SISSA
Scuola
Internazionale
Superiore di
Studi Avanzati

Mathematics Area - PhD course in
Geometry and Mathematical Physics

Metric and Probabilistic Aspects of Grassmann and Tensor Geometry

Candidate:
Andrea Rosana

Advisor:
Antonio Lerario

Academic Year 2024-25

Abstract

Metric Algebraic Geometry is an emerging field that connects algebraic and differential geometry through the study of metric properties of real algebraic varieties, motivated by applications to the sciences. In this thesis we address three different problems in tensor and Grassmann geometry that showcase the interplay between differential, analytic and algebraic aspects typical of Metric Algebraic Geometry.

The first problem deals with estimating the probability that a random real symmetric tensor is close to rank–one tensors. We show how this problem is equivalent to estimating the volume of a neighborhood of the real Veronese variety. We study metric invariants of the Veronese variety and give explicit formulae for its reach and curvature coefficients with respect to the Bombieri–Weyl metric. A description of its second fundamental form is also obtained in terms of matrices from the Gaussian Orthogonal Ensemble. We give a closed formula for the probability of being close to rank–one and show that in the case of rational normal curves it has an exponential decay with respect to the order of the tensor.

In the second part we study typical ranks of real $m \times n \times \ell$ tensors. We introduce a new geometric framework through which we express rank probabilities for Gaussian tensors in terms of the probabilities of having enough points in the intersection between a random linear subspace and the Segre variety of rank–one matrices. We show that in the range $(m-1)(n-1)+1 \leq \ell \leq mn$ typical ranks are contained in $\{\ell, \ell+1\}$ and ℓ is always typical. Based on asymptotic results on the average number of real points in a random slice of the Segre variety with a complementary dimension subspace, we give some heuristics on how the rank probabilities behave. For $3 \times 3 \times 5$ tensors, for which typical ranks are known to be 5 and 6, we relate the rank probabilities to the probability that a random determinantal cubic surface contains only real lines. As a by–product, we obtain bounds on the expected number of real lines on such a cubic.

The last problem constitutes the main project carried out during my PhD studies. Motivated by the notion of Euclidean Distance Degree, which is a measure of the complexity of the nearest point problem to an algebraic variety in a Euclidean space, we introduce the concept of *Grassmann Distance Complexity* (GDC) of a subanalytic set in the Grassmannian. This measures the complexity of solving the nearest point problem when the subset is subanalytic and the distance used is the Riemannian one. As this distance is not smooth nor semialgebraic, we employ tools from subgradient theory for locally Lipschitz functions and o–minimial geometry. We prove key properties of the GDC, such as its finiteness, and compute explicit bounds for real algebraic hypersurfaces. In the last part we explicitly solve the nearest point problem for a class of simple Schubert varieties through a non–linear version of Eckart–Young theorem.

The research we present in this thesis can be found in the following publications and preprints:

- A. Cazzaniga, A. Lerario, A. Rosana, *What Is the Probability That a Random Symmetric Tensor Is Close to Rank–One?*, SIAM Journal on Applied Algebra and Geometry, 8(2):227-258, 2024.
- P. Breiding, S. Eggleston, A. Rosana, *Typical Ranks of Random Order–Three Tensors*, International Mathematics Research Notices, Volume 2025, Issue 4, 2025.
- A. Lerario, A. Rosana, *The Grassmann distance complexity*, arxiv:2411.16589

