Metric Algebraic Geometry of Grassmannians

by

Hannah Friedman

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Mathematics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Bernd Sturmfels, Co-chair
Professor Serkan Hoşten, Co-chair
Professor David Eisenbud
Professor Nikhil Srivastava

Spring 2026

The dissertation of Hannah Friedman, titled Metric Algebraic Geometry of Grassmannians, is approved:

Co-chair	____________________________	Date	____________
Co-chair	____________________________	Date	____________
	____________________________	Date	____________
	____________________________	Date	____________

University of California, Berkeley

Metric Algebraic Geometry of Grassmannians

Abstract

Metric Algebraic Geometry of Grassmannians

by

Hannah Friedman

Doctor of Philosophy in Mathematics

University of California, Berkeley

Professor Bernd Sturmfels, Co-chair

Professor Serkan Hoşten, Co-chair

The Grassmannian $\mathrm{Gr}(k, n)$ is the set of k-dimensional subspaces of the vector space $\mathbb{R}^n$. It is an important object in algebraic geometry, but also in applications, particularly statistics and data science. The Grassmannian is an algebraic variety, and therefore has equational representations. The equational representations one uses in algebraic geometry and in applications are different, and the interplay between these different representations will be an important tool and object of study for us. The equational representation used in algebraic geometry is called the Plücker embedding, and we introduce the projection Grassmannian for the equational representation from applications.

Metric algebraic geometry is a relatively new field which bridges algebraic and differential geometry by investigating metric aspects of algebraic varieties. One important aspect of metric algebraic geometry is the study of algebraic optimization: given a data point and a variety, what point on the variety is closest to the data point? We answer this question by computing critical points of the optimization problem. The number of complex critical points is constant for generic data and is called the algebraic degree of the problem.

We develop a framework for algebraic optimization and describe the prime ideals of the Grassmannian and related varieties in the first part of the thesis. In Parts II and III, we study optimization problems on $\mathrm{Gr}(k, n)$ from an algebraic perspective.

In the second part of the thesis, we explore various polynomial optimization problems on the Grassmannian from the perspective of metric algebraic geometry. The multi-eigenvector problem, canonical correlation analysis, and correspondence analysis are all polynomial optimization problems on Grassmannians and flag varieties. Their critical points are all real, and come from the spectral theorem and the singular value decomposition. We then turn to a constrained eigenvalue problem in quantum chemistry, namely the optimization of the Rayleigh quotient on tensor train varieties. We prove that tensor train varieties are parametrized by products of Grassmannians. We compute critical points and data discriminants for this

problem. We conclude with a study of Euclidean distance optimization for subvarieties of the Grassmannian with data also in the Grassmannian. This leads us to introduce the Grassmann distance degree as an algebraic measure of the difficulty of this problem. We characterize the Schubert varieties for which this problem is particularly simple.

In the third part, we take the perspective of algebraic statistics and study maximum likelihood estimation on a family of probability distributions called determinantal point processes. We prove that, for a dense open subset of determinantal point processes, the maximum likelihood estimation problem can be solved by recursion. We illuminate the relationship between projection determinantal point processes and the Grassmannian via a new variety called the squared Grassmannian $\mathrm{sGr}(k,n)$. We prove that, for $k=2$, all complex critical points of the maximum likelihood estimation problem on the squared Grassmannian are real and positive. We introduce squared linear models and show that they share this remarkable property.

Contents

List of Figures

List of Tables

Acknowledgments

I am extremely grateful for my advisors, Bernd Sturmfels and Serkan Hoşten. I feel very lucky to have worked with you these past few years and have benefitted so much from your mentorship. Thank you for introducing me to the wonderful nonlinear algebra community, and for teaching me how to be a mathematician. I am grateful for all of your advice and encouragement, and I have been continuously inspired by you these past few years.

Next, I want to thank my collaborators: Yulia Alexandr, Renato G. Bettiol, Viktoriia Borovik, Paul Breiding, Taylor Brysiewicz, Kristen Dawson, Karel Devriendt, Fatemeh Mohammadi, Max Pfeffer, Bernhard Reinke, Andrea Rosana, Pardis Semnani, Maximilian Wiesmann, Teresa Yu, and Maksym Zubkov. I have learned so much from each of you and have enjoyed working with you immensely.

Thank you to my teachers at UC Berkeley: Ian Agol, David Eisenbud, Mark Haiman, Peter Haine, and Andrés Vindas Meléndez. Thank you, David Eisenbud and Nikhil Srivastava for taking the time to be on my dissertation committee. For your mentorship and helpful mathematical conversations, I additionally want to thank Aldo Conca, Claudia Fevola, Christian Gaetz, Alheydis Geiger, Vadim Gorin, Jon Hauenstein, Ben Hollering, Julia Lindberg, Jose Israel Rodriguez, Otto Schmidt, Anna Seigal, Luca Sodomaco, Simon Telen, Cynthia Vinzant, and Claudia Yun. Thank you, Clay, Christian, Saskia, and Mirke for your help with administrative tasks.

I am grateful for the support and camaraderie of the combinatorics community at UC Berkeley and I thank the algebraic statistics community for being so welcoming. Thank you to my friends and colleagues at UC Berkeley and at the Max Planck Institute for Mathematics in the Sciences, and my academic siblings, Lizzie, Svala, and Smita. I am so thankful that I've gotten to know and work with all of you these past few years.

Finally, I want to thank my friends and family. I am so grateful to have you in my life, and I would not be the same person without you. To my friends, thank you for all the little things: the hotpot, the hikes, the emotional guarding, the improv, the book recommendations, the dancing, the yarn, the long drives, and the beverages. To my family, thank you for always supporting and believing in me. Mom and Dad, thank you for your enduring encouragement and enthusiasm.

Part I

Foundations

Chapter 1

Overview

The theme of this thesis is optimization on Grassmannians. The Grassmannian, denoted $\mathrm{Gr}(k,n)$, is the set of k-dimensional vector spaces of $\mathbb{R}^n$. It is a central object in mathematics (algebraic geometry, combinatorics, differential geometry) and in applications (statistics, machine learning, quantum chemistry, physics). The aim of an optimization problem on the Grassmannian is to find the best k-dimensional subspace, for some notion of "best." For instance, given a symmetric matrix, the problem of finding its best rank-k approximation is an optimization problem on $\mathrm{Gr}(k,n)$. We view the Grassmannian as an algebraic variety and adopt the approach of metric algebraic geometry [19] to solve these optimization problems. This turns optimization problems on $\mathrm{Gr}(k,n)$ into systems of polynomial equations. In Part I, we introduce and give background on algebraic optimization and on Grassmannians separately. In Part II, we investigate various polynomial optimization problems on the Grassmannian that arise in applications. In Part III, we turn our focus another optimization problem, namely, maximum likelihood estimation for determinantal point processes; these are statistical models closely related to the Grassmannian. We now give an overview of the contents of each chapter, followed by a summary of the main results in this thesis.

1.1 Outline and Contributions

In Chapter 2, we present a general introduction to the algebraic perspective on distance optimization. Most importantly, we define the *algebraic degree* of an optimization problem as the number of complex critical points of that problem for generic data, and give conditions under which this degree is finite. We introduce three kinds of general optimization problems: optimizing a polynomial on an affine variety, optimizing a degree zero rational function on a projective variety, and maximum likelihood estimation on a projective variety. Maximum likelihood estimation is an important optimization problem in algebraic statistics, the aim of which is to find the point in a statistical model that best explains a given data point. In each of these scenarios, we introduce the *critical correspondence*, which is the incidence variety of data and critical points, describe its prime ideal, and give formulas for the degrees of generic models. We conclude the chapter with parametric analogs of the above concepts and a few words on numerical computations.

In Chapter 3, we collect background information on the Grassmannian and related varieties. This chapter is largely based on Section 2 of joint work with Serkan Hoşten [47], which has been submitted for publication. In addition, Sections 3.3 and 3.6 are based on Sections 5 and 2, respectively, of joint work with Karel Devriendt, Bernhard Reinke, and Bernd Sturmfels [34], which was published in *Acta Universitatis Sapientiae, Mathematica*. In particular, we discuss different embeddings of the Grassmannian and flag varieties and how they relate to one another. The effect of this interplay on optimization is a central theme of our work. We prove that these embeddings are smooth, and we describe their prime ideals.

Chapter 4 is based on joint work with Serkan Hoşten [47]. We study the algebraic degrees of various optimization problems from linear algebra and statistics, namely the multi-eigenvector problem, the heterogeneous quadrics minimization problem, canonical correlation, and correspondence analysis. We prove that all complex critical points are real for the multi-eigenvector, canonical correlation, and correspondence analysis problems. In fact, we give complete descriptions of the critical points for each of these problems.

In Chapter 5, we study a constrained eigenvalue problem on tensor train varieties from the perspective of Chapter 2. We prove that tensor train varieties are parametrized by products of Grassmannians. We introduce and study the discriminant of this problem. We conclude with numerical experiments on small tensor train varieties. This is based on joint work with Viktoriia Borovik, Serkan Hoşten, and Max Pfeffer. This work has been submitted for publication and a preprint is available [16].

We turn to distance minimization within the Grassmannian of lines in Chapter 6. This is based on joint work with Andrea Rosana and Bernd Sturmfels. We study the image of the rational map sending a skew-symmetric matrix to its square. This map is closely connected to the different realizations of the Grassmannian in Chapter 3. We introduce the Grassmann distance degree, which is the algebraic degree of Euclidean distance minimization on a subvariety of the Grassmannian where the data point is also in the Grassmannian. We describe how to compute the critical points of such a problem, with a particular focus on examples in the Grassmannian of lines in 3-space and Schubert varieties. It has been submitted for publication and a preprint is available [48].

In Part III, we focus on maximum likelihood estimation for determinantal point processes. We consider a subclass of determinantal point processes called *L-ensembles* in Chapter 7. We prove that the critical points with a given block structure can be found by solving a series of smaller maximum likelihood estimation problems. We conclude with numerical experiments, showing that the maximum likelihood degree, which is defined as the algebraic degree of the maximum likelihood estimation problem, grows rapidly with the model. This is based on joint work with Bernd Sturmfels and Maksym Zubkov. It was published in *Algebraic Statistics* [50].

In Chapter 8, we consider a subclass of determinantal point processes that is disjoint from the subclass of L-ensembles; these are called *projection* determinantal point processes. We uncover the relationship between projection determinantal point processes and the Grassmannian via a new variety called the *squared Grassmannian*. We find the maximum likelihood degree of the squared Grassmannian and prove that all critical points are real and positive. We also give a formula for the maximum likelihood degree of the Grassmannian. Section 8.2 is based on [34, Section 3], and Section 8.6 is partially based on [34, Section 4]. The remainder of the chapter is based on [46], which was published in *Proceedings of the American Mathematical Society*.

In Chapter 9, we study statistical models that are parametrized by squares of linear forms. All critical points of the likelihood function are real and positive. There is one critical point in each region of the projective hyperplane arrangement defined by the linear forms. We examine the ideal and singular locus of the model, and we describe generators of its likelihood correspondence. We characterize tropical degenerations of the MLE, describe the log-normal polytopes, and explore connections to determinantal point processes. This is based on joint work [49] with Bernd Sturmfels and Maximilian Wiesmann, which was published in *SIAM Journal on Applied Algebra and Geometry*.

1.2 Main Results

The main results of this thesis come in three flavors: (1) describing maps between varieties, (2) describing prime ideals of varieties, and (3) describing critical points of optimization problems. For the reader's convenience, we list the main results here; we explain and state a few of these results below.

The main results describing maps between varieties are

- Theorem 3.6.2: Plücker embedding to projection Grassmannian;
- Theorem 5.4.3: product of Grassmannians to tensor train variety.

The following results are descriptions of prime ideals for various important varieties:

- Theorem 2.2.3: critical correspondence;
- Theorem 3.3.2: projection Grassmannian;
- Theorem 3.4.2: projection flag variety;
- Theorem 3.5.2: isospectral model;
- Theorem 8.2.3: squared Grassmannian;
- Theorem 9.3.2: parametric likelihood correspondence for squared linear models.

The remaining results have to do with the critical points of various optimization problems:

- Theorem 4.1.6: critical points of the multi-eigenvector problem are eigenspaces of the data matrix;

- Theorem 6.3.7: the ED degree is strictly greater than the GD degree for subvarieties of the Grassmannian of lines having dimension at least 4;
- Theorem 6.5.6: certain Schubert varieties in the Grassmannian have GD degree 1;
- Theorem 7.2.1: the parametric maximum likelihood estimation problem for L-ensembles can be solved recursively;
- Theorem 8.1.4: ML degree of the squared Grassmannian;
- Theorem 8.1.5: all complex critical points of the maximum likelihood estimation problem for the squared Grassmannian are real and positive;
- Theorem 8.6.2: ML degree of the Grassmannian;
- Theorem 9.1.1: all complex critical points of the maximum likelihood estimation problem for squared linear models are real and positive.

We now explain and state four of these results in more detail.

The Projection Grassmannian

A central theme in this work is the interplay of the two "lives" of the Grassmannian; see Chapter 3. The first life is familiar to algebraic geometers: it is the Plücker embedding, which realizes the Grassmannian as a smooth projective variety. The prime ideal of this variety is generated by the Plücker relations; see Theorem 3.2.1.

The second life of the Grassmannian is more prevalent in applications. If A is a real $n \times k$ matrix whose columns span a point in the Grassmannian, then we can represent the column span of A by the unique orthogonal projection matrix $P = A(A^T A)^{-1} A^T$ onto that k-dimensional vector space. The matrix P is symmetric and satisfies $P^2 = P$ and $\mathrm{rank}(P) = k$. Because the minimal polynomial of P divides $z^2 - z$, the eigenvalues of P are all either 0 or 1, and so we can replace the rank condition with the equation $\mathrm{trace}(P) = k$. We therefore define the projection Grassmannian as the zero set of the equations $P^2 - P$ and $\mathrm{trace}(P) = k$. In fact, these equations generate the prime ideal of the projection Grassmannian.

Theorem 3.3.2 *The ideal $\langle P^2 - P\rangle$ is radical, and the ideal $\langle P^2 - P, \mathrm{trace}(P) - k\rangle$ is prime for all k. The projection Grassmannian is smooth.*

The proof uses the transitive action of the orthogonal group on the Grassmannian along with the Jacobian criterion. Similar arguments work for the Stiefel manifold (Theorem 3.1.2) and for generalizations of the projection Grassmannian to the projection flag variety (Theorem 3.4.2) and to the isospectral model (Theorem 3.5.2).

The ML Degree of the Positive Grassmannian

We now turn to the Plücker embedding of the Grassmannian; see Section 3.2. We will view the positive Grassmannian $\mathrm{Gr}(k,n)_{>0}$ as a statistical model. The positive Grassmannian is the set of points (x_I) in $\mathrm{Gr}(k,n)$ in the Plücker embedding with $x_I \geq 0$ for all k-subsets I of $[n] = \{1,\ldots,n\}$. A point (x_I) in Plücker coordinates is a distribution on k-subsets of $[n]$ where the probability of choosing a set I is $x_I / \sum_J x_J$. Given a data point $u = (u_I)$, the log-likelihood function is

$$L_u(x) = \sum_I u_I \log(x_I) - (\sum_I u_I) \log(\sum_I x_I). \tag{1.1}$$

Here, each sum is over the k-subsets of $[n]$. The point (x_I) which maximizes L_u is called the maximum likelihood estimate, and it is the distribution that was most likely sampled from if one observes the data u. This is easier to see from the likelihood function $\exp(L_u(X)) = \prod_I x_I^{u_I} / (\sum_I x_I)^{\sum_I u_I}$, which gives the probability of observing the data point u when sampling from the distribution x. The log-likelihood function and likelihood function have the same maximizer. The maximum likelihood degree (ML degree) of the positive Grassmannian is the number of complex critical points of the constrained optimization problem

$$\max_{x \in \mathrm{Gr}(k,n)} L_u(x).$$

The ML degree is an algebraic measure of complexity of the optimization problem, and it is an important invariant of a variety in algebraic statistics [27, 68, 122]. The ML degrees of the Grassmannian were first computed in [68] for $k = 2$ and $n = 4, 5$. We will now present a general formula for the ML degree for $k = 2$; this is a partial answer to [68, Problem 13].

Theorem 8.6.2 *The positive Grassmannian* $\mathrm{Gr}(2,n)_{>0}$ *has ML degree* $(2^{n-1} - n)(n-3)!$.

Positive Critical Points

Once one has computed the ML degree of a variety, one might ask how many critical points are statistically relevant, i.e., how many correspond to actual probability distributions. It can happen that even though the ML degree is high, the number of relevant local optima is low. This happens for linear and toric models; see [122, Corollary 7.3.8]. In both cases, there is exactly one critical point in the probability simplex. In these cases, maximum likelihood estimation is easy, even though the ML degree might be high. The two main results we present here describe the opposite phenomenon. To the best of our knowledge, these are the first examples of models for which all complex critical points of the log-likelihood function are real and positive.

The first model is called the squared Grassmannian $\mathrm{sGr}(2,n)$; see Section 8.2. This variety is the Zariski closure of the image of the coordinatewise squaring map on the Grassmannian in Plücker coordinates:

$$\mathrm{Gr}(2,n) \longrightarrow \mathbb{P}^{\binom{n}{2}-1}, \qquad (x_I) \mapsto (q_I) = (x_I^2).$$

The state space of the model is the set of 2-subsets of $[n]$, and the probability of observing a set I is $x_I^2/\sum_J x_J^2 = q_I/\sum_J q_J$. The corresponding log-likelihood function is (1.1), which we evaluate at points $q \in \mathrm{sGr}(2, n)$.

Theorem 8.1.5. *For $u \in \mathbb{N}^{\binom{n}{2}}$, each complex critical point of the log-likelihood function $L_u(q)$ over $\mathrm{sGr}(2, n)$ is real, positive, and a local maximum.*

The statistical model we have just described is an example of a determinantal point process; see Section 8.1. Theorem 8.1.4 states that ML degree of this model is $(n-1)!/2$. Statistically, these results imply that maximum likelihood estimation is very difficult for this model. In particular, using a local gradient ascent method is very likely to become trapped in a spurious local maximum. To find the true global maximum, one would need to compute all local maxima and evaluate the objective function at each one.

The second example is the family of squared linear models. The variety $\mathcal{M}$ which underlies the model is the image-closure of the rational map

$$\mathbb{P}^{s-1} \dashrightarrow \mathbb{P}^{n-1}, \qquad\qquad x \mapsto y = (\ell_1^2(x) : \cdots : \ell_n^2(x))$$

where $\ell_1(x), \ldots, \ell_n(x)$ are linear forms with real coefficients. The points in the model are probability distributions with state space $[n]$ where the probability of observing the event i is $\ell_i^2(x)/\sum_j \ell_j^2(x) = y_i/\sum_j y_j$. The corresponding log-likelihood function is

$$L_u(y) = \sum_{i=1}^n u_i y_i - (\sum_{i=1}^n u_i)\log(\sum_{i=1}^n y_i).$$

Theorem 9.1.1. *For generic data $u \in \mathbb{R}^n_{>0}$, every complex critical point of $L_u(y)$ on the model $\mathcal{M} \subset \mathbb{P}^{n-1}$ is real, positive and a local maximum.*

For this model, the ML degree is the number of regions in the projective hyperplane arrangement defined by $\ell_1, \ldots, \ell_2$; this is computed as $|\chi(-1)|/2$ where χ is the characteristic polynomial of the matroid represented by the hyperplane arrangement; see Corollary 9.1.4. As the ML degree of squared linear models grows rapidly, and all critical points are positive, maximum likelihood esitmation is very challenging for squared linear models as well.

Chapter 2

The Algebraic Degree of an Optimization Problem

Given polynomials $f_1, \ldots, f_r \in \mathbb{R}[x_1, \ldots, x_n]$, let $\mathcal{M} = V(f_1, \ldots, f_r)$ be an algebraic variety of codimension c whose real points $\mathcal{M}_{\mathbb{R}}$ are Zariski dense in $\mathcal{M}$. If $f_1, \ldots, f_r$ are homogeneous, then we view $\mathcal{M}$ as a projective variety in complex projective space $\mathbb{P}^{n-1}$. Otherwise, $\mathcal{M} \subseteq \mathbb{C}^n$ is an affine variety. We write $d_i = \deg(f_i)$ for $i = 1, \ldots, r$ and $I_{\mathcal{M}} = \langle f_1, \ldots, f_r \rangle$ for the ideal of $\mathcal{M}$. We consider optimization problems of the form

$$\min_{x \in \mathcal{M}_{\mathbb{R}}}/\max g_u(x_1, \ldots, x_n) \tag{2.1}$$

where $g_u : \mathcal{M} \to \mathbb{C}$ is a function that is real-valued on $\mathcal{M}_{\mathbb{R}}$, i.e., $g_u(\mathcal{M}_{\mathbb{R}}) \subseteq \mathbb{R}$. We also require that the partial derivatives of g_u are algebraic functions of the variables $x_1, \ldots, x_n$, and linear functions of the parameters $u_1, \ldots, u_m$. We write $\nabla g_u(x)$ for the gradient vector of g_u.

In order to exactly solve the constrained optimization problem (2.1), one computes the critical points. These are smooth points x such that $\nabla g_u(x)$ is orthogonal to the tangent space $T_x(\mathcal{M})$ of $\mathcal{M}$ at x. We can find this set of critical points using Lagrange multipliers by solving the following system of polynomial equations

$$\nabla(g_u(x) + \lambda_1 f_1(x) + \cdots + \lambda_r f_r(x)) = 0.$$

where partial derivatives are taken with respect to $x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_r$. An equivalent formulation that avoids introducing new variables is that the gradient $\nabla g_u(x)$ is in the rowspan of the $r \times n$ Jacobian

$$\mathcal{J}_{\mathcal{M}}(x) = \begin{pmatrix} -\nabla f_1(x)- \\ \vdots \\ -\nabla f_r(x)- \end{pmatrix} \tag{2.2}$$

After finding all the critical points, we plug them back into the objective function g_u to find the minimum or maximum. The complexity of this process can be measured by the number of critical points. This is the motivation for the following definition.

Definition 2.0.1. The *algebraic degree* of the optimization problem (2.1) is the number of complex critical points for generic data u.

The algebraic degree is a measure of the difficulty of the optimization problem (2.1). For instance, if (2.1) has one critical point for generic data, then (2.1) may be solved correctly and efficiently using local methods. Furthermore, any optimal solution is an algebraic function of the input data of this degree [99, 100, 108]. Finally, there is a paradigm in numerical algebraic geometry, that computing the solutions to a system of polynomial equations is faster if the number of solutions is known a priori; see Section 2.5. We will show that the algebraic degree is often finite.

To do this, we introduce the *critical correspondence* of (2.1). This is the set

$$\mathcal{C}_{\mathcal{M}} = \overline{\{(x,u) \in \mathcal{M}^\circ \times D : x \text{ is critical for } u\}}$$

where $\mathcal{M}^\circ$ is an open subset of $\mathcal{M}$ to be specified and D is the data space $\mathbb{C}^m$ or $\mathbb{P}^{m-1}$. The critical correspondence naturally projects onto the first and second factors:

$$\begin{tikzcd} & \mathcal{C}_{\mathcal{M}} \arrow[dl, "\pi_1"'] \arrow[dr, "\pi_2"] & \\ \mathcal{M} & & D \end{tikzcd}$$

Given data $u \in D$, the fiber $\pi_2^{-1}(u)$ is the set of critical points for (2.1). In Euclidean distance optimization, the critical correspondence is called the ED correspondence [37], and in likelihood geometry, it is called the likelihood correspondence [70]. We present here a general framework that encompasses both of these perspectives.

This chapter is organized as follows. In Section 2.1, we introduce the three concrete scenarios that appear in this paper, namely optimization of a polynomial over an affine variety, optimization of a rational function on a projective variety, and optimization of the log-likelihood function on a projective variety. In each scenario, we will give an explicit description of the critical correspondence $\mathcal{C}_{\mathcal{M}}$, and we introduce a running example for each of these scenarios. In Section 2.2, we prove that the critical correspondence is a vector bundle over $\mathcal{M}^\circ$ of dimension $\dim(D)$ in the three scenarios covered in Section 2.1. In particular, this proves that the algebraic degree is finite in these scenarios. We also describe the prime ideal of $\mathcal{C}_{\mathcal{M}}$ (Theorem 2.2.3), which in turn gives a description of the critical ideal for fixed data u (Corollary 2.2.4). In Section 2.3, we collect formulae from the literature for the algebraic degree of a generic complete intersection (Theorem 2.3.2). In Section 2.4, we give parametric analogs of the above: we define a parametric correspondence which is also a vector bundle (Theorem 2.4.2), describe the prime ideal of this correspondence (Theorem 2.4.4), and describe the critical ideal for fixed data (Corollary 2.4.5). We give an upper bound on the number of critical points of a general rational function (Theorem 2.4.7). In Section 2.5, we conclude with comments about how one solves these optimization problems numerically.

2.1 Scenarios in Algebraic Optimization

The three important examples for us will be when g_u is a polynomial and $\mathcal{M}$ is an affine variety (Chapter 4), when g_u is a rational function and $\mathcal{M}$ is a projective variety (Chapters 5 and 6), and when g_u is the log-likelihood function and $\mathcal{M}$ is a projective variety (Chapters 7, 8, and 9). We introduce these in detail now.

Scenario 1 (Affine)**.** We now take $g_u(x)$ to be a polynomial in x with $\deg(g_u) = e$ and $\mathcal{M} \subseteq \mathbb{C}^n$ to be an affine variety. Examples from the literature include the Euclidean distance $g_u(x) = ||x-u||^2$ [37] and the linear function $g_u(x) = u_1x_1 + \cdots + u_nx_n$ [96]. The critical points are smooth points for which $\nabla g_u(x)$ is in the row span of the Jacobian matrix $\mathcal{J}_{\mathcal{M}}(x)$ in (2.2). Equivalently, we look for smooth points x where the augmented Jacobian

$$\mathcal{J}^u_{\mathcal{M}}(x) = \begin{pmatrix} \nabla g_u(x) \\ \mathcal{J}_{\mathcal{M}}(x) \end{pmatrix} \tag{2.3}$$

drops rank. We therefore write the critical correspondence explicitly as

$$\mathcal{C}_{\mathcal{M}} = \overline{\{(x,u) \in \mathcal{M}_{\text{reg}} \times \mathbb{C}^m : \operatorname{rank}(\mathcal{J}^u_{\mathcal{M}}(x)) = c\}} \subseteq \mathbb{C}^n \times \mathbb{C}^m. \tag{2.4}$$

Example 2.1.1 (Cuspidal Cubic)**.** Given a data point $(u,v) \in \mathbb{R}^2$, what is the closest point on the cuspidal cubic defined by $f = x^3 - y^2$? The answer is given by the solution of

$$\min_{(x,y)\in V(f)} (x-u)^2 + (y-v)^2.$$

The augmented Jacobian equals $\mathcal{J}^{(u,v)}_{\mathcal{M}}(x,y) = \begin{pmatrix} x-u & y-v \\ 3x^2 & -2y \end{pmatrix}$ and the critical correspondence $\mathcal{C}_{V(f)} \subseteq \mathbb{C}^2 \times \mathbb{C}^2$ is found by imposing the condition that the augmented Jacobian has rank at most one and saturating by the singularity:

$$\langle x^3 - y^2, \det(\mathcal{J}^{(u,v)}_{\mathcal{M}}(x,y))\rangle : \langle x, y\rangle^\infty.$$

Given data $(u,v) = (2,1)$, the critical points are the four solutions to

$$\begin{aligned} 2x^2 - 4x + 3y^2 - 3y &= 0 & \qquad 9xy - 41x + 18y^2 - 12y + 16 &= 0 \\ -16xy + 12x + 9y^3 - 18y^2 + 17y &= 0 & \qquad 9xy^2 - 18xy + 17x + 6y^2 - 6y - 16 &= 0. \end{aligned}$$

Two of these four critical points are real, as seen in Figure 2.1. ⋄

Scenario 2 (Projective)**.** We now wish to optimize the quotient $g_u = h_u/q$ where h_u and q are homogeneous polynomials of degree e in x and q does not depend on the data $u \in \mathbb{P}^{m-1}$. One important example is the Rayleigh quotient $g_u(x) = x^TUx/x^Tx$ where U is an $n \times n$ symmetric matrix; see Chapter 5. In practice, $V(q)$ usually has no real points, so g_u is well-defined on real projective space $\mathbb{P}^{n-1}_{\mathbb{R}}$. For instance, in this work, q will always be the

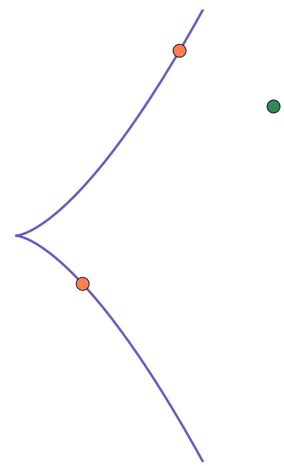

Figure 2.1: Real critical points of the ED problem for the cuspidal cubic $x^3 = y^2$ with data $(u, v) = (2, 1)$ (green). The global minimizer is $(0.51798, 1.27026)$ and the other critical point is $(-0.3728, 1.43165)$ (orange).

isotropic quadric $x_1^2 + \cdots + x_n^2$. Here $\mathcal{M} \subseteq \mathbb{P}^{n-1}$ is a projective variety that is not contained in $V(q)$. We then define the augmented Jacobian matrix

$$\mathcal{J}_{\mathcal{M}}^u(x) = \begin{pmatrix} \nabla h_u(x) \\ \nabla q(x) \\ \mathcal{J}_{\mathcal{M}}(x). \end{pmatrix} \tag{2.5}$$

In addition to excluding the singular points of $\mathcal{M}$, we now also exclude points where g_u is not well-defined. Hence, the critical correspondence is now

$$\mathcal{C}_{\mathcal{M}} = \overline{\{(x, u) \in \mathcal{M}_{\text{reg}} \backslash V(q) \times \mathbb{P}^{m-1} : \operatorname{rank}(\mathcal{J}_{\mathcal{M}}^u(x)) = c + 1\}}. \tag{2.6}$$

Lemma 2.1.2. *For every point $(x, u) \in \mathcal{C}_{\mathcal{M}} \cap (\mathcal{M}_{\text{reg}} \backslash V(q) \times \mathbb{P}^{m-1})$, the matrix* $\begin{pmatrix} \nabla q \\ \mathcal{J}_{\mathcal{M}}(x) \end{pmatrix}$ *has rank $c + 1$ and x is a critical point of (2.1) for the data u.*

Proof. Since x is smooth, $\mathcal{J}_{\mathcal{M}}(x)$ has rank c, so we just need to show that $\nabla q(x)$ is not in the rowspan. Suppose there exists $\lambda_1, \ldots, \lambda_r$ with $\nabla q(x) = \lambda_1 \nabla f_1(x) + \cdots + \lambda_r \nabla f_r(x)$. Multiplying on the right by the column vector x and applying Euler's Theorem, we find

$$e \cdot q(x) = \lambda_1 d_1 f_1(x) + \cdots + \lambda_r d_r f_r(x) = 0$$

where the second equality holds because $x \in \mathcal{M} = V(f_1, \ldots, f_r)$. Since $x \notin V(q)$, the matrix has rank $c + 1$.

We now argue that x is critical. By the quotient rule, the gradient of the objective function is

$$\nabla g_u(x) = \nabla \left(\frac{h_u}{q} \right)(x) = \frac{1}{q(x)^2}(q(x) \nabla h_u(x) - h_u(x) \nabla q(x)).$$

Thus x is critical if and only if $q(x)^2\nabla g_u(x) = q(x)\nabla h_u(x) - h_u(x)\nabla q(x)$ is in the rowspan of $\mathcal{J}_{\mathcal{M}}(x)$. Since $(x,u) \in \mathcal{C}_{\mathcal{M}} \cap (\mathcal{M}_{\text{reg}}\backslash V(q) \times \mathbb{P}^{m-1})$, there exist $\mu_1 \neq 0, \mu_2$ such that $\mu_1\nabla h_u(x) - \mu_2\nabla q(x)$ is in the rowspan of $\mathcal{J}_{\mathcal{M}}(x)$. We argue that $\mu_2/\mu_1 = g_u(x)$. Indeed, writing $\mu_1\nabla h_u(x) - \mu_2\nabla q(x)$ as a linear combination of the rows of $\mathcal{J}_{\mathcal{M}}(x)$, multiplying by x, and applying Euler's Theorem as above, we find that $\mu_1 h_u(x) = \mu_2 q(x)$. Since $\mu_1, q(x) \neq 0$ we have the desired equality $\mu_2/\mu_1 = g_u(x)$ and thus $\nabla g_u(x)$ is in the rowspan of $\mathcal{J}_{\mathcal{M}}$ and x is critical for u. □

Example 2.1.3 (Circle)**.** Set $\mathcal{M} = V(x^2+y^2-z^2) \subseteq \mathbb{P}^2$, let $h_U(x,y,z) = \begin{pmatrix} x & y & z\end{pmatrix} U \begin{pmatrix} x & y & z\end{pmatrix}^T$, and fix $q(x,y,z) = x^2+y^2+z^2$. Here the data space $\mathbb{P}(\mathrm{Sym}^2\mathbb{C}^n)$ is the space of symmetric 3×3 matrices. For $U \in \mathbb{P}(\mathrm{Sym}^2\mathbb{R}^n)$, the optimization problem

$$\min_{(x,y,z)\in V(x^2+y^2-z^2)}/\max \frac{h_U(x,y,z)}{q(x,y,z)}$$

is the best approximation of the eigenvector belonging to the smallest (resp. largest) eigenvalue of U while staying on the circle. The augmented Jacobian is

$$\mathcal{J}^U_{V(f)}(x,y,z) = \begin{pmatrix} u_{11}x + u_{12}y + u_{13}z & u_{12}x + u_{22}y + u_{23}z & u_{13}x + u_{23}y + u_{33}z \\ x & y & z \\ x & y & -z \end{pmatrix}. \tag{2.7}$$

Since the model is smooth, the critical correspondence $\mathcal{C}_{\mathcal{M}} \subseteq \mathbb{P}^2 \times \mathbb{P}^5$ is found by imposing a rank condition on the Jacobian and saturating by the denominator:

$$\langle x^2+y^2-z^2, \det(\mathcal{J}^U_{V(f)}(x,y,z))\rangle : \langle x^2+y^2+z^2\rangle^\infty.$$

For the data matrix

$$U = \begin{pmatrix} 1 & 3 & 5 \\ 3 & 7 & 11 \\ 5 & 11 & 13 \end{pmatrix}$$

the four critical points are the solutions to

$$6xy + 11xz - 6y^2 - 5yz + 3z^2 = 0 \qquad 4xz^2 + 6y^3 + 5y^2z - 7yz^2 - 4z^3 = 0$$
$$x^2+y^2-z^2 = 0$$

The two real critical points are shown in Figure 2.2. ⋄

Scenario 3 (Maximum Likelihood)**.** For maximum likelihood estimation, we set $n = m$ and

$$g_u(x) = \sum_{i=1}^n u_i \log(x_i) - (u_1 + \cdots + u_n)\log(x_1 + \cdots + x_n).$$

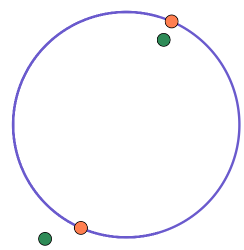

Figure 2.2: Real critical points of the constrained eigenvalue problem on the circle. The picture is in the affine chart $z = 1$. The orange points are the critical values. The green points represent the true eigenvectors of the matrix U corresponding to the largest and smallest eigenvalues. The point $[0.40102 : 0.91607 : 1]$ is the global maximizer and $[-0.40102 : -0.91607 : 1]$ is the global minimizer.

The following perspective on maximum likelihood estimation is called *likelihood geometry*. Likelihood geometry is an important topic in algebraic statistics [68, 69, 70, 122] and it is the setting of Part III.

The denominators that appear in the partial derivatives of g_u are $x_1, \ldots, x_n, x_1 + \cdots + x_n$. We therefore define the hyperplane arrangement $\mathcal{H} = \{x \in \mathbb{P}^{n-1} : x_1 \cdots x_n(x_1 + \cdots + x_n) = 0\}$ where these partial derivatives are not well-defined. By [122, Theorem 7.2.3], the appropriate augmented Jacobian is now the $(c+2) \times n$ matrix

$$\mathcal{J}^u_{\mathcal{M}}(x) = \begin{pmatrix} u_1 & \cdots & u_n \\ x_1 & \cdots & x_n \\ x_1 \frac{\partial f_1}{\partial x_1} & \cdots & x_n \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ x_1 \frac{\partial f_r}{\partial x_1} & \cdots & x_n \frac{\partial f_r}{\partial x_n} \end{pmatrix}. \tag{2.8}$$

Note the appearance of $x_1, \ldots, x_n$ in the last r rows. The critical correspondence is called the likelihood correspondence in this case:

$$\mathcal{C}_{\mathcal{M}} = \overline{\{(x, u) \in \mathcal{M}_{\text{reg}} \backslash \mathcal{H} \times \mathbb{P}^{n-1} : \operatorname{rank}(\mathcal{J}^u_{\mathcal{M}}(x)) = c + 1\}}. \tag{2.9}$$

In this situation, the algebraic degree is called the *maximum likelihood degree* (ML degree).

Example 2.1.4 (Binomial Distribution)**.** Suppose you are flipping a biased coin that has probability t of landing on heads and probability $1 - t$ of landing on tails. If you flip the coin twice, the probability of observing two heads is $x = t^2$, two tails is $z = (1 - t)^2$, and one of each is $y = 2t(1 - t)$. The set of probability distributions that can arise from this process is the intersection of $f = 4xz - y^2$ with the probability simplex $\Delta_2 = \{(x, y, z) \in \mathbb{R}^3 : x + y + z = 1, x \geq 0, y \geq 0, z \geq 0\}$. Flipping a coin with unknown bias, we observe that we see two heads u times, two tails w times and one of each v times. To find the bias t of the coin, we solve the maximum likelihood estimation problem

$$\max_{[x:y:z]\in V(f)} u\log(x) + v\log(y) + w\log(z) - (u + v + w)\log(x + y + z).$$

The augmented Jacobian is

$$\mathcal{J}_{V(f)}^{(u,v,w)}(x, y, z) = \begin{pmatrix} u & v & w \\ x & y & z \\ 4xz & -2y^2 & 4xz \end{pmatrix}$$

and the critical correspondence $\mathcal{C}_{\mathcal{M}} \subseteq \mathbb{P}^2 \times \mathbb{P}^2$ is cut out by

$$\begin{aligned} &\langle 4xz - y^2,\ \det(\mathcal{J}_{V(f)}^{(u,v,w)}(x, y, z))\rangle : \langle xyz(x + y + z)\rangle^{\infty} \\ = &\langle 4xz - y^2,\ yv + 2yw - 4zu - 2zv,\ 2xv + 4xw - 2yu - yv\rangle \end{aligned}$$

The ML degree of this model is 1. The data point $(u, v, w) = (2, 3, 5)$ has a real critical point in the simplex, namely $(x, y, z) = \frac{1}{400}(49, 182, 169)$; see Figure 2.3. This critical point has $t = \frac{7}{20}$, so the probability of the coin landing on heads was most likely $\frac{7}{20}$. $\diamond$

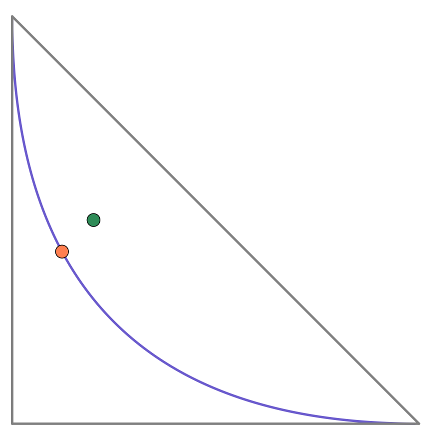

Figure 2.3: Maximum likelihood estimation for binomial distribution. Here we work in the affine chart $x+y+z = 1$. The horizontal axis is the x-axis and the vertical axis is the z-axis. The diagonal represents the third boundary component $y = 0$ of the probability simplex. The data point $(2, 3, 5)$ has a single critical point $\frac{1}{400}(49, 182, 169)$.

We summarize the above in the following table: the type of g_u, the open set $\mathcal{M}^{\circ}$ of $\mathcal{M}$ where we look for critical points, the data space D, the augmented Jacobian $\mathcal{J}_{\mathcal{M}}^{u}$, the critical correspondence $\mathcal{C}_{\mathcal{M}}$, and finally the ideal $I_{\mathcal{M}\setminus\mathcal{M}^{\circ}}$ defining the Zariski closed subset of $\mathcal{M}$ that we remove in the definition of $\mathcal{C}_{\mathcal{M}}$.

Scenario	g_u	$\mathcal{M}^\circ$	D	$\mathcal{J}^u_{\mathcal{M}}$	$\mathcal{C}_{\mathcal{M}}$	$I_{\mathcal{M}\setminus\mathcal{M}^\circ}$
Affine	polynomial	$\mathcal{M}_{\text{reg}}$	$\mathbb{C}^m$	(2.3)	(2.4)	$I_{\mathcal{M}_{\text{sing}}}$
Projective	rational function	$\mathcal{M}_{\text{reg}}\setminus V(q)$	$\mathbb{P}^{m-1}$	(2.5)	(2.6)	$I_{\mathcal{M}_{\text{sing}}} \cdot \langle q\rangle$
Maximum Likelihood	log-likelihood function	$\mathcal{M}_{\text{reg}}\setminus\mathcal{H}$	$\mathbb{P}^{n-1}$	(2.8)	(2.9)	$I_{\mathcal{M}_{\text{sing}}} \cdot I_{\mathcal{H}}$

The singular locus of $\mathcal{M}$ is the vanishing set of the ideal

$$I_{\mathcal{M}_{\text{sing}}} \;:=\; I_{\mathcal{M}} \;+\; \quad c\times c \text{ minors of } \mathcal{J}_{\mathcal{M}}(X) \tag{2.10}$$

and $I_{\mathcal{H}} = \langle x_1 \cdots x_n(x_1 + \cdots + x_n)\rangle$.

2.2 The Critical Correspondence

In this section, we will prove that the algebraic degree of a model is finite under the following assumption:

Assumption 2.2.1. *The linear map $u \mapsto \nabla g_u(x)$ is surjective for all $x \in \mathcal{M}$, and open set $\mathcal{M}^\circ$ is nonempty.*

In particular, this ensures that for every $x \in \mathcal{M}^\circ$, there exists some data u such that x is critical for u.

Theorem 2.2.2. *Suppose Assumption 2.2.1 holds. The correspondence $\mathcal{C}_{\mathcal{M}}$ is an irreducible variety of dimension $\dim(D)$. The projection $\pi_1 : \mathcal{C}_{\mathcal{M}} \to \mathcal{M}$ onto the first factor is an (affine or projective) vector bundle over $\mathcal{M}^\circ$. The projection $\pi_2 : \mathcal{C}_{\mathcal{M}} \to D$ is generically finite-to-one. The cardinality of the fiber $\pi_2^{-1}(u)$ is the algebraic degree of (2.1).*

Proof. For the maximum likelihood case, see [70, Theorem 1.6]. We prove the affine and projective cases. For all $x \in \mathcal{M}^\circ$, the fiber $\pi_1^{-1}(x)$ of the first projection is an (affine) linear space of dimension $c + m - n$ that varies smoothly with x. In the affine case, this is true because x being smooth implies that the rowspan of $\mathcal{J}^u_{\mathcal{M}}(x)$ is c-dimensional. For the projective case, one needs Lemma 2.1.2 for this statement. Since $\mathcal{M}$ is irreducible, the correspondence is as well and its dimension is $(n - c) + (c + m - n) = m$ in the affine case and $(n - c - 1) + (c + m - n) = m - 1$ in the projective case. Since the projection onto the second factor is dominant, the fibers are generically finite. □

We now describe the prime ideal of this critical correspondence.

Theorem 2.2.3. *Under Assumption 2.2.1, the critical correspondence $\mathcal{C}_{\mathcal{M}}$ has prime ideal*

$$I_{\mathcal{C}_{\mathcal{M}}} = (I_{\mathcal{M}} + \langle (c+i) \times (c+i) \textit{ minors of } \mathcal{J}^u_{\mathcal{M}}(X)\rangle) \;:\; I^\infty_{\mathcal{M}\setminus\mathcal{M}^\circ} \subseteq \mathbb{C}[x,u] \tag{2.11}$$

where $i \in \{1, 2\}$ is the number of rows added to form the augmented Jacobian $\mathcal{J}^u_{\mathcal{M}}(X)$. This ideal has dimension $\dim(D)$ in the affine case, and $\dim(D) + 1$ otherwise.

Before we prove this theorem, we state an important consequence. In practice, computing the critical correspondence is difficult. However, computing the defining ideal of the variety of critical points for fixed data u is much faster. We call this the *critical ideal*. This corollary gives a method for computing the critical points of (2.1) using computer algebra. This is reminiscent of [68, Algorithm 6].

Corollary 2.2.4. *For generic data $u \in D$, the critical ideal of (2.1) is*

$$I_{\mathcal{M}}^u = (I_{\mathcal{M}} + \langle (c+i) \times (c+i) \text{ minors of } \mathcal{J}_{\mathcal{M}}^u(x) \rangle) : I_{\mathcal{M}\setminus\mathcal{M}^\circ}^\infty \subseteq \mathbb{C}[x] \tag{2.12}$$

under Assumption 2.2.1, where i is the number of additional rows in $\mathcal{J}_{\mathcal{M}}^u(X)$. This ideal is zero-dimensional in the affine case, and one-dimensional otherwise. Its degree is equal to the algebraic degree of (2.1).

Proof. The dimension claim follows Theorems 2.2.3 and from the fact that for a generic point $\bar{u}$, the vector $u - \bar{u}$ forms a regular sequence in $\mathbb{C}[x,u]/I_{\mathcal{C}_{\mathcal{M}}}$. Since the ring $\mathbb{C}[x,u]/I_{\mathcal{C}_{\mathcal{M}}}$ is reduced, so is $\mathbb{C}[x,u]/(I_{\mathcal{C}_{\mathcal{M}}} + \langle u - \bar{u}\rangle) \cong \mathbb{C}[x]/I_{\mathcal{M}}^{\bar{u}}$ by Bertini's Theorem. Thus $I_{\mathcal{M}}^{\bar{u}}$ is the vanishing ideal of the fiber $\pi_2^{-1}(\bar{u})$ and therefore has degree equal to the algebraic degree of (2.1) by Theorem 2.2.2. □

Proof of Theorem 2.2.3. For simplicity of notation, we work in the affine case where $I_{\mathcal{M}\setminus\mathcal{M}^\circ} = I_{\mathcal{M}_{\text{sing}}}$, but the argument for the other cases is analogous. Write $R = \mathbb{C}[x]/I_{\mathcal{M}}$ for the coordinate ring of $\mathcal{M}$. Write $J = \langle (c+1) \times (c+1) \text{ minors of } \mathcal{J}_{\mathcal{M}}^u(x)\rangle$ for the ideal in $R[u]$. Since $\mathcal{C}_{\mathcal{M}}$ is irreducible, to show that $I_{\mathcal{M}}^u$ is prime, it suffices to prove that $J : I_{\mathcal{M}_{\text{sing}}}^\infty$ is radical. To do this, we pass to the localization $S = R_{I_{\mathcal{M}_{\text{sing}}}}$ at the ideal defining the singular locus. We prove that $J \cdot S[u]$ is radical, which implies that its preimage $J : I_{\mathcal{M}_{\text{sing}}}^\infty$ in $R[u]$ under the localization map is also radical.

To do so, we apply Lemma 2.2.5. Since S is a domain, it is also reduced. The gradient of the objective function is an affine linear function $\nabla g_u(x) = C \cdot u - a$ with $C \in S^{n\times m}$ and $a \in S^n$. By Assumption 2.2.1, C is surjective when evaluated at $x \in \mathcal{M}_{\text{reg}}$. The generators of J are linear combinations of the entries of $\nabla g_u(x)$ where the coefficients are minors of $\mathcal{J}_{\mathcal{M}}(x)$. Hence there exists some matrix D such that $J = \langle D \cdot (C \cdot u - a)\rangle = \langle (DC) \cdot u - D \cdot a\rangle$. Since C is surjective, $b = D \cdot a$ is in the rowspan of $L = DC$. For all smooth points in $\mathcal{M}_{\text{reg}}$, the matrix $L = D \cdot C$ has rank c. Thus, writing $M \cong S^m/\text{im}(L)$ for the cokernel of the map L, for all maximal ideals $\mathfrak{m} \subset S$, we have $\dim_{\kappa(\mathfrak{m})} M_{\mathfrak{m}}/\mathfrak{m}M_{\mathfrak{m}} = m - c$ where $\kappa(\mathfrak{m}) = S_{\mathfrak{m}}/\mathfrak{m} = \mathbb{C}$. Therefore by Lemma 2.2.5, $S[u]/J$ is reduced and hence (2.11) is radical and prime. □

Lemma 2.2.5. *Suppose that S a reduced noetherian ring and that $J = \langle L \cdot u - b\rangle \subseteq S[u]$ is an ideal of $S[u]$ with $L : S^n \to S^m$ and $b \in \text{im}(L)$. Write M for the cokernel of the map L. If $\dim_{\kappa(\mathfrak{m})} M_{\mathfrak{m}}/(\mathfrak{m}M)$ is constant for all maximal ideals $\mathfrak{m} \subseteq S$, then $S[u]/J$ is reduced.*

Proof. We first consider the homogeneous case $b = 0$ and begin by proving that $S[u]/J$ is flat over $S[u]$. By [6, Exercise 7.16], M is flat over S. Since taking the symmetric algebra of a module is a right exact functor [42, Proposition A2.2], it preserves cokernels and hence the symmetric algebra of M is precisely $S[u]/J$. By Lazard's Theorem, M is a filtered colimit of free modules, and so by right exactness, $S[u]/J$ is a filtered colimit of polynomial rings. Since these are also free, $S[u]/J$ is flat over S by another application of Lazard's theorem.

We check reducedness of $S[u]/J$ on maximal ideals. Since $\varphi : S \to S[u]/J$ is flat, the localization $S_{\varphi^{-1}(\mathfrak{m})} \to (S[u]/J)_{\mathfrak{m}}$ is flat. Note that the fiber rings $\kappa(\mathfrak{p}) \otimes_S S[y]/J$ are reduced for all primes $\mathfrak{p}$ of $S_{\mathfrak{m}}$. Since S is reduced, we may apply [95, Theorem 23.9], to obtain that $(S[u]/J)_{\mathfrak{m}}$ is reduced. Since $S[u]/J$ is reduced at every maximal ideal, it is reduced.

If $b \neq 0$, then there exists $u_0 \in S^n$ with $L \cdot u_0 = b$. There is an isomorphism $S[u]/\langle Lu - b\rangle \cong S[u]/\langle Lu\rangle$ given by $u \mapsto u + u_0$ and hence $S[u]/\langle Lu - b\rangle$ is reduced as well. □

2.3 Algebraic Degrees of Generic Models

We now state the algebraic degrees of these problems when the model is generic, i.e., when the coefficients of $f_1, \ldots, f_r$ are random. For this, we restrict to the projective and maximum likelihood scenarios. There is currently no general formula for the affine case, although one is given in the ED case in [37, Proposition 2.6].

For the projective case, we define the *Lagrangian locus* of any model $\mathcal{M}$ and data u, inspired by [107]:

$$\mathcal{L}^u_{\mathcal{M}} = \overline{\{x \in \mathcal{M}_{\text{reg}} : \operatorname{rank}(\mathcal{J}^u_{\mathcal{M}}(x)) < c + 2\}}$$

The critical locus is $\overline{\mathcal{L}^u_{\mathcal{M}} \setminus V(q)}$. We now show that $\mathcal{L}^u_{\mathcal{M}} \cap V(q)$ does not depend on the data u.

Proposition 2.3.1. *The set of isotropic points $\mathcal{L}^u_{\mathcal{M}} \cap V(q)$ in the Lagrangian locus contains the set of points at which $\mathcal{M}$ intersects $V(q)$ nontransversely for all u. For generic data u, every point in $L^u_{\mathcal{M}} \cap V(q)$ is a nontransverse intersection point.*

Proof. If $\mathcal{M}$ and $V(q)$ intersect nontransversely at x, then the matrix $\begin{pmatrix} \nabla q(x) \\ \mathcal{J}_{\mathcal{M}}(x) \end{pmatrix}$ drops rank and hence $x \in \mathcal{L}^u_{\mathcal{M}}$. This proves the first claim. For the second claim, suppose that $x \in \mathcal{L}^u_{\mathcal{M}} \cap V(q)$. Then there exist μ_1, μ_2 such that $\mu_1 \nabla h_u(x) - \mu_2 \nabla q(x)$ is in the rowspan of $\mathcal{J}_{\mathcal{M}}(x)$. Then as in the proof of Lemma 2.1.2, we can use Euler's relation to see that $\mu_1 h_u(x) = \mu_2 q(x) = 0$. For generic u, this implies μ_1 equals zero and so $\nabla q(x)$ is in the rowspan of $\mathcal{J}_{\mathcal{M}}$. Thus the intersection at x is nontransverse. □

If the model $\mathcal{M}$ has a transverse intersection with $V(q)$ in the projective case, then the critical locus equals the Lagrangian locus. For generic data u, we write $\Lambda^u_{\mathcal{M}}$ for the Lagrangian ideal:

$$\Lambda^u_{\mathcal{M}} = I_{\mathcal{M}} + \langle (c + 2 \times c + 2) \text{ minors of } \mathcal{J}^u_{\mathcal{M}}(x)\rangle : I^{\infty}_{\mathcal{M}_{\text{sing}}}. \tag{2.13}$$

Note that we are not saturating with respect to q. If $\mathcal{M}$ is a general complete intersection, then it is smooth, and the prime ideal $\Lambda^u_{\mathcal{M}} = I^u_{\mathcal{M}}$ can be computed without saturation.

Because we have an explicit description of this ideal without saturation, we can compute its degree using Bezout's Theorem and the Giambelli-Thom-Porteous formula [51, Theorem 14.4].

Theorem 2.3.2. *Suppose that $f_1, \dots, f_c$ form a regular sequence. The algebraic degree of (2.1) is at most*

(i) $d_1 \cdots d_c \sum_{j_0+\cdots+j_c=n-c-1} (j_0+1)(e-1)^{j_0}(d_1-1)^{j_1}\cdots(d_c-1)^{j_c}$ *in the projective scenario [107, 2.2.1],*

(ii) and $d_1 \cdots d_c \cdot \sum_{j_1+\cdots+j_c \leq n-c-1} d_1^{j_1}\cdots d_c^{j_c}$ *in the maximum likelihood scenario [68, Theorem 5].*

Equality holds if $\mathcal{M}$ is a generic complete intersection.

Proof. The references given show equality when $\mathcal{M}$ is a generic complete intersection. To see that the inequalities hold, we apply the argument in [113, Theorem 3.1]. Observe that $\mathcal{M}$ is an irreducible component of the complete intersection $V(f_1, \dots, f_c)$ and hence its algebraic degree must be less than or equal to the algebraic degree of $V(f_1, \dots, f_c)$. $\square$

Corollary 2.3.3. *Suppose $\mathcal{M}$ is an irreducible hypersurface of degree d. Then the algebraic degree of (2.1) is at most*

(i) $d\frac{(n-1)(e-1)^n - n(e-1)^{n-1}(d-1)+(d-1)^n}{(d-e)^2}$ *if $d \neq e$ and* $\binom{n}{2}e(e-1)^{n-2}$ *if $d = e$ in the projective scenario,*

(ii) and $d\frac{d^{n-1}-1}{d-1}$ *for $d > 1$ in the maximum likelihood scenario.*

If $e = 2$, the bound in the projective scenario is $d\frac{n(d-2)+(d-1)^n-1}{(d-2)^2}$ *for $d > 2$.*

Example 2.3.4 (Circle)**.** For the optimization problem in Example 2.1.3, the Lagrangian ideal is generated $x^2 + y^2 - z^2$ and the determinant of the Jacobian (2.7). By Bezout's Theorem, this ideal has degree 6, which matches the bound given by Corollary 2.3.3(i). We saw in Example 2.1.3 that the algebraic degree of the constrained eigenvalue problem on the circle is 4. The remaining two points are the nontransverse intersection points $[1 : i : 0], [1 : -i : 0]$ of the circle $V(x^2 + y^2 - z^2)$ with the isotropic quadric $V(x^2 + y^2 + z^2)$. $\diamond$

In specific cases, there are other formulae for degrees of specific optimization problems, for example, in terms of the polar classes of the model [37, 96]. We state here one such result that is specific to the maximum likelihood degree. The maximum likelihood degree of a variety is a topological quantity. The set $\mathcal{M}\backslash\mathcal{H} \subseteq (\mathbb{C}^*)^n$ is a *very affine variety*, meaning that it is a closed subvariety of the algebraic torus $(\mathbb{C}^*)^n$. The following theorem first appeared as [27, Theorem 19] under stronger hypotheses.

Theorem 2.3.5 ([69], Theorem 1)**.** *If the very affine variety $\mathcal{M}\backslash\mathcal{H}$ is smooth of dimension d, then the ML degree of $\mathcal{M}$ equals the signed Euler characteristic $(-1)^d\chi(\mathcal{M}\backslash\mathcal{H})$.*

Example 2.3.6 (Binomial Distribution)**.** Corollary 2.3.3(ii) gives a bound of 6 on the ML degree of the scaled rational quadric, but Theorem 2.3.5 allows us to compute the ML degree exactly. The conic $V(4xz - y^2)$ has Euler characteristic 2. It intersects the hyperplane arrangement $\mathcal{H} = V(xyz(x+y+z))$ in 3 points, namely $[0:0:1], [1:0:0], [1:0:1]$, so the Euler characteristic of the very affine variety $V(4xz-y^2)\backslash\mathcal{H}$ is $2-3=-1$. This is the negative of the ML degree. $\diamond$

2.4 Parametrization

The varieties one encounters in applications often come with parametrizations. Let $\phi : \mathbb{C}^s \dashrightarrow \mathcal{M}$ be a parametrization of $\mathcal{M}$ that is generically δ-to-one. Then the constrained optimization problem (2.1) can be replaced by the unconstrained optimization problem

$$\min_{t\in\mathbb{R}^s}/\max g_u(\phi(t)).$$

Almost all of the optimization problems we consider in this thesis will have parametric analogs. We fix the following notation. Let $\mathcal{J}_\phi(t)$ denote the $n \times s$ Jacobian of the map ϕ:

$$\mathcal{J}_\phi(t) = \begin{pmatrix} -\nabla\phi_1(t)- \\ \vdots \\ -\nabla\phi_n(t)- \end{pmatrix} \tag{2.14}$$

We assume that the set $\{\phi(t) : \operatorname{rank}(\mathcal{J}_\phi(t)) = s\}$ is Zariski dense in $\mathcal{M}$.

In a slight abuse of notation, we write $g_u(t)$ for $g_u(\phi(t))$. The chain rule gives $\nabla g_u(t) = \mathcal{J}_\phi(t)^T \cdot \nabla_x g_u(\phi(t))$, where $\nabla_x g_u(\phi(t))$ denotes vector whose entries are partial derivatives of g_u with respect to $x_1, \ldots, x_n$ evaluated at $\phi(t)$.

We will use the following result implicitly when we move between parametric and implicit formulations of a problem.

Lemma 2.4.1. *Let $\mathcal{W} \subset \mathbb{C}^s$ be an affine variety. Let $\phi : \mathcal{W} \dashrightarrow \mathcal{M}$ be a dominant rational map with base locus $\mathcal{B}$, and let g be a regular function. Then $t \in \mathcal{W}_{\mathrm{reg}}\backslash\mathcal{B}$ and $x = \phi(t) \in \mathcal{M}_{\mathrm{reg}}$ are critical points of $g' = g \circ \phi$ and g, respectively, if the differential $d\phi_t : T_t(\mathcal{W}) \to T_x(\mathcal{M})$ is surjective.*

Proof. The differential $d\phi_t$ is, in coordinates, equal to the linear map given by the Jacobian of ϕ restricted to $T_t(\mathcal{W})$. Suppose $t \in \mathcal{W}_{\mathrm{reg}}\backslash\mathcal{B}$ is a critical point of g'. This means that $\nabla_t(g') \in T_t(\mathcal{W})^\perp$, i.e., $\nabla_t(g') \cdot v = 0$ for all $v \in T_t(\mathcal{W})$. Equivalently, $\nabla_x(g) \cdot \mathcal{J}_\phi(t) \cdot v = 0$ for all $v \in T_t(\mathcal{W})$. This implies that $\nabla_x(g)$ is orthogonal to the subspace of the tangent space $T_x(\mathcal{M})$ that is the image of $T_t(\mathcal{W})$ under the differential $d\phi_t$. Now, if $d\phi_t$ is surjective, then the gradient $\nabla_x(g)$ must be in $T_x(\mathcal{M})^\perp$, and hence x is a critical point of g.

For the converse, pick a critical point $x = \phi(t) \in \mathcal{M}_{\mathrm{reg}}$ of g, i.e., x with $\nabla_x(g)$ in $T_x(\mathcal{M})^\perp$. By the surjectivity of the differential $d\phi_t$ the image of $T_t(\mathcal{W})$ under the differential is $T_x(\mathcal{M})$, and since $\nabla_t(g') \cdot v = \nabla_x(g) \cdot \mathcal{J}_\phi(t) \cdot v = 0$ for all $v \in T_x(\mathcal{W})$, we conclude that t is a critical point of g'. $\square$

If we restrict to the case $\mathcal{W} = \mathbb{C}^s$, then the above result states that critical points are preserved by ϕ when $\mathcal{J}_\phi(t)$ has rank s. We therefore define the *parametric critical correspondence* as

$$\mathcal{P}_\mathcal{M} = \overline{\{(t,u) \in \phi^{-1}(\mathcal{M}^\circ)\backslash\mathcal{V}_{\phi,s} \times D : \phi(t) \text{ is critical for } u\}} \tag{2.15}$$

where $\mathcal{V}_{\phi,s} = \{t \in \mathbb{C}^s : \operatorname{rank}(\mathcal{J}_\phi(t)) < s\}$.

Theorem 2.4.2. *The parametric critical correspondence $\mathcal{P}_\mathcal{M}$ of $\mathcal{M} = \overline{\operatorname{im}(\phi)}$ is an irreducible variety of dimension $\dim(D)$. The first projection $\gamma_1 : \mathcal{P}_\mathcal{M} \to \mathbb{C}^s$ is an (affine or projective) vector bundle over $\phi^{-1}(\mathcal{M}^\circ)\backslash\mathcal{V}_{\phi,s}$. The second projection $\gamma_2 : \mathcal{P}_\mathcal{M} \to D$ is generically finite-to-one. The size of the generic fiber $\gamma_2^{-1}(u)$ is δ times the algebraic degree of the optimization problem.*

Proof. By Lemma 2.4.1 and the definitions of $\mathcal{P}_\mathcal{M}$ and $\mathcal{C}_\mathcal{M}$, the following diagram commutes:

$$\begin{tikzcd}
\mathcal{P}_\mathcal{M} \arrow[r, dashed, "\phi\times\mathrm{Id}"] \arrow[d, "\gamma_1"] & \mathcal{C}_\mathcal{M} \arrow[d, "\pi_1"] \\
\mathbb{C}^s \arrow[r, dashed, "\phi"] & \mathcal{M}
\end{tikzcd}$$

This implies that the fibers $\gamma_1^{-1}(t)$ and $\pi_1^{-1}(\phi(t))$ are isomorphic. Thus, the vector bundle property, irreducibility, and dimension all follow from Theorem 2.2.2. For the second projection, we have the commutative diagram

$$\begin{tikzcd}
\mathcal{P}_\mathcal{M} \arrow[r, dashed, "\phi\times\mathrm{Id}"] \arrow[dr, "\gamma_2"] & \mathcal{C}_\mathcal{M} \arrow[d, "\pi_2"] \\
& D
\end{tikzcd}$$

Thus $\gamma_2^{-1}(\hat{u}) = (\phi \times \mathrm{Id})^{-1}(\pi_2^{-1}(\hat{u}))$. The size of $\pi_2^{-1}(\hat{u})$ is the algebraic degree of the optimization problem, which we multiply by $\deg(\phi \times \mathrm{Id}) = \delta$ to find the size of $\gamma_2^{-1}(\hat{u})$. $\square$

The parametric likelihood correspondence was studied in detail in [75]. We therefore focus on the affine and projective scenarios. Our next goal will be to describe the ideals of the parametric correspondence. It is not obvious how one writes down equations for $\mathcal{P}_\mathcal{M}$ from the definition above. In particular, the regularity of a point cannot be detected from a parametrization. In the projective case, we will make use of the $2 \times (s+1)$ matrix

$$\begin{pmatrix} h_u & \frac{\partial h_u}{\partial t_1} & \frac{\partial h_u}{\partial t_2} & \cdots & \frac{\partial h_u}{\partial t_s} \\ q & \frac{\partial q}{\partial t_1} & \frac{\partial q}{\partial t_2} & \cdots & \frac{\partial q}{\partial t_s} \end{pmatrix} \tag{2.16}$$

Indeed, it follows from the quotient rule that if $q(t) \neq 0$, then $\nabla g_u = 0$ precisely when (2.16) has rank 1.

Proposition 2.4.3. *The parametric critical correspondence* (2.15) *is equal to*

$$\overline{\{(t,u)\in\mathbb{C}^s\backslash(\mathcal{B}\cup\mathcal{V}_{\phi,s})\times\mathbb{C}^m : \nabla g_u(t)=0\}} \text{ in the affine scenario and}$$
$$\overline{\{(t,u)\in\mathbb{C}^s\backslash(\mathcal{B}\cup\mathcal{V}_{\phi,s}\cup V(q))\times\mathbb{P}^{m-1} : \operatorname{rank}(2.16)\leq 1\}} \text{ in the projective scenario.} \tag{2.17}$$

Proof. We will argue the projective case as the affine case follows from a similar argument. Observe that we have the containment

$$\begin{aligned}\mathcal{P}_{\mathcal{M}} &= \overline{\{(t,u)\in\phi^{-1}(\mathcal{M}_{\mathrm{reg}})\backslash(\mathcal{V}_{\phi,s}\cup V(q))\times\mathbb{P}^{m-1} : \phi(t) \text{ is critical for } u\}}\\ &\subseteq \overline{\{(t,u)\in\phi^{-1}(\mathcal{M})\backslash(\mathcal{V}_{\phi,s}\cup V(q))\times\mathbb{P}^{m-1} : \phi(t) \text{ is critical for } u\}}\\ &= \overline{\{(t,u)\in\mathbb{C}^s\backslash(\mathcal{B}\cup\mathcal{V}_{\phi,s}\cup V(q))\times\mathbb{P}^{m-1} : \operatorname{rank}(2.16)\leq 1\}}.\end{aligned}$$

Since $\mathcal{P}_{\mathcal{M}}$ is irreducible of dimension $m-1$ it suffices to prove that (2.17) has the same dimension. Indeed, (2.17) is also a vector bundle over $\mathcal{U}=\mathbb{C}^s\backslash(\mathcal{B}\cup\mathcal{V}_{\phi,s}\cup V(q))$. For a given $t\in\mathcal{U}$, consider the fiber $\gamma_1^{-1}(u)$ of the projection. This is the set of u such that (2.16) has rank 1. Note that its rank is at least 1 because $q(t)\neq 0$. By Euler's relation and the chain rule, we have

$$\begin{pmatrix} h_u(t) & \nabla h_u(t)\end{pmatrix} = \begin{pmatrix}\frac{\partial h_u}{\partial x_1}(\phi(t)) & \cdots & \frac{\partial h_u}{\partial x_n}(\phi(t))\end{pmatrix}\begin{pmatrix}\phi_1(t) & \frac{\partial\phi_1}{\partial t_1}(t) & \cdots & \frac{\partial\phi_1}{\partial t_s}(t)\\ \vdots & \vdots & \ddots & \vdots\\ \phi_n(t) & \frac{\partial\phi_n}{\partial t_1}(t) & \cdots & \frac{\partial\phi_n}{\partial t_s}(t)\end{pmatrix}. \tag{2.18}$$

Without loss of generality, we may assume that $\phi_1=1$ identically, since $\phi(t)$ can be replaced by $\phi(t)/\phi_1(t)$. Since $\mathcal{J}_\phi(t)$ has rank s, the matrix on the right hand side of the above equation has rank $s+1$. In particular, the set of preimages of the span of $\begin{pmatrix}q(t) & \nabla q(t)\end{pmatrix}$ under the map $\begin{pmatrix}\phi(t) & \mathcal{J}_\phi(t)\end{pmatrix}$ has dimension $n-s$. Since the map $u\mapsto\nabla_x h_u(\phi(t))$ is surjective by Assumption 2.2.1, every preimage of $\begin{pmatrix}q(t) & \nabla q(t)\end{pmatrix}$ under $\begin{pmatrix}\phi(t) & \mathcal{J}_\phi(t)\end{pmatrix}$ can be written as $\nabla_x h_u(\phi(t))$ for some u. The fibers of $u\mapsto\nabla_x h_u(\phi(t))$ have dimension $m-n$, so the dimension of the projective linear space of u for which t is critical is $m-n+n-s-1=m-s-1$. This holds for all t in the open set $\mathcal{U}$, so the vector bundle property follows. Furthermore, the dimension of (2.17) is $m-1$, and thus it is equal to $\mathcal{P}_{\mathcal{M}}$. □

We can now give descriptions of the prime ideals of the parametric correspondences. Let $I_{\phi,s}=I(\mathcal{V}_{\phi,s})\subseteq\mathbb{C}[t]$.

Theorem 2.4.4. *The prime ideal of* $\mathcal{P}_{\mathcal{M}}$ *is* $\langle\nabla g_u(t)\rangle : (I_{\phi,s}\cdot I_{\mathcal{B}})^\infty\subseteq\mathbb{C}[t,u]$ *in the affine case and* $\langle 2\times 2$ *minors of* (2.16)$\rangle : (I_{\phi,s}\cdot I_{\mathcal{B}}\cdot\langle q(t)\rangle)^\infty\subseteq\mathbb{C}[t,u]$ *in the projective case.*

Proof. One applies Proposition 2.4.3 along with the same argument as in the proof of Theorem 2.2.3 to show that the ideal is prime. □

Corollary 2.4.5. *The parametric critical ideals are given by fixing u to be a generic $u\in\mathbb{C}^m$ in the ideals in Theorem 2.4.4. These ideals are zero-dimensional in the affine case and one-dimensional otherwise. Their degrees are δ times the algebraic degree of* (2.1).

Example 2.4.6 (Cuspidal Cubic)**.** The cuspidal cubic admits the parametrization $t \mapsto (t^2, t^3)$. Parametrically, our objective function is $g_u(t) = (t^2 - u)^2 + (t^3 - v)^2$. The base locus of the parametrization is empty, and the Jacobian $\begin{pmatrix} 2t & 3t^2 \end{pmatrix}$ drops rank only at $t = 0$. The parametric critical correspondence is a hypersurface in $\mathbb{C} \times \mathbb{C}^2$ and its prime ideal is $\langle 2t(t^2 - u) + 3t^2(t^3 - v)\rangle : t^\infty$. Its defining polynomial is $2t^2 - 2u + 3t^4 - 3vt$. For general u, v, this polynomial is irreducible of degree 4, which matches the ED degree of 4 we found in Example 2.1.1. $\diamond$

We conclude with a formula for the number of critical points of a rational function. In particular, this can be used to compute the algebraic degree of (2.1) in the scenario where $\mathcal{M}$ is a projective variety, $g_u = h_u/q$ is a rational function, and the finite-to-one parametrization ϕ is inhomogeneous and the polynomials have general coefficients.

Theorem 2.4.7. *If $h(t_1, \ldots, t_s)$ and $q(t_1, \ldots, t_s)$ are general inhomogeneous polynomials of degrees d and e, respectively, then the rational function $g(t) = h(t)/q(t)$ has $\frac{d(d-1)^s - e(e-1)^s}{d-e}$ complex critical points if $d \neq e$, and it has $(s + 1)(d - 1)^s$ complex critical points if $d = e$.*

Proof. The critical ideal is generated by the 2×2 minors of the matrix (2.16). For $d \neq e$, the Giambelli-Thom-Porteous formula says that the degree of this ideal equals

$$\sum_{k=0}^{s} (-1)^{s-k} \binom{s}{k} h_k(d, e). \tag{2.19}$$

Here h_k is the homogeneous symmetric function of degree k. This simplifies to $\frac{d(d-1)^s - e(e-1)^s}{d-e}$.

When $d = e$, we need a correction because the matrix obtained from (2.16) by taking the entries' leading forms drops rank at some points on the hyperplane at infinity in $\mathbb{P}^{s-1}$. This follows from Euler's relation, which states that $(-d, t_1, t_2, \ldots, t_s)^T$ is in the kernel of that matrix of homogeneous polynomials. We apply Giambelli-Thom-Porteous to the last s columns of that matrix to see that the number of these solutions at infinity is $s(d-1)^{s-1}$. We subtract $s(d-1)^{s-1}$ from (2.19) evaluated at $e = d$. The result simplifies to $(s+1)(d-1)^s$. $\square$

Example 2.4.8 (Circle)**.** The circle $x^2 + y^2 - z^2$ admits the parametrization $t \mapsto [t^2 - 1 : 2t : t^2+1]$. The base locus is empty, and the Jacobian $\begin{pmatrix} 2t & 2 & 2t \end{pmatrix}$ has full rank for all $t \in \mathbb{C}$. The objective function is given by $h_u(t) = \begin{pmatrix} t^2 - 1 & 2t & t^2 + 1 \end{pmatrix} U \begin{pmatrix} t^2 - 1 & 2t & t^2 + 1 \end{pmatrix}^T$ and denominator $(t^2+1)^2$. Since numerator and denominator both have degree 4, Theorem 2.4.7 gives the bound of $(2)(4-1)^1 = 6$. The parametric critical correspondence of the constrained eigenvalue problem is a hypersurface in $\mathbb{C} \times \mathbb{P}^5$. The determinant of the 2×2 matrix (2.16) is

$$((u_{12} + u_{23})t^4 - 2(u_{11} + u_{13} - u_{22})t^3 - 6t^2 u_{12} + 2(u_{11} - u_{13} - u_{22})t + u_{12} - u_{23})(t^2 + 1)$$

After saturating by $t^2 + 1$, we are left with the degree four factor, which corresponds to the algebraic degree we saw in Example 2.1.3. Note that the extraneous points $i \mapsto [1 : -i : 0]$ and $-i \mapsto [1 : i : 0]$ are precisely the isotropic points we saw in Example 2.3.4. $\diamond$

2.5 Numerical Computation

We conclude this chapter with a few words on the computations in this thesis. The results presented in this chapter are most useful for symbolic computations with a computer algebra system. While these symbolic computations are very important, one can go much further with numerical computations. We begin with a brief overview of the numerical homotopy continuation, and then explain how one can use Corollaries 2.2.4 and 2.4.5 for numerical computations.

