joswig. *Essentials of Tropical Combinatorics*, volume 219. American Mathematical Society, 2021.
- [Kar17] Steven N Karp. Sign Variation, the Grassmannian, and Total Positivity. *Journal of Combinatorial Theory, Series A*, 145:308–339, 2017.
- [KM23] Steven N Karp and John Machacek. Shelling the  $m = 1$  Amplituhedron. *Combinatorial Theory*, 3(1), 2023.
- [KW19] Steven N Karp and Lauren K Williams. The  $m = 1$  Amplituhedron and Cyclic Hyperplane Arrangements. *International Mathematics Research Notices*, 2019(5):1401–1462, 2019.
- [Kit17] Yoshiyuki Kitaoka. Notes on the Distribution of Roots Modulo a Prime of a Polynomial. *Uniform distribution theory*, 12(2):91–117, 2017.
- [KPR<sup>+</sup>21] Kathlén Kohn, Ragni Piene, Kristian Ranestad, Felix Rydell, Boris Shapiro, Rainer Sinn, Miruna-Stefana Sorea, and Simon Telen. Adjoints and Canonical Forms of Polypols. *arXiv:2108.11747*, 2021.
- [KR20] Kathlén Kohn and Kristian Ranestad. Projective geometry of Wachspress coordinates. *Foundations of Computational Mathematics*, 20:1135–1173, 2020.
- [KSS20] Kathlén Kohn, Boris Shapiro, and Bernd Sturmfels. Moment Varieties of Measures on Polytopes. *Annali della Scuola Normale Superiore di Pisa (Classe Scienze), Serie V*, 21:739–770, 2020.
- [KF09] Daphne Koller and Nir Friedman. *Probabilistic Graphical Models: Principles and Techniques*. MIT press, 2009.
- [Lam16] Thomas Lam. Totally Nonnegative Grassmannian and Grassmann Polytopes. In *Current Developments in Mathematics 2014*, pages 51–152. Int. Press, Somerville, MA, 2016.
- [Lam22] Thomas Lam. An Invitation to Positive Geometries. *arXiv:2208.05407*, 2022.
- [Lam23] Thomas Lam. Personal communication, 2023.

- [Lan19] Joseph M Landsberg. A Very Brief Introduction to Quantum Computing and Quantum Information Theory for Mathematicians. In Edoardo Ballico, Alessandra Bernardi, Iacopo Carusotto, Sonia Mazzucchi, and Valter Moretti, editors, *Quantum Physics and Geometry*, pages 5–41. Springer, 2019.
- [Lau96] Steffen L Lauritzen. *Graphical Models*, volume 17. Clarendon Press, 1996.
- [LP08] Matthew S Leifer and David Poulin. Quantum Graphical Models and Belief Propagation. *Annals of Physics*, 323(8):1899–1946, 2008.
- [LR73] Elliott H Lieb and Mary Beth Ruskai. Proof of the Strong Subadditivity of Quantum-Mechanical Entropy. *Les rencontres physiciens-mathématiciens de Strasbourg-RCP25*, 19:36–55, 1973.
- [LUSB14] Shaowei Lin, Caroline Uhler, Bernd Sturmfels, and Peter Bühlmann. Hyper-surfaces and Their Singularities in Partial Correlation Testing. *Foundations of Computational Mathematics*, 14(5):1079–1116, 2014.
- [LPW23] Tomasz Lukowski, Matteo Parisi, and Lauren K Williams. The Positive Tropical Grassmannian, the Hypersimplex, and the  $m = 2$  Amplituhedron. *International Mathematics Research Notices*, 03 2023. rrad010.
- [MDLW18] Marloes Maathuis, Mathias Drton, Steffen Lauritzen, and Martin Wainwright. *Handbook of Graphical Models*. CRC Press, 2018.
- [MPP23] Yelena Mandelshtam, Dmitrii Pavlov, and Elizabeth Pratt. Combinatorics of  $m = 1$  Grasstopes. *arXiv:2307.09603*, 2023.
- [MM23] Léo Mathis and Chiara Meroni. Fiber Convex Bodies. *Discrete & Computational Geometry*, 70(4):1451–1475, 2023.
- [MHMT23] Saiee-Jaeyeong Matsubara-Heo, Sebastian Mizera, and Simon Telen. Four Lectures on Euler Integrals. *SciPost Phys. Lect. Notes*, page 75, 2023.
- [MW98] Mathieu Meyer and Elisabeth Werner. The Santaló-Regions of a Convex Body. *Transactions of the American Mathematical Society*, 350(11):4569–4591, 1998.
- [MS21a] Mateusz Michałek and Bernd Sturmfels. *Invitation to Nonlinear Algebra*, volume 211 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, 2021.
- [MS21b] Joaquín Moraga and Hendrik Süß. Bounding Toric Singularities with Normalized Volume. *arXiv:2111.01738*, 2021.
- [Nem06] Arkadi Nemirovski. Advances in Convex Optimization: Conic Programming. In *International Congress of Mathematicians*, volume 1, pages 413–444, 2006.
- [NN94] Yurii Nesterov and Arkadii Nemirovskii. *Interior-Point Polynomial Algorithms in Convex Programming*. SIAM, 1994.
- [NRS10] Jiawang Nie, Kristian Ranestad, and Bernd Sturmfels. The Algebraic Degree of Semidefinite Programming. *Mathematical Programming*, 122:379–405, 2010.
- [NGKG13] Sönke Niekamp, Tobias Galla, Matthias Kleinmann, and Otfried Gühne. Computing Complexity Measures for Quantum States Based on Exponential Families. *Journal of Physics A: Mathematical and Theoretical*, 46(12):125301, 2013.