Contents

1 Introduction **1**

1.1 Symmetric tensors close to rank–one . . . 2

1.1.1 Overview of main contributions and sections . . . 2

1.1.2 Statement of the results . . . 3

1.2 Typical ranks of order–three tensors . . . 11

1.2.1 Overview of main contributions and sections . . . 13

1.2.2 Statement of the results . . . 14

1.3 Nearest point problems in the Grassmannian . . . 17

1.3.1 Overview of main contributions and sections . . . 18

1.3.2 Statement of the results . . . 20

2 Background material **27**

2.1 The second fundamental form and Weingarten operator . . . 27

2.2 Tubular neighborhoods . . . 29

2.3 The reach of a submanifold . . . 31

2.4 Weyl's tube formula . . . 34

2.5 Pfaffian functions . . . 35

2.6 Low–rank matrix approximation: the Eckart–Young theorem . . . 37

2.7 The geometry of Grassmannians . . . 39

2.7.1 Grassmannians as homogeneous spaces . . . 39

2.7.2 Isomorphisms of tangent spaces . . . 40

2.7.3 The Grassmann metric and its geodesics . . . 41

2.7.4 Principal angles, length minimizing geodesics and the Grassmann distance . . . 44

2.7.5 O–minimality of principal angles . . . 46

3 Real random tensors **49**

Basic notions on tensors . . . 49

3.1 Symmetric tensors close to rank–one . . . 52

3.1.1 Symmetric tensors and homogeneous polynomials . . . 54

3.1.2 The Veronese variety . . . 58

3.1.3 A restatement with polynomials . . . 61

3.1.4 The reach of the spherical Veronese variety . . . 63

3.1.5 The volume of tubular neighborhoods . . . 67
3.1.6 Applications . . . 76
Eigenvalues of random symmetric matrices . . . 76
Geometric measure of entanglement . . . 77
3.2 Typical ranks of order–three tensors . . . 79
3.2.1 A geometric perspective: (randomly) slicing the Segre variety . . . 81
3.2.2 Probabilistic Friedland's theorem . . . 84
3.2.3 Typical ranks of tall tensors . . . 87
3.2.4 $3 \times 3 \times 5$ tensors and cubic surfaces . . . 88
3.2.5 Asymptotics and heuristics . . . 92

4 Nearest point problems in the Grassmannian 95
4.1 The Euclidean Distance Degree . . . 96
4.2 The Grassmann Distance Complexity . . . 100
4.2.1 The subgradient of the Grassmann distance function . . . 102
4.2.2 Restricting the Grassmann distance to a submanifold . . . 109
4.2.3 Ignoring the cut locus . . . 111
4.2.4 Discreteness of critical points . . . 113
4.2.5 Uniform boundedness: the Grassmann Distance Complexity . . . 114
4.2.6 Non-semialgebraicity of the critical equations . . . 116
4.2.7 Pfaffian bounds . . . 117
4.3 Nearest point problems to simple Schubert varieties . . . 122
4.3.1 Schubert varieties . . . 123
4.3.2 Eckart–Young theorem and critical points . . . 126
4.3.3 The local structure around the deepest singularity . . . 130
4.3.4 Global minimum . . . 132
4.3.5 Global maximum . . . 135

Table of notations 139

Acknowledgments 143

Bibliography 143

Chapter 1

Introduction

Nowadays, differential geometry and algebraic geometry are presented and regarded as different areas of mathematics; the tools and approaches of the two subjects have diverged to the point that specialists of one area often speak a completely different language from those in the other. This was not always the case in the history of geometry. The 19th century geometers used to study various aspects of plane curves and surfaces in the three–dimensional space, such as their singularities and curvature, without any explicit distinction between algebraic and differential properties. This is exemplified in the classical books of that period, such as Salmon's treatises [75, 76]. In those works, many classical objects related to algebraic curves are defined in terms of distances and angles, for instance polars of algebraic curves, their foci, envelopes, and so on. See [17, Chapter 1] for the history of these objects.

It was only later, as mathematics evolved, that differential and algebraic geometry grew far apart. However, in modern science, data have become ubiquitous and data science now plays a key role in a wide variety of different settings. Geometry is used as an important tool for understanding and analyzing the shapes in which data present themselves. Typical tasks involving geometry and data are of a metric nature, such as distance optimization from a data point to a model. In many applications, polynomial equations arise naturally from the structure of the problem as defining equations of our models; their solutions over the real numbers are *real algebraic varieties*. We are thus faced with the study of metric properties of algebraic models: this is how differential and algebraic geometry come together once again. For this reason, in recent years a new subject has emerged as a vital area of research: *Metric Algebraic Geometry*. The term was coined by Madeleine Weinstein, who used it for the first time in the title of her PhD dissertation (UC Berkeley 2021). In the introduction, she writes the following:

> "*Metric algebraic geometry arose out of a desire to bring the perspective and tools of algebraic geometry to the objects of differential geometry that arise in applications concerning the geometry of data*".