Our main computational tool is *numerical homotopy continuation*, see [119]. This is a numerical method for finding all isolated solutions to a system of polynomial equations. Concretely, let $x = (x_1, \dots, x_n)$ be variables and $u = (u_1, \dots, u_m)$ be parameters, and consider a family of polynomial systems

$$\mathcal{F}(x; u) = \big(f_1(x; u), \dots, f_n(x; u)\big) \in \mathbb{C}[x]^n, \quad \text{with } \ u \in \mathbb{C}^m.$$

Choose two parameter values $u, v \in \mathbb{C}^m$ such that the target system is $F(x) = \mathcal{F}(x; u)$, and the start system $G(x) = \mathcal{F}(x; v)$ has easily computable solutions. The *parameter homotopy*

$$H(x, t) := \mathcal{F}\big(x; (1-t)u + tv\big), \quad t \in [0, 1],$$

deforms $G(x)$ at $t = 1$ to $F(x)$ at $t = 0$. The parameter continuation theorem [15, 98] guarantees that the number of regular isolated solutions in $\mathbb{C}^n$ is constant along $t \in (0, 1]$. Starting from solutions of $G(x)$, we track solution paths from $t = 1$ to $t = 0$. If the start system $G(x)$ has at least as many regular isolated solutions as the target $F(x)$ and all solution paths arrive at $t = 0$, then we obtain all solutions of $F(x)$. When $G(x)$ has exactly as many solutions as $F(x)$, the homotopy is called *optimal*.

Another numerical method for solving polynomial systems is *monodromy*, introduced in [40]. Starting from a general parameter value $\hat{u}$ and a set of known solutions to $\mathcal{F}(x; \hat{u}) = 0$, one traces each solution along closed loops in the parameter space based at $\hat{u}$ using parameter homotopy. These loops induce permutations of the solution set, and running random loops often uncovers previously unknown solutions. When the total number of solutions to a general system in the family is known, monodromy has a stopping criterion. This is one reason that it is helpful to know the algebraic degree of a problem a priori. Otherwise, suitable heuristics may be applied. The monodromy method is powerful because it does not require a start system, just a single start solution. Consequently, it may be used to solve systems whose equations are not all polynomials. In particular, it can be used to solve systems of rational functions. Both methods are implemented in the `Julia` [8] package `HomotopyContinuation.jl` [20].

We now discuss how to use Corollaries 2.2.4 and 2.4.5 to solve numerical systems in `HomotopyContinuation.jl`. When one has a δ-to-1 parametrization of the model $\mathcal{M}$ for small δ, it is often more computationally efficient to solve the problem parametrically. If Theorem 2.4.2 applies, namely the incidence correspondence is irreducible, then monodromy is typically the fastest method. For instance, one solves Example 2.1.4 with the following code:

```
using HomotopyContinuation
@var t, u[1:3]
gu = u[1]*log(t^2) + u[2]*log((1-t)^2) + u[3]*log(2*t*(1-t))
                         - sum(u[1:3])*log(t^2 + (1-t)^2 + 2*t*(1-t))
F = System(differentiate(gu, [t]), parameters=u)
mon_res = monodromy_solve(F)
```

If one wants to solve the problem for specific data point `data`, one can then run

```
S = solve(F, solutions(mon_res); start_parameters=parameters(mon_res),
                                 target_parameters=data)
solutions(S)
```

We have observed that in the case of optimizing a rational function numerically, this method is superior to the determinantal description in Corollary 2.4.5.

We now consider the implicit case. Rather than using the determinantal equations in Corollary 2.2.4, it is preferable to use Lagrange multipliers. To obtain a finite dimensional solution set, one imposes $r - c$ linear equations on the Lagrange multipliers. In the affine case, the system has $n + r$ variables $x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_r$ and $n + 2r - c$ equations:

$$\begin{cases} \nabla(g_u + \sum_{i=1}^r \lambda_i f_i) = 0, \\ M \cdot \lambda = v. \end{cases}$$

Here M is a random $(r - c) \times r$ matrix and v is a random length r vector. When $\mathcal{M}$ is a complete intersection, then one does not need additional linear equations. One sets up the system in Example 2.4.6 as follows, and solves as above:

```
using HomotopyContinuation
@var x[1:2], u[1:2], a
gu = sum((x - u).^2)
f = x[1]^3-x[2]^2
L = gu - a*f
F = System(differentiate(L, [x;a]), parameters=u)
```

In this chapter, we introduced the algebraic degree of an optimization problem and proved that it is finite for a large class of optimization problems. We did this via the critical correspondence, which is the incidence variety of data and critical points. One can compute this critical correspondence, as well the critical ideal, using the descriptions of prime ideals we gave. We gave parametric analogs of the above for rational models, and concluded with methods for solving these systems numerically. We made this concrete for three scenarios: optimization of a polynomial on a affine variety, optimization of a degree 0 rational function on a projective variety, and maximum likelihood estimation of a projective variety. We will study each of these scenarios in detail in the following chapters.

Chapter 3

The Grassmannian and Friends

Linear subspaces of dimension k in the vector space $\mathbb{R}^n$ correspond to points in the Grassmannian $\mathrm{Gr}(k, n)$, which is a real manifold of dimension $k(n-k)$. This manifold is also an algebraic variety. This means that it can be described by polynomial equations in finitely many variables. Such a description is useful for working with linear spaces in applications, such as optimization on manifolds in data science [56] and scattering amplitudes in physics [129].

One way to represent a k-dimensional vector space in $\mathbb{R}^n$ is by choosing a basis. This basis can be represented by a matrix $A \in \mathbb{R}^{n\times k}$ with linearly independent columns. In fact, using Gram-Schmidt, we may choose them to be orthonormal. This brings us to our first "friend" of the Grassmannian, namely the set of orthogonal k-frames, or the *real Stiefel manifold* $\mathrm{St}(k, n)_{\mathbb{R}} = \{Z \in \mathbb{R}^{n\times k} \ : \ ZZ^T = \mathrm{Id}_k\}$. However, the Stiefel manifold alone is not a good model for the Grassmannian, as a given vector space has many choices of orthonormal bases. The *real orthogonal group* $\mathrm{O}(k)_{\mathbb{R}} = \{Z \in \mathbb{R}^{k\times k} \ : \ ZZ^T = \mathrm{Id}_k\}$ acts transitively on these bases. We therefore identify the Grassmannian with the quotient space $\mathrm{St}(k, n)_{\mathbb{R}}/\mathrm{O}(k)_{\mathbb{R}}$.

The previous discussion generalizes to *flag manifolds*. A flag manifold is determined by natural numbers $0 < k_1 < \cdots < k_r \leq n$. Each point in the flag manifold represents a chain of nested vectors subspaces of $\mathbb{R}^n$ whose dimensions are given by the vector $\mathbf{k} = (k_1, \ldots, k_r)$:

$$\mathrm{Fl}(\mathbf{k}; n) = \mathrm{Fl}(k_1, \ldots, k_r; n) = \{W_1 \subset W_2 \subset \cdots \subset W_r : W_i \subseteq \mathbb{R}^n,\ \dim(W_i) = k_i,\ 1 \leq i \leq r\}.$$

We observe that the partial flags $0 = W_0 \subset W_1 \subset W_2 \subset \cdots \subset W_r \subset \mathbb{R}^n$ with $\dim(W_i) = k_i$ are in bijection with the sequence of subspaces $(W_1/W_0, W_2/W_1, W_3/W_2, \ldots, W_r/W_{r-1})$ where W_i/W_{i-1} is viewed as a subspace in $\mathbb{R}^n/W_{i-1}$ for $i = 1, \ldots, r$. This gives a formula for the dimension of $\mathrm{Fl}(\mathbf{k}; n) = \mathrm{Fl}(k_1, \ldots, k_r; n)$.

Proposition 3.0.1. *The dimension of* $\mathrm{Fl}(\mathbf{k}; n)$ *equals* $\sum_{i=1}^{r}(k_i - k_{i-1})(n - k_i)$. *In particular, the dimension of the complete flag variety* $\mathrm{Fl}(1, 2, \ldots, n-1; n)$ *is* $n(n-1)/2$.

Proof. The above bijection implies that there is a bijection from $\mathrm{Fl}(\mathbf{k}; n)$ to the points in

$$\mathrm{Gr}(n, k_1) \times \mathrm{Gr}(k_2 - k_1, n - k_1) \times \cdots \times \mathrm{Gr}(k_r - k_{r-1}, n - k_{r-1}).$$

Since the ith Grassmannian above has dimension $(k_i - k_{i-1})(n - k_i)$ the formula follows. For the complete flag variety, $k_i - k_{i-1} = 1$ for all $i = 1, \ldots, r$, and again the formula follows. $\square$

To give $\mathrm{Fl}(\mathbf{k}, n)$ a manifold structure, we associate to a flag $W_1 \subset W_2 \subset \cdots \subset W_r \subset \mathbb{R}^n$ with $\dim(W_i) = k_i$ an orthonormal k_r-frame Z together with the orthogonal groups $\mathrm{O}(k_i - k_{i-1})$ acting on the columns $Z_{k_{i-1}+1}, \ldots, Z_{k_i}$ by right multiplication. In other words,

$$\mathrm{Fl}(k_1, \ldots, k_r; n) \simeq \mathrm{St}(k_r, n)_{\mathbb{R}} \,/\, (\mathrm{O}(k_1)_{\mathbb{R}} \times \mathrm{O}(k_2 - k_1)_{\mathbb{R}} \times \cdots \times \mathrm{O}(k_r - k_{r-1})_{\mathbb{R}}).$$

Equivalently, a flag manifold may be realized as the quotient of an orthogonal group. We can associate to a flag $W_1 \subset W_2 \subset \cdots \subset W_r \subset V$ an orthogonal matrix Q whose first k_i columns comprise an orthonormal basis of W_i for $i = 1, \ldots, r$. The smaller orthogonal group $\mathrm{O}(k_i - k_{i-1})$ acts on the columns $Q_{k_{i-1}+1}, \ldots, Q_{k_i}$ by right multiplication without altering the flag. Therefore

$$\mathrm{Fl}(k_1, \ldots, k_r; n) \simeq \mathrm{O}(n)_{\mathbb{R}} \,/\, (\mathrm{O}(k_1)_{\mathbb{R}} \times \mathrm{O}(k_2 - k_1)_{\mathbb{R}} \times \cdots \times \mathrm{O}(n - k_r)_{\mathbb{R}}).$$

In particular, the Grassmannian is realized as $\mathrm{O}(n)_{\mathbb{R}}/\mathrm{O}(k)_{\mathbb{R}} \times \mathrm{O}(n-k)_{\mathbb{R}}$.

The above initial introduction views $\mathrm{Fl}(\mathbf{k}; n)$ and $\mathrm{Gr}(k, n)$ primarily as manifolds. As algebraic varieties, they have different embeddings. We collect answers to fundamental questions about these varieties, beginning with their dimensions. After the dimension, the second most important invariant of an embedded variety is its degree. This is the number of complex intersection points with a generic linear subspace of complementary dimension. Finally, we give descriptions of defining equations. We remark that so far we have defined our varieties over the real numbers $\mathbb{R}$. This is because data is usually real, not complex. However, to describe the algebraic properties of these varieties, we will study their complexifications.

Section 3.1 focuses on the Stiefel manifold and orthogonal group. In Section 3.2, we turn to the Plücker embedding of Grassmann and flag varieties. We introduce the projection Grassmannian, a realization of the Grassmannian as an affine variety, in Section 3.3. In Sections 3.4 and 3.5, we give two realizations of flag varieties as affine varieties that generalize the projection Grassmannian. Finally, in Section 3.6, we explain how to move between different realizations of Grassmann and flag varieties.

3.1 Stiefel Manifold and the Orthogonal Group

In applications, the Stiefel manifold is used as a model for the Grassmannian by representing points by orthonormal bases.

Proposition 3.1.1 (Stiefel Manifold). *The Stiefel manifold* $\mathrm{St}(k, n)$ *is the set of orthonormal* k*-frames* $\mathrm{St}(k, n) = \{Z \in \mathbb{C}^{n \times k} : Z^T Z = \mathrm{Id}_k\}$. *For* $k < n$, *it is the* $\mathrm{SO}(n)$*-orbit of the* $n \times k$ *matrix* $(e_1 \cdots e_k)$. *In symbols,* $\mathrm{St}(k, n) = \{R \cdot (e_1 \cdots e_k) : R \in \mathrm{SO}(n)\}$. *In particular,* $\mathrm{St}(k, n)$ *is irreducible when* $k < n$.

The Stiefel manifold $\mathrm{St}(k,n)$ is an affine algebraic variety in $\mathbb{C}^{n\times k}$. The orthogonal group $\mathrm{St}(n,n) = \mathrm{O}(n)$ has two irreducible components, including the special orthogonal group $\mathrm{SO}(n) = \{R \in \mathrm{St}(n,n) : \det(R) = 1\}$.

We now summarize the properties of $\mathrm{St}(k,n)$ as a variety. Some of the contents of the following theorem are known, for example in [25, Lemma 2.4].

Theorem 3.1.2. *The Stiefel manifold* $\mathrm{St}(k,n)$ *has dimension*

$$\dim(\mathrm{St}(k,n)) = nk - \binom{k+1}{2} = \binom{n}{2} - \binom{n-k}{2}.$$

The Stiefel ideal $I_{\mathrm{St}(k,n)} = I(\mathrm{St}(k,n)) = \langle Z^T Z - \mathrm{Id}_k\rangle$ *is a radical complete intersection. The variety* $\mathrm{St}(k,n)$ *is smooth. When* $k < n$*, then* $I_{\mathrm{St}(k,n)}$ *is prime.*

Proof. For the dimension, we observe that $\mathrm{Fl}(1,2,\ldots,k;n) \simeq \mathrm{St}(k,n) / \prod_{i=1}^k \mathrm{O}(1)$, since each orthonormal frame Z corresponds to a flag $W_1 \subset W_2 \subset \cdots \subset W_k$ with $\dim(W_i) = i$ up to the sign of each column of Z. This gives a finite map from $\mathrm{St}(k,n)$ to $\mathrm{Fl}(1,2,\ldots,k;n)$ of degree 2^k. Hence, the dimension of $\mathrm{St}(k,n)$ is equal to the dimension of this flag manifold which is $\binom{n}{2} - \binom{n-k}{2} = nk - \binom{k+1}{2}$ by Proposition 3.0.1. The affine variety defined by $I_{\mathrm{St}(k,n)}$ is precisely $\mathrm{St}(k,n)$. Since $I_{\mathrm{St}(k,n)}$ has $\binom{k+1}{2}$ generators, it is a complete intersection.

The claim that $I_{\mathrm{St}(k,n)}$ is prime will follow from irreducibility of $\mathrm{St}(k,n)$ (Proposition 3.1.1) and radicality of $I_{\mathrm{St}(k,n)}$. To prove smoothness and radicality, we apply Lemma 3.1.3. There is a transitive $\mathrm{O}(n)$ action on the closed points of the affine scheme defined by $I_{\mathrm{St}(k,n)}$. It therefore suffices to prove smoothness at a single closed point, namely $(e_1 \cdots e_k)$. We use the Jacobian criterion [57, Lemma 6.27]. The Jacobian of $I_{\mathrm{St}(k,n)}$ is the $kn \times \binom{k+1}{2}$ matrix

$$\mathcal{J}_{\mathrm{St}(k,n)}(Z)^T = \begin{pmatrix} 2Z_1 & & & & & Z_2 & Z_3 & \cdots & Z_k & & & & \cdots & \\ & 2Z_2 & & & & Z_1 & & & & Z_3 & \cdots & Z_k & \cdots & \\ & & 2Z_3 & & & & Z_1 & & & Z_2 & & & \cdots & \\ & & & \ddots & & & & \ddots & & & \ddots & & & Z_k \\ & & & & 2Z_k & & & & Z_1 & & & Z_2 & \cdots & Z_{k-1} \end{pmatrix}, \tag{3.1}$$

where Z_i denotes the ith column of Z for $i = 1,\ldots,k$. The columns of $\mathcal{J}_{\mathrm{St}(k,n)}((e_1 \cdots e_k))^T$ have disjoint supports, and hence $\mathcal{J}_{\mathrm{St}(k,n)}((e_1 \cdots e_k))^T$ has rank $\binom{k+1}{2}$. □

Lemma 3.1.3. *Let* $I \subseteq \mathbb{C}[x] = \mathbb{C}[x_1,\ldots,x_n]$ *be an ideal. Suppose that an orthogonal group* $\mathrm{O}(m)$ *or* $\mathrm{SO}(m)$ *acts transitively on the closed points of the scheme* $X = \mathrm{Spec}(\mathbb{C}[x]/I) \subseteq \mathbb{A}^n$*. If* X *has a single smooth closed point, then* $\mathrm{Spec}(\mathbb{C}[x]/I)$ *is smooth and* I *is radical.*

Proof. Every $R \in \mathrm{O}(m)$ induces an isomorphism on X that sends smooth points to smooth points. Since the action is transitive, every closed point must be smooth. Smooth schemes are reduced [57, Theorem 6.28, Proposition B.77] and affine schemes are reduced if and only if their defining ideals are radical [57, Proposition 3.27]. □

A combinatorial rule for the degree of the Stiefel manifold was derived by Brysiewicz and Gesmundo [25]. We state a simplified version for small k.

Theorem 3.1.4 ([25, Theorem 1.1])**.** *If* $n \geq 2k-1$*, then* $\deg(\mathrm{St}(k,n)) = 2^{\binom{k+1}{2}}$.

The degree of the special orthogonal group was computed by Brandt, Bruce, Brysiewicz, Krone and Robeva.

Theorem 3.1.5 ([18, Theorem 1.1])**.** *The degree of* $\mathrm{SO}(n)$ *is*

$$2^{n-1}\det\left(\left(\binom{2n-2i-2j}{n-2i}\right)\right)_{1\leq i,j\leq \lfloor n/2\rfloor}.$$

Both the Stiefel manifold and the orthogonal group can serve as models for the Grassmannian. Indeed, the Grassmannian may be realized as $\mathrm{St}(k,n)/\mathrm{O}(k)$ or as $\mathrm{O}(n)/(\mathrm{O}(k)\times\mathrm{O}(n-k))$. One disadvantage, computationally speaking, is that one has to work with equivalent classes of matrices in either of these representations. In this next section, we explore two realizations of Grassmannians which do not require a choice of basis or an equivalence class.

3.2 Plücker Coordinates

In classical algebraic geometry, Grassmannians and flag varieties are embedded in projective space via the Plücker embedding. We begin with the case of the Grassmannian. A k-dimensional subspace W of $\mathbb{R}^n$ is identified with an $n \times k$ matrix A of rank k whose columns form a basis of W. This identification is up to the action of the general linear group $\mathrm{GL}(k) = \{M \in \mathbb{C}^{k\times k} : \det(M) \neq 0\}$ by right multiplication to account for all possible bases of W. Then we consider the map

$$\mathrm{Gr}(k,n) \longrightarrow \mathbb{P}^{\binom{n}{k}-1} \qquad A \mapsto \det(A_I)_{I\in\binom{[n]}{k}}$$

where A_I is the $k \times k$ submatrix of A whose rows are indexed by I. The image is an irreducible projective variety of dimension $k(n-k)$ defined by the quadratic *Plücker relations* [52, Section 9.1] in the coordinates x_I of $\mathbb{P}^{\binom{n}{k}-1}$. In similar fashion, we embed $\mathrm{Fl}(k_1,\ldots,k_r;n)$ into $\mathbb{P}^{\binom{n}{k_1}-1}\times\cdots\times\mathbb{P}^{\binom{n}{k_r}-1}$. This time we consider the map

$$\mathrm{Fl}(\mathbf{k};n) \to \mathbb{P}^{\binom{n}{k_1}-1}\times\cdots\times\mathbb{P}^{\binom{n}{k_r}-1}, \quad A \mapsto \left(\det(A_{I_1})_{I_1\in\binom{[n]}{k_1}}, \cdots, \det(A_{I_r})_{I_r\in\binom{[n]}{k_r}}\right) \tag{3.2}$$

where A is a $n \times k_r$ matrix of rank k_r and A_{I_j} is the $k_j \times k_j$ submatrix of the first k_j columns of A with rows indexed by I_j. The image is an irreducible projective variety of dimension $\sum_{i=1}^{r}(k_i - k_{i-1})(n-k_i)$, which is also defined by quadratic relations in the coordinates x_I where I is a subset of n of cardinality equal to one of $k_1, k_2, \ldots, k_r$. The equations below express the fact that $W_t \subseteq W_s$ for every pair of subspaces in the flag whenever $t \leq s$.

Theorem 3.2.1 ([52, Proposition 9.1.1])**.** *The variety* $\mathrm{Fl}(k_1, \ldots, k_r; n) \subseteq \mathbb{P}^{\binom{n}{k_1}-1} \times \cdots \times \mathbb{P}^{\binom{n}{k_r}-1}$ *in Plücker coordinates is defined by the prime ideal generated by the quadrics*

$$x_{i_1,\ldots,i_{k_s}} x_{j_1,\ldots,j_{k_t}} - \sum x_{i'_1,\ldots,i'_{k_s}} x_{j'_1,\ldots,j'_{k_t}}$$

for every pair of $1 \leq t \leq s \leq r$ *and* $m \in [t]$*. Here, the sum is over all* (i', j') *obtained by exchanging the first* m *of the* j*-subscripts with* m *of the* i*-subscripts while preserving their order and given a permutation* σ *on* s_k *elements,* $x_{i_1,\ldots,i_{k_s}} = \mathrm{sgn}(\sigma) x_{\sigma(i_1),\ldots,\sigma(i_{k_s})}$.

We now turn our focus to the Grassmannian. Every point in $\mathrm{Gr}(k, n)$ can be encoded by a skew-symmetric tensor X. The $\binom{n}{k}$ entries of X are denoted $x_{i_1 i_2 \ldots i_k}$ where $1 \leq i_1, i_2, \ldots, i_k \leq n$. Using the sign flips coming from skew-symmetry, we obtain linearly independent coordinates x_I where $I = (i_1, i_2, \ldots, i_k)$ satisfies $i_1 < i_2 < \cdots < i_k$. When $k = 2$, then X is a skew-symmetric matrix, and $\mathrm{Gr}(2, n)$ is realized as the locus of rank 2 skew-symmetric matrices.

Proposition 3.2.2. *The* 4×4 *Pfaffians of* X*, namely the polynomials*

$$\underline{x_{ij} x_{\ell m}} - x_{i\ell} x_{jm} - x_{im} x_{j\ell} \qquad \text{for} 1 \leq i < \ell < j < m \leq n$$

form a reduced Gröbner basis for the prime ideal of $\mathrm{Gr}(2, n)$*, for any monomial order that selects the underlined leading term, e.g., the lexicographic order with* $\mathrm{weight}(x_{ij}) = j - i$.

Proof. The proof is identical to that of [97, Theorem 5.8]. ☐

Example 3.2.3 $(n = 5, k = 2)$**.** We consider the span of a 5×2 matrix that has rank 2:

$$A \;=\; \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \end{bmatrix}^T.$$

This image is a point in the 6-dimensional manifold $\mathrm{Gr}(2, 5)$. The embedding into $\mathbb{P}^9_{\mathbb{R}}$ is given by the ten 2×2 minors $x_{ij} = a_{1i}a_{2j} - a_{1j}a_{2i}$. These satisfy the Plücker equations

$$\begin{array}{rcccccc} & x_{12}x_{34} - x_{13}x_{24} + x_{14}x_{23} & = & x_{12}x_{35} - x_{13}x_{25} + x_{15}x_{23} & = & x_{12}x_{45} - x_{14}x_{25} + x_{15}x_{24} \\ = & x_{13}x_{45} - x_{14}x_{35} + x_{15}x_{34} & = & x_{23}x_{45} - x_{24}x_{35} + x_{25}x_{34} & = & 0. \end{array}$$

We note that these five quadrics are the 4×4 Pfaffians in the skew-symmetric 5×5 matrix

$$X \;=\; \begin{bmatrix} 0 & x_{12} & x_{13} & x_{14} & x_{15} \\ -x_{12} & 0 & x_{23} & x_{24} & x_{25} \\ -x_{13} & -x_{23} & 0 & x_{34} & x_{35} \\ -x_{14} & -x_{24} & -x_{34} & 0 & x_{45} \\ -x_{15} & -x_{25} & -x_{35} & -x_{45} & 0 \end{bmatrix}. \qquad (3.3)$$

Hence, $\mathrm{Gr}(2, 5)$ is the projective variety in $\mathbb{P}^9_{\mathbb{R}}$ whose points are the above matrices X of rank 2, up to scaling. The subspace represented by X is the row space (or column space) of X. This embedding realizes $\mathrm{Gr}(2, 5)$ as a 6-dimensional variety of degree 5 in $\mathbb{P}^9_{\mathbb{R}}$. ◇

Theorem 3.2.4 ([97, Theorem 5.12]). *The degree of the Grassmannian* $\mathrm{Gr}(k,n)$ *in its Plücker embedding is the number of standard Young tableaux of rectangular shape* $k\times(n-k)$*.*

These degrees are listed for small values of n and k in following table.

$n \backslash k$	0	1	2	3	4	5	6	7	8	9	10
3	1	1	1	1							
4	1	1	2	1	1						
5	1	1	5	5	1	1					
6	1	1	14	42	14	1	1				
7	1	1	42	462	462	42	1	1			
8	1	1	132	6 006	24 024	6 006	132	1	1		
9	1	1	429	87 516	1 662 804	1 662 804	87 516	429	1	1	
10	1	1	1 430	1 385 670	140 229 804	701 149 020	140 229 804	1 385 670	1 430	1	1

3.3 The Projection Grassmannian

A common approach in applications is to represent a linear subspace of $\mathbb{R}^n$ by the orthogonal projection matrix onto it [30, 56, 76, 83, 87]. This realizes $\mathrm{Gr}(k,n)$ as an affine variety. Its ambient space is the space $\mathbb{R}^{\binom{n+1}{2}}$ of symmetric $n\times n$ matrices P. In that embedding, $\mathrm{Gr}(k,n)$ is described by the quadratic equations $P^2=P$ and the linear equation $\mathrm{trace}(P)=k$. Given a full rank matrix $A\in\mathbb{R}^{n\times k}$, the orthogonal projection onto its column span is given by

$$P = A(A^TA)^{-1}A^T.$$

Note that the above works only for projections onto real vector spaces. An orthogonal projection onto a complex vector space should instead be Hermitian. The next result gives a set-theoretic description of a complexification of the projection Grassmannian, along with a description of the projection Grassmannian as an orbit. While the real points of this variety are identified with the real Grassmannian, the variety cannot be identified with the complex Grassmannian. Write E_k for the diagonal matrix $\mathrm{diag}(1,\dots,1,0,\dots,0)$ with k ones.

Proposition 3.3.1 (Projection Grassmannian). *The projection Grassmannian*

$$\mathrm{pGr}(k,n) = \{P\in\mathrm{Sym}^2\mathbb{C}^n : P^2=P,\ \mathrm{trace}(P)=k\}$$

is an irreducible affine variety of dimension $k(n-k)$*. It is equal to an orthogonal group orbit:*

$$\mathrm{pGr}(k,n) = \{R^TE_kR : R\in\mathrm{SO}(n)\}.$$

Proof. For the set-theoretic description, we observe that an orthogonal projection matrix P onto a real k-dimensional vector space must satisfy $P=P^T$, $P^2=P$, and $\mathrm{rank}(P)=k$. Since $P^2=P$ and $P\neq 0,\mathrm{Id}_n$, the minimal polynomial of P is z^2-z, so the eigenvalues of P are in $\{0,1\}$ and hence $\mathrm{rank}(P)=\mathrm{trace}(P)$.

The points P in the $\mathrm{O}(n)$ orbit satisfy the relations $P^2 - P$ and $\mathrm{trace}(P) = k$. Conversely, every $P \in \mathrm{pGr}(k, n)$ is diagonalizable to E_k, since its minimal polynomial is $z^2 - z$. By [53, Theorem XI.4], P can be orthogonally diagonalizable to E_k. Irreducibility follows from this description. The dimension follows from the injectivity of the map

$$\mathrm{SO}(n)/(\mathrm{SO}(k) \times \mathrm{SO}(n-k)) \to \mathrm{pGr}(k, n), \qquad R \mapsto R^T E_k R. \qquad \square$$

With the aim of understanding the prime ideal of $\mathrm{pGr}(k, n)$, we now examine the affine variety in $\mathbb{R}^{\binom{n+1}{2}}$ that is defined by the entries of $P^2 - P$. Let I_n denote the ideal generated by these $\binom{n+1}{2}$ quadratic polynomials in the $\binom{n+1}{2}$ entries p_{ij} of the symmetric matrix P. By fixing the trace of P, we obtain the projection Grassmannian. We now prove that the given equations generate prime ideals in the polynomial ring $\mathbb{C}[P]$ whose unknowns are the $\binom{n+1}{2}$ entries of $P = (p_{ij})$. In particular, $I_n + \langle \mathrm{trace}(P) - k \rangle$ is the ideal of $\mathrm{pGr}(k, n)$.

Theorem 3.3.2. *The ideal I_n is radical, and $I_{n,k} \coloneqq I_n + \langle\, \mathrm{trace}(P) - k \,\rangle$ is prime for all k. The projection Grassmannian $\mathrm{pGr}(k, n)$ is smooth.*

Proof. The ideal I_n is generated by the following quadratic polynomials in the ring $\mathbb{C}[P]$:

$$f_{ij} \;\coloneqq\; p_{ij} - p_{i1}p_{1j} - p_{i2}p_{2j} - \cdots - p_{1n}p_{nj} \quad \text{for } 1 \leq i \leq j \leq n. \tag{3.4}$$

Let X_n denote the subscheme of the affine space $\mathrm{Spec}(\mathbb{C}[P]) = \mathbb{A}_{\mathbb{C}}^{\binom{n+1}{2}}$ defined by I_n. Moreover, let $X_{n,k}$ denote the subscheme of X_n defined by the ideal $I_{n,k}$.

By Proposition 3.3.1, the orthogonal group $\mathrm{O}(n)$ acts transitively on closed points of $X_{k,n}$, namely $\mathrm{pGr}(k, n)$. Thus by Lemma 3.1.3, the scheme $X_{n,k}$ is smooth and $I_{n,k}$ is radical if $X_{n,k}$ is smooth at the closed point E_k. Since $X_n = \sqcup_{k=0}^n X_{n,k}$ is the disjoint union of the $X_{n,k}$, this will imply that X_n is also smooth and that I_n is also radical. Finally, since $X_{n,k}$ is irreducible by Proposition 3.3.1, we conclude that $I_{n,k}$ is a prime ideal in $\mathbb{C}[P]$.

All that remains to be shown is that the affine scheme $X_{n,k}$ is smooth at the closed point E_k. The scheme $X_{n,k}$ has dimension $k(n-k)$ by Proposition 3.3.1. Hence, using the Jacobian criterion, it suffices to show that the gradients ∇f_{ij}, evaluated at the point E_k, span a space of dimension at least $\binom{n+1}{2} - k(n-k)$.

For all pairs $i, j \in \{1, \ldots, k\}$ and $i, j \in \{k+1, \ldots, n\}$, we have

$$\left.\frac{\partial f_{ij}}{\partial p_{i'j'}}\right|_{E_{\mathbf{k}}} = \begin{cases} \pm 1 & i' = i, j' = j, \\ 0 & \text{else} \end{cases}$$

where the sign is positive if $i, j \leq k$ and negative if $k < i, j$. There are $\binom{n+1}{2} - k(n-k)$ choices for such pairs $\{i, j\}$ and the vectors ∇f_{ij} have disjoint support and are therefore linearly independent. This shows that $X_{n,k}$ is smooth at E_k. $\square$

Corollary 3.3.3. *The radical ideal I_n has $n+1$ minimal primes, one for each Grassmannian:*

$$I_n \;=\; \bigcap_{k=0}^{n} I_{n,k}. \tag{3.5}$$

Proof. From the previous proof we have the following disjoint decomposition of varieties:

$$X_n \;=\; V(I_n) \;=\; \bigcup_{k=0}^{n} X_{n,k} \;=\; \bigcup_{k=0}^{n} V\left(I_{n,k}\right).$$

We apply the ideal operator on both sides. This turns the union into an intersection of ideals. Since all ideals are radical, by Theorem 3.3.2, Hilbert's Nullstellensatz implies (3.5). ☐

The Grassmannian $\mathrm{Gr}(k,n)$ is isomorphic to the Grassmannian $\mathrm{Gr}(n-k,n)$. We can see this isomorphism very nicely for the projection Grassmannians $\mathrm{pGr}(k,n)$ and $\mathrm{pGr}(n-k,n)$:

Remark 3.3.4. The linear involution $P \mapsto \mathrm{Id}_n - P$ fixes the radical ideal I_n, and it switches the prime ideals $I_{n,k}$ and $I_{n,(n-k)}$ that define $\mathrm{pGr}(k,n)$ and $\mathrm{pGr}(n-k,n)$.

There is one more linear coordinate change in matrix space which we wish to point out.

Remark 3.3.5. The projection Grassmannian $\mathrm{pGr}(k,n)$ is linearly isomorphic to the variety of orthogonal symmetric matrices of trace $2k-n$. Indeed, if we replace the matrix P with $Q = 2P - \mathrm{Id}_n$, then the equation $P^2 = P$ becomes $Q^2 = \mathrm{Id}_n$, so Q is a symmetric matrix that is orthogonal. Lai, Lim and Ye [83] proposed this embedding of $\mathrm{pGr}(k,n)$ into the orthogonal group $\mathrm{O}(n)$ in order to improve numerical stability in Grassmannian optimization [87].

Lim and Ye derived a formula for the degree of the projection Grassmannian in terms of Jack symmetric functions [88, Theorem 4.3]. We state here closed forms for $k = 0,1,2,3,4$.

Proposition 3.3.6. *The projection Grassmannians* $\mathrm{pGr}(k,n)$ *and* $\mathrm{pGr}(n-k,n)$ *have the same degree. For* $n \geq 3$*, we have*

$$\deg(\mathrm{pGr}(0,n)) = 1 \qquad \deg(\mathrm{pGr}(1,n)) = 2^{n-1} \qquad \deg(\mathrm{pGr}(2,n)) = 2\binom{2n-4}{n-2}.$$

For $n \geq 5$*, we have*

$$\deg(\mathrm{pGr}(3,n)) = \frac{(8n-25)(2n-9)!!}{(n-2)!}2^{2n-6}.$$

For $n \geq 7$*, we have*

$$\deg(\mathrm{pGr}(4,n)) = \frac{(32n^2-288n+634)(2n-13)!!(2n-9)!!}{(n-2)!(n-4)!}2^{2n-6}.$$

Proof. The first sentence follows from Remark 3.3.4. The $k = 0$ case holds because $\mathrm{pGr}(0,n)$ is just the zero matrix. Similarly, $\mathrm{pGr}(n,n) = \{\mathrm{Id}_n\}$. For $k = 1$, we note that $\mathrm{pGr}(1,n)$ is the $(n-1)$-dimensional variety consisting of all symmetric rank 1 matrices of trace 1. This is an affine-linear section of the quadratic Veronese variety $\nu_2(\mathbb{R}^n)$, so its degree is 2^{n-1}. The remaining claims are the statements of Corollaries 4.4 and 4.5 in [88]. ☐

These degree are listed for small values of n and k in following table. Rows 3-8 were created with `Macaulay2` [60]. We used `HomotopyContinuation.jl` [20] for rows 9-10.

$n \backslash k$	0	1	2	3	4	5	6	7	8	9	10
3	1	4	4	1							
4	1	8	12	8	1						
5	1	16	40	40	16	1					
6	1	32	140	184	140	32	1				
7	1	64	504	992	992	504	64	1			
8	1	128	1 848	5 824	7 056	5 824	1 848	128	1		
9	1	256	6 864	36 096	60 864	60 864	36 096	6 864	256	1	
10	1	512	25 740	232 320	587 664	672 288	587 664	232 320	25 740	512	1

3.4 Projection Flag Variety

The projection flag variety embeds the flag variety as a tuple of projection matrices of different ranks. The relations between these matrices encode the containment of subspaces. Indeed, if P and P' are projection matrices with $\mathrm{rank}(P) \leq \mathrm{rank}(P')$, then $\mathrm{im}(P) \subseteq \mathrm{im}(P')$ precisely when $P'P = P$. The following set-theoretic description may be found in [131, Proposition 17]. As before, let $\mathbf{k} = (k_1, \ldots, k_r)$.

Proposition 3.4.1 (Projection Flag Variety)**.** *The projection flag variety*

$$\mathrm{pFl}(\mathbf{k}; n) = \{(P_1, \ldots, P_r) \in (\mathrm{Sym}^2\mathbb{C}^n)^r : P_iP_j = P_j, \mathrm{trace}(P_i) = k_i \text{ for } 1 \leq j \leq i \leq r\}. \quad (3.6)$$

is an irreducible affine variety of dimension $\sum_{i=1}^r (k_i - k_{i-1}(n - k_i))$*. It is equal to an orthogonal group orbit:*

$$\mathrm{pFl}(\mathbf{k}; n) = \{(R^T E_{k_1} R, \cdots, R^T E_{k_r} R) : R \in \mathrm{SO}(n)\}.$$

Proof. We begin by proving the second claim. It is not hard to verify that the tuples in the orbit satisfy the relations in (3.6). We argue the reverse containment. Since each P_i is a point in the projection Grassmannian, it is diagonalizable to E_{k_i}. Note that $P_jP_i = (P_iP_j)^T = P_j^T = P_j$ and hence all pairs P_i, P_j commute and all P_i are simultaneously diagonalizable. The polar decomposition of the diagonalizing matrix can be used to simultaneously orthogonally diagonalize the P_i to the E_{k_i}; see the proof of [53, Theorem XI.4]. The sign of the last column of R can be flipped to ensure that the matrix is in $\mathrm{SO}(n)$. Since $\mathrm{SO}(n)$ is irreducible, the irreducibility of $\mathrm{pFl}(\mathbf{k}; n)$ follows. The dimension follows from the injectivity of the map

$$\begin{aligned} \mathrm{SO}(n)/\mathrm{SO}(k_1) \times \cdots \times \mathrm{SO}(n - k_r) &\to \mathrm{pFl}(\mathbf{k}; n) \\ R &\mapsto (R^T E_{k_1} R, \cdots, R^T E_{k_r} R). \end{aligned}$$

□

The equations in (3.6) define the projection flag variety set-theoretically. We now prove that they generate the prime ideal $I(\mathrm{pFl}(\mathbf{k}; n))$. We note that in (3.6) it suffices to retain the relations $P_i^2 = P_i$ for $i = 1, \ldots, r$ and $P_iP_{i-1} = P_{i-1}$ for $i = 2, \ldots, r$.

Theorem 3.4.2. *Let* $\mathrm{Fl}(k_1,\dots,k_r;n)$ *be a flag variety and let* $P_1,\dots,P_r$ *be* $n\times n$ *symmetric matrices of unknowns. Then* $\mathrm{pFl}(\mathbf{k};n)$ *is smooth and has prime ideal*

$$\langle P_iP_{i-1}-P_{i-1} : 2\le i\le r\rangle + \langle P_i^2-P_i, \mathrm{trace}(P_i)-k_i : 1\le i\le r\rangle \subset \mathbb{C}[P_1,\dots,P_r].$$

Proof. By Proposition 3.4.1 and Lemma 3.1.3, it suffices to check that $\mathrm{pFl}(\mathbf{k};n)$ is smooth at $E_{\mathbf{k}}=(E_{k_1},\dots,E_{k_r})$. The codimension of $\mathrm{pFl}(\mathbf{k};n)$ is $N=\sum_{i=1}^r\binom{n+1}{2}-n(n-k_i)+\sum_{i=2}^r k_{i-1}(n-k_i)$. We identify N linearly independent rows of the Jacobian evaluated at $E_{\mathbf{k}}$. Fix $i\in\{1,\dots,r\}$ and choose $\ell,j\in\{1,\dots,k_i\}$ or $\ell,j\in\{k_i+1,\dots,n\}$. Then

$$\left.\frac{\partial(P_i^2-P_i)_{\ell j}}{\partial(P_{i'})_{\ell' j'}}\right|_{E_{\mathbf{k}}} = \begin{cases}\pm 1 & i'=i,\ell'=\ell,j'=j,\\ 0 & \text{else}\end{cases}$$

where the sign is positive if $\ell,j\le k_i$ and negative if $k_i<\ell,j$. There are $\binom{n+1}{2}-k_i(n-k_i)$ choices for such pairs $\{\ell,j\}$.

Next fix $i\in\{2,\dots,r\}$, let $j\in\{1,\dots,k_{i-1}\}$, and let $\ell\in\{k_i+1,\dots,n\}$. Then

$$\left.\frac{\partial(P_iP_{i-1}-P_{i-1})_{\ell j}}{\partial(P_{i'})_{\ell' j'}}\right|_{E_{\mathbf{k}}} = \begin{cases}1 & \text{if } i'=i,\ell'=\ell,j'=j,\\ -1 & \text{if } i'=i-1,\ell'=\ell,j'=j,\\ 0 & \text{else.}\end{cases}$$

There are $k_{i-1}(n-k_i)$ choices for pairs $\{\ell,j\}$. Ranging over all choices of i yields N rows of the Jacobian with disjoint supports. Thus the Jacobian has the desired rank. $\square$

3.5 Isospectral Flag Variety

We now discuss the *isospectral model* which was first introduced by Lim and Ye [89]. This model realizes the flag variety $\mathrm{Fl}(\mathbf{k};n)$ as the set of real symmetric matrices S with a fixed spectrum $\mathbf{c}=(c_1\ge\cdots\ge c_n)$ such that $c_1=\cdots=c_{k_1}$, $c_{k_1+1}=\cdots=c_{k_2}$, etc. The isospectral model contains a unique diagonal matrix $S_0=\mathrm{diag}(c_1,\dots,c_n)$. The nested subspaces are identified with the nested eigenspaces of S, i.e., the smallest subspace of the flag is given by the eigenspace of c_{k_1}, the next subspace is given by the union of the eigenspaces of c_{k_1} and c_{k_2}. To recover the Grassmannian, one sets $c_1=\cdots=c_k=1$ and $c_{k+1}=\cdots=c_n=0$. The eigenvalues of a matrix can be recovered from its minimal polynomial. To find the multiplicities, the trace is usually enough, unless the values of $\mathbf{c}$ are very special. We therefore have the following equational description of the isospectral model for generic $\mathbf{c}$.

Proposition 3.5.1 (Isospectral Model)**.** *The isospectral model*

$$\mathrm{Fl}_{\mathbf{c}}(\mathbf{k};n)=\{S\in\mathrm{Sym}^2\mathbb{C}^n : \prod_{j=1}^r(S-c_{k_j}\mathrm{Id}_n)=0,\ \mathrm{trace}(S)=\sum_{j=1}^n c_j\} \tag{3.7}$$

is an irreducible affine variety of dimension $\sum_{i=1}^{r}(k_i - k_{i-1}(n-k_i))$*. It is the following orthogonal group orbit:*

$$\mathrm{Fl}_{\mathbf{c}}(\mathbf{k};n) = \{R^T S_0 R : R \in \mathrm{SO}(n)\}.$$

Proof. It is clear that all matrices in the $\mathrm{SO}(n)$-orbit satisfy the conditions in (3.7). Conversely, every $S \in \mathrm{Fl}_{\mathbf{c}}(\mathbf{k},n)$ has the correct eigenvalues and multiplicities by [89, Proposition 4.5] and is diagonalizable to S_0 since it has a squarefree minimal polynomial. We again apply [53, Theorem XI.4] to conclude that it is orthogonally diagonalizable to S_0. Irreducibility follows from the description of the variety as an $\mathrm{SO}(n)$ orbit and the dimension follows from the injectivity of the map

$$\mathrm{SO}(n)/\mathrm{SO}(k_1) \times \cdots \times \mathrm{SO}(n-k_r) \to \mathrm{Fl}_{\mathbf{c}}(k,n), \qquad R \mapsto R^T S_0 R.$$ □

Theorem 3.5.2. *Let* $\mathrm{Fl}(\mathbf{k};n)$ *be a flag variety and let* S *be a symmetric matrix of unknowns. Given a generic choice of* $c_1,\ldots,c_n$ *satisfying* $c_{k_j+1} = \cdots = c_{k_{j+1}}$ *for* $j = 0,\ldots,r$*, the variety* $\mathrm{Fl}_{\mathbf{c}}(\mathbf{k};n)$ *is smooth and its prime ideal is*

$$I(\mathrm{Fl}_{\mathbf{c}}(\mathbf{k};n)) = \langle \prod_{j=1}^{r}(S - c_{k_j}\mathrm{Id}_n),\ \mathrm{trace}(S) - \sum_{j=1}^{n} c_j \rangle \subseteq \mathbb{C}[S].$$

Proof. Write X for the scheme defined by the proposed ideal $I(\mathrm{Fl}_{\mathbf{c}}(\mathbf{k};n))$. By Proposition 3.5.1 and Lemma 3.1.3, it suffices to show that X is smooth at S_0. We use the Jacobian criterion, so our task is to show that $\mathcal{J}_X(S_0)$ has rank at least $n + \binom{k_1}{2} + \cdots + \binom{n-k_r}{2}$. The rows and columns of $\mathcal{J}_X(S_0)$ are both indexed by entries (i,j) of the matrix S. We will show that the entry $\mathcal{J}_X(S_0)_{(i,j),(i',j')}$ is nonzero precisely when $i = i', j = j'$, and $c_i = c_j$. With the correct ordering of variables, $\mathcal{J}_X(S_0)$ is diagonal with diagonal entries of the form $\mathcal{J}_X(S_0)_{(i,j),(i,j)}$. It is then clear from counting that this matrix has the correct rank.

We prove off-diagonal entries of $\mathcal{J}_X(S_0)$ are 0. If $\{i,j\} \cap \{i',j'\} = \emptyset$, by the product rule

$$\frac{\partial (S^m)_{ij}}{\partial s_{i'j'}}\bigg|_{S=S_0} = \sum_{\ell=1}^{n} s_{i\ell} \frac{\partial (S^{m-1})_{\ell j}}{\partial s_{i'j'}}\bigg|_{S=S_0} = c_i \frac{\partial (S^{m-1})_{ij}}{\partial s_{i'j'}}\bigg|_{S=S_0} = 0$$

where the second equality comes from evaluating at $S = S_0$ and the last equality is induction on m. Since derivatives of all powers of X vanish, $\mathcal{J}_X(S_0)_{(i,j),(i',j')} = 0$ if $\{i,j\} \cap \{i',j'\} = \emptyset$.

Next suppose $i = i'$ and $j \neq j'$. Again using the product rule,

$$\frac{\partial (S^m)_{ij}}{\partial s_{ij'}}\bigg|_{S=S_0} = (S_0^{m-1})_{j'j} + \sum_{\ell=1}^{n} s_{i\ell} \frac{\partial (S^{m-1})_{\ell j}}{\partial s_{ij'}}\bigg|_{S=S_0} = c_i \frac{\partial (S^{m-1})_{ij}}{\partial s_{ij'}}\bigg|_{S=S_0} = 0$$

where the last equality is by induction on m. As above, we have $\mathcal{J}_X(S_0)_{(i,j),(i,j')} = 0$ for $j \neq j'$. This concludes the proof that $\mathcal{J}_X(S_0)$ is a diagonal matrix.

Finally, we prove that the diagonal entries $\mathcal{J}_X(S_0)_{(i,j),(i,j)}$ are nonzero if $c_i = c_j$. By induction, we compute

$$\frac{\partial (S^m)_{ij}}{\partial s_{ij}}\bigg|_{S=S_0} = c_j^{m-1} + c_i \frac{\partial (S^{m-1})_{ij}}{\partial s_{ij}}\bigg|_{S=S_0} = \sum_{p=0}^{m-1} c_i^p c_j^{m-1-p} = m c_i^{m-1}.$$

Therefore the entry $\mathcal{J}_X(S_0)_{(i,j),(i,j)} = f'(c_i)$ where $f(z) = (z - c_{k_1}) \cdots (z - c_{k_r})$ is the minimal polynomial of X. Since c_i is a root of f and f has no double roots $\mathcal{J}_X(S_0) \neq 0$.

Since $\mathcal{J}_X(S_0)$ is diagonal with at least $n + \binom{k_1}{2} + \cdots + \binom{n-k_r}{2}$ nonzero entries, we conclude that $\mathcal{J}_X(S_0)$ has rank at least $n + \binom{k_1}{2} + \cdots + \binom{n-k_r}{2}$. □

We believe that the degree of the isospectral model can be derived from the representation theory of the orthogonal group using the same methods that were used to find the degree of the projection Grassmannian [88]. We present here degrees for the isospectral models whose matrices have sizes 4 and 5. We index the degrees by the partitions λ that give the multiplicities of the eigenvalues c. Note that $k_i = \lambda_1 + \cdots + \lambda_i$.

	$n=4$					$n=5$						
λ	4^1	$3^1 1^1$	2^2	$2^1 1^2$	1^4	5^1	$4^1 1^1$	$3^1 2^1$	$3^1 1^2$	$2^2 1^1$	$2^1 1^3$	1^5
Degrees	1	8	12	20	24	1	16	40	70	92	120	120

3.6 Moving Between Different Lives

This section is dedicated to understanding the relationships between the different lives of the flag variety (and the Grassmannian) described above. The results are collected in Figure 3.1.

The most difficult relationship is the one between the projection and Plücker coordinates of the Grassmannian. Given a projection matrix P with $\mathrm{rank}(P) = k$, one obtains the Plücker coordinates by taking maximal minors of any k linearly independent rows of P. For the other direction, we define the *cocircuit matrix* of $x \in \mathrm{Gr}(k,n)$ as the $k \times \binom{n}{k-1}$ matrix $X = (x_{iK})$. The image of this matrix is the linear space represented by x. If $k = 2$ and $n = 5$ then the cocircuit matrix is the skew-symmetric $n \times n$ matrix X we saw in (3.3). This generalizes to $k = 2$ and $n \geq 6$. For $k \geq 3$, the cocircuit matrix has more columns than rows.

Example 3.6.1 ($n = 6, k = 3$)**.** The Grassmannian $\mathrm{Gr}(3,6)$ has dimension 9. As a projective variety it has degree 42 in $\mathbb{P}^{19}_{\mathbb{R}}$ and its ideal is generated by 35 quadrics in the 20 Plücker coordinates x_{ijk}. The 15 cocircuits of the subspace are the columns of the cocircuit matrix

$$X = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & x_{123} & x_{124} & \cdots & x_{146} & x_{156} \\ 0 & -x_{123} & -x_{124} & -x_{125} & -x_{126} & 0 & 0 & \cdots & x_{246} & x_{256} \\ x_{123} & 0 & -x_{134} & -x_{135} & -x_{136} & 0 & -x_{234} & \cdots & x_{346} & x_{356} \\ x_{124} & x_{134} & 0 & -x_{145} & -x_{146} & x_{234} & 0 & \cdots & 0 & x_{456} \\ x_{125} & x_{135} & x_{145} & 0 & -x_{156} & x_{235} & x_{245} & \cdots & -x_{456} & 0 \\ x_{126} & x_{136} & x_{146} & x_{156} & 0 & x_{236} & x_{246} & \cdots & 0 & 0 \end{bmatrix}.$$

As an affine variety, the Grassmannian $\mathrm{Gr}(3,6)$ lives in $\mathbb{R}^{21}$. This embedded manifold consists of all 6×6 symmetric matrices P with $P^2 = P$ and $\mathrm{trace}(P) = 3$. ⋄

A version of the following result was first proved by Bloch and Karp; see [10, Lemma 4.11].

Theorem 3.6.2. *The entries of the projection matrix P are ratios of quadratic forms in Plücker coordinates:*

$$P = \frac{k}{\mathrm{trace}(XX^T)} XX^T.$$

Proof. Our point of departure is the formula $P = A(A^TA)^{-1}A^T$ where $A\in\mathbb{R}^{n\times k}$. By the Cauchy-Binet formula,

$$\det(A^TA) \;=\; \sum_I x_I^2.$$

If we multiply the inverse of A^TA by this sum of squares then we obtain the adjoint of A^TA. This is a matrix whose entries are the $(k-1)\times(k-1)$ minors of A^TA. We can identify the adjoint with the $(k-1)$st exterior power, using the isomorphism $\mathbb{R}^k \simeq \wedge_{k-1}\mathbb{R}^k$ that sends the basis vector e_i to the basis vector $(-1)^i e_1\wedge\cdots\wedge e_{i-1}\wedge e_{i+1}\wedge\cdots\wedge e_k$. In symbols, we have

$$\mathrm{adj}(A^TA) \;=\; \wedge_{k-1}(A^TA) \;=\; \wedge_{k-1}A^T\cdot\wedge_{k-1}A \;=\; (\wedge_{k-1}A)^T\cdot\wedge_{k-1}A.$$

This implies that the scaled projection matrix equals

$$\Bigl(\sum_I x_I^2\Bigr)\cdot P \;=\; A\mathrm{adj}(A^TA)A^T \;=\; (A\wedge_{k-1}A^T)\cdot(A\wedge_{k-1}A^T)^T. \tag{3.8}$$

It remains to analyze the matrix $A(\wedge_{k-1}A)^T$. This matrix has format $n\times\binom{n}{k-1}$ and its rank is k. We claim that its entry in row i and column K is equal to x_{iK}. Indeed, x_{iK} is the $k\times k$ minor of A with row indices iK. The (i,K) entry in the matrix product $A(\wedge_{k-1}A)^T$ is computed by multiplying the ith row of A with the Kth column of $(\wedge_{k-1}A)^T$. This is precisely the Laplace expansion of the iK minor of A with respect to ith column. This implies that the entry of (3.8) in (i,j) is equal to $\sum_K x_{iK}\,x_{jK}$. This completes the proof. □

We introduced two generalizations of the projection coordinates for flag varieties. We illustrate the relationship between Plücker coordinates and projection coordinates for the flag variety with the following example.

Example 3.6.3. We show the different ways of representing the flag variety $\mathrm{Fl}(1,2;3)$. In Stiefel coordinates, this flag variety is realized as $\mathrm{St}(2,3)/\mathrm{O}(1)\times\mathrm{O}(1)$ via a 3×2 matrix

$$Z = \begin{pmatrix} z_{11} & z_{12}\\ z_{21} & z_{22}\\ z_{31} & z_{32}\end{pmatrix}$$

with orthonormal columns. We identify such matrices up to the sign of each column.

To obtain the Plücker coordinates for the flag represented by Z, we take minors. This yields six Plücker coordinates that realize this complete flag variety in $\mathbb{P}^2 \times \mathbb{P}^2$:

$$\begin{array}{ccc} x_1 = z_{11}, & x_2 = z_{21}, & x_3 = z_{31}, \\ x_{12} = z_{11}z_{22} - z_{12}z_{21}, & x_{13} = z_{11}z_{32} - z_{12}z_{31}, & x_{23} = z_{21}z_{32} - z_{22}z_{31}. \end{array}$$

By Theorem 3.2.1 these coordinates satisfy a single relation $x_1x_{23} - x_2x_{13} + x_3x_{12} = 0$.

To represent the flag in projection coordinates, let P_1 be the unique orthogonal projection matrix onto the space spanned by the first column of Z and P_2 be the unique orthogonal matrix onto the column space of Z:

$$P_1 = \begin{pmatrix} z_{11} \\ z_{21} \\ z_{31} \end{pmatrix} \begin{pmatrix} z_{11} & z_{21} & z_{31} \end{pmatrix} \qquad P_2 = ZZ^T.$$

One checks that these matrices satisfy the relations

$$\begin{array}{ccc} P_1^2 = P_1 & P_2^2 = P_2 & P_2P_1 = P_1 \\ \mathrm{trace}(P_1) = 1 & \mathrm{trace}(P_2) = 2. & \end{array}$$

Finally, we can also represent this flag with a single matrix via isospectral coordinates. We append a column to Z to form a 3×3 orthogonal matrix $\tilde{Z}$ and fix the generic vector $\mathbf{c} = (c_1, c_2, c_3)$. The isospectral coordinates for Z with respect to $\mathbf{c}$ are given by the matrix

$$S = \tilde{Z}\,\mathrm{diag}(c_1, c_2, c_3)\,\tilde{Z}^T$$

whose eigenvalues are c_1, c_2, c_3 and whose eigenvectors are the columns of $\tilde{Z}$. This matrix satisfies its minimal polynomial and has the correct trace:

$$(S - c_1\mathrm{Id}_3)(S - c_2\mathrm{Id}_3)(S - c_3\mathrm{Id}_3) = 0 \quad \mathrm{trace}(S) = c_1 + c_2 + c_3. \qquad \diamond$$

From this example, we see that it is not difficult to write the Plücker, projection, and isospectral coordinates once we have Stiefel coordinates for a flag. The following proposition summarizes what is known about how to write different coordinates for the flag variety in terms of other coordinates.

Proposition 3.6.4. *The relationships between the lives of the flag variety* $\mathrm{Fl}(k_1, k_2, \ldots, k_r; n)$ *are captured in Figure 3.1. In the figure,* Z *is an* $n \times k_r$ *matrix in* $\mathrm{St}(k_r, n)$*,* $\tilde{Z}$ *is an* $n \times n$ *orthogonal matrix whose first* k_r *columns comprise* Z*, and* $Z_{I,J}$ *denotes the submatrix of* Z *with rows indexed by the set* I *and columns indexed by the set* J*. The variables* $x_{i_1,\ldots,i_{k_s}}$ *indicate Plücker coordinates and* X_i *denotes the cocircuit matrix for* $\mathrm{Gr}(k_i, n)$*. The matrix* P_i *is a projection matrix in the tuple of* $(P_1, \ldots, P_r)$*. Finally,* S *is an* $n \times n$ *symmetric matrix with eigenvalues* $\mathbf{c} = (c_1, \ldots, c_n)$*.*

$$\begin{tikzcd}
& \text{Plücker} \arrow[ddl, "P_i = \frac{k_i}{\operatorname{trace}(X_i X_i^T)} X_i X_i^T"'] & \\
& \text{Stiefel} \arrow[u, "x_{i_1,\dots,i_{k_s}} = \det(Z_{i_1\cdots i_{k_s},1\cdots k_s})"'] \arrow[dl, leftrightarrow, "P_i = Z_{1\cdots n,1\cdots k_i} Z_{1\cdots n,1\cdots k_i}^T"] \arrow[dr, leftrightarrow, "S = \tilde{Z}\,\operatorname{diag}(c_1,\dots,c_n)\,\tilde{Z}^T"] & \\
\text{Projection} \arrow[rr, "S = c_n \mathrm{Id}_n + (c_{k_r} - c_n)P_{k_r} + \cdots + (c_{k_1} - c_{k_2})P_1"'] & & \text{Isospectral}
\end{tikzcd}$$

Figure 3.1: Diagram explaining how to move from one life of the flag variety to another. If $A \to B$ in the diagram, the edge label explains how to write the B coordinates in terms of the A coordinates. Two of the arrows are bidirectional, meaning that one direction comes from matrix multiplication and the other comes from a matrix factorization.

Proof. The map from Stiefel to Plücker coordinates is as in (3.2). The map from Plücker to projection coordinates for the Grassmannian in Theorem 3.6.2 extends to flag varieties.

To move from Stiefel coordinates to projection coordinates, one takes projection matrices onto the spans of the appropriate submatrices of Z. To go the other way, one takes a spectral decomposition of $P_r = \tilde{Z} E_{k_r} \tilde{Z}^T$ where E_i is the $n \times n$ matrix whose $(1,1), \dots, (i,i)$ entries are 1 and all other entries are zero. Set Z to be the first k_r columns of $\tilde{Z}$.

The map from isospectral to Stiefel coordinates is similar: take the spectral decomposition of $S = \tilde{Z}\,\operatorname{diag}(c_1, \dots, c_n)\,\tilde{Z}^T$ and remove the last $n - k_r$ columns from $\tilde{Z}$. Given Stiefel coordinates Z, one forms $\tilde{Z}$ by extending the columns of Z to an orthonormal basis for $\mathbb{C}^n$. The equivalence between these coordinates follows from the definitions of the Stiefel and isospectral coordinates.

To move from projection coordinates to isospectral coordinates one composes the maps passing through the Stiefel coordinates. Write the diagonal matrix

$$\operatorname{diag}(c_1, \dots, c_n) = c_n \mathrm{Id}_n + (c_{k_r} - c_n)E_{k_r} + \cdots + (c_{k_1} - c_{k_2})E_{k_1}.$$

Conjugating with the augmented Stiefel coordinates $\tilde{Z}$ yields the expression for S in terms of projection coordinates. $\square$

In this chapter, we gave dimension formulas for and described the prime ideals of the Stiefel manifold, Grassmannians and flag varieties in the Plücker embedding, the projection Grassmannian, the projection flag variety, and the isospectral model. We collected results on the degrees on these varieties and concluded by describing the maps between these different embeddings. We will now turn our attention to optimization problems on these varieties.

Part II

Grassmannians in Applications

We now turn to polynomial optimization problems on the Grassmannian that arise in applications. In Chapter 4, we consider problems from linear algebra and statistics that can naturally be formulated as optimization problems on the Grassmannian. In particular, we consider generalizations of the eigenvalue problem (Sections 4.1 and 4.2) as well as canonical correlation analysis and correspondence analysis (Section 4.3). For each of these optimization problems, we give complete descriptions of the complex critical points, all of which are real.

In Chapter 5, we consider a constrained version of the eigenvalue problem that arises in quantum chemistry: the minimization of the Rayleigh quotient on a tensor train variety. We prove that tensor train varieties are parametrized by a product of Grassmannians (Theorem 5.4.3). We discuss the critical ideal and the discriminant of the optimization problem in Sections 5.1 and 5.2, and we conclude with numerical experiments in Section 5.5.

In Chapter 6, we turn to Euclidean distance minimization inside the Grassmannian of lines $\mathrm{Gr}(2, n)$. Our starting point is the map between the Plücker embedding of the Grassmannian and the projection Grassmannian introduced in Section 3.6. We study the projective version of this map in Section 6.1. We introduce the *Grassmann distance degree*, the algebraic degree of Euclidean distance minimization for a subvariety of the Grassmannian when the data is also on the Grassmannian (Section 6.3). We explore this idea in detail for subvarieties of $\mathrm{Gr}(2, 4)$ (Section 6.4) and for Schubert varieties (Section 6.5).

Chapter 4

Linear Algebra, Statistics, and Optimization

We now turn to optimization problems from linear algebra and statistics that involve Grassmannians and flag varieties. Our starting point is the eigenvalue problem for symmetric matrices. Let H be a real symmetric $n \times n$ matrix and consider the problem of minimizing or maximizing the Rayleigh quotient:

$$\min_{z \neq 0}/\max \ \frac{z^T H z}{z^T z}. \tag{4.1}$$

It is known that the optimal value for this eigenvalue problem is the minimum (resp. maximum) eigenvalue of H. A formulation which avoids the fraction in the objective function is the quadratic optimization problem over S^{n-1}, the unit sphere of dimension $n-1$:

$$\min_{z \in S^{n-1}}/\max \ z^T H z. \tag{4.2}$$

The optimal solution to the above *eigenvalue problem* (4.1) and (4.2) is a unit eigenvector corresponding to the minimum/maximum eigenvalue. The critical points of this constrained optimization problem are precisely the unit eigenvectors of H. The choice of sign for each eigenvector gives us $2n$ critical points.

The eigenvalue problem has the following generalization to compute multiple eigenvectors at once where we replace the vector z in (4.2) with an $n \times k$ matrix Z consisting of orthonormal columns. We solve a quadratic optimization problem over the Stiefel manifold called the *multi-eigenvector problem*:

$$\min_{Z^T Z = \mathrm{Id}_k}/\max \ \operatorname{trace}(Z^T H Z). \tag{4.3}$$

The critical points of this problem are the sets of orthogonal k-frames consisting of unit eigenvectors of H, up to the action of the orthogonal group $\mathrm{O}(k)$ over the complex numbers; this is the content of Theorem 4.1.2.

Because the multi-eigenvector problem is $\mathrm{O}(k)$-invariant, it is more naturally considered over the Grassmannian; see Chapter 3. In Section 4.1, we will restate this problem over the Grassmannian in projection coordinates (see Section 3.3), where it has $\binom{n}{k}$ critical points (Theorem 4.1.3). To formulate this problem over the flag variety, one can use isospectral coordinates (see Section 3.5). In the isospectral formulation, the set of critical points is a disjoint union of products of smaller orthogonal groups (Theorem 4.1.5). In both Theorems 4.1.3 and 4.1.5, we give explicit descriptions of the critical points.

In Section 4.2, we study the heterogeneous quadratics minimization problem, which is used as a benchmark for testing numerical methods for Riemannian optimization [105]. Given symmetric $n \times n$ matrices $H_1, \ldots, H_k$, we seek to find $Z = (Z_1 \cdots Z_k)$ such that

$$\min/\max_{Z^T Z = \mathrm{Id}_k} \sum_{i=1}^{k} Z_i^T H_i Z_i.$$

If $H_1 = \cdots = H_k$, one recovers the multi-eigenvector problem (4.3). However, for a generic choice of $H_1, \ldots, H_k$, this problem is not an optimization problem over the Grassmannian, but rather over the flag variety $\mathrm{Fl}(1, \ldots, k; n)$. We formulate the problem in projection coordinates and observe that in this formulation the heterogeneous quadratics minimization problem is a generic linear optimization problem over the flag variety. We use numerical methods to compute the critical point counts for small values of k (Table 4.1). We also report our observation based on our computational experiments that taking $H_1, \ldots, H_k$ to be generic diagonal matrices does not change the number of critical points. For $n = 3$ and $k = 2$, we show that this number is 40 (Proposition 4.2.2). The code for these computations is available at `https://github.com/hannahfriedman/flag_optimization_algebraic_perspective`.

In Section 4.3, we examine two problems from statistics [131]: canonical correlation analysis and correspondence analysis. In these problems, we consider a rectangular matrix H instead of a symmetric matrix and the critical points of the problems are given by the singular value decomposition of H. In the first case, we give a formula for the finite number of critical points which correspond to pairs of left and right singular vectors of H (Theorem 4.3.1). The second problem can be naturally formulated over a product of Grassmannians, and we give a complete description of the critical points in this case as well (Theorem 4.3.2).

4.1 The Multi-Eigenvector Problem

In this section, we give three different formulations of the multi-eigenvector problem (4.3). The first two are formulated as optimization problems over a Grassmannian in Stiefel and projection coordinates. The third one utilizes a flag variety in its isospectral formulation. In all three cases we describe the sets of critical points of these optimization problems, where in the Stiefel case we describe the $\mathrm{O}(k)$-orbit of the critical points in $\mathrm{St}(k, n)$.

Critical points in Stiefel coordinates

The goal of this section is to prove Theorem 4.1.2 which characterizes the critical points of the multi-eigenvector problem in Stiefel coordinates. We first prove the following lemma.

Lemma 4.1.1. *Suppose $H \in \mathbb{C}^{n\times n}$ is orthogonally diagonalizable and has full rank. Let $Z \in \mathrm{St}(k,n)$. If $M \in \mathbb{C}^{k\times k}$ is such that $HZ = ZM$, then M is also orthogonally diagonalizable.*

Proof. Let $H = U\Lambda U^T$ be a factorization of H with Λ diagonal and $U \in \mathrm{O}(n)$. Then $AZ = U\Lambda U^T Z = ZM$ and rearranging we have $M = (Z^TU)\Lambda(U^TZ)$ where $(Z^TU)(U^TZ) = \mathrm{Id}_k$. This factorization implies that for every diagonal entry λ_i of Λ, $M(Z^TU)_i = \lambda_i(Z^TU)_i$. Since M has at most k eigenvalues and the λ_i are nonzero, $n-k$ of the columns of Z^TU are forced to be zero. This yields an orthogonal diagonalization for M. $\square$

Theorem 4.1.2. *Let H be a generic real symmetric $n\times n$ matrix and let Z be an $n\times k$ variable matrix. The set of complex critical points of the multi-eigenvector problem (4.3) is*

$$\bigsqcup_{\{i_1,\dots,i_k\}\in\binom{[n]}{k}} \{(u_{i_1}\, u_{i_2}\, \cdots\, u_{i_k})Q \,:\, Q\in \mathrm{O}(k)\}$$

where $u_1,\dots,u_n$ is an orthonormal eigenbasis of H. This algebraic set is a disjoint union of $\binom{n}{k}$ varieties isomorphic to $\mathrm{O}(k)$; it has dimension $\dim(\mathrm{O}(k)) = \binom{k}{2}$ and degree $\deg(\mathrm{O}(k))\binom{n}{k}$.

The degree of the orthogonal group $\mathrm{O}(k)$ is stated in Theorem 3.1.5. Notice in particular that the critical point set is positive dimensional. This is because the problem does not satisfy Assumption 2.2.1. The gradient of the objective function is the row vector

$$\nabla\mathrm{trace}(Z^THZ) = 2\begin{pmatrix} Z_1^TH & Z_2^TH & \cdots & Z_k^TH\end{pmatrix}$$

with nk entries where Z_i denotes the ith column of Z. For a fixed $Z\in\mathrm{St}(k,n)$ and any $H\in \mathrm{Sym}^2\mathbb{C}^n$, the relation $(Z_1^TH)Z_2 = (Z_2^TH)Z_1$ must hold. Since Z is fixed, this is a nontrivial linear condition on the entries of $\nabla\mathrm{trace}(Z^THZ)$, and hence the map $H\mapsto \nabla\mathrm{trace}(Z^THZ)$ is not surjective onto $\mathbb{C}^{nk}$.

Proof of Theorem 4.1.2. The Jacobian of the polynomials defined by $Z^TZ = \mathrm{Id}_k$ is $\mathcal{J}_{\mathrm{St}(k,n)}(Z)$ as in (3.1). A critical point is a vector Z satisfying $Z^TZ = \mathrm{Id}_k$ where $\nabla\mathrm{trace}(Z^THZ)$ is in the row span of $\mathcal{J}_{\mathrm{St}(k,n)}(Z)$. In particular, there exist $\mu_{ij}\in\mathbb{C}$, $1\le i,j\le k$ with $\mu_{ij} = \mu_{ji}$ such that $HZ_i = \sum_{j=1}^k \mu_{ij}Z_j$. These μ_{ij} can be assembled into a $k\times k$ symmetric matrix M satisfying $HZ = ZM$. So Z is a critical point if and only if it satisfies $Z^TZ = \mathrm{Id}_k$ and there exists a symmetric matrix $M\in\mathrm{Sym}^2\mathbb{C}^k$ such that $HZ = ZM$.