- [NC02] Michael A Nielsen and Isaac Chuang. *Quantum Computation and Quantum Information*. American Association of Physics Teachers, 2002.
- [Oja90] Manuel Ojanguren. The Witt group and the problem of Lüroth. 1990.
- [OSC24] OSCAR – Open Source Computer Algebra Research system, Version 0.14.0, 2024.
- [PS05] Lior Pachter and Bernd Sturmfels. *Algebraic Statistics for Computational Biology*, volume 13. Cambridge university press, 2005.
- [PSBW23] Matteo Parisi, Melissa Sherman-Bennett, and Lauren K Williams. The  $m = 2$  Amplituhedron and the Hypersimplex: Signs, Clusters, Tilings, Eulerian Numbers. *Communications of the American Mathematical Society*, 3(07):329–399, 2023.
- [Pav23] Dmitrii Pavlov. Logarithmically Sparse Symmetric Matrices. *arXiv:2301.10042*, 2023.
- [PST23] Dmitrii Pavlov, Bernd Sturmfels, and Simon Telen. Gibbs Manifolds. *Information Geometry*, pages 1–27, 2023.
- [PT24] Dmitrii Pavlov and Simon Telen. Santaló Geometry of Convex Polytopes. *arXiv:2402.18955*, 2024.
- [Pet86] Dénes Petz. Sufficient Subalgebras and the Relative Entropy of States of a von Neumann Algebra. *Communications in Mathematical Physics*, 105:123–131, 1986.
- [Pos06] Alexander Postnikov. Total Positivity, Grassmannians, and Networks. *arXiv math/0609764*, 2006.
- [PH11] David Poulin and Matthew B Hastings. Markov Entropy Decomposition: a Variational Dual for Quantum Belief Propagation. *Physical Review Letters*, 106(8):080403, 2011.
- [RST24] Kristian Ranestad, Rainer Sinn, and Simon Telen. Adjoints and Canonical Forms of Tree Amplituhedra. *arXiv:2402.06527*, 2024.
- [RS88] Jean-Pierre Roudneff and Bernd Sturmfels. Simplicial Cells in Arrangements and Mutations of Oriented Matroids. *Geometriae Dedicata*, 27(2), August 1988.
- [Sch22] Claus Scheiderer. Extreme Points of Gram Spectrahedra of Binary Forms. *Discrete & Computational Geometry*, 67(4):1174–1190, 2022.
- [Sha13] Igor R Shafarevich. *Basic Algebraic Geometry*. Springer, Berlin, 2013.
- [SW<sup>+</sup>05] Andrew J Sommese, Charles W Wampler, et al. *The Numerical Solution of Systems of Polynomials Arising in Engineering and Science*. World Scientific, 2005.
- [Stu88] Bernd Sturmfels. Totally Positive Matrices and Cyclic Polytopes. *Linear Algebra and its Applications*, 107:275–281, 1988.