Instances of this interplay between the two branches of geometry appear in optimal transport, machine learning, computer vision, statistics and many other fields.

In this dissertation we will discuss three problems that can be cast within the framework and scope of Metric Algebraic Geometry. The first two problems, though being very different from each other, are both related to *real random tensors*; for this reason they are grouped together in Chapter 3. In Section 3.1 we will analyze, from a probabilistic point

of view, the task of estimating how far a real symmetric tensor is from its best rank–one approximation; remark that rank–one symmetric tensors form an algebraic variety. In Section 3.2 we will instead turn our attention to *typical ranks* of real tensors of order three. Even though the subject of tensor rank has an algebraic nature, the methods we will use are a mixture of algebraic tools, such as the General Position Theorem, and differential ones, such as probability theory and Howard's kinematic formula. Lastly, Chapter 4 is devoted to the subject which constitutes the main work of the author of this dissertation during his PhD studies, that is the study of nearest point problems in the Grassmannian. In this work, we aim to extend the well–known notion of *Euclidean Distance Degree* to a Grassmannian, non–algebraic setting. This will be achieved through the notion of *Grassmann Distance Complexity*.

Chapter 2 contains a miscellanea of several distinct tools and results needed in the later chapters. We chose to collect them in a separate chapter since we deem these to be part of a more general background, less specific to the problems treated in Chapter 3 and Chapter 4.

We now proceed to provide, for each of the three mentioned problems, a brief introduction, an overview of the main contributions and a detailed account of the main results. Precise references and credits are given whenever a result is not original work of the author of this dissertation and his collaborators.

1.1 Symmetric tensors close to rank–one

Over the last decades, symmetric tensors have been used as very flexible way to encode data across numerous fields. In applications, tensor decomposition and rank–one approximations often play a crucial role. These include psychology and medical diagnostics, where they are used as a way to analyze multi–way data ([60],[3]); blind source separation ([28]); statistics, where they can be used to represent contingency tables ([68]); signal processing ([83]) and image analysis, where they are used for outlier removal, image compression and face recognition ([74, Ch.3, 4]); chemistry ([84]) and phylogenetics ([74, Ch.5],[63]); quantum computing ([46]; see also Section 3.1.6); the list could go further on.

In some of these applications, the distance between a tensor and its best rank–one approximation gives a measure of how accurately the approximation represents the tensor; in other words, it quantifies the error introduced when you substitute the original data with its rank approximation. For this reason, adopting the point of view of random tensors, we are drawn to the following question:

"What is the probability for a real symmetric tensor to be "close" to rank–one?"

The results we are going to describe here have been obtained as joint work with Alberto Cazzaniga and Antonio Lerario and were published in [25].

1.1.1 Overview of main contributions and sections

In order to give a precise meaning to the question above, we introduce the natural Frobenius scalar product on the space of real symmetric tensors and we consider its induced distance and associated Gaussian probability measure. In Section 3.1.1 we recall a well–known isomorphism between symmetric tensors and homogeneous polynomials and show that the Frobenius scalar product at the level of tensors corresponds to the Bombieri–Weyl

one at the level of polynomials. In Section 3.1.2 we describe the classical *Veronese variety* and its variants, in particular the *spherical Veronese variety*, as subvarieties of the space of homogeneous polynomials. We discuss the well–known correspondence, under the isomorphism from Section 3.1.1, between the Veronese variety and rank–one symmetric tensors. We end that section by highlighting that Veronese varieties are orbits for an action of the orthogonal group on the space of homogeneous polynomials by isometries. The computation of their volumes in the relevant metric is also provided. Thanks to the results from the previous sections, in Section 3.1.3 we reinterpret our original task in terms of homogeneous polynomials and show how the answer is provided by the computation of the volume of a tubular neighborhood of the spherical Veronese variety.

For the computation of the volume, the tool we use is *Weyl's tube formula*, which we introduce in Section 2.4 as part of the background material. The use of such result requires to compute the *reach* (see Section 2.3) and the *second fundamental form* of the spherical Veronese variety. In Section 3.1.4 we show how to simplify the computation of the reach, which is usually a very hard task, exploiting the invariance of the geometry of the Veronese variety. This way we are able to perform the explicit computation of the *bottleneck width* and the *maximal curvature*, obtaining an explicit evaluation of the reach. As a by–product, we obtain an interesting condition ensuring that a symmetric tensor has a unique best rank–one approximation.