If such a matrix M exists, then by Lemma 4.1.1, it is orthogonally diagonalizable. Therefore we replace the condition $HZ = ZM$ with the condition $HZ = ZQ^T\Lambda Q$, where $Q\in\mathrm{O}(k)$ and Λ is diagonal. Thus Z is a critical point if and only if it satisfies $Z^TZ = \mathrm{Id}_k$ and $HZQ^T = ZQ^T\Lambda$. These conditions hold precisely when the columns of ZQ^T are a subset of an orthonormal eigenbasis for H, which proves that Z has the desired form. $\square$

Critical points in projection coordinates

In Stiefel coordinates, we work with an orbit of critical points in $\mathrm{St}(k,n)$ which corresponds to a k-dimensional subspace spanned by k eigenvectors of H. It is desirable to formulate the multi-eigenvector problem so that it has finitely many critical points, because equivalence classes are difficult to compute with. We therefore transform the multi-eigenvector problem into projection coordinates for the Grassmannian as in Section 3.6. If Z is $n \times k$ and $Z^T Z = \mathrm{Id}_k$, then $P = ZZ^T$ projects onto the column space of Z. Then using the trace trick, we have $\mathrm{trace}(Z^T HZ) = \mathrm{trace}(HP)$ and we can reformulate our problem as

$$\min_{\substack{P^2=P \\ \mathrm{trace}(P)=k}}/\max \ \mathrm{trace}(HP). \tag{4.4}$$

Theorem 4.1.3. *The degree of the multi-eigenvector problem (4.4) over the projection Grassmannian* $\mathrm{pGr}(k,n)$ *is* $\binom{n}{k}$. *The critical points are the k-dimensional subspaces spanned by k eigenvectors of H.*

Proof. Consider the parametrization ϕ from the Stiefel variety $\mathrm{St}(k,n)$ to the projection Grassmannian which sends $Z \mapsto ZZ^T$. We apply Lemma 2.4.1 to prove that the critical points are preserved by this map. By Theorems 3.1.2 and 3.3.2 and Proposition 3.6.4, this is a surjective morphism between the two smooth varieties where an entire orbit in Theorem 4.1.2 is sent to a single projection matrix. Now we prove that the differential $d\phi_Z$ is surjective. Since $\mathrm{O}(n)$ acts transitively on $\mathrm{St}(k,n)$, we consider the tangent space $T_{(e_1\cdots e_k)}(\mathrm{St}(k,n))$. We note that the columns of $\mathcal{J}_{\mathrm{St}(k,n)}(Z)$ as in (3.1) are indexed by z_{ij} for $i = 1,\dots,n$ and $j = 1,\dots,k$. We see that the standard unit vectors e_{ij} for $i = k+1,\dots,n$ and $j = 1,\dots,k$ are in the kernel of $\mathcal{J}_{\mathrm{St}(k,n)}((e_1\cdots e_k))$. We next compute the Jacobian of ϕ. For this, we order the kn columns of $\mathcal{J}_\phi(Z)$ by the entries of the rows of Z. It will be more convenient to look at blocks of entries in each row corresponding to the rows of Z which we denote by $\tilde{Z}_1,\dots,\tilde{Z}_n$. The rows of $\mathcal{J}_\phi(Z)$ correspond to the $\binom{n+1}{2}$ entries of the symmetric matrix ZZ^T. We order the rows by first the diagonal entries of this matrix, then the first super diagonal, followed by the second super diagonal, etc. Then the submatrix consisting of the first n rows of $\mathcal{J}_\phi(Z)$ is $\mathrm{diag}(2\tilde{Z}_1,\dots,2\tilde{Z}_n)$. The next $n-1$ rows will be of the form

$$(0\cdots 0\,\tilde{Z}_{i+1}\,\tilde{Z}_i\,0\cdots 0)$$

for $i = 1,\dots,n-1$ where $\tilde{Z}_i$ appears at block $i+1$. The following $n-2$ rows are of the form

$$(0\cdots 0\,\tilde{Z}_{i+2}\,0\,\tilde{Z}_i\,0\cdots 0)$$

for $i = 1,\dots,n-2$ where $\tilde{Z}_i$ appears at block $i+2$, etc. The last row will be $(\tilde{Z}_n\,0\cdots 0\,\tilde{Z}_1)$. Evaluating this Jacobian at $(e_1\cdots e_k)$ produces $k(n-k)$ distinct standard unit vectors of $\mathbb{C}^{kn}$ in the last $k(n-k)$ columns indexed by the entries of $\tilde{Z}_{k+1},\dots,\tilde{Z}_n$. Hence, multiplying $\mathcal{J}_\phi(Z)$ with the vectors e_{ij} for $i = k+1,\dots,n$ and $j = 1,\dots,k$ that are tangent vectors to $\mathrm{St}(k,n)$ at Z gives $k(n-k)$ linearly independent vectors in the tangent space to $\mathrm{pGr}(k,n)$ at ZZ^T. Since the dimension of this smooth Grassmannian is $k(n-k)$, we conclude that $d\phi_Z$ is surjective. We finish the proof by applying Lemma 2.4.1. □

The trace of the objective function is a generic linear function in the entries of H:

$$\mathrm{trace}(HP) = \sum_i h_{ii}p_{ii} + 2\sum_{i<j} h_{ij}p_{ij}.$$

Thus solving the multi-eigenvector problem on the projection Grassmannian is equivalent to solving a generic linear optimization problem on the same variety. The number of critical points of optimizing a generic linear form on a variety is called the *linear optimization degree* (LO degree) of the variety [96].

Corollary 4.1.4. *The LO degree of* $\mathrm{pGr}(k,n)$ *is* $\binom{n}{k}$.

Critical points in isospectral coordinates

We now present the most general formulation of the multi-eigenvector problem. While the Stiefel and projection formulations are optimization problems over the Grassmannian, writing this problem in isospectral coordinates allows us to generalize to an optimization problem over a flag variety $\mathrm{Fl}(\mathbf{k};n) = \mathrm{Fl}(k_1,\ldots,k_r;n)$. We proceed as in Section 3.5. We fix $c_1,\ldots,c_n \in \mathbb{C}$ with $c_1 = \cdots = c_{k_1}$, $c_{k_1+1} = \cdots = c_{k_2}$, …, $c_{k_r+1} = \cdots = c_n$. We let $S_0 = \mathrm{diag}(c_1,\ldots,c_n)$. With this, we introduce the following linear optimization problem over the isospectral model $\mathrm{Fl}_{\mathbf{c}}(\mathbf{k};n)$:

$$\min_{S\in \mathrm{Fl}_{\mathbf{c}}(\mathbf{k};n)}/\max \ \mathrm{trace}(HS). \tag{4.5}$$

If the c_i are distinct, then the optimization problem is over the complete flag variety and, as we will see, it has $n!$ critical points. On the other hand, if $c_1 = \cdots = c_k = 1$ and $c_{k+1} = \cdots = c_n = 0$, then $S = QS_0Q^T$ is a point in $\mathrm{pGr}(k,n)$ and we recover the optimization problem (4.4). The following theorem interpolates between these extreme cases.

Theorem 4.1.5. *The linear optimization problem* (4.5) *over the isospectral flag variety* $\mathrm{Fl}_{\mathbf{c}}(\mathbf{k};n)$ *has* $\binom{n}{k_1,k_2-k_1,\ldots,n-k_r}$ *critical points.*

We first prove the following which is an analog of Theorem 4.1.2. Let $\mathfrak{S}_n$ denote the symmetric group on n elements.

Theorem 4.1.6. *Let H be a generic real symmetric $n\times n$ matrix and $S_0 = \mathrm{diag}(c_1,\ldots,c_n)$ with $c_{k_j+1} = \cdots = c_{k_{j+1}}$ for $j = 0,\ldots,r$ where $k_0 = 0$ and $k_{r+1} = n$. The algebraic set of complex critical points of the optimization problem*

$$\min_{Q^TQ=\mathrm{Id}_n}/\max \ \mathrm{trace}(HQS_0Q^T)$$

is equal to

$$\bigsqcup_{\sigma\in\mathfrak{S}_n/(\mathfrak{S}_{k_1}\times\mathfrak{S}_{k_2-k_1}\cdots\times\mathfrak{S}_{n-k_r})} \{U(\sigma\cdot R) \,:\, R\in \mathrm{O}(k_1)\times\mathrm{O}(k_2-k_1)\times\cdots\times\mathrm{O}(n-k_r)\} \tag{4.6}$$

where $U = (u_1\, u_2\, \cdots\, u_n)$ for $u_1, \ldots, u_n$ an orthonormal eigenbasis of H and σ acts by permuting the rows of R. This algebraic set is a disjoint union of $\binom{n}{k_1, k_2-k_1, \ldots, n-k_r}$ varieties isomorphic to $\mathrm{O}(k_1) \times \mathrm{O}(k_2 - k_1) \times \cdots \times \mathrm{O}(n - k_r)$.

Proof. Similar to the proof of Theorem 4.1.2, a matrix $Q \in \mathrm{O}(n)$ is a critical point of the problem if and only if there exists a symmetric matrix M such that $HQS_0 = QM$. The condition that $Q^T HQS_0$ is symmetric is equivalent to requiring that $Q^T HQ$ and S_0 commute, which occurs precisely when $Q^T HQ$ is a block diagonal matrix with blocks of size $k_1, k_2 - k_1, \ldots, n - k_r$. If Q is in the disjoint union (4.6), then

$$Q^T HQ = (R^T \cdot \sigma^{-1})U^T HU(\sigma \cdot R) = R^T \Lambda R$$

where Λ is the diagonal matrix with eigenvalues of H on the diagonal. Since Λ is diagonal and R has the desired block structure, the points in the disjoint union are critical points.

Conversely, suppose $Q^T HQ$ has the desired block structure. Since H is diagonalizable, every block of $Q^T HQ$ is diagonalizable. Assembling these blocks gives a diagonalization $Q^T HQ = R\Lambda R^T$ where $R \in \mathrm{O}(k_1) \times \mathrm{O}(k_2 - k_1) \times \cdots \times \mathrm{O}(n - k_r)$, up to a permutation of the rows of R. But then $H = QR\Lambda R^T Q^T$, which implies that $Q = UR^T$, as desired. □

Proof of Theorem 4.1.5. The parametrization $\mathrm{SO}(n) \to \mathrm{Fl}_{\mathbf{c}}(\mathbf{k}; n)$ can naturally be extended to $\phi\colon \mathrm{O}(n) \to \mathrm{Fl}_{\mathbf{c}}(\mathbf{k}; n)$. The fiber of a point in $\mathrm{Fl}_{\mathbf{c}}(\mathbf{k}; n)$ is isomorphic to a copy of $\mathrm{O}(k_1) \times \cdots \times \mathrm{O}(n-k_r)$. Once we prove that the parametrization takes critical points to critical points, we will be finished as in the proof of Theorem 4.1.3. Indeed, we may apply Lemma 2.4.1, since the parametrization is surjective and $\mathrm{Fl}_{\mathbf{c}}(\mathbf{k}; n)$ is smooth by Theorem 3.5.2.

Since there is a transitive $\mathrm{O}(n)$-action on $\mathrm{O}(n)$, it suffices to compute the dimension of the image of the Jacobian evaluated at the tangent space of the identity matrix Id_n. For all $1 \le i < j \le n$ we have $e_{ij} - e_{ji} \in T_{\mathrm{Id}_n}(\mathrm{O}(n))$. These $\binom{n}{2}$ vectors are linearly independent.

We now compute $\mathcal{J}_\phi(\mathrm{Id}_n)$. This matrix has rows of the form $2c_i e_{ii}^T$ for all $i = 1, \ldots, n$ and of the form $c_j e_{ij}^T + c_i e_{ji}^T$ for $1 \le i < j \le n$. Therefore $\mathcal{J}_\phi(\mathrm{Id}_n) \cdot (e_{ij} - e_{ji})$ is zero when $c_i = c_j$ and otherwise has one nonzero entry. Since the nonzero entry has a different index for distinct pairs i, j, this process produces $\dim(\mathrm{Fl}_{\mathbf{c}}(\mathbf{k}; n))$ linearly independent vectors. □

Corollary 4.1.7. *The LO degree of the isospectral model $\mathrm{Fl}_{\mathbf{c}}(\mathbf{k}; n)$ is $\binom{n}{k_1, k_2-k_1, \ldots, n-k_r}$.*

We conclude with the observation that the degree of the multi-eigenvector problem over the isospectral flag variety is equal to its ED degree. Indeed, the Euclidean distance in the Frobenius norm is

$$||H - S||^2 = \mathrm{trace}((H - S)^2) = \mathrm{trace}(H^2) - 2\mathrm{trace(HS)} + \mathrm{trace(S)}^2.$$

Since both $\mathrm{trace}(H^2)$ and $\mathrm{trace}(S^2) = \mathrm{trace}(S_0^2)$ are fixed, minimizing the Euclidean distance is equivalent to maximizing the trace of HS.

Proposition 4.1.8 (Proposition 3.1, [82])**.** *The Euclidean distance degree of the isospectral model $\mathrm{Fl}_{\mathbf{c}}(\mathbf{k}; n)$ is $\binom{n}{k_1, k_2-k_1, \ldots, n-k_r}$.*

4.2 Heterogeneous Quadratics Minimization Problem

The heterogeneous quadratics minimization problem generalizes the multi-eigenvector problem considered in the previous section. Its most natural formulation is in Stiefel coordinates:

$$\min_{Z^T Z = \mathrm{Id}_k} \sum_{i=1}^{k} Z_i^T H_i Z_i \tag{4.7}$$

where $H_i \in \mathrm{Sym}^2 \mathbb{R}^n$ and Z_i denotes the ith column of Z for $i = 1, \ldots, k$. The multi-eigenvector problem is recovered by choosing $H_i = H$ of all $i = 1, \ldots, k$.

In Stiefel coordinates, the critical points can be computed using

$$\nabla \left(\sum_{i=1}^{k} Z_i^T H_i Z_i \right) = 2 \begin{pmatrix} Z_1^T H_1 & Z_2^T H_2 & \cdots & Z_k^T H_k \end{pmatrix}$$

and $\mathcal{J}_{\mathrm{St}(k,n)}(Z)$ as in (3.1). A more convenient system of equations whose solutions are the same critical points is the following (see [105, Lemma 1]):

$$\begin{pmatrix} H_1 Z_1 & H_2 Z_2 & \cdots & H_k Z_k \end{pmatrix} Z^T - Z \begin{pmatrix} H_1 Z_1 & H_2 Z_2 & \cdots & H_k Z_k \end{pmatrix}^T = 0 \quad \text{and} \quad Z^T Z = \mathrm{Id}_k.$$

The optimization problem is invariant under the action of $\mathrm{O}(1)^k$: if $(Z_1\, Z_2 \cdots Z_k)$ is a critical point, so is $(\pm Z_1 \pm Z_2 \cdots \pm Z_k)$. Computing the number of complex critical points of (4.7) is a challenging open problem. Our numerical computations, produced with `HomotopyContinuation.jl` [20], are summarized in Table 4.1.

$k \backslash n$	2	3	4	5	6	7	8	9
2	8	40	112	240	440	728	1 120	1 632
3		80	960	5 536	21 440	64 624		
4			1 920	57 216				

Table 4.1: Degrees of the heterogeneous quadratics minimization problem for small k, n.

Conjecture 4.2.1. The number of critical points of the heterogeneous quadratics minimization problem for $k = 2$ is $2\binom{2n}{3}$.

Diagonal case

Our computational experiments indicate that the number of critical points of the heterogeneous quadratics minimization problem stays stable if we take the input matrices $H_1, \ldots, H_k$ to be generic *diagonal* matrices. While we do not have a general proof for this observation, we present a result addressing the first nontrivial case.

Proposition 4.2.2. *Let* $H_1 = \operatorname{diag}(a_{11}, a_{12}, a_{13})$ *and* $H_2 = \operatorname{diag}(a_{21}, a_{22}, a_{23})$ *be generic diagonal matrices. Then the algebraic degree of the corresponding heterogeneous quadratics minimization problem* (4.7) *is* 40.

Proof. We will explicitly describe these 40 critical points. The critical points are defined by the Lagrange multiplier equations

$$H_1 Z_1 = q_{11} Z_1 + q_{12} Z_2,\ H_2 Z_2 = q_{12} Z_1 + q_{22} Z_2, \quad Z_1^T Z_1 = 1,\ Z_1^T Z_2 = 0,\ Z_2^T Z_2 = 1$$

where q_{11}, q_{12}, q_{22} are Lagrange multipliers. This is a square system with 9 variables and 9 equations. We obtain $2^2 \cdot 3 \cdot 2 = 24$ solutions by taking $Z_1^* = \pm e_i$, $Z_2^* = \pm e_j$, $q_{11}^* = a_{1i}$, $q_{12}^* = 0$, and $q_{22}^* = a_{2j}$ for all $i \neq j$. By computing a Gröbner basis of the ideal given by the above equations over the rational function field $\mathbb{Q}(a_{1j}, a_{2j} : j = 1, 2, 3)$ we see that this ideal is zero dimensional and has 40 solutions. The remaining 16 solutions come from flipping the signs of rows and columns of the Z^* whose rows we list below:

$$(Z^{*T})_1 = \frac{\sqrt{a_{12} - a_{13} + a_{23} - a_{22}}}{\alpha} \begin{pmatrix} \sqrt{-(a_{12} - a_{13})(a_{21} - a_{22})(a_{21} - a_{23})} \\ \sqrt{(a_{22} - a_{23})(a_{11} - a_{12})(a_{11} - a_{13})} \end{pmatrix}$$

$$(Z^{*T})_2 = \frac{\sqrt{a_{11} - a_{13} + a_{23} - a_{21}}}{\alpha} \begin{pmatrix} \sqrt{-(a_{11} - a_{13})(a_{22} - a_{21})(a_{22} - a_{23})} \\ \sqrt{(a_{21} - a_{23})(a_{12} - a_{11})(a_{12} - a_{13})} \end{pmatrix}$$

$$(Z^{*T})_3 = \frac{\sqrt{a_{11} - a_{12} + a_{22} - a_{21}}}{\alpha} \begin{pmatrix} \sqrt{-(a_{11} - a_{12})(a_{23} - a_{21})(a_{23} - a_{22})} \\ \sqrt{(a_{21} - a_{22})(a_{13} - a_{11})(a_{13} - a_{12})} \end{pmatrix}$$

Here $\alpha = a_{11}a_{22} - a_{11}a_{23} + a_{12}a_{23} - a_{12}a_{21} + a_{13}a_{21} - a_{13}a_{22}$. The corresponding Lagrange multipliers are

$$q_{11}^* = \frac{1}{\alpha}\big(- a_{11}a_{12}(a_{21} - a_{22}) + a_{11}a_{13}(a_{21} - a_{23}) - a_{12}a_{13}(a_{22} - a_{23})\big)$$

$$q_{12}^* = \frac{2}{\alpha}\sqrt{-(a_{11} - a_{12})(a_{11} - a_{13})(a_{12} - a_{13})(a_{21} - a_{22})(a_{21} - a_{23})(a_{22} - a_{23})}$$

$$q_{22}^* = \frac{1}{\alpha}\big(a_{21}a_{22}(a_{11} - a_{12}) - a_{21}a_{23}(a_{11} - a_{13}) + a_{22}a_{23}(a_{12} - a_{13})\big).$$

These computations were performed in `Oscar.jl` [102]. □

The above discussion implies that critical ideal in the sense of Corollary 2.2.4 is not prime. For generic data, it has twenty five primary components: twenty four of degree 1 and one of degree 16. The results on primeness fail because the map $(H_1, H_2) \mapsto (H_1 Z_1, H_2 Z_2)$ is not surjective onto $\mathbb{C}^6$ for every $Z \in \operatorname{St}(2, 3)$, e.g., $Z_1 = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}^T$ and $Z_2 = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix}^T$.

Projection coordinates

We now formulate the heterogeneous quadratics minimization problem as an optimization problem over a flag variety in projection coordinates. We rewrite the objective function as

$$\sum_{i=1}^{n} Z_i^T H_i Z_i = \sum_{i=1}^{n} \operatorname{trace}(Z_i^T H_i Z_i) = \sum_{i=1}^{k} \operatorname{trace}(G_i P_i)$$

where $P_i = \sum_{j=1}^{i} Z_j Z_j^T$ for $i = 1, \dots, k$ and $G_i = H_i - H_{i+1}$ for $i = 1, \dots, k-1$ and $G_k = H_k$. Over $\mathrm{pFl}(1, 2, \dots, k; n)$ we can reformulate our problem as follows:

$$\begin{aligned} \text{minimize } & \sum_{i=1}^{k} \operatorname{trace}(G_i P_i) \\ \text{subject to } & P_i P_j = P_j, \ \operatorname{trace}(P_i) = i \text{ for } 1 \leq j \leq i \leq k. \end{aligned} \tag{4.8}$$

Proposition 4.2.3. *If the heterogeneous quadratics minimization problem (4.8) has m critical points, then (4.7) has $2^k m$ critical points in $\mathrm{St}(k, n)$. Both problems have finite algebraic degree.*

Proof. The map from the Stiefel formulation of the flag variety $\mathrm{Fl}(1, 2, \dots, k; n)$ to the projection formulation $\mathrm{pFl}(1, 2, \dots, k; n)$ given by $Z \mapsto (P_1, \dots, P_k)$ where $P_i = \sum_{j=1}^{i} Z_j Z_j^T$ is 2^k to 1 since $(\pm Z_1 \pm Z_2 \cdots \pm Z_k)$ map to the same point. We proceed as in the proof of Theorem 4.1.3. A basis for the tangent space of $\mathrm{St}(k, n)$ at $Z = (e_1 \cdots e_k)$ consists of $k(n-k)$ standard unit vectors e_{ij} for $i = k+1, \dots, n$ and $j = 1, \dots, k$, and the vectors $e_{ij} - e_{ji}$ for $1 \leq i < j \leq k$. The Jacobian of the above 2^k-to-1 parametrization map consists of a stack of k Jacobians where each individual Jacobian is the Jacobian of the parametrization map from $\mathrm{St}(j, n)$ for $j = 1, \dots, k$ as we computed in the proof of Theorem 4.1.3. A careful computation shows that the images of these $k(n-k) + \binom{k}{2}$ vectors under the Jacobian of the parametrization map stay linearly independent. Hence, this image has dimension $\dim(\mathrm{pFl}(1, 2, \dots, k; n))$. Therefore, the differential of the parametrization map is surjective. Lemma 2.4.1 implies that the critical points on $\mathrm{St}(k, n)$ are mapped to the critical points on $\mathrm{pFl}(1, 2, \dots, k; n)$.

To see that the algebraic degree is finite, we observe that the gradient vector of the objective function consists of the entries in the data matrices, and hence the map in Assumption 2.2.1 is surjective. Finiteness then follows from Theorem 2.2.2. □

Corollary 4.2.4. *The degree of the heterogeneous quadratics minimization problem over the flag variety $\mathrm{pFl}(1, 2, \dots, k; n)$ is equal to the LO degree of $\mathrm{pFl}(1, 2, \dots, k; n)$.*

4.3 Two Problems from Statistics

In this section, we discuss two optimization problems from statistics which involve flags, namely *canonical correlation analysis* and *correspondence analysis*. Our goal is to describe and count the number of complex critical points of these optimization problems. The formulations are taken from [131, Section 1.2].

Canonical correlation analysis

Canonical correlation analysis is a statistical technique for detecting common features in a pair of data sets [94]. The problem is formulated as follows. Let $X \in \mathbb{R}^{n\times p}$ and $Y \in \mathbb{R}^{n\times q}$ be data matrices normalized to have sample mean 0. Here $n \geq p, q$ is the common sample size. Let $S_X = X^T X$ and $S_Y = Y^T Y$ denote the sample covariance matrices, and let $S_{XY} = X^T Y$ denote the sample cross-covariance. The kth pair $(a_k, b_k) \in \mathbb{R}^p \times \mathbb{R}^q$ of canonical correlation loadings is defined inductively

$$\begin{aligned}(a_k, b_k) = \operatorname{argmax}\{&a^T S_{XY} b \colon a^T S_X a = b^T S_Y b = 1,\\ &a^T S_X a_j = a^T S_{XY} b_j = b^T S_{YX} a_j = b^T S_Y b_j = 0,\ j = 1, \ldots, k-1\}.\end{aligned}$$

We perform a standard simplification via the Cholesky factorization $S_X = P^T P$ and $S_Y = Q^T Q$ where $P \in \mathbb{R}^{p\times p}$ and $Q \in \mathbb{R}^{q\times q}$ are upper triangular. We substitute $H = P^{-T} S_{XY} Q^{-1}$, $u = Pa$, and $v = Qb$ to obtain the simpler problem

$$\begin{aligned}(u_k, v_k) = \operatorname{argmax}\{&u^T H v \colon u^T u = v^T v = 1,\\ &u^T u_j = u^T H v_j = v^T H^T u_j = v^T v_j = 0,\ j = 1, \ldots, k-1\}.\end{aligned}$$

If we collect $u_1, \ldots, u_k$ and $v_1, \ldots, v_k$ into the $p \times k$ matrix U and $q \times k$ matrix V, respectively, then (U, V) represents a pair of flags in $\mathrm{Fl}(1, 2, \ldots, k; p) \times \mathrm{Fl}(1, 2, \ldots, k; q)$ [131].

We say the pair $(U, V) \in \mathbb{C}^{p\times k} \times \mathbb{C}^{q\times k}$ is a critical point of the canonical correlation problem if for all $i = 1, \ldots, k$, the pair (u_i, v_i) is a critical point of the optimization problem

$$\begin{aligned}\text{maximize } & u^T H v\\ \text{subject to } & u^T u = v^T v = 1\\ & u^T u_j = u^T H v_j = 0,\ j = 1, \ldots, i-1\\ & v^T H^T u_j = v^T v_j = 0,\ j = 1, \ldots, i-1.\end{aligned} \tag{4.9}$$

There are two quadric constraints, and $4(i-1)$ linear constraints on u and v. The critical points are given by the singular value decomposition of H [31].

Theorem 4.3.1. *The critical points of the canonical correlation analysis problem are the pairs $(U, V) \in \mathbb{C}^{p\times k} \times \mathbb{C}^{q\times k}$ where the columns of U are the left singular unit vectors of H, the columns of V are the right singular unit vectors of H, and corresponding columns have the same singular value. There are $\binom{\min(p,q)}{k} k! 2^k$ critical points.*

Proof. The count follows from the description of the critical points: choose k unit left and right singular vectors of H, permute the corresponding vectors simultaneously, and flip signs as desired. We will prove by induction that for each i, there exist λ_i, η_i such that $Hv_i = \lambda_i u_i$ and $H^T u_i = \eta_i v_i$ if and only if the pair (u_i, v_i) is a critical point of (4.9).

We proceed by induction on i. When $i = 1$, the optimization problem simplifies to $\max_{u^T u = v^T v = 1} v^T H u$. By computing the transpose of the Jacobian, we find that the critical points of this problem are characterized by the leftmost column of the matrix

$$\begin{pmatrix} Hv & u & 0 \\ H^T u & 0 & v \end{pmatrix}$$

being in the span of the rest of the columns. Therefore (u, v) is a critical point if and only if there exist λ_1, η_1 such that $Hv = \lambda_1 u$ and $H^T u = \eta_1 v$.

Suppose now that for $j = 1, \ldots, i-1$, we have that $Hv_j = \lambda_j u_j$ and $H^T u_j = \eta_j v_j$. The transpose of the augmented Jacobian matrix of (4.9) is

$$\begin{aligned} &\begin{pmatrix} Hv_i & u_1 & \cdots & u_i & 0 & \cdots & 0 & Hv_1 & \cdots & Hv_{i-1} & 0 & \cdots & 0 \\ H^T u_i & 0 & \cdots & 0 & v_1 & \cdots & v_i & 0 & \cdots & 0 & H^T u_1 & \cdots & H^T u_{i-1} \end{pmatrix} \\ = &\begin{pmatrix} Hv_i & u_1 & \cdots & u_i & 0 & \cdots & 0 & \lambda_1 u_1 & \cdots & \lambda_{i-1} u_{i-1} & 0 & \cdots & 0 \\ H^T u_i & 0 & \cdots & 0 & v_1 & \cdots & v_i & 0 & \cdots & 0 & \eta_1 v_1 & \cdots & \eta_{i-1} v_{i-1} \end{pmatrix} \end{aligned}$$

Proving that this matrix drops rank is equivalent to proving that the matrix

$$\begin{pmatrix} Hv_i & u_1 & \cdots & u_i & 0 & \cdots & 0 \\ H^T u_i & 0 & \cdots & 0 & v_1 & \cdots & v_i \end{pmatrix}$$

drops rank. It is clear that the matrix drops rank if $Hv_i \in \operatorname{span}(u_i)$ and $H^T u_i \in \operatorname{span}(v_i)$. Conversely, suppose that

$$\begin{aligned} Hv_i &= \alpha_1 u_1 + \cdots + \alpha_i u_i, \\ H^T u_i &= \beta_1 v_1 + \cdots + \beta_i v_i. \end{aligned}$$

Multiplying the first equation by u_j^T for $j = 1, \ldots, i-1$ gives $0 = u_j^T H v_i = \alpha_j$. By symmetry, $\alpha_j = \beta_j = 0$ for $j = 1, \ldots, i-1$. Thus $Hv_i = \alpha_i u_i$ and $H^T u_i = \beta_i v_i$. □

Correspondence Analysis

Correspondence analysis (CA) is an analog of principal component analysis for categorical data. The data for CA come in the form of an $n \times p$ matrix X known as a contingency table. We let $\mathbf{1}$ be the all-ones vector of appropriate size and set $t = \mathbf{1}^T X \mathbf{1} \in \mathbb{R}$. The row and column weights are defined as

$$\begin{aligned} r &= \frac{1}{t} X \mathbf{1} \in \mathbb{R}^n & \qquad c &= \frac{1}{t} X^T \mathbf{1} \in \mathbb{R}^p \\ D_r &= \frac{1}{t} \operatorname{diag}(r) \in \mathbb{R}^{n \times n} & \qquad D_c &= \frac{1}{t} \operatorname{diag}(c) \in \mathbb{R}^{p \times p}. \end{aligned}$$

For $k = 1, \dots, p$, we seek a pair of matrices $(U_k, V_k) \in \mathbb{R}^{n\times k} \times \mathbb{R}^{p\times k}$ such that

$$(U_k, V_k) = \operatorname{argmax}\{\operatorname{trace}(U^T(\frac{1}{t}X - rc^T)V) \colon U^T D_r U = V^T D_c V = \mathrm{Id}_k\}.$$

We begin with two simplifications of this problem. The first is standard in correspondence analysis: since D_r, D_c are diagonal, they can be factored into U, V, respectively, by replacing U with $\sqrt{D_r}U$ and V with $\sqrt{D_c}V$. The second is to replace the matrix $\frac{1}{t}X - rc^T$ from statistics with a generic matrix $H \in \mathbb{R}^{n\times p}$. The new optimization problem is

$$\max_{U^TU=V^TV=\mathrm{Id}_k} \operatorname{trace}(U^T H V). \tag{4.10}$$

The solution to this problem is given by the singular value decomposition of H; see [31]. The problem is invariant under a simultaneous $\mathrm{O}(k)$-action on U and V.

Theorem 4.3.2. *Let H be a generic real $n \times p$ matrix with $p > n$, let U be an $n \times k$ variable matrix and let V be a $p \times k$ variable matrix. The algebraic set of complex critical points of the optimization problem (4.10) is equal to*

$$\bigsqcup_{\{i_1,\dots,i_k\}\in\binom{[n]}{k}} \{((u_{i_1}\, u_{i_2}\, \cdots\, u_{i_k})Q,\ (v_{i_1}\, v_{i_2}\, \cdots\, v_{i_k})Q) \,:\, Q \in \mathrm{O}(k)\} \tag{4.11}$$

where $u_1, \dots, u_n$ is an orthonormal basis of left singular vectors for H and $v_1, \dots, v_n$ is an orthogonal basis of right singular vectors for H such that u_i, v_i share a common singular value for $i = 1, \dots, n$. This algebraic set is a disjoint union of $\binom{n}{k}$ varieties isomorphic to $\mathrm{O}(k)$.

Proof. As in the proof of Theorem 4.1.2, a pair (U, V) is a critical point if and only it satisfies $U^TU = V^TV = \mathrm{Id}_k$ and there exist symmetric matrices M, N such that $HV = UM$ and $H^TU = VN$. We must have $M = N$, since $U^THV = U^TUM = M$ and $V^TH^TU = V^TVN = N$ and M, N are symmetric. Thus (U, V) is a critical point if and only if $U^TU = V^TV = \mathrm{Id}_k$ and there exists a single symmetric matrix M with $HV = UM$ and $H^TU = VM$. The matrix M satisfies $HH^TU = UM^2$ where HH^T is full rank so by Lemma 4.1.1, M^2 is orthogonally diagonalizable, which implies M is orthogonally diagonalizable. We write $M = Q^T\Lambda Q$ for the spectral decomposition of M. Then the constraints become $HVQ^T = UQ^T\Lambda$ and $H^TUQ^T = VQ^T\Lambda$ which is precisely what it means for (UQ^T, VQ^T) to be in the set (4.11). □

The following problem is equivalent to (4.10):

$$\max_{MM^T\in\mathrm{pGr}(k,p)} \operatorname{trace}(HM) \tag{4.12}$$

where we use the map $(U, V) \mapsto M = VU^T \in \mathbb{R}^{p\times n}$. Note that if a matrix $M \in \mathbb{R}^{n\times p}$ is such that MM^T is in $\mathrm{Gr}(k, p)$, then the singular values of M consist of k ones. Therefore M has a truncated singular value decomposition $M = V\mathrm{Id}_kU^T$ where $(U, V) \in \mathrm{St}(k, n) \times \mathrm{St}(k, p)$.

Corollary 4.3.3. *The algebraic degree of* (4.12) *is* $\binom{n}{k}$. *The critical points are matrices*

$$M = (v_{i_1} v_{i_2} \cdots v_{i_k})(u_{i_1} u_{i_2} \cdots u_{i_k})^T$$

where $u_1, \ldots, u_n$ *is an orthonormal basis of left singular vectors for* H *and* $v_1, \ldots, v_n$ *is an orthogonal basis of right singular vectors for* H *such that* u_i, v_i *share a common singular value for* $i = 1, \ldots, n$.

Proof. We proceed using Lemma 2.4.1. Because there is a transitive $\mathrm{O}(n) \times \mathrm{O}(p)$-action on the product of Stiefel manifolds $\mathrm{St}(k, n) \times \mathrm{St}(k, p)$, it suffices to consider the tangent space of $\mathrm{St}(k, n) \times \mathrm{St}(k, p)$ at the point $(U, V) = ((\mathrm{Id}_k \ 0_{k \times n-k})^T, (\mathrm{Id}_k \ 0_{k \times p-k})^T)$. By computing the Jacobian, one verifies that this tangent space contains the linearly independent set

$$\{e_{ij} : i \in \{k+1, \ldots, n\}, j \in [k]\} \cup \{f_{ij} : i \in \{k+1, \ldots, p\}, j \in [k]\} \cup \{e_{ij} - e_{ji} : i \neq j \in [k]\}$$

where $[k] = \{1, \ldots, k\}$ and $e_{ij}, f_{ij} \in \mathbb{R}^{nk+pk}$ are the indicator vectors for the variables u_{ij} and v_{ij}, respectively. One verifies by computation that these $nk + pk - k^2 - \binom{k+1}{2}$ vectors remain linearly independent after applying the differential. The dimension of the image of the map $(U, V) \mapsto VU^T$ is the dimension of the domain $\mathrm{St}(k, n) \times \mathrm{St}(k, p)$ minus the dimension of the fibers each of which is a copy of $\mathrm{O}(k)$; this dimension matches the number of linearly independent vectors, so the differential is surjective. By Lemma 2.4.1, the critical points of (4.12) are precisely the images of the critical points of (4.10) under the map $(U, V) \mapsto VU^T$ and the statement follows from Theorem 4.3.2. □

Chapter 5

Constrained Rayleigh Quotient Optimization

In this chapter, our main focus is the minimization problem

$$\min_{\psi \in \mathcal{M}_{\mathbb{R}}} \frac{\psi^T H \psi}{\psi^T \psi} \tag{5.1}$$

of the Rayleigh quotient on a projective variety $\mathcal{M} \subseteq \mathbb{P}^{n-1}$. This is inspired by a problem from quantum chemistry, where H is a Hamiltonian and ψ is constrained to lie in a given tensor train variety. We begin by giving some background on this problem.

Since its derivation in 1926 by Erwin Schrödinger, the electronic Schrödinger equation has been studied extensively, mainly in order to predict molecular bindings between atoms or the structure of larger molecules. In numerical mathematics and theoretical physics/chemistry, the main task is to compute the ground state energy of the equation, that is, the smallest eigenvalue of the Hamiltonian operator. While this might seem straightforward, the problem size scales exponentially with the number of electrons involved, and therefore, standard eigensolvers fail even for small systems of 10-20 electrons. Before the advent of cryptocurrencies and AI, computations aiming at solving the Schrödinger problem commonly used about a third of the computation power of some large supercomputers [61].

Many approximation methods for energy minimization of the electronic Schrödinger equation have been proposed by physicists and mathematicians, among them the *Hartree-Fock* method, which treats the involved electrons as uncorrelated, and *density functional theory*, where the two-particle term of the Hamiltonian is replaced by a one-particle potential. See [90] for a mathematical introduction into both. For correlated many-body systems, widely used methods are (1) the *coupled cluster* method [44, 45, 62], where a surrogate problem is solved on a lower-dimensional algebraic variety, and (2) the *density matrix renormalization group* (DMRG) algorithm, which approximates the eigenstates by a low-rank tensor [123]. The latter is the perspective taken in this article. It was proposed by theoretical physicists along with *matrix product states* to approximate solutions of the electronic Schrödinger equation in second quantization [128]. The realization that this approach is actually a low-rank tensor method emerged slowly and separately for physicists [114] and mathematicians [103]. In mathematics, the matrix product states have been dubbed *tensor trains* (TT) corresponding to the tensor train rank (TT rank) of a tensor, and the DMRG algorithm has been recognized as a two-site version of the *alternating linear scheme* (ALS) [65]. However, while this method often yields accurate results in practice (empirically validated using chemical measurements), theoretical guarantees for convergence and accuracy are lacking. In particular, the DMRG algorithm converges at best to local minima of the energy function. The number of such local minima on the set of low-rank tensors as well as their approximation quality is unknown. We pursue an understanding of the critical points of (5.1) with this motivation in mind.

We study the constrained Rayleigh quotient optimization problem (5.1) from the perspective of Chapter 2. We will define the *Rayleigh–Ritz degree* (RR degree) of a projective variety in Section 5.1 as the algebraic degree of this optimization problem. Here, we will present various ways of computing these critical points ((5.2), Proposition 5.1.2, and (5.10). In Section 5.2 we define the RR discriminant, which is the set of symmetric matrices H for which the number of isolated critical points is less than the RR degree.

In Section 5.3, we will introduce tensor train varieties and identify instances when they are Segre products of projective spaces (Theorem 5.3.6). We also report what we know about the defining ideals of tensor trains. A parametrization of TT varieties from a product of Grassmannians will be presented in Section 5.4. From this parametrization, we obtain the formula for the dimension of a TT variety (Corollary 5.4.5).

We present our numerical experiments in Section 5.5. We compute all critical points of our optimization problem for small cases using the numerical algebraic geometry package `HomotopyContinuation.jl` [20]. We use this to test the correctness of the ALS and DMRG algorithms on these examples. In small cases, we observe that the ALS method can converge to any of the local minima and that the corresponding energy value can be far from optimal.

5.1 Rayleigh-Ritz Optimization

We now introduce the constrained Rayleigh quotient optimization problem. Given a real symmetric matrix $H \in \mathrm{Sym}^2\mathbb{R}^n$ and a complex projective algebraic variety $\mathcal{M} \subseteq \mathbb{P}^{n-1}$ not contained in the isotropic quadric $V(\psi^T\psi)$, we seek to solve the optimization problem (5.1). We will study this problem using the tools from Chapter 2, and we adopt the notation of that chapter. The problem (5.1) is a relaxation of the problem "Find eigenvectors of H in a variety $\mathcal{M}$," which is over-constrained. The set of matrices having eigenvalues in $\mathcal{M}$ is the *Kalman variety* of $\mathcal{M}$; see [113].

When $\mathcal{M} = \mathbb{P}^{n-1}$ as in (4.1), the optimal value for (5.1) is the smallest eigenvalue of H, and a corresponding eigenvector is an optimal solution. When $\mathcal{M}$ is a linear space, then (5.1) is a *Rayleigh-Ritz optimization problem* over $\mathcal{M}$ [125]. We maintain this terminology in the nonlinear case, and we refer to the algebraic degree of (5.1) as the *Rayleigh-Ritz degree* (RR degree). We saw our first nonlinear example of (5.1) in Example 2.1.3.

Up to scaling, the gradient of the objective function is $\psi^T H$ and the gradient of $\psi^T\psi$ is ψ^T. Hence the augmented Jacobian (2.5) is

$$\mathcal{J}^H_{\mathcal{M}}(\psi) = \begin{pmatrix} \psi^T H \\ \psi^T \\ \mathcal{J}_{\mathcal{M}}(\psi) \end{pmatrix}.$$

We refer to the critical correspondence of (5.1) as the *RR correspondence*:

$$\mathcal{C}_{\mathcal{M}} = \overline{\{(\psi, H) \in \mathcal{M}_{\text{reg}} \backslash V(\psi^T\psi) : \operatorname{rank} \mathcal{J}^H_{\mathcal{M}}(\psi) \le c+2\}}.$$

Proposition 5.1.1. *Theorem 2.2.2 holds for the RR problem. In particular, every model $\mathcal{M} \subseteq \mathbb{P}^{n-1}$ that is not contained in the isotropic quadric $V(\psi^T\psi)$ has finite RR degree.*

Proof. The map $H \mapsto H\psi$ is surjective for any $\psi \in \mathbb{P}^{n-1}$. Since Assumption 2.2.1 holds, the result follows from Theorem 2.2.2. □

By Corollary 2.2.4, the critical ideal for a given $H \in \operatorname{Sym}^2\mathbb{R}$ is

$$(I_{\mathcal{M}} + \langle (c+2)\times(c+2) \text{ minors of } \mathcal{J}^H_{\mathcal{M}}(\psi)\rangle) : (I_{\mathcal{M}_{\text{sing}}} \cdot \psi^T\psi)^{\infty}. \tag{5.2}$$

Recall that when $\mathcal{M} \subseteq \mathbb{P}^{n-1}$ admits a rational δ-to-one parametrization $\psi : \mathbb{C}^s \dashrightarrow \mathcal{M}$, we can consider the unconstrained RR optimization problem

$$\min_{t\in\mathbb{R}^s} = \frac{\psi(t)^T H \psi(t)}{\psi(t)^T\psi(t)}, \tag{5.3}$$

which has δ times the RR degree many critical points for generic H. Note that this includes tensor train varieties; see Algorithm 2. By Lemma 2.4.1, a point $t \in \psi^{-1}(\mathcal{M}_{\text{reg}}\backslash V(\psi^T\psi))$ with $\operatorname{rank}(\mathcal{J}_\psi(t)) = s$ (2.14) is a critical point of (5.3) if and only if $\psi(t)$ is a critical point of (5.1). Recalling the notation of Chapter 2, we write $\mathcal{J}_\psi(t)$ for the $n\times s$ Jacobian of ψ and $\mathcal{V}_{\phi,s}$ for the locus where this matrix has rank less than s. We assume that $\mathcal{V}_{\phi,s} \subsetneq \mathbb{C}^s$. From Corollary 2.4.5 and (2.18), we have the following parametric description of the critical locus and ideal.

Proposition 5.1.2. *Let $\psi : \mathbb{C}^s \dashrightarrow \mathcal{M}$ be a δ-to-one parametrization of $\mathcal{M} \subseteq \mathbb{P}^{n-1}$ with base locus $\mathcal{B}$. The parametric critical locus is*

$$\overline{\{t \in \mathbb{C}^s\backslash(\mathcal{B}\cup\mathcal{V}_{\psi,s}\cup V(\psi^T\psi) : \operatorname{rank} \begin{pmatrix}\psi(t) & \mathcal{J}_\psi(t)\end{pmatrix}^T \cdot \begin{pmatrix}\psi(t) & H\psi(t)\end{pmatrix} \le 1\}}$$

and the critical ideal is

$$\langle 2\times 2 \text{ minors of } \begin{pmatrix}\psi(t) & \mathcal{J}_\psi(t)\end{pmatrix}^T \cdot \begin{pmatrix}\psi(t) & H\psi(t)\end{pmatrix}\rangle$$
$$: (I_{\mathcal{B}} \cdot \langle (n-c)\times(n-c) \text{ minors of } \mathcal{J}_\psi(t)\rangle \cdot \langle(\psi^T\psi)(t)\rangle)^{\infty}$$

where $I_{\mathcal{B}}$ is the ideal of the base locus $\mathcal{B}$.

Example 5.1.3 (Rational Quadric)**.** The RR correspondence of the rational quadric $V(\psi_1\psi_3 - \psi_2^2) \subseteq \mathbb{P}^2$ is a 5-dimensional subvariety of $\mathbb{P}^2 \times \mathbb{P}^5$ and is minimally generated by $\psi_1\psi_3 - \psi_2^2$ and the 3×3 determinant

$$\begin{vmatrix} h_{11}\psi_1 + h_{12}\psi_2 + h_{13}\psi_3 & h_{12}\psi_1 + h_{22}\psi_2 + h_{23}\psi_3 & h_{13}\psi_1 + h_{23}\psi_2 + h_{33}\psi_3 \\ \psi_1 & \psi_2 & \psi_3 \\ \psi_3 & -2\psi_2 & \psi_1 \end{vmatrix}.$$

This ideal is computed without saturation because the model is smooth and intersects the isotropic quadric $V(\psi^T\psi)$ transversely. Its degree is $2 \cdot 3 = 6$. The rational normal curve admits a parametrization $t \mapsto [1 : t : t^2]$. This parametrization has empty base locus $\mathcal{B} = \emptyset$ and its Jacobian $\mathcal{J}_\psi(t) = \begin{pmatrix} 0 & 1 & 2t \end{pmatrix}$ has rank 1 for all $t \in \mathbb{C}$. Therefore the parametric RR correspondence is the hypersurface in $\mathbb{C} \times \mathbb{P}^5$ that is cut out by the determinant of

$$\begin{pmatrix} h_{11} + h_{12}t + h_{13}t^2 & h_{12} + h_{22}t + h_{23}t^2 & h_{13} + h_{23}t + h_{33}t^2 \\ 1 & t & t^2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ t & 1 \\ t^2 & 2t \end{pmatrix}$$

which evaluates to

$$\begin{aligned} h_{12} + t(h_{11} + 2h_{13} + h_{22}) + t^2(-h_{12} + 3h_{23}) + t^3(2h_{11} + 2h_{33}) \\ + t^4(-3h_{12} + h_{23}) + t^5(-2h_{13} - h_{22} + h_{33}) - t^6 h_{23}. \end{aligned}$$

This shows that the RR degree of the rational quadric is six. ◇

Example 5.1.4 (Twisted Cubic)**.** Unlike the rational quadric, the twisted cubic $V(\psi_1\psi_3 - \psi_2^2, \psi_2\psi_4 - \psi_3^2, \psi_1\psi_4 - \psi_2\psi_3) \subseteq \mathbb{P}^3$ is not a complete intersection. It intersects the isotropic quadric transversely, so its RR correspondence is generated by its defining polynomials, along with the 4×4 minors of the augmented Jacobian

$$\begin{pmatrix} (H\psi)_1 & (H\psi)_2 & (H\psi)_3 & (H\psi)_4 \\ \psi_1 & \psi_2 & \psi_3 & \psi_4 \\ \psi_3 & -2\psi_2 & \psi_1 & 0 \\ 0 & \psi_4 & -2\psi_3 & \psi_2 \\ \psi_4 & -\psi_3 & -\psi_2 & \psi_1 \end{pmatrix}$$

Theorem 2.3.2(i) with $n = 4$, $c = 2$, and $e = d_1 = d_2 = 2$ gives an upper bound of 16 for the RR degree of the twisted cubic, but the true RR degree is 10. This is apparent from looking at the parametric RR correspondence given by the parametrization $t \mapsto [1 : t : t^2 : t^3]$. The parametric RR correspondence $\mathcal{P}_\mathcal{M} \subseteq \mathbb{C} \times \mathbb{P}^9$ is cut out by the 2×2 determinant:

$$\begin{vmatrix} \psi(t)^T H\psi(t) & (H\psi(t))_2 + 2t(H\psi(t))_3 + 3t^2(H\psi(t))_4 \\ 1 + t^2 + t^4 + t^6 & t + 2t^3 + 3t^5 \end{vmatrix}$$

which has degree 10. For the specific instance

$$H = \begin{pmatrix} 100 & 78 & 76 & 42 \\ 78 & 170 & 111 & 67 \\ 76 & 111 & 85 & 54 \\ 42 & 67 & 54 & 41 \end{pmatrix}, \tag{5.4}$$

it is equal to

$$78 + 222t + 381t^2 + 238t^3 + 189t^4 - 280t^5 - 381t^6 - 562t^7 - 405t^8 - 178t^9 - 54t^{10}. \qquad \diamond$$

One obtains the critical ideal by evaluating at specific values of H. Recall the definition of the Lagrangian ideal (2.13)

$$\Lambda_{\mathcal{M}}^H = (I_{\mathcal{M}} + \langle (c+2)\times(c+2) \text{ minors of } \mathcal{J}_{\mathcal{M}}^H(x)\rangle) : I_{\mathcal{M}_{\text{sing}}}^{\infty}.$$

The critical ideal is $I_{\mathcal{M}}^H = \Lambda_{\mathcal{M}}^H : (\psi^T\psi)^\infty$. By Proposition 2.3.1, the critical and Lagrangian ideals are equal for the rational quadric and cubic curves, because these two varieties intersect the isotropic quadric $V(\psi^T\psi)$ transversely. We now see an example where the critical and Lagrangian ideals are not equal.

Example 5.1.5 ($\mathbb{P}^1\times\mathbb{P}^1$)**.** The simplest tensor train variety (see Section 5.3) is $\mathcal{M} = \mathbb{P}^1\times\mathbb{P}^1$. Here, $I_{\mathcal{M}} = \langle \psi_{00}\psi_{11} - \psi_{01}\psi_{10}\rangle \subset \mathbb{R}[\psi_{00},\psi_{01},\psi_{10},\psi_{11}]$. The augmented Jacobian is

$$\begin{pmatrix} \psi_{00} & \psi_{01} & \psi_{10} & \psi_{11} \\ (H\psi)_{00} & (H\psi)_{01} & (H\psi)_{10} & (H\psi)_{11} \\ \psi_{11} & -\psi_{10} & -\psi_{01} & \psi_{00} \end{pmatrix}.$$

The Lagrangian ideal (2.13), generated by the 3×3 minors of the augmented Jacobian along with $\psi_{00}\psi_{11} - \psi_{01}\psi_{10}$, is a zero-dimensional ideal of degree 12. This agrees with the upper bound on the RR degree given by Corollary 2.3.3(i). Its primary decomposition has three components over $\mathbb{Q}$:

$$I_{\mathcal{M}}^H \cap \langle \psi_{01}-\psi_{10}, \psi_{00}+\psi_{11}, \psi_{10}^2+\psi_{11}^2\rangle \cap \langle \psi_{01}+\psi_{10}, \psi_{00}-\psi_{11}, \psi_{10}^2+\psi_{11}^2\rangle$$

where the critical ideal $I_{\mathcal{M}}^H$ is

$$\begin{aligned} I_{\mathcal{M}}+\langle & 3307\psi_{00}\psi_{01} + 5577\psi_{01}^2 - 6399\psi_{00}\psi_{10} - 5016\psi_{10}^2 - 1188\psi_{00}\psi_{11} + 783\psi_{01}\psi_{11} - 4295\psi_{10}\psi_{11} - 561\psi_{11}^2, \\ & 6614\psi_{00}^2 - 16624\psi_{01}^2 + 24459\psi_{00}\psi_{10} + 13582\psi_{10}^2 + 606\psi_{00}\psi_{11} - 14379\psi_{01}\psi_{11} + 3978\psi_{10}\psi_{11} - 3572\psi_{11}^2\rangle \end{aligned}$$

for the data matrix (5.4). By Proposition 2.3.1, the extraneous points in the Lagrangian locus are precisely the transverse intersection points of $\mathbb{P}^1\times\mathbb{P}^1$ with $V(\psi^T\psi)$. The Lagrangian locus has four extraneous isotropic critical points that are not critical:

$$[-1 : i : i : 1],\ [-1 : -i : -i : 1],\ [1 : i : -i : 1],\ [1 : -i : i : 1]. \tag{5.5}$$

This critical ideal $I_{\mathcal{M}}^H$ is a zero-dimensional complete intersection and hence it has degree 8. Thus that RR degree of $\mathbb{P}^1 \times \mathbb{P}^1$ is 8. Of the eight critical points, six are real, listed below as unit vectors, with the underlined one attaining the global minimum value 4.66885:

$$\begin{array}{ll}
(-0.05582, -0.43762, 0.11355, 0.89020) & (-0.49237, -0.35933, 0.64035, 0.46732) \\
(-0.82393, 0.56305, -0.05284, 0.03611) & (0.55151, 0.59762, 0.39467, 0.42767) \\
(-0.44158, 0.21247, 0.78549, -0.37795) & \underline{(-0.01170, 0.01709, -0.56495, 0.82486).}
\end{array}$$

We parametrize $\mathbb{P}^1 \times \mathbb{P}^1 \subseteq \mathbb{P}^3$ by $(t_1, t_2) \mapsto [1 : t_1 : t_2 : t_1t_2]$. The base locus of this parametrization is empty, and its Jacobian has rank 2 for all (t_1, t_2). The parametric RR correspondence is defined by

$$\left\langle 2 \times 2 \text{ minors of } \begin{pmatrix} 1 & t_1 & t_2 & t_1t_2 \\ 0 & 1 & 0 & t_2 \\ 0 & 0 & 1 & t_1 \end{pmatrix} \begin{pmatrix} 1 & h_{11} + h_{12}t_1 + h_{13}t_2 + h_{14}t_1t_2 \\ t_1 & h_{12} + h_{22}t_1 + h_{23}t_2 + h_{24}t_1t_2 \\ t_2 & h_{13} + h_{23}t_1 + h_{33}t_2 + h_{34}t_1t_2 \\ t_1t_2 & h_{14} + h_{24}t_1 + h_{34}t_2 + h_{44}t_1t_2 \end{pmatrix} \right\rangle : \langle 1 + t_1^2 + t_2^2 + t_1^2t_2^2 \rangle.$$

The excess component of the determinantal ideal is cut out by $\langle 1 + t_1^2, 1 + t_2^2 \rangle$ and consists of the four points $t_1 = \pm i$ and $t_2 = \pm i$. These are precisely the four points we saw in (5.5). $\diamond$

The RR degrees of Segre-Veronese varieties were recently derived in [112], where, in particular, explicit formulae for the RR degrees of rank-one matrices of any size and of rank-one binary tensors were obtained.

Proposition 5.1.6 ([112, Corollaries 5.2 and 5.3])**.** *The RR degree of the variety of rank-one binary tensors of order n ($\mathcal{M} = (\mathbb{P}^1)^n$) is $2^n n!$. The RR degree of the variety of rank-one $n \times m$ matrices ($\mathcal{M} = \mathbb{P}^{n-1} \times \mathbb{P}^{m-1}$) with $n \leq m$ is $\sum_{i=1}^n 4^{i-1} \binom{n}{i} \binom{m}{i}$.*

In the remainder of this section, we present an alternative perspective on the optimization problem (5.1). For this, we will consider $\mathcal{M}$ in its second Veronese embedding $\nu_2(\mathcal{M})$ and transform the Rayleigh-Ritz optimization problem (5.1) to a distance optimization problem using the *Bombieri-Weyl* distance. The equivalence of these two optimization problems has been introduced and fully developed in [112, Section 2]. We give a very brief introduction to the Bombieri-Weyl distance optimization.

Note that the Rayleigh-Ritz problem (5.1) is equivalent to the optimization problem

$$\min_{\psi \in C(\mathcal{M}_{\mathbb{R}}) \cap S^{n-1}} \psi^T H \psi \tag{5.6}$$

where $C(\mathcal{M}_{\mathbb{R}}) \subset \mathbb{R}^n$ is the cone over the projective variety $\mathcal{M} \subset \mathbb{P}^{n-1}$ and S^{n-1} is the $(n-1)$-dimensional unit sphere in $\mathbb{R}^n$. We also point out that each critical point ψ of (5.1) corresponds to two critical points $\pm \frac{1}{(\psi)^T \psi} \psi$ of (5.6).

A homogeneous form f of degree d in n variables with real coefficients can be represented by an element of $\mathrm{Sym}^d \mathbb{R}^n$. This representation as a symmetric tensor can be written as

$$f(x) = \sum_{|\alpha|=d} \binom{d}{\alpha} f_\alpha x^\alpha$$

where $\binom{d}{\alpha}$ is the multinomial coefficient $\frac{d!}{\alpha_1!\cdots\alpha_n!}$. Hence f is identified with $(f_\alpha)_{|\alpha|=d}$. The case we will work with is $d = 2$. A homogeneous quadratic polynomial in n variables and the corresponding symmetric matrix of this quadratic form are

$$\sum_{i=1}^{n} h_{ii}x_i^2 + \sum_{i<j} 2h_{ij}x_ix_j \qquad H = \begin{pmatrix} h_{11} & h_{12} & \cdots & h_{1n} \\ h_{12} & h_{22} & \cdots & h_{2n} \\ \vdots & & \ddots & \vdots \\ h_{1n} & h_{2n} & \cdots & h_{nn} \end{pmatrix}. \tag{5.7}$$

From now on we will identify a symmetric matrix H with the quadratic form $f_H = x^T Hx$. The *Bombieri-Weyl inner product* between two homogeneous forms $f, g \in \mathbb{R}[x_1, \ldots, x_n]_d$ is

$$\langle f, g\rangle_{\mathrm{BW}} := \sum_{|\alpha|=d} \binom{d}{\alpha} f_\alpha g_\alpha.$$

The corresponding *Bombieri-Weyl norm* of f is the square root of

$$||f||^2_{\mathrm{BW}} := \sum_{|\alpha|=d} \binom{d}{\alpha} f_\alpha^2.$$

If f is a quadratic polynomial as in (5.7) then $||f||^2_{\mathrm{BW}} = \mathrm{trace}(H^2)$.

Proposition 5.1.7 ([112, Proposition 2.6])**.** *Let $\mathcal{M} \subset \mathbb{P}^{n-1}$ be an irreducible variety. A point $\lambda\nu_2(\psi) \in C(\nu_2(\mathcal{M}))$ with $\psi \in S^{n-1}$ and $\lambda \in \mathbb{C}$ is a critical point of*

$$\min_{M\in C(\nu_2(\mathcal{M}))} \; ||f_H - M||^2_{\mathrm{BW}} \tag{5.8}$$

if and only if ψ is a critical point of (5.6) and $\lambda = f_H(\psi) = \psi^T H\psi$.

Example 5.1.8 ($\mathbb{P}^1 \times \mathbb{P}^1$)**.** In the case of $\mathbb{P}^1 \times \mathbb{P}^1$, the optimization problem (5.8) is

$$\min_{\substack{\psi_{00}\psi_{11}=\psi_{10}\psi_{01} \\ \psi_{00}^2+\psi_{01}^2+\psi_{10}^2+\psi_{11}^2=1}} \quad \mathrm{trace}((H - \nu_2(\psi))^2)$$

where

$$\nu_2(\psi) = \begin{pmatrix} \psi_{00}^2 & \psi_{00}\psi_{01} & \psi_{00}\psi_{10} & \psi_{00}\psi_{11} \\ \psi_{00}\psi_{01} & \psi_{01}^2 & \psi_{01}\psi_{10} & \psi_{01}\psi_{11} \\ \psi_{00}\psi_{10} & \psi_{01}\psi_{10} & \psi_{10}^2 & \psi_{10}\psi_{11} \\ \psi_{00}\psi_{11} & \psi_{01}\psi_{11} & \psi_{10}\psi_{11} & \psi_{11}^2 \end{pmatrix}.$$

The critical ideal is generated by $\psi_{00}\psi_{11} - \psi_{10}\psi_{01}$, $\psi_{00}^2 + \psi_{01}^2 + \psi_{10}^2 + \psi_{11}^2 - 1$, and the 3×3 minors of the 3×4 augmented Jacobian. No saturation is needed to compute this zero dimensional ideal of degree 16. This agrees with the RR degree 8 we computed in Example 5.1.5. ⋄

Note that the trace of $\nu_2(\psi)$ is $\operatorname{trace}(\nu_2(\psi)) = \psi_1^2 + \cdots + \psi_n^2 = q(\psi)$. Therefore constraining ψ to lie on the sphere is equivalent to constraining $\nu_2(\psi)$ to have trace 1. Therefore

$$\operatorname{trace}((H - \nu_2(\psi))^2) = \operatorname{trace}(H^2) - 2\operatorname{trace}(H\nu_2(\psi)) + 1$$

and hence maximizing the trace of $(H - \nu_2(\psi))^2$ is equivalent to minimizing $\operatorname{trace}(H\nu_2(\psi))$. We therefore consider the linear optimization problem

$$\max_{\substack{M \in C(\nu_2(\mathcal{M})) \\ \operatorname{trace}(M)=1}} \operatorname{trace}(MH). \tag{5.9}$$

We have shown that the Rayleigh-Ritz problem may be realized as a linear optimization problem over the Veronese square $\nu_2(\mathcal{M})$. This is reminiscent of the reformulations of the multi-eigenvector problem in Section 4.1.

Corollary 5.1.9. *The linear optimization degree of the affine variety* $C(\nu_2(\mathcal{M})) \cap V(\operatorname{trace}(M) - 1)$ *is twice the RR degree of* $\mathcal{M}$.

The linear problem (5.9) has an equivalent projective formulation:

$$\max_{M \in \nu_2(\mathcal{M}))} \frac{\operatorname{trace}(MH)}{\operatorname{trace}(M)}.$$

We will study the optimization of the same objective function over the projection Grassmannian in Chapter 6. By Corollary 2.2.4, the critical ideal of this optimization problem is

$$\left(I_{\nu_2(\mathcal{M})} + \left\langle (c+2) \times (c+2) \text{ minors of } \begin{pmatrix} h_{11} & 2h_{12} & 2h_{13} & \cdots & h_{nn} \\ 1 & 0 & 0 & \cdots & 1 \\ & & \mathcal{J}_{\nu_2(\mathcal{M})}(M) & & \end{pmatrix} \right\rangle \right) : \left(I_{\nu_2(\mathcal{M})_{\text{sing}}} \cdot \operatorname{trace}(M) \right)^{\infty}. \tag{5.10}$$

This provides another method of computing the critical equations of the Rayleigh-Ritz optimization problem.

5.2 The Rayleigh-Ritz Discriminant

We now consider the case of degenerate data, that is, of data matrices H for which the critical locus is badly behaved. This data locus is called the *Rayleigh-Ritz discriminant*. Analogous discriminants were defined in [77] for the maximum likelihood estimation problem and in [37] for the Euclidean distance optimization problem.

As in Section 5.1, we write $\mathcal{C}_{\mathcal{M}}$ for the critical correspondence of (5.1), which comes equipped with projections $\pi_1 : \mathcal{C}_{\mathcal{M}} \to \mathcal{M}$ onto the model and $\pi_2 : \mathcal{C}_{\mathcal{M}} \to \mathbb{P}(\operatorname{Sym}^2\mathbb{C}^n)$ onto the data. By Proposition 5.1.1, the fiber $\pi_2^{-1}(H)$ of a general matrix H has RR degree many points. By "general," we mean that this statement holds outside of a subvariety of the data space $\mathbb{P}(\mathrm{Sym}^2\mathbb{C}^n)$. We call this subvariety, which is typically a hypersurface, the *Rayleigh-Ritz discriminant* (RR discriminant):

$$\Sigma_{\mathcal{M}} = \overline{\{H \in \mathbb{P}(\mathrm{Sym}^2\mathbb{C}^n) : \text{ the critical locus of (5.1) is infinite or nonreduced}\}}.$$

We use the Jacobian criterion to determine when the critical locus of a matrix H is singular or infinite by treating the variables $h_{11}, h_{12}, \ldots, h_{nn}$ as parameters. Given the ideal $I_{\mathcal{C}_{\mathcal{M}}} = \langle g_1, \ldots, g_\ell \rangle \subseteq \mathbb{C}[\psi_1, \ldots, \psi_n, h_{11}, h_{12}, \ldots, h_{nn}]$ of the correspondence, a point in the critical locus is nonreduced or lies on a positive dimensional component precisely when the following $\ell \times n$ matrix drops rank:

$$\mathcal{J}_{\mathcal{C}_{\mathcal{M}},\psi}(\psi, H) = \begin{pmatrix} \frac{\partial g_1}{\partial \psi_1} & \frac{\partial g_1}{\partial \psi_2} & \cdots & \frac{\partial g_1}{\partial \psi_n} \\ \frac{\partial g_2}{\partial \psi_1} & \frac{\partial g_2}{\partial \psi_2} & \cdots & \frac{\partial g_2}{\partial \psi_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_\ell}{\partial \psi_1} & \frac{\partial g_\ell}{\partial \psi_2} & \cdots & \frac{\partial g_\ell}{\partial \psi_n} \end{pmatrix}. \tag{5.11}$$

The RR discriminant can therefore be written as

$$\Sigma_{\mathcal{M}} = \pi_2(\{(\psi, H) \in \mathcal{C}_{\mathcal{M}} : \operatorname{rank} \mathcal{J}_{\mathcal{C}_{\mathcal{M}},\psi}(\psi, H) \leq n-2\}).$$

The RR discriminant can also be computed from the parametric RR correspondence. If $I_{\mathcal{P}_{\mathcal{M}}} = \langle f_1, \ldots, f_p \rangle \subseteq \mathbb{C}[t_1, \ldots, t_s, h_{11}, h_{12}, \ldots, h_{nn}]$, we define the $p \times s$ Jacobian

$$\mathcal{J}_{\mathcal{P}_{\mathcal{M}},t}(t, H) = \begin{pmatrix} \frac{\partial f_1}{\partial t_1} & \frac{\partial f_1}{\partial t_2} & \cdots & \frac{\partial f_1}{\partial t_s} \\ \frac{\partial f_2}{\partial t_1} & \frac{\partial f_2}{\partial t_2} & \cdots & \frac{\partial f_2}{\partial t_s} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_p}{\partial t_1} & \frac{\partial f_p}{\partial t_2} & \cdots & \frac{\partial f_p}{\partial t_s} \end{pmatrix}.$$

Proposition 5.2.1. *Suppose $\mathcal{M}$ admits a rational parametrization $\psi : \mathbb{C}^s \dashrightarrow \mathcal{M}$. Then the RR discriminant is*

$$\Sigma_{\mathcal{M}} = \pi_2(\{(t, H) \in \mathcal{P}_{\mathcal{M}} : \operatorname{rank} \mathcal{J}_{\mathcal{P}_{\mathcal{M}},t}(t, H) \leq s-1\}).$$

Proof. Let

$$\mathcal{C}_{\mathcal{M}}^{\mathrm{ram}} = \{(\psi, H) \in \mathcal{C}_{\mathcal{M}} : \operatorname{rank} \mathcal{J}_{\mathcal{C}_{\mathcal{M}},\psi}(\psi, H) \leq n-2\}$$
$$\text{and } \mathcal{P}_{\mathcal{M}}^{\mathrm{ram}} = \{(x, H) \in \mathcal{P}_{\mathcal{M}} : \operatorname{rank} \mathcal{J}_{\mathcal{P}_{\mathcal{M}},t}(t, H) \leq s-1\}.$$

It suffices to prove that the image of $\psi \times \mathrm{Id} : \mathcal{P}_{\mathcal{M}}^{\mathrm{ram}} \dashrightarrow \mathcal{C}_{\mathcal{M}}$ is dense in $\mathcal{C}_{\mathcal{M}}^{\mathrm{ram}}$. In this case, $\psi \times \mathrm{Id}$ gives a rational map $\mathcal{P}_{\mathcal{M}}^{\mathrm{ram}} \dashrightarrow \mathcal{C}_{\mathcal{M}}^{\mathrm{ram}}$ sending $(t, H) \mapsto (\psi(t), H)$ and so the projections onto the second factor of these sets will be equal.

Suppose (t, H) lies in the dense open set $\mathcal{P}_{\mathcal{M}}^{\mathrm{ram}} \backslash (\mathcal{V}_{\psi,s} \times \mathbb{P}(\mathrm{Sym}^2\mathbb{C}^n)$. By the definition of $\mathcal{P}_{\mathcal{M}}^{\mathrm{ram}}$, we have $\operatorname{rank} \mathcal{J}_{\mathcal{P}_{\mathcal{M}},t}(t, H) \leq s-1$, and consequently the tangent space $T_t(\gamma_2^{-1}(H))$

has dimension at least one. Since t is not in $\mathcal{V}_{\psi,s}$, the Jacobian $\mathcal{J}_\psi(t)$ defines an injective map from the parametric critical locus $T_t(\gamma_2^{-1}(H))$ to the implicit critical locus $T_{\psi(t)}(\pi_2^{-1}(H))$. Consequently, the domain has dimension at least one, the map is injective, and the codomain $T_{\psi(t)}(\pi_2^{-1}(H))$ is at least one-dimensional. Thus the Jacobian $\mathcal{J}_{\mathcal{P}_\mathcal{M},\psi}(\psi(t), H)$ has rank at most $n-2$. This proves that $\psi \times \mathrm{Id}$ maps $\mathcal{P}_\mathcal{M}^{\mathrm{ram}}$ into $\mathcal{C}_\mathcal{M}^{\mathrm{ram}}$. By construction, the image of our dense open set under $\psi \times \mathrm{Id}$ is dense in $\mathcal{C}_\mathcal{M}^{\mathrm{ram}}$. □

We introduced $\Sigma_\mathcal{M}$ as the set of matrices H whose critical locus is infinite or singular. This is the perspective taken in [77] when defining the discriminant of maximum likelihood estimation. The authors of [37] take a different perspective, defining the discriminant of the Euclidean distance optimization problem as the branch locus of the projection π_2 onto the second factor. We now argue that these two definitions agree for RR discriminants.

Proposition 5.2.2. *The RR discriminant of $\mathcal{M}$ is the branch locus of $\pi_2 : \mathcal{C}_\mathcal{M} \to \mathbb{P}(\mathrm{Sym}^2\mathbb{C}^n)$.*

Proof. The branch locus consists of the matrices H for which there exists $\psi \in \pi_1^{-1}(H)$ such that the differential $\mathcal{J}_{\pi_2}(\psi, H)$ of the projection π_2 is not surjective. We work in the affine chart where $\psi_1 = h_{11} = 1$. The Jacobian of $\mathcal{C}_\mathcal{M}$ at a point (ψ, H) with $\psi \in \mathcal{M}_{\mathrm{reg}}$ may be written as

$$\mathcal{J}_{\mathcal{C}_\mathcal{M}}(\psi, H) = \begin{pmatrix} \mathcal{J}_{\mathcal{C}_\mathcal{M},\psi_{-1}}(\psi, H) & \mathcal{J}_{\mathcal{C}_\mathcal{M},H_{-11}}(\psi, H) \end{pmatrix}$$

where $\mathcal{J}_{\mathcal{C}_\mathcal{M},\psi_{-1}}(\psi, H)$ consists of the last $n-1$ columns of (5.11) and $\mathcal{J}_{\mathcal{C}_\mathcal{M},H_{-11}}(\psi, H)$ is the analogous $\ell \times (\binom{n+1}{2} - 1)$ matrix where derivatives are taken with respect to $h_{12}, h_{13}, \ldots, h_{nn}$. There exists a $(n + \binom{n+1}{2} - 2) \times (\binom{n+1}{2} - 1)$ matrix $A = \begin{pmatrix} A_\psi \\ A_H \end{pmatrix}$ whose columns form a basis for the kernel of $\mathcal{J}_{\mathcal{C}_\mathcal{M}}$, i.e., they form a basis for the tangent space of $\mathcal{C}_\mathcal{M}$ at (ψ, H). The column span of A_H is the image of the tangent space $T_{(\psi,H)}\mathcal{C}_\mathcal{M}$ under the differential map $\mathcal{J}_{\pi_2}(\psi, H)$. For general H, this map is surjective and A_H has full rank.

If H is in the branch locus, then A_H is not full rank, and we may choose A_H so that its last column is zero. This choice forces the last column of A_ψ to be in the kernel of $\mathcal{J}_{\mathcal{C}_\mathcal{M},\psi_{-1}}(\psi, H)$. Hence $\mathcal{J}_{\mathcal{C}_\mathcal{M},\psi_{-1}}(\psi, H)$ has rank less than $n-1$ and H is contained in $\Sigma_\mathcal{M}$. Conversely, if $H \in \Sigma_\mathcal{M}$, then the kernel of $\mathcal{J}_{\mathcal{C}_\mathcal{M},\psi_{-1}}(\psi, H)$ contains a nonzero vector v. Then the vector $\begin{pmatrix} v \\ 0 \end{pmatrix}$ is in the column span of A. Hence A_H drops rank and H is in the branch locus of π_2. Thus A_H drops rank precisely when $J_{\mathcal{C}_\mathcal{M},\psi}(\psi, H)$ drops rank, and the discriminants match. □

We now examine the RR discriminants of projective spaces and rational normal curves, before turning to the components of RR discriminants.

Proposition 5.2.3. *The RR discriminant of $\mathbb{P}^{n-1}$ is the discriminant of the characteristic polynomial of a symmetric matrix. The defining polynomial has degree $n(n-1)$.*

Proof. The RR correspondence of $\mathbb{P}^{n-1}$ is

$$\mathcal{C}_{\mathbb{P}^{n-1}} = \overline{\{(\psi, H) \in \mathbb{P}^{n-1} \times \mathbb{P}(\mathrm{Sym}^2(\mathbb{C}^n)) : \psi \text{ is an eigenvector of } H \text{ and } \psi^T\psi \neq 0\ \}}.$$

It follows that the fiber of a complex symmetric matrix H is

$$\pi_2^{-1}(H) = \overline{\{\psi \in \mathbb{P}^{n-1} : \psi \text{ is an eigenvector of } H \text{ and } \psi^T\psi \neq 0\}}.$$

We seek conditions on H under which this set is singular. If the characteristic polynomial of a matrix H has no double roots, then H is diagonalizable and has n distinct eigenvectors. Thus $\pi_2^{-1}(H)$ consists of n isolated points if the characteristic polynomial of H has no double root. Therefore the discriminant of the characteristic polynomial of H contains $\Sigma_{\mathbb{P}^{n-1}}$.

Conversely, suppose H has a repeated eigenvalue λ. If the corresponding eigenspace has dimension greater than 1, then $\pi_2^{-1}(H)$ has a positive dimensional component, so H is in $\Sigma_{\mathbb{P}^{n-1}}$. If λ has geometric multiplicity 1, then the corresponding eigenspace is spanned by a single vector ψ. Consider a curve $H : [0,1] \to \mathbb{P}(\mathrm{Sym}^2(\mathbb{C}^n))$ such that $H = H(1)$ and $H(t)$ is diagonalizable for $t \in [0,1)$. The eigenvalues and eigenvectors of $H(t)$ vary continuously, so there exist functions $\lambda_1, \lambda_2 : [0,1] \to \mathbb{C}$ and $\psi_1, \psi_2 : [0,1] \to \mathbb{P}^{n-1}$ such that

- $H(t)\psi_1(t) = \lambda_1(t)\psi_1(t)$ and $H(t)\psi_2(t) = \lambda_2(t)\psi_2(t)$ for all $t \in [0,1]$,
- $\lambda_1(t) \neq \lambda_2(t)$ and $\psi_1(t) \neq \psi_2(t)$ for $t \in [0,1)$, and
- $\lambda_1(1) = \lambda_2(1) = \lambda$.

At $t = 1$, we have $H\psi_1(1) = \lambda\psi_1(1)$ and $H\psi_2(1) = \lambda\psi_2(1)$. Since the eigenspace of λ is one-dimensional, $\psi_1(1) = \psi_2(1) = \psi$ and hence ψ is a nonreduced point of $\pi^{-1}(H)$. Therefore, if the characteristic polynomial of H has a double root, then $H \in \Sigma_{\mathbb{P}^{n-1}}$. □

Remark 5.2.4. If H is not contained in $\Sigma_{\mathbb{P}^{n-1}}$ then no points in fiber $\pi_2^{-1}(H)$ lie on the isotropic quadric $V(\psi^T\psi)$. This is because a symmetric matrix has an isotropic eigenvector if and only if its characteristic polynomial has a repeated eigenvalue [115]. Consequently, for every $H \notin \Sigma_{\mathbb{P}^{n-1}}$, the problem (5.1) has exactly RR degree many critical points.