- [STVvR24] Bernd Sturmfels, Simon Telen, François-Xavier Vialard, and Max von Renesse. Toric Geometry of Entropic Regularization. *Journal of Symbolic Computation*, 120:102221, 2024.
- [SU10] Bernd Sturmfels and Caroline Uhler. Multivariate Gaussians, Semidefinite Matrix Completion, and Convex Algebraic Geometry. *Annals of the Institute of Statistical Mathematics*, 62(4):603–638, 2010.
- [Sul18] Seth Sullivant. *Algebraic Statistics*, volume 194. American Mathematical Soc., 2018.
- [Syl83] James Joseph Sylvester. XXXIX. On the Equation to the Secular Inequalities in the Planetary Theory. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 16(100):267–269, 1883.
- [Tel22] Simon Telen. Introduction to Toric Geometry. *arXiv:2203.01690*, 2022.
- [TV15] Tomáš Tyc and Jan Vlach. Quantum Marginal Problems. *The European Physical Journal D*, 69:1–6, 2015.
- [URBY13] Caroline Uhler, Garvesh Raskutti, Peter Bühlmann, and Bin Yu. Geometry of the Faithfulness Assumption in Causal Inference. *The Annals of Statistics*, 41(2):436–463, 2013.
- [Vig99] Eric Vigoda. *Sampling from Gibbs Distributions*. PhD Dissertation, Computer Science Dept., UC Berkeley, 1999.
- [War96] Joe Warren. Barycentric Coordinates for Convex Polytopes. *Advances in Computational Mathematics*, 6:97–108, 1996.
- [WG23] Stephan Weis and João Gouveia. The Face Lattice of the Set of Reduced Density Matrices and its Coatoms. *Information Geometry*, pages 1–34, 2023.
- [Wil13] Mark M Wilde. *Quantum Information Theory*. Cambridge University Press, 2013.
- [Wil21] Lauren K Williams. The Positive Grassmannian, the Amplituhedron, and Cluster Algebras. In *International Congress of Mathematicians*, 10 2021.
- [Zas75] Thomas Zaslavsky. *Facing up to Arrangements: Face-Count Formulas for Partitions of Space by Hyperplanes*, volume 154 of *Memoirs of the American Mathematical Society*. 1975.
- [Zho08] Duanlu Zhou. Irreducible Multiparty Correlations in Quantum States without Maximal Rank. *Physical Review Letters*, 101:180505, Oct 2008.
- [Zie12] Günter M Ziegler. *Lectures on Polytopes*, volume 152. Springer Science & Business Media, 2012.

## Bibliographische Daten

---

Real Algebraic Geometry for Physics and Optimization  
(Reale Algebraische Geometrie für Physik und Optimierung)  
Pavlov, Dmitrii  
Universität Leipzig, Dissertation, 2024  
120 Seiten, 14 Abbildungen, 129 Referenzen



## **Selbstständigkeitserklärung**

Hiermit erkläre ich, die vorliegende Dissertation selbständig und ohne unzulässige fremde Hilfe angefertigt zu haben. Ich habe keine anderen als die angeführten Quellen und Hilfsmittel benutzt und sämtliche Textstellen, die wörtlich oder sinngemäß aus veröffentlichten oder unveröffentlichten Schriften entnommen wurden, und alle Angaben, die auf mündlichen Auskünften beruhen, als solche kenntlich gemacht. Ebenfalls sind alle von anderen Personen bereitgestellten Materialien oder erbrachten Dienstleistungen als solche gekennzeichnet.

Leipzig, den 16. August 2024

.....  
(Dmitrii Pavlov)

## Daten zum Autor

---

**Name:** Dmitrii Pavlov  
**Geburtsdatum:** 05.08.1998 in Chabarowsk (Russland)  
**09/2016 - 06/2022** Spezialist (M.Sc.) in Mathematik  
Staatliche Lomonossow-Universität Moskau  
**seit 08/2022** Doktorand der Mathematik