We open Section 3.1.5 with the computation of the second fundamental form of the spherical Veronese variety. Thanks to the geometry invariance, this will be the same at every point. We prove that the normal bundle splits as an orthogonal direct sum of two subspaces: along normal directions in the first one, the second fundamental form is null; along a random normal direction in the other distributed as a Gaussian, the second fundamental form is a multiple of a GOE matrix. The computation is achieved via an explicit parametrization of the variety. Exploiting the structure of the normal bundle and the connection with random matrix theory, we reduce the computation of the curvature coefficients of the Veronese variety, which are the building blocks of Weyl's tube formula, to the computation of the expected value of the determinant of a random matrix. We solve this task, which is of a combinatorial nature, and thus provide a closed exact formula answering our question. In other words, we explicitly compute the probability that a random symmetric tensor is close, in the Frobenius metric, to the set of rank–one tensors, for any small enough threshold; how small is dictated by the reach we computed in Section 3.1.4.

We end this part by showing in Section 3.1.6 two applications of our probability formula. The first one deals with eigenvalues of random symmetric matrices. In particular, we obtain a condition, expressed in terms of singular values, under which a symmetric matrix has a unique best rank–one approximation. The second one discusses the relation between the distance from rank–one and quantum entanglement in quantum information theory. We show how our result can be interpreted in that framework to give information on how likely a quantum state is to have a low level of entanglement.

1.1.2 Statement of the results

The first step is to define a precise setting for our problem. We thus have to introduce a notion of distance and a probability measure on the space of real symmetric tensors of order d on $\mathbb{R}^n$; we denote this space by $\mathcal{S}^d(n)$. These are the elements of the tensor space $(\mathbb{R}^n)^{\otimes d}$ that are invariant under permutations of the factors (see Definition 3.3). The set

$\mathcal{S}^d_1(n)$ of symmetric rank–one tensors consists of the elements $T \in \mathcal{S}^d(n)$ such that

$$T = \pm v^{\otimes d} = \pm v \otimes \cdots \otimes v,$$

for some $v \in \mathbb{R}^n$.

The space of tensors $(\mathbb{R}^n)^{\otimes d}$ is endowed with the *Frobenius scalar product*: given two d–tensors $T = (T_{i_1,\ldots,i_d})$, $U = (U_{i_1,\ldots,i_d}) \in (\mathbb{R}^n)^{\otimes d}$ their Frobenius product is given by

$$\langle T, U \rangle_F = \sum_{i_1,\ldots,i_d=1}^{n} T_{i_1,\ldots,i_d} U_{i_1,\ldots,i_d} \,.$$

It is the product obtained by declaring the standard basis of $(\mathbb{R}^n)^{\otimes d}$ to be orthonormal. It can be interpreted as a standard Euclidean scalar product and is therefore the most natural one to consider. We use the same notation $\langle \cdot, \cdot \rangle_F$ for the restriction of the Frobenius scalar product to symmetric tensors $\mathcal{S}^d(n)$. We also denote by $\| \cdot \|_F$ the associated Frobenius norm and by $\widehat{\delta}(\cdot, \cdot)$ the induced distance function.

The choice of a scalar product on a finite-dimensional vector space naturally leads to the definition of a Gaussian measure on it (this comes for free without introducing further structure). In the case of our interest, the space of symmetric tensors with the Frobenius scalar product can be turned into a *Gaussian probability space* by defining for every Borel set $\mathcal{U} \subseteq \mathcal{S}^d(n)$

$$\mathbb{P}\Big\{ T \in \mathcal{U} \Big\} := \frac{\displaystyle\int_{\mathcal{U}} \exp\left(-\frac{1}{2} \|T\|_F^2 \right) \mathrm{d}\mu}{\displaystyle\int_{\mathcal{S}^d(n)} \exp\left(-\frac{1}{2} \|T\|_F^2 \right) \mathrm{d}\mu} \,, \tag{1.1}$$

where $\mathrm{d}\mu$ denotes integration with respect to the Lebesgue measure on $\mathcal{S}^d(n)$; the denominator serves as a normalization constant.