Proposition 5.2.5. *The RR degree of the rational normal curve $C \subseteq \mathbb{P}^d$ of degree d is $2(2d-1)$. The parametric RR correspondence of the rational normal curve is a hypersurface in $\mathbb{C} \times \mathbb{P}(\mathrm{Sym}^2(\mathbb{C}^{d+1}))$ which is linear in $h_{11}, h_{12}, \ldots, h_{d+1,d+1}$. The RR discriminant is a hypersurface of degree $2(4d-3)$.*

Proof. The first claim follows from [112, Proposition 3.6]. For the RR correspondence we use the parametrization $t \mapsto [1\colon t\colon \cdots : t^d]$. Since the dimension of the parametric RR correspondence is $\binom{d+2}{2} - 1$, it is a hypersurface in $\mathbb{C} \times \mathbb{P}(\mathrm{Sym}^2(\mathbb{C}^{d+1}))$. For each $t \in \mathbb{C}$, the fiber $\gamma_1^{-1}(t) \subseteq \mathcal{P}_C$ is a linear space. Therefore the multidegree of this hypersurface is $(2(2d-1), 1)$. By Proposition 5.2.1, the RR discriminant is the discriminant of this univariate polynomial of degree $2(2d-1)$ whose coefficients are linear in H. Therefore its degree is $2(2(2d-1)) - 2 = 2(4d-3)$. □

We now turn to the structure of the discriminant. The RR discriminant is generally not irreducible. Let $\mathcal{C}_{\mathcal{M}}^{\mathrm{ram}} = \{(\psi, H) \in \mathcal{C}_{\mathcal{M}} : \operatorname{rank} \mathcal{J}_{\mathcal{C}_{\mathcal{M}},\psi}(\psi, H) \leq n-2\}$ denote the ramification locus of the projection π_2. We can write $\mathcal{C}_{\mathcal{M}}^{\mathrm{ram}}$ as the union

$$\mathcal{C}_{\mathcal{M}}^{\mathrm{ram}} = (\mathcal{C}_{\mathcal{M}}^{\mathrm{ram}} \cap V(\psi^T\psi)) \cup \overline{(\mathcal{C}_{\mathcal{M}}^{\mathrm{ram}} \setminus V(\psi^T\psi))}$$

where $V(\psi^T\psi) \subseteq \mathbb{P}^{n-1} \times \mathbb{P}(\mathrm{Sym}^2\mathbb{C}^n)$. The projections of these components $\pi_2(\mathcal{C}_{\mathcal{M}}^{\mathrm{ram}} \cap V(\psi^T\psi))$ and $\pi_2(\overline{(\mathcal{C}_{\mathcal{M}}^{\mathrm{ram}} \setminus V(\psi^T\psi))})$ are often, but not always, hypersurfaces. We call the set $\pi_2(\mathcal{C}_{\mathcal{M}}^{\mathrm{ram}} \cap V(\psi^T\psi))$ the *isotropic part* of the discriminant and the set $\pi_2(\overline{(\mathcal{C}_{\mathcal{M}}^{\mathrm{ram}} \setminus V(\psi^T\psi))})$ the *non-isotropic part* of the discriminant. We illustrate this phenomenon with some examples.

Example 5.2.6. The nonisotropic part of $\Sigma_{\mathbb{P}^{n-1}}$ is a proper subvariety of $\Sigma_{\mathbb{P}^{n-1}}$ with codimension greater than 1 and describes matrices with diagonal blocks. The isotropic part of $\Sigma_{\mathbb{P}^{n-1}}$ is equal to $\Sigma_{\mathbb{P}^{n-1}}$. Its defining polynomial is a sum of squares; see [72]. For example, $\Sigma_{\mathbb{P}^1}$ is the hypersurface cut out by $(h_{11}-h_{22})^2+4h_{12}^2$. The nonisotropic part $V(h_{12}, h_{11}-h_{22})$ of $\Sigma_{\mathbb{P}^1}$ has codimension 2 and consists of scalar multiples of the 2×2 identity matrix. ⋄

Example 5.2.7 (Normal quadric)**.** The rational normal quadric has RR degree 6. The RR discriminant is a degree-10 hypersurface. Its defining polynomial has three irreducible factors over $\mathbb{Q}$. The two isotropic factors are

$$\begin{aligned}&3(2h_{12}-2h_{13}-h_{22}+h_{33})^2+(2h_{11}-2h_{12}-2h_{13}-h_{22}+4h_{23}-h_{33})^2 \text{ and}\\ &3(2h_{12}+2h_{13}+h_{22}-h_{33})^2+(2h_{11}+2h_{12}-2h_{13}-h_{22}-4h_{23}-h_{33})^2.\end{aligned}$$

The nonisotropic factor has 206 terms, the first few of which are

$$\begin{gathered}3456h_{11}^4h_{23}^2-6912h_{11}^3h_{12}h_{13}h_{23}-3456h_{11}^3h_{12}h_{22}h_{23}+3456h_{11}^3h_{12}h_{23}h_{33}+2048h_{11}^3h_{13}^3\\ +3072h_{11}^3h_{13}^2h_{22}-3072h_{11}^3h_{13}^2h_{33}+1536h_{11}^3h_{13}h_{22}^2-3072h_{11}^3h_{13}h_{22}h_{33}-19008h_{11}^3h_{13}h_{23}^2+1536h_{11}^3h_{13}h_{33}^2\\ +256h_{11}^3h_{22}^3-768h_{11}^3h_{22}^2h_{33}-9504h_{11}^3h_{22}h_{23}^2+768h_{11}^3h_{22}h_{33}^2-4320h_{11}^3h_{23}^2h_{33}+\cdots\end{gathered}$$

The full polynomial is available at [17]. The discriminant divides $\mathbb{P}(\mathrm{Sym}^2\mathbb{R}^3)$ into two regions. We computed the regions in the complement of the nonisotropic factor using the `Julia` software `HypersurfaceRegions.jl` [22]. In one region, (5.1) has 2 real critical points; in the other it has 4 real critical points. For example, (5.1) has 2 real critical points for the left matrix and 4 real critical points for the right matrix:

$$H = \begin{pmatrix} 1 & 3 & 5\\ 3 & 7 & 9\\ 5 & 9 & 11\end{pmatrix}, \qquad H = \begin{pmatrix} 8 & 1 & 13\\ 1 & 11 & 7\\ 13 & 7 & 5\end{pmatrix}$$

This discriminant also appears in [112, Example 5.8]. ⋄

In the following examples, we were not able to compute the full discriminant polynomials. The degrees were computed symbolically by intersecting the discriminant hypersurface with a line. We can give a lower bound on the number of regions in the discriminant complement by computing all critical points for (5.3) for random real matrices and counting the numbers of real solutions that arise.

Example 5.2.8 (Twisted Cubic)**.** The twisted cubic has RR degree 10. The RR discriminant has degree 18 and its defining polynomial factors into three components over $\mathbb{Q}$. The first isotropic factor is a sum of squares of rank 2:

$$(h_{11} - 2h_{13} - h_{22} + 2h_{24} + h_{33} - h_{44})^2 + 4(h_{12} - h_{14} - h_{23} + h_{34})^2.$$

The second isotropic factor

$$\begin{gathered} h_{11}^4 - 8h_{11}^3h_{24} - 4h_{11}^3h_{33} + 16h_{11}^2h_{12}h_{14} + 16h_{11}^2h_{12}h_{23} + 8h_{11}^2h_{13}^2 + 8h_{11}^2h_{13}h_{22} - 8h_{11}^2h_{13}h_{44} \\ -16h_{11}^2h_{14}h_{34} + 2h_{11}^2h_{22}^2 - 4h_{11}^2h_{22}h_{44} - 16h_{11}^2h_{23}h_{34} + 24h_{11}^2h_{24}^2 + 24h_{11}^2h_{24}h_{33} + 6h_{11}^2h_{33}^2 \\ +2h_{11}^2h_{44}^2 - 32h_{11}h_{12}^2h_{13} - 16h_{11}h_{12}^2h_{22} + 16h_{11}h_{12}^2h_{44} + 64h_{11}h_{12}h_{13}h_{34} - 64h_{11}h_{12}h_{14}h_{24} \\ -32h_{11}h_{12}h_{14}h_{33} + 32h_{11}h_{12}h_{22}h_{34} - 64h_{11}h_{12}h_{23}h_{24} - 32h_{11}h_{12}h_{23}h_{33} - 32h_{11}h_{12}h_{34}h_{44} + \cdots \end{gathered}$$

has 187 terms (see [17]). It is a sum of squares of rank at most 4. We checked this using the `Julia` packages `COSMO.jl` [54], `JuMP.jl` [91], and `SumOfSquares.jl` [85]. The nonisotropic factor has degree 12 and divides $\mathbb{P}(\mathrm{Sym}^2\mathbb{R}^4)$ into at least three regions, exhibited by the fact that (5.1) can have 2, 4, or 6 real critical points. ◇

Conjecture 5.2.9. The nonisotropic part of the RR discriminant of the rational normal curve of degree $d \geq 2$ is a hypersurface of degree $6(d-1)$. The isotropic part has degree $2d$.

We have verified Conjecture 5.2.9 numerically for $d \leq 10$; see [17] for computations.

Example 5.2.10 ($\mathbb{P}^1 \times \mathbb{P}^1$)**.** The RR degree of $\mathbb{P}^1 \times \mathbb{P}^1$ is 8. The RR discriminant of $\mathbb{P}^1 \times \mathbb{P}^1$ is a hypersurface of degree 32. The isotropic part has degree 8. The nonisotropic part has degree 24. The discriminant divides $\mathbb{P}(\mathrm{Sym}^2\mathbb{R}^4)$ into at least three regions: (5.1) can have 4, 6, or 8 real critical points for $\mathbb{P}^1 \times \mathbb{P}^1$. ◇

Example 5.2.11 (Hirzebruch Surface)**.** Consider the Hirzebruch surface parametrized by $\psi : \mathbb{C}^2 \to \mathbb{P}^4$ defined by $(a, b) \mapsto [1 : a : b : ab : a^2b]$. The RR degree of this variety is 20. The RR correspondence is a 14-dimensional subvariety of $\mathbb{P}^4 \times \mathbb{P}^{14}$. The RR discriminant is a hypersurface of degree 74. Its isotropic part has degree 26 and its nonisotropic part has degree 48. The discriminant divides $\mathbb{P}(\mathrm{Sym}^2\mathbb{R}^5)$ into at least 5 regions: the optimization problem (5.1) can have 4, 6, 8, 10, or 12 real critical points. ◇

It is not clear whether there can exist $(\psi, H) \in \mathcal{C}_{\mathcal{M}}^{\mathrm{ram}} \cap V(\psi^T\psi)$ with real H. The vector ψ cannot be real, since $\psi^T\psi = 0$ is a sum of squares. In the examples we have computed, the variety $\pi_2(\mathcal{C}_{\mathcal{M}}^{\mathrm{ram}} \cap V(\psi^T\psi))$ has no real points and its defining polynomial factors as a sum of squares over $\mathbb{Q}$. We therefore make the following conjecture, which we have verified for the rational quadric (Example 5.2.7), the twisted cubic (Example 5.2.8), and projective space (Example 5.2.6).

Conjecture 5.2.12. Every irreducible isotropic factor of the defining polynomial of the RR discriminant is a sum of squares.

One motivation for studying the RR discriminant of a variety is to compute the average RR degree, that is, the expected number of real critical points. The role of the discriminant in this computation is explained in the Euclidean distance case in [37]. In particular, in this application we are only interested in the real locus of the discriminant. Thus, if Conjecture 5.2.12 holds, then only the nonisotropic part of the discriminant is relevant. We therefore present degrees of the nonisotropic parts of RR discriminants of various varieties in Tables 5.2 and 5.4, and make the following conjecture.

Conjecture 5.2.13. The nonisotropic part of the RR discriminant of $\mathbb{P}^1 \times \mathbb{P}^{n-1}$ is a hypersurface of degree $24\binom{n+1}{3}$. The nonisotropic part of the RR discriminant of $\mathbb{P}^2 \times \mathbb{P}^{n-1}$ is a hypersurface of degree $24n^2\binom{n}{2}$.

5.3 Geometry of Tensor Train Varieties

In this section, we introduce *tensor train varieties* (TT varieties) – projective varieties of tensors in a tensor train format. These varieties can be described by the combinatorial structure of a graph following the tensor network viewpoint as, for example, in [84, 116, 123].

Let Γ_n be an undirected path graph on vertices $\{0, 1, \dots, n\}$, with edges $\{i-1, i\}$ for $i = 1, \dots, n$. We assign positive integers $\mathbf{k} = (k_0, k_1, \dots, k_n)$ to the vertices, known as *physical dimensions*, and $\mathbf{r} = (r_1, r_2, \dots, r_n)$ to the edges, called *bond dimensions*; by convention, we set $r_0 = r_{n+1} = 1$. We identify tensors in the ambient projective space $\mathbb{P}\big(\bigotimes_{i=0}^n \mathbb{C}^{k_i}\big)$ with points having homogeneous coordinates $\psi_{j_0,\dots,j_n}$, where $j_i \in [k_i] := \{1, \dots, k_i\}$. For each index $i \in \{0, 1, \dots, n\}$ and each $j_i \in [k_i]$, let $A_i^{(j_i)}$ be a general $r_i \times r_{i+1}$ matrix of parameters.

Definition 5.3.1. The *tensor train variety* $\mathcal{T}_{\mathbf{k},\mathbf{r}}$ is the projective variety in $\mathbb{P}\left(\bigotimes_{i=0}^n \mathbb{C}^{k_i}\right)$ defined as the Zariski closure of the image of the following rational map:

$$\begin{aligned} \Psi_{\mathbf{k},\mathbf{r}} : \prod_{i=0}^{n} \Big(\mathbb{P}\big(\mathbb{C}^{r_i \times r_{i+1}}\big)\Big)^{k_i} &\dashrightarrow \mathbb{P}\left(\bigotimes_{i=0}^{n} \mathbb{C}^{k_i}\right) \\ \left(A_i^{(j_i)}\right)_{0 \le i \le n,\ 1 \le j_i \le k_i} &\longmapsto \left(\psi_{j_0,j_1,\dots,j_n}\right)_{(j_0,j_1,\dots,j_n) \in \prod_{i=0}^n [k_i]}, \end{aligned} \tag{5.12}$$

where $\psi_{j_0,j_1,\dots,j_n} = A_0^{(j_0)} A_1^{(j_1)} \cdots A_n^{(j_n)}$. In symbols, we have $\mathcal{T}_{\mathbf{k},\mathbf{r}} = \overline{\operatorname{im}(\Psi_{\mathbf{k},\mathbf{r}})}$.

It is often convenient to characterize $\mathcal{T}_{\mathbf{k},\mathbf{r}}$ by rank constraints on certain matrix flattenings of a tensor. For $i \in \{0, 1, \dots, n-1\}$, let ψ^i denote the flattening of ψ according to the partition $\{0, 1, \dots, i\} \sqcup \{i+1, \dots, n\}$, viewed as a $(k_0 \cdots k_i) \times (k_{i+1} \cdots k_n)$ matrix.

Lemma 5.3.2. *The tensor train variety $\mathcal{T}_{\mathbf{k},\mathbf{r}}$ is equal to*

$$\mathcal{T}_{\mathbf{k},\mathbf{r}} = \left\{ \psi \in \mathbb{P}\left(\bigotimes_{i=0}^{n} \mathbb{C}^{k_i}\right) : \operatorname{rank}(\psi^i) \le r_{i+1} \quad \textit{for } i = 0, \dots, n-1 \right\}. \tag{5.13}$$

In particular, set-theoretically $\mathcal{T}_{\mathbf{k},\mathbf{r}}$ is cut by all $(r_{i+1}+1) \times (r_{i+1}+1)$ minors of flattenings ψ^i.

Proof. Let ψ be in the image of $\Psi_{\mathbf{k},\mathbf{r}}$. Fix $i \in \{0, \ldots, n-1\}$ and consider two matrices

$$L_i := A_0^{(j_0)} \cdots A_i^{(j_i)} \in \mathbb{C}^{1\times r_{i+1}}, \qquad R_i := A_{i+1}^{(j_{i+1})} \cdots A_n^{(j_n)} \in \mathbb{C}^{r_{i+1}\times 1}.$$

Then the flattening ψ^i has entries $\psi^i[j_0, \ldots, j_i\,; j_{i+1}, \ldots, j_n] = L_i\, R_i$. Thus ψ^i factors through an r_{i+1}-dimensional space, hence $\operatorname{rank}(\psi^i) \le r_{i+1}$.

Conversely, let $\operatorname{rank}(\psi^i) \le r_{i+1}$ for all $i = 0, \ldots, n-1$. Start with $\psi^0 \in \mathbb{C}^{k_0\times(k_1\cdots k_n)}$ and take a rank factorization $\psi^0 = L_0 R_0$ with $L_0 \in \mathbb{C}^{k_0\times r_1}$ and $R_0 \in \mathbb{C}^{r_1\times(k_1\cdots k_n)}$. Let $A_0^{(j_0)}$ be the j_0-th row of L_0, and reshape R_0 as a tensor $\psi^{(1)} \in \mathbb{C}^{r_1} \otimes \mathbb{C}^{k_1} \otimes \cdots \otimes \mathbb{C}^{k_n}$.

Using $\operatorname{rank}(\psi^1) \le r_2$, the flattening of $\psi^{(1)}$ across $\{r_1, k_1\} \sqcup \{k_2, \ldots, k_n\}$ also has rank at most r_2, so we can factor it as $L_1 R_1$ with $L_1 \in \mathbb{C}^{(r_1 k_1)\times r_2}$ and $R_1 \in \mathbb{C}^{r_2\times(k_2\cdots k_n)}$. Read the k_1 slices of L_1 as $A_1^{(j_1)} \in \mathbb{C}^{r_1\times r_2}$, and reshape R_1 to $\psi^{(2)} \in \mathbb{C}^{r_2} \otimes \mathbb{C}^{k_2} \otimes \cdots \otimes \mathbb{C}^{k_n}$. We proceed inductively and at step $i = n$ (with $r_{n+1} = 1$) the remainder $\psi^{(n+1)}$ is scalar, and we obtain

$$\psi_{j_0,\ldots,j_n} = A_0^{(j_0)} A_1^{(j_1)} \cdots A_n^{(j_n)}.$$

Rank constraints are given by vanishing minors, so right hand side of (5.13) is closed. □

Corollary 5.3.3. *We have the following constraints on sequences* $\mathbf{k}$ *and* $\mathbf{r}$*:*

$$r_i \le r_{i-1}k_{i-1} \quad \textit{and} \quad r_{i-1} \le r_i k_{i-1} \quad \textit{for} \;\; i = 1, \ldots, n. \tag{5.14}$$

Proof. The left inequalities follow from the proof of Lemma 5.3.2. The condition on the right in (5.14) is symmetric: apply the same rank factorizations starting from $\psi^{n-1} \in \mathbb{C}^{(k_0\cdots k_{n-1})\times k_n}$ and run the construction backwards. □

If either of the inequalities in (5.14) for r_i is an equality or fails to hold, then the corresponding rank condition in (5.13) is automatically satisfied, as in the example below.

Example 5.3.4. Take $\mathbf{k} = (2,2,2,2)$ and $\mathbf{r} = (1,2,1)$. The parametrization (5.12) yields

$$\psi_{ijkl} = A_0^{(i)} A_1^{(j)} A_2^{(k)} A_3^{(\ell)}, \qquad i, j, k, \ell \in \{1, 2\},$$

where $A_0^{(i)}$ and $A_3^{(\ell)}$ are scalars, $A_1^{(j)} \in \mathbb{C}^{1\times 2}$, and $A_2^{(k)} \in \mathbb{C}^{2\times 1}$. Writing

$$A_0^{(i)} = (a_0^i), \quad A_1^{(j)} = \begin{pmatrix} a_{11}^j & a_{12}^j \end{pmatrix}, \quad A_2^{(k)} = \begin{pmatrix} a_{21}^k \\ a_{22}^k \end{pmatrix}, \quad A_3^{(\ell)} = (a_3^\ell),$$

one checks that $\mathcal{T}_{\mathbf{k},\mathbf{r}}$ is isomorphic to $\mathbb{P}^3 \times (\mathbb{P}^1)^2$. The vanishing ideal of this Segre product is minimally generated by the 2×2 minors of the flattenings ψ^0 and ψ^2. Minors of ψ^1 don't appear since $r_2 = 2 = 1\cdot 2 = r_1 k_1$. The corresponding tensor network is in Figure 5.1. ⋄

From now on, we use the shorthand $(1)^n := (\underbrace{1, \ldots, 1}_{n})$ for sequences of repeated indices.

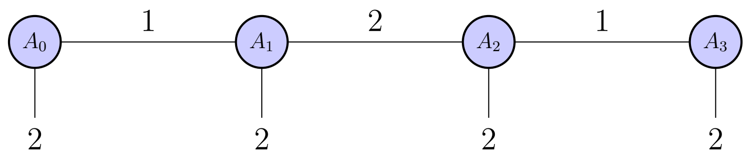

Figure 5.1: A tensor network diagram for binary tensors of order 4 and TT rank $(1, 2, 1)$.

Proposition 5.3.5. *For $\mathbf{k} = (k_0, k_1, \ldots, k_n)$ and $\mathbf{r} = (1)^n$ we have $\mathcal{T}_{\mathbf{k},\mathbf{r}} = \mathbb{P}^{k_0-1} \times \cdots \times \mathbb{P}^{k_n-1}$, embedded by the Segre map.*

Proof. The matrices $A_i^{(j_i)}$ in the parametrization are 1×1 matrices: $A_i^{(j_i)} = (a_i^{j_i})$. Therefore, $\psi_{j_0,j_1,\ldots,j_n} = a_0^{j_0} a_1^{j_1} \cdots a_n^{j_n}$. This is precisely the Segre embedding of $\mathbb{P}^{k_0-1} \times \cdots \times \mathbb{P}^{k_n-1}$. $\square$

We next isolate a large, explicit family of tensor train varieties that are again Segre products. We say that indices i with $r_i > 1$ are *separated* if no two such indices are adjacent; likewise, we call a *separated block* a consecutive sequence $(i_1, \ldots, i_t)$ with $r_{i_1}, \ldots, r_{i_t} > 1$ such that no other $r_j > 1$ is adjacent to the block.

Theorem 5.3.6. *Let $\mathbf{k} = (k_0, \ldots, k_n)$ and $\mathbf{r} = (r_1, \ldots, r_n)$, with $r_0 = r_{n+1} = 1$. Suppose $\mathbf{r}$ has ℓ separated blocks of consecutive indices $I_p = (a_p, \ldots, b_p)$ such that $r_i > 1$ if and only if $i \in \bigcup_p I_p$. Assume furthermore that for every $i \in \bigcup_p I_p$, one of the inequalities $r_i \geq r_{i-1}k_{i-1}$ or $r_i \geq r_{i+1}k_i$ holds. Then*

$$\mathcal{T}_{\mathbf{k},\mathbf{r}} \simeq \mathbb{P}^{K_1-1} \times \cdots \times \mathbb{P}^{K_\ell-1} \times \mathbb{P}^{k_{j_1}-1} \times \cdots \times \mathbb{P}^{k_{j_s}-1},$$

where $K_p = \prod_{v=a_p-1}^{b_p} k_v$ for $p = 1, \ldots, \ell$ and j_q are such that $r_{j_q+1} = 1$ is not adjacent to two separated blocks for $q = 1, \ldots, s$. Moreover,

$$\dim \mathcal{T}_{\mathbf{k},\mathbf{r}} = \sum_{p=1}^{\ell}(K_p - 1) + \sum_{q=1}^{s}(k_{j_q} - 1), \quad \deg \mathcal{T}_{\mathbf{k},\mathbf{r}} = \frac{\left(\sum_{p=1}^{\ell}(K_p - 1) + \sum_{q=1}^{s}(k_{j_q} - 1)\right)!}{\prod_{p=1}^{l}(K_p - 1)! \, \prod_{q=1}^{s}(k_{j_q} - 1)!}.$$

Proof. For a block I_p, the corresponding TT cores have shape $(r_i \times k_i \times r_{i+1})$ for $i \in I_p$, and $r_{a_p-1} = r_{b_p+1} = 1$. Since these ranks exceed 1 and one of the inequalities (5.14) fails, the block parameters span all coordinates in $\mathbb{P}^{K_p-1}$, giving a surjective rational reparametrization onto $\mathbb{P}^{K_p-1}$.

Because blocks are separated, these local reparametrizations are independent, so the global parametrization factors as the Segre product of the $\mathbb{P}^{K_p-1}$ for $p \in \{1, \ldots, \ell\}$ with the remaining $\mathbb{P}^1$ factors. The dimension is the sum of the factor dimensions, and the degree of $\mathbb{P}^{d_1} \times \cdots \times \mathbb{P}^{d_s}$ is the multinomial coefficient $\binom{d_1+\cdots+d_s}{d_1 \cdots d_s}$, yielding the claimed formulas. $\square$

We are mostly interested in binary tensors ($k_i = 2$) due to their prevalence in applications.

Example 5.3.7. The tensor train variety $\mathcal{T}_{(2,2),(2)}$ is the variety of 2×2 matrices of rank at most 2, and consequently it is $\mathbb{P}^3$.

When $(k_0, k_1, k_2) = (2, 2, 2)$ and $(r_0, r_1, r_2, r_3) = (1, 2, 2, 1)$, the inequalities $2 = r_1 \geq r_0 k_0 = 2$ and $2 = r_2 \geq r_3 k_2 = 2$ in Theorem 5.3.6 hold. We can see that the parametrization is surjective onto $\mathbb{P}^7$ explicitly by setting $A_0^{(i)} = e_i$ and $A_2^{(i)} = e_i$ for $i = 1, 2$.

Finally, if $(k_0, k_1, k_2, k_3) = (2, 2, 2, 2)$ and $(r_1, r_2, r_3) = (2, 2, 2)$, Theorem 5.3.6 does not apply. One of the inequalities in the statement holds for $i = 1, 3$, but they both fail for $i = 2$. The only nontrivial rank condition is that the 4×4 matrix from the $i = 2$ flattening ψ^2 should have rank at most 2. Therefore this tensor train variety is the determinantal variety $\mathcal{T}_{(4,4),(2)}$ of 4×4 matrices of rank at most 2. $\diamond$

The following corollary generalizes the above example.

Corollary 5.3.8. *Let $\mathbf{k} = (2)^{n+1}$ and $\mathbf{r} = (r_1, \dots, r_n)$. Suppose there are ℓ separated indices i with $r_i > 1$, t separated pairs $(i, i+1)$ with $r_i, r_{i+1} > 1$, and all remaining r_j equal 1. Then*

$$\mathcal{T}_{\mathbf{k},\mathbf{r}} \simeq \left(\mathbb{P}^3\right)^l \times \left(\mathbb{P}^7\right)^t \times \left(\mathbb{P}^1\right)^{n-\ell-2t-s},$$

where s is the number of $r_j = 1$ adjacent to two separated blocks. In particular, the dimension of $\mathcal{T}_{\mathbf{k},\mathbf{r}}$ is $n + 2\ell + 5t - s$ and its degree is the multinomial coefficient

$$\frac{(n + 2\ell + 5t - s)!}{(3!)^l (7!)^t}.$$

We now turn to equations. Lemma 5.3.2 shows that $\mathcal{T}_{\mathbf{k},\mathbf{r}}$ is set-theoretically cut out by determinantal equations. The following conjecture, first suggested to us by Bernd Sturmfels in personal communication, asserts that these suffice scheme-theoretically as well.

Conjecture 5.3.9. For fixed $\mathbf{k}$ and $\mathbf{r}$, let ψ be a $k_0 \times \cdots \times k_n$ tensor of indeterminates and let ψ^i be the flattenings as above. Then the homogeneous prime ideal of $\mathcal{T}_{\mathbf{k},\mathbf{r}}$ in $\mathbb{C}[\psi]$ is

$$I_{\mathcal{T}_{\mathbf{k},\mathbf{r}}} = \sum_{i=0}^{n-1} \left\langle (r_{i+1} + 1) \times (r_{i+1} + 1) \text{ minors of } \psi^i \right\rangle.$$

Moreover, the union of $(r_{i+1} + 1) \times (r_{i+1} + 1)$ minors of flattenings ψ^i for $i = 0, \dots, n-1$ forms a Gröbner basis for the ideal $I_{\mathcal{T}_{\mathbf{k},\mathbf{r}}}$.

We collect some positive evidence in the case of binary tensors, in particular, in the case of Segre varieties whose Gröbner bases have been studied widely. We start with the following lemma; a proof is available in [14, Chapter 3].

Lemma 5.3.10. *Let $\mathbf{k} = (2)^{n+1}$ and $\mathbf{r} = (r_1, \dots, r_n)$ with ℓ separated indices where $r_i > 1$, t separated pairs $(i, i+1)$ where $r_i, r_{i+1} > 1$, and all remaining $r_j = 1$. Let*

$$\{i_1, \dots, i_{n-\ell-2t}\} = \{1, \dots, n\} \setminus \left(\{i : \; r_i > 1\} \cup \{i, i+1 : \; r_i, r_{i+1} > 1\}\right).$$

Then the union

$$G_{i_1} \cup \cdots \cup G_{i_{n-\ell-2t}},$$

where G_j denotes the set of 2×2 minors of ψ^{j-1}, is a minimal reduced Gröbner basis of $I_{\mathcal{T}_{\mathbf{k},\mathbf{r}}}$ with respect to any lexicographic or reverse lexicographic term order that respects the row–column structure of the flattenings ψ^j for $j = 0, \dots, n-1$.

Remark 5.3.11. In general the union above is not a universal lexicographic or universal reverse lexicographic Gröbner basis, see the example below.

Example 5.3.12 (Segre products of two projective spaces)**.** Let $\mathbf{k} = (k_0, k_1)$ and $\mathbf{r} = (1)$. Then $I_{\mathcal{T}_{\mathbf{k},\mathbf{r}}}$ is generated by the 2×2 minors of a $k_0 \times k_1$ matrix. By [110, Theorem 3.9], these minors form a quadratic universal reverse-lex Gröbner basis. For $k_0, k_1 > 2$, they are not a universal Gröbner basis. For instance, for $(k_0, k_1) = (3, 3)$ the Gröbner fan has 108 maximal cones, but only 96 correspond to reduced Gröbner bases consisting of 2×2 minors. $\diamond$

Example 5.3.13 (Segre products of several projective spaces)**.** For $n > 1$, the 2×2 minors need not form a reverse-lex Gröbner basis under an arbitrary permutation of the variables. For $\mathbf{k} = (2, 2, 2)$ and $\mathbf{r} = (1)_2$, order the variables as

$$\psi_{111} \succ \psi_{122} \succ \psi_{212} \succ \psi_{221} \succ \psi_{222} \succ \psi_{121} \succ \psi_{211} \succ \psi_{112}.$$

The reduced reverse-lex Gröbner basis contains the cubic $\psi_{222}\,\psi_{121}\,\psi_{211} - \psi_{221}^2\,\psi_{112}$, showing that 2×2 minors alone do not form a Gröbner basis in this case. $\diamond$

We record a conjecture that represents a key step towards the proof of the Conjecture 5.3.9.

Conjecture 5.3.14. Let $\mathbf{k} = (2)^{n+1}$ and $\mathbf{r} = (1, r_2, \dots, r_{n-1}, 1)$. Then the union

$$G_1^{(r_1+1)} \cup \cdots \cup G_n^{(r_n+1)}$$

of $(r_i+1) \times (r_i+1)$ minors of the flattenings ψ^{i-1} forms a Gröbner basis of $I_{\mathcal{T}_{\mathbf{k},\mathbf{r}}}$ with respect to the reverse-lexicographic order induced by a variable order compatible with the flattenings.

Example 5.3.15. The analogous statement for lexicographic orders generally fails unless $\mathcal{T}_{\mathbf{k},\mathbf{r}}$ is a Segre variety. A minimal counterexample is $\mathbf{k} = (2)^6$, $\mathbf{r} = (1, 2, 2, 2, 1)$. $\diamond$

5.4 Tensor Train Varieties as Products of Grassmannians

In this section, we present a one-to-one parametrization of a dense open set of $\mathcal{T}_{\mathbf{k},\mathbf{r}}$ from a product of Grassmannians

$$\mathcal{Z} := \mathrm{Gr}(r_1, k_0) \times \mathrm{Gr}(r_2, r_1 k_1) \times \cdots \times \mathrm{Gr}(r_n, r_{n-1} k_{n-1}) \times \mathbb{P}(\mathbb{C}^{r_n \times k_n}).$$

This open set is contained in the *tensor train manifold* denoted by $\mathcal{T}^{=}_{\mathbf{k},\mathbf{r}}$, which is the set of tensors in $\mathcal{T}_{\mathbf{k},\mathbf{r}}$ whose i-th flattening has rank *exactly* r_{i+1} for all $i = 0, \ldots, n-1$. The construction is inspired by the TT-cross algorithm [104], itself an adaptation of the TT-SVD [58, 103] in which successive SVDs are replaced by so-called skeleton decompositions of matrices [126]. Skeleton (or CUR) decompositions factor an $m \times n$ matrix A by selecting r linearly independent columns and rows of A as columns and rows of $C \in \mathbb{C}^{m \times r}$ and $R \in \mathbb{C}^{r \times n}$, respectively. In the literature, this decomposition is often generalized to the case where only rows or columns are chosen, resulting in an XR-decomposition or CX-decomposition, respectively [38]. In our algorithm, we use a special case of the XR-decomposition for a rank-r matrix A, under the following assumption.

Assumption 5.4.1. *For a rank-r matrix $A \in \mathbb{C}^{m \times n}$, the first r rows are linearly independent.*

If Assumption 5.4.1 holds, the XR-decomposition is the factorization $A = XR$ where $R \in \mathbb{C}^{r \times n}$ is the first r rows of A and $X = AR^{\dagger}(RR^{\dagger})^{-1}$. We denote the output by:

$$\mathtt{XR}(A, r) := (X, R).$$

Invertibility of $RR^{\dagger}$ is guaranteed by Assumption 5.4.1. This decomposition satisfies

$$XR = AR^{\dagger}(RR^{\dagger})^{-1}R = AP,$$

where P is the orthogonal projection onto the rowspan of A; see Section 3.3. Projecting the rows of A onto the rowspan of A leaves them unchanged, and hence $XR = AP = A$. Moreover, X has the form

$$X = \begin{pmatrix} \mathrm{Id}_r \\ B \end{pmatrix} \Bigg($$

with $B \in \mathbb{C}^{(m-r) \times r}$. Such matrices parametrize a large open subset of $\mathrm{Gr}(r, m)$ via their column spans, with the first block fixed to the identity to ensure a one-to-one parametrization.

Our algorithm gives a bijection from an open set $\mathcal{U} \subseteq \mathcal{T}^{=}_{\mathbf{k},\mathbf{r}}$ to an open set $\mathcal{W} \subseteq \mathcal{Z}$. In order to write down our algorithm, we will first need to identify these open subsets. The description of the open set $\mathcal{W}$ is straightforward. Let W_i denote the open set

$$W_i := \{\mathrm{span}(A) : A \in \mathbb{C}^{r_{i-1}k_{i-1} \times r_i},\ \text{first } r_i \text{ rows linearly independent}\} \subseteq \mathrm{Gr}(r_i, r_{i-1}k_{i-1})$$

for $i = 1, \ldots, n$ and let $\mathbb{C}^{r_n \times k_n}_{*} \subseteq \mathbb{P}(\mathbb{C}^{r_n \times k_n})$ denote the open set of full rank matrices that have a 1 in the first entry. Then the following set is open in $\mathcal{Z}$:

$$\mathcal{W} := W_1 \times \cdots \times W_n \times \mathbb{C}^{r_n \times k_n}_{*}. \tag{5.15}$$

The set $\mathcal{U}$ consists of tensors for which Assumption 5.4.1 holds in every step of the algorithm below; its precise description is in the proof of Theorem 5.4.3, where we show it is Zariski dense.

Similar to Section 5.3, we denote the i-th flattening of a tensor $\psi \in \bigotimes_{i=0}^{n} \mathbb{C}^{k_i}$ by

$$\psi^i = \texttt{unfold}(\psi, [k_0 \cdots k_i, k_{i+1} \cdots k_n]) \in \mathbb{C}^{(k_0 \cdots k_i) \times (k_{i+1} \cdots k_n)}$$

and the tensorization of a matrix $\psi^i \in \mathbb{C}^{(k_0 \cdots k_i) \times (k_{i+1} \cdots k_n)}$ is denoted by

$$\psi = \texttt{tensorize}(\psi^i, [k_0, \ldots, k_n]) \in \bigotimes_{i=0}^{n} \mathbb{C}^{k_i}.$$

We are now equipped to present the parametrization algorithm for a tensor with given TT-rank $\mathbf{r} = (r_1, \ldots, r_n)$. The procedure is similar to the one in Lemma 5.3.2; see Algorithm 1.

Algorithm 1: Factorization $\texttt{factorize}(\psi, \mathbf{r})$

Input : $\psi \in \mathcal{U} \subseteq \mathcal{T}^{=}_{\mathbf{k},\mathbf{r}}$, TT-rank $\mathbf{r} = (r_1, \ldots, r_n)$
Output: $(X_0, \ldots, X_n) \in \mathcal{W} \subseteq \mathcal{Z}$

$r_0 = 1$, $\phi = \psi$;
for $i = 0$ **to** $n - 1$ **do**
  $\psi^i = \texttt{unfold}(\phi, [r_i k_i,\ k_{i+1} \cdots k_n])$;
  $(X_i, R_i) = \texttt{XR}(\psi^i, r_{i+1})$;
  $\phi = \texttt{tensorize}(R_i, [r_{i+1}, k_{i+1}, \ldots, k_n])$;
end
$X_n = \phi/\phi_1$;
return $(X_0, \ldots, X_n)$

Remark 5.4.2. Algorithm 1 works only if the rank conditions in (5.14) are satisfied. For any i with $r_i > r_{i-1}k_{i-1}$, we therefore replace r_i by $r_{i-1}k_{i-1}$ before running the algorithm. Note that the procedure can also be performed starting from the right-hand side, in which case one uses CX-decompositions instead of XR-decompositions.

We now describe the inverse of Algorithm 1, namely the *decompression* procedure that reconstructs a TT tensor from its parametrized form; see Algorithm 2.

Algorithm 2: Decompression $\texttt{decompress}(X_0, \ldots, X_n)$

Input : $(X_0, \ldots, X_n) \in \mathcal{W} \subseteq \mathcal{Z}$
Output: $\psi \in \mathcal{U} \subseteq \mathcal{T}^{=}_{\mathbf{k},\mathbf{r}}$

$\mathbf{r} := (1, r_1, \ldots, r_n, 1)$, $\psi := X_n$;
for $i = n - 1$ **to** 0 **do**
  $R_i = \texttt{unfold}(\psi, [r_{i+1},\ k_{i+1} \cdots k_n])$;
  $\psi^i = X_i R_i$;
  $\psi = \texttt{tensorize}(\psi^i, [r_i,\ k_i, \ldots, k_n])$;
end
return ψ

Theorem 5.4.3. *Algorithm 2 defines a one-to-one parametrization from $\mathcal{Z}$ to $\mathcal{T}^{=}_{\mathbf{k},\mathbf{r}}$. The restriction $\mathcal{Z}_{\mathbb{R}} \dashrightarrow \mathcal{T}^{=}_{\mathbf{k},\mathbf{r},\mathbb{R}}$ to the real points is birational.*

Proof. The subset $\mathcal{U} \subseteq \mathcal{T}^{=}_{\mathbf{k},\mathbf{r}}$ consists of all tensors satisfying Assumption 5.4.1 for each XR-factorization in Algorithm 1, i.e., the first r_{i+1} rows of ψ^i (Line 3) are linearly independent for $i = 0, \dots, n-1$. In other words, the matrix R_i (Line 4) must have full rank, that is, it must have some nonvanishing maximal minor, for all $i = 0, \dots, n-1$. In addition, the first entry of $X_n = R_{n-1}$ should be nonzero. We argue that this condition is algebraic in the entries of the input tensor ψ.

Indeed, the matrix R_0 is defined as the first r_1 rows of ψ^0, which is just a reshaping of the input tensor ψ. Hence all entries of R_0 are entries of the input tensor ψ. Similarly, since ψ^1 and R_1 are both flattenings of the same tensor, all entries of R_1 are also entries of R_0, and thus of the input tensor ψ. Proceeding in this manner, we observe that the entries of $R_0, \dots, R_{n-1}$ are always entries in the original ψ. Thus the condition that R_i has a nonvanishing maximal minor for $i = 0, \dots, n-1$ is an algebraic condition in the entries of the input tensor ψ. Hence, the open set $\mathcal{U}$ is Zariski dense in $\mathcal{T}^{=}_{\mathbf{k},\mathbf{r}}$.

We now prove that Algorithm 2 is a bijection from the open set $\mathcal{U} \subseteq \mathcal{T}^{=}_{\mathbf{k},\mathbf{r}}$ to the open set $\mathcal{W} \subseteq \mathcal{Z}$. We show that it is the inverse of Algorithm 1. We proceed by induction on n. For $n = 1$, the claim follows from the XR-decomposition. For general n, let $(X_0, \dots, X_n) \in \mathcal{W}$ and let $\psi = \texttt{decompress}(X_0, \dots, X_n)$. After the first iteration of the **for**-loop in Algorithm 1, we have recovered X_0 and ψ is now $\texttt{decompress}(X_1, \dots, X_n)$; this is because this first iteration of the **for**-loop in Algorithm 1 undoes the last iteration of the **for**-loop in Algorithm 2. By induction, the output of $\texttt{factorize}(\texttt{decompress}(X_1, \dots, X_n))$ matches $(X_1, \dots, X_n)$, and we have reached our conclusion. The argument for the reverse direction is analogous.

Both maps are algebraic when we restrict to the reals. Since they are inverses, they define birational maps. □

Remark 5.4.4. We believe that the statement that $\mathcal{Z}$ parametrizes $\mathcal{T}^{=}_{\mathbf{k},\mathbf{r}}$ is contained in [43], somewhat enigmatically. However, to the best of our knowledge, the resulting 1-to-1 parametrization of tensor train manifolds in Algorithm 2 is derived here for the first time.

The dimension of the manifold $\mathcal{T}^{=}_{\mathbf{k},\mathbf{r}}$ is known in the tensor decomposition literature; see, for example, [65]. It can also be derived from [26], where more general tensor network varieties were studied. We present it here as a corollary of Theorem 5.4.3.

Corollary 5.4.5. *The tensor train variety $\mathcal{T}_{\mathbf{k},\mathbf{r}}$ has dimension*

$$\dim(\mathcal{T}_{\mathbf{k},\mathbf{r}}) = \sum_{i=0}^{n-1} \left((r_i k_i - r_{i+1}) \cdot r_{i+1} \right) + r_n k_n - 1.$$

Example 5.4.6. We parameterize the tensor train variety $\mathcal{T}_{(3,2,2),(2,1)} \cong \mathbb{P}^5 \times \mathbb{P}^1$. We map from the open set $\mathcal{W} \subseteq \mathrm{Gr}(2,3) \times \mathrm{Gr}(1,4) \times \mathbb{P}^1$. The image of a point

$$X_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ a & b \end{pmatrix}, \quad X_1 = \begin{pmatrix} 1 \\ c \\ d \\ e \end{pmatrix}, \quad X_2 = \begin{pmatrix} 1 & f \end{pmatrix}$$

in $\mathcal{W}$ under Algorithm 2 is ψ with slices

$$\psi_{(1,:,:)} = \begin{pmatrix} 1 & f \\ c & cf \end{pmatrix} \qquad \psi_{(2,:,:)} = \begin{pmatrix} d & df \\ e & ef \end{pmatrix} \qquad \psi_{(3,:,:)} = \begin{pmatrix} a+bd & (a+bd)f \\ ac+be & (ac+be)f \end{pmatrix}.$$

We illustrate Algorithm 1 with an explicit numerical example. Consider the $3 \times 2 \times 2$ tensor ψ of TT-rank $(r_0, r_1, r_2, r_3) = (1, 2, 1, 1)$ whose slices are

$$\psi_{(1,:,:)} = \begin{pmatrix} 2 & 3 \\ 4 & 6 \end{pmatrix}, \quad \psi_{(2,:,:)} = \begin{pmatrix} 6 & 9 \\ 8 & 12 \end{pmatrix}, \quad \psi_{(3,:,:)} = \begin{pmatrix} -2 & -3 \\ -4 & -6 \end{pmatrix}.$$

Its first unfolding and decomposition in Algorithm 1 yield

$$\psi^0 = \begin{pmatrix} 2 & 3 & 4 & 6 \\ 6 & 9 & 8 & 12 \\ -2 & -3 & -4 & -6 \end{pmatrix} \quad X_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad R_0 = \begin{pmatrix} 2 & 3 & 4 & 6 \\ 6 & 9 & 8 & 12 \end{pmatrix}.$$

We then tensorize R_0 to obtain the $2 \times 2 \times 2$ tensor ϕ with slices

$$\phi_{(:,:,1)} = \begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix}, \quad \phi_{(:,:,2)} = \begin{pmatrix} 3 & 6 \\ 9 & 12 \end{pmatrix}.$$

This tensor unfolds and factors as

$$\psi^1 = \begin{pmatrix} 2 & 3 \\ 4 & 6 \\ 6 & 9 \\ 8 & 12 \end{pmatrix}, \quad X_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}, \quad R_1 = \begin{pmatrix} 2 & 3 \end{pmatrix}.$$

After normalizing so that the first entry of $R_1 = X_2$ is one, the output of the algorithm is

$$X_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & -1 \end{pmatrix} \quad X_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}, \quad X_2 = \begin{pmatrix} 1 & 3/2 \end{pmatrix}.$$

The first matrix is a point in $\mathrm{Gr}(2,3) \cong \mathbb{P}^2$, the second is a point in $\mathrm{Gr}(1,4) \cong \mathbb{P}^3$, and the third is a point in $\mathbb{P}^1$. $\diamond$

5.5 Numerical Experiments

In our last section, we compute the complex critical points of (5.1) over a tensor train variety using numerical algebraic geometry. For this, we use parameter homotopy and monodromy methods. We report the results of our computations for small determinantal varieties as well as for the first few instances of TT varieties that are Segre products by Theorem 5.3.6; see Tables 5.1 and 5.2. In many of these cases, the RR degrees are given by Proposition 5.1.6. However, we compute not only the degrees, but the actual critical points. One of our observations is that the number of local extrema is typically much smaller than the corresponding RR degree.

We also report the degree of the nonisotropic part of the RR discriminant (see Section 5.2) for some determinantal and TT varieties; see Tables 5.2 and 5.4. The take-away message there is that these degrees are high, implying that the computation of the actual discriminant polynomials are currently out of reach.

Finally, we try to gauge the quality of the low energy state solutions produced by the DMRG and ALS algorithms by comparing to the complete list of local and global minima produced by our numerical methods. We conclude that ALS frequently gets stuck in local minima with suboptimal energy, although this happens less often for physical Hamiltonians; see Table 5.5. The DMRG method does not converge to critical points of (5.1) at all.

Hamiltonians

In our experiments, whenever we look at a specific Hamiltonian H, we consider two types:

1. **Random symmetric real matrices.**

   We generate a symmetric real matrix $H \in \mathbb{R}^{(k_0\cdots k_n)\times(k_0\cdots k_n)}$ by drawing entries from the standard normal distribution and symmetrizing. For the ALS and DMRG experiments, H is then expressed in tensor operator form; see [103].

2. **Random Hamiltonians in second quantization.**

   In second-quantized form, the Schrödinger operator has the structure (see [123]):

$$H = \sum_{i,j} t_{ij}\mathbf{a}_i^\dagger \mathbf{a}_j + \sum_{i,j,k,l} v_{ijkl}\mathbf{a}_i^\dagger \mathbf{a}_j^\dagger \mathbf{a}_k \mathbf{a}_l, \tag{5.16}$$

   where $\mathbf{a}_i^\dagger = \mathbf{s}\otimes\cdots\otimes\mathbf{s}\otimes\mathbf{a}^\dagger\otimes\mathbf{I}\otimes\cdots\otimes\mathbf{I}$ and $\mathbf{a}_i = \mathbf{s}\otimes\cdots\otimes\mathbf{s}\otimes\mathbf{a}\otimes\mathbf{I}\otimes\cdots\otimes\mathbf{I}$ are the *creation* and *annihilation operators*, respectively. Here, we have

$$\mathbf{s} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \mathbf{a} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad \mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{5.17}$$

   The coefficients $t_{ij}, v_{ijkl} \in \mathbb{R}$ for $i,j,k,l \in \{0,1,\ldots,n\}$ are chosen randomly from the standard normal distribution. In this case, $\mathbf{k} = (2)^{n+1}$, and H is symmetric.

Homotopy Continuation Computations

A central task in our setting is to solve polynomial systems that encode the optimization problem (5.3) on tensor trains. There are many ways to write down such a system, for example by using (5.2), Proposition 5.1.2, or (5.10)

We combine the two homotopy continuation methods introduced in Section 2.5 to solve the optimization problem (5.3), in particular to compute RR degrees and to numerically determine the degrees of RR discriminants. All algorithms are implemented in our `Julia` package `TensorTrainOptimization.jl` using `HomotopyContinuation.jl` and `Oscar.jl` [102], which is provided together with the complete dataset of our experiments at:

https://zenodo.org/records/17777072.

r	$m\backslash n$	2	3	4	5	6	7	8
1	2	8	18	32	50	72	98	128
1	3		61	148	295	518	833	1 256
1	4			480	1 220	2 624	5 012	8 768
1	5				3 881	10 166	23 051	46 856
2	2	4	6	8	10	12	14	16
2	3		154	448	970	1 784	2 954	4 544
2	4			5 840	24 924	74 775	182 306	386 716
3	3		9	12	15	18	21	24
3	4			2 368	7 340	17 552	35 755	65 280
3	5				460 351			

Table 5.1: RR degrees of determinantal varieties of $m \times n$ matrices with rank $\leq r$.

Example 5.5.1 (RR degrees). We work with the Rayleigh–Ritz optimization problem (5.6) formulated on the sphere. Specifically, we apply Algorithm 1 to parametrize a tensor ψ in tensor train format for given $\mathbf{k}$ and $\mathbf{r}$. The function `get_T` returns this tensor as a vector of polynomials in $\dim \mathcal{T}_{\mathbf{k},\mathbf{r}}$ parameters. The critical equations of (5.6) form a polynomial system, which we solve via the monodromy method using the entries of H as parameters. The monodromy computation produces twice the RR degree many solutions; see the end of Section 5.1.

We perform these computations for low-rank matrices up to size 5 and for binary tensor train varieties up to order 6, with results reported in Tables 5.1 and 5.2. For instance, the following command computes the first entry of Table 5.1:

```
get_RRdegree_monodromy([2,2], [1])
```

The largest number 460 351 in Table 5.1 is the RR degree of the variety of 5×5 matrices of rank 3. It was computed in approximately 205 hours, using 29.465 GB of memory.

Proposition 5.5.2. *The smallest tensor train variety which is not a Segre product or determinantal variety is* $\mathbf{k} = (2)^6$ *and* $(r_1, r_2, r_3, r_4, r_5) = (1, 2, 2, 2, 1)$*. Its RR degree is* $\geq 691\,127$*.*

We computed this lower bound by running the monodromy method as described above. The computation did not terminate after 60 days. ◇

k	**r**	codim	RR deg	time	Disc. deg	time
$(2)^3$	$(1)^2$	5	48	$14s$	360	$1.8m$
$(2)^4$	$(1)^3$	12	384	$90s$	5 438	$3h$
$(2)^4$	$(1,2,1)$	10	352	$59s$	6 000	$6.4h$
$(2)^5$	$(1)^4$	27	3 840	$26m$		
$(2)^5$	$(1,2,1,1)$	25	4 608	$33m$		
$(2)^5$	$(1,2,2,1)$	22	2 752	$1h$		
$(2)^6$	$(1)^5$	58	46 080	$24h$		
$(2)^6$	$(1,2,1,1,1)$	56	69 120	$49h$		
$(2)^6$	$(1,1,2,1,1)$	56	69 120	$37h$		
$(2)^6$	$(1,2,2,1,1)$	53	67 583	$153h$		
$(2)^6$	$(1,2,1,2,1)$	55	134 656	$78h$		

Table 5.2: RR degrees and RR discriminant degrees for small tensor train varieties.

Example 5.5.3 (Optimization on Tensor Trains)**.** The computation proceeds as in Example 5.5.1: we first compute twice the RR degree many critical points for a random symmetric Hamiltonian via monodromy, and then use parameter homotopy to track them to the critical points of a given Hamiltonian H. For random real symmetric H, this parameter homotopy is still optimal, meaning that there are RR degree many reduced solutions to the target system.

The function `TT_optimization_monodromy` returns the RR degree, the number of real critical points, and vectors of all complex and real solutions. Applying `get_global_extrema` to the real solutions yields the global minimum and the corresponding ground state energy. The function `return_local_min_max` identifies all real critical points that are local minima or maxima by checking whether the Hessian of the Rayleigh–Ritz quotient is positive or negative definite. We find that the number of local extrema is typically much smaller than the RR degree. For example, the following command

```
TT_optimization_monodromy([2,2,2,2], [1,2,1]; H = M1, RRdeg = 352)
```

computes 352 critical points for the tensor train variety $\mathbb{P}^3 \times (\mathbb{P}^1)^2$ and the given Hamiltonian M_1. The global minimum value is -3.5692, with only 8 local extrema among the 352 critical points. These extremal values, computed using `return_local_min_max`, are listed in Table 5.3. We repeated this for four different real symmetric matrices M_2-M_5 and observe the same pattern of very few local extrema. When we compute the local extrema of five random Hamiltonians H_1-H_5 in second quantization, even fewer extrema can be observed. In fact, for H_1, H_2 and H_4 there is only one global minimum and maximum. For the last matrix H_5, we did not obtain any extrema and expect that the solution lies in the complement of the open set $\mathcal{W}$ defined in Algorithm 1, so our homotopy algorithm cannot reach it. All ten matrices in this example are available at the `Zenodo` website [17]. ◇

M	**Extremal values**
M_1	min: -3.5692, -3.2364, -2.4927, -1.9931; max: 2.4383, 3.4325, 3.4346, 3.4475
M_2	min: -3.0762, -3.023; max: 1.1842, 2.352, 2.4979, 2.5777
M_3	min: -3.8648, -2.8834; max: 1.7402, 1.9833, 2.0849, 3.2976
M_4	min: -4.2191, -2.4659, -2.2213; max: 1.9474, 2.6817, 3.1394
M_5	min: -4.1527, -3.2442; max: 0.7862, 1.9889, 2.1986, 3.7494, 3.8863

H	**Extremal values**
H_1	min: -11.0834; max: 1.2403
H_2	min: -5.7333; max: 8.0021
H_3	min: -2.8502, -2.3294; max: 5.4079, 6.4597
H_4	min: -2.1645; max: 6.0279
H_5	min: - max: -

Table 5.3: Minimum and maximum values on $\mathbb{P}^3 \times (\mathbb{P}^1)^2$ for Hamiltonians M_1-M_5 and H_1-H_5.

Physical Hamiltonians are typically very sparse, with many zero entries. Specializing parameters cannot increase the number of complex solutions, so the number of complex critical points for physical Hamiltonians is typically smaller than the RR degree.

Example 5.5.4 (Physical Hamiltonians)**.** Let $\mathbf{k} = (2)^3$ and $\mathbf{r} = (1)^2$. Consider the following Hamiltonian which was derived by Otto Schmidt from the Bose-Hubbard model [73]:

$$H = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -0.5 & -0.1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -0.1 & -0.5 & -0.1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -0.1 & -0.5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & -0.1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -0.1 & -1 & -0.1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -0.1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1.5 \end{pmatrix}.$$

The RR degree of the corresponding $\mathcal{T}_{\mathbf{k},\mathbf{r}}$ is 48, and for this particular Hamiltonian, (5.1) has 48 critical points. However, only 44 of them are in the open set $\mathcal{W}$ defined in Algorithm 1. In particular, the global minimum is one of the four points not in $\mathcal{W}$. More computations with physical Hamiltonians are available at the `Zenodo` website [17]. ◇

Example 5.5.5 (RR discriminant degrees)**.** We compute the degrees of the nonisotropic component of the RR discriminant from Section 5.2. To this end, we model the parametric ramification locus using the Lagrange multiplier equations from Example 5.5.1, augmented by the condition that their Jacobian is singular. The resulting polynomial system has a positive-dimensional solution set, so we intersect it with a general affine line in the H-space parametrized by t. The projection of the solutions onto the t-coordinate yields a linear section of the main component of the RR discriminant, and the number of such points equals its degree. This allows us to compute the degree of the discriminant without computing the discriminant itself, thereby avoiding a costly symbolic computation. We report these degrees for higher-order tensors in Table 5.2 and for determinantal varieties in Table 5.4. $\diamond$

r	$m\backslash n$	2	3	4	5	6	7
1	2	24	96	240	480	840	1 344
1	3		648	2 304	6 000	12 960	2 4696
2	3		2 430	10 944			

Table 5.4: Degree of the nonisotropic part of the RR discriminant $m \times n$ matrices of rank $\leq r$.

DMRG and ALS

In practice, the ground state energy of the Hamiltonian H is computed by solving (5.1) using the DMRG method [123]. For its derivation, we note that Algorithm 2 can be naturally generalized to take any tensor train $(A_0, \ldots, A_n) \in \prod_{i=0}^{n} \mathbb{C}^{r_i k_i \times r_{i+1}}$ as an input, where A_i is the collection of the $(r_i \times r_{i+1})$ parameter matrices $A_i^{(j_i)}$. We denote this generalized map by

$$\tau : \prod_{i=0}^{n} \mathbb{C}^{r_i k_i \times r_{i+1}} \to \mathbb{C}^{k_0 \times \cdots \times k_n}.$$

The DMRG is a two-site version of the ALS [64], which we introduce first. A random TT tensor is chosen as an initial guess. In each outer iteration loop (often called a *sweep*), every component A_i is updated independently and successively, starting at A_0 and ending at A_n. Typically, the sweep is then performed backwards (from A_n to A_0). These forward and backward sweeps are repeated until convergence.

In the i-th step of each sweep, only the component A_i is updated. The algorithm ensures that the left components $A_0, \ldots, A_{i-1}$ are left-unitary, that is, $A_j^\dagger A_j = I_{r_{j+1}}$ for $j = 0, \ldots, i-1$, and the right components $A_{i+1}, \ldots, A_n$ are right-unitary: $B_j B_j^\dagger = I_{r_j}$, where $B_j = \texttt{unfold}(A_j, [r_j, k_j r_{j+1}])$ for $j = i+1, \ldots, n$. The solution of the subproblem

$$\tilde{A}_i = \underset{Z \in \mathbb{C}^{r_i k_i \times r_{i+1}}}{\operatorname{argmin}} \frac{\langle \tau(A_0, \ldots, Z, \ldots, A_n),\, H\, \tau(A_0, \ldots, Z, \ldots, A_n) \rangle}{\langle \tau(A_0, \ldots, Z, \ldots, A_n),\, \tau(A_0, \ldots, Z, \ldots, A_n) \rangle} \tag{5.18}$$

can be found by computing the eigenvectors of a significantly reduced Hamiltonian using conventional methods. In a forward sweep, $\tilde{A}_i$ is subsequently left-orthogonalized (for example, via a QR decomposition, setting $A_i \leftarrow Q$) to prepare for the update of A_{i+1}.

Example 5.5.6 (Correctness of ALS)**.** For the ten matrices from Example 5.5.3, we computed all critical points and classified the local minima and maxima (Table 5.3). We then ran the ALS algorithm 1000 times per matrix, and recorded the approximations of the ground energy attained at convergence. For $\mathbf{k} = (2)^4$ and $\mathbf{r} = (1, 2, 1)$, these outcomes are summarized in Table 5.5, where in the second column we also list the smallest and largest eigenvalues on the full space. Here, V_i means an approximation to the lowest eigenvalue returned by the ALS method and C_i is the number of runs out of 1000 that produced V_i.

M	Extremal eigenvalues	V_1	C_1	V_2	C_2	V_3	C_3	V_4	C_4
1	-4.7543, 2.9994	-3.5692	336	-3.2364	332	-2.4927	204	-1.9931	128
2	-4.6467, 2.5529	-3.0231	568	-3.0762	432				
3	-5.4243, 3.3603	-3.8648	616	-2.8834	384				
4	-5.6478, 3.0117	-4.2191	710	-2.2213	146	-2.4659	144		
5	-4.7984, 3.9884	-4.1527	628	-3.2443	372				

Table 5.5: ALS statistics for $\mathbf{k} = (2, 2, 2, 2)$ and $\mathbf{r} = (1, 2, 1)$ for symmetric matrices M_1-M_5.

Comparing Tables 5.3 and 5.5 shows that ALS can converge to any local minimum with nonzero probability. Notably, the global minimum is not necessarily the most frequently observed outcome (see the second row of Table 5.5) and its energy can be very far from the ground state energy in the full space. Moreover, as the problem size increases, the number of local minima increases, and hence, empirically, ALS is increasingly likely to become trapped in a spurious local minimum. However, for physical Hamiltonians in second quantization, the spectrum typically contains fewer local minima, and ALS correspondingly performs better. For the matrices H_1, H_2 and H_4, there is only one global minimum, resulting in a 100% success rate of ALS. Experiments for larger cases are available online at [17]. $\diamond$

Since the initial guess determines the size of each component, the ALS algorithm is automatically constrained to the TT manifold $\mathcal{T}^{=}_{\mathbf{k},\mathbf{r}}$ and the TT rank remains fixed. In contrast, the DMRG method is rank adaptive. The initial guess is typically chosen to have TT rank $\mathbf{1}$, i.e., $r_0 = \cdots = r_n = 1$. Here, all but two neighboring components (say A_i and A_{i+1}) are fixed and orthogonalized as above. The subproblem (5.18) is then solved for these two components simultaneously and *without imposing the any constraint on* r_{i+1}. This yields a merged double component $\tilde{A}_{i,i+1} \in \mathbb{C}^{r_i k_i \times k_{i+1} r_{i+2}}$ that is subsequently decomposed to produce new A_i (left-orthogonal) and A_{i+1}. This is typically done using a truncated SVD. The updated rank r_{i+1} is chosen to meet prescribed error tolerances.

In theory, for a general initial guess and without rank restrictions, the DMRG method will eventually find the global optimum of the unconstrained Rayleigh quotient. However, for large systems, a maximum rank must be chosen to maintain computational feasibility. Once this maximum rank is reached, the DMRG suffers from the same issues as ALS: it might get stuck in local minima whose quality (in terms of the resulting energy) is unknown.

In our experiments, we have observed another issue with the DMRG. Since we want to compare the output of the DMRG with critical points on the manifold $\mathcal{T}_{\mathbf{k},\mathbf{r}}^{=}$, we must truncate the ranks to the maximum rank $\mathbf{r}$ in each step. While the solution to the subproblem is optimal in terms of energy, the rank truncation of the solution is not. The truncated SVD yields the optimal approximation in the Frobenius norm but not in energy. Hence, the DMRG does not converge to critical points of the Rayleigh quotient on the manifold, but to other points with suboptimal energy. We do not report on the solutions of DMRG, as they are not critical points on the manifold. The DMRG analog of Table 5.5 is available in [17].

Chapter 6

Distance Minimization in the Grassmannian of Lines

We now turn our attention to the Grassmannian $\mathrm{Gr}(2,n)$ of lines in projective $(n-1)$-space $\mathbb{P}^{n-1}$. It is embedded in the projective space in $\mathbb{P}^{\binom{n}{2}-1}$ as the variety of skew-symmetric $n \times n$ matrices $X = (x_{ij})$ of rank 2. The matrix entries x_{ij} are the Plücker coordinates. By Proposition 3.2.2, the homogeneous prime ideal of $\mathrm{Gr}(2,n)$ is generated by the 4×4 Pfaffians of our skew-symmetric matrix:

$$\mathrm{Pf}\left(\begin{pmatrix} 0 & x_{ij} & x_{ik} & x_{il} \\ -x_{ij} & 0 & x_{jk} & x_{jl} \\ -x_{ik} & -x_{jk} & 0 & x_{kl} \\ -x_{il} & -x_{jl} & -x_{kl} & 0 \end{pmatrix}\right) = x_{ij}x_{kl} - x_{ik}x_{jl} + x_{il}x_{jk} \quad \text{for } 1 \le i < j < k < l \le n. \tag{6.1}$$

Recall from Section 3.3 that the projection Grassmannian is an affine variety. By Theorem 3.6.2, one obtains the projection matrix from the Plücker coordinates by scaling the matrix $P = X^2$ to have trace 2. We now distinguish between the image of the squaring map $X \mapsto X^2 = P$ in $\mathbb{P}^{\binom{n+1}{2}}$ and the point in the projection Grassmannian

$$\tilde{P} \;\coloneqq\; \frac{2}{\mathrm{trace}(P)} P \;=\; \frac{2}{||X||^2} X^2 \tag{6.2}$$

where $||X||^2$ denotes the Frobenius norm of X:

$$\mathrm{trace}(P) \;=\; p_{11} + \cdots + p_{nn} \;=\; 2(x_{12}^2 + x_{13}^2 + \cdots + x_{n-1,n}^2) \;=\; ||X||^2.$$

Given two points in $\mathrm{Gr}(2,n)$, encoded by projection matrices $\tilde{P} = (\tilde{p}_{ij})$ and $\tilde{Q} = (\tilde{q}_{ij})$ as in (6.2), we measure their distance in the Frobenius norm on the space $\mathrm{Sym}^2\mathbb{R}^n$:

$$||\tilde{P} - \tilde{Q}|| \;=\; \left(\sum_{i=1}^{n} (\tilde{p}_{ii} - \tilde{q}_{ii})^2 \;+\; 2\sum_{i<j} (\tilde{p}_{ij} - \tilde{q}_{ij})^2\right)^{1/2}. \tag{6.3}$$

We consider the Euclidean distance problem for a subvariety $\mathcal{M}$ in $\mathrm{Gr}(2, n)$ from the perspective of Chapter 2. The model $\mathcal{M}$ is a family of lines in $\mathbb{P}^{n-1}$. We are given a data point Q which is also a line in $\mathbb{P}^{n-1}$. This differs from the setting of Chapter 2, where the data was generic. Our task is to compute a line P in the model $\mathcal{M}$ which best approximates the given line Q. Thus we wish to solve the optimization problem

$$\text{Minimize } ||\tilde{P} - \tilde{Q}||^2 \;\text{ subject to } \tilde{P} \in \mathcal{M}. \tag{6.4}$$

The key parameter is the *Grassmann distance degree* (GD degree) of $\mathcal{M}$. We define this as the number of complex critical points of (6.4) for generic $Q \in \mathrm{Gr}(2, n)$.

The GD degree provides an algebraic alternative to the Grassmann distance complexity due to Lerario and Rosana [86]. Their work rests on the *geodesic distance*, which is not an algebraic function [86, Section 3.6].

Conway, Hardin and Sloane [30] refer to our metric (6.3) as the *chordal distance*, and they report: *"We have made extensive computations on ... optimal packings. These computations have led us to conclude that the best definition of distance on Grassmannian space is the chordal distance."* We note that the chordal distance between two lines in $\mathbb{P}^{n-1}$ is equal to

$$||\tilde{P} - \tilde{Q}||^2 \;=\; \mathrm{trace}\big(\,(\tilde{P} - \tilde{Q})^2\,\big) \;=\; 4 - 2 \cdot \mathrm{trace}(\tilde{P}\tilde{Q}) \;=\; 4 - 2\lambda - 2\mu, \tag{6.5}$$

where $\lambda \geq \mu$ are the non-zero eigenvalues of the product $\tilde{P}\tilde{Q}$ of the two projection matrices.

Many metrics in Grassmannian optimization are functions of the eigenvalues λ and μ; see [130, Theorem 2]. The following table, derived from [130, Table 2], shows nine options:

	Distance squared in λ, μ	Minimizer
Chordal	$4 - 2\lambda - 2\mu$	$\star$
Geodesic	$\arccos(\sqrt{\lambda})^2 + \arccos(\sqrt{\mu})^2$	$\circ$
Procrustes	$4 - 2\sqrt{\lambda} - 2\sqrt{\mu}$	$\diamond$
Binet-Cauchy	$1 - \lambda\mu$	$\triangle$
Fubini-Study	$\arccos(\sqrt{\lambda\mu})^2$	$\triangle$
Martin	$-\log(\lambda\mu)$	$\triangle$
Asimov	$\arccos(\sqrt{\mu})^2$	$\square$
Projection	$1 - \mu$	$\square$
Spectral	$2 - 2\sqrt{\mu}$	$\square$

While our focus is on the chordal distance (6.3)-(6.5), the algebraic tools introduced in this chapter are applicable to any of the metrics above. The following example demonstrates this.

Example 6.0.1 (Schubert surface)**.** Let $n = 5$. For our model, we take the surface

$$\mathcal{S}_{25} \;=\; \{X \in \mathrm{Gr}(2, 5) \,:\, x_{12} = x_{13} = x_{14} = x_{15} = x_{23} = x_{24} = x_{34} \;=\; 0\}.$$

We write the model as a projection matrix $\tilde{P}$, and we fix a projection matrix $\tilde{Q}$ for the data:

$$\tilde{P} = \frac{1}{x_{25}^2+x_{35}^2+x_{45}^2}\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & x_{25}^2 & x_{25}x_{35} & x_{25}x_{45} & 0 \\ 0 & x_{25}x_{35} & x_{35}^2 & x_{35}x_{45} & 0 \\ 0 & x_{25}x_{45} & x_{35}x_{45} & x_{45}^2 & 0 \\ 0 & 0 & 0 & 0 & x_{25}^2+x_{35}^2+x_{45}^2 \end{pmatrix}, \quad \tilde{Q} = \frac{1}{10}\begin{pmatrix} 6 & 4 & 2 & 0 & -2 \\ 4 & 3 & 2 & 1 & 0 \\ 2 & 2 & 2 & 2 & 2 \\ 0 & 1 & 2 & 3 & 4 \\ -2 & 0 & 2 & 4 & 6 \end{pmatrix}.$$

We define the *spectral region* of the pair $(\mathcal{M}, \tilde{Q})$ as the set of all pairs $\lambda \geq \mu$ of eigenvalues of the matrices $\tilde{P}\tilde{Q}$, where $\tilde{Q}$ is fixed and $\tilde{P}$ ranges over the real points of the model $\mathcal{M}$.

This spectral region is shown in Figure 6.1. Its lower left corner is $(3/5, 0)$. The eigenvalues are bounded above by $\lambda \leq 1$ and $\mu \leq 2/5$. The nonlinear boundary curve is the quartic

$$100\lambda^2\mu^2 - 100\lambda^2\mu + 49\lambda^2 - 100\lambda\mu^2 + 170\lambda\mu - 84\lambda + 49\mu^2 - 84\mu + 36 \;=\; 0. \tag{6.6}$$

This specifies the *Pareto frontier*. The map from the surface $\mathcal{M}$ onto the (λ, μ)-plane is 4-to-1. The nine metrics give rise to five distinct minimizers on the Pareto frontier. For instance, the circle is for the geodesic distance in [86]. The minimizer for our distance (6.3)-(6.5), indicated by a star, satisfies $x_{25} = x_{35} = x_{45} = 1$, and $\lambda = \frac{3+\sqrt{3}}{5} = 0.9464$, $\mu = \frac{3-\sqrt{3}}{5} = 0.2536$, with optimal value $8/5$. The curve (6.6) can be viewed as the universal solution for all nine metrics. ⋄

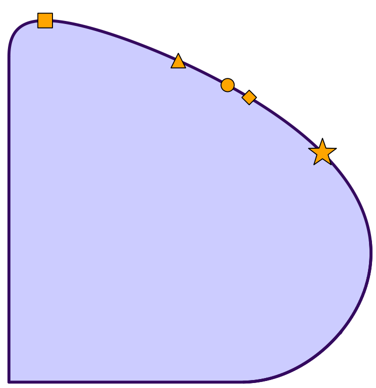

Figure 6.1: The spectral region comprises all eigenvalue pairs (λ, μ) as X ranges over $\mathcal{M}$. The marked points are the optimal solutions for the nine metrics on $\mathrm{Gr}(2, 5)$ in the table above.

We now discuss the organization and contributions of this chapter. In Section 6.1 we study the squaring map $X \mapsto X^2$ for skew-symmetric $n \times n$ matrices X. Theorem 6.1.1 characterizes its image and fibers. The restriction to the rank 2 locus identifies $\mathrm{Gr}(2, n)$ with the projection Grassmannian $\mathrm{pGr}(2, n)$, as shown in Theorem 3.6.2. Theorems 6.1.9 and 6.1.13 feature varieties that arise naturally from multiplying skew-symmetric matrices. In Section 6.2 we introduce the polynomial ideals that characterize the critical points of (6.4), where $\tilde{Q}$ is an arbitrary symmetric $n \times n$ matrix. Theorem 6.2.3 identifies what is special about this ED problem. We use Theorem 2.4.7 to give an upper bound on the ED degree of a unirational model in $\mathrm{Gr}(2, n)$; see Corollary 6.2.6.

In Section 6.3 we turn to the GD problem. Now the data point $\tilde{Q}$ itself lies in $\mathrm{pGr}(2, n)$. The critical ideal and the GD degree are introduced in Definition 6.3.2. Their computation requires the removal of the algebraic cut locus. Theorem 6.3.7 explains the difference between the ED degree and the GD degree. Section 6.4 is devoted to lines in 3-space. Models $\mathcal{M}$ in $\mathrm{Gr}(2, 4)$ have dimension one (ruled surfaces), two (congruences) and three (complexes). They appear in computer vision [106], physics [63], and many other applications. Degree formulas for generic complete intersection are presented in Theorem 6.4.1 and Conjecture 6.4.2.

In Section 6.5 we discuss our optimization problem for Schubert varieties in $\mathrm{Gr}(2, n)$. Conjectures 6.5.1 and 6.5.4 concern the embeddings of Schubert varieties in varieties in projection coordinates. Formulas for their ED and GD degrees are given in Theorem 6.5.6 and Conjectures 6.5.5 and 6.5.7.

6.1 Squaring Skew-Symmetric Matrices

Our starting point for this section is the map in Section 3.6 from the Grassmannian $\mathrm{Gr}(2, n)$ to the projection Grassmannian $\mathrm{pGr}(2, n)$. A point in $\mathrm{Gr}(2, n)$ is the row span of a $2 \times n$ matrix $A = (a_{ij})$, and the cocircuit matrix may be written as:

$$X = (x_{ij}) \;=\; A^T \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} A. \tag{6.7}$$

This is a skew-symmetric $n \times n$ matrix of rank 2. In what follows, we examine this passage from Plücker coordinates to projection coordinates from a more general perspective. Namely, we temporarily drop the hypothesis that X has rank two.

Taking the square is a rational map from skew-symmetric matrices to symmetric matrices:

$$(-)^2 \,:\, \mathbb{P}(\wedge^2 \mathbb{C}^n) \dashrightarrow \mathbb{P}(\mathrm{Sym}^2 \mathbb{C}^n), \qquad X \mapsto X^2. \tag{6.8}$$

This map is rational but it is not a morphism. Its base locus contains complex matrices like

$$X \;=\; \begin{pmatrix} 0 & 1 & i & 0 \\ -1 & 0 & 0 & i \\ -i & 0 & 0 & -1 \\ 0 & -i & 1 & 0 \end{pmatrix} \in \mathrm{Gr}(2, 4) \subset \mathbb{P}(\wedge^2 \mathbb{C}^4). \tag{6.9}$$

We write $\mathcal{V}_{X^2}$ for the Zariski closure in $\mathbb{P}(\mathrm{Sym}^2 \mathbb{C}^n)$ of the image of the squaring map $(-)^2$.

Theorem 6.1.1. *The variety $\mathcal{V}_{X^2}$ has the expected dimension $\binom{n}{2} - 1$. It is the Zariski closure of the set of symmetric diagonalizable matrices whose nonzero eigenvalues have multiplicity 2. The fiber of $(-)^2$ containing a generic rank $2r$ skew-symmetric $n \times n$ matrix X has size 2^{r-1}.*

The proof rests on the conjugation action of the orthogonal group $\mathrm{O}(n)$ on the space $\wedge^2\mathbb{C}^n$. Normal forms for the orbits are given in [36, 53]. The eigenvalues can be read off of the normal form in Gantmacher's book [53, §XI.4]. We prefer it for that reason. The nonzero eigenvalues of skew-symmetric matrices come in pairs $\pm\lambda$. This is clear from the elementary divisors.

Lemma 6.1.2 ([53, Theorem 7])**.** *If $\lambda \neq 0$ and $(z-\lambda)^p$ is an elementary divisor of a skew-symmetric matrix X, then so is $(z+\lambda)^p$. Moreover, even powers of z come in pairs.*

The normal form uses $2p \times 2p$ blocks for the pair of elementary divisors $(z \pm \lambda)^p$, and $q \times q$ matrices for elementary divisors z^q with q odd. A generic rank $2r$ skew-symmetric $n \times n$ matrix X has $2r+1$ distinct eigenvalues $\pm\lambda_1 i, \ldots, \pm\lambda_r i, 0$ or, if $n = 2r$, it has $2r$ eigenvalues $\pm\lambda_1 i, \ldots, \pm\lambda_r i$. Its normal form is a direct sum of 2×2 blocks and 1×1 zero-blocks:

$$X \;=\; \begin{pmatrix} 0 & \lambda_1 \\ -\lambda_1 & 0 \end{pmatrix} \oplus \begin{pmatrix} 0 & \lambda_2 \\ -\lambda_2 & 0 \end{pmatrix} \oplus \cdots \oplus \begin{pmatrix} 0 & \lambda_r \\ -\lambda_r & 0 \end{pmatrix} \oplus [0] \oplus \cdots \oplus [0]. \tag{6.10}$$

The example in (6.9) is a skew-symmetric matrix that cannot be put in the form (6.10). In fact, it is already in Gantmacher's normal form; compare with the Jordan canonical form.

Proof of Theorem 6.1.1. The squaring map commutes with the $\mathrm{O}(n)$-action. Given a generic point of rank $2r$ in the domain, we consider its normal form (6.10). Its image under (6.8) is

$$-X^2 \;=\; \begin{pmatrix} \lambda_1^2 & 0 \\ 0 & \lambda_1^2 \end{pmatrix} \oplus \begin{pmatrix} \lambda_2^2 & 0 \\ 0 & \lambda_2^2 \end{pmatrix} \oplus \cdots \oplus \begin{pmatrix} \lambda_r^2 & 0 \\ 0 & \lambda_r^2 \end{pmatrix} \oplus [0] \oplus \cdots \oplus [0].$$

Hence $P = X^2$ has $r+1$ distinct eigenvalues (or r if $n = 2r$), and all nonzero eigenvalues have multiplicity 2. The fiber over P is given by the 2^r matrices which are obtained by multiplying each λ_i with $+1$ or -1. This gives rise to 2^{r-1} elements X in $\mathbb{P}(\wedge^2\mathbb{C}^n)$.

Since the set of full rank skew-symmetric matrices with normal form (6.10) is Zariski dense in the domain $\mathbb{P}(\wedge^2\mathbb{C}^n)$, the image of this set is Zariski dense in the image variety $\mathcal{V}_{X^2}$. $\square$

Over the real numbers $\mathbb{R}$, every symmetric $n \times n$ matrix is diagonalizable. It is equivalent to a diagonal matrix under the $\mathrm{O}(n)$-action. This is not true over the complex numbers $\mathbb{C}$. It thus follows from Theorem 6.1.1 that the complex variety $\mathcal{V}_{X^2}$ is the Zariski closure of the set of real symmetric matrices whose nonzero eigenvalues come in pairs. The varieties in $\mathbb{P}(\mathrm{Sym}^2\mathbb{R}^n)$ given by prescribed eigenvalue multiplicities were studied by Weinstein in [127].

Corollary 6.1.3. *The image of the real locus $\mathbb{P}(\wedge^2\mathbb{R}^n)$ under (6.8) consists of negative semidefinite matrices $P \in \mathrm{Sym}^2\mathbb{R}^n$ whose nonzero eigenvalues have even multiplicity. If such a matrix P has maximal rank then every complex matrix $X \in \wedge^2\mathbb{C}^n$ with $X^2 = P$ is real.*

Proof. The first claim holds because every real skew-symmetric matrix has normal form (6.10). Consider a real symmetric matrix P of maximal rank in the image of the squaring map. All elementary divisors of P are linear. If $X^2 = P$, then the nonzero elementary divisors of X are linear as well. Since P has at most one 0 eigenvalue, X has normal form

(6.10). Furthermore, since P is negative semidefinite, the eigenvalues of X are of the form $\pm\lambda_j i$ with λ_j real, so X has real normal form. The matrix X is obtained from the normal form by conjugating with a real orthogonal matrix that diagonalizes P. Hence every matrix in the fiber is real. $\square$

We now discuss the ideal of the image variety $\mathcal{V}_{X^2}$. The first non-trivial case is $n = 3$.

Example 6.1.4 ($n = 3$)**.** Here (6.8) is the Veronese embedding of $\mathbb{P}^2$ into $\mathbb{P}^5$. Thus $\mathcal{V}_{X^2}$ is a surface of degree 4. Its prime ideal is generated by the 2×2 minors of the symmetric matrix

$$2P - \mathrm{trace}(P) \cdot \mathrm{Id}_3 \;=\; 2\begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{12} & p_{22} & p_{23} \\ p_{13} & p_{23} & p_{33} \end{pmatrix} - (p_{11} + p_{22} + p_{33}) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \qquad \diamond$$

Example 6.1.5 ($n = 4$)**.** The variety $\mathcal{V}_{X^2}$ has dimension 5 and degree 6 in the $\mathbb{P}^9$ of symmetric 4×4 matrices $P = (p_{ij})$. Its prime ideal is generated by the 2×2 minors of the matrix

$$\begin{pmatrix} p_{11} - p_{22} - p_{33} + p_{44} & 2p_{13} + 2p_{24} & 2p_{12} - 2p_{34} \\ 2p_{13} - 2p_{24} & -p_{11} - p_{22} + p_{33} + p_{44} & 2p_{14} + 2p_{23} \\ 2p_{12} + 2p_{34} & -2p_{14} + 2p_{23} & -p_{11} + p_{22} - p_{33} + p_{44} \end{pmatrix}.$$

Hence $\mathcal{V}_{X^2}$ is a cone over the Segre variety $\mathbb{P}^2 \times \mathbb{P}^2$. Its apex is the 4×4 identity matrix. $\diamond$

Example 6.1.6 ($n = 5, 6, 7$)**.** The variety $\mathcal{V}_{X^2}$ has dimension 9 and degree 46 in the $\mathbb{P}^{14}$ of symmetric 5×5 matrices. Its ideal is minimally generated by 15 cubics. For $n = 6$, the degree is 92, and its ideal contains 20 cubics. For $n = 7$, it has dimension 20 and degree 1 072. $\diamond$

Conjecture 6.1.7. The prime ideal of the variety $\mathcal{V}_{X^2}$ is generated in degree $\lceil n/2 \rceil$.

The Grassmannian $\mathrm{Gr}(2, n)$ is the rank 2 locus in $\mathbb{P}(\wedge^2 \mathbb{C}^n)$. The rank $2r$ locus is the secant variety $\mathrm{Sec}_r(\mathrm{Gr}(2, n))$. We know from Theorem 6.1.1 that the degree of the map (6.8) is 2^{r-1} on this secant variety. For $r = 1$, the image of $\mathrm{Gr}(2, n)$ is the projective closure of the affine variety $\mathrm{pGr}(2, n)$ from [34, Section 5]. We denote this by $\overline{\mathrm{p}}\mathrm{Gr}(2, n) \subset \mathbb{P}(\mathrm{Sym}^2 \mathbb{C}^n)$ and we also call it the *projection Grassmannian*.

Remark 6.1.8. The variety $\overline{\mathrm{p}}\mathrm{Gr}(2, n)$ has dimension $2n - 4$ and degree $2\binom{2n-4}{n-2}$; see Proposition 3.3.6. Its prime ideal has $\binom{n+1}{2}$ quadratic generators, namely the entries of $2P^2 - \mathrm{trace}(P) \cdot P$. For $n \geq 6$, we also need the 3×3 minors of P. A general statement appears in Conjecture 6.5.1.

The singular locus of $\overline{\mathrm{p}}\mathrm{Gr}(2, n)$ has dimension $n - 2$ and degree 2^{n-1}. It consists of all traceless symmetric matrices of rank 1. This is a complex variety with no real points, namely it is the image of the Fermat quadric $V(x_1^2 + x_2^2 + \cdots + x_n^2)$ under the second Veronese map. This singular locus is the Zariski closure of the image under (6.8) of the $\mathrm{O}(n)$-orbit of (6.18).

The map most relevant for our application to distance optimization is the composition

$$\mathbb{P}(\wedge^2\mathbb{C}^n) \dashrightarrow \mathbb{P}(\mathrm{Sym}^2\mathbb{C}^n) \dashrightarrow \mathrm{Sym}^2\mathbb{C}^n, \quad X \mapsto X^2 = P, \quad P \mapsto \frac{2}{\mathrm{trace}(P)}P = \tilde{P}. \quad (6.11)$$

The second map is the identity on the affine open chart defined by $\mathrm{trace}(P) \neq 0$. The base locus of the composition consists of all skew-symmetric matrices X such that $\mathrm{trace}(X^2) = 0$. This is a hypersurface in $\mathbb{P}(\wedge^2\mathbb{C}^n)$ which has no real points. For a complex point see (6.18). The first map is (6.8). Its base locus is the set of skew-symmetric matrices X with $X^2 = 0$.

Theorem 6.1.9. *The* $\mathrm{O}(n)$*-orbit of a matrix* B *with* k *blocks* (6.9) *and* $r = n - 4k$ *blocks* $(\,0\,)$ *has codimension* $4k^2 + 2kr + \binom{r}{2}$. *The base locus of* (6.8) *is the union of all such orbits. It has two irreducible components if* n *is a multiple of* 4 *and is irreducible otherwise.*

Proof. The codimension of the orbit of B is equal to the dimension of the stabilizer $\{R \in \mathrm{SO}(n) : R^TBR = B\}$. We compute the dimension of the corresponding Lie algebra, which is $\{X \in \wedge^2\mathbb{C}^n : [X, B] = 0\}$. Let M be the direct sum of k copies of (6.9). We write skew-symmetric $n \times n$ matrices as $X = \begin{pmatrix} Y & Z \\ Z^T & W \end{pmatrix}$ where Y and W are skew-symmetric and Z is any $4k \times r$ matrix. The condition $[X, B] = 0$ is equivalent to $[M, Y] = 0$ and $AZ = 0$. We have $\binom{r}{2}$ degrees of freedom to choose W. Since (6.9) has rank 2, M has nullity $2k$, so we have $2kr$ degrees of freedom to choose Z. We partition Y into 4×4 blocks that commute with (6.9), with skew-symmetric blocks on the diagonal and general blocks on the off-diagonal. The space of skew-symmetric and general matrices that commute with (6.9) have dimensions 4 and 8 respectively, so we have $4k + 8\binom{k}{2} = 4k^2$ degrees of freedom to choose Y.

The only block in the normal form [53, §XI.4] whose square is the zero matrix is (6.9); this proves the second claim. The last claim follows from the closure ordering of nilpotent orbits [29, Theorem 6.2.5]. □

Example 6.1.10. By Theorem 6.1.9, the base locus of (6.8) is empty for $n = 2, 3$. For $n = 4$, it is the Zariski closure of the $\mathrm{O}(4)$-orbit of (6.9) and has two components. Its ideal is

$$\langle x_{14} - x_{23}, x_{13} + x_{24}, x_{12} - x_{34}, x_{23}^2 + x_{24}^2 + x_{34}^2\rangle \cap \langle x_{14} + x_{23}, x_{13} - x_{24}, x_{12} + x_{34}, x_{23}^2 + x_{24}^2 + x_{34}^2\rangle.$$

For $n = 5, 6, 7$, the base locus is irreducible. It is the orbit-closure of (6.9) padded by zeros. ⋄

We are especially interested in the restriction of the base locus to matrices X of rank 2.

Corollary 6.1.11. *The base locus of* (6.8) *on the Grassmannian* $\mathrm{Gr}(2, n)$ *is the* $\mathrm{O}(n)$*-orbit closure of the matrix* (6.9) *padded with zeros. It is irreducible of dimension* $2n - 7$ *for* $n \geq 5$.

Proof. This follows from Theorem 6.1.9 by restricting to $k = 1$ and $r = n - 4$. □

Remark 6.1.12. The base locus has codimension 3 in $\mathrm{Gr}(2, n)$. Its prime ideal is minimally generated by the $\binom{n}{4}$ Plücker quadrics and the $\binom{n+1}{2}$ entries of X^2. Computations show that its degree equals $\frac{4}{n-2}\binom{2n-6}{n-3}$ for $n = 4, 5, \ldots, 14$. The first values are $4, 8, 20, 56, 168, 528, 1716$.

A natural extension of (6.8) is the map which multiplies two skew-symmetric matrices X and Y. This operation was studied in 1922 by Stenzel [120]. This was followed up in the linear algebra literature throughout the 20th century. We refer to the bibliography in [71]. We define $\mathcal{V}_{XY}$, $\mathcal{V}_{X^2Y^2}$ and $\mathcal{V}_{XY^2X}$ to be the varieties respectively parametrized by all products XY, X^2Y^2 and XY^2X, where $X, Y \in \mathbb{P}(\wedge^2\mathbb{C}^n)$. The first two varieties lie in the space $\mathbb{P}(\mathbb{C}^{n\times n})$ of all $n \times n$ matrices, while $\mathcal{V}_{XY^2X} \subset \mathbb{P}(\mathrm{Sym}^2\mathbb{C}^n)$. We call $\mathcal{V}_{XY}$ the *Stenzel variety*. For any of these three varieties $\mathcal{V}_\bullet$ we write $\mathcal{V}_{\bullet,\leq 2}$ for the subvariety consisting of matrices of rank at most two. These are obtained by taking X and Y from the Grassmannian $\mathrm{Gr}(2, n)$.

Theorem 6.1.13. *The Stenzel variety $\mathcal{V}_{XY}$ is the Zariski closure of diagonalizable $n \times n$ matrices whose nonzero eigenvalues have multiplicity 2. The varieties $\mathcal{V}_{XY}$ and $\mathcal{V}_{X^2Y^2}$ have dimension $n(n-1) - \lfloor n/2 \rfloor - 1$. The varieties $\mathcal{V}_{XY,\leq 2}$ and $\mathcal{V}_{X^2Y^2,\leq 2}$ have dimension $4n - 8$.*

Proof. Following Stenzel [120, Satz VII in Section 4], the elementary divisors of the product XY come in pairs. The generic point in $\mathcal{V}_{XY}$ is a diagonalizable matrix with nonzero eigenvalues of multiplicity 2. We derive the dimension of the Stenzel variety $\mathcal{V}_{XY}$ from this description. Consider first the case where n is even. Then $\mathcal{V}_{XY}$ is the $\mathrm{GL}(n)$-orbit closure of the set of diagonal matrices $\mathrm{diag}(\lambda_1, \lambda_1, \ldots, \lambda_{n/2}, \lambda_{n/2})$. The group $\mathrm{GL}(n)$ has dimension n^2 and the stabilizer of the action, namely the set of invertible matrices with 2×2 blocks, has dimension $4n/2 = 2n$. Adding the $n/2$ degrees of freedom for choosing the λ_i, this set has dimension $n^2 - 2n + n/2 - 1 = n(n-1) - \lfloor n/2 \rfloor - 1$ in projective space. For odd n, we take the orbit of $\mathrm{diag}(\lambda_1, \lambda_1, \ldots, \lambda_{n/2}, \lambda_{n/2}, 0)$. The stabilizer now has size $4(n-1)/2+1 = 2n-1$, so the projective dimension is $n^2 - (2n-1) + (n-1)/2 - 1 = n(n-1) - \lfloor n/2 \rfloor - 1$.

The dimension of $\mathcal{V}_{XY,\leq 2}$ is found in the same way. We take the $\mathrm{GL}(n)$-orbit closure of $\mathrm{diag}(\lambda_1, \lambda_1, 0, \ldots, 0)$. The dimension is $n^2 - 4 - (n-2)^2 + 1 - 1 = 4n - 8$. Finally, $\mathcal{V}_{XY}$ and $\mathcal{V}_{X^2Y^2}$ have the same dimension, as do their restrictions to rank 2. We omit the proof. □

Example 6.1.14 ($n = 3$)**.** The Stenzel variety $\mathcal{V}_{XY}$ has dimension four, and it is the Segre variety $\mathbb{P}^2 \times \mathbb{P}^2 \subset \mathbb{P}^8$. Its prime ideal is generated by the 2×2 minors of the matrix

$$2P - \mathrm{trace}(P) \cdot \mathrm{Id}_3 \;=\; \begin{pmatrix} p_{11} - p_{22} - p_{33} & 2p_{12} & 2p_{13} \\ 2p_{21} & -p_{11} + p_{22} - p_{33} & 2p_{23} \\ 2p_{31} & 2p_{32} & -p_{11} - p_{22} + p_{33} \end{pmatrix}.$$

The fourfold $\mathcal{V}_{X^2Y^2}$ has degree 20 in $\mathbb{P}^8$, with ideal generated by 10 cubics and 13 quartics. ⋄

Example 6.1.15 ($n = 4$)**.** The Stenzel variety $\mathcal{V}_{XY}$ has dimension 9 and degree 14 in $\mathbb{P}^{15}$. Its ideal is generated by 15 quadrics, which look like the 4×4 Pfaffians of a skew-symmetric 6×6 matrix of linear forms. Its subvariety $\mathcal{V}_{XY,\leq 2}$ has dimension 8 and degree 28, with prime ideal generated by 16 quadrics. Both varieties are arithmetically Gorenstein.

The varieties $\mathcal{V}_{X^2Y^2}$ and $\mathcal{V}_{X^2Y^2,\leq 2}$ have the same dimensions as above, namely 9 and 8 respectively. The latter is defined by 28 cubics and 16 quartics, and it has degree 196. ⋄

The variety $\mathcal{V}_{XY^2X}$ arises because the symmetric matrix XY^2X has the same eigenvalues as the non-symmetric matrix X^2Y^2. These eigenvalues are real whenever X and Y are real.

Example 6.1.16. Suppose $n \geq 5$. If n is odd then $\mathcal{V}_{XY^2X}$ is the hypersurface defined by the symmetric determinant. For even n, it is the ambient space $\mathbb{P}(\mathrm{Sym}^2\mathbb{C}^n)$. The variety $\mathcal{V}_{XY^2X,\leq 2}$ consists of $n \times n$ matrices of rank ≤ 2; it has dimension $2n-2$ and degree $\frac{1}{2}\binom{2n-2}{n-1}$. For $n = 4$, the dimension is one less than expected. Here, $\mathcal{V}_{XY^2X}$ is a hypersurface in the $\mathbb{P}^9$ of symmetric 4×4 matrices. Its defining polynomial has degree 6 and 260 terms. ⋄

6.2 Critical Equations

Our main goal is to solve the optimization problem (6.4). We seek exact solutions, to be reached by algebraic methods. This rests on polynomial equations for the critical points as in Chapter 2. We now introduce these critical equations, both for implicit models and parametric models. In this section, the data Q is a generic symmetric matrix, and the number of critical points is the ED degree of $\mathcal{M}$. In Section 6.3, the data will come from the projection Grassmannian.

The model $\mathcal{M}$ is an irreducible subvariety of $\mathrm{Gr}(2,n)$, which is defined over $\mathbb{R}$ and whose real points are Zariski dense. Its ideal is generated by polynomials in the Plücker coordinates:

$$I_{\mathcal{M}} \;=\; \langle f_1, f_2, \ldots, f_r \rangle \;+\; I_{\mathrm{Gr}(2,n)} \;\subset\; \mathbb{R}\big[\, x_{ij} : 1 \leq i < j \leq n \,\big]. \tag{6.12}$$

Here $I_{\mathrm{Gr}(2,n)}$ is the ideal generated by the $\binom{n}{4}$ Plücker quadrics (6.1). We write $c = \mathrm{codim}(\mathcal{M})$.

Since the objective function factors through the squaring map (6.8), we can also work with the variety $\mathcal{M}^2$ in the projection Grassmannian $\overline{\mathrm{p}}\mathrm{Gr}(2,n)$. In this setting, the prime ideal of the model is generated by polynomials in the $\binom{n+1}{2}$ entries p_{ij} of a symmetric $n \times n$ matrix.

In practice, many models $\mathcal{M}$ are unirational: they are given parametrically. Ideally, this parametrization will be birational. In this case, we consider the critical equations in the ring of polynomials in these parameters. We illustrate this by introducing our running example.

Example 6.2.1 (Chow threefold of the twisted cubic)**.** Let $\mathcal{M} \subset \mathrm{Gr}(2,4)$ be parametrized by

$$A \;=\; \begin{pmatrix} 1 & t_1 & t_1^2 & t_1^3 \\ 0 & 1 & t_2 & t_3 \end{pmatrix} \tag{6.13}$$

The first row defines the twisted cubic curve in $\mathbb{P}^3$, and $\mathcal{M}$ is the threefold of all lines that intersect this curve. The Chow form is the defining equation of $\mathcal{M}$ in Plücker coordinates:

$$\det \begin{pmatrix} x_{12} & x_{13} & x_{23} \\ x_{13} & x_{14}+x_{23} & x_{24} \\ x_{23} & x_{24} & x_{34} \end{pmatrix} \;=\; 0. \tag{6.14}$$

The ideal of $\mathcal{M}^2$ is more complicated. It has 16 minimal generators, starting with the 10 quadrics that define $\overline{\mathrm{p}}\mathrm{Gr}(2,n)$ in $\mathbb{P}(\mathrm{Sym}^2\mathbb{C}^4)$. In addition, we have six more quartics, like

$$\begin{gathered}
2p_{23}^2p_{24}p_{33} - 2p_{22}p_{24}p_{33}^2 - 2p_{23}^3p_{34} + 2p_{14}p_{22}p_{24}p_{34} - 2p_{13}p_{23}p_{24}p_{34} - 2p_{12}p_{24}^2p_{34} + 2p_{22}p_{23}p_{33}p_{34} \\
+2p_{12}p_{24}p_{33}p_{34} - 4p_{13}p_{22}p_{34}^2 + 4p_{12}p_{23}p_{34}^2 - 2p_{23}^2p_{34}^2 + p_{11}p_{24}p_{34}^2 + p_{22}p_{24}p_{34}^2 - p_{11}p_{33}p_{34}^2 \\
+p_{22}p_{33}p_{34}^2 - p_{24}p_{33}p_{34}^2 - p_{33}^2p_{34}^2 + 2p_{23}p_{34}^3 - 2p_{34}^4 - 2p_{14}p_{22}p_{23}p_{44} + 2p_{13}p_{23}^2p_{44} + 2p_{12}p_{23}p_{24}p_{44} \\
+4p_{13}p_{22}p_{33}p_{44} - 6p_{12}p_{23}p_{33}p_{44} + 2p_{23}^2p_{33}p_{44} - p_{11}p_{24}p_{33}p_{44} - p_{22}p_{24}p_{33}p_{44} + p_{11}p_{33}^2p_{44} \\
-p_{22}p_{33}^2p_{44} + p_{24}p_{33}^2p_{44} + p_{33}^3p_{44} - 2p_{23}p_{33}p_{34}p_{44} + p_{24}p_{34}^2p_{44} + 3p_{33}p_{34}^2p_{44} - p_{24}p_{33}p_{44}^2 - p_{33}^2p_{44}^2.
\end{gathered}$$

For our optimization problem, we prefer the cubic (6.14) or the parametrization (6.13). ⋄

While our objective function $||\tilde{P} - Q||^2$ is quadratic in the entries of $\tilde{P}$, the relations $\tilde{P}^2 = \tilde{P}$ and $\mathrm{trace}(\tilde{P}) = 2$ may be used to rewrite this function so that it is linear in $\tilde{P}$:

$$\mathrm{trace}((\tilde{P} - Q)^2) \;=\; \mathrm{trace}(\tilde{P}^2) - 2\,\mathrm{trace}(\tilde{P}Q) + \mathrm{trace}(Q^2) \;=\; 2 + \mathrm{trace}(Q^2) - 2\,\mathrm{trace}(\tilde{P}Q).$$

This allows us to replace (6.4) with an equivalent linear optimization problem over the model:

$$\max_{\tilde{P} \in \mathcal{M}^2 \subseteq \mathrm{pGr}(2,n)} \mathrm{trace}(\tilde{P}Q). \tag{6.15}$$

If $\mathcal{M}^2$ equals $\mathrm{pGr}(2,n)$, then all $\binom{n}{2}$ critical points are real. They are the projections onto the 2-dimensional Q-invariant subspaces of $\mathbb{R}^n$; see Theorem 4.1.2 and [82, Section 4].

Proposition 6.2.2. *Theorem 2.2.2 holds for the optimization problem (6.4). In particular, the ED degree of a subvariety of $\mathcal{M}^2 \subseteq \mathrm{pGr}(2,n)$ of the projection Grassmannian is finite.*

Proof. The map $Q \mapsto \nabla \mathrm{trace}(\tilde{P}Q) = (q_{11}\ 2q_{12}\ \cdots\ q_{nn})$ is surjective. Therefore Assumption 2.2.1 is satisfied, and the theorem holds. □

Written in Plücker coordinates, our objective function in (6.15) is the rational function

$$\phi_Q(X) \;:=\; 2\,\frac{\mathrm{trace}(X^2 Q)}{\mathrm{trace}(X^2)}. \tag{6.16}$$

Let $\mathcal{J}_\mathcal{M}(X)$ denote the Jacobian matrix of the model $\mathcal{M}$. This $\left(n+\binom{n}{4}\right) \times \binom{n}{2}$ matrix has rows indexed by the generators of $I_\mathcal{M}$, and columns indexed by the variables x_{ij}. The entries are the partial derivatives. The rank of $\mathcal{J}_\mathcal{M}(X)$ at a generic point $X \in \mathcal{M}$ is $c = \mathrm{codim}(\mathcal{M})$.

The singular locus of the model $\mathcal{M}$ is the subvariety of $\mathbb{P}(\wedge^2\mathbb{C}^n)$ defined by the ideal $I_{\mathcal{M}_{\mathrm{sing}}}$ as in (2.10). The augmented Jacobian matrix $\mathcal{J}^Q_\mathcal{M}(X)$ in (2.5) is the $\left(n+\binom{n}{4}+2\right) \times \binom{n}{2}$ matrix

$$\mathcal{J}^Q_\mathcal{M}(X) \;=\; \begin{pmatrix} (XQ+QX)_{12} & (XQ+QX)_{13} & \cdots & (XQ+QX)_{n-1,n} \\ x_{12} & x_{13} & \cdots & x_{n-1,n} \\ & & \mathcal{J}_\mathcal{M}(X) & \end{pmatrix} \tag{6.17}$$

Up to scaling, the first row in (6.17) is the gradient of the numerator $\mathrm{trace}(X^2Q)$ in (6.16). The second row is the gradient of the denominator $\mathrm{trace}(X^2)$. We say a model $\mathcal{M} \subseteq \mathrm{Gr}(2,n)$ is *ED-general* if it intersects the hypersurface $\{\mathrm{trace}(X^2) = 0\}$ transversely. Equivalently, this is when the Lagrangian ideal $\Lambda^Q_\mathcal{M}$ as in (2.13) is equal to the critical ideal $I^Q_\mathcal{M}$ (2.12). When this is the case, the number of critical points may be found from Theorem 2.3.2. Models in the Grassmannian are never ED-general, unless they are curves or surfaces:

Theorem 6.2.3. *Fix a subvariety $\mathcal{M} \subseteq \mathrm{Gr}(2,n)$. If $\dim(\mathcal{M}) \geq 3$, then $\mathcal{M}$ is not ED-general. If $\dim(\mathcal{M}) \leq 2$ and $\mathcal{M}$ is defined by general polynomials f_i, then $\mathcal{M}$ is ED-general.*

Proof. The intersection $\mathrm{Gr}(2,n)\cap\{\mathrm{trace}(X^2)=0\}$ is the $\mathrm{O}(n)$-orbit closure of

$$\begin{pmatrix} 0 & 1+i & 0 \\ -1-i & 0 & -1+i \\ 0 & 1-i & 0 \end{pmatrix} \oplus (0) \oplus \cdots \oplus (0). \tag{6.18}$$

It is also the union of the $\mathrm{O}(n)$-orbits of (6.9) and (6.18). It follows from [79, Theorem 2′] that the singular locus of this union is the $\mathrm{O}(n)$-orbit closure of (6.9). Corollary 6.1.11 states that this singular locus equals $\mathrm{Gr}(2,n)\cap\{X^2=0\}$ and has codimension 3 in $\mathrm{Gr}(2,n)$. If $\dim(\mathcal{M})\geq 3$ then $\mathcal{M}$ intersects the base locus and is hence tangent to the isotropic quadric. If $\dim(\mathcal{M})\leq 2$ then this intersection is empty, unless the model $\mathcal{M}$ is very special. □

Example 6.2.4 ($n=4$)**.** The tangent curves in Proposition 6.4.3 fail to be ED-general. ⋄

Theorem 6.2.3 allows us to derive formulas for the ED degrees of generic curves and surfaces in the Grassmannian $\mathrm{Gr}(2,n)$. This will be carried out in Theorem 6.4.1 for the case $n=4$.

We now return to our running example. This is a threefold and hence not ED-general.

Example 6.2.5 (Chow threefold of the twisted cubic)**.** Let f denote the Bézout determinant in (6.14). The augmented Jacobian of the threefold $\mathcal{M}$ it defines in $\mathrm{Gr}(2,4)$ is the $4{\times}6$ matrix

$$\mathcal{J}^Q_{\mathcal{M}}(X) \;=\; \begin{pmatrix} \frac{\partial \mathrm{trace}(X^2Q)}{\partial x_{12}} & \frac{\partial \mathrm{trace}(X^2Q)}{\partial x_{13}} & \frac{\partial \mathrm{trace}(X^2Q)}{\partial x_{14}} & \frac{\partial \mathrm{trace}(X^2Q)}{\partial x_{23}} & \frac{\partial \mathrm{trace}(X^2Q)}{\partial x_{24}} & \frac{\partial \mathrm{trace}(X^2Q)}{\partial x_{34}} \\ x_{12} & x_{13} & x_{14} & x_{23} & x_{24} & x_{34} \\ x_{34} & -x_{24} & x_{23} & x_{14} & -x_{13} & x_{12} \\ \frac{\partial f}{\partial x_{12}} & \frac{\partial f}{\partial x_{13}} & \frac{\partial f}{\partial x_{14}} & \frac{\partial f}{\partial x_{23}} & \frac{\partial f}{\partial x_{24}} & \frac{\partial f}{\partial x_{34}} \end{pmatrix}.$$

This matrix has linear entries in rows 1, 2, and 3, and quadratic entries in row 4. Computation of the ideal $\Lambda^Q_{\mathcal{M}}$, and saturation by $\mathrm{trace}(X^2)$, shows that this model has ED degree 42. ⋄

The computation of critical equations is easier and faster when the model $\mathcal{M}$ is given by a birational parametrization. This usually takes the form of a $2\times n$ matrix $A(t)$ which depends on s parameters $t=(t_1,t_2,\ldots,t_s)$, and t is identifiable from the rowspan of $A(t)$.

The matrix $X=X(t)$ is now a function of t, and hence so is $\phi_Q(X(t))$. The aim is to compute critical points of the rational function $\phi_Q(X(t))$ with no constraints on t; see Section 2.4. The critical ideal is generated by the 2×2 minors of the matrix

$$\begin{pmatrix} \mathrm{trace}(X(t)^2Q) & \frac{\partial \mathrm{trace}(X(t)^2Q)}{\partial t_1} & \frac{\partial \mathrm{trace}(X(t)^2Q)}{\partial t_2} & \cdots & \frac{\partial \mathrm{trace}(X(t)^2Q)}{\partial t_s} \\ \mathrm{trace}(X(t)^2) & \frac{\partial \mathrm{trace}(X(t)^2)}{\partial t_1} & \frac{\partial \mathrm{trace}(X(t)^2)}{\partial t_2} & \cdots & \frac{\partial \mathrm{trace}(X(t)^2)}{\partial t_s} \end{pmatrix}$$

By applying Theorem 2.4.7 to our situation, we obtain the following bound on the ED degree.

Corollary 6.2.6. *Suppose that the model $\mathcal{M}$ is parametrized by the $2 \times n$ matrix*

$$A(t) \;=\; \begin{pmatrix} a_{11}(t) & a_{12}(t) & \cdots & a_{1n}(t) \\ a_{21}(t) & a_{22}(t) & \cdots & a_{2n}(t) \end{pmatrix}$$

and that $t = (t_1, \dots, t_s)$ is identifiable from the rowspan of $A(t)$. If d is the maximal degree among the 2×2 minors of $A(t)$, then the ED degree of $\mathcal{M}$ is at most $(s+1)(2d-1)^s$.

Proof. This is a direct consequence of Theorem 2.4.7. □

Example 6.2.7 ($s \le 2, n = 4$)**.** It is possible for curves and surfaces to attain the bound in Corollary 6.2.6. Consider a parametric model $\mathcal{M}$ for lines in 3-space, given by

$$t \;\mapsto\; \begin{pmatrix} 1 & a_{12}(t) & a_{13}(t) & a_{14}(t) \\ 0 & 1 & a_{23}(t) & a_{24}(t) \end{pmatrix} \tag{6.19}$$

where a_{12}, a_{13}, a_{14} are generic polynomials of degree 3 and a_{23}, a_{24} are generic polynomials of degree 1. The ED degree is $2 \cdot 7 = 14$ for $s = 1$ and $3 \cdot 7^2 = 147$ for $s = 2$. ⋄

Example 6.2.8 ($s = 3, n = 4$)**.** Consider the threefold defined by (6.19) where a_{12}, a_{23}, a_{24} are linear, $\deg(a_{13}) = 2$, and $\deg(a_{14}) = 3$. The rightmost 2×2 minor has the largest degree, namely 4, so the upper bound in Corollary 6.2.6 is $4 \cdot 7^3 = 1372$. If the polynomials $a_{12}, \dots, a_{24}$ are generic, then the ED degree of this model is 337. For the Chow threefold of the twisted cubic, with parametrization (6.13), the ED degree drops to 42. Indeed, the bound in Corollary 6.2.6 cannot be attained for models of dimension greater than 2 by Theorem 6.2.3. ⋄

6.3 The Grassmann Distance Degree

Our primary objective is distance optimization inside the Grassmannian. The given data is a real $2 \times n$ matrix B of rank 2. This represents a line in $\mathbb{P}^{n-1}$. From B, we compute a skew-symmetric matrix Y as in (6.7), and we define $\tilde{Q}$ as the image of Y under the map (6.11):

$$\tilde{Q} \;=\; \frac{2}{\operatorname{trace}(Y^2)} Y^2 \qquad \text{where} \qquad Y \;=\; B^T \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} B. \tag{6.20}$$

Thus, in this section, the matrix $\tilde{Q}$ is special. It is a projection matrix in $\operatorname{pGr}(2, n)$. As in the ED case, the critical points are given by a rank condition on $\mathcal{J}^{\tilde{Q}}_{\mathcal{M}}(X)$ as in (6.17). However, for data $\tilde{Q}$ as in (6.20), there are now extraneous critical points $X \in \operatorname{Gr}(2, n)$, where $\mathcal{J}^{\tilde{Q}}_{\mathcal{M}}(X)$ drops rank regardless of what the model is. These are critical points of the optimization problem

$$\text{Maximize} \quad \frac{\operatorname{trace}(X^2\tilde{Q})}{\operatorname{trace}(X^2)} \quad \text{subject to} \quad X \in \operatorname{Gr}(2, n). \tag{6.21}$$

As in Section 6.2, they are defined by the constraint that the augmented Jacobian matrix $\mathcal{J}^{\tilde{Q}}_{\operatorname{Gr}(2,n)}(X)$ has rank at most $\operatorname{codim}(\operatorname{Gr}(2, n)) + 1$. Here is a geometric characterization that is reminiscent of Theorem 4.1.3:

Proposition 6.3.1. *For general real data $\tilde{Q}$ in $\mathrm{pGr}(2,n)$, the critical points of (6.21) are precisely the $\tilde{Q}$-invariant subspaces of dimension 2 in $\mathbb{R}^n$. The dimension of the variety*

$$V\left(I^{\tilde{Q}}_{\mathrm{Gr}(2,n)}\right) \setminus \{Y\} \tag{6.22}$$

is $n-2$ for $n = 4, 5$, and $2n-8$ for $n \geq 6$. It is contained in the hypersurface $V(\mathrm{trace}(XY))$.

Proof. Let $\tilde{P} \in \mathrm{pGr}(2,n)$ be as in (6.2). Because $\mathrm{O}(n)$ acts transitively on $\mathrm{pGr}(2,n)$, we may assume that $\tilde{P} = \begin{pmatrix} \mathrm{Id}_2 & 0 \\ 0 & 0 \end{pmatrix}$. The tangent vectors to $\mathrm{pGr}(2,n)$ at $\tilde{P}$ have the block structure $\dot{\tilde{P}} = \begin{pmatrix} 0 & A^T \\ A & 0 \end{pmatrix}$ for $(n-2) \times 2$ matrices A. A subspace X is critical for (6.21) if and only if $\mathrm{trace}(\dot{\tilde{P}}\tilde{Q}) = 0$ for every $\dot{\tilde{P}}$. Decomposing $\tilde{Q} = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{12}^T & Q_{22} \end{pmatrix}$, this is equivalent to $Q_{12} = 0$. Hence the matrices $\tilde{P}$ and $\tilde{Q}$ commute. This implies the first statement.

Let $Y = \mathrm{im}(\tilde{Q})$ and $Y^\perp = \ker(\tilde{Q})$. By the first part of the proof, a critical subspace X is spanned by vectors in Y and $Y^\perp$. The subspaces $X \subset Y^\perp$ form a Schubert variety of codimension 4 in $\mathrm{Gr}(2,n)$. The next case, when $\dim(X \cap Y) = 1$ and $\dim(X \cap Y^\perp) = 1$, is a transverse intersection of two Schubert varieties of codimensions $n-3$ and 1 respectively. This variety has codimension $n-2$ in $\mathrm{Gr}(2,n)$. The remaining case is $X = Y$. The dimension claim follows. Finally, $\mathrm{trace}(XY)$ vanishes if and only if X intersects $Y^\perp$ non-trivially. □

The Zariski closure of the set in (6.22) is called the *extraneous critical locus*. We refer to its points as the *extraneous critical points*. This locus is independent of the choice of model $\mathcal{M}$. It consists of all critical points on the Grassmannian other than the data point.

Now, let us fix a model $\mathcal{M}$ and data Y, both in the Grassmannian $\mathrm{Gr}(2,n)$. The intersection between $\mathcal{M}$ and the extraneous critical locus is contained in the variety defined by the ED critical ideal $I^{\tilde{Q}}_{\mathcal{M}}$. That intersection can have a positive-dimensional component. Any such component would be contained in the *algebraic cut locus* of the point Y, which is

$$\mathcal{D}_Y \;=\; \{ X \in \mathrm{Gr}(2,n) : \mathrm{trace}(XY) = 0 \} .$$

The term *cut locus* comes from Riemannian geometry. It refers to the set of points in a manifold that are connected to a given point Y by two or more length-minimizing geodesics; see [86, Section 3]. The cut locus for the geodesic distance on $\mathrm{Gr}(2,n)$ is precisely the set of real points in $\mathcal{D}_Y$. Indeed, for the geodesic distance problem in [86], a generic data point in $\mathrm{Gr}(2,n)$ can have infinitely many critical points in $\mathcal{M}$ that lie in the cut locus. However, for smooth models, a (local) minimizer cannot be among them. This is [86, Theorem 36].

A similar phenomenon occurs for the chordal distance (6.5). Proposition 6.3.1 shows that the hypersurface $\mathcal{D}_Y$ can contain infinitely many critical points for (6.4), even for generic data $Y \in \mathrm{Gr}(2,n)$. It is remarkable that, even though the metrics used in [86] and in this paper enjoy very different properties, the same locus plays a crucial role in both optimization settings. For this reason we call the hypersurface $\mathcal{D}_Y \subset \mathrm{Gr}(2,\mathrm{n})$ the algebraic cut locus of Y.

Definition 6.3.2. We define the *GD critical ideal* to be the saturation of the Lagrangian ideal $\Lambda_{\mathcal{M}}^{Q}$ by $\langle \mathrm{trace}(XY)\rangle^{\infty}$. If the model $\mathcal{M}$ is not ED-general, then we also saturate by $\langle \mathrm{trace}(X^2)\rangle^{\infty}$; this will be always implicit in the following. By Proposition 6.3.1, saturating by $\langle \mathrm{trace}(XY)\rangle^{\infty}$ removes all extraneous critical points from the variety defined by the ED critical ideal. We define the *GD degree* of the model $\mathcal{M}$, denoted $\mathrm{GDdegree}(\mathcal{M})$, to be the number of complex points in the variety defined by the GD critical ideal for generic data points $Y \in \mathrm{Gr}(2, \mathrm{n})$.

Example 6.3.3 (Chow threefold of the twisted cubic)**.** The model in (6.14) has ED degree 42. By taking data $\tilde{Q}$ in the Grassmannian as in (6.20), we find that the GD degree is 10. ⋄

Remark 6.3.4. We believe that the GD degree of every model $\mathcal{M}$ in $\mathrm{Gr}(2, n)$ is finite. If this holds, then $\mathrm{GDdegree}(\mathcal{M}) \leq \mathrm{EDdegree}(\mathcal{M})$ follows. At present, we do not have a proof.

A model $\mathcal{M}$ in $\mathrm{Gr}(2, n)$ is *GD-general* if it satisfies the following for generic $Y \in \mathrm{Gr}(2, n)$:

- the model $\mathcal{M}$ intersects the extraneous critical locus transversely, and
- the extraneous critical locus contains the intersection of the algebraic cut locus $\mathcal{D}_Y$ with the variety defined by the ED critical ideal.

The first condition implies that (6.22) has the expected codimension in $\mathcal{M}$. The second condition ensures that saturating by $\mathrm{trace}(XY)$ only removes extraneous critical points and not genuine ones. The model $\mathcal{S}_{25}$ from Example 6.0.1 does not satisfy the second condition: it has a critical point in $\mathcal{D}_Y$ that is not extraneous in the sense of Proposition 6.3.1.

If the model $\mathcal{M}$ is GD-general and small enough, then it will not intersect the extraneous critical locus. In this case, the GD degree of the model agrees with the ED degree.

Example 6.3.5 ($n = 4$)**.** For GD-general curves, the ED degree equals the GD degree. For GD-general surfaces, the ED critical variety has isolated points that are extraneous. Their number is the difference between ED and GD degrees. For GD-general threefolds, the ED critical variety has extraneous curves. All of these extraneous points are critical for (6.21). ⋄

Example 6.3.6 ($n = 5$)**.** For GD-general curves and surfaces in $\mathrm{Gr}(2, 5)$, the ED degree equals the GD degree. For GD-general threefolds, the ED critical variety has isolated points that are extraneous. For GD-general fourfolds, the ED critical variety contains extraneous curves. For GD-general fivefolds, the ED critical variety contains extraneous surfaces. ⋄

The phenomena seen in the previous two examples have the following general description.

Theorem 6.3.7. *Fix $n \geq 6$ and assume that the polynomials $f_1, \ldots, f_r$ in (6.12) are generic. If $r > 2n - 8$ then* $\mathrm{EDdegree}(\mathcal{M}) = \mathrm{GDdegree}(\mathcal{M})$. *If $r \leq 2n - 8$, then* $\mathrm{EDdegree}(\mathcal{M}) > \mathrm{GDdegree}(\mathcal{M})$ *and the ED critical variety has extraneous components of dimension* $2n-8-r$.

Proof. By the genericity of its defining polynomials $f_1, \ldots, f_r$, the model $\mathcal{M}$ is GD-general. The conclusion then follows from the dimension statement in Proposition 6.3.1. □

An immediate consequence is that $\mathrm{GDdegree}(\mathcal{M}) < \mathrm{EDdegree}(\mathcal{M})$ if $\dim(\mathcal{M}) \geq 4$. A discrepancy between the ED and GD degrees means that the projection Grassmannian $\overline{\mathrm{p}}\mathrm{Gr}(2,n)$ is contained in the *ED discriminant* $\Sigma_{\mathcal{M}}$ of the given model $\mathcal{M}$. Recall from [37, Section 7] that $\Sigma_{\mathcal{M}}$ is the hypersurface in $\mathbb{P}(\mathrm{Sym}^2\mathbb{C}^n)$ for which this drop occurs, c.f. Section 5.2. In fact, based on computations, the following stronger statement seems to hold for many models $\mathcal{M}$:

$$\mathcal{M} \subset \overline{\mathrm{p}}\mathrm{Gr}(2,n) \subset \mathcal{V}_{X^2} \subset \Sigma_{\mathcal{M}} \subset \mathbb{P}(\mathrm{Sym}^2\mathbb{C}^n). \tag{6.23}$$

We use the term *SD degree* for the number of critical points when Q is generic in $\mathcal{V}_{X^2}$. The inclusion $\mathcal{V}_{X^2} \subset \Sigma_{\mathcal{M}}$ in (6.23) means that the SD degree is less than the ED degree for $\mathcal{M}$.

In our running example, $\mathcal{M}$ has SD degree 20. This is strictly between $\mathrm{GDdegree}(\mathcal{M}) = 10$ and $\mathrm{EDdegree}(\mathcal{M}) = 42$. In fact, we observed that $\mathrm{SDdegree}(\mathcal{M}) = 2 \cdot \mathrm{GDdegree}(\mathcal{M})$ whenever $\mathcal{M}$ is the Chow threefold of a toric curve in $\mathbb{P}^3$; see Section 6.4. For the Schubert varieties $\mathcal{S}_{ij}$ in Section 6.5, we conjecture that $\mathcal{V}_{X^2} \subset \Sigma_{\mathcal{S}_{ij}}$ if and only if $i = 1$ or $(i,j) = (2,3)$. These families of models will be introduced in slow motion in the remaining two sections.

6.4 Lines in 3-Space

This section is devoted to the GD problem for lines in 3-space. The model $\mathcal{M}$ is a subvariety of $\mathrm{Gr}(2,4)$. If $\dim(\mathcal{M}) = 1$ then the union of its lines is a *ruled surface* in $\mathbb{P}^3$. If $\dim(\mathcal{M}) = 2$ then it represents a *line congruence*, i.e. a surface of lines whose union covers $\mathbb{P}^3$. These are important in computer vision [106]. The term *line complex* is used classically for a hypersurface in $\mathrm{Gr}(2,4)$. Such a threefold $\mathcal{M}$ is defined by one polynomial in the Plücker coordinates $x_{12}, x_{13}, x_{14}, x_{23}, x_{24}, x_{34}$. We begin with the case of generic complete intersections in $\mathrm{Gr}(2,4)$.

Theorem 6.4.1. *A surface in* $\mathrm{Gr}(2,4)$ *defined by generic polynomials* f_1 *and* f_2 *of degrees* d_1 *and* d_2 *has ED degree* $2d_1d_2(d_1^2+d_2^2+d_1d_2+3)$. *A curve in* $\mathrm{Gr}(2,4)$ *defined by three generic polynomials* f_1, f_2, f_3 *of degrees* d_1, d_2, d_3 *has ED degree and GD degree* $2d_1d_2d_3(d_1+d_2+d_3)$.

Proof. These surfaces and curves are ED-general by Theorem 6.2.3, so the number of critical points is given by Theorem 2.3.2(i) substituting $e = 2$, $n = 5$, and $d_3 = 2$ (resp. $d_4 = 2$) for the Plücker quadric:

$$\deg(I_{\mathcal{M}}^Q) \;=\; (2d_1 \cdots d_c) \cdot \textstyle\sum_{i_0+\cdots+i_c=4-c} \binom{i_0+2}{2} (d_1-1)^{i_1} \cdots (d_c-1)^{i_c}.$$

One obtains the formulas by plugging in $r = 2$ for surfaces and $r = 3$ for curves. By Example 6.3.5, the ED and GD degrees are equal for a general curve in $\mathrm{Gr}(2,4)$. ☐

Conjecture 6.4.2. The GD degree of a surface defined by generic polynomials of degrees d_1, d_2 is $2d_1d_2(d_1^2+d_2^2+d_1d_2+2)$. This is $2d_1d_2$ less than the ED degree. A threefold defined by a generic polynomial f of degree d has ED degree $2d(d^3+3d+2)$ and GD degree $2d(d^3+2d)$.

The following tables show these data for surfaces and threefolds of low degree in $\mathrm{Gr}(2,4)$:

$d_1 \backslash d_2$	1	2	3	4
1	12	40	96	192
2		120	264	496
3			540	960
4				1632

ED degree of surfaces

$d_1 \backslash d_2$	1	2	3	4
1	10	36	90	184
2		112	252	480
3			522	936
4				1600

GD degree of surfaces

d	ED degree	GD degree
1	12	6
2	64	48
3	228	198
4	624	576
5	1420	1350

Threefolds

Certain ruled surfaces, line congruences and line complexes arise from a given curve $\mathcal{C}$ of degree d in $\mathbb{P}^3$. The *Chow threefold* of $\mathcal{C}$ is the set of lines that intersect $\mathcal{C}$. Its defining Plücker polynomial has degree d. If $d = 1$ then the Chow threefold is the Schubert variety $\mathcal{S}_{13}$ in Example 6.5.2. For any given line L, our GD problem asks for the line closest to L among those that intersect $\mathcal{C}$. The *secant congruence* of $\mathcal{C}$ is the set of lines that intersect $\mathcal{C}$ at two points. The best known ruled surfaces are quadrics in $\mathbb{P}^3$. These have two rulings of lines. Our GD problem asks for the line in a ruling that is closest to the data line L. The set of lines tangent to $\mathcal{C}$ is the *tangent curve* in $\mathrm{Gr}(2,4)$. Its lines form the tangential surface of $\mathcal{C}$.

We present an experimental study for toric curves $\mathcal{C} = \{\, (1 : t^{u_1} : t^{u_2} : t^{u_3}) \in \mathbb{P}^3 \,:\, t \in \mathbb{C}\,\}$. The exponents are integers which satisfy $0 < u_1 < u_2 < u_3$ and $\gcd(u_1,u_2,u_3) = 1$. We consider the three parametrically presented models $\mathcal{M} \subset \mathrm{Gr}(2,4)$ that are derived from $\mathcal{C}$:

- The Chow threefold: $A(t,x,y) = \begin{pmatrix} 1 & t^{u_1} & t^{u_2} & t^{u_3} \\ 0 & 1 & x & y \end{pmatrix}.$
- The secant surface: $A(t,x) = \begin{pmatrix} 1 & t^{u_1} & t^{u_2} & t^{u_3} \\ 1 & x^{u_1} & x^{u_2} & x^{u_3} \end{pmatrix}.$
- The tangent curve: $A(t) = \begin{pmatrix} 1 & t^{u_1} & t^{u_2} & t^{u_3} \\ 0 & u_1 t^{u_1-1} & u_2 t^{u_2-1} & u_3 t^{u_3-1} \end{pmatrix}.$

These models have dimensions $3, 2, 1$. For the secant surface, the parametrization is two-to-one. For the first and last, it is birational. For the curve, we found the following result:

Proposition 6.4.3. *The tangent curve of $\mathcal{C}$ has the same ED degree and GD degree. This degree equals $4u_3 + 2u_2 - 2u_1$, unless $u_3 = u_1 + u_2$ and u_3 is even, in which case we subtract 2.*

Proof. The numerator of the derivative of the objective function, denoted $f(t)$, has degree $4u_3 + 3u_2 + u_1 - 5$. The zeros of f parametrize the critical points. This tangent curve is not ED general. Indeed, f has a factor of $t^{3u_1+u_2-5}$. The remaining $4u_3 + 2u_3 - 2u_1$ points are critical. If $u_3 = u_1 + u_2$ and u_3 is even, then $\mathrm{trace}(X(t)^2)$ has a factor of $(t^2+1)^2$; therefore f has a factor of $t^2 + 1$. These points are extraneous because $t^2 + 1$ divides $\mathrm{trace}(X(t)^2)$. $\square$

The ED, SD, and GD degrees for some toric curves are shown in the following table:

	Chow threefold			Secant surface			Tangent curve		
(u_1, u_2, u_3)	ED	SD	GD	ED	SD	GD	ED	SD	GD
$(1, 2, 3)$	42	20	10	19	15	15	14	14	14
$(1, 2, 4)$	55	26	13	62	54	54	18	18	18
$(1, 2, 5)$	68	32	16	111	98	98	22	22	22
$(1, 2, 6)$	81	38	19	170	151	151	26	26	26
$(1, 2, 7)$	94	44	22	235	209	209	30	30	30
$(1, 3, 4)$	54	24	12	59	50	50	18	18	18
$(1, 3, 5)$	71	34	17	122	108	108	24	24	24
$(1, 3, 6)$	84	40	20	179	158	158	28	28	28
$(1, 3, 7)$	97	46	23	262	234	234	32	32	32
$(1, 4, 5)$	74	36	18	113	97	97	26	26	26
$(1, 4, 6)$	87	42	21	212	189	189	30	30	30

We also tested whether the optimal line intersects the data line. For the examples above, this happens for the Chow threefolds, but not for the secant surfaces or tangent curves. Based on these computations, we venture the following conjectures for every curve $\mathcal{C} \subset \mathbb{P}^3$:

Conjecture 6.4.4. The optimal line in the Chow threefold always intersects the data line. This does not hold for the secant surface and the tangent curve.

Conjecture 6.4.5. We have SDdegree $= 2 \cdot$ GDdegree for the Chow threefold. For the secant surface, SDdegree = GDdegree.

6.5 Schubert Varieties

The most prominent subvarieties of a Grassmannian are its Schubert varieties. Their metric algebraic geometry was studied in [82, 86]. For $1 \leq i < j \leq n$, write $\mathcal{L}_{ij}$ for the linear space of skew-symmetric matrices $X = (x_{rs})$ such that $x_{rs} = 0$ unless $i \leq r$ and $j \leq s$. The *Schubert variety* $\mathcal{S}_{ij}$ is the intersection $\mathrm{Gr}(2, n) \cap \mathcal{L}_{ij}$. We use this term for any variety obtained from $\mathcal{S}_{ij}$ by an orthogonal transformation. Note that $\dim(\mathcal{S}_{ij}) = 2n - i - j - 1$. The two extreme cases are $\mathcal{S}_{12} = \mathrm{Gr}(2, n)$ and $\mathcal{S}_{n-1,n} =$ one point. See also Example 6.0.1. We write $\mathcal{S}^2_{ij} \subset \overline{\mathrm{p}}\mathrm{Gr}(2, n) \subset \mathbb{P}(\mathrm{Sym}^2\mathbb{C}^n)$ for the image of the Schubert variety $\mathcal{S}_{ij}$ under (6.8).

The implicit representation of the Schubert variety in Plücker coordinates is the ideal

$$\text{ideal of } \mathcal{S}_{ij} \;=\; \text{ideal of } \mathrm{Gr}(2, n) \;+\; \quad x_{rs} : r < i \text{ or } s < j \;\; .$$

We conjecture that the ideal of $\mathcal{S}^2_{ij}$ is generated by quadrics and cubics; see Conjecture 6.5.1. We write P_{rs} for the $2 \times n$ submatrix of P given by the rows r and s.

Conjecture 6.5.1. The prime ideal of $\mathcal{S}^2_{ij}$ is generated by the entries of $2P^2 - \mathrm{trace}(P) \cdot P$, the 3×3 minors of P, and the 2×2 minors of the submatrices P_{rs} where $r < i$ or $s < j$.

This conjecture was verified for $n \leq 8$. We next present a census for $n = 4, 5$.

Example 6.5.2 ($n = 4$)**.** There are six Schubert varieties in $\mathrm{Gr}(2,4)$, summarized as follows:

	$\mathcal{S}_{12}$	$\mathcal{S}_{13}$	$\mathcal{S}_{23}$	$\mathcal{S}_{14}$	$\mathcal{S}_{24}$	$\mathcal{S}_{34}$
dim	4	3	2	2	1	0
mingens	$(0,10)$	$(0,16)$	$(4,6)$	$(4,6)$	$(7,1)$	$(9,0)$
degree	12	12	4	4	2	1
ED degree	6	10	3	3	2	1
GD degree	1	2	1	1	2	1

The rows labeled "degree" and "mingens" refer to the variety $\mathcal{S}^2_{ij}$ in $\mathbb{P}(\mathrm{Sym}^2\mathbb{C}^4) = \mathbb{P}^9$. The ideal generators have degrees 1 and 2. We list their numbers. For instance, the ideal of the surface $\mathcal{S}^2_{23}$ is generated by $p_{11}, p_{12}, p_{13}, p_{14}$ and the entries of $2{P_{-1}}^2 - \mathrm{trace}(P_{-1})P_{-1}$, where P_{-i} is the 3×3 matrix obtained from P by deleting row i and column i. The ideal of the surface $\mathcal{S}^2_{14}$ is generated by p_{14}, p_{24}, p_{34} and $p_{11}+p_{22}+p_{33}-p_{44}$ plus the 2×2 minors of P_{-4}. ◇

Example 6.5.3 ($n = 5$)**.** There are ten Schubert varieties in $\mathrm{Gr}(2,5)$, summarized as follows:

	$\mathcal{S}_{12}$	$\mathcal{S}_{13}$	$\mathcal{S}_{23}$	$\mathcal{S}_{14}$	$\mathcal{S}_{24}$	$\mathcal{S}_{15}$	$\mathcal{S}_{34}$	$\mathcal{S}_{25}$	$\mathcal{S}_{35}$	$\mathcal{S}_{45}$
dim	6	5	4	4	3	3	2	2	1	0
mingens	$(0,15)$	$(0,25)$	$(5,10)$	$(0,40)$	$(5,16)$	$(5,20)$	$(9,6)$	$(9,6)$	$(12,1)$	$(14,0)$
degree	40	40	12	28	12	8	4	4	2	1
ED degree	10	24	6	16	10	4	3	3	2	1
GD degree	1	2	1	2	2	1	1	2	2	1

◇

The degree of $\mathcal{S}^2_{ij}$ depends only on the differences $n-j$ and $n-i$. Here are some values:

$n-j$ \ $n-i$	1	2	3	4	5	6	7	8	9
0	1	2	4	8	16	32	64	128	256
1		4	12	28	60	124	252	508	1020
2			12	40	100	224	476	984	2004
3				40	140	364	840	1824	3828
4					140	504	1344	3168	6996
5						504	1848	5016	12 012
6							1848	6864	18 876
7								6864	25 740

When $n = j$, the variety $\mathcal{S}^2_{in}$ is the Veronese square $\nu_2(\mathbb{P}^{n-i-1})$. This explains the powers of two in the $n - j = 0$ row. Rows 1-7 match the OEIS sequence `A201385` up to a factor of 4.

Conjecture 6.5.4. For $j < n$, the varieties $\mathcal{S}^2_{ij}$ satisfy $\deg(\mathcal{S}^2_{i,n-1}) = 4(2^{n-i-1} - 1)$ and

$$\deg(\mathcal{S}^2_{ij}) \;=\; \deg(\mathcal{S}^2_{i+1,j}) + \deg(\mathcal{S}^2_{i,j+1}) \qquad \text{where } \deg(\mathcal{S}^2_{jj}) = 0.$$

The solution to this recursion is $\deg(\mathcal{S}^2_{ij}) = \sum_{k=n-j+1}^{n-i-1} \binom{2n-i-j}{k} + 2\binom{2n-i-j-1}{n-j}$.

The ED degree was found in [82, Corollary 4.5] for a special class of Schubert varieties:

$$\mathrm{EDdegree}(\mathcal{S}_{i,i+1}) \;=\; \binom{n-i+1}{2} \quad \text{and} \quad \mathrm{EDdegree}(\mathcal{S}_{in}) \;=\; n-i.$$

For the general case we have the following conjecture:

Conjecture 6.5.5. If $i < j-1$, then the ED degree of the Schubert variety $\mathcal{S}_{ij}$ equals

$$\mathrm{EDdegree}(\mathcal{S}_{ij}) \;=\; 2(n-j+1)^2(j-i) - (n-j+1)(n-i).$$

We verified Conjecture 6.5.5 for $n \leq 12$ with Gröbner basis computations over finite fields in `Maple`. For $n = 13, 14$ we verified the conjecture in `HomotopyContinuation.jl` [20].

There is a dramatic drop from the ED degree to the GD degree for Schubert varieties.

Theorem 6.5.6. *The n Schubert varieties $\mathcal{S}_{12}, \mathcal{S}_{23}, \ldots, \mathcal{S}_{n-1,n}, \mathcal{S}_{1n}$ have GD degree 1. The minimizer is obtained from the data point by zeroing out appropriate Plücker coordinates.*

The proof of Theorem 6.5.6 is given at the end of this section. For all other Schubert varieties we put forth the following remarkable conjecture. This was verified for $n \leq 12$.

Conjecture 6.5.7. If $1 < j - i < n-1$, then the Schubert variety $\mathcal{S}_{ij}$ has GD degree 2.

Example 6.5.8 (Schubert curves)**.** Consider the curve $\mathcal{S}_{n-2,n}$ and write the data as

$$B \;=\; \begin{pmatrix} b_1 & b_2 & \cdots & b_{n-3} & 1 & \beta & 0 \\ c_1 & c_2 & \cdots & c_{n-3} & 0 & \gamma & 1 \end{pmatrix} \text{and set } \; b^2 := \sum_{i=1}^{n-3} b_i^2\,,\; bc := \sum_{i=1}^{n-3} b_i c_i\,,\; c^2 := \sum_{i=1}^{n-3} c_i^2.$$

This model has GD degree 2. A computation shows that the optimal solutions are

$$A(t) \;=\; \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 & t & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & 1 \end{pmatrix}$$

where $\big((-c^2-1)\beta + bc\gamma\big)\cdot t^2 + \big((c^2+1)\beta^2 - 2bc\beta\gamma + b^2\gamma^2 - c^2 - 1\big)\cdot t + (c^2+1)\beta - bc\gamma \;=\; 0.$ $\diamond$

For fixed data $\tilde{Q} \in \mathrm{pGr}(2,n)$, define a map $\sigma_{\tilde{Q}} : \mathcal{M}_{\mathbb{R}} \to \mathbb{R}^2$ from the real locus of the model that sends X to the pair (λ, μ), where $\lambda \geq \mu$ are the nonzero eigenvalues of $\tilde{P}\tilde{Q}$ where $\tilde{P}$ is as in (6.2). We call the image of $\sigma_{\tilde{Q}}$ the *spectral region* of the pair $(\mathcal{M}, \tilde{Q})$; see Figure 6.1.

Remark 6.5.9. To compute the spectral region, we consider the variety in $\mathcal{M} \times \mathbb{C}^2$ given by the ideal $I_{\mathcal{M}} + \det(\lambda\,\mathrm{Id}_n - \tilde{P}\tilde{Q}), \det(\mu\,\mathrm{Id}_n - \tilde{P}\tilde{Q})$. Consider the map from that variety onto the second factor $\mathbb{C}^2$. The branch locus of this map can be computed using elimination. It is the algebraic boundary of the spectral region. This is how the quartic in (6.6) was computed.

Our problem (6.4) asks for the maximum coordinate sum $\lambda+\mu$ over the spectral region. There are many ways to assign reasonable metrics to homogeneous spaces; see, e.g., [7]. For other metrics on $\mathrm{Gr}(2,n)$, like that in [86], other monotone functions in λ,μ are maximized. Multi-objective distance optimization on $\mathcal{M}\subset\mathrm{Gr}(2,n)$ means identifying the Pareto frontier of the spectral region. That frontier is a subset of the algebraic boundary in Remark 6.5.9. We say that the model $\mathcal{M}$ is *snug* if the Pareto frontier is a single point for all data $\tilde{Q}$.

We now give bounds on the spectral region and characterize the snug Schubert varieties. Fix $E_j=\mathrm{diag}(0,0,\ldots,0,1,1,\ldots,1)$, where $j-1$ entries are zero. For an $n\times n$ matrix M with real eigenvalues, we write the eigenvalues in decreasing order, i.e., $\lambda_1(M)\geq\cdots\geq\lambda_n(M)$.

Theorem 6.5.10. *If $\tilde{P}\in\mathcal{S}_{ij}$, then $\lambda\leq\lambda_1(\tilde{Q}E_i)$ and $\mu\leq\min(\lambda_2(\tilde{Q}E_i),\lambda_1(\tilde{Q}E_j))$. Both bounds are attained for all $\tilde{Q}\in\mathrm{pGr}(2,n)$ if and only if $\mathcal{S}_{ij}$ is snug if and only if $i=1$ or $j-i=1$.*

The last condition shows that among the $\binom{n}{2}$ Schubert varieties, only $2n-3$ are snug. To prove Theorem 6.5.10 we need the lemma below, which follows from [124, Theorem 1].

Lemma 6.5.11. *Let $P_1\in\mathrm{pGr}(k_1,n)$, $P_2\in\mathrm{pGr}(k_2,n)$, and let $S\in\mathrm{pGr}(k,n)$ be the orthogonal projection onto the intersection of the linear spaces represented by P_1 and P_2. Let $Q\in\mathrm{Sym}^2\mathbb{R}^n$ be a projection matrix, i.e., $Q^2=Q$. Then for $i=1,\ldots,n-(k_2-k)$ we have*

$$\lambda_i(QP_1)\;\geq\;\lambda_i(QS)\;\geq\;\lambda_{i+k_2-k}(QP_2).$$

Proof of Theorem 6.5.10. Taking $P_1=E_i$, $P_2=P$ and $Q=\tilde{Q}$ in Lemma 6.5.11 yields the inequalities $\lambda\leq\lambda_1(\tilde{Q}E_i)$ and $\mu\leq\lambda_2(\tilde{Q}E_i)$. Taking instead $P_1=E_j$ in Lemma 6.5.11 yields the inequality $\mu\leq\lambda_1(\tilde{Q}E_j)$. This proves the first assertion.

We turn to the second assertion. First let $i=1, j<n$, so $\mathcal{S}_{1j}=\{\tilde{P}:\dim(\tilde{P}\cap E_j)=1\}$. Let $\{q_1,q_2\}$ and $\{f_1,f_2,\ldots\}$ be orthonormal bases for the subspaces $\tilde{Q}$ and E_j such that $\langle q_1,f_1\rangle=\sqrt{\lambda_1(\tilde{Q}E_j)}$ and $\langle q_2,f_2\rangle=\sqrt{\lambda_2(\tilde{Q}E_j)}$. Let $\tilde{P}$ denote the orthogonal projection onto the span of f_1 and q_2, so $\lambda=\lambda_1(\tilde{Q}\tilde{P})$ and $\mu=\lambda_2(\tilde{Q}\tilde{P})$. This implies $\lambda=1=\langle q_2,q_2\rangle^2$ and $\mu=\lambda_1(\tilde{Q}E_j)=\langle q_1,f_1\rangle^2$, so the upper bounds are attained.

Next consider the Schubert variety $\mathcal{S}_{i,i+1}=\{\tilde{P}:\tilde{P}\subseteq E_i\}$. Let $\{q_1,q_2\}$ and $\{f_1,f_2,\ldots\}$ be orthonormal bases for $\tilde{Q}$ and E_i respectively such that $\langle q_1,f_1\rangle=\sqrt{\lambda_1(\tilde{Q}E_i)}$ and $\langle q_2,f_2\rangle=\sqrt{\lambda_2(\tilde{Q}E_i)}$. The desired minimizer $\tilde{P}$ is the orthogonal projection onto $\langle f_1,f_2\rangle$.

It remains to be shown that none of the other Schubert varieties are snug. Consider

$$\mathcal{S}_{ij}\;=\;\big\{\tilde{P}\;:\;\dim(\tilde{P}\cap E_j)=1\text{ and }\tilde{P}\subseteq E_i\big\}\;=\;\mathcal{S}_{1j}\cap\mathcal{S}_{i,i+1}.$$

There exists a matrix $\tilde{Q}$ such that $\lambda_2(\tilde{Q}E_i)<\lambda_1(\tilde{Q}E_j)$. The inequalities $\lambda\leq\lambda_1(\tilde{Q}E_i)$ and $\mu\leq\lambda_2(\tilde{Q}E_i)$ are attained by the minimizer of (6.4) over $\mathcal{S}_{i,i+1}$, which strictly contains $\mathcal{S}_{ij}$. Generically, the minimizer intersects E_j trivially, so it is not contained in $\mathcal{S}_{ij}$. By the uniqueness of the principal vectors, the bounds cannot be simultaneously achieved. $\square$

Proof of Theorem 6.5.6. We prove the claim for $\mathcal{S}_{i,i+1}$. The argument is similar for $\mathcal{S}_{1n}$. Fix

$$B \;=\; \begin{pmatrix} b_{11} & \cdots & b_{1,i-1} & 1 & 0 & b_{1,i+2} & \cdots & b_{1n} \\ b_{21} & \cdots & b_{2,i-1} & 0 & 1 & b_{2,i+2} & \cdots & b_{2n} \end{pmatrix}\Bigg($$

By the proof of Theorem 6.5.10, the solution of (6.4) for $\mathcal{M} = \mathcal{S}_{i,i+1}$ is the orthogonal projection of the row span of B onto E_i. This is found by setting $b_{1j} = b_{2j} = 0$ for $j = 1, \ldots, i-1$. Hence the minimizer of (6.4) is a rational function of the data, and so the GD degree is one. □

In summary, we have introduced the Grassmann distance degree of a subvariety of the Grassmannian. We have described the conditions under which the ED degree of a model agrees with its GD degree. We have studied the GD degrees of generic complete intersections and geometrically meaningful subvarieties in $\mathrm{Gr}(2,4)$ and classified the Schubert varieties of GD degree one.

Over the last three chapters, we have studied polynomial optimization problems on Grassmannians and related varieties. In Chapter 4, we saw that the critical points sets of optimization problems from linear algebra and statistics can be described in terms of the spectra of the data matrices. In Chapter 5, we studied a constrained Rayleigh quotient optimization problem. We described the discriminant of this optimization problem for small varieties; this is a promising direction for further study. Finally, in Chapter 6, we considered the problem of minimizing a linear function on a subvariety of the projection Grassmannian, with data in the Grassmannian. There are many interesting directions for future work, including studying ED discriminants of subvarieties of the Grassmannians and extending our results to other Grassmannians. One more direction is the development a general theory for distance optimization on a subvariety of a variety with data in that same variety.

Part III

Likelihood Geometry of Determinantal Point Processes

The determinantal point process (DPP) for discrete random variables is a statistical model whose states are the subsets of a finite set $[n] = \{1, 2, \ldots, n\}$. DPPs elegantly model negative correlation, meaning that they naturally select for diverse subsets of the ground set. This model is ubiquitous in probability theory [12, 67, 92], statistical physics [93], algebraic combinatorics [13], and machine learning [80, 81].

A DPP is a random variable Z on the power set $2^{[n]}$ of the finite set such that $\mathrm{Prob}(I \subseteq Z) = \det(K_I)$, where K is an $n \times n$ symmetric matrix with eigenvalues in $[0, 1]$ and K_I is the principal submatrix of K whose rows and columns are indexed by I. The matrix K is the *kernel* of the DPP and encodes the similarity of the elements of the ground set $[n]$.