Remark 1.1. The Frobenius scalar product is invariant under the action of the orthogonal group on $\mathcal{S}^d(n)$. Therefore the corresponding Gaussian probability distribution (1.1) inherits the same invariance.

With these definitions, that make explicit what distance and what probability measure we are considering on $\mathcal{S}^d(n)$, we can state our problem as follows: we are required to compute, for a given threshold $\delta > 0$, the quantity:

$$\mathbb{P}\left\{ T \in \mathcal{S}^d(n) \,\middle|\, \widehat{\delta}(T, \mathcal{S}^d_1(n)) \leq \delta \|T\|_F \right\}. \tag{1.2}$$

Notice that we have turned the question into a conic problem that takes into account also the norm of the tensor. This is a common procedure in numerical algebraic geometry [20]. If we interpret the distance as an approximation error, this corresponds to a switch from an *absolute error* to a *relative* one.

Symmetric tensors and homogeneous polynomials

It will be useful for us to identify symmetric tensors with homogeneous polynomials. For notational reasons, we work with tensors on $\mathbb{R}^{n+1}$ rather than on $\mathbb{R}^n$. More precisely,

to every symmetric tensor $T \in \mathcal{S}^d(n+1)$ we associate a homogeneous polynomial $p_T \in \mathbb{R}[x_0, \dots, x_n]_{(d)}$ of degree d in $n+1$ variables defined by

$$p_T(x) := \langle T, x^{\otimes d} \rangle_F,$$

where $x = (x_0, \dots, x_n)$. This gives a linear isomorphism

$$\phi_d : \mathcal{S}^d(n+1) \xrightarrow{\simeq} \mathbb{R}[x_0, \dots, x_n]_{(d)}, \tag{1.3}$$

which, when $d = 2$, is the familiar identification between a symmetric matrix Q and the associated quadratic form $p_Q(x) = \langle x, Qx \rangle$. The advantage of this approach is that it connects to well–studied objects in the theory of random polynomials and it allows to give a neat description of the set of rank–one tensors, under the linear isomorphism ϕ.

On the space of complex polynomials we consider the *Bombieri–Weyl* Hermitian product, defined by

$$\langle p_1, p_2 \rangle_{\mathrm{BW}} := \frac{1}{\pi^{n+1}} \int_{\mathbb{C}^{n+1}} p_1(z)\overline{p_2(z)} e^{-\|z\|^2} \mathrm{d}z, \tag{1.4}$$

where $\mathrm{d}z := (i/2)^n \mathrm{d}z_0 \mathrm{d}\overline{z_0} \dots \mathrm{d}z_n \mathrm{d}\overline{z_n}$ is the Lebesgue measure. The Frobenius scalar product corresponds to the restriction to $\mathbb{R}[x_0, \dots, x_n]_{(d)}$ of the real part of (1.4) under the identification (1.3), which therefore defines an isometry with respect to these structures (Proposition 3.7). We call this restriction the *Bombieri–Weyl* scalar product. The expression (1.4) defines the unique, up to multiples, Hermitian product on the space of complex polynomials that is invariant under the action of the unitary group by change of variables. This implies the invariance of the Bombieri–Weyl scalar product under the action of the orthogonal group by change of variables. In the same way we defined a Gaussian probability measure on $\mathcal{S}^d(n+1)$, we consider on the space of homogeneous polynomials $\mathbb{R}[x_0, \dots, x_n]_{(d)}$ the Gaussian measure associated to the Bombieri–Weyl scalar product; this can also be obtained as the pushforward of the Frobenius one via (1.3).

We have thus explained how the isomorphism ϕ_d allows to identify symmetric tensors with homogeneous polynomials, and how this identification extends to the relevant probability distributions. We now need to understand how the correspondence behaves at the level of rank–one tensors; to do this, we have to introduce *Veronese varieties*.