We study maximum likelihood estimation for determinantal point processes. Given some data $(u_I)_{I\in 2^{[n]}}$, we seek the maximizer of the log-likelihood function

$$L_u(q) = \sum_{I\subseteq[n]} u_I \log(q_I) - \Big(\sum_{I\subseteq[n]} u_I\Big) \log\Big(\sum_{I\subseteq[n]} q_I\Big) \tag{7.1}$$

where the vector $q = (q_I)_{I\in 2^{[n]}}$ is an unknown probability distribution coming from a DPP, up to scaling. The maximizer of L_u, called the *maximum likelihood estimate* (MLE), is the point on the DPP model that best explains the data u. The number of complex critical points of (7.1) for generic data u is the ML degree of the model; see Section 2.1.

To count the critical points of (7.1), we need a precise description of the constraints on $(q_I)_{I\in 2^{[n]}}$. In particular, we need the probability $q_I = \mathrm{Prob}(I = Z)$ of observing a set I, rather than $\mathrm{Prob}(I \subseteq Z)$. We discuss two scenarios in which there are nice formulae for q_I.

In order for K to define a probability distribution, its eigenvalues must lie in the interval $[0, 1]$. If the eigenvalues of K lie in the interior $(0, 1)$, then the corresponding DPP is called an *L-ensemble* and it is parametrized by a positive-definite symmetric $n\times n$ matrix $\Theta = (\theta_{ij})$. This is the setting of Chapter 7. The matrices K and Θ are related by

$$\Theta = K(\mathrm{Id}_n - K)^{-1} \qquad\qquad K = \Theta(\mathrm{Id}_n + \Theta)^{-1};$$

see [23]. Note that the matrix Θ is typically denoted L in the statistics literature. In our model, the probability $\mathrm{Prob}(Z = I)$ of observing a subset I is proportional to the principal minor $\det(\Theta_I)$ indexed by that subset. In symbols,

$$p_I \;=\; \det(\Theta_I)/Z \qquad \text{for } I \subseteq [n]. \tag{7.2}$$

Here $\det(\Theta_\emptyset) = 1$, and the partition function is given by the sum of all 2^n principal minors:

$$Z \;=\; \sum_{I\subseteq[n]} \det(\Theta_I) \;=\; \det(\Theta + \mathrm{Id}_n),$$

where Id_n denotes the identity matrix of size $n \times n$. The maximum likelihood estimation problem can then be solved by maximizing (7.1) subject to the constraint that q is in the image of the principal minor map; see [101]. Practitioners use the unconstrained, parametric log-likelihood function

$$L_u(\Theta) = \sum_{I \subseteq [n]} u_I \log(\det(\Theta_I)) - \Big(\sum_{I \subseteq [n]} u_I\Big) \log(\det(\Theta + \mathrm{Id}_n)). \tag{7.3}$$

In Theorem 7.2.1, we express the number of critical points of the parametric log-likelihood function in terms of the ML degree of the hyperdeterminantal variety.

We now discuss the second scenario. If the eigenvalues of K are on the boundary $\{0, 1\}$ of the interval $[0, 1]$, then K parameterizes a *projection DPP*. This is the setting of Chapter 8. Since K satisfies its minimal polynomial and is real symmetric we have $K^2 = K$ and hence K is the unique orthogonal projection matrix onto its image. In other words, K is a point in the projection Grassmannian $\mathrm{Gr}(\mathrm{rank}(K), n)$. In Chapter 8, we denote K by P for "projection." The state space of a projection DPP with kernel P is the set of k-subsets of $[n]$, denoted $\binom{[n]}{k}$, and the probability of observing a k-subset I is $q_I = \det(P_I)$. Now the maximum likelihood estimation problem can be solved by maximizing (7.1) subject to the constraint that q is in the image of a degeneration of the principal minor map; see [2, Section 6]. When $k = 2$, all complex critical points of the log-likelihood function are real, positive, and local maxima (Theorem 8.1.5). We also give a formula for the ML degree of the Grassmannian (Theorem 8.6.2).

In Chapter 9, we introduce *squared linear models*, which share the property that all complex critical points of the log-likelihood function are local maxima. We go further for these models, by describing the ideal of the likelihood correspondence (Section 9.3), degenerate solutions (Section 9.4), and log-Voronoi cells (Section 9.5). We conclude with a discussion of projection DPPs that are also squared linear models, called *linear DPPs*, (Section 9.6).

Chapter 7

Likelihood Geometry of L-Ensembles

The motivation for this chapter is the work on likelihood inference by Brunel, Moitra, Rigollet and Urschel [23]. We shall answer their question [23, Conjecture 12] about critical points of the log-likelihood function for L-ensembles.

The model $\mathcal{M}_n \cap \Delta_{2^n-1}$ for L-ensembles on $[n]$ is the image of real symmetric positive-definite matrices under the principal minor map. The variety $\mathcal{M}_n$ is cut out by the *hyperdeterminantal ideal* that was studied by Holtz-Sturmfels [66], Oeding [101], and Al Ahmadieh-Vinzant [2]. In the first non-trivial case $n = 3$, our model is the zero set (in Δ_7 or in $\mathbb{P}^7$) of the hyperdeterminant of format $2 \times 2 \times 2$, which is the quartic

$$\begin{aligned}\mathrm{Det} \ = \ & q_{000}^2 q_{111}^2 + q_{001}^2 q_{110}^2 + q_{011}^2 q_{100}^2 + q_{010}^2 q_{101}^2 + 4q_{000}q_{011}q_{101}q_{110} + 4q_{001}q_{010}q_{100}q_{111} \\ & - 2q_{000}q_{001}q_{110}q_{111} - 2q_{000}q_{010}q_{101}q_{111} - 2q_{000}q_{011}q_{100}q_{111} \\ & - 2q_{001}q_{010}q_{101}q_{110} - 2q_{001}q_{011}q_{100}q_{110} - 2q_{010}q_{011}q_{100}q_{101}.\end{aligned} \tag{7.4}$$

In (7.4), binary strings of length 3 represent subsets of $[3] = \{1, 2, 3\}$. For $n \geq 4$, one takes all occurrences of the hyperdeterminant (7.4) in a tensor of format $2 \times 2 \times \cdots \times 2$ to cut out $\mathcal{V}_n$ [101]. A different representation of the variety $\mathcal{M}_n$ was found by Al Ahmadieh-Vinzant [2].

We consider the restriction of (7.1) to the hyperdeterminantal variety $\mathcal{M}_n$. For small n,

$$\mathrm{MLdegree}(\mathcal{M}_2) = 1 \quad \text{and} \quad \mathrm{MLdegree}(\mathcal{M}_3) = 13. \tag{7.5}$$

The number 13 for the hyperdeterminant was computed in [39, Example 2.2.10]. The number 1 arises because $\mathcal{M}_2 = \mathbb{P}^3$. But even this tiny case is interesting, as we shall see in Example 7.0.1. The DPP model is a semialgebraic set $\mathcal{M}_n$ of dimension $\binom{n+1}{2}$ in the simplex Δ_{2^n-1} whose points are the probability distributions on $2^{[n]}$.

In machine learning [23, 59, 80, 111], one uses the parametric form of the log-likelihood (7.3). Grigorescu, Juba, Wimmer and Xie [59] showed that computing the maximum of this log-likelihood function is an NP-complete problem. This was conjectured by Kulesza [80].

The next example illustrates the distinction between the formulations in (7.1) and (7.3).

Example 7.0.1 ($n = 2$)**.** We write $u_\emptyset, u_1, u_2, u_{12}$ for the observed counts of subsets of $\{1, 2\}$. The model $\mathcal{M}_2$ is given by the principal minors of a 2×2 matrix $\Theta = (\theta_{ij})$, namely

$$q_\emptyset = \frac{1}{Z},\ q_1 = \frac{\theta_{11}}{Z},\ q_2 = \frac{\theta_{22}}{Z},\ q_{12} = \frac{\theta_{11}\theta_{22} - \theta_{12}^2}{Z}, \quad \text{where} \quad Z = \theta_{11}\theta_{22} - \theta_{12}^2 + \theta_{11} + \theta_{22} + 1.$$

We view $L_u(q) = u_\emptyset \log(q_\emptyset) + u_1 \log(q_1) + u_2 \log(q_2) + u_{12} \log(q_{12})$ as a function of $\theta_{11}, \theta_{12}, \theta_{22}$. Setting its partial derivatives to zero gives three rational function equations in three unknowns. The solutions Θ^* of these equations are the critical points of L_u. A computation reveals

$$\theta_{11}^* = \frac{u_1}{u_\emptyset},\ \theta_{12}^* = \pm\frac{\sqrt{u_1 u_2 - u_\emptyset u_{12}}}{u_\emptyset},\ \theta_{22}^* = \frac{u_2}{u_\emptyset} \ \text{ or } \ \theta_{11}^* = \frac{u_1 + u_{12}}{u_\emptyset + u_2},\ \theta_{12}^* = 0,\ \theta_{22}^* = \frac{u_2 + u_{12}}{u_\emptyset + u_1}. \tag{7.6}$$

From the implicit perspective emphasized in [68], there is one solution. Its two preimages under the 2-1 parametrization of $\mathcal{M}_2$ are shown on the left in (7.6). On the right in (7.6) is a ramification point of that 2-1 map. It is not a critical point of (7.1) on $\mathcal{M}_2 = \mathbb{P}^3$. ◇

In Section 7.1, we offer a detailed investigation of the case $n = 3$. In particular, we present our counterexample to [23, Conjecture 12]. The number of critical points of the parametric log-likelihood (7.3) is found to be $4 \cdot 13 + 2 \cdot 1 + 2 \cdot 1 + 2 \cdot 1 + 1 = 59$. This is the $n = 3$ analogue to the count $2 \cdot 1 + 1$ for the three solutions (7.6) in Example 7.0.1. Section 7.2 features a general formula for counting complex critical points when n is arbitrary. This involves contributions from all possible block decompositions of the matrix Θ. These decompositions are called *partial decouplings* in the DPP literature. Our Theorem 7.2.1 generalizes [23, Theorem 11], where it is assumed that the data vector u lies on the model $\mathcal{M}_n$. In Section 7.3, we apply numerical methods to our problem. Going well beyond (7.5), we compute some solutions for $n = 4, 5$ with the software `HomotopyContinuation.jl` [20]. After replacing (7.2) with a birational parametrization, we apply the monodromy method (see Section 2.5) for rational likelihood equations in [1, 121]. The focus is on solutions that are real and positive-definite.

7.1 Three-by-three Matrices

In this section, we examine the likelihood geometry of the DPP model with $n = 3$. The model parameters are the six entries θ_{ij} of the symmetric 3×3 matrix Θ. Fix any data vector $u = (u_\emptyset, u_1, u_2, u_3, u_{12}, u_{13}, u_{23}, u_{123})$. The sum of its eight coordinates u_I is the sample size of the data, here denoted $|u|$. The parametric log-likelihood function (7.3) equals

$$\begin{aligned} L_u \ = \ & u_1 \log(\theta_{11}) + u_2 \log(\theta_{22}) + u_3 \log(\theta_{33}) + \ u_{12} \log(\theta_{11}\theta_{22} - \theta_{12}^2) + u_{13} \log(\theta_{11}\theta_{33} - \theta_{13}^2) \\ & + u_{23} \log(\theta_{22}\theta_{33} - \theta_{23}^2) \ + \ u_{123} \log(\det(\Theta)) \ - \ |u| \log\big(\det(\Theta + \mathrm{Id}_3)\big). \end{aligned} \tag{7.7}$$

Setting the partial derivatives of L_u to zero gives six rational function equations in six unknowns. The solutions to these equations are the critical points of L_u. Our theory in Section 7.3 predicts 59 critical points when the data vector u is generic. Among these are $4 \cdot 13 = 52$ solutions arising from critical points of (7.1) on the hyperdeterminantal hypersurface $\mathcal{M}_3$. These critical points Θ^* have $\theta^*_{ij} \neq 0$ for all i, j. The clusters of four arise by multiplying two of the three off-diagonal entries θ^*_{ij} by -1. This does not change the distribution $p^* \in \mathcal{M}_3$.

The other $7 = 1 + 3(2 \cdot 1)$ critical points of (7.3) are extraneous, in the sense that they do not come from critical points of (7.1) on $\mathcal{M}_3$. This is analogous to the point on the right in (7.6). One of the seven special solutions is the diagonal matrix $\Theta^* = \mathrm{diag}(\theta^*_{11}, \theta^*_{22}, \theta^*_{33})$, with

$$\theta^*_{11} = \frac{u_1+u_{12}+u_{13}+u_{123}}{u_\emptyset + u_2 + u_3 + u_{23}}, \;\; \theta^*_{22} = \frac{u_2+u_{12}+u_{23}+u_{123}}{u_\emptyset + u_1 + u_3 + u_{13}}, \;\; \theta^*_{33} = \frac{u_3+u_{13}+u_{23}+u_{123}}{u_\emptyset + u_1 + u_2 + u_{12}}. \tag{7.8}$$

The other six special solutions come in three pairs, one for each decomposition of a 3×3 matrix into a 2×2 block and a 1×1 block. One such pair of critical points has the form

$$\Theta^* = \begin{bmatrix} \theta^*_{11} & \theta^*_{12} & 0 \\ \theta^*_{12} & \theta^*_{22} & 0 \\ 0 & 0 & \theta^*_{33} \end{bmatrix}, \tag{7.9}$$

where θ^*_{33} is the expression on the right in (7.8), and the other three entries in (7.9) are

$$\theta^*_{11} = \frac{u_1 + u_{13}}{u_\emptyset + u_3}, \;\; \theta^*_{12} = \pm\frac{\sqrt{(u_1+u_{13})(u_2+u_{23}) - (u_\emptyset+u_3)(u_{12}+u_{123})}}{u_\emptyset + u_3}, \;\; \theta^*_{22} = \frac{u_2 + u_{23}}{u_\emptyset + u_3}. \tag{7.10}$$

We now move away from the hypothesis that the data are generic. Namely, we make the assumption that u lies on the variety $\mathcal{M}_n$. In fact, we even assume that $\frac{1}{u_\emptyset}u$ lies in the model, meaning that it is the vector of principal minors of some real symmetric $n \times n$ matrix. This is the standing assumption on the data in the article [23] to which we shall turn shortly.

Example 7.1.1 (Data in the model)**.** We consider the following data for $n = 3$:

$$(u_\emptyset, u_1, u_2, u_3, u_{12}, u_{13}, u_{23}, u_{123}) \;=\; (1, 8, 22, 18, 151, 135, 360, 2412) \tag{7.11}$$

This vector lies in the model because its entries are the principal minors of any of the matrices

$$\begin{bmatrix} 8 & 5 & 3 \\ 5 & 22 & 6 \\ 3 & 6 & 18 \end{bmatrix}, \begin{bmatrix} 8 & -5 & -3 \\ -5 & 22 & 6 \\ -3 & 6 & 18 \end{bmatrix}, \begin{bmatrix} 8 & -5 & 3 \\ -5 & 22 & -6 \\ 3 & -6 & 18 \end{bmatrix}, \begin{bmatrix} 8 & 5 & -3 \\ 5 & 22 & -6 \\ -3 & -6 & 18 \end{bmatrix}. \tag{7.12}$$

By construction, these are the four global maxima of the log-likelihood function in (7.7), and they map to the global maximum of (7.1) on $\mathcal{M}_3$. Among the other 12 complex critical points on $\mathcal{M}_3$, six are real and lie on $\mathcal{M}_3$. Four of these correspond to the positive-definite matrices

$$\begin{array}{ll} \theta^*_{11} = 7.72799090116006 & \theta^*_{12} = 4.14366972540362 \\ \theta^*_{13} = 1.87300176302618 & \theta^*_{22} = 20.1464857136673 \\ \theta^*_{23} = 0.82526924316919 & \theta^*_{33} = 16.4735825997691 \end{array} \qquad \begin{array}{ll} \theta^*_{11} = 6.92478592243203 & \theta^*_{12} = 0.42796700405714 \\ \theta^*_{13} = 1.80374923458180 & \theta^*_{22} = 19.2531487326101 \\ \theta^*_{23} = 4.44778298807768 & \theta^*_{33} = 17.4175047796638 \end{array}$$

$$\begin{array}{ll} \theta^*_{11} = 7.56880693316022 & \theta^*_{12} = 4.28496510066628 \\ \theta^*_{13} = 0.86306207237349 & \theta^*_{22} = 21.3776618810445 \\ \theta^*_{23} = 4.79253523095731 & \theta^*_{33} = 17.0982120638953 \end{array} \qquad \begin{array}{ll} \theta^*_{11} = 7.57820456385679 & \theta^*_{12} = -3.8397212783772 \\ \theta^*_{13} = 0.98046698151082 & \theta^*_{22} = 20.9281938578911 \\ \theta^*_{23} = 3.86656494286390 & \theta^*_{33} = 17.1007249363163. \end{array}$$

In addition, the parametric log-likelihood (7.7) has seven more critical points with a block structure. These are obtained by substituting (7.11) into the formulas (7.8), (7.9), (7.10). ⋄

We now turn to the approach of Brunel, Moitra, Rigollet and Urschel [23], and we present a counterexample to [23, Conjecture 12], which states that there are no critical points other than those obtained from *partial decouplings*. Partial decouplings correspond to the block decompositions we saw in (7.6), (7.8) and (7.9). We discuss these in Section 7.2 for general n.

The set-up in [23] uses the parametric form (7.3) of the log-likelihood, i.e., L_u is a function on the cone of positive-definite $n \times n$ matrices. Furthermore, [23] assumes that the data vector u is sampled from the model $\mathcal{M}_n$. When u is given by the principal minors of some positive-definite $n \times n$ matrix, [23, Theorem 11] shows that all partial decouplings are critical points of L_u. This includes the empirical distribution $p^* = \frac{1}{|u|}u$, which is the only critical point from partial decouplings with full support. There are exponentially many critical points from partial decouplings, and the question is whether these are all. We show that the answer is no.

Proposition 7.1.2. *For $n = 3$, the log-likelihood function in Θ given by some $u \in \mathcal{M}_3$ has critical points that do not correspond to partial decouplings. This resolves [23, Conjecture 12].*

Proof. The proof is furnished by Example 7.1.1. The log-likelihood function L_u for u in (7.11) has 28 fully-supported real critical points, 20 of which are positive-definite. The four matrices in (7.12) are the global maxima. Below that, we show matrix representatives for four positive-definite critical points not corresponding to partial decouplings. □

7.2 Partial Decouplings

We saw that some of the critical points of the parametric log-likelihood (7.3) are matrices Θ^* with a block decomposition. These were called partial decouplings in [23]. Such critical points were characterized in [23, Theorem 11], under the hypothesis that the data vector u lies in the model $\mathcal{M}_n$. In what follows we offer a generalization of that result. We no longer assume $u \in \mathcal{M}_n$. From now on, we allow $u = (u_I : I \subseteq [n])$ to be any complex vector of length 2^n. If u is generic, then all critical points of (7.1) on $\mathcal{M}_n$ have fully supported preimages under the principal minor map. We write μ_n for the ML degree of the projective variety $\mathcal{M}_n$. We know from (7.5) that $\mu_1 = \mu_2 = 1$ and $\mu_3 = 13$. In the next section we shall show that $\mu_4 = 3526$.

Recall that a *set partition* of $[n] = \{1, 2, \dots, n\}$ is a set $\pi = \{\pi_1, \dots, \pi_k\}$, where the π_i are non-empty pairwise disjoint subsets of $[n]$ whose union equals $[n]$. We write Π_n for the set of all set partitions of $[n]$. The cardinality $|\Pi_n|$ is the *Bell number*, which equals $2, 5, 15, 52, 203, 877$ for $n = 2, 3, 4, 5, 6, 7$. See Examples 7.2.2 and 7.2.3 for the cases $n = 3, 4$.

Theorem 7.2.1. *The critical points Θ^* of the parametric log-likelihood function L_u in (7.3) are found by solving various likelihood equations on submodels $\mathcal{M}_r$ for $r \le n$. If u is generic,* *in the sense of algebraic geometry, then the total number of complex critical points of L_u is*

$$\sum_{\pi\in\Pi_n}\prod_{i=1}^{|\pi|}\left(2^{|\pi_i|-1}\,\mu_{|\pi_i|}\right). \tag{7.13}$$

Given a partition $\pi = \{\pi_1,\ldots,\pi_k\}$, the upper limit $|\pi| = k$ is the number of parts of π. The phrase "generic in the sense of algebraic geometry" means that there exists a proper subvariety in the data space $\mathbb{R}^{2^n}$ such that the statement holds for all vectors u outside that variety. In particular, it holds with probability one for a randomly selected vector $u \in \mathbb{R}^{2^n}$.

Example 7.2.2. For $n = 3$, we have $\Pi_3 = \{\{1,2,3\},\{12,3\},\{13,2\},\{23,1\},\{123\}\}$. Thus the number (7.13) of critical points is $1\cdot 1 + 2\cdot 1 + 2\cdot 1 + 2\cdot 1 + 4\cdot 13 \;=\; 59$. These 59 solutions were described in Section 7.1, with an explicit numerical instance in Example 7.1.1. $\diamond$

Example 7.2.3. For $n = 4$, there are 28 441 critical points. The sum (7.13) is over the 15 set partitions of $[4]$. The biggest summand is $2^{4-1}\cdot 3\,526 = 28\,208$, for $\pi = \{1234\}$ with $k = 1$. The partitions with $k \geq 2$ contribute the summands $52, 52, 52, 52, 4, 4, 4, 2, 2, 2, 2, 2, 2, 1$. $\diamond$

Proof of Theorem 7.2.1. We fix one partition $\pi = \{\pi_1,\ldots,\pi_k\}$ in Π_n. Suppose that Θ is a symmetric $n\times n$ matrix that has the block structure π, so some of the entries are zero. However, all nonzero entries of Θ are distinct unknowns. We write $\Theta = \Theta_{\pi_1}\oplus\Theta_{\pi_2}\oplus\cdots\oplus\Theta_{\pi_k}$.

Let $L_u(\Theta)$ denote the evaluation of the log-likelihood function at the block matrix Θ. Then, $L_u(\Theta)$ is a function in $\sum_{i=1}^k\binom{|\pi_i|+1}{2}$ unknowns, namely the entries of the k blocks Θ_{π_i}. Assuming u to be generic, we count critical points Θ^* for which all entries in the blocks $\Theta^*_{\pi_i}$ are nonzero. We claim that the total number of these critical points is equal to

$$\prod_{i=1}^{k}\left(2^{|\pi_i|-1}\mu_{|\pi_i|}\right). \tag{7.14}$$

Our π-restricted log-likelihood function admits an additive decomposition

$$L_u(\Theta) \;=\; \sum_{i=1}^{k} L_{v^{(i)}}(\Theta_{\pi_i}). \tag{7.15}$$

Here $v^{(i)}$ is a vector in $\mathbb{R}^{2^{|\pi_i|}}$, indexed by subsets of π_i, that is obtained from u by a linear transformation. To see this, we use the following identity for the minors of our block matrix:

$$\log(\det(\Theta_I)) \;=\; \log\Big(\prod_{i=1}^{k}\det(\Theta_{I\cap\pi_i})\Big) \;=\; \sum_{i=1}^{k}\log(\det(\Theta_{I\cap\pi_i})) \qquad \text{for all } I\subseteq[n].$$

The analogous decomposition holds for the log-partition function $\log(Z) = \log(\det(\Theta + \mathrm{Id}_n))$. From this we conclude that (7.15) holds if we define the restricted data vector $v^{(i)}$ as follows:

$$v^{(i)}_J \;=\; \sum\big\{u_I : I\subseteq[n] \text{ and } I\cap\pi_i = J\big\} \qquad \text{for all } J\subseteq\pi_i.$$

The number of fully supported critical points of $L_{v^{(i)}}(\Theta_{\pi_i})$ is equal to $2^{|\pi_i|-1}\mu_{|\pi_i|}$. Indeed, the data vector $v^{(i)}$ is still generic, and we are computing critical points on the variety $\mathcal{M}_{|\pi_i|}$. Each critical point on $\mathcal{M}_{|\pi_i|}$ comes from a cluster of $2^{|\pi_i|-1}$ critical matrices $\Theta^*_{\pi_i}$. Since the summands in (7.15) involve disjoint sets of unknowns, these critical points combine for $i = 1, \ldots, k$. Therefore, the total number of critical points of $L_u(\Theta)$ is the product in (7.14).

The next step is to show that the points above are critical points of $L_u(\Theta)$, where Θ is now an $n\times n$ matrix with all $\binom{n+1}{2}$ entries distinct unknowns. To see this, consider the partial derivative of $L_u(\Theta)$ with respect to any off-diagonal parameter θ_{ij}. This partial derivative is an element of the ring R that is obtained by localizing the polynomial ring $\mathbb{R}[\Theta]$ at the product of Z and all principal minors of Θ. We claim that $\partial L_u/\partial\theta_{ij}$ lies in the following ideal of R, where the intersection is over all 2^{n-2} partitions $K \cup L = [n]$ with $i \in K$ and $j \in L$:

$$\partial L_u/\partial\theta_{ij} \;\in\; \bigcap_{K\ni i, L\ni j} \big\langle\, \theta_{kl} \,:\, k\in K, l \in L\,\big\rangle. \tag{7.16}$$

Let $\mathcal{I}$ be the ideal generated by the monomials in $\partial\mathrm{det}(\Theta)/\partial\theta_{ij}$. We claim that the ideals on the right side of (7.16) are associated primes of $\mathcal{I}$. Namely, $\langle\, \theta_{kl} \,:\, k \in K, l \in L\rangle = (\mathcal{I} : a)$ where $a = \prod\{\theta_{h_1,h_2} \,:\, (h_1, h_2) \in K\times K \cup L\times L\}$. Since $\mathcal{I}$ is a monomial ideal, it suffices to show the following: if m is a monomial, then there exists a monomial in $\partial\mathrm{det}(\Theta)/\partial\theta_{ij}$ which divides ma if and only if θ_{kl} divides m from some $k \in K$ and $l \in L$. Every monomial in the determinant has the form $s = \prod_{h=1}^{n} \theta_{h\sigma(h)}$ for some permutation σ. Suppose that $\partial s/\partial\theta_{ij}$ is nonzero and divides ma. Then $\sigma(i)$ must be j. Since $i \in K$ and $j \in L$, for σ to be a bijection, there must be some $k \in K$ and $l \in L$ such that $\sigma(l) = k$. Since θ_{lk} cannot divide a, but θ_{lk} divides $\partial s/\partial\theta_{ij}$, which, in turn, divides ma, it follows that $\theta_{lk} = \theta_{kl}$ divides m.

Now suppose θ_{kl} divides m and let $s = \prod_{h=1}^{n} \theta_{h\sigma(h)}$ where $\sigma = (i\ j\ l\ k)$ if $k \neq i$ and $l \neq j$, $\sigma = (i\ j\ l)$ if $k = i$ and $l \neq j$, $\sigma = (i\ j\ k)$ if $k \neq i$ and $l = j$, and $\sigma = (i\ j)$ if $k = i$ and $l = j$. Up to scaling, $\partial s/\partial\theta_{ij} = s/\theta_{ij}$, which divides ma, as θ_{kl} divides m and $s/(\theta_{ij}\theta_{kl})$ divides a. This concludes the proof of our claim that $\langle\, \theta_{kl} \,:\, k \in K, l \in L\rangle = (\mathcal{I} : a)$.

The same statement holds when Θ is replaced by $\Theta + \mathrm{Id}_n$ or any principal submatrix Θ_I. Since $\partial L_u/\partial\theta_{ij}$ is in the ideal of these determinants, we have established the inclusion (7.16).

Now, fix any set partition $\pi \in \Pi_n$, and suppose that i and j lie in distinct blocks of π. The partial derivative $\partial L_u/\partial\theta_{ij}$ vanishes identically when the full matrix Θ is replaced by the block matrix $\Theta_{\pi_1} \oplus \cdots \oplus \Theta_{\pi_k}$. This follows from (7.16). This vanishing property shows that the block matrices $\Theta^*_{\pi_1} \oplus \cdots \oplus \Theta^*_{\pi_k}$ derived above are, in fact, critical points of $L_u(\Theta)$.

At this point, we know that (7.14) is a lower bound for the number of critical points. The final step in our proof is to show that no further critical points exist. To see this, let Θ^* be any critical point of the parametric log-likelihood (7.3). Suppose that the support of Θ^* is not contained in any proper block structure. Then its fiber over $\mathcal{M}_n$ consists of 2^{n-1} distinct matrices, which are reduced points in that fiber. This implies that the common image in $\mathcal{M}_n$ of the 2^{n-1} matrices is a critical point of (7.1). Hence Θ^* was counted in (7.14), by the summand for $\pi = \{[n]\}$. If u is generic then we can conclude that Θ^* has no zero coordinates.

It remains to consider critical points Θ^* that conform to the block structure for some partition $K \cup L = [n]$, i.e. $\theta^*_{kl} = 0$ for all $k \in K$ and $l \in L$. We apply the previous argument inductively to the respective blocks Θ^*_K and Θ^*_L, and eventually we arrive at $\Theta^* = \Theta^*_{\pi_1} \oplus \cdots \oplus \Theta^*_{\pi_k}$ for some partition π of $[n]$. This means that Θ^* was counted in (7.14). □

We have shown that, for generic data vectors u, the support of each critical point Θ^* is precisely given by one of the block structures. This property can fail when u is not generic.

Example 7.2.4. Fix $n = 3$ and $u = (2, 1, 3, 7, 9, 10, 19, 22)$. Then L_u has 59 distinct critical points, as in Example 7.2.2, with 52 from the trivial partition $\pi = \{123\}$. One of these is

$$\Theta^* = \begin{bmatrix} 2 & 0 & 2 \\ 0 & 4 & 3 \\ 2 & 3 & 7 \end{bmatrix}.$$

The zero entry $\theta^*_{12} = 0$ is accidental, not due to any block structure. Here, u is not generic. ⋄

It is now instructive to revisit the implicit formulation of our MLE problem. We seek points p on the hyperdeterminant $V(\mathrm{Det}) \subset \mathbb{P}^7$ such that the following matrix has rank ≤ 2:

$$\begin{bmatrix} u_\emptyset & u_1 & u_2 & u_3 & u_{12} & u_{13} & u_{23} & u_{123} \\ p_{000} & p_{100} & p_{010} & p_{001} & p_{110} & p_{101} & p_{011} & p_{111} \\ p_{000}\frac{\partial \mathrm{Det}}{\partial p_{000}} & p_{100}\frac{\partial \mathrm{Det}}{\partial p_{100}} & p_{010}\frac{\partial \mathrm{Det}}{\partial p_{010}} & p_{001}\frac{\partial \mathrm{Det}}{\partial p_{001}} & p_{110}\frac{\partial \mathrm{Det}}{\partial p_{110}} & p_{101}\frac{\partial \mathrm{Det}}{\partial p_{101}} & p_{011}\frac{\partial \mathrm{Det}}{\partial p_{011}} & p_{111}\frac{\partial \mathrm{Det}}{\partial p_{111}} \end{bmatrix}.$$

We require each coordinate p_{ijk} to be non-zero, and also $\sum_{ijk} p_{ijk} \neq 0$. We further disallow p to lie in the singular locus of $V(\mathrm{Det})$, i.e. the three flattenings of the $2\times 2\times 2$ tensor p are 2×4 matrices of rank 2. This system has 13 solutions, even for the special u in Example 7.2.4.

7.3 Numerical Computations

We now discuss the solution of the likelihood equations using methods from numerical algebraic geometry. For our computations we use the software `HomotopyContinuation.jl` due to Breiding and Timme [20] with the certification feature in [21]. Our approach is based on the monodromy method (see Section 2.5) for rational likelihood equations that was developed in [1, 121].

The underlying idea is as follows. We consider the likelihood equations $\nabla L_u(\Theta) = 0$ where both u and Θ are unknowns. These define the parametric likelihood correspondence. In contrast, with Section 2.4, we do not saturate out the ramification points:

$$\{\Theta^* \in \mathrm{Sym}^2\mathbb{C}^n \backslash \mathcal{H} : \partial L_u / \partial \theta_{ij}(\Theta^*) = 0\}. \tag{7.17}$$

where $\mathcal{H} = \{\Theta^* \in \mathrm{Sym}^2\mathbb{C}^n : \prod_{I\subseteq[n]} \det(\Theta^*_I) = 0\}$. This likelihood correspondence has many irreducible components, one for each set partition $\pi \in \Pi_n$. This is the geometric interpretation of Theorem 7.2.1. The parametric likelihood correspondence $\mathcal{P}_{\mathcal{M}_n}$ in the sense of Section 2.4 is the main component of (7.17), corresponding to the partition $\pi = \{[n]\}$.

These are the preimages of the critical points of (7.1) restricted to $\mathcal{M}_n$. Numerical algebraic geometry restricts to this component automatically.

The likelihood equations are linear in u. We thus can fix a random complex matrix Θ, and then solve for a matching u. Afterwards, we fix u and we vary Θ. By running monodromy loops in `HomotopyContinuation.jl`, one eventually finds all solutions Θ^* to $\nabla L_u(\Theta) = 0$ for that fixed u. Here "all" means all critical points of (7.1) on $\mathcal{M}_n$, because the monodromy loops stay on the main irreducible component of the likelihood correspondence.

The program terminates after a heuristic criterion is satisfied. If this happens, then we can be confident that all solutions have been found, and that the number of solutions is equal to $\mu_n = \mathrm{MLdegree}(\mathcal{M}_n)$. However, there is still a tiny chance that some solutions have been missed, which would mean that the true μ_n is a little larger than the current count. At this stage, we apply the command `certify` which generates a proof, based on interval arithmetic, that all floating-point approximations that were found are, in fact, distinct solutions [21].

The pipeline described above proves that the number we found is a lower bound for μ_n. To prove that it is also an upper bound, one would need some insights from intersection theory. But this is still missing for many statistical models, including the one treated in this paper. The process described above is quite fast for $n = 4$, and it yields the following result.

Proposition 7.3.1. *The ML degree of the L-ensemble model $\mathcal{M}_4$ satisfies $\mu_4 \geq 3\,526$. Based on our numerical computation, we are confident that $\mu_4 = 3\,526$.*

The principal minor map is 2^{n-1}-to-1. For our computations we use a reparametrization which makes the map 1-to-1, reducing the number of paths we need to track by a factor of 2^{n-1}. We first show the new coordinates for $n = 3$.

Example 7.3.2 (Birational Reparametrization)**.** We reparametrize our matrix

$$\Theta \;=\; \begin{bmatrix} \theta_{11} & \theta_{12} & \theta_{13} \\ \theta_{12} & \theta_{22} & \theta_{23} \\ \theta_{13} & \theta_{23} & \theta_{33} \end{bmatrix} \tag{7.18}$$

so that the principal minor map $\Theta \mapsto \big(\theta_{11}, \theta_{22}, \theta_{33}, \theta_{11}\theta_{22}-\theta_{12}^2, \theta_{11}\theta_{33}-\theta_{13}^2, \theta_{22}\theta_{33}-\theta_{23}^2, \det(\Theta)\big)$ becomes injective by replacing the monomials $\theta_{12}^2, \theta_{13}^2, \theta_{23}^2, \theta_{12}\theta_{13}\theta_{23}$ with new variables. These four monomials are algebraically dependent, so we introduce three new variables: $x_{12} = \theta_{12}^2$, $x_{13} = \theta_{13}^2$, and $x_{23} = \theta_{12}\theta_{13}\theta_{23}$. Solving for the θ_{ij}, we now substitute the following into (7.18):

$$\theta_{12} \;=\; \sqrt{x_{12}} \qquad\qquad \theta_{13} \;=\; \sqrt{x_{13}} \qquad\qquad \theta_{23} \;=\; x_{23}/\sqrt{x_{12}x_{13}}.$$

This yields a birational map between $\mathbb{C}^6$ and the hypersurface $V(\mathrm{Det})$ in $\mathbb{P}^7$. $\diamond$

The general case is similar. For $n \geq 4$, we replace all off-diagonal parameters as follows:

$$\text{For } i \neq j \text{ we set} \quad \theta_{ij} \;=\; \begin{cases} \sqrt{x_{ij}} & \text{if } i = 1, \\ x_{ij}/\sqrt{x_{1i}x_{1j}} & \text{otherwise.} \end{cases}$$

After this, the log-likelihood L_u is a function in the n diagonal entries θ_{ii} and the $\binom{n}{2}$ new variables x_{ij}. Its partial derivatives give a system of $\binom{n+1}{2}$ rational function equations in $\binom{n+1}{2}$ variables. We now use the command `monodromy_solve` to solve this system. We find and certify 13 complex solutions for $n = 3$ and 3 526 complex solutions for $n = 4$. The $n = 3$ computations run in under a second. For $n = 4$, the computation takes about 20 minutes. These times can be improved significantly by using multiple threads in `Julia`.

In the statistical application to DPP, we seek critical points that are real, not complex. Ideally, we want Θ^* to be positive-definite. We ran the above computation on many data vectors u, with the aim of maximizing the number of real critical points. The reality of the critical points is governed by the discriminant of the optimization problem; see Section 5.2 and [77]. We found data with the property that 11 of the 13 critical points are positive definite.

Example 7.3.3 (11 Positive-Definite Critical Points)**.** Fix the data $u = (1, 5, 5, 5, 5, 5, 5, 1)$. The likelihood function L_u has two complex critical points. The remaining 11 critical points correspond to positive-definite matrices. Nine come from by permuting indices on three Θ^*:

$$\begin{array}{llllll}
\theta^*_{11} = 6.0 & \theta^*_{12} = 4.4721360 & \theta^*_{11} = 2.142857 & \theta^*_{12} = 0.8571429 & \theta^*_{11} = 5.0 & \theta^*_{12} = 3.4641016 \\
\theta^*_{13} = 4.472136 & \theta^*_{22} = 3.8888889 & \theta^*_{13} = 2.236068 & \theta^*_{22} = 2.1428571 & \theta^*_{13} = 3.4641016 & \theta^*_{22} = 3.0 \\
\theta^*_{23} = 3.111111 & \theta^*_{33} = 3.8888889 & \theta^*_{23} = 2.236068 & \theta^*_{33} = 3.5 & \theta^*_{23} = 2.0 & \theta^*_{33} = 3.0.
\end{array}$$

The remaining two critical points $p^* \in \mathcal{M}_3$ are invariant under permuting the indices $1, 2, 3$:

$$\begin{array}{llll}
\theta^*_{11} = 5.652906131 & \theta^*_{12} = 5.265758657 & \theta^*_{11} = 1.742592619 & \theta^*_{12} = -0.840402407 \\
\theta^*_{13} = 5.265758657 & \theta^*_{22} = 5.652906131 & \theta^*_{13} = -0.840402407 & \theta^*_{22} = 1.742592619 \\
\theta^*_{23} = 5.265758657 & \theta^*_{33} = 5.652906131 & \theta^*_{23} = -0.840402407 & \theta^*_{33} = 1.742592619.
\end{array}$$

Among the 11 critical points, five are local maxima, two of which are global maxima. The global maxima come from the matrices that are invariant under permuting indices, i.e. the matrices with constant diagonal and off-diagonal entries. The log-likelihood evaluates to -63.46051485 at these two points. The other local maxima come from permutations of the matrix with $\theta^*_{11} = 5.0$. The log-likelihood evaluates to -63.63109767 at these three points. $\diamond$

We now turn to the case $n = 4$, where numerical accuracy is already a notable challenge.

Example 7.3.4 ($n = 4$)**.** We fix $u = (u_\emptyset, u_1, \ldots, u_4, u_{12}, \ldots, u_{34}, u_{123}, \ldots, u_{234}, u_{1234})$ to be

$$u \;=\; (1, 12, 12, 12, 12, 12, 12, 12, 12, 12, 12, 12, 12, 12, 12, 1).$$

We certified 3 221 complex critical points p^* for the function L_u on the 10-dimensional variety $\mathcal{M}_4$. Among these 3 221, the global maximum is given by the principal minors of the matrix

$$\Theta^* \;=\; \begin{bmatrix} 6.5 & 5.5 & 5.97913 & 5.97913 \\ 5.5 & 6.5 & 5.97913 & 5.97913 \\ 5.97913 & 5.97913 & 6.5 & 5.5 \\ 5.97913 & 5.97913 & 5.5 & 6.5 \end{bmatrix}$$

For the other $305 = 3\,526 - 3\,221$ solutions, more careful path-tracking with homotopy methods is needed. We found 315 of our critical points to be real. Only 180 have real preimages Θ^* under the maximal minor map. Among these, 104 come from positive-definite matrices. These 104 are the statistically meaningful critical points. They include five local maxima. ⋄

We conclude this chapter by reporting on our computations for $n = 5$. We use the birational parametrization in Example 7.3.2. Our system consists of 15 rational function equations in 15 unknowns, namely the variables x_{ij} of our birational parametrization and the diagonal entries θ_{ii} of a 5×5 matrix Θ. Using 256 threads, we apply `monodromy_solve` to the 15×15 system of partial derivatives of L_u. In six days, we already found 29.5 million solutions. Hence the ML degree satisfies $\mu_5 \geq 29\,500\,000$.

The ML degree of the model of L-ensembles grows very rapidly with n, meaning that maximum likelihood estimation is likely very difficult for these models. One challenge in continuing this line of research is that the hyperdeterminantal variety $\mathcal{M}_n$ is complicated. We will turn to projection determinantal point processes in the next section. L-ensembles can be written as mixtures of projection determinantal point processes [67]. The model for projection DPPs has deep connections to the Grassmannian, which we will use to show that maximum likelihood estimation is still extremely difficult.

Chapter 8

Likelihood Geometry of the Squared Grassmannian

In Section 8.1, we explain the connection between projection determinantal point processes and the Grassmannian via the squared Grassmannian. We turn our attention to the squared Grassmannian in Section 8.2. We give a description of the prime ideal of $\mathrm{sGr}(2,n)$ in Theorem 8.2.3. In Section 8.3, we outline the approach for the proof of Theorem 8.1.4 and study the topology of the parametric model arising from (8.1). In Section 8.4, we study the stratification of the parametric model by a deletion map and prove Theorem 8.1.4. In Section 8.5, we turn to real critical points and prove Theorem 8.1.5. We compute the MLE for random data for $4 \leq n \leq 9$ and we record the runtimes. In Section 8.6, we turn to two other statistical models on the Grassmannian, one on the configuration space $X(k,n)$, and the other on the positive Grassmannian $\mathrm{Gr}(k,n)_{>0}$. We give a formula for the ML degree of the positive Grassmannian in Theorem 8.6.2.

Code for the numerical experiments in Section 8.5 is available in the `MathRepo` collection via `https://mathrepo.mis.mpg.de/likelihood_geometry_squared_grassmannian`.

8.1 Projection Determinantal Point Processes

The states of a projection determinantal point processes are the k-subsets of a finite set $[n]$. The model is a subset of $\Delta_{\binom{n}{k}-1}$ and is parametrized by orthogonal projection matrices. In particular, the model is the image of the projection Grassmannian under the map $P \mapsto (\det(P_I) : |I| = k)$. In the spirit of Section 3.6, we write these distributions in Plücker coordinates $(x_I)_{I\in\binom{[n]}{k}}$; see Section 3.2.

Lemma 8.1.1. *The $k\times k$ principal minors of P are proportional to squared Plücker coordinates:*

$$\det(P_I) \;=\; \frac{x_I^2}{\sum_{J\in\binom{[n]}{k}} x_J^2}.$$

Proof. Let A be any $n \times k$ matrix such that $P = A(A^TA)^{-1}A^T$. We write A_I for the $k \times k$ submatrix of A with row indices I. Then we have $P_I = (A_I) \cdot (A^TA)^{-1} \cdot A_I^T$. This is a product of three $k \times k$ matrices, so its determinant is the product of the three determinants. We find

$$\det(P_I) \;=\; \frac{\det(A_I)^2}{\det(A^TA)} \;=\; \frac{\det(A_I)^2}{\sum_J \det(A_J)^2} \;=\; \frac{x_I^2}{\sum_J x_J^2}.$$

Here, I and J are index sets in $\binom{[n]}{d}$. The middle equation uses the Cauchy-Binet formula. □

Lemma 8.1.1 offers a link between the two lives of the Grassmannian in Sections 3.2 and 3.3. We call the Zariski closure of our model the squared Grassmannian; see [2].

Definition 8.1.2 (Squared Grassmannian)**.** The *squared Grassmannian* $\mathrm{sGr}(k, n)$ is the image of the Grassmannian $\mathrm{Gr}(k, n) \subset \mathbb{P}^{\binom{n}{k}-1}$ in its Plücker embedding under the map

$$\mathrm{Gr}(k,n) \to \mathbb{P}^{\binom{n}{d}-1}, \;\; (x_I)_{I \in \binom{[n]}{k}} \mapsto (x_I^2)_{I \in \binom{[n]}{k}}.$$

Corollary 8.1.3. *The projection DPP is the discrete statistical model on the state space* $\binom{[n]}{k}$ *whose underlying algebraic variety is the squared Grassmannian* $\mathrm{sGr}(k, n)$*.*

Thus the MLE of a projection DPP is found by maximizing the implicit log-likelihood function (7.1) over $\mathrm{sGr}(k, n)$. We discuss the algebraic property of the variety $\mathrm{sGr}(k, n)$ in Section 8.2. The main result of this section is the description of the prime ideal of $\mathrm{sGr}(2, n)$ (Theorem 8.2.3).

We mainly focus on the case where $k = 2$, so our model has state space $\binom{[n]}{2}$. We determine the number of critical points of the implicit log-likelihood function:

Theorem 8.1.4. *The ML degree of* $\mathrm{sGr}(2, n)$ *is* $\frac{(n-1)!}{2}$ *for* $n \geq 3$*.*

In Section 8.5, we prove that all of these critical points are relevant from a statistical perspective. This is reminiscent of Varchenko's Theorem [27, Theorem 13].

Theorem 8.1.5. *For* $u \in \mathbb{N}^{\binom{n}{2}}$*, each critical point of the log-likelihood function* (7.1) *over* $\mathrm{sGr}(2, n)$ *is real, positive, and a local maximum when restricted to real points of* $\mathrm{sGr}(2, n)$*.*

The Grassmannian may be parametrized by the minors of $k \times n$ matrix $M_{k,n}$. This parametrization extends to the squared Grassmannian as

$$q_I = \frac{(M_{k,n})_I^2}{\sum_{J \in \binom{[n]}{k}} (M_{k,n})_J^2} \tag{8.1}$$

where $x_I = (M_{k,n})_I$ is the maximal minor of $M_{d,n}$ whose columns are indexed by I. We then have the parametric log-likelihood function for projection DPPs:

$$L_u(M_{k,n}) = \sum_{I \in \binom{[n]}{k}} u_I \log(\det((M_{k,n})_I)^2) - \Big(\sum_{I \in \binom{[n]}{k}} u_I\Big) \log\Big(\sum_{I \in \binom{[n]}{k}} \det((M_{k,n})_I)^2\Big). \tag{8.2}$$

If we restrict to the matrices $M_{k,n}$ whose first $k \times k$ square is the identity, we get a 2^{n-1}-to-1 parametrization; see [34, Section 4]. Therefore the number of critical points of (8.2) is 2^{n-1} times the ML degree of the squared Grassmannian.

Example 8.1.6 ($d = 2, n = 3$)**.** The implicit log-likelihood function is

$$L_u(q) = u_{12}\log(q_{12}) + u_{13}\log(q_{13}) + u_{23}\log(q_{23}) - (u_{12}+u_{13}+u_{23})\log(q_{12}+q_{13}+q_{23}).$$

Because $\mathrm{sGr}(2,3) = \mathbb{P}^2$, the MLE is u and the ML degree is one. We can find the same answer parametrically: if $M_{2,3} = \begin{pmatrix} 1 & 0 & t_{13} \\ 0 & 1 & t_{23} \end{pmatrix}$, then

$$L_u(M_{2,3}) = u_{12}\log(1) + u_{13}\log(t_{23}^2) + u_{23}\log(t_{13}^2) - (u_{12}+u_{13}+u_{23})\log(1+t_{13}^2+t_{23}^2)$$

whose partial derivatives are

$$\frac{\partial L_u}{\partial t_{13}} = \frac{2u_{23}}{t_{13}} - \frac{2(u_{12}+u_{13}+u_{23})t_{13}}{1+t_{13}^2+t_{23}^2} = 0, \qquad \frac{\partial L_u}{\partial t_{23}} = \frac{2u_{13}}{t_{23}} - \frac{2(u_{12}+u_{13}+u_{23})t_{23}}{1+t_{13}^2+t_{23}^2} = 0.$$

We can solve these equations by hand to find $t_{13} = \pm\sqrt{\frac{u_{23}}{u_{12}}}$ and $t_{23} = \pm\sqrt{\frac{u_{13}}{u_{12}}}$ for a total of four solutions. Our parametrization sends the parametric solution to the implicit one:

$$\begin{pmatrix} 1 & 0 & \pm\frac{\sqrt{u_{23}}}{\sqrt{u_{12}}} \\ 0 & 1 & \pm\frac{\sqrt{u_{13}}}{\sqrt{u_{12}}}. \end{pmatrix} \mapsto (1 : \frac{u_{13}}{u_{12}} : \frac{u_{23}}{u_{12}}). \qquad \diamond$$

8.2 The Squared Grassmannian

In this section, we study the dimension, degree, and homogeneous prime ideal of $\mathrm{sGr}(k,n)$. The two most basic invariants of any projective variety are its dimension and its degree.

Proposition 8.2.1. *The squared Grassmannian* $\mathrm{sGr}(d,n)$ *satisfies*

$$\dim(\mathrm{sGr}(k,n)) = k(n-k) \quad \text{and} \quad \mathrm{degree}(\mathrm{sGr}(k,n)) = 2^{(k-1)(n-k-1)} \cdot \mathrm{degree}(\mathrm{Gr}(k,n)).$$

Proof. The self-map of $\mathbb{P}^m$ given by squaring all coordinates is 2^m-to-one. It preserves the dimension of any subvariety, because all of its fibers are finite. This yields the first equation.

For the second equation, we factor the squaring morphism $\mathrm{Gr}(k,n) \to \mathrm{sGr}(k,n)$ through the quadratic Veronese map, which is an isomorphism from $\mathrm{Gr}(k,n)$ to its Veronese square $\nu_2(\mathrm{Gr}(k,n))$. Note that the degree of $\nu_2(\mathrm{Gr}(k,n))$ is $2^{k(n-k)} \cdot \mathrm{degree}(\mathrm{Gr}(k,n))$. The projection from $\nu_2(\mathrm{Gr}(k,n))$ onto $\mathrm{sGr}(k,n)$ deletes all mixed coordinates. This map is 2^{n-1}-to-one since each fiber is given by switching the signs independently in the n columns of the matrix A. So, the degree of $\mathrm{sGr}(k,n)$ is the degree of $\nu_2(\mathrm{Gr}(k,n))$ divided by 2^{n-1}. ∎

The familiar formula [97, Theorem 5.13] for the degree of the Grassmannian implies:

Corollary 8.2.2. *An explicit formula for the degree of the squared Grassmannian is*

$$\text{degree}(\text{sGr}(k,n)) \;=\; \frac{2^{(k-1)(n-k-1)}\cdot(k(n-k))!}{\prod_{j=1}^{d} j(j+1)(j+2)\cdots(j+n-k-1)}.$$

For the special case $k=2$, these degrees are scalings of the Catalan numbers:

$$\text{degree}(\text{sGr}(2,n)) \;=\; \tfrac{2^{n-3}}{n-1}\tbinom{2n-4}{n-2} \;=\; 2^{n-3}\cdot C_{n-2}. \tag{8.3}$$

We now turn to the ideal of the squared Grassmannian, beginning with the case $d=2$. The squared Plücker coordinates are denoted by $q_{ij}=x_{ij}^2$ for $1\le i<j\le n$. We write these coordinates as the entries of a symmetric $n\times n$ matrix that has zeros on the main diagonal.

Theorem 8.2.3. *The prime ideal $I_{\text{sGr}(2,n)}$ is generated by the 4×4 minors of the matrix*

$$Q \;=\; \begin{bmatrix} 0 & q_{12} & q_{13} & \cdots & q_{1n}\\ q_{12} & 0 & q_{23} & \cdots & q_{2n}\\ q_{13} & q_{23} & 0 & \cdots & q_{3n}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ q_{1n} & q_{2n} & q_{3n} & \cdots & 0 \end{bmatrix}.$$

We thank Aldo Conca for suggesting the following proof.

Proof. Let Q be a generic symmetric $n\times n$ matrix. Let J be the ideal generated by the 4×4 minors of Q, and let I be the ideal $\langle q_{11},\ldots,q_{nn}\rangle$, in the polynomial ring $\mathbb{C}[Q]$ with complex coefficients. Then $V(I+J)=V(I)\cap V(J)$ is the variety of $n\times n$ matrices of rank at most 3 with zeros on the diagonal. We will show that $I+J$ is prime, that its projective variety contains $\text{sGr}(2,n)$, and that the projective varieties have the same dimension.

We first make some observations about the ring $\mathbb{C}[Q]/J$. We will show that $\mathbb{C}[Q]/J\cong\mathbb{C}[Y^TY]$ where Y is a generic $3\times n$ matrix. This isomorphism is the map $[q_{ij}]\mapsto(Y^TY)_{ij}$, where $[q_{ij}]$ is the image of q_{ij} under the quotient. The fact that $\mathbb{C}[Y^TY]$ is the image of this map is then clear. By the first isomorphism theorem for rings, it is then sufficient to show that J is equal to the kernel of this map. The containment of J within the kernel is relatively straightforward: the image of a 4×4 minor of Q in $\mathbb{C}[Y^TY]$ is a 4×4 minor of Y^TY, which is zero as Y^TY has rank at most 3. The containment of the kernel into J is the Second Fundamental Theorem of Invariant Theory for the orthogonal group; see [32, Theorem 5.7].

Now given some $3\times n$ matrix Y, the product Y^TY is invariant under the action of the orthogonal group $\mathrm{O}(3)$ on Y by matrix multiplication on the left. Thus we have

$$\mathbb{C}[Q]/J \;\cong\; \mathbb{C}[Y^TY] \;=\; \mathbb{C}[Y]^{\mathrm{O}(3)} \;\subset\; \mathbb{C}[Y],$$

where $\mathbb{C}[Y]^{\mathrm{O}(3)}$ is the invariant ring of the action of $\mathrm{O}(3)$ on Y by left multiplication. We have justified the inclusion $\mathbb{C}[Y^TY]\subseteq\mathbb{C}[Y]^{\mathrm{O}(3)}$. The other inclusion is given by the First Fundamental Theorem of Invariant Theory for the orthogonal group; see [32, Theorem 5.6].

As O(3) is linearly reductive in characteristic zero, $\mathbb{C}[Y]^{\mathrm{O}(3)} = \mathbb{C}[Y^TY]$ is a direct summand of $\mathbb{C}[Y]$. In summary, $\mathbb{C}[Q]/J$ is isomorphic to $\mathbb{C}[Y^TY]$, which is a direct summand of $\mathbb{C}[Y]$.

We now turn to the image of $(I+J)/J$ under this isomorphism. Recall that $\{[q_{11}], \dots, [q_{nn}]\}$ generates $(I+J)/J$. The image of $(I+J)/J$ in $\mathbb{C}[Y^TY]$ is $I_{Y^TY} = \langle (Y^TY)_{11}, \dots, (Y^TY)_{nn}\rangle \subseteq \mathbb{C}[Y^TY]$. We now turn to the extension of this ideal to $\mathbb{C}[Y]$, which we denote $I_{Y^TY}\mathbb{C}[Y] = \langle (Y^TY)_{11}, \dots, (Y^TY)_{nn}\rangle \subseteq \mathbb{C}[Y]$. For all i, the quadric $(Y^TY)_{ii} = y_{1i}^2 + y_{2i}^2 + y_{3i}^2$ is irreducible. As the $(Y^TY)_{ii}$ use disjoint sets of variables, we conclude that $I_{Y^TY}\mathbb{C}[Y]$ is prime.

We now show I_{Y^TY} is prime. Since $\mathbb{C}[Y^TY]$ is a direct summand of $\mathbb{C}[Y]$, we have $I_{Y^TY} = I_{Y^TY}\mathbb{C}[Y] \cap \mathbb{C}[Y^TY]$. The inclusion $\subseteq$ holds, because $I_{Y^TY}\mathbb{C}[Y] \cap \mathbb{C}[Y^TY]$ is the contraction of the extension of I_{Y^TY}. For the inclusion $\supseteq$, suppose $a = \sum_{i=1}^m a_i b_i \in I_{Y^TY}\mathbb{C}[Y] \cap \mathbb{C}[Y^TY]$ where $b_i \in \mathbb{C}[Y]$ and $a_i \in I_{Y^TY}$. The $\mathbb{C}[Y^TY]$-module homomorphisms $\mathbb{C}[Y^TY] \hookrightarrow \mathbb{C}[Y] \xrightarrow{\pi} \mathbb{C}[Y^TY]$ compose to the identity, so $a = \pi(a) = \sum_{i=1}^m a_i \pi(b_i) \in I_{Y^TY}$. This proves the other containment. To conclude that I_{Y^TY} is prime, recall that I_{Y^TY} is the inverse image of $I_{Y^TY}\mathbb{C}[Y]$ under the map $\mathbb{C}[Y^TY] \hookrightarrow \mathbb{C}[Y]$, and that inverse images of primes are prime.

We now show that $V(I+J)$ has the correct dimension. The dimension of the projective variety $V(J)$ is $3n-4$ because any symmetric matrix of rank ≤ 3 is determined by the upper triangle of the upper left 3×3 block and the $3(n-3)$ entries in the upper right corner. To find the dimension of $V(I+J)$, we note that $[q_{11}], \dots, [q_{nn}]$ is a regular sequence in the ring $\mathbb{C}[Q]/J$. One proves this by applying the arguments above to show that $\langle [q_{11}], \dots, [q_{ii}]\rangle$ is prime for any i. Therefore, the dimension of the projective variety $V(I+J)$ is $(3n-4)-n = 2n-4$. In particular, $V(I+J)$ has the same dimension as the Grassmannian $\mathrm{Gr}(2,n)$.

We now argue that the squared Grassmannian $\mathrm{sGr}(2,n)$ is contained in $V(I+J)$. For any point q in $\mathrm{sGr}(2,n)$, there is a $2\times n$ matrix $A = (a_{kl})$ such that $q_{ij} = a_{1i}^2 a_{2j}^2 - 2a_{1i}a_{2i}a_{1j}a_{2j} + a_{1j}^2 a_{2i}^2$ for $1 \leq i < j \leq n$. Hence, Q has zeros on the diagonal and is a sum of three matrices of rank one, meaning that its 4×4 minors vanish. This shows that $q \in V(I+J)$.

Because $V(I+J)$ has the same dimension as the squared Grassmannian, $\mathrm{sGr}(2,n) \subseteq V(I+J)$, and both varieties are irreducible, it follows that $\mathrm{sGr}(2,n) = V(I+J)$. Finally, because $I+J$ is prime, we may apply the Nullstellensatz to see that $I_{\mathrm{sGr}(2,n)} = I+J$. $\square$

Example 8.2.4 ($d = 2, n \leq 5$)**.** The Grassmannian $\mathrm{Gr}(2,4)$ is defined in $\mathbb{P}^5$ by the quadric $x_{12}x_{34} - x_{13}x_{24} + x_{14}x_{23}$. By setting $q_{ij} = x_{ij}^2$ and eliminating the x-variables, we obtain

$$q_{12}^2 q_{34}^2 + q_{13}^2 q_{24}^2 + q_{14}^2 q_{23}^2 - 2q_{12}q_{13}q_{24}q_{34} - 2q_{12}q_{14}q_{23}q_{34} - 2q_{13}q_{14}q_{23}q_{24}.$$

This quartic is the determinant of a symmetric 4×4 matrix (q_{ij}) with zeros on the diagonal.

The squared Grassmannian $\mathrm{sGr}(2,5)$ has dimension 6 in $\mathbb{P}^9$. Its degree is 20, by (8.3). The ideal of $\mathrm{sGr}(2,5)$ is generated by 15 quartics, namely the 4×4 minors in Theorem 8.2.3. $\diamond$

Conjecture 8.2.5. For all $n \geq k \geq 2$, the prime ideal $I_{\mathrm{sGr}(k,n)}$ is generated by quartics.

We verified this conjecture for $k = 3, n = 6$ using `Macaulay2` [60]. The variety $\mathrm{sGr}(3,6)$ has dimension 9 and degree 672 in $\mathbb{P}^{19}$. Its ideal is minimally generated by 285 quartics.

It is proved in [2, Section 6] that $\mathrm{sGr}(k,n)$ is cut out by quartic polynomials that are derived from the $2\times 2\times 2$ hyperdeterminant. We state their result as follows.

Theorem 8.2.6 ([2, Corollary 6.4])**.** *The set-theoretic version of Conjecture 8.2.5 is true.*

8.3 Parametric Model

We now turn to the likelihood geometry of the squared Grassmannian. Our approach to computing the ML degree of $\mathrm{sGr}(2,n)$ is topological, namely, we will apply Theorem 2.3.5. As in Chapter 2, we write $\mathcal{H} = \{(q_I) \in \mathbb{P}^{\binom{n}{k}} : (\sum_I q_I) \cdot \prod_I q_I = 0\}$ for the hyperplane arrangement where the log-likelihood function is undefined. Theorem 2.3.5 states that the ML degree of a model $\mathcal{M} \subseteq \mathbb{P}^n$, is equal to the signed Euler characteristic of the very affine variety $\mathcal{M} \backslash \mathcal{H} \subseteq (\mathbb{C}^*)^{\binom{n}{k}}$.

The Euler characteristic is a topological invariant that can be defined as the alternating sum of the Betti numbers of the space. We will often make use of the excision property of the Euler characteristic: if $\mathcal{Z} = \mathcal{X} \sqcup \mathcal{Y}$, then $\chi(\mathcal{Z}) = \chi(\mathcal{X}) + \chi(\mathcal{Y})$. We also rely on the fibration property of the Euler characteristic; if $\phi : \mathcal{X} \to \mathcal{Y}$ is a fibration, then $\chi(\mathcal{X}) = \chi(\mathcal{Y})\chi(\mathcal{F})$, where $\mathcal{F}$ is the fiber of a point in $\mathcal{Y}$. For the maps we consider, proving that the fibers of a map are homeomorphic suffices to show that the map is a fibration.

We apply Theorem 2.3.5 to the squared Grassmannian. The *open squared Grassmannian* is the very affine variety

$$\mathrm{sGr}(k,n)^\circ = \Big\{(q_I)_{I \in \binom{[n]}{k}} \in \mathrm{sGr}(k,n) : \Big(\sum_{I \in \binom{[n]}{k}} q_I\Big) \cdot \prod_{I \in \binom{[n]}{k}} q_I \neq 0 \Big\} \subseteq (\mathbb{C}^*)^{\binom{n}{k}}.$$

The variety $\mathrm{sGr}(k,n)^\circ$ is smooth because $\mathrm{Gr}(k,n)$ is smooth and the Jacobian of the squaring map has full rank on points in the preimage of $\mathrm{sGr}(k,n)^\circ$. Because $\dim(\mathrm{sGr}(k,n)) = k(n-k)$ by Proposition 8.2.1, the ML degree of $\mathrm{sGr}(k,n)$ is $(-1)^{k(n-k)}\chi(\mathrm{sGr}(k,n)^\circ)$ by Theorem 2.3.5.

We now turn to the case $k = 2$ and parameterize the open squared Grassmannian by

$$\mathcal{X}_n \to \mathrm{sGr}(2,n)^\circ, \quad M_n = \begin{pmatrix} 1 & 0 & t_{13} & t_{14} & \cdots & t_{1n} \\ 0 & 1 & t_{23} & t_{24} & \cdots & t_{2n} \end{pmatrix} \mapsto (x_{ij}^2)_{1 \le i < j \le n} \tag{8.4}$$

where x_{ij} is the 2×2 minor formed from columns i and j of M_n and X_n is the subset of $\mathbb{C}^{2(n-2)}$ such that the image of the parametrization is $\mathrm{sGr}(2,n)^\circ$. To be specific, we define

$$Q_n = \sum_{1 \le i < j \le n} x_{ij}^2 \qquad \text{and} \qquad \mathcal{X}_n = \Big\{M_n \in \mathbb{C}^{2(n-2)} : Q_n \cdot \Big(\prod_{1 \le i < j \le n} x_{ij}\Big) \neq 0\Big\}.$$

We can explicitly write the parametric log-likelihood function in the $k = 2$ case as

$$L_u(M_n) = \sum_{i=3}^{n} \big(u_{1i}\log(t_{2i}^2) + u_{2i}\log(t_{1i}^2)\big) + \sum_{3 \le i < j \le n} u_{ij}\log((t_{1i}t_{2j} - t_{2i}t_{1j})^2) - \sum_{1 \le i < j \le n} u_{ij}\log(Q_n). \tag{8.5}$$

The map in (8.4) has degree 2^{n-1} with no ramification points because we can flip the signs of any row and column of the matrix M_n independently; see [34, Remark 4.4]. By multiplicativity of the Euler characteristic, the implicit and parametric models are related by:

$$\chi(\mathcal{X}_n)=2^{n-1}\chi(\mathrm{sGr}(2,n)^\circ) = 2^{n-1}\mathrm{MLdegree}(\mathrm{sGr}(2,n)) = \#\{\text{critical points of (8.5)}\}. \quad (8.6)$$

By (8.6), we can compute the number of critical points of both the implicit and parametric log-likelihood functions from $\chi(\mathcal{X}_n)$. We compute $\chi(\mathcal{X}_n)$ in Section 8.4 by an argument similar to those in [1, 28, 41]. In particular, we wish to define a map $\mathcal{X}_{n+1} \to \mathcal{X}_n$ which deletes the last column and use the fact that this map is a stratified fibration to compute the Euler characteristic inductively. However, this deletion map is not well defined, because Q_n vanishes on points in $\mathcal{X}_{n+1}$, so we define this map only on the open subset of $\mathcal{X}_{n+1}$ where Q_n does not vanish, denoted $\mathcal{X}_{n+1}^\circ = \{M_{n+1} \in \mathcal{X}_{n+1} \colon Q_n \neq 0\}$. We define the projection

$$\pi_{n+1} : \mathcal{X}_{n+1}^\circ \to \mathcal{X}_n, \qquad \begin{pmatrix} 1 & 0 & t_{13} & t_{14} & \cdots & t_{1n} & t_{1,n+1} \\ 0 & 1 & t_{23} & t_{24} & \cdots & t_{2n} & t_{2,n+1} \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & t_{13} & t_{14} & \cdots & t_{1n} \\ 0 & 1 & t_{23} & t_{24} & \cdots & t_{2n} \end{pmatrix}$$

and argue that the change from $\mathcal{X}_{n+1}$ to $\mathcal{X}_{n+1}^\circ$ does not affect the Euler characteristic. To do so, we observe that the fiber of a point $M_n \in \mathcal{X}_n$ under π_{n+1} is the complement in $(\mathbb{C}^*)^2$ of the lines $x_{i(n+1)}$ and conic Q_{n+1}. The lines $x_{i(n+1)}$ all pass through the origin, so the key to understanding the geometry of the fibers is to understand the conic Q_{n+1}. We can write

$$Q_{n+1} = \begin{pmatrix} 1 & t_{1,n+1} & t_{2,n+1} \end{pmatrix} \begin{pmatrix} Q_n & 0 & 0 \\ 0 & 1+\sum_{i=3}^n t_{2,i}^2 & -\sum_{i=3}^n t_{1i}t_{2i} \\ 0 & -\sum_{i=3}^n t_{1i}t_{2i} & 1+\sum_{i=3}^n t_{1i}^2 \end{pmatrix} \begin{pmatrix} 1 \\ t_{1,n+1} \\ t_{2,n+1} \end{pmatrix} \quad (8.7)$$

Taking the determinant of the matrix, we find that the discriminant of Q_{n+1} as a conic in $t_{1,n+1}$ and $t_{2,n+1}$ is Q_n^2. We prove that restricting to $\mathcal{X}_{n+1}^\circ$ does not change the Euler characteristic.

Lemma 8.3.1. *For $n \geq 3$, we have $\chi(X_{n+1}) = \chi(\mathcal{X}_{n+1}^\circ)$.*

Proof. By the excision property, $\chi(\mathcal{X}_{n+1}) = \chi(\mathcal{X}_{n+1}^\circ) + \chi(\mathcal{X}_{n+1} \backslash \mathcal{X}_{n+1}^\circ)$. We shall prove that $\chi(\mathcal{X}_{n+1} \backslash \mathcal{X}_{n+1}^\circ) = 0$ by showing that the map below is a fibration on its image:

$$\mathcal{X}_{n+1} \backslash \mathcal{X}_{n+1}^\circ \to (\mathbb{C}^*)^{2(n-2)}, \quad \begin{pmatrix} 1 & 0 & t_{13} & t_{14} & \cdots & t_{1n} & t_{1,n+1} \\ 0 & 1 & t_{23} & t_{24} & \cdots & t_{2n} & t_{2,n+1} \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & t_{13} & t_{14} & \cdots & t_{1n} \\ 0 & 1 & t_{23} & t_{24} & \cdots & t_{2n} \end{pmatrix}$$

It suffices to show that the fibers are homeomorphic. Every nonempty fiber $\mathcal{F}$ is the complement in $(\mathbb{C}^*)^2$ of $n-1$ lines through the origin: each minor $x_{i,n+1}$ for $i \geq 3$ contributes one line and Q_{n+1} degenerates into a double line since $Q_n = 0$. The Euler characteristic of finitely many lines through the origin in $(\mathbb{C}^*)^2$ is zero because each line is a $\mathbb{P}^1$ with two points removed and the lines do not intersect. Because $\chi((\mathbb{C}^*)^2) = 0$, the Euler characteristic of the complement of $n-1$ lines through the origin in $(\mathbb{C}^*)^2$ is zero, too. By the fibration property of the Euler characteristic and because $\chi(\mathcal{F}) = 0$, we conclude that $\chi(\mathcal{X}_{n+1} \backslash \mathcal{X}_{n+1}^\circ) = 0$. $\square$

8.4 Stratification of the Parametric Model

We now prove that the map π_{n+1} is a stratified fibration and compute $\chi(\mathcal{X}_n)$. A map $\phi : \mathcal{X} \to \mathcal{Y}$ between complex algebraic varieties with a stratification $\{\mathcal{S}_\alpha\}_{\alpha=1}^m$ of $\mathcal{Y}$ by closed sets is a *stratified fibration* if the restriction of ϕ to each open stratum $\mathcal{S}_\alpha^\circ = \mathcal{S}_\alpha \setminus \bigcup_{\mathcal{S}_\beta \subsetneq \mathcal{S}_\alpha} \mathcal{S}_\beta$ is a fibration with fiber denoted $\mathcal{F}_{\mathcal{S}_\alpha}$. We use Möbius inversion with the fibration property of the Euler characteristic, to compute Euler characteristics along stratified fibrations.

Lemma 8.4.1 ([1], Lemma 2.3)**.** *Let $\phi : \mathcal{X} \to \mathcal{Y}$ be a stratified fibration with stratification $\{\mathcal{S}_\alpha\}_{\alpha=1}^m$. Write $\mathcal{F}_{\mathcal{S}_\alpha}$ for the fiber of a generic point in $\mathcal{S}_\alpha$, and let μ be the Möbius function of the poset $\{\mathcal{S}_\alpha\}_{\alpha=1}^m$ ordered by inclusion. Then*

$$\chi(\mathcal{X}) = \chi(\mathcal{Y}) \cdot \chi(\mathcal{F}_\mathcal{Y}) + \sum_{\alpha=1}^m \chi(\mathcal{S}_\alpha) \cdot \sum_{\mathcal{S}_\beta \supseteq \mathcal{S}_\alpha} \mu(\mathcal{S}_\alpha, \mathcal{S}_\beta) \cdot (\chi(\mathcal{F}_{\mathcal{S}_\beta}) - \chi(\mathcal{F}_\mathcal{Y})).$$

In our situation, only the generic fiber contributes: the sums vanish and $\chi(\mathcal{X}) = \chi(\mathcal{Y})\, \chi(\mathcal{F}_\mathcal{Y})$. We first describe the stratification of $\mathcal{X}_n$ and the fibers of π_{n+1}.

Lemma 8.4.2. *The map π_{n+1} is a stratified fibration with stratification*

$$\{\mathcal{X}_n\} \cup \{\mathcal{S}_i : 1 \le i \le n\} \cup \{\mathcal{S}_i \cap \mathcal{S}_j : 1 \le i < j \le n\},$$

where $\mathcal{S}_i = \left\{ M_n \in \mathcal{X}_n : \sum_{j=1}^n x_{ij}^2 = 0 \right\}$. The fibers of π_{n+1} are the complements of n lines $p_{i(n+1)}$ and a conic Q_{n+1} in $\mathbb{C}^2$ where the intersections of the lines with the conic are

$$\begin{aligned}
Q_{n+1} \cap x_{\ell,n+1} &= \textit{two points for all } \ell && \textit{in } \mathcal{F}_{\mathcal{X}_n}.\\
Q_{n+1} \cap x_{\ell,n+1} &= \begin{cases} \textit{two points for } \ell \neq i \\ \emptyset \textit{ for } \ell = i. \end{cases} && \textit{in } \mathcal{F}_{\mathcal{S}_i}.\\
Q_{n+1} \cap x_{\ell,n+1} &= \begin{cases} \textit{two points for } \ell \neq i, j \\ \emptyset \textit{ for } \ell \in \{i, j\} \end{cases} && \textit{in } \mathcal{F}_{\mathcal{S}_i \cap \mathcal{S}_j}.
\end{aligned}$$

Proof. To prove π_{n+1} is a fibration, it suffices to show that the fibers of π_{n+1} are homeomorphic on each open stratum $\mathcal{S}^\circ$. The fiber in $\mathbb{C}^2$ of a point $M_n \in \mathcal{X}_n$ is the complement of the lines $x_{i,n+1}$ and the conic Q_{n+1}. Since all the lines $x_{i,n+1}$ pass through the origin, their intersection is the same for any $M_n \in \mathcal{X}_n$. The variation in the fibers comes from the intersection with the conic. We claim that every line $x_{i,n+1}$ either intersects the conic Q_{n+1} in two distinct points or not at all. In the intersection locus of $x_{i,n+1} = 0$ and $Q_{n+1} = 0$,

$$Q_{n+1} = Q_n + \sum_{j=1}^n t_{j,n+1}^2 = Q_n + \frac{t_{1,n+1}^2}{t_{1i}^2} \sum_{j=1}^n t_{ij}^2 = 0. \tag{8.8}$$

If $\sum_{j=1}^n x_{ij}^2 \neq 0$, then there are two distinct intersection points. If $\sum_{j=1}^n x_{ij}^2 = 0$, then (8.8) becomes $Q_n = 0$, which is a contradiction, so the intersection of Q_{n+1} and $x_{i,n+1}$ is empty.