The Veronese variety and its variations

We give the following key definition (Definition 3.8). For $n \geq 1$ and $d \geq 1$, the *real Bombieri–Weyl Veronese embedding* is the map

$$\begin{aligned} \nu_{n,d} : \mathbb{RP}^n &\longrightarrow \mathbb{RP}^N \\ [a] &\longmapsto \left[\binom{d}{\alpha}^{1/2} a^\alpha \right], \end{aligned}$$

where $\mathbb{RP}^N \simeq \mathbb{P}(\mathbb{R}[x_0, \dots, x_n]_{(d)})$ and $\alpha \in \mathbb{Z}^{n+1}_{\geq 0}$ satisfy $\alpha_0 + \cdots + \alpha_n = d$, with $\binom{d}{\alpha}$ being the multinomial coefficient. Notice that on the right we have the evaluations of the polynomials of an orthonormal basis for the Bombieri–Weyl scalar product (see (3.6)); this is also the basis with respect to which we express coordinates in $\mathbb{R}[x_0, \dots, x_n]_{(d)}$. We call the image of this map the *Bombieri–Weyl projective Veronese variety* and denote it by $\mathcal{P}V_{n,d} := \mathrm{im}(\nu_{n,d})$. The corresponding cone over $\mathcal{P}V_{n,d}$ is denoted by $\mathcal{V}_{n,d}$ and called the

Bombieri–Weyl Veronese variety. The main object we consider in this work is the *real spherical Veronese variety* $V_{n,d}$. This is the intersection of the Veronese variety with the unit sphere in the Bombieri–Weyl norm:

$$V_{n,d} := \mathcal{V}_{n,d} \cap S^N.$$

We regard the set $V_{n,d}$ as the image of the *spherical Veronese map* associated to the Bombieri–Weyl basis, or its double copy, depending on the parity of d. This is the smooth map $\widehat{\nu}_{n,d} : S^n \to S^N$ given by

$$x \overset{\widehat{\nu}_{n,d}}{\longmapsto} \left(\binom{d}{\alpha}^{\frac{1}{2}} x^\alpha \right),$$

where S^n is the Euclidean sphere in $\mathbb{R}^{n+1}$. Denoting the image $\widehat{\nu}_{n,d}(S^n)$ by $\mathcal{S}V_{n,d}$, we see that

$$V_{n,d} = \mathcal{S}V_{n,d} \cup -\mathcal{S}V_{n,d},$$

where $\mathcal{S}V_{n,d} = -\mathcal{S}V_{n,d}$ if d is odd and $\mathcal{S}V_{n,d} \cap (-\mathcal{S}V_{n,d}) = \emptyset$ if d is even. For this reason, we will call $\mathcal{S}V_{n,d}$ the *spherical Veronese surface*, to distinguish it from the spherical Veronese variety $V_{n,d}$, in the case d is even. In the projective picture, all these objects coincide:

$$\mathcal{P}V_{n,d} := \mathbb{P}(\mathcal{S}V_{n,d}) = \mathbb{P}(V_{n,d}) \subseteq \mathbb{R}\mathrm{P}^N.$$

It follows from an easy computation that $\mathcal{P}V_{n,d}$ parametrizes projective classes of d–th powers of linear forms on $\mathbb{R}^{n+1}$; the Veronese variety $\mathcal{V}_{n,d}$ then parametrizes signed such powers. Observe that given a rank–one tensor $T = \pm\, v^{\otimes d}$, its image through the isomorphism ϕ_d is given by

$$p_T(x) = \langle \pm\, v^{\otimes d}, x^{\otimes d} \rangle_F = \pm \langle v, x \rangle^d,$$

which is the d–th power of the linear form defined by v. Therefore, under the isomorphism ϕ_d, the set of rank–one tensors can be identified with the Veronese variety $\mathcal{V}_{n,d}$.

In the same way, $\mathcal{S}V_{n,d}$ parametrizes d–th powers of norm–1 linear forms, which correspond to rank–one norm–1 symmetric tensors, up to sign. Hence, the spherical Veronese surface $\mathcal{S}V_{n,d}$ corresponds to an orbit for the action of the orthogonal group $O(n+1)$ on homogeneous polynomials by change of variables (Proposition 3.11). Even more is true: when turning $\mathcal{S}V_{n,d}$ into a Riemannian manifold with the metric induced by the Bombieri–Weyl scalar product, the transitive action of $O(n+1)$ on $\mathcal{S}V_{n,d}$ is given by isometries that are induced by ambient isometries of S^N; this is due to the invariance property of the Bombieri–Weyl structure. The immediate, yet crucial, consequence is that the extrinsic geometry of the isometric embedding $\mathcal{S}V_{n,d} \hookrightarrow S^N$ is the same at every point; the same conclusion holds also for the spherical Veronese variety $V_{n,d} \hookrightarrow S^N$. This geometry invariance will be one of the key ingredients for our results.