Thus, in the fiber $\mathcal{F}_{\mathcal{X}_n}$, every line intersects the conic in two points and in the fiber $\mathcal{F}_{\mathcal{S}_i}$, every line intersects the conic in two points, except $x_{i,n+1}$ which does not intersect the conic.

It remains to be shown that there can be at most two lines $x_{i,n+1}$ which do not intersect Q_{n+1}. When $x_{i,n+1}$ does not intersect Q_{n+1} in $\mathbb{C}^2$, the line and conic intersect at a double point at infinity. In other words, $x_{i,n+1}$ is tangent to the projectivization of Q_{n+1} at infinity. In projective space, $p_{i(n+1)}$ intersects Q_{n+1} tangentially if and only if $x^\vee_{i,n+1}$ lies in the dual conic $Q^\vee_{n+1}$, i.e., $\begin{pmatrix}0 & -t_{2i} & t_{1i}\end{pmatrix} A_n^{-1} \begin{pmatrix}0 & -t_{2i} & t_{1i}\end{pmatrix}^T = \sum_{j=1}^n x_{ij}^2 = 0$, where A_n is the matrix in (8.7). Since at most two lines passing through the origin can be tangent to a conic, only two of the lines $x_{i,n+1}$ can have an empty intersection with the conic Q_{n+1} at once. Thus, in the fiber $\mathcal{F}_{\mathcal{S}_i \cap \mathcal{S}_j}$, every line intersects the conic in two points, except $x_{i,n+1}$ and $x_{j,n+1}$ which do not intersect the conic. □

The combinatorics of the stratification is simple, allowing us to evaluate the formula for the Euler characteristic in Lemma 8.4.1. To apply the lemma, we need the Möbius function of the poset of each closed stratum; see Figure 8.1 for the posets of each of the special strata along with the values of their Möbius functions. The remaining computations necessary to prove Theorem 8.1.4 are Euler characteristics of the fibers and of the strata.

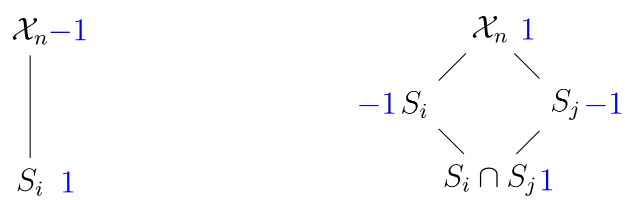

Figure 8.1: Posets for the special strata. The values of the Möbius function $\mu(-, S_i)$ and $\mu(-, S_i \cap S_j)$, respectively, are shown in blue.

Lemma 8.4.3 (Euler characteristics of fibers)**.** *The Euler characteristics of the fibers are*

$$\chi(\mathcal{F}_{\mathcal{X}_n}) = 2n \qquad \chi(\mathcal{F}_{\mathcal{S}_i}) = 2n-2 \qquad \chi(\mathcal{F}_{\mathcal{S}_i \cap \mathcal{S}_j}) = 2n-4 \qquad \text{for } 1 \le i < j \le n.$$

Proof. In $\mathbb{C}^2 \backslash \{(0,0)\}$, we have $\chi(Q_{n+1}) = 0$, since the conic Q_{n+1} has nonzero constant term Q_n and is therefore topologically equivalent to a $\mathbb{P}^1$ with two points at infinity removed. Similarly, in $\mathbb{C}^2 \backslash \{(0,0)\}$, we have $\chi(p_{i(n+1)}) = 0$, since the line $x_{i,n+1}$ is topologically equivalent to a $\mathbb{P}^1$ with the origin and a point at infinity removed. Suppose that $M_{n+1} \in \mathcal{X}_{n+1}$. Because $\chi(\mathbb{C}^2 \backslash \{(0,0)\}) = 0$, by the excision property,

$$\begin{aligned}\chi(\pi_{n+1}^{-1}(M_{n+1})) &= -\chi(Q_{n+1}) - \sum_{i=1}^n \chi(x_{i,n+1}) + \sum_{i=1}^n \chi\left(x_{i,n+1} \cap Q_{n+1}\right) \\ &= \sum_{i=1}^n \chi(x_{i,n+1} \cap Q_{n+1}) = \bigcup_{i=1}^n \left(x_{i,n+1} \cap Q_{n+1}\right) .\end{aligned}$$

Here, the last equality holds because the intersection of a line with a conic is zero dimensional. We computed these intersections in Lemma 8.4.2. In the generic fiber, each of the n lines intersects the conic in two points, for a total of $2n$ intersection points. The intersections for $\mathcal{S}_i$ and $\mathcal{S}_i \cap \mathcal{S}_j$ are the same, but with $n-1$ and $n-2$ lines, respectively. □

We now show that the Euler characteristics of the closed, codimension one strata vanish.

Lemma 8.4.4 (Euler characteristics of strata)**.** *We have $\chi(\mathcal{S}_i) = 0$ for all i.*

Proof. It is sufficient to prove the claim for $\mathcal{S}_n$ by symmetry. We first remark that $\mathcal{S}_n \subset \mathcal{X}_n^\circ$. We argue that $\mathcal{S}_n$ has Euler characteristic zero by showing that $\pi_n|_{\mathcal{S}_n}$ is a fibration whose fiber is the union of two lines through the origin in $(\mathbb{C}^*)^2$. Since the fiber has Euler characteristic zero, the result follows from the fibration property.

Because $\sum_{j=1}^{n-1} x_{jn}^2$ is a homogeneous polynomial in t_{1n} and t_{2n}, the fiber of a point in $\pi_n(\mathcal{S}_n)$ is a degenerate conic in the variables t_{1n}, t_{2n}, so each fiber is the union of two lines through the origin in the complement of the lines x_{jn} and conic Q_n in $(\mathbb{C}^*)^2$. Since the x_{jn} all pass through the origin, they do not intersect the degenerate conic in $(\mathbb{C}^*)^2$. The conic Q_n also does not intersect $\sum_{j=1}^{n-1} x_{jn}^2$, because $Q_n - \sum_{j=1}^{n-1} x_{jn}^2 = Q_{n-1} = 0$ on the intersection, contradicting $\mathcal{S}_n \subset \mathcal{X}_n^\circ$. Since the conic $\sum_{j=1}^{n-1} x_{jn}^2$ does not intersect the lines x_{jn} or the conic Q_n, the fiber is the union of two lines through the origin in $(\mathbb{C}^*)^2$. □

We do not need to compute $\chi(\mathcal{S}_i \cap \mathcal{S}_j)$, because the inside sum in Lemma 8.4.1 vanishes for $\mathcal{S}_i \cap \mathcal{S}_j$. In fact, we can count the number of critical points of the parametric log-likelihood function. Theorem 8.1.4 then follows from Proposition 8.4.5 and (8.6).

Proposition 8.4.5. *The parametric log-likelihood function* (8.5) *has* $2^{n-2}(n-1)!$ *critical points.*

Proof. By (8.6), the number of critical points is equal to the Euler characteristic $\chi(\mathcal{X}_n)$, so it suffices to show that $\chi(\mathcal{X}_n) = 2^{n-2}(n-1)!$. We proceed by induction on n. The base case is furnished by Example 8.1.6. By Lemmas 8.3.1, 8.4.1, and 8.4.2, we have

$$\begin{aligned}\chi(\mathcal{X}_{n+1}) = \chi(\mathcal{X}_{n+1}^\circ) = \chi(\mathcal{F}_{\mathcal{X}_n})\chi(\mathcal{X}_n) &+ \sum_{i=1}^{n} \chi(\mathcal{S}_i) \sum_{\mathcal{S} \in \{\mathcal{S}_i, \mathcal{X}_n\}} \mu(\mathcal{S}_i, \mathcal{S})(\chi(\mathcal{F}_{\mathcal{S}}) - \chi(\mathcal{F}_{\mathcal{X}_n})) \\ &+ \sum_{1 \le i < j \le n} \chi(\mathcal{S}_i \cap \mathcal{S}_j) \sum_{\mathcal{S} \in \{\mathcal{S}_i \cap \mathcal{S}_j, \mathcal{S}_i, \mathcal{S}_j, \mathcal{X}_n\}} \mu(\mathcal{S}_i \cap \mathcal{S}_j, \mathcal{S})(\chi(\mathcal{F}_{\mathcal{S}}) - \chi(\mathcal{F}_{X_n}))\end{aligned}$$

By Lemma 8.4.4, $\chi(\mathcal{S}_i) = 0$, so the first sum is zero. By Lemma 8.4.3 and the values of the Möbius function in Figure 8.1, the second sum is zero:

$$\sum_{\mathcal{S} \in \{\mathcal{S}_i \cap \mathcal{S}_j, \mathcal{S}_i, \mathcal{S}_j, \mathcal{X}_n\}} \mu(\mathcal{S}_i \cap \mathcal{S}_j, \mathcal{S})(\chi(\mathcal{F}_{\mathcal{S}}) - \chi(\mathcal{F}_{\mathcal{X}_n})) = (1)(-4) + (-1)(-2) + (-1)(-2) = 0.$$

By induction and Lemma 8.4.3, $\chi(\mathcal{X}_{n+1}) = \chi(\mathcal{F}_{\mathcal{X}_n})\chi(\mathcal{X}_n) = 2n(2^{n-2}(n-1)!) = 2^{n-1}n!$. □

A formula for the ML degree of $\mathrm{sGr}(k,n)$ for $k > 2$ has not yet been conjectured. In our calculation of the ML degree of $\mathrm{sGr}(2,n)$, we relied heavily on the nice geometry of conics in the complex projective plane, which we do not have for higher k. There are some computations for $k = 3$ in Table 8.1, but the numbers do not follow an obvious pattern, so it is likely that the combinatorics of the stratification is complicated and that we do not have the nice cancellation we get in the proof of Proposition 8.4.5.

8.5 Real Solutions

From a statistical perspective, the only relevant critical points of the implicit log-likelihood function (7.1) are real, nonnegative ones. Because each q_I is a square, this condition is equivalent to the critical points of the parametric log-likelihood function (8.2) being real. We will give a lower bound for all k, n on the number of local maxima of the parametric log-likelihood function (8.2). Furthermore, we prove that when $k = 2$, all critical points of the parametric log-likelihood function (8.5) are real and local maxima. This is not true for larger k.

Example 8.5.1 ($k = 3, n = 6$)**.** This parametric log-likelihood function generically has 17 664 critical points. For 100 trials with data vectors whose entries were sampled uniformly at random from [1000], each parametric log-likelihood function had 11 904 real critical points, all of which were local maxima. These computations suggest that the parametric log-likelihood function with $k = 3, n = 6$ generically has 11 904 real solutions. These computations were performed using the `Julia` package `HomotopyContinuation.jl` [20]. ◇

Let $M_{k,n}$ be a $k \times n$ matrix whose first $k \times k$ block is the identity and let x_I be the maximal minor whose columns are indexed by I. For real solutions we turn from the log-likelihood function to the likelihood function

$$\frac{\prod_{I \in \binom{[n]}{k}} x_I^{2u_I}}{\left(\sum_{I \in \binom{[n]}{k}} x_I^2\right)^{\sum_I u_I}}, \tag{8.9}$$

which shares its critical points with the log-likelihood function. We optimize (8.9) on the real open set $\mathcal{X}_{k,n}^{\mathbb{R}} = \mathcal{X}_{k,n} \cap \mathbb{R}^{k(n-k)}$ where

$$Q_{k,n} = \sum_{I \in \binom{[n]}{k}} x_I^2 \qquad \text{and} \qquad \mathcal{X}_{k,n} = \left\{ M_{k,n} \in \mathbb{C}^{k(n-k)} : Q_{k,n} \cdot \Big(\prod_{I \in \binom{[n]}{k}} x_I \Big) \neq 0 \right\}.$$

Since the quadric $V(Q_{k,n})$ has no real solutions, $\mathcal{X}_{k,n}^{\mathbb{R}}$ is the complement of the real algebraic variety $\bigcup_{I \in \binom{[n]}{k}} V_{\mathbb{R}}(x_I)$. The irreducible hypersurfaces $V_{\mathbb{R}}(x_I)$ all pass through the origin and divide $\mathcal{X}_{k,n}^{\mathbb{R}}$ into unbounded regions. It is known that every bounded region contains at least one critical point, because the function is either positive or negative on the region and must therefore achieve a local minimum or maximum; see [27, Proposition 10]. The following result extends this idea to unbounded regions.

Lemma 8.5.2. *Let $f_1, \dots, f_n \in \mathbb{R}[x_1, \dots, x_s]$ such that $\deg(f_1) = 0$ and $f_+ = f_1 + \cdots + f_n$ is positive on $\mathbb{R}^s$. If $u_1, \dots, u_n \in \mathbb{N}_{>0}$ is generic, then*

$$\#\Big\{\textit{regions of } \mathbb{R}^s \backslash \bigcup_{i=2}^{n} \{f_i = 0\}\Big\} \leq \#\Big\{\textit{critical points of } \frac{f_1^{u_1} \cdots f_n^{u_n}}{(f_+)^{u_1 + \cdots + u_n}} \textit{ in } \mathbb{R}^s\Big\} \leq \mathrm{MLdegree}(\mathcal{M})$$

where $\mathcal{M} \subseteq \mathbb{P}^{n-1}$ is the Zariski closure of the image of the map $x \mapsto [f_1(x) : \cdots : f_n(x)]$. Further, $L = \frac{f_1^{u_1} \cdots f_n^{u_n}}{(f_1 + \cdots + f_n)^{u_1 + \cdots + u_n}}$ has a local maximum for every region of $\mathbb{R}^s \backslash \bigcup_{i=2}^n \{f_i = 0\}$ where $L > 0$ and a local minimum for every region where $L < 0$.

Proof. Since $f_1 + \cdots + f_n > 0$, the function L is smooth on $\mathbb{R}^s$. If $\deg(f_i) = d_i$ for $i = 1, \dots, n$, then numerator of L has degree $\sum_{i=2}^n d_i u_i$ and the denominator has degree $(\max_{i \in [n]} d_i) \sum_{i=1}^n u_i$. Because the denominator of L has larger degree than the numerator, L approaches 0 at infinity and is therefore bounded on all regions. Because L is smooth and bounded, it attains a local maximum or minimum on each region, depending on the sign of L. This proves the first inequality and second statement. The second inequality follows from the definition of ML degree. □

A similar technique was used to show that all critical points of a linear model are real. When Lemma 8.5.2 is applied to our problem, we get a lower bound on the number of real critical points and on the number of local maxima. This lower bound is given by the number of connected regions in the real open Grassmannian $\mathrm{Gr}_{\mathbb{R}}(k, n)^{\circ}$. Because the sign vectors of the Plücker coordinates are fixed on a given region, the number of possible sign vectors is a lower bound on the number of regions. We denote this set of sign vectors as $\mathrm{sgn}(\mathrm{Gr}_{\mathbb{R}}(k, n)^{\circ}) = \{(\mathrm{sgn}(x_I))_{I \in \binom{[n]}{k}} : x \in \mathrm{Gr}_{\mathbb{R}}(k, n)^{\circ}, x_{1 \cdots k} = 1\}$. This yields the following corollary:

Corollary 8.5.3. *For any k and n, the number of local maxima of (8.9) is bounded below by the number of possible sign vectors of Plücker coordinates:*

$$\begin{aligned} \#\mathrm{sgn}(\mathrm{Gr}_{\mathbb{R}}(k, n)^{\circ}) \leq & \#\{\textit{regions of } \mathcal{X}_{k,n}^{\mathbb{R}}\} \\ & \leq \#\{\textit{local maxima of } (8.9)\} \leq \#\{\textit{real critical points of } (8.9)\}. \end{aligned}$$

Proof. The result follows from Lemma 8.5.2 with $f_0 = x_{1 \cdots k}^2 = 1$ and the fact that (8.9) is nonnegative on $\mathbb{R}^{k(n-k)}$. □

When $k = 2$, the inequalities in Corollary 8.5.3 become equalities. We prove this by showing that the number of sign vectors in $\mathrm{sgn}(\mathrm{Gr}_{\mathbb{R}}(2, n)^{\circ})$ matches the number of critical points $2^{n-2}(n-1)!$ of the parametric log-likelihood function (8.5). Theorem 8.1.5 is an immediate corollary, since squares of nonzero real numbers are positive.

Theorem 8.5.4. *If $k = 2$, then every critical point of the parametric likelihood function (8.9) is real and a local maximum.*

Proof. We will prove that $\#\mathrm{sgn}(\mathrm{Gr}_{\mathbb{R}}(2,n)^{\circ}) = 2^{n-2}(n-1)!$. By Proposition 8.4.5, the total number of critical points is $2^{n-2}(n-1)!$, so the result follows from Corollary 8.5.3.

To count the sign vectors, we fix $2 \leq \ell \leq n$ and begin with a matrix

$$M_n = \begin{pmatrix} 1 & 0 & -t_{13} & \cdots & -t_{1\ell} & t_{1,\ell+1} & \cdots & t_{1n} \\ 0 & 1 & t_{23} & \cdots & t_{2\ell} & t_{2,\ell+1} & \cdots & t_{2n} \end{pmatrix}\Bigg($$

in $\mathcal{X}_{2,n}^{\mathbb{R}}$ such that $t_{13}, \ldots, t_{2n} > 0$. We take any permutation of the last $n-2$ columns and flip the sign of any of the last $n-2$ columns. This process yields $2^{n-2}(n-2)!$ different matrices. We argue that given a fixed M_n, each of these matrices has a distinct sign vector. Because every matrix in $\mathcal{X}_{2,n}^{\mathbb{R}}$ can arise in this way, there are $2^{n-2}(n-2)!$ possible sign vectors for a fixed $\ell \in \{2, \ldots, n\}$. Hence, in total, we have $2^{n-2}(n-1)!$ total possible sign vectors.

Let x be the Plücker vector of M_n and y be the Plücker vector of a matrix produced from M_n by the permutation and sign changes described above. We identify the permutation and columns whose signs were flipped. The columns i with $y_{1i} < 0$ had their signs flipped. Assuming that $y_{1i} > 0$ for all i, we uniquely identify the permutation from its inversions: for $i, j \geq 3$ the signs of x_{ij} and y_{ij} agree if and only if the pair ij is not an inversion. $\square$

The practical implication of this result is that likelihood inference for projection DPPs is difficult. In particular, for any data u, the number of local maximizers of both the implicit and parametric log-likelihood function grows exponentially. We will see that this holds for squared linear models as well in Theorem 9.1.1.

Example 8.5.1 shows that for larger k the critical points are not necessarily all real. However, in the $k = 3, n = 6$ case, the quantities in Corollary 8.5.3 are all equal to 11 904. We can prove this with the help of a computer: sample 1 million 3×3 matrices A and compute the Plücker coordinates of the 3×6 matrix $(\mathrm{Id}_3 \ A)$. The number of sign vectors we get from this process is a lower bound of 11 904 on the number of sign vectors that can arise in this way. By Example 8.5.1 all quantities in Corollary 8.5.3 are therefore equal to 11 904.

Conjecture 8.5.5. For $k \geq 3$, the last two inequalities in Corollary 8.5.3 are equalities, i.e.,

$$\#\{\text{regions of } X_{k,n}^{\mathbb{R}}\} = \#\{\text{local maxima of } (8.9)\} = \#\{\text{real critical points of } (8.9)\}.$$

Computational Experiments $(k = 2)$**.** To compute the true MLE for some data u, one needs to compute all critical points of the parametric log-likelihood function (8.5), evaluate (8.5) at the critical points, and select the one which yields the largest value. We give runtimes for computing the MLE of our model for data selected uniformly at random from [1000]. We use the numerical algebraic geometry software `HomotopyContinuation.jl` [20] to compute the critical points of (8.5). We use the strategy outlined in Section 7.3 to find solutions to the rational equations $\nabla L_u(M_n) = 0$. We first use the monodromy method to compute the solutions to $\nabla L_{u'}(M_n) = 0$ for some complex start parameters u'. We then use a coefficient parameter homotopy to move the start parameters u' to the target parameters $u \in [1000]^{\binom{n}{2}}$, simultaneously moving each solution of $\nabla L_{u'}(M_n) = 0$ to a solution of $\nabla L_u(M_n) = 0$.

	$n=4$	$n=5$	$n=6$	$n=7$	$n=8$	$n=9$
Runtime	180 ms	235 ms	962 ms	14.287 s	330 s	2 h
Number of Solutions	24	192	1 920	23 040	322 560	5 160 960

The runtime grows exponentially, as expected. The bulk of the time is spent computing the solutions to $\nabla L_{u'}(M_n) = 0$ with the monodromy method. As discussed in Section 2.5, knowing the number of solutions a priori makes the computation significantly faster than if we did not know the number of solutions.

8.6 The ML Degree of the Positive Grassmannian

We now turn to likelihood inference for other models on the Grassmannian. Perhaps the easiest way to obtain a statistical model from the Grassmannian is to restrict to the positive Grassmannian $\mathrm{Gr}(k,n)_{>0}$, which is defined by requiring that all Plücker coordinates x_I are positive. This semialgebraic set plays a prominent role at the interface of combinatorics and physics [129]. It is naturally a statistical model, via the usual identification of the positive projective space $\mathbb{P}^{\binom{n}{k}-1}_{\mathbb{R}_{>0}}$ with the probability simplex $\Delta_{\binom{n}{k}-1}$. In this model, the probability of observing a k-set $I = \{i_1, i_2, \ldots, i_k\}$ equals $x_I \,/\, \sum_J x_J$. The second model is the configuration space $X(k,n)$ which is obtained from $\mathrm{Gr}(k,n)_{>0}$ by taking the quotient modulo the natural torus action by the multiplicative group $\mathbb{R}^n_{>0}$. This model plays a prominent role in the study of scattering amplitudes; see [121]. In the special case $k = 2$, this Grassmannian model is the moduli space $\mathcal{M}_{0,n}$ of n distinct labeled points on the line $\mathbb{P}^1_{\mathbb{R}}$.

Proposition 8.6.1 ([121, Proposition 1]). *The ML degree of $\mathcal{M}_{0,n}$ is $(n-3)!$. For generic real data $u \in \mathbb{R}^n_{>0}$, all critical points are real.*

	$k=2$						$k=3$			
n	4	5	6	7	8	9	6	7	8	9
$\mathrm{sGr}(k,n)$	3	12	60	360	2 520	20 160	552	73 440	??	??
$X(k,n)$	1	2	6	24	120	720	26	1 272	188 112	74 570 400
$\mathrm{Gr}(k,n)_{>0}$	4	22	156	1 368	14 400	177 840	1 937	$\geq$ 499 976	??	??

Table 8.1: Maximum likelihood degrees of three models on small Grassmannians: the squared Grassmannian, the configuration space, and the positive Grassmannian.

The ML degrees of these two models and of the squared Grassmannian are summarized in Table 8.1 for small values of k, n. For $k = 2$, the first row follows from Theorem 8.1.4 and the second row follows from Proposition 8.6.1. The values for $X(3, n)$ come from [1, Theorem 5.1], and the first two entries in the last row appear in [68, Problem 13]. The last row follows from the main result of this section, Theorem 8.6.2, which gives a partial answer to [68, Problem 13]. The remaining values were computed numerically via `HomotopyContinuation.jl`.

Theorem 8.6.2. *The positive Grassmannian* $\mathrm{Gr}(2, n)_{>0}$ *has ML degree* $(2^{n-1} - n)(n - 3)!$*.*

We thank Simon Telen for his help with the following proof.

Proof. Since $\mathrm{Gr}(2, n)$ is smooth, we apply Theorem 2.3.5 and show that $\chi(\mathrm{Gr}(2, n)\backslash\mathcal{H}) = (2^{n-1} - n)(n - 3)!$. Consider the map

$$\begin{aligned}\mathrm{Gr}(2, n)\backslash\mathcal{H} = \mathrm{Gr}(2, n)^\circ\backslash V\Big(\sum_{1\le i<j\le n} x_{ij}\Big) &\to \mathcal{M}_{0,n} \cong \mathrm{Gr}(2, n)^\circ/(\mathbb{C}^*)^n,\\ x &\mapsto [x].\end{aligned}$$

We first observe that this map is surjective, meaning that for every $x \in \mathrm{Gr}(2, n)^\circ$, there is some scaling $(\lambda_i\lambda_j x_{ij})_{ij}$ such that the sum $\sum_{ij} \lambda_i\lambda_j x_{ij}$ is nonzero. If $\sum_{ij} \lambda_i\lambda_j x_{ij} = 0$ for all choices of $\lambda_1, \dots, \lambda_n$, then the columns of the matrix

$$H = \begin{pmatrix} 0 & x_{12} & \cdots & x_{1n} \\ x_{12} & 0 & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1n} & x_{2n} & \cdots & 0 \end{pmatrix} \tag{8.10}$$

must sum to zero. Then H is singular, and so by Lemma 8.6.3 at least one of the x_{ij} is zero. Therefore $x \notin \mathrm{Gr}(2, n)^\circ$, and we conclude that the map is surjective.

Since $\chi(\mathcal{M}_{0,n}) = (-1)^n(n-3)!$ by Proposition 8.6.1, it suffices to show that the fibers are isomorphic and have Euler characteristic $2^{n-1} - n$ by Lemma 8.4.1. To compute the fiber of $[x] \in \mathcal{M}_{0,n}$, choose a representative x. Then the fiber is isomorphic to $(\mathbb{C}^*)^n\backslash V(\sum_{ij} \lambda_i\lambda_j x_{ij})$. The Newton polytope of the polynomial $\sum_{ij} \lambda_i\lambda_j x_{ij}$ is the hypersimplex $\Delta_{2,n}$. By [55, Theorem 6.2.4], the fiber of $[x]$ has Euler characteristic $(-1)^n\mathrm{Vol}(\Delta_{2,n}) = (-1)^n(2^{n-1} - n)$, provided that x is not in the principal A-determinant of $\Delta_{2,n}$. Furthermore, as long as x and x' are not in the principal A-determinantal, the fibers of $[x]$ and $[x']$ are isomorphic. We now argue that the open Grassmannian $\mathrm{Gr}(2, n)^\circ$ does not intersect this principal A-determinant.

By [28, Theorem 3.6] the principal A-determinant of $\Delta_{2,n}$ is the product of principal minors of size at least 4 of (8.10). It therefore suffices to prove that if any of these minors vanish on $x \in \mathrm{Gr}(2, n)$, then $x_{ij} = 0$ for some i, j. As the principal submatrices of (8.10) can be identified with points in smaller Grassmannians, the claim follows from Lemma 8.6.3. □

Lemma 8.6.3. *Suppose* X *is an* $n \times n$ *skew-symmetric matrix whose* (i, j) *entry is* x_{ij} *if* $i > j$ *and* H *is as in* (8.10)*. If* $\mathrm{rank}\, X \le 2$*, then* $\det(H) = 2^{n-2}x_{12}x_{23}\cdots x_{1n}$*. Thus* H *is singular if and only if* $x_{i,i+1}$ *is zero for some* $i \in [n]$*.*

Proof. Since X has rank 2, there exist $u, v \in \mathbb{C}^n$ such that we can write the (i,j)-entry of X as $u_i v_j - u_j v_i$. Then the entries of H are $H_{ij} = (-1)^{i<j}(u_i v_j - u_j v_i)$ where $(-1)^{i<j}$ is -1 if $i < j$ and 1 otherwise. Therefore the determinant of H is

$$\det(H) = \sum_{\sigma \in \mathfrak{S}_n} (-1)^{\mathrm{inv}(\sigma)} \prod_{i=1}^{n} (-1)^{i<\sigma(i)} (u_i v_{\sigma(i)} - u_{\sigma(i)} v_i).$$

where $\mathfrak{S}_n$ denotes the symmetric group on n elements and $\mathrm{inv}(\sigma)$ denotes the number of inversions (pairs $i < j$ with $\sigma(i) > \sigma(j)$) of σ. An index i with $i < \sigma(i)$ is called an *excedance*. Writing $u_S = \prod_{i \in S} u_i$, we can write this determinant as

$$\begin{aligned}\det(H) &= \sum_{S \subseteq [n]} (-1)^{|S|} \sum_{\sigma \in \mathfrak{S}_n} (-1)^{\mathrm{inv}(\sigma)+\mathrm{exc}(\sigma)} u_S v_{\sigma(S)} u_{\sigma(S)^c} v_{S^c} \\ &= \sum_{S \subseteq [n]} \sum_{\substack{T \subseteq [n] \\ |T| = |S|}} (-1)^{|S|} \left(\sum_{\substack{\sigma \in \mathfrak{S}_n \\ \sigma(S) = T}} (-1)^{\mathrm{inv}(\sigma)+\mathrm{exc}(\sigma)} \right) u_S v_T u_{T^c} v_{S^c}\end{aligned}$$

where S^c denotes the complement $[n] \backslash S$ of S.

We will compute the coefficients $\sum_{\sigma \in \mathfrak{S}_n, \sigma(S) = T} (-1)^{\mathrm{inv}(\sigma)+\mathrm{exc}(\sigma)}$. A determinantal form of the *signed excedance enumerator* $\sum_{\sigma \in \mathfrak{S}_n} (-1)^{\mathrm{inv}(\sigma)} q^{\mathrm{exc}(\sigma)} = (1-q)^{n-1}$ is given in [118, pp. 785], namely, it is the determinant of the matrix

$$\begin{pmatrix} 1 & q & q & \cdots & q \\ 1 & 1 & q & \cdots & q \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 1 & 1 & \cdots & 1 \end{pmatrix} \tag{8.11}$$

It follows that if $S = T = [n]$ or $S = T = \emptyset$, then the coefficient $\sum_{\sigma \in \mathfrak{S}_n} (-1)^{\mathrm{inv}(\sigma)+\mathrm{exc}(\sigma)}$ is 2^{n-1}. We will modify the matrix to restrict the sum to permutations that map S to T for fixed proper subsets $\emptyset \subsetneq S, T \subsetneq [n]$. To do this, we zero out the entries $(i,j) \in (S \times T^c) \cup (S^c \times T)$. We will show that the determinant in question is zero unless $T = S$, in which case it is equal to $(1-q)^{n-2}$ or $T = S + 1 := \{s + 1 : s \in S\}$, in which case it is $-q(1-q)^{n-2}$. Note that if $n \in S$, we include 1 in $S + 1$. In the $T = S$ case, the matrix with zeros is block diagonal with two blocks of the form (8.11) having sizes $|S|$ and $n - |S|$. Thus the determinant is $(1-q)^{|S|-1}(1-q)^{n-|S|-1} = (1-q)^{n-2}$ (see [118, Theorem 1]). In the $T = S + 1$ case, the zeroed out version of (8.11) has two blocks. One has size $|S|$ and occupies the entries $S \times T$; the other has size $n - |S|$ and occupies the entries $S^c \times T^c$. Depending on whether n is in S or S^c, one block has the form (8.11), and the other has the form

$$\begin{pmatrix} q & q & q & \cdots & q \\ 1 & q & q & \cdots & q \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 1 & 1 & \cdots & q \end{pmatrix}.$$

In either case, the determinant is $-q(1-q)^{n-2}$ (see [118, Theorem 6]).

Now suppose that $T \notin \{S, S+1\}$. The zeroed out matrix (8.11) still has two blocks, but now one of them is singular. To see this, observe that the number of q-entries in a given column $t \in T$ is the number of $s \in S$ with $s < t$. For the matrix to be nonsingular, these numbers must be distinct for every column. For this to be true for both the (S, T) block and the (S^c, T^c) block, the T must be either S or $S+1$. Thus the matrix is singular and has determinant 0.

Plugging in $q = -1$, we see that the coefficients in our sum are 2^{n-2}:

$$\det(H) = \left(2^n + \sum_{\emptyset \subsetneq S \subsetneq [n]} \left((-1)^{|S|} 2^{n-2}\right) u_{[n]} v_{[n]} + \sum_{S \subseteq [n]} \left(-1\right)^{|S|} 2^{n-2} u_S v_{S+1} u_{S^C+1} v_{S^c}\right.$$
$$= 2^{n-2} \prod_{i=1}^{n} (u_i v_{i+1} - u_{i+1} v_i).$$

where $u_{n+1} = u_1$ and $v_{n+1} = v_1$. □

The ML degrees are now known for all three models on $\mathrm{Gr}(2, n)$. From the perspective of likelihood geometry, the two models that come from applications, namely $\mathcal{M}_{0,n}$ and $\mathrm{sGr}(2, n)$, are simpler models than the positive Grassmannian. Another observation is that, unlike the first two models, the critical points of the likelihood function on the Grassmannian are not all real. The key to understanding the real behavior is the logarithmic discriminant [77]. This is analogous to the RR discriminant in Section 5.2, and is computed in the same way.

Example 8.6.4. The logarithmic discriminant of $\mathrm{Gr}(2, 4)$ is an irreducible hypersurface of degree 14 in the data space $\mathbb{P}^5$. Its defining polynomial has 8 427 terms; the first few are

$$\begin{gathered} u_{12}^8 u_{13}^2 u_{14}^2 u_{23}^2 + 2u_{12}^8 u_{13} u_{14}^3 u_{23}^2 + u_{12}^8 u_{14}^4 u_{23}^2 + 8u_{12}^7 u_{13}^3 u_{14}^2 u_{23}^2 + 16u_{12}^7 u_{13}^2 u_{14}^3 u_{23}^2 \\ +8u_{12}^7 u_{13} u_{14}^4 u_{23}^2 + 26u_{12}^6 u_{13}^4 u_{14}^2 u_{23}^2 + 48u_{12}^6 u_{13}^3 u_{14}^3 u_{23}^2 + 16u_{12}^6 u_{13}^2 u_{14}^4 u_{23}^2 + 44u_{12}^5 u_{13}^5 u_{14}^2 u_{23}^2 \\ +64u_{12}^5 u_{13}^4 u_{14}^3 u_{23}^2 - 16u_{12}^5 u_{13}^3 u_{14}^4 u_{23}^2 + 41u_{12}^4 u_{13}^6 u_{14}^2 u_{23}^2 + 26u_{12}^4 u_{13}^5 u_{14}^3 u_{23}^2 - 95u_{12}^4 u_{13}^4 u_{14}^4 u_{23}^2 \\ +20u_{12}^3 u_{13}^7 u_{14}^2 u_{23}^2 - 24u_{12}^3 u_{13}^6 u_{14}^3 u_{23}^2 + 4u_{12}^2 u_{13}^8 u_{14}^2 u_{23}^2 - 28u_{12}^2 u_{13}^7 u_{14}^3 u_{23}^2 - 8u_{12} u_{13}^8 u_{14}^3 u_{23}^2 + \cdots \end{gathered}$$

Based on numerical computations, when we sample the entries of u uniformly at random from [1000], we observe that the number of real critical points is either 2 or 4. ◇

The main results of this section are formulae for the ML degree of the squared Grassmannian $\mathrm{sGr}(2, n)$ and the Grassmannian $\mathrm{Gr}(2, n)$, along with the fact that for positive data, all critical points of the maximum likelihood estimation problem on $\mathrm{sGr}(2, n)$ are positive. It would be interesting to study the discriminants of these optimization problems, and to extend these results to other Grassmannians.

Chapter 9

Squared Linear Models

In this chapter, we study an infinite family of models for which an analog of Theorem 8.1.5 holds. Consider a discrete statistical model on n states which is given by linear forms $\ell_1, \ldots, \ell_n$ in real variables $x_1, \ldots, x_s$, where $n > s > 1$. In this model, the probability of observing the i-th state is

$$p_i(x) \;=\; \frac{\ell_i^2(x)}{\sum_{j=1}^n \ell_j^2(x)} \qquad \text{for } i = 1, 2, \ldots, n. \tag{9.1}$$

Here $x = (x_1, \ldots, x_s)$ are the model parameters. The denominator is the partition function. We write $\mathcal{M}$ for the projective variety in $\mathbb{P}^{n-1} = \mathbb{P}^{n-1}_{\mathbb{C}}$ parametrized by (9.1) and $I_{\mathcal{M}}$ for its prime ideal in $\mathbb{C}[p_1, p_2, \ldots, p_n]$. Our aim is to explain the likelihood geometry of $\mathcal{M}$.

Example 9.0.1 ($s = 3, n = 4$)**.** Consider the model $\mathcal{M}$ defined by four lines in $\mathbb{P}^2$, namely

$$\ell_1 = x_1,\;\; \ell_2 = x_2,\;\; \ell_3 = x_3 \;\;\text{ and }\;\; \ell_4 = x_1 + x_2 + x_3.$$

Then $\mathcal{M}$ is a *Steiner surface* in $\mathbb{P}^3$ [35, Example 3.5] with ideal $I_{\mathcal{M}}$ generated by the quartic

$$\begin{gathered} p_1^4 + p_2^4 + p_3^4 + p_4^4 \;+\; 6\,(p_1^2p_2^2 + p_1^2p_3^2 + p_1^2p_4^2 + p_2^2p_3^2 + p_2^2p_4^2 + p_3^2p_4^2) \\ -\,4(p_1^3p_2 + p_1^3p_3 + \cdots + p_3p_4^3) \;+\; 4(p_1^2p_2p_3 + p_1^2p_2p_4 + \cdots + p_2p_3p_4^2) \;-\; 40p_1p_2p_3p_4. \end{gathered}$$

The surface $\mathcal{M}_{\mathbb{R}}$ is singular along three lines. Its smooth part has $16 = 4 + 3 \cdot 4$ connected components in the tetrahedron $\Delta_3 = \mathbb{P}^3_{>0}$. These components are formed by the $7 = 4 + 3$ regions of the line arrangement $\mathcal{A} = \{\ell_1, \ell_2, \ell_3, \ell_4\}$ in the plane $\mathbb{P}^2_{\mathbb{R}}$. The likelihood function

$$p_1(x)^{u_1} p_2(x)^{u_2} p_3(x)^{u_3} p_4(x)^{u_4}, \qquad \text{for given data } u_1, u_2, u_3, u_4 \in \mathbb{N},$$

is positive on $\mathbb{P}^2_{\mathbb{R}} \backslash \mathcal{A}$. It has seven complex critical points, all real, one in each region. $\diamond$

In general, the likelihood function of a squared linear model (9.1) can have arbitrarily many local maxima inside the probability simplex. We next present an example to show this.

Example 9.0.2 (The braid arrangement)**.** Fix $c = d+1$, $n = \binom{c}{2}$, and let $\mathcal{M}$ be the model

$$p_{ij}(x) \;=\; \frac{(x_i - x_j)^2}{c\,(\sum_{k=1}^{c} x_k^2) \,-\, (\sum_{k=1}^{c} x_k)^2} \qquad \text{for } 1 \le i < j \le c. \tag{9.2}$$

The ideal $I_\mathcal{M}$ is generated by the 2×2 minors of the symmetric $s \times s$ matrix with diagonal entries $2p_{ic}$ and off-diagonal entries $p_{ic} + p_{jc} - p_{ij}$ for $1 \le i, j \le d$. Thus $\mathcal{M}$ is a *Veronese variety*. It has dimension $c-2$ and degree 2^{c-2} in $\mathbb{P}^{\binom{c}{2}-1}$. The likelihood function has $c!/2$ complex critical points. All of them are real and positive. There is one such point in each region of the braid arrangement in $\mathbb{P}^{c-1}_\mathbb{R}$. These regions are indexed by the $c!$ permutations of $\{1, 2, \ldots, c\}$, modulo reversal involution. We view $\mathcal{M}$ as a subvariety of the squared Grassmannian $\mathrm{sGr}(2, c+1)$ (Section 8.2). Both models have the same ML degree. $\diamond$

Given the model (9.1), we fix data $u_1, \ldots, u_n \in \mathbb{R}_{>0}$. We are interested in the log-likelihood function

$$L_u(x) = u_1 \log\, \ell_1^2(x) + \cdots + u_n \log\, \ell_n^2(x) - (u_1 + \cdots + u_n) \log\bigl(\ell_1^2(x) + \cdots + \ell_n^2(x)\bigr). \tag{9.3}$$

In Section 9.1 we prove that all complex critical points of (9.3) are real, and there is precisely one critical point in each connected component of $\mathbb{P}^{d-1}_\mathbb{R} \backslash \mathcal{A}$. Hence the ML degree equals the number of such regions, which is $\sum_{i=0}^{d-1} \binom{n-1}{i}$ when the hyperplane arrangement $\mathcal{A}$ is generic; see, e.g., (9.9). This result is Theorem 9.1.1. It rests on the work of Reinke and Wang [109].

In Section 9.2 we study the variety $\mathcal{M} \subset \mathbb{P}^{n-1}$, focusing on generic arrangements $\mathcal{A}$. For large n, the ideal $I_\mathcal{M}$ is generated by linear forms and the 2×2 minors of a symmetric $s \times s$ matrix (Proposition 9.2.1). For small n, the variety $\mathcal{M}$ is a projection of the Veronese variety $\nu_2(\mathbb{P}^{s-1})$. The ideal $I_\mathcal{M}$ is complicated, and tough computations are needed. Proposition 9.2.4 summarizes what we know. The singular locus of $\mathcal{M}$ is determined in Theorem 9.2.5.

Section 9.3 is devoted to the parametric likelihood correspondence. This is the variety in $\mathbb{P}^{s-1} \times \mathbb{P}^{n-1}$ which parametrizes pairs (x, u) where x is a critical point of the log-likelihood function (9.3) for the data u. We study its bihomogeneous prime ideal using the methods in [75, Section 2]. Theorem 9.3.2 expresses the likelihood ideal for generic arrangements as a determinantal ideal. The ideal is minimally generated by $\binom{n}{s-2}$ polynomials of bidegree $(1, n-s+2)$ in $\mathbb{C}[x, u]$.

In Section 9.4 we study likelihood degenerations [1] when $\mathcal{A}$ is generic. Theorem 9.4.1 shows that the critical points are distinct even when u is a unit vector e_i. We then turn to tropical maximum likelihood estimation (MLE) for squared linear models. Tropical MLE for linear models appeared in [1, Section 7], and it was developed for toric models in [11]. Corollary 9.4.3 is an analogue to [1, Theorem 7.1], but now for all regions of the arrangement $\mathcal{A}$, not just bounded regions.

In Section 9.5 we investigate log-normal polytopes and log-Voronoi cells, as defined in [4]. For each distribution p in the model $\mathcal{M}$, the former parametrizes data u such that p is critical for (9.3), while the latter parametrizes data u such that p is the MLE. These convex bodies agree for linear models [3]. They disagree for squared linear models, even when both are polytopes (Example 9.5.5). The log-normal polytopes are characterized in Theorem 9.5.1.

In Section 9.6, we show how squared linear models arise in applications. Namely, we prove that when the hyperplane arrangement $\mathcal{A}$ is a discriminantal arrangement, then the corresponding model is a linear projection determinantal point process (Proposition 9.6.2).

Software and data for computational results in this paper are presented on a Zenodo page, available at https://doi.org/10.5281/zenodo.15997158.

9.1 Real and Positive

We will show that all complex critical points x of (9.3) are real. Therefore, $\ell_1^2(x), \ldots, \ell_n^2(x)$ are positive, and every critical point of the log-likelihood function yields a probability distribution in Δ_{n-1}. This is reminiscent of the proof of Theorem 8.1.5. By Theorem 2.3.5, the number of critical points of (9.3) is the signed Euler characteristic of the very affine variety underlying our statistical model. This variety is the complement in $\mathbb{P}^{s-1}$ of the n hyperplanes $V(\ell_1), \ldots, V(\ell_n)$ and the quadric $V(q)$ defined by $q = \ell_1^2 + \cdots + \ell_n^2$. Its Euler characteristic is an upper bound on the number of real critical points. One argues that there is at least one critical point per region of $\mathbb{P}^{s-1}_{\mathbb{R}} \backslash \cup_{i=1}^n V(\ell_i)$. This gives a lower bound on the number of real critical points. The punchline is that the two bounds agree.

Theorem 9.1.1. *For generic data $u \in \mathbb{R}^n_{>0}$, all complex critical points of the log-likelihood function (9.3) are real, and there is one critical point in each region of $\mathbb{P}^{s-1}_{\mathbb{R}} \backslash \mathcal{A}$. Hence every critical point on the implicit model $\mathcal{M} \subset \mathbb{P}^{n-1}$ is positive and is a local maximum.*

As a consequence, any local optimization method, e.g., gradient ascent, used to find the MLE is likely to fail, since there is an abundance of spurious local maxima. Instead, computing the MLE via homotopy continuation might be more appropriate.

To prove Theorem 9.1.1, we apply [109, Theorem 1.2], which states that, if the partition function $q = \ell_1^2 + \cdots + \ell_n^2$ is replaced with a generic quadric g, then all critical points of (9.3) are real. By generic, we mean that the hypersurface $V(g)$ is in general position relative to the hyperplane arrangement $\mathcal{A} = \{V(\ell_1), \ldots, V(\ell_n)\}$, i.e., no nonempty intersection $\bigcap_{i\in I} V(\ell_i)$ for $I \subseteq [n]$ is tangent to $V(g)$. We will prove that the partition function q satisfies this.

Lemma 9.1.2. *The quadratic hypersurface defined by $q = \ell_1^2 + \cdots + \ell_n^2$ in $\mathbb{P}^{s-1}$ is in general position with respect to the hyperplane arrangement $\mathcal{A}$ that underlies the squared linear model.*

Proof. We prove that, for any nonempty subset $I \subseteq [n]$, the intersection $\bigcap_{i\in I} V(\ell_i)$ is not tangent to the hypersurface $V(q)$. Let A be the $n \times s$ matrix satisfying $Ax = (\ell_i(x))_{i\in[n]}$. We assume that A has rank s, so the arrangement $\mathcal{A}$ is essential. Otherwise, $s - \mathrm{rank}(A)$ of the variables can be eliminated, resulting in an essential arrangement in $\mathbb{P}^{\mathrm{rank}(A)-1}$. If A_i denotes the ith row of A, then the gradient of the partition function is

$$\nabla(q(x)) \;=\; 2\ell_1(x)A_1 + \cdots + 2\ell_n(x)A_n \;=\; 2x^T A^T A.$$

We proceed by induction on the cardinality $|I|$. We first claim that $V(\ell_i)$ is not tangent to $V(q)$ for $i = 1, \ldots, n$. This is equivalent to proving that the $2 \times s$ matrix

$$C \;=\; \begin{pmatrix} x^T A^T A \\ A_i \end{pmatrix}$$

has rank 2 at every point $x \in V(q) \cap V(\ell_i)$. We now show the contrapositive.

Suppose there exists $\alpha \in \mathbb{C}^*$ such that $\alpha x^T A^T A = A_i$. Since A_i is real and nonzero and $A^T A$ is a real invertible matrix, αx is also real and nonzero. Since q is a sum of squares and $\alpha \neq 0$, we have $q(\alpha x) = \alpha^2 q(x) \neq 0$. Thus the matrix C has rank 2 for any $x \in V(q) \cap V(\ell_i)$.

We now consider an index set I of cardinality r. After relabeling, $I = \{1, 2, \ldots, r\}$, and we assume that $\bigcap_{i=1}^r V(\ell_i)$ is non-empty. We must show that this intersection is not tangent to $V(q)$. If it were tangent, there would exist $\lambda_1, \ldots, \lambda_r$ such that $\lambda_1 A_1 + \cdots + \lambda_r A_r = x^T A^T A$. Letting $\mathrm{proj}_{\ell_i}(v)$ denote the orthogonal projection of v onto the hyperplane $V(\ell_i)$, this implies

$$0 + \lambda_2 \mathrm{proj}_{\ell_1} A_2 + \cdots + \lambda_r \mathrm{proj}_{\ell_1} A_r \;=\; \mathrm{proj}_{\ell_1}(x^T A^T A). \tag{9.4}$$

We now view the intersections $V(\ell_2) \cap V(\ell_1), \ldots,\ V(\ell_n) \cap V(\ell_1),\ V(q) \cap V(\ell_1)$ as subvarieties of $V(\ell_1) \cong \mathbb{P}^{d-2}$. In that ambient space, the normal vector of $V(q) \cap V(\ell_1)$ is $\mathrm{proj}_{\ell_1}(x^T A^T A)$ at x. Similarly, the normal vector of $V(\ell_i) \cap V(\ell_1)$ is $\mathrm{proj}_{\ell_1}(A_i)$. The condition (9.4) means that the intersection $\bigcap_{i=2}^r (V(\ell_i) \cap V(\ell_1))$ is tangent to $V(q) \cap V(\ell_1) = V(\ell_2^2 + \cdots + \ell_n^2) \cap V(\ell_1)$ inside $\mathbb{P}^{s-2}$. This is a contradiction to our induction hypothesis. We conclude that the ℓ_i are in general position relative to the quadric q in the complex projective space $\mathbb{P}^{d-1}$. $\square$

Remark 9.1.3. Lemma 9.1.2 can fail if the coefficients of the $\ell_1, \ldots, \ell_n$ are complex. For example, if $s = 2, n = 3$ and $\ell_1 = \sqrt{-1} \cdot x,\ \ell_2 = y,\ \ell_3 = x + y$, then the quadric factors as $q = \ell_1^2 + \ell_2^2 + \ell_3^2 = 2y(x + y)$. The intersection $V(q) \cap V(\ell_2) = V(\ell_2)$ is not transverse.

We will now prove the main result in this section.

Proof of Theorem 9.1.1. By Lemma 9.1.2, the quadric $V(\ell_1^2 + \cdots + \ell_n^2)$ is in general position with respect to the arrangement $\mathcal{A}$. As in the previous proof, we assume that $\mathcal{A}$ is essential, since we can always reduce the number of variables to obtain an essential arrangement.

We work in an affine chart where $\ell_1(x) = 1$; the quadric is still general with respect to $V(\ell_2), \ldots, V(\ell_n)$ and the arrangement $\mathcal{A}' = \{V(\ell_2), \ldots, V(\ell_n)\}$ is essential. Indeed, since $n > s$, we may assume that the normal of $V(\ell_1)$ is in the span of the normals of $V(\ell_2), \ldots, V(\ell_n)$.

On our affine chart $\{\ell_1(x) = 1\} \cong \mathbb{C}^{d-1}$, we consider the affine log-likelihood function

$$2u_2 \log\, \ell_2(x) + \cdots + 2u_n \log\, \ell_n(x) \,-\, (u_1 + \cdots + u_n) \log(1 + \ell_2(x)^2 + \cdots + \ell_n(x)^2). \tag{9.5}$$

We do not lose any critical points when moving to the affine chart, since all critical points of (9.3) have $\ell_1(x) \neq 0$. By [109, Theorem 1.2], the affine log-likelihood function (9.5) has only real critical points, and there is exactly one critical point per region of the affine hyperplane arrangement $\mathbb{R}^{s-1} \backslash \mathcal{A}'$. Hence the log-likelihood function (9.3) has only real critical points. It has exactly one critical point per region of the projective hyperplane arrangement $\mathbb{P}^{s-1}_{\mathbb{R}} \backslash \mathcal{A}$.

The implicit model $\mathcal{M}$ is a subvariety of $\mathbb{P}^{n-1}$. For a critical point p^* in $\mathcal{M}$, there is a critical point $x^* \in \mathbb{P}^{s-1}$ of (9.3) such that $p_i^* = \ell_i^2(x^*)$ for all i. Since x^* is real, all complex critical points p^* on X are real with positive coordinates. The likelihood function $\prod_{i=1}^n \ell_i(x)^{2u_i}$ is positive on $\mathbb{P}_{\mathbb{R}}^{s-1} \backslash \mathcal{A}$ and zero on $\mathcal{A}$, so all critical points are local maxima. $\square$

The number of regions of a hyperplane arrangement $\mathcal{A}$ in real projective space $\mathbb{P}_{\mathbb{R}}^{s-1}$ is computed by means of the *characteristic polynomial* $\chi_{\mathcal{A}}(t)$ of the matroid represented by $\mathcal{A}$. See [109, Section 2] for definitions and [24] for state of the art on computations. Zaslavsky's Theorem states that the number of regions of the central arrangement in $\mathbb{R}^s$ equals $|\chi_{\mathcal{A}}(-1)|$.

Corollary 9.1.4. *The ML degree of the squared linear model* (9.1) *is equal to* $|\chi_{\mathcal{A}}(-1)|/2$.

Example 9.1.5. The braid arrangement in Example 9.0.2 has the characteristic polynomial

$$\chi_{\mathcal{A}}(t) \;=\; (t-1)(t-2)\cdots(t-(c-1)).$$

Hence the ML degree of the squared linear model (9.2) is equal to $|\chi_{\mathcal{A}}(-1)|/2 \;=\; c!/2$. $\diamond$

Remark 9.1.6. The variety $\mathcal{M}$ has degree 2^{s-1} in $\mathbb{P}^{n-1}$ provided $\mathcal{A}$ contains $s+1$ hyperplanes in general position. In this case, the map $\mathbb{P}^{s-1} \to \mathcal{M}$ is birational, and the ML degree of the parametric model agrees with that of the implicit model. To prove this, consider a generic fiber $\{(\pm x_1 : \pm x_2 : \cdots : \pm x_s)\}$ of the coordinatewise squaring map $\mathbb{P}^{s-1} \to \mathbb{P}^{s-1}$. These 2^{s-1} points have distinct images in $\mathbb{P}^s$ if we tag on the extra coordinate $(\pm x_1 + \pm x_2 + \cdots + \pm x_n)^2$.

A birational inverse from $\mathcal{M}$ to $\mathbb{P}^{s-1}$ can be found using Gröbner bases, but its coordinates are generally complicated. For instance, here is such an inversion formula for Example 9.0.1:

$$\frac{x_1}{x_3} \;=\; \frac{p_1^2+p_2^2+p_3^2+p_4^2-2p_1p_2+6p_1p_3-2p_1p_4-2p_2p_3-2p_2p_4-2p_3p_4}{4p_3(-p_1+p_2-p_3+p_4)}.$$

If the matroid of $\mathcal{A}$ is not connected, then the map $\mathbb{P}^{s-1} \to \mathcal{M}$ is not birational, and the degree of $\mathcal{M}$ in $\mathbb{P}^{n-1}$ is strictly smaller than 2^{s-1}. For instance, if we replace Example 9.0.1 by

$$\ell_1 = x_1,\ \ell_2 = x_2,\ \ell_3 = x_3 \ \text{ and } \ \ell_4 = x_1 + x_2,$$

then $\mathcal{M}$ is a quadratic cone in $\mathbb{P}^3$, and its parametrization $\mathbb{P}^2 \to \mathcal{M}$ is two-to-one.

Let $\mathcal{M}_{>0}$ denote the intersection of the variety $\mathcal{M}$ with the open probability simplex $\Delta_{n-1} = \mathbb{P}_{>0}^{n-1}$. In statistics, $\mathcal{M}_{>0}$ is the natural domain of the log-likelihood function.

Proposition 9.1.7. *If the variety* $\mathcal{M}$ *is smooth and its parametrization* $\mathbb{P}^{s-1} \to \mathcal{M}$ *is birational, then* $\mathcal{M}_{>0}$ *is a manifold with precisely* $|\chi_{\mathcal{A}}(-1)|/2$ *connected components. This holds for generic arrangements* $\mathcal{A}$ *when* $n \geq 2s-1$*, but it does not hold for* $n = s+1 \geq 4$.

Proof. Here we make a forward reference to the results on singularities in Theorem 9.2.5. Since $\mathcal{M}$ is smooth, the parametrization is one-to-one outside the n hyperplanes $V(\ell_i)$ of the arrangement $\mathcal{A}$. Hence the parametrization restricts to a homeomorphism between $\mathbb{P}^{s-1}_{\mathbb{R}}\backslash\mathcal{A}$ and $\mathcal{M}_{>0}$. The inverse is given by identifying the correctly signed square root for each $p_i(x) = \ell_i(x)^2$. The result hence follows from Zaslavsky's Theorem. The failure for $n = s+1$ was seen for the Steiner surface in Example 9.0.1. Here the singular locus has codimension one, and the map attaches pairs of different regions of $\mathbb{P}^{s-1}_{\mathbb{R}}\backslash\mathcal{A}$ along the singular locus. □

9.2 Implicit Model Description

In this section, we study the implicit representation of a squared linear model in $\mathbb{P}^{n-1}$. Our aim is to describe its homogeneous prime ideal. We first focus on the generic case. Let $\mathcal{M}_{s,n}$ denote the model for n generic linear forms in s variables, and write $I_{s,n}$ for its prime ideal. Thus $I_{s,n}$ is the ideal of algebraic relations among the squares of n generic linear forms in $x_1, \ldots, x_s$. That ideal is easy to describe in the regime when the number n is large enough relative to s.

Fix $m = \binom{s+1}{2}$. Let M be the $n \times m$ matrix whose ith row consists of the coefficients of $\ell_i^2(x)$. The columns of M are indexed by quadratic monomials in $x_1, \ldots, x_s$. Let $(M\ p)$ denote the $n \times (m+1)$ matrix obtained from M by appending a column of variables $p = (p_1 \cdots p_n)^T$. To construct the quadratic generators of $I_{s,n}$, we define M' to be the $m \times m$ matrix formed by the first m rows of M. The matrix M' is invertible because the linear forms $\ell_1, \ldots, \ell_m$ are generic, so their squares span the space of all quadratic forms. We define the column vector

$$r \;=\; M'^{-1} \cdot (p_1\ p_2\ \cdots\ p_m)^T.$$

The entries of r are indexed by the quadratic monomials in $x_1, \ldots, x_s$. These are naturally identified with the entries of a symmetric $s \times s$ matrix $R = (r_{ij})$. Hence R is a $s \times s$ matrix whose entries are linear forms in the first $m = \binom{s+1}{2}$ coordinates $p_1, \ldots, p_m$ of $\mathbb{P}^{n-1}$. Given any matrix M and $t \in \mathbb{N}$, we write $I_t(M)$ for the ideal generated by the $t \times t$ minors of M.

Proposition 9.2.1. *Let $n \geq m$. The prime ideal of $\mathcal{M}_{s,n}$ equals $I_{s,n} = I_{m+1}((M\ p)) + I_2(R)$. This ideal is minimally generated by $n - \binom{s}{2}$ linear forms and $\frac{1}{12}(s+1)s^2(s-1)$ quadrics. The variety $\mathcal{M}_{s,n}$ is isomorphic to the quadratic Veronese embedding of $\mathbb{P}^{s-1}$ into $\mathbb{P}^{m-1}$.*

Proof. The $(m+1) \times (m+1)$ minors of $(M\ p)$ give linear forms in p that vanish on the variety $\mathcal{M}_{d,n}$. Since M has rank m, precisely $n - m$ of these linear forms are linearly independent.

By construction, the matrix R has rank one on the variety $\mathcal{M}_{s,n}$. This means that the 2×2 minors of R lie in the ideal $I_{s,n}$. Together with the $n-m$ independent linear forms from $(M\ p)$ they generate a prime ideal of the correct codimension. Hence this ideal equals $I_{s,n}$.

A symmetric $s \times s$ matrix has $\binom{\binom{s}{2}+1}{2}$ minors of size two. These are linearly dependent for $s \geq 4$. Namely, they obey $\binom{s}{4}$ linear relations. Hence the space spanned by the 2×2

minors of a generic symmetric $s \times s$ matrix has dimension $\frac{1}{12}(s+1)s^2(s-1)$. This is also the dimension of the Schur module $S_{2,2}(\mathbb{C}^d)$, so our count follows from [84, Proposition 6.10.4.1].

We claim that $\mathcal{M}_{s,n}$ is isomorphic to $\nu_2(\mathbb{P}^{s-1})$. Indeed, the coordinate ring of $\mathcal{M}_{s,n}$ is

$$\mathbb{C}[p_1,\ldots,p_n]/I_{s,n} \;=\; \mathbb{C}[p_1,\ldots,p_n]/(I_2(R) + I_{m+1}([M\ \ p])) \;\cong\; \mathbb{C}[p_1,\ldots,p_m]/I_2(R).$$

The isomorphism follows from the fact that the maximal minors of $(M\ \ p)$ express $p_{m+1},\ldots,p_n$ as linear combinations of $p_1,\ldots,p_m$. Now, apply Proj to these isomorphic graded rings. $\square$

The braid arrangements in Example 9.0.2 yield an infinite family of models with $n = m$. The first instance in this family of quadratic Veronese varieties is the plane curve $\nu_2(\mathbb{P}^1) \subset \mathbb{P}^2$.

Example 9.2.2. Let $s = 2, n = m = 3$ and consider the linear forms $\{x, y, x+y\}$. We have

$$M \;=\; M' \;=\; \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 2 & 1 \end{pmatrix} \quad\text{and hence}\quad M'^{-1}p \;=\; \frac{1}{2}\begin{pmatrix} 2 & 0 & 0 \\ -1 & -1 & 1 \\ 0 & 2 & 0 \end{pmatrix} p \;=\; \frac{1}{2}\begin{pmatrix} 2p_1 \\ p_3 - p_1 - p_2 \\ 2p_2 \end{pmatrix}.$$

The model $\mathcal{M}_{2,3} \cong \nu_2(\mathbb{P}^1)$ is the conic in the plane $\mathbb{P}^2$ defined by $4p_1p_2 = (p_3 - p_1 - p_2)^2$. $\diamond$

Example 9.2.3. Consider the seven lines in $\mathbb{P}^2$ defined by $\{x, y, z, x+y+z, x+2y+3z, x+5y+7z, x+11y+13z\}$. Here, $s = 3, n = 7$ and $m = \binom{3+1}{2} = 6$. We consider the 7×7 matrix

$$(M\ \ p) \;=\; \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & p_1 \\ 0 & 0 & 1 & 0 & 0 & 0 & p_2 \\ 0 & 0 & 0 & 0 & 0 & 1 & p_3 \\ 1 & 2 & 1 & 2 & 2 & 1 & p_4 \\ 1 & 4 & 4 & 6 & 12 & 9 & p_5 \\ 1 & 10 & 25 & 14 & 70 & 49 & p_6 \\ 1 & 22 & 121 & 25 & 286 & 169 & p_7 \end{pmatrix}.$$

The first six columns are labeled by $(x^2, xy, y^2, xz, yz, z^2)$. The upper left 6×6 block of $(M\ \ p)$ is the matrix M'. Its inverse yields the vector $r = M'^{-1}p$, whose coordinates are the entries in

$$R \;=\; \begin{pmatrix} r_1 & r_2 & r_4 \\ r_2 & r_3 & r_5 \\ r_4 & r_5 & r_6 \end{pmatrix}.$$

The prime ideal $I_{s,n}$ of the model $\mathcal{M}_{3,6} \subset \mathbb{P}^5$ is generated by one linear form and six quadrics. The linear form is the determinant of $(M\ \ p)$, and the quadrics are the 2×2 minors of R. $\diamond$

We now turn to the regime where n is smaller than $m = \binom{s+1}{2}$. Here the ideal generators are much more complicated than in Proposition 9.2.1. Not much is known about the ideals of such projections in general. The following result collects what we know in small dimensions.

Proposition 9.2.4. *Table 9.1 gives the degrees of the generators of $I_{s,n}$ for $s \leq 6$ and $n < m$.*

(s,n)	Degrees	(s,n)	Degrees	(s,n)	Degrees	(s,n)	Degrees
$(3,4)$	4^1	$(5,6)$	16^1	$(5,14)$	2^{35}	$(6,13)$	4^{533}
$(3,5)$	3^7	$\mathbf{(5,7)}$	$\mathbf{9^{71}10^{35}11^1}$			$(6,14)$	3^{98}
		$(5,8)$	$5^1 6^{98}$	$(6,7)$	32^1	$(6,15)$	3^{218}
$(4,5)$	8^1	$(5,9)$	5^{286}	$(6,8)$	$15^1 16^{178} 17^{3708} 18^{11}$	$(6,16)$	$2^{10}3^{194}$
$(4,6)$	$4^1 5^{10} 6^2$	$(5,10)$	$3^{10}4^{120}$	$(6,9)$	$10^{1338}11^{1668}$	$(6,17)$	$2^{27}3^{48}$
$(4,7)$	4^{45}	$(5,11)$	3^{76}	$(6,10)$	$7^{694}8^{1158}$	$(6,18)$	2^{45}
$(4,8)$	$2^2 3^{20} 4^1$	$(5,12)$	$2^8 3^{58}$	$(6,11)$	6^{1820}	$(6,19)$	2^{64}
$(4,9)$	2^{10}	$(5,13)$	$2^{21}3^2$	$(6,12)$	$4^{78}5^{429}$	$(6,20)$	2^{84}

Table 9.1: Degrees of the minimal generators (multiplicities in the exponents) for the prime ideals $I_{s,n}$. The gray entries refer not to minimal generators but to Gröbner basis elements.

Proof and Discussion. Proposition 9.2.4 was obtained by computer algebra, utilizing the Gröbner basis implementation in `Oscar.jl` [102]. Computations were performed over the rational numbers or finite fields. We computed Gröbner bases and extracted minimal generators. The cases $(6,8),(6,9),(6,10)$ are more challenging and will be completed in the future.

The ideal $I_{s,n}$ comprises relations among squares ℓ_i^2 of generic linear forms ℓ_i. If we replace these squares with generic quadrics then most entries of Table 9.1 remain unchanged. The smallest exception is $(s,n)=(5,7)$. Here, generic quadrics yield different numbers of minimal generators. The prime ideal $I_{5,7}$ has 107 minimal generators. There are 71 generators in degree 9, and 35 in degree 10, and one in degree 11. By contrast, the prime ideal of relations among seven generic quadrics in five variables requires 112 minimal generators, namely 70 in degree 9 and 42 in degree 10. The $(5,7)$ entry would be $\mathbf{9^{70}10^{42}}$ for generic quadrics. $\square$

What follows is a brief discussion of the prime ideal I for the squared linear model $\mathcal{M}$ given by an arbitrary hyperplane arrangement $\mathcal{A}$. The linear forms $\ell_1,\ell_2,\ldots,\ell_n$ are allowed to be special. The ideal I was studied for small s by Dey, Görlach and Kaihnsa in [35, Section 4.3]. Their analysis rests on the point configuration dual to the arrangement $\mathcal{A}$. Let $A_i=\nabla_x\ell_i(x)\in\mathbb{R}^s$ be the vector of coefficients of ℓ_i. We view $A_1,\ldots,A_n$ as points in $\mathbb{P}^{s-1}$.

The ideal I depends on the space of quadrics in $\mathbb{P}^{s-1}$ that pass through the points $A_1,\ldots,A_n$. If there are no such quadrics then we are in the generic situation of Proposition 9.2.1. If the A_i span a unique quadric that is irreducible then I is generated by $n-m-1$ linear forms together with the nonlinear equations of $\mathcal{M}_{s,m-1}$. For instance, if $s=3$ and $A_1,A_2,\ldots,A_n$ lie on a conic in $\mathbb{P}^2$ then I is generated by $n-5$ linear forms and 7 cubics. This is case (a) in [35, Theorem 4.9 (ii)]. If the A_i lie on several quadratic hypersurfaces then we are led to a case distinction as in [35, Theorem 4.9 (iii)]. See also [35, Section 4.2] for an interesting connection to the variety of symmetric matrices with degenerate eigenvalues, namely the symmetric matrices in the discriminant in Example 5.2.6.

We now turn to the question whether the generic model $\mathcal{M}_{s,n}$ is smooth. This happens when n is large relative to s. For instance, if $s = 3$ then the model is singular for $n = 4$, by Example 9.0.1, but it is smooth for $n \geq 5$. This threshold is explained by the following theorem.

Theorem 9.2.5. *If $n > 2s-2$ then the generic squared linear model $\mathcal{M}_{s,n}$ is smooth in $\mathbb{P}^{n-1}$. If $n \leq 2s-2$ then $\mathcal{M}_{s,n}$ is singular, and the singular locus has dimension $2s-n-1$. It is the image of all linear subspaces in $\mathbb{P}^{s-1}$ on which the map $x \mapsto (\ell_1^2(x) : \cdots : \ell_n^2(x))$ is not injective. In particular, for $n = 2s-2$ the singular locus of $\mathcal{M}_{s,n}$ consists of $\frac{1}{2}\binom{2s-2}{s-1}$ lines.*

Proof. For $n \geq m = \binom{s+1}{2}$, the variety $\mathcal{M}_{s,n}$ is smooth by Proposition 9.2.1. Next suppose $2s-2 < n < m$. Then $\mathcal{M}_{s,n}$ is the image of $\nu_2(\mathbb{P}^{s-1}) \subset \mathbb{P}^{m-1}$ under a generic linear projection $\pi\colon \mathbb{P}^{m-1} \dashrightarrow \mathbb{P}^{n-1}$. The center $\mathcal{E}$ of π has dimension $m-n-1$. The secant variety $\sigma_2(\nu_2(\mathbb{P}^{s-1}))$ has dimension $2s-2$, by [84, Exercise 5.1.2.4]. Therefore, since $2s-2 < n$, a general linear space $\mathcal{E}$ does not meet the secant variety of $\nu_2(\mathbb{P}^{s-1})$. Thus, by [117, Corollary 2.7], π defines an isomorphic embedding of $\mathbb{P}^{s-1}$ into $\mathbb{P}^{n-1}$, and smoothness is preserved.

Now, let $n \leq 2s-2$. Since $\mathcal{A}$ is generic, the rank of the Jacobian of our map is constant on $\mathbb{P}^{s-1} \backslash \mathcal{A}$. All singularities of $\mathcal{M}_{s,n}$ arise from pairs $\{x, x'\}$ of distinct points in $\mathbb{P}^{s-1}$ which have the same image in $\mathcal{M}_{s,n} \subset \mathbb{P}^{n-1}$. Since the linear map $A : \mathbb{P}^{s-1} \to \mathbb{P}^{n-1}, x \mapsto (\ell_i(x))_{i \in [n]}$ is injective, these points satisfy $Ax \neq Ax'$. Hence Ax and Ax' differ only by sign flips.

Consider any partition $I \sqcup J$ of $[n] = \{1, \ldots, n\}$. Let $A_{I,J}$ be the matrix whose ith row is A_i if $i \in I$ and $-A_i$ if $i \in J$. Then $(Ax)^2 = (Ax')^2$ exactly when $Ax' = A_{I,J}\,x$ for some partition $I \sqcup J = [n]$. Note that if $|I| \geq s$ or $|J| \geq s$, then x' is forced to equal x. Therefore the model parametrization is non-injective precisely on the subspaces $\ker(BA_{I,J}) \subset \mathbb{P}^{s-1}$, where the partition $I \sqcup J = [n]$ satisfies $|I|, |J| \leq s-1$. Here B denotes an $(n-s) \times s$ matrix with $\ker(B) = \mathrm{im}(A)$. The subspaces $\ker(BA_{I,J})$ have dimension $2s-n-1$ as desired.

In the boundary case $n = 2s-2$, we have $|I| = |J| = d-1$. There are precisely $\frac{1}{2}\binom{2s-2}{s-1}$ such partitions $I \sqcup J = [n]$. Each of them contributes a line to the singular locus of $\mathcal{M}_{s,n}$. $\square$

Example 9.2.6 ($s = 4$)**.** The variety $\mathcal{M}_{4,n}$ is smooth for $n \geq 7$. For $n = 6$, our model is a threefold of degree 8 in $\mathbb{P}^5$, defined by one quartic, ten quintics and two sextics. The singular locus of $\mathcal{M}_{4,6}$ consists of ten lines, one for each partition of $\{1, 2, \ldots, 6\}$ into two triples.

For $n = 5$, we take $\ell_i = x_i$ for $i = 1, 2, 3, 4$ and $\ell_5 = x_1 + x_2 + x_3 + x_4$. Then $I_{4,5}$ is generated by one polynomial of degree 8 with 495 terms. The singular locus of the threefold $\mathcal{M}_{4,5}$ decomposes into ten irreducible surfaces in $\mathbb{P}^4$. Each of these is a quadratic cone, like

$$V(p_1 - p_2,\, p_3^2 + p_4^2 + p_5^2 - 2p_3p_4 - 2p_3p_5 - 2p_4p_5) \quad = \quad \text{image of } \big[V(\ell_1, \ell_2) + V(\ell_3, \ell_4, \ell_5)\big].$$

It would be interesting to derive a general combinatorial rule for the singular locus of $\mathcal{M}$. $\diamond$

9.3 Likelihood Correspondence

In this section, we study the likelihood correspondence of the squared linear model given by n linear forms $\ell_1, \ldots, \ell_n$ in s variables $x = (x_1, \ldots, x_s)$. The likelihood correspondence (Chapter 2) captures the relationship between data and critical points of the log-likelihood function

$$L_{\mathcal{A},u}(x) \;=\; \sum_{i=1}^{n} u_i \log(p_i(x)) \;=\; \sum_{i=1}^{n} u_i \log(\ell_i^2(x)) \;-\; \log(q(x)) \sum_{j=1}^{n} u_j. \tag{9.6}$$

As before, $q = \ell_1^2 + \cdots + \ell_n^2$ is the partition function, and the parameters $u_1, \ldots, u_n$ represent the data. The partial derivatives of $L_{\mathcal{A},u}$ with respect to $x_1, \ldots, x_d$ are homogeneous rational functions. These depend linearly on $u_1, \ldots, u_n$. Our object of interest is the variety in the product space of all pairs (x, u) which is defined by setting these rational functions to zero:

Definition 9.3.1. The *likelihood correspondence* $\mathcal{P}_\mathcal{A}$ is the Zariski closure in $\mathbb{P}^{s-1} \times \mathbb{P}^{n-1}$ of

$$\left\{ (x, u) \in \mathbb{P}^{s-1} \times \mathbb{P}^{u-1} : \ \ell_1(x) \cdots \ell_n(x) q(x) \neq 0, \ p(x) \in \mathcal{M}_{\mathrm{reg}} \text{ and } \ \frac{\partial L_{\mathcal{A},u}}{\partial x_i}(x) \;=\; 0 \ \forall i \in [s] \right\}.$$

The *likelihood ideal* $I_\mathcal{A} \subset \mathbb{C}[u_1, \ldots, u_n, x_1, \ldots, x_d]$ is the bihomogeneous prime ideal of $\mathcal{P}_\mathcal{A}$.

The implicit likelihood correspondence was covered in Chapter 2; here we use the parametric likelihood correspondence, in the spirit of Section 2.4. The projection $\mathbb{P}^{s-1} \times \mathbb{P}^{n-1} \to \mathbb{P}^{n-1}$, $(x, u) \mapsto u$ induces a finite-to-one map $\mathcal{P}_\mathcal{A} \to \mathbb{P}^{n-1}$. The degree of this map is the ML degree. By Corollary 9.1.4, the ML degree is equal to the number of regions of the hyperplane arrangement. Thus, all fibers are fully real. Each connected component of $\mathbb{P}^{s-1}_\mathbb{R} \backslash \mathcal{A}$ is represented by one point in each fiber of $\mathcal{P}_\mathcal{A} \to \mathbb{P}^{n-1}$.

Let A denote the $n \times s$ Jacobian matrix of our linear forms. Thus, the rows of A are the gradient vectors $A_i = \nabla_x \ell_i(x)$ for $i = 1, \ldots, n$. In the following theorem, we assume that A is generic. In particular, the rank s matroid of A is uniform, and q is transverse to the generic arrangement $\mathcal{A}$ in $\mathbb{P}^{s-1}$. Fix any $(n-s) \times n$ matrix B whose kernel equals the image of A. The $n-s$ rows of B span the linear relations among the linear forms $\ell_1(x), \ell_2(x), \ldots, \ell_n(x)$.

Our main result in this section is the following explicit description of the prime ideal $I_\mathcal{A}$.

Theorem 9.3.2. *The likelihood ideal $I_\mathcal{A}$ for the generic model $\mathcal{M}_{s,n}$ is minimally generated by the maximal minors of the matrix*

$$C \;=\; \begin{pmatrix} u_1 & u_2 & \cdots & u_n \\ \ell_1^2 & \ell_2^2 & \cdots & \ell_n^2 \\ \multicolumn{4}{c}{B \cdot \mathrm{diag}(\ell_1, \ldots, \ell_n)} \end{pmatrix}. \tag{9.7}$$

This matrix has $n-s+2$ rows and n columns, and its $\binom{n}{s-2}$ minors all have degree $(1, n-s+2)$.

Proof. We use the description of the likelihood ideal for parametric models appearing in [74, Definition 2.3] and [75, Section 1]. Therein, the authors consider polynomials $f_1, \ldots, f_m$. When these define a strict normal crossing (SNC) arrangement, the construction in [75, equation (7)] gives a matrix $Q = Q^u_{\backslash 1}$ whose maximal minors generate the likelihood ideal. In that construction, the first polynomial f_1 is special. Namely, the subscript "\1" indicates that the column for f_1 was deleted from another matrix Q^u. In what follows, the partition function q plays the role of f_1, and the linear forms $\ell_1, \ldots, \ell_n$ play the role of $f_2, \ldots, f_m$.

The log-likelihood function of the squared linear model $\mathcal{M}_{s,n}$ is shown in (9.6). This function coincides with the log-likelihood function of the (unsquared) arrangement $\mathcal{A}' = \{q, \ell_1, \ell_2, \ldots, \ell_n\}$ after substituting $u_0 \mapsto -(u_1 + \cdots + u_n)$ and $u_i \mapsto 2u_i$ for $i = 1, \ldots, n$. We claim that the arrangement $\mathcal{A}'$ is SNC in the sense of [75]. This means that all intersections of hypersurfaces in $\mathcal{A}'$ are smooth of the expected dimension.

Indeed, by Lemma 9.1.2, no intersection of the hyperplanes is tangent to $V(q)$. It remains to show that the hyperplanes meet $V(q)$ in the expected dimension. Without loss of generality, we may work in an affine chart and transform coordinates so that $\ell_i(x) = x_i$ for $i = 1, \ldots, s$. Assume there exists $x \in V(q) \cap \bigcap_{i=1}^{s-1} V(\ell_i)$. Then $x = (0, \ldots, 0, x_s)$ and $q(x) = \alpha x_s^2$ for some nonzero constant α. Hence, $x \in V(\ell_s)$, so x is in the intersection of s hyperplanes. This is a contradiction to the ℓ_i being in general position. Therefore, $\mathcal{A}'$ is strict normal crossing.

It now follows from [75, Corollary 2.4] that the likelihood ideal $I_{\mathcal{A}}$ is generated by the maximal minors of the following matrix, which has $n+2$ rows and $n+d$ columns:

$$Q \;=\; \begin{pmatrix} 0 & 0 & \ldots & 0 & \partial_{x_1} q(x) & \partial_{x_2} q(x) & \ldots & \partial_{x_s} q(x) \\ \ell_1(x) & 0 & \ldots & 0 & \partial_{x_1}\ell_1(x) & \partial_{x_2}\ell_1(x) & \ldots & \partial_{x_s}\ell_1(x) \\ 0 & \ell_2(x) & \ldots & 0 & \partial_{x_1}\ell_2(x) & \partial_{x_2}\ell_2(x) & \ldots & \partial_{x_s}\ell_2(x) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \ell_n(x) & \partial_{x_1}\ell_n(x) & \partial_{x_2}\ell_n(x) & \ldots & \partial_{x_s}\ell_n(x) \\ u_1 & u_2 & \ldots & u_n & 0 & 0 & \ldots & 0 \end{pmatrix}. \tag{9.8}$$

On the right, we see the $n \times s$ Jacobian matrix $A = \big(\partial_{x_i}\ell_j(x)\big)$ of the n linear forms. And, the row above this is the gradient $\nabla q(x) = 2x^T A^T A$ of the quadratic form $q(x)$.

The image of A spans the kernel of the $(n-s) \times n$ matrix B, so the $n \times n$ matrix $\big(B^T\ A\big)^T$ is invertible. The following $(n+2) \times (n+2)$ matrix has an inverse with polynomial entries:

$$T \;=\; \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 \\ -1 & \ell_1(x) & \cdots & \ell_n(x) & 0 \\ 0 & & B & & 0 \\ 0 & & A^T & & 0 \end{pmatrix}.$$

Hence the ideal of maximal minors of Q agrees with that of the $(n+2) \times (n+s)$ matrix

$$T \cdot Q \;=\; \begin{pmatrix} u_1 \;\; u_2 \;\; \cdots \;\; u_n & 0 \;\; 0 \;\; \cdots \;\; 0 \\ \ell_1^2 \;\; \ell_2^2 \;\; \cdots \;\; \ell_n^2 & -x^T A^T A \\ B \cdot \mathrm{diag}(\ell_1, \ldots, \ell_n) & 0 \;\; 0 \;\; \cdots \;\; 0 \\ A^T \cdot \mathrm{diag}(\ell_1, \ldots, \ell_n) & A^T \cdot A \end{pmatrix}.$$

Since the rows of A^T are linearly independent, the symmetric $s \times s$ matrix $A^T A$ is invertible. Hence we can replace $T \cdot Q$ by its submatrix given by the first n columns and the first

$2+n-s$ rows. That submatrix is precisely the $(n-s+2)\times n$ matrix in (9.7). We have thus shown that $I_\mathcal{A}$ is generated by the maximal minors of (9.7). All generators have the same degree $(1, n-s+2)$. And, unlike Q, the matrix (9.7) has no constant entries. This ensures that the $\binom{n}{s-2}$ maximal minors of (9.7) are minimal generators for the prime ideal $I_\mathcal{A}$. □

Our argument confirms that the ML degree of $\mathcal{M}_{s,n}$ equals the ML degree of the arrangement $\mathcal{A}'$. According to [75], the latter is the coefficient of z^{s-1} in the generating function

$$\frac{1}{(1-z)^{n-s}\,(1-2z)}. \tag{9.9}$$

This coefficient is found to be $\mu = \sum_{i=0}^{s-1}\binom{n-1}{i}$, which is the number of regions in $\mathbb{P}^{s-1}_{\mathbb{R}}\backslash\mathcal{A}$.

Example 9.3.3 $(s=3, n=4)$**.** Let $\mathcal{A}$ be the arrangement of four lines in Example 9.0.1. Then

$$A \;=\; \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}^T \quad \text{and} \quad B \;=\; \begin{pmatrix} 1 & 1 & 1 & -1 \end{pmatrix}.$$

The likelihood ideal $I_\mathcal{A}$ in $\mathbb{C}[x_1,x_2,x_3,u_1,u_2,u_3,u_4]$ is generated by the 3×3 minors of

$$C \;=\; \begin{pmatrix} u_1 & u_2 & u_3 & u_4 \\ \ell_1^2 & \ell_2^2 & \ell_3^2 & \ell_4^2 \\ \ell_1 & \ell_2 & \ell_3 & -\ell_4 \end{pmatrix}.$$

The likelihood correspondence $\mathcal{P}_\mathcal{A}$ is a threefold in $\mathbb{P}^3\times\mathbb{P}^2$. The map onto $\mathbb{P}^3$ is 7-to-1. ⋄

We conclude with a remark on the non-generic case. If $\mathcal{A}'$ is not SNC then the maximal minors of the matrices (9.7) and (9.8) generate the same ideal I in $\mathbb{C}[x,u]$. However, that ideal is strictly contained in the likelihood ideal $I_\mathcal{A}$. We can compute $I_\mathcal{A}$ from I by saturation with respect to $\ell_1\ell_2\cdots\ell_n$. This follows from Proposition 2.9 and Remark 2.10 of [74].

Example 9.3.4 (The braid arrangement)**.** Let $s=3, n=6, c=4$ and consider the model in Example 9.0.2. Setting $x_1=x$, $x_2=y$, $x_3=z$, $x_4=0$ and relabeling, the matrix (9.7) equals

$$C \;=\; \begin{pmatrix} u_{12} & u_{13} & u_{14} & u_{23} & u_{24} & u_{34} \\ x^2 & y^2 & z^2 & (x-y)^2 & (x-z)^2 & (y-z)^2 \\ -x & y & 0 & x-y & 0 & 0 \\ -x & 0 & z & 0 & x-z & 0 \\ 0 & -y & z & 0 & 0 & y-z \end{pmatrix}.$$

The six maximal minors of C generate a radical ideal I. It has the prime decomposition

$$I \;=\; I_\mathcal{A}\cap\langle x,y\rangle\cap\langle x,z\rangle\cap\langle y,z\rangle\cap\langle x-y, x-z\rangle.$$

The likelihood ideal $I_\mathcal{A}$ has three minimal generators, of bidegrees $(1,3)$, $(1,4)$ and $(1,5)$. ⋄

9.4 Likelihood Degenerations

In this section, we examine the likelihood correspondence of squared linear models over degenerate data points u. In particular, we determine tropical limits of the critical points of (9.3). The framework of tropical likelihood degenerations was developed for linear models in [1, 5] and for toric models in [11]. It allows one to study maximum likelihood estimation as model parameters or data points approach special values, particularly model zeros. We extend these recent advances on the tropical geometry of likelihood inference.

The following formulation of our MLE problem will be used in this section. Our model is specified by two matrices $A \in \mathbb{R}^{n\times s}$ and $B \in \mathbb{R}^{(n-s)\times n}$ such that the image of A equals the kernel of B. We shall assume that the pair (A, B) is generic, in a sense to be made precise.

For our computations, we use the n unknowns $y_i = \sqrt{p_i} = \ell_i(x)$, $i = 1, 2, \ldots, n$. Thus we work in projective space $\mathbb{P}^{n-1}_{\mathbb{R}}$ with coordinates $y = (y_1 : y_2 : \cdots : y_n)$. The hyperplane arrangement $\mathcal{A}$ is the restriction of the coordinate hyperplanes in $\mathbb{P}^{n-1}_{\mathbb{R}}$ to $\mathrm{im}(A) \simeq \mathbb{P}^{s-1}_{\mathbb{R}}$.

According to Theorem 9.3.2, our task is to solve the system of polynomial equations

$$B \cdot y \,=\, 0 \quad \text{and} \quad \mathrm{rank}\left(\begin{pmatrix} u_1 & u_2 & \cdots & u_n \\ y_1^2 & y_2^2 & \cdots & y_n^2 \\ \multicolumn{4}{c}{B \cdot \mathrm{diag}(y_1, \ldots, y_n)} \end{pmatrix}\right) \leq n - s + 1. \tag{9.10}$$

For generic data $u = (u_1, \ldots, u_n) \in \mathbb{P}^{n-1}_{\mathbb{R}}$, the number of solutions to (9.10) is $\mu = \sum_{i=0}^{s-1} \binom{n-1}{i}$.

We set $[n] = \{1, 2, \ldots, n\}$ and we write e_i for the ith standard basis vector in $\mathbb{R}^n$. In what follows next, we specialize $u = e_i$. This means that $u_i = 1$ and $u_j = 0$ for $j \in [n]\backslash\{i\}$.

Theorem 9.4.1. *Suppose that (A, B) is generic and set $u = e_i$. Then (9.10) defines a radical ideal in $\mathbb{R}[y_1, \ldots, y_n]$, which has μ zeros y in $\mathbb{P}^{n-1}_{\mathbb{R}}$. The supports of the μ zeros are distinct, i.e., $J = \{j \in [n] : y_j = 0\}$ ranges over all subsets of cardinality at most $s - 1$ in $[n]\backslash\{i\}$.*

Proof. Let $\mathcal{A}^{(i)}$ be the affine arrangement of $n - 1$ hyperplanes obtained by setting $y_i = 1$. Let b_j denote the jth column of the matrix B. We consider the $(n - s + 1) \times (n - 1)$ matrix

$$B^{(i)} \;=\; \begin{pmatrix} y_1 & \cdots & y_{i-1} & y_{i+1} & \cdots & y_n \\ b_1 & \cdots & b_{i-1} & b_{i+1} & \cdots & b_n \end{pmatrix}.$$

We also introduce the diagonal matrix $Y^{(i)} = \mathrm{diag}(y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_n)$. Since $u = e_i$, the maximal minors of the matrix in (9.10) are precisely the maximal minors of $B^{(i)}Y^{(i)}$.

After relabeling we may assume $i = n$. Each maximal minor of $B^{(n)}Y^{(n)}$ factors as $\det(B^{(n)}_c)\, y_{c_1} y_{c_2} \cdots y_{c_{n-d+1}}$ where $B^{(n)}_c$ is the submatrix of $B^{(n)}$ with column indices $c = \{c_1, \ldots, c_{n-d+1}\} \subseteq [n-1]$. We claim that the ideal of maximal minors $I_{n-s+1}(B^{(n)}Y^{(n)})$ is radical and that its prime decomposition is given by

$$I_{n-s+1}(B^{(n)}Y^{(n)}) \;=\; \bigcap_J \left[\left(\langle y_j : j \in J\rangle \;+\; I_{n-s+1}(B^{(n)}_{[n-1]\backslash J})\right)\right]. \tag{9.11}$$

Here, the intersection is taken over all flats $J \subseteq [n-1]$ of the affine arrangement $\mathcal{A}^{(n)}$. Since B is generic, there are μ flats, given by the subsets J of cardinality $\leq s-1$. The factorization of minors above shows that the radical of the ideal on the left in (9.11) equals the intersection on the right. In particular, the ideal $I_{n-s+1}(B^{(n)}Y^{(n)})$ in (9.11) has the expected codimension $s-1$. Since it is a determinantal ideal, this implies it is Cohen–Macaulay [42, Theorem 18.18] of degree μ. Therefore, both ideals in (9.11) are Cohen–Macaulay of the same degree and agree up to radical, hence they are equal.

We now impose the $n-s$ linear equations $B \cdot y = 0$ on each of the μ linear spaces of codimension $s-1$ on the right in (9.11). Since B is generic, this has a unique solution $y \in \mathbb{P}^{n-1}$, and this solution satisfies $y_j \neq 0$ for $j \notin J$. More precisely, the coordinates are

$$y_n = 1, \;\; y_j = 0 \;\text{ for }\; j \in J, \quad \text{and} \quad y_{[n-1]\backslash J} \;=\; -B^T_{[n-1]\backslash J}\,(B_{[n-1]\backslash J} B^T_{[n-1]\backslash J})^{-1}\, b_n.$$

For generic B, $y_{[n-1]\backslash J}$ has non-zero coordinates, and $B \cdot y = B_{[n-1]\backslash J} \cdot y_{[n-1]\backslash J} + b_n = 0$.

We have shown that the linear space $\{y \in \mathbb{P}^{n-1} : B \cdot y = 0\}$ intersects the μ irreducible components in distinct reduced points. Therefore, the entries of $B \cdot y$ form a regular sequence modulo $I_{n-s+1}(B^{(n)}Y^{(n)})$. Adding this regular sequence yields a Cohen–Macaulay ideal of Krull dimension one in $\mathbb{R}[y_1, \ldots, y_n]$. This ideal is radical, and it equals the ideal in (9.10). □

Example 9.4.2 ($s = 2, n = 4$)**.** The rank 2 matroid on [4] is uniform for the model given by

$$A^T \;=\; \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 2 & 1 \end{pmatrix} \quad \text{and} \quad B \;=\; \begin{pmatrix} 1 & -2 & 1 & 0 \\ 1 & -1 & 0 & 1 \end{pmatrix}.$$

Theorem 9.4.1 concerns the special data $s = (1, 0, 0, 0)$. The relevant minor of C in (9.7) is

$$\det(B^{(1)}Y^{(1)}) \;=\; \det \begin{pmatrix} y_2^2 & y_3^2 & y_4^2 \\ -2y_2 & y_3 & 0 \\ -y_2 & 0 & y_4 \end{pmatrix}.$$

The radical ideal $I_3(B^{(1)}Y^{(1)}) \;=\; \langle y_2 \rangle \,\cap\, \langle y_3 \rangle \,\cap\, \langle y_4 \rangle \,\cap\, \langle y_2 + 2y_3 + y_4 \rangle$ gives $\mu = 4$ planes in $\mathbb{P}^3$, but two of these planes meet the line $\{y \in \mathbb{P}^3 : B \cdot y = 0\}$ in the same point. So the model is not generic in the sense of Theorem 9.4.1. The likelihood ideal at $u = e_1$ is not radical:

$$I_3(B^{(1)}Y^{(1)}) + \langle B \cdot y \rangle = \langle y_2,\, y_1 + y_4,\, y_3 - y_4 \rangle \cap \langle y_3^2,\, y_1 - y_3 + 2y_4,\, y_2 - y_3 + y_4 \rangle \cap \langle y_4,\, y_1 - y_3,\, y_2 - y_3 \rangle.$$

One checks that this ideal becomes radical for any small perturbation of the model (A, B). ⋄

We now discuss tropical MLE. By Theorem 9.1.1, all solutions to (9.10) are real, and there is one solution in each region of $\mathbb{P}^{s-1}_{\mathbb{R}} \backslash \mathcal{A}$. This statement remains valid for any real closed field, such as the field $R = \mathbb{R}\{\{\epsilon\}\}$ of real Puiseux series. If our data vector u has coordinates in R, the system (9.10) has μ solutions with coordinates in R. The coordinatewise valuation $w = \mathrm{val}(u) \in \mathbb{Q}^n$ of u is the *tropical data vector*. For any solution y of (9.10) in $\mathbb{P}^{n-1}_R$, we also record its valuation $z = \mathrm{val}(y) \in \mathbb{Q}^n/\mathbb{Q}\mathbf{1}$. We present a formula that writes z in terms of w.

Corollary 9.4.3. *Fix a generic model (A, B) and $i \in [n]$. Given data $u \in R^n$, where $w = \mathrm{val}(u)$ satisfies $w_i < w_j$ for $j \in [n] \backslash \{i\}$, the μ tropical critical points z are distinct. There is one for each $J \subseteq [n] \backslash \{i\}$ of cardinality at most $s - 1$, and it is*

$$z \;=\; \sum_{j \in J} (w_j - w_i)\, e_j. \tag{9.12}$$

Before we come to the proof, we go over a detailed example that explains the assertion.

Example 9.4.4 ($s = 3, n = 4$)**.** Let $B = \begin{pmatrix} 1 & 1 & 1 & -1 \end{pmatrix}$. This is the Steiner surface model in Examples 9.0.1 and 9.3.3. Its ML degree is $\mu = 7$. We examine Theorem 9.4.1 for $i = 1$, so the data vector is $u = (1, 0, 0, 0)$. The ideal in (9.10) is radical. Its seven solutions in $\mathbb{P}^3_{\mathbb{R}}$ are

$$\begin{aligned} y(0) \;\;=\;\; & (1:0:0:1),\ (1:0:-1:0),\ (1:-1:0:0), \\ & (2:0:-1:1),\ (2:-1:0:1),\ (2:-1:-1:0),\ (3:-1:-1:1). \end{aligned}$$

We now introduce a small parameter $\epsilon > 0$, and we fix the data vector $u = \big(1, \epsilon^3, \epsilon^4, \epsilon^5\big)$. Each solutions $y(\epsilon) \in \mathbb{P}^3_R$ converges to one of the $y(0)$ above. The seven solutions $y(\epsilon)$ are

$$\begin{gathered} \big(1 \,:\, 2\epsilon^3 + 6\epsilon^6 - 2\epsilon^7 + 2\epsilon^8 \,:\, 2\epsilon^4 - 2\epsilon^7 - 6\epsilon^8 + 2\epsilon^9 \,:\, 1 + 2\epsilon^3 + 2\epsilon^4 - 6\epsilon^6 - 4\epsilon^7\big), \\ \big(1 \,:\, 2\epsilon^3 - 6\epsilon^6 + 2\epsilon^7 - 2\epsilon^8 \,:\, -1 - 2\epsilon^3 - 2\epsilon^5 + 6\epsilon^6 - 2\epsilon^7 \,:\, -2\epsilon^5 + 2\epsilon^8 - 2\epsilon^9 + 6\epsilon^{10}\big), \\ \big(1 \,:\, -1 - 2\epsilon^4 - 2\epsilon^5 - 2\epsilon^7 + 4\epsilon^8 \,:\, 2\epsilon^4 + 2\epsilon^7 - 6\epsilon^8 - 2\epsilon^9 \,:\, -2\epsilon^5 - 2\epsilon^8 + 2\epsilon^9 + 6\epsilon^{10}\big), \\ \big(2 \,:\, 6\epsilon^3 - 48\epsilon^6 + 12\epsilon^7 + 12\epsilon^8 + 816\epsilon^9 \,:\, -1 - 3\epsilon^3 - 3\epsilon^4 + 3\epsilon^5 + 24\epsilon^6 \,:\, 1 + 3\epsilon^3 - 3\epsilon^4 + 3\epsilon^5 - 24\epsilon^6\big), \\ \big(2 \,:\, -1 - 3\epsilon^3 - 3\epsilon^4 + 3\epsilon^5 + 12\epsilon^6 \,:\, 6\epsilon^4 + 12\epsilon^7 - 48\epsilon^8 + 12\epsilon^9 - 12\epsilon^{10} \,:\, 1 - 3\epsilon^3 + 3\epsilon^4 + 3\epsilon^5 + 12\epsilon^6\big), \\ \big(2 \,:\, -1 - 3\epsilon^3 + 3\epsilon^4 - 3\epsilon^5 + 12\epsilon^6 \,:\, -1 + 3\epsilon^3 - 3\epsilon^4 - 3\epsilon^5 - 12\epsilon^6 \,:\, -6\epsilon^5 - 12\epsilon^8 - 12\epsilon^9 + 48\epsilon^{10}\big), \\ \big(3 \,:\, -1 - 8\epsilon^3 + 4\epsilon^4 + 4\epsilon^5 + 72\epsilon^6 \,:\, -1 + 4\epsilon^3 - 8\epsilon^4 + 4\epsilon^5 - 36\epsilon^6 \,:\, 1 - 4\epsilon^3 - 4\epsilon^4 + 8\epsilon^5 + 36\epsilon^6\big). \end{gathered}$$

Each coordinate is a convergent power series in $\mathbb{Z}[\![\epsilon]\!]$. We computed these with the command `puiseux` in `Maple`. Passing to valuations, the tropical data vector is $w = \mathrm{val}(u) = (0, 3, 4, 5)$.

The tropical MLE is given by the valuations of the seven solutions above. We see that

$$\begin{aligned} z \;=\; \mathrm{val}(y(\epsilon)) \;=\; & (0:3:4:0),\ (0:3:0:5),\ (0:0:4:5), \\ & (0:3:0:0),\ (0:0:4:0),\ (0:0:0:5),\ (0:0:0:0). \end{aligned}$$

These are the $\mu = 7$ tropical critical points. These were written in terms of w in (9.12). $\diamond$

Proof of Corollary 9.4.3. We may assume $u_i = 1$ and $w_i = 0$. Then $u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_n$ have positive valuations. We write $w = \sum_{j \in [n]} w_j e_j = \sum_{j \in [n]} \mathrm{val}(u_j) e_j$ for the tropical data. The system (9.10) has μ distinct solutions $y(\epsilon)$ in $\mathbb{P}^{n-1}_R$. Their special fibers $y(0) \in \mathbb{P}^{n-1}_{\mathbb{R}}$ are the zeros of the prime components in (9.11). Hence there are μ distinct tropical critical points.

Suppose that $y(\epsilon)$ is a tropical critical point with valuation $z = \mathrm{val}(y(\epsilon))$ and let $J = \{j \in [n] : z_j > 0\}$. Because $J = \{j \in [n] : (y(0))_j = 0\}$, Theorem 9.4.1 implies that $|J| \leq s - 1$. By the definition of J, $z_j = 0$ if $j \notin J$. Our claim states that $z_j = w_j$ if $j \in J$.

Fix $j \in J$. To prove $z_j = w_j$, we tropicalize a determinantal equation in (9.10). Let c be the union of $\{i, j\}$ with an $(n-s)$-subset of $[n]\backslash(J \cup \{i\})$. The minor of C indexed by s is

$$\prod_{c_i \in c} y_{c_i} \cdot \sum_{k \neq \ell \in c} \pm (u_k/y_k)\, y_\ell \, \det(B_{c\backslash\{k,\ell\}}).$$

The coordinates y_{c_i} of y are nonzero for $c_i \in c$, so the sum vanishes. After tropicalization, the vanishing of the sum becomes the condition that the minimum $\min_{k \neq \ell \in c}(w_k - z_k + z_\ell)$ is attained twice. Since $z_k = 0$ for $k \in c\backslash\{j\}$, we have

$$w_k - z_k + z_\ell \;=\; \begin{cases} w_k & \text{if } k, \ell \neq j, \\ w_j - z_j & \text{if } k = j, \\ w_k + z_j & \text{if } \ell = j. \end{cases}$$

Since $z_j > 0$, the minimum is never attained in the last case. By assumption, $w_k > w_i$. This means that w_k cannot attain the minimum unless $k = i$. Hence the tropical equation simplifies to $\min(w_i, w_j - z_j)$. This yields $w_j - z_j = w_i = 0$, and we conclude that $z_j = w_j$. $\square$

The field R of real Puiseux series is both valued and ordered, and it is interesting to combine both structures when studying the likelihood geometry of squared linear models. In the setting of Corollary 9.4.3, there is one critical point y in each region of $\mathbb{P}^{n-1}_R\backslash\mathcal{A}$. The region is a polytope in $\mathbb{P}^{n-1}_{\mathbb{R}}$, and $y = y(\epsilon)$ is an arc whose limit $y(0)$ lies on one of the faces of that polytope. All arcs are repelled by the hyperplane $\{y_i = 0\}$ at infinity, so they converge to a face of the affine arrangement $\mathcal{A}^{(i)}$. Note that $\mathcal{A}^{(i)}$ has μ flats, while $\mathcal{A}$ has μ regions. The μ critical arcs $y = y(\epsilon)$ therefore specify a bijection between the flats and the regions.

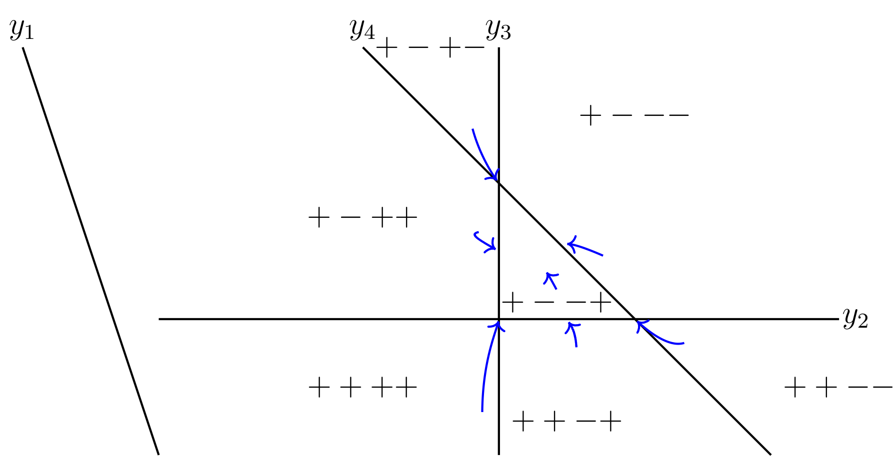

Figure 9.1: Tropical MLE for $s = 3, n = 4$ gives a bijection between the seven regions of $\mathbb{P}^2_{\mathbb{R}}\backslash\mathcal{A}$ and the seven faces of the triangle. Each arc $y(\epsilon)$ travels from its region to the triangle.

Example 9.4.5 ($n = s+1$)**.** Here $\mathcal{A}$ has $\mu = 2^s - 1$ regions. Only one of them is bounded in $\mathcal{A}^{(i)}$. The bounded region is a $(s-1)$-simplex, so it has μ faces. Each region meets the $(s-1)$-simplex in a distinct face, made manifest by an arc $y(\epsilon)$ from the region to that face.

We visualize the case $s = 3$ with $i = 1$ in Figure 9.1. The arrangement $\mathcal{A}$ has four lines and seven regions in $\mathbb{P}^2_{\mathbb{R}}$. The seven arcs $y(\epsilon)$ are given algebraically in Example 9.4.4. Each limit point $y(0)$ lies in the closed triangle: three at the vertices, three on the edges, and one in the interior. The seven tropical solutions z reveal how each arc approaches its limit point. $\diamond$

9.5 Log-Normal Polytopes

The likelihood correspondence can be viewed as the logarithmic normal bundle of the $(s-1)$-dimensional variety $\mathcal{M} \subset \mathbb{P}^{n-1}$. This geometric perspective was emphasized in [70]. Its fiber over a regular point p in $\mathcal{M}_{>0}$ is a linear space of dimension $n-s$ inside the space $\mathbb{P}^{n-1}$ of data u. The *log-normal polytope* is the intersection of this fiber with the simplex $\overline{\Delta}_{n-1} = \mathbb{P}^{n-1}_{\geq 0}$. Each log-normal polytope has dimension $n-s$.

The model is given by a pair (A, B) as in Section 9.4. We fix $p = (y_1^2, \ldots, y_n^2)$ in $\mathcal{M}_{>0}$. Then $y = Ax$ for some $x \in \mathbb{P}^{s-1}_{\mathbb{R}} \backslash \mathcal{A}$, and we have $By = 0$. The rank constraints in (9.10) give $s-1$ independent linear equations in the unknowns $u = (u_1, \ldots, u_n)$. We seek the solutions to these equations in the simplex $\overline{\Delta}_{n-1} = \mathbb{P}^{n-1}_{\geq 0}$. In symbols, the log-normal polytope at y is

$$\Pi(y) \;:=\; \left\{ u \in \mathbb{R}^n_{\geq 0} \,:\, s \text{ satisfies (9.10) and } u_1 + u_2 + \cdots + u_n = 1 \right\}.$$

Thus, $\Pi(y)$ consists of all probability distributions in the row span of the $(n-k+1) \times n$ matrix

$$\begin{pmatrix} y_1^2 & y_2^2 & \cdots & y_n^2 \\ & B \cdot Y & & \end{pmatrix} \;=\; \begin{pmatrix} y_1 & y_2 & \cdots & y_n \\ & B & & \end{pmatrix} \cdot Y \;=\; \begin{pmatrix} 1 & 1 & \cdots & 1 \\ & B \cdot Y^{-1} & & \end{pmatrix} \cdot Y^2. \tag{9.13}$$

Here we abbreviate $Y = \mathrm{diag}(y_1, \ldots, y_n)$. The $\binom{n}{s-1}$ maximal minors of (9.13) factor as

$$y_{j_0} y_{j_1} \cdots y_{j_{n-d}} \cdot \det\begin{pmatrix} y_{j_0} & y_{j_1} & \cdots & y_{j_{n-d}} \\ b_{j_0} & b_{j_1} & \cdots & b_{j_{n-d}} \end{pmatrix} \quad \text{for} \quad 1 \leq y_{j_0} < y_{j_1} < \cdots < y_{j_{n-d}} \leq n. \tag{9.14}$$

These determinants define $\binom{n}{s-1}$ hyperplanes in $\mathbb{P}^{s-1} = \mathbb{P}(\ker(B))$, in addition to the n given hyperplanes $y_i = \ell_i(x)$. The *chamber arrangement* $\mathcal{A}^{\mathrm{ch}}$ consists of all $n + \binom{n}{s-1}$ hyperplanes.

Using this enlarged arrangement, we obtain the following characterization of our polytope.

Theorem 9.5.1. *The polytope $\Pi(y)$ is combinatorially dual to the convex hull Ξ of the columns of the $(n-s) \times n$ matrix BY^{-1}. Each log-normal polytope $\Pi(y)$ is simple for $y \in \mathbb{P}^{s-1}_{\mathbb{R}} \backslash \mathcal{A}^{\mathrm{ch}}$. Its combinatorial type is fixed for y in a region of the chamber arrangement $\mathcal{A}^{\mathrm{ch}}$.*

Proof. Let Ξ^{Δ} denote the polar dual of Ξ. We prove that Ξ^{Δ} is combinatorially equivalent to $\Pi(y)$. Since $y_i^2 > 0$ for all i, removing the factor Y^2 from (9.13) does not change the combinatorial type of the polytope given by its columns. Setting $\tilde{B} = \begin{pmatrix} \mathbf{1} \\ BY^{-1} \end{pmatrix}$, we have

$$\Xi^{\Delta} = \{z \in \mathbb{R}^{n-d} \colon z^T BY^{-1} \le \mathbf{1}\} \simeq \{z \in \mathbb{R}^{n-s+1} \colon z_1 = 1,\ z^T \tilde{B} \ge 0\}, \qquad \text{and}$$

$$\Pi(y) = \{z^T \tilde{B} \in \mathbb{R}^n \colon z^T \tilde{B} \ge 0, \sum_{i=1}^n (z^T \tilde{B})_i = 1\} \simeq \{z \in \mathbb{R}^{n-s+1} \colon z^T \tilde{B} \ge 0, \sum_{i=1}^n (z^T \tilde{B})_i = 1\}.$$

This shows that the cones over Ξ^{Δ} and $\Pi(y)$ are equal. We then apply the argument in the proof of [3, Theorem 4] to conclude that the polytopes are combinatorially equivalent.

The combinatorial type of Ξ is determined by the rank $n-s+1$ oriented matroid of $\tilde{B}$; see [9, Section 9.1]. When y is in $\mathbb{P}^{s-1}_{\mathbb{R}} \backslash \mathcal{A}^{\mathrm{ch}}$, the minors of $\tilde{B}$ are nonzero and hence the underlying matroid of the oriented matroid is uniform; hence Ξ is simplicial. Thus the polar dual Ξ^{Δ} and $\Pi(y)$ are both simple. For y in a fixed region of $\mathbb{P}^{s-1}_{\mathbb{R}} \backslash \mathcal{A}^{\mathrm{ch}}$, the oriented matroid is fixed, since changing the matroid requires crossing one of the hyperplanes. □

Example 9.5.2 $(d = 2, n = 6)$**.** We consider the 1-dimensional model $\mathcal{M}$ in $\mathbb{P}^5$ defined by

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 & 6 \end{pmatrix}^T \quad \text{and} \quad B = \begin{pmatrix} 1 & -2 & 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 1 & -2 & 1 \end{pmatrix}.$$

The chamber arrangement $\mathcal{A}^{\mathrm{ch}}$ consists of 12 points on the circle $\mathbb{P}^1_{\mathbb{R}}$. In addition to the six points given by $y_1, \ldots, y_6$, there are six new points from the determinants in (9.14). For instance, for $\{j_0, \ldots, j_4\} = \{1, 2, 3, 5, 6\}$, we obtain $3y_1 + 2y_2 + y_3 - y_5 - 2y_6 = 3x_1 - 7x_2$, which reveals the transition point $(70 : 30)$ in $\mathcal{A}^{\mathrm{ch}} \subset \mathbb{P}^1_{\mathbb{R}}$. For $x \in \mathbb{P}^1_{\mathbb{R}} \backslash \mathcal{A}^{\mathrm{ch}}$, the log-normal polytope is a product of simplices, namely $\Delta_2 \times \Delta_2$ or $\Delta_1 \times \Delta_3$ or Δ_4, depending on the oriented matroid of (9.13). For an explicit transition, try the points $x = (69 : 30), (70 : 30), (71 : 30)$. ⋄

We now turn to the *log-Voronoi cell* of a point $p = (y_1^2, \ldots, y_n^2)$ in the squared linear model $\mathcal{M}_{>0}$. This is the subset of all data points u in $\Pi(y)$ such that p is the MLE for u. Log-Voronoi cells for discrete models were introduced by Alexandr and Heaton in [4]. Theorems 8, 9 and 10 in [4] identify various models for which each log-Voronoi cell coincides with its ambient log-normal polytope. This holds for all linear models. Their log-Voronoi cells were studied in detail by Alexandr in [3]. However, in general, the log-Voronoi cells are non-linear convex bodies that are strictly contained in their log-normal polytopes. For an illustration see [4, Figure 2]. In what follows, we initiate the study of log-Voronoi cells for squared linear models. We shall see that their geometry is more complicated than that for linear models.

We begin by discussing the simplest model, which is a circle inscribed in a triangle.

Example 9.5.3 $(n = 3, s = 2)$**.** Let $\mathcal{A}$ be the arrangement of three points in $\mathbb{P}^1$ defined by $\ell_1 = x_1$, $\ell_2 = x_2$, $\ell_3 = x_1 + x_2$. The model is the conic $X = V(p_1^2 + p_2^2 + p_3^2 - 2p_1p_2 - 2p_1p_3 - 2p_2p_3)$ in $\mathbb{P}^2$. In the probability triangle $\overline{\Delta}_2$, this is an inscribed circle, so it has three connected components in Δ_2. We see this in Figure 9.2, where the model is drawn in blue.

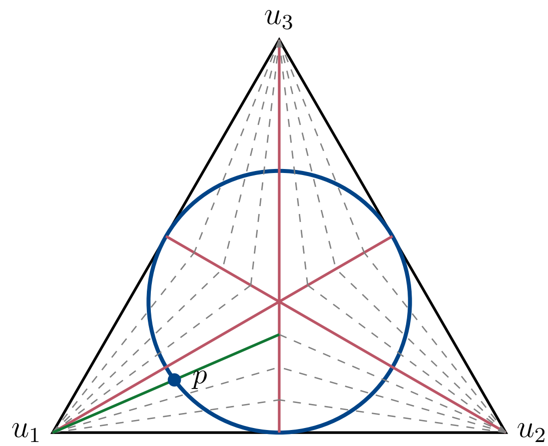

Figure 9.2: The squared linear model (blue) shown inside the triangle Δ_2 of data (black), together with its logarithmic normal bundle (gray dashed lines). The triangle is divided into six Weyl chambers (red lines). The log-Voronoi cell of the point p is the intersection of the fiber of the logarithmic normal bundle with the corresponding Weyl chamber (green line).

Fix $p = \left(x_1^2 : x_2^2 : (x_1+x_2)^2\right) \in \mathcal{M}_{>0}$. The fiber of the log-normal bundle over p is the line

$$au_1 + bu_2 + cu_3 \ := \ \det \begin{pmatrix} u_1 & u_2 & u_3 \\ x_1^2 & x_2^2 & (x_1+x_2)^2 \\ -x_1 & -x_2 & x_1+x_2 \end{pmatrix}.$$

This is a linear form in (u_1, u_2, u_3), whose coefficients depend cubically on the model point

$$a \ = \ x_2(x_1 + 2x_2)(x_1 + x_2), \ \ b \ = \ -x_1(x_1 + x_2)(2x_1 + x_2), \ \ c \ = \ -x_1x_2(x_1 - x_2).$$

The triangle Δ_2 is divided into six Weyl chambers, depending on the ordering of u_1, u_2, u_3. These are the six red triangles in Figure 9.2. The log-Voronoi cell at p is the green line segment through p. It is the intersection of the log-normal line with the Weyl chamber. The log-Voronoi segments interpolate between a red boundary and a half-edge of the triangle.

This example also shows that each log-Voronoi cell is strictly contained in its log-normal polytope. The latter is the intersection of the triangle with the line spanned by the dashed segment. Thus, squared linear models are more complicated than linear models and toric models, for which these $(n-s)$-dimensional convex bodies coincide [4, Theorems 9 and 10]. $\diamond$

The topological boundary ∂S of a log-Voronoi cell S consists of data points with at least one additional MLE with another sign pattern. The boundary ∂S is defined by a piecewise analytic function. We next identify a setting in which the boundary has linear pieces. Given $i, j \in \{1, \ldots, n\}$ and $\sigma \in \{-1, +1\}^n$, let $\tau_{ij\sigma} : \mathbb{P}^{n-1} \to \mathbb{P}^{n-1}$ denote the map exchanging coordinates i and j followed by coordinatewise multiplication by σ.

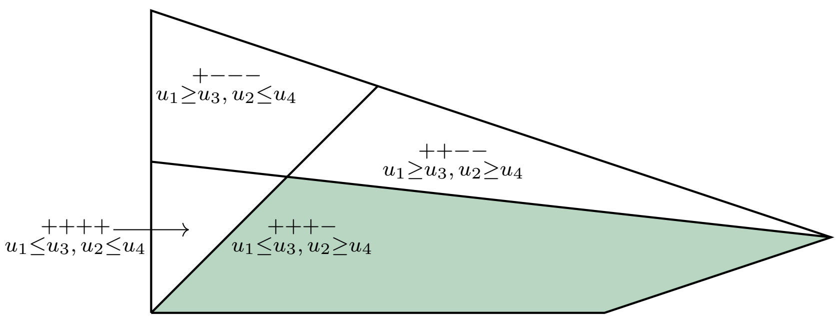

Figure 9.3: The log-normal polygon in Example 9.5.5. The log-Voronoi cell is marked in green.

Proposition 9.5.4. *Fix a point y^2 in a squared linear model $\mathcal{M}$ and its log-Voronoi cell S. If y and $\tau_{ij\sigma}(y)$ have different sign vectors, $\tau_{ij\sigma}(y)^2 \in \mathcal{M}$, and $V(u_i - u_j) \cap S \neq \emptyset$, then $V(u_i - u_j) \cap S$ is a connected, linear piece of the boundary ∂S of the log-Voronoi cell.*

Proof. Let $S' \subset \Pi(y)$ be the set of data points s in the log-normal polytope $\Pi(y)$ such that s has an MLE with sign vector equal to the sign vector of $\tau_{ij\sigma}(y)$. If $s \in S \cap V(u_i - u_j)$, then the likelihood function for s has the same value at y and at $\tau_{ij\sigma}(y)$. Therefore $s \in S \cap S'$. Since S is convex, $S \cap V(u_i - u_j) \subseteq S \cap S' \subseteq \partial S$ is a connected, linear piece of the boundary. $\square$

If the proposition holds for sufficiently many triples i, j, σ, then ∂S is piecewise linear.

Example 9.5.5 ($s = 2, n = 4$)**.** The model in Example 9.4.2 has $\mathcal{A} = \{x_1, x_1 + x_2, x_1 + 2x_2, x_2\}$. Fix $y = (3, 2, 1, -1)$. The log-normal polygon $\Pi(y)$ is the intersection of the tetrahedron Δ_3 with the plane $V(u_1 - u_2 - 3u_3 - 2u_4)$. This is a quadrilateral, divided into four cells based on the sign vector of the MLE; see Figure 9.3. Their boundaries are $u_1 = u_3$ and $u_2 = u_4$, because the vectors $\tau_{13,+++-}(y) = (1, 2, 3, 1)$ and $\tau_{24,+---}(y) = (3, 1, -1, -2)$ are in the kernel of B.

Consider the slightly modified arrangement $\{x_1, x_1 + x_2, x_1 - 2x_2, x_2\}$. The log-normal polytope of $y = (3, 2, 5, -1)$ is also a quadrilateral with four nonempty cells. The log-Voronoi cell meets all other cells in a codimension 1 boundary. This boundary is nonlinear. $\diamond$

We found that the log-Voronoi cells of squared linear models exhibit a wide range of behaviors. Experiments suggest that, for a generic model and point, the boundary of the log-Voronoi cell is nonlinear. As in Section 9.4, being generic is stronger than having a uniform matroid. Example 9.5.5 underscores this. As in [4, page 9], we think that these boundaries are generally not algebraic. Experiments indicate that it is also possible to have linear and nonlinear boundary pieces at the same time. Finally, it is possible that all boundaries are linear, so the log-Voronoi cell is a polytope (see Figure 9.3). In conclusion, the detailed study of log-Voronoi cells for squared linear models is a promising direction for future research.

9.6 Revisiting Determinantal Point Processes

In this final section we show how our squared linear models appear in the applied context of Chapter 8. We characterize the squared linear models that can be realized as projection DPPs.

Example 9.6.1. The braid arrangement in Example 9.0.2 may be realized as a projection DPP whose state space is 2-subsets of $[n]$. The corresponding parameter matrix Θ is

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \end{pmatrix}. \qquad \diamond$$

We now consider the submodel given by a linear subspace $\mathcal{L}$ of the Plücker space $\mathbb{P}^{\binom{n}{k}} - 1$ which is contained in the Grassmannian $\mathrm{Gr}(k, n)$. For instance, we can take $\mathcal{L}$ to be the set of all k-planes in $\mathbb{R}^n$ that contain a fixed $(k-1)$-plane. Or, dually, we can consider all k-planes in $\mathbb{R}^n$ that lie in a fixed $(k+1)$-plane. In terms of the parametrization above, we take Θ to be a matrix whose last row consists of unknowns $\theta_1, \theta_2, \ldots, \theta_n$ and whose other entries are fixed real numbers $a_{i,j}$. We refer to this model as a *linear projection DPP*, and we denote it by $\mathcal{L}^2 \subset \mathrm{sGr}(k, n)$. We say that $\mathcal{L}^2$ is *generic* if the numbers $a_{i,j}$ are chosen generically.

Using row operations, we can eliminate $k-1$ of the model parameters, and we can write

$$\Theta = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n-k+1} & 1 & 0 & \cdots & 0 \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n-k+1} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & 0 & 0 & \ddots & 0 \\ a_{k-1,1} & a_{k-1,2} & \cdots & a_{k-1,n-k+1} & 0 & 0 & \cdots & 1 \\ \theta_1 & \theta_2 & \cdots & \theta_{n-k+1} & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

The probabilities in the model $\mathcal{L}^2$ are the squares of all maximal minors of Θ. Equivalently, the probabilities are the squares of maximal minors containing the last column in the matrix

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & -a_{1,1} & -a_{2,1} & \cdots & -a_{k-1,1} & \theta_1 \\ 0 & 1 & \cdots & 0 & -a_{1,2} & -a_{2,2} & \cdots & -a_{k-1,2} & \theta_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{1,n-k+1} & -a_{2,n-k+1} & \cdots & -a_{k-1,n-k+1} & \theta_{n-k+1} \end{pmatrix}. \tag{9.15}$$

These $\binom{n}{k}$ minors are precisely the hyperplanes spanned by any $n-k$ of the first n columns.

Proposition 9.6.2. *A squared linear model is a linear projection DPP if and only its hyperplane arrangement $\mathcal{A}$ is a* discriminantal arrangement*. This means that the hyperplanes in $\mathcal{A}$ are those spanned by subsets of the n points in $\mathbb{P}^{n-k}$ seen in the first n columns in (9.15).*

Let us consider generic models $\mathcal{L}^2$ where $\ell = n - k$ is fixed. These are given by the $\binom{n}{\ell}$ hyperplanes which are spanned by all ℓ-subsets of n generic points in $\mathbb{P}^\ell$. It is known that the number of regions of this discriminantal arrangement is a polynomial in n of degree ℓ^2 [1].

Corollary 9.6.3. *The ML degree of the generic linear projection DPP is a polynomial in n of degree ℓ^2. For instance, if $\ell = 2$ then the ML degree of the generic model $\mathcal{L}^2$ equals*

$$\frac{1}{8}(n-1)(n^3 - 5n^2 + 14n - 8). \tag{9.16}$$

Proof. The first sentence is a direct consequence of [1, Theorem 3.3]. In [78, p. 793] the authors compute the characteristic polynomial $\chi_{\mathcal{A}}(t)$ for the *non-central* discriminantal arrangement. Then $\chi_{\mathcal{A}}(1)$ is the number of bounded regions in $\mathbb{R}^n \backslash \mathcal{A}$, while $\chi_{\mathcal{A}}(-1)$ is the number of unbounded regions in $\mathbb{R}^n \backslash \mathcal{A}$. To obtain the total number of regions in $\mathbb{P}^{n-1}_{\mathbb{R}} \backslash \mathcal{A}$, i.e., the ML degree, one then computes $\frac{1}{2}(\chi_{\mathcal{A}}(-1) + \chi_{\mathcal{A}}(1))$. This is the polynomial (9.16). $\square$

Example 9.6.4 ($n = 6, k = 4$)**.** Let Θ be a 4×6 matrix whose first three rows are random real numbers and whose last row consists of parameters $\theta_1, \theta_2, \ldots, \theta_6$. The probability of any quadruple in [6] is the squared determinant of the corresponding maximal minor of Θ, divided by the sum of all 15 squared minors. The variety $\mathcal{L}^2$ is a 3-dimensional Veronese surface inside the 8-dimensional squared Grassmannian $\mathrm{sGr}(4, 6) \subseteq \mathbb{P}^{14}$. The ML degree of this DPP model equals 70. This the value of (9.16) for $n = 6$. It counts the regions in the discriminantal arrangement for six points in $\mathbb{P}^2$. These points are derived by the row operations on Θ that yield (9.15). All complex critical points are real, and there is one critical point in each region. $\diamond$

In this chapter, we initiated the study of the likelihood geometry of squared linear models. We showed that squared linear models are isomorphic to projections of Veronese varieties, that all critical points of maximum likelihood estimation are real and positive, and we gave an explicit description of the parametric likelihood ideal. We gave a formula for the tropical maximum likelihood estimate, and described the log-normal polytopes of squared linear models. We concluded by characterizing the squared linear models that are projection DPPs.

In the last three chapters, we have conducted a study of the likelihood geometry of determinantal point processes. We showed that the likelihood equations of L-ensembles have a recursive structure, and that the critical points of maximum likelihood estimation are all real and positive for projection DPPs and squared linear models. In addition, we gave formulas for the ML degrees of the squared Grassmannian and the Grassmannian.

In this thesis, we studied optimization problems on Grassmannians in a variety of contexts. For our objective functions, we primarily used general linear functions and the log-likelihood function. There are many avenues for further work, including the study of discriminants of optimization problems. Several chapters focused exclusively on the Grassmannian of lines $\mathrm{Gr}(2, n)$, and should be extended to other Grassmannians. It would also be interesting to extend this research to flag varieties. One more important direction to further develop metric algebraic geometry for the Grassmannian is the study of distance minimization using other metrics.

Bibliography

[1] D. Agostini, T. Brysiewicz, C. Fevola, L. Kühne, B. Sturmfels and S. Telen: *Likelihood degenerations*, Advances in Mathematics **414** (2023) 108863.

[2] A. Al Ahmadieh and C. Vinzant: *Characterizing principal minors of symmetric matrices via determinantal multiaffine polynomials*, Journal of Algebra **638** (2024) 255–278.

[3] Y. Alexandr: *Logarithmic Voronoi polytopes for discrete linear models*, Algebraic Statistics **15** (2024) 1–13.

[4] Y. Alexandr and A. Heaton: *Logarithmic Voronoi cells*, Algebraic Statistics **12** (2021) 75–95.

[5] F. Ardila-Mantilla, C. Eur and R. Penaguiao: *The tropical critical points of an affine matroid*, SIAM Journal on Discrete Mathematics **38** (2024) 1930–1942.

[6] M.F. Atiyah and I.C. MacDonald: *Introduction to Commutative Algebra*, Addison-Wesley Series in Mathematics, Westview Press, Boulder, CO, 2016.

[7] R. Bettiol and H. Friedman: *Counting Homogeneous Einstein Metrics*, `arXiv:2509.09830`.

[8] J. Bezanson, A. Edelman, S. Karpinski and V. B. Shah: *Julia: A fresh approach to numerical computing*, SIAM Review **59** (2017) 65–98.

[9] A. Björner, M. Las Vergnas, B. Sturmfels, N. White and G. Ziegler: *Oriented Matroids*, Encyclopedia of Mathematics and its Applications, Vol. 46, Cambridge University Press, Cambridge, 1999.

[10] A. Bloch and S. Karp: *Gradient flows, adjoint orbits, and the topology of totally nonnegative flag varieties*, Communications in Mathematical Physics **398** (2023) 1213–1289.

[11] E. Boniface, K. Devriendt and S. Hoşten: *Tropical toric maximum likelihood estimation*, SIAM Journal on Applied Algebra and Geometry **9** (2025) 525–553.

[12] A. Borodin: *Determinantal point processes* in *The Oxford Handbook of Random Matrix Theory*, Oxford University Press, Oxford, 2011, 231–249.

[13] A. Borodin, P. Diaconis and J. Fulman: *On adding a list of numbers (and other one-dependent point processes)*, Bulletin of the American Mathematical Society **47** (2010) 639–670.

[14] V. Borovik: *Flat Degenerations and Polynomial System Solving in Physics and Chemistry*, PhD thesis, Universität Osnabrück, 2025.

[15] V. Borovik and P. Breiding: *A short proof for the parameter continuation theorem*, Journal of Symbolic Computation **127** (2025) 102373.

[16] V. Borovik, H. Friedman, S. Hoşten and M. Pfeffer: *Numerical algebraic geometry for energy computations on tensor train varieties*, `arXiv:2512.06939`.

[17] V. Borovik, H. Friedman, S. Hoşten and M. Pfeffer: *Numerical algebraic geometry for energy computations on tensor train varieties: Supplementary material*, `doi:10.5281/zenodo.17777072`, 2025.

[18] M. Brandt, J. Bruce, T. Brysiewicz, R. Krone and E. Robeva: *The degree of* $\mathrm{SO}(n, \mathbb{C})$, Fields Institute Communications **80** (2017) 229–246.

[19] P. Breiding, K. Kohn and B. Sturmfels: *Metric Algebraic Geometry*, Oberwolfach Seminars, Vol. 53, Birkhäuser/Springer, Cham, 2024.

[20] P. Breiding and S. Timme: *HomotopyContinuation.jl: A package for homotopy continuation in Julia*, Math. Software – ICMS 2018, 458–465, Springer International Publishing (2018).

[21] P. Breiding, K. Rose and and S. Timme: *Certifying zeros of polynomial systems using interval arithmetic*, Association for Computing Machinery. Transactions on Mathematical Software **49** (2023) 11.

[22] P. Breiding, B. Sturmfels and K. Wang. *Computing arrangements of hypersurfaces*, Journal of Software for Algebra and Geometry **15** (2025) 11–27.

[23] V.-E. Brunel, A. Moitra, P. Rigollet and J. Urschel: *Maximum likelihood estimation of determinantal point processes*, `arXiv:1701.06501`.

[24] T. Brysiewicz, H. Eble and L. Kühne: *Computing characteristic polynomials of hyperplane arrangements with symmetries*, Discrete and Computational Geometry **70** (2023) 1356–1377.

[25] T. Brysiewicz and F. Gesmundo. *The degree of Stiefel manifolds*, Enumerative Combinatorics and Applications **1** (2021) S2R20, 18.

[26] W. Buczyńska, J. Buczyński and M. Michałek. *The Hackbusch conjecture on tensor formats*, Journal de Mathématiques Pures et Appliquées **104** (2015) 749–761.

[27] F. Catanese, S. Hoşten, A. Khetan and B. Sturmfels: *The maximum likelihood degree*, American Journal of Mathematics **128** (2006) 671–697.

[28] O. Clarke, S. Hoşten, N. Kushnerchuk and J. Oldekop: *Matroid stratifications of ML degrees of independence models*, Algebraic Statistics **15** (2024) 199–223.

[29] D.H. Collingwood and W.M. McGovern: *Nilpotent Orbits in Semisimple Lie Algebras*, Van Nostrand Reinhold Mathematics Series, Van Nostrand Reinhold Co., New York, 1993.

[30] J. Conway, R. Hardin and N. Sloane: *Packing lines, planes, etc.: Packings in Grassmannian spaces*, Experimental Mathematics **5** (1996) 139–159.

[31] C.M. Cuadras and M. Greenacre: *A short history of statistical association: from correlation to correspondence analysis to copulas*, Journal of Multivariate Analysis **188** (2022) 104901.

[32] C. De Concini and C. Procesi: *A characteristic free approach to invariant theory*, Advances in Mathematics **21** (1976) 330–354.

[33] J. De Loera, J. Rambau and F. Santos: *Triangulations*, Algorithms and Computation in Mathematics, Vol. 25, Springer-Verlag, Berlin, 2010.

[34] K. Devriendt, H. Friedman, B. Reinke and B. Sturmfels: *The two lives of the Grassmannian* Acta Universitatis Sapientiae, Mathematica **17** (2025) 8.

[35] P. Dey, P. Görlach and N. Kaihnsa: *Coordinate-wise powers of algebraic varieties*, Beiträge zur Algebra und Geometrie **61** (2020) 473–505.

[36] D. Đoković, K. Rietsch and K. Zhao: *Normal forms for orthogonal similarity classes of skew-symmetric matrices*, Journal of Algebra **308** (2007) 686–703.

[37] J. Draisma, E. Horobeţ, G. Ottaviani, B. Sturmfels, and R. Thomas: *The Euclidean distance degree of an algebraic variety*, Foundations of Computational Mathematics **16** (2016) 99–149.

[38] P. Drineas, M. Mahoney and S. Muthukrishnan: *Subspace sampling and relative-error matrix approximation: column-row-based methods* in *Algorithms—ESA 2006*, Lecture Notes in Computer Science, Vol. 4168, Springer, Berlin, 2006, 304–314.

[39] M. Drton, B. Sturmfels and S. Sullivant: *Lectures on Algebraic Statistics*, Oberwolfach Seminars, Vol. 39, Birkhäuser Verlag, Basel, 2009.

[40] T. Duff, C. Hill, A. Jensen, K. Lee, A. Leykin and J. Sommars: *Solving polynomial systems via homotopy continuation and monodromy*, IMA Journal of Numerical Analysis **39** (2018) 1421–1446.

[41] N. Early, A. Geiger, M. Panizzut, B. Sturmfels and C. Yun: *Positive del Pezzo geometry*, Combinatorics, Physics and their Interactions **12** (2025) 819–868.

[42] D. Eisenbud: *Commutative Algebra*, Graduate Texts in Mathematics, Vol. 150, Springer-Verlag, New York, 1995.

[43] A. Falcó, W. Hackbusch and A. Nouy: *Geometry of tree-based tensor formats in tensor Banach spaces*, Annali di Matematica Pura ed Applicata. Series IV **202** (2023) 2127–2144.

[44] F. Faulstich and M. Oster: *Coupled cluster theory: Toward an algebraic geometry formulation*, SIAM Journal on Applied Algebra and Geometry **8** (2024) 138–188.

[45] F. Faulstich, B. Sturmfels and S. Sverrisdóttir: *Algebraic varieties in quantum chemistry*, Foundations of Computational Mathematics **25** (2024) 1167–1198.

[46] H. Friedman: *Likelihood geometry of the squared Grassmannian*, Proceedings of the American Mathematical Society **153** (2025) 4463–4474.

[47] H. Friedman and S. Hoşten: *Grassmann and flag varieties in linear algebra, optimization, and statistics: an algebraic perspective*, `arXiv:2505.15969`.

[48] H. Friedman, A. Rosana and B. Sturmfels: *Distance optimization in the Grassmannian of lines*, `arXiv:2601.22843`.

[49] H. Friedman, B. Sturmfels and M. Wiesmann: *Squared linear models*, SIAM Journal on Applied Algebra and Geometry **10** (2026) 153–173.

[50] H. Friedman, B. Sturmfels and M. Zubkov: *Likelihood geometry of determinantal point processes* Algebraic Statistics **15** (2024) 15–25.

[51] W. Fulton: *Intersection Theory*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Vol. 2, Springer-Verlag, Berlin, 1998.

[52] W. Fulton: *Young Tableaux*, Mathematical Society Student Texts, Vol. 35, Cambridge University Press, Cambridge, 1996.

[53] F.R. Gantmacher: *The Theory of Matrices*, Vols. 1–2, Chelsea Publishing, New York, 1959.

[54] M. Garstka, M. Cannon and P. Goulart: *COSMO: A conic operator splitting method for convex conic problems*, Journal of Optimization Theory and Applications **190** (2021) 779–810.

[55] I.M. Gel′fand, M.M. Kapranov and A.V. Zelevinsky: *Discriminants, resultants, and multidimensional determinants*, Mathematics: Theory & Applications, Birkhäuser Boston, Inc., Boston, MA, 1994.

[56] K. Gilman, D. Tarzanagh and L. Bolzano: *Grassmannian optimization for online tensor completion and tracking with the t-SVD*, IEEE Transactions on Signal Processing **70** (2022) 2152–2167.

[57] U. Görtz and T. Wedhorn: *Algebraic Geometry I. Schemes—with Examples and Exercises*, Springer Studium Mathematik—Master, Springer Spektrum, Wiesbaden, 2020.

[58] L. Grasedyck: *Hierarchical singular value decomposition of tensors*, SIAM Journal on Matrix Analysis and Applications **31** (2010) 2029–2054.

[59] E. Grigorescu, B. Juba, K. Wimmer and N. Xie: *Hardness of maximum likelihood learning of DPPs*, Proceedings of 35^{th} Conference on Learning Theory, PMLR **178** (2022) 3800–3819.

[60] D. Grayson and M. Stillman: *Macaulay2, a software system for research in algebraic geometry*, available at `https://macaulay2.com`.

[61] S. Heinen, M. Schwilk, G. F. von Rudorff and O. A. von Lilienfeld: *Machine learning the computational cost of quantum chemistry*, Machine Learning: Science and Technology **1** (2020) 025002.

[62] T. Helgaker, P. Jorgensen and J. Olsen: *Molecular Electronic-Structure Theory*. John Wiley & Sons, Ltd, 2000.

[63] B. Hollering, E. Mazzucchelli, M. Parisi and B. Sturmfels: *Varieties of lines in 3-space*, `arXiv:2511.21333`.

[64] S. Holtz, T. Rohwedder and R. Schneider: *The alternating linear scheme for tensor optimization in the tensor train format*, SIAM Journal on Scientific Computing **34** (2012) A683–A713.

[65] S. Holtz, T. Rohwedder and R. Schneider: *On manifolds of tensors of fixed TT-rank*, Numerische Mathematik **120** (2012) 701–731.

[66] O. Holtz and B. Sturmfels: *Hyperdeterminantal relations among symmetric principal minors*, Journal of Algebra **316** (2007) 634–648.

[67] J.B. Hough, M. Krishnapur, Y. Peres and B. Virág: *Determinantal processes and independence*, Probability Surveys **3** (2005) 206–229.

[68] S. Hoşten, A. Khetan and B. Sturmfels: *Solving the likelihood equations*, Foundations of Computational Mathematics **5** (2005) 389–407.

[69] J. Huh: *The maximum likelihood degree of a very affine variety*, Compositio Mathematica **149** (2013) 1245–1266.

[70] J. Huh and B. Sturmfels: *Likelihood geometry* in *Combinatorial Algebraic Geometry*, Lecture Notes in Mathematics, Vol. 2108, Springer, Cham, 2014, 63–117.

[71] K. Ikramov and H. Fassbender: *On the product of two skew-Hamiltonian or two skew-symmetric matrices*, Journal of Mathematical Sciences **157** (2009) 697–700.

[72] N. V. Ilyushechkin: *The discriminant of the characteristic polynomial of a normal matrix* Matematicheskie Zametki **51** (1992) 16–23.

[73] D. Jaksch and P. Zoller: *The cold atom Hubbard toolbox*, Annals of Physics **315** (2005) 52–79.

[74] T. Kahle, L. Kühne, L. Mühlherr, B. Sturmfels and M. Wiesmann: *Arrangements and likelihood*, Vietnam Journal of Mathematics (to appear).

[75] T. Kahle, H. Schenck, B. Sturmfels and M. Wiesmann: *The likelihood correspondence*, Foundations of Computational Mathematics (to appear).

[76] A. Kassel and T. Lévy: *Determinantal probability measures on Grassmannians*, Annales de l'Institut Henri Poincaré D. Combinatorics, Phsyics and their Interactions **9** (2022) 659–732.

[77] L. Kayser, A. Kretschmer and S. Telen: *Logarithmic discriminants of hyperplane arrangements*, Le Matematiche **80** (2025) 325–346.

[78] H. Koizumi, Y. Numata and A. Takemura: *On intersection lattices of hyperplane arrangements generated by generic points*, Annals of Combinatorics **16** (2012) 789–813.

[79] H. Kraft and C. Procesi: *On the geometry of conjugacy classes in classical groups*, Commentarii Mathematici Helvetici **57** (1982) 539–602.

[80] A. Kulesza: *Learning with Determinantal Point Processes*, PhD Thesis, University of Pennsylvania, 2012.

[81] A. Kulesza and B. Taskar: *Determinantal point processes for machine learning*, Foundations and Trends in Machine Learning **5** (2012) 123–286.

[82] Z. Lai, L.-H. Lim and K. Ye: *Euclidean distance degree in manifold optimization* SIAM Journal on Optimization **35** (2025) 2402–2422.

[83] Z. Lai, L.-H. Lim and K. Ye: *Simpler Grassmannian optimization*, `arXiv:2009.13502`.

[84] J.M. Landsberg: *Tensors: Geometry and Applications*, Graduate Studies in Mathematics, Vol. 128, American Mathematical Society, Providence, RI 2012.

[85] B. Legat, C. Coey, R. Deits, J. Huchette and A. Perry: *Sum-of-squares optimization in Julia* in *The First Annual JuMP-dev Workshop*, 2017.

[86] A. Lerario and A. Rosana: *The Grassmann distance complexity*, `arXiv:2411.16589`.

[87] L.-H. Lim, K.S. Wong and K. Ye: *Numerical algorithms on the affine Grassmannian*, SIAM Journal on Matrix Analysis and Applications **40** (2019) 371–393.

[88] L.-H. Lim and K. Ye: *Degree of the Grassmannian as an affine variety*, Advances in Mathematics **480** (2025) 1439–1962.

[89] L.-H. Lim and K. Ye: *Simple matrix models for the flag, Grassmann, and Stiefel manifolds*, Foundations of Comptuational Mathematics (to appear).

[90] L. Lin and J. Lu: *A Mathematical Introduction to Electronic Structure Theory*, Society for Industrial and Applied Mathematics, Philadelphia, 2019.

[91] M. Lubin, O. Dowson, J. Dias Garcia, J. Huchette, Benoît Legat and J.P. Vielma: *JuMP 1.0: Recent improvements to a modeling language for mathematical optimization*, Mathematical Programming Computation **15** (2023) 581–589.

[92] R. Lyons: *Determinantal probability measures*, Publications Mathématiques de l'IHÉS **98** (2003) 167–212.

[93] O. Macchi: *The coincidence approach to stochastic point processes*, Advances in Applied Probability **7** (1975) 83–122.

[94] K.V. Mardia, J.T. Kent and J.M. Bibby: *Multivariate Analysis*, Probability and Mathematical Statistics: A Series of Monographs and Textbooks, Academic Press, London-New York-Toronto, 1979.

[95] H. Matsumura: *Commutative Ring Theory*, translated by Miles Reid, Cambridge studies in Advance Mathematics, Vol. 8, Cambridge University Press, Cambridge, 1986.

[96] L.G. Maxim, J.I. Rodriguez, B. Wang, and L. Wu: *Linear optimization on varieties and Chern-Mather classes*, Advances in Mathematics **437** (2024) 109443, 22.

[97] M. Michałek and B. Sturmfels: *Invitation to Nonlinear Algebra*, Graduate Studies in Mathematics, Vol. 211, American Mathematical Society, Providence, RI 2021.

[98] A.P. Morgan and A.J. Sommese: *Coefficient-parameter polynomial continuation*, Applied Mathematics and Computation **29** (1989) 123–160.

[99] J. Nie and K. Ranestad: *Algebraic degree of polynomial optimization*, SIAM Journal on Optimization **20** (2009) 485–502.

[100] J. Nie, K. Ranestad and B. Sturmfels: *The algebraic degree of semidefinite programming*, Mathematical Programming **122** (2010) 379–405.

[101] L. Oeding: *Set-theoretic defining equations of the variety of principal minors of symmetric matrices*, Algebra & Number Theory **5** (2011) 75–109.

[102] OSCAR – Open Source Computer Algebra Research system, Version 1.3.1, The OSCAR Team, 2025. `https://www.oscar-system.org`.

[103] I. Oseledets: *Tensor-train decomposition*, SIAM Journal on Scientific Computing **33** (2011) 2295–2317.

[104] I. Oseledets and E. Tyrtyshnikov: *TT-cross approximation for multidimensional arrays*, Linear Algebra and its Applications **432** (2010) 70–88.

[105] H. Oviedo: *Implicit steepest descent algorithm for optimization with orthogonality constraints*, Optimization Letters **16** (2022) 1773–1797.

[106] J. Ponce, B. Sturmfels and M. Trager: *Congruences and concurrent lines in multi-view geometry*, Advances in Applied Mathematics **88** (2017) 62–91.

[107] K. Ranestad: *Algebraic degree in nonlinear models* in *Combinatorial, Computational, and Applied Algebraic Geometry: A Tribute to Bernd Sturmfels*, Proceedings of Symposia in Pure Mathematics, Vol. 111, American Mathematical Society, Providence, RI 2025, 287–305.

[108] K. Ranestad: *Algebraic degree in semidefinite and polynomial optimization*, in *Handbook on Semidefinite, Conic and Polynomial Optimization*, International Series Operations Research & Management Science, Vol. 166, Springer, New York, 2012, 61–75.

[109] B. Reinke and K. Wang: *Hypersurface arrangements with generic hypersurfaces added*, `arXiv:2412.20869`.

[110] G. Restuccia and G. Rinaldo: *On certain classes of degree reverse lexicographic Gröbner bases*, International Mathematical Forum **2** (2007) 1053–1068.

[111] K. Sadeghi and A. Rinaldo: *Markov properties of discrete determinantal point processes* in *Proceedings of the Twenty-Second International Conference on Artificial Intelligence and Statistics*, Vol. 89, PMLR, 2019, 1313–1321.

[112] F. Salizzoni, L. Sodomaco and J. Weigert: *Nonlinear Kalman varieties*, `arXiv:2512.16540`.

[113] F. Salizzoni, L. Sodomaco and J. Weigert: *Nonlinear Rayleigh quotient optimization*, `arXiv:2510.17760`.

[114] U. Schollwöck: *The density-matrix renormalization group in the age of matrix product states*, Annals of Physics **326** (2011) 96–192.

[115] N. H. Scott: *A theorem on isotropic null vectors and its application to thermoelasticity*, Proceedings of the Royal Society **440** (1993) 431–442.

[116] T. Seynnaeve, J. Jovcheva and N. Vannieuwenhoven: *Minimality of tree tensor network ranks*, `arXiv:2509.09463`.

[117] I. Shafarevich: *Basic Algebraic Geometry. 1*, Springer-Verlag, 1994.

[118] S. Sivasubramanian: *Signed excedance enumeration via determinants*, Advances in Applied Mathematics **47** (2011) 783–794.

[119] A.J. Sommese and C.W. Wampler: *The Numerical Solution of Systems of Polynomials Arising in Engineering and Science*, World Scientific, 2005.

[120] H. Stenzel: *Über die Darstellbarkeit einer Matrix als Produkt von zwei symmetrischen Matrizen, als Produkt von zwei alternierenden Matrizen und als Produkt von einer symmetrischen und einer alternierenden Matrix*, Mathematische Zeitschrift **15** (1922) 1–25.

[121] B. Sturmfels and S. Telen: *Likelihood equations and scattering amplitudes*, Algebraic Statistics **12** (2021) 167–186.

[122] S. Sullivant: *Algebraic Statistics*, Graduate Studies in Mathematics, Vol. 194, American Mathematical Society, Providence, RI, 2018.

[123] S. Szalay, M. Pfeffer, V. Murg, G. Barcza, F. Verstraete, R. Schneider, and Ö. Legeza: *Tensor product methods and entanglement optimization for ab initio quantum chemistry*, International Journal of Quantum Chemistry **115** (2015) 1342–1391.

[124] R.C. Thompson: *Principal submatrices. IX. Interlacing inequalities for singular values of submatrices*, Linear Algebra and its Applications **5** (1972) 1–12.

[125] L.N. Trefethen and D. Bau: *Numerical Linear Algebra*, Society of Industrial and Applied Mathematics, Philadelphia, PA, 1997.

[126] E. Tyrtyshnikov: *Incomplete cross approximation in the mosaic-skeleton method*, Computing **64** (2000) 367–380.

[127] M. Weinstein: *Real symmetric matrices with partitioned eigenvalues*, Linear Algebra and its Applications **633** (2022) 281–289.

[128] S.R. White: *Density matrix formulation for quantum renormalization groups*, Physical Review Letters **69** (1992) 2863–2866.

[129] L. Williams: *The positive Grassmannian, the amplituhedron, and cluster algebras*, in *ICM–International Congress of Mathematicians*, Vol. 6, EMS Press, Berlin, 2023, 4710–4737.

[130] K. Ye and L.-H. Lim: *Schubert varieties and distances between subspaces of different dimensions*, SIAM Journal on Matrix Analysis and Applications **37** (2016) 1176–1197.

[131] K. Ye, K.S.-W. Wong and L.-H. Lim. *Optimization on flag manifolds*, Mathematical Programming **194** (2022) 621–660.