A geometric formulation via random polynomials

At this point we are in the position of giving a more geometric formulation of our problem. Denoting by d^{BW} and $\|\cdot\|_{\mathrm{BW}}$ the distance and the norm induced by the Bombieri–Weyl scalar product respectively, (1.2) is equivalent, via the identifications outlined before, to

$$\mathbb{P}\left\{ p \in \mathbb{R}[x_0, \dots, x_n]_{(d)} \;\middle|\; d^{\mathrm{BW}}(p, \mathcal{V}_{n,d}) \leq \delta \|p\|_{\mathrm{BW}} \right\}, \tag{1.5}$$

where the homogeneous polynomial $p \in \mathbb{R}[x_0, \dots, x_n]_{(d)}$ is distributed according to the Gaussian measure associated to the Bombieri–Weyl structure, and $\mathcal{V}_{n,d}$ is the Veronese variety (notice the notational shift from n to $n+1$ variables). Using the conic nature of the problem, we can reduce to consider only points on the unit sphere S^N. Then, since a centered Gaussian restricts to the uniform measure on a sphere, by elementary geometry we can see that

$$(1.5) = \frac{\mathrm{Vol}(\mathcal{U}(V_{n,d}, \varepsilon))}{\mathrm{Vol}(S^N)}, \tag{1.6}$$

where $\mathcal{U}(V_{n,d}, \varepsilon)$ is the tubular neighborhood of the spherical Veronese variety $V_{n,d} \hookrightarrow S^N$ of radius $\varepsilon = \arcsin(\delta)$ (see Fig. 3.1); this is the set of points at (spherical) distance less than ε from $V_{n,d}$ (see Eq. (2.2)). The computation of the probability in (1.2) is thus equivalent to the knowledge of the volume of the tubular neighborhood of $V_{n,d}$. This is our new goal.

Remark 1.2. The space of homogeneous polynomials with the Bombieri–Weyl Gaussian measure is sometimes called the *Kostlan Ensemble* or the *Shub-Smale Ensemble*. The properties of this probability space have been intensively studied, starting from the influential works of A. Edelman, E. Kostlan, M. Shub and S. Smale [36, 81, 80, 82]. The point of view of random tensors has been adopted first by E. Horobet and J. Draisma in [33] and by P. Breiding in [15] for the study of the expected number of eigenvalues of a random symmetric tensor, with respect to the Bombieri–Weyl distribution. Under the identification between symmetric tensors and homogeneous polynomials, eigenvalues correspond to critical values of the restriction of the polynomial to the unit sphere. Eigenvectors correspond to critical points of the polynomial: under the Veronese embedding, these critical points give rank–one tensors that are critical points of the distance function from the given tensor restricted to the Veronese variety. Among these critical points, the closest to the original tensor are its best rank–one approximations. In [15] the average number of such critical points is computed. In our work, as a by–product of our computations, we will instead give the size and estimate (from below) the probability of the set of tensors which admit a unique best rank–one approximation.

The use of *Weyl's tube formula* is fairly standard to obtain results of the type of (1.6) ([20], [8]). We describe this tool in Section 2.4: it allows to deduce an exact expression, for small enough radii, of the volume of tubular neighborhoods of smooth submanifolds of a Euclidean space or a round sphere. This is obtained as a function of some differential–geometric quantities of the submanifold, which are called *curvature coefficients* (see (2.5) and Definition 2.10). In the case of the spherical Veronese variety Weyl's tube formula has the following expression

$$\mathrm{Vol}(\mathcal{U}(V_{n,d}, \varepsilon)) = \sum_{\substack{0 \le e \le n \\ e \text{ even}}} K_{c+e}(V_{n,d}) J_{N,c+e}(\varepsilon),$$

where $c := N - n$ is the codimension of $V_{n,d} \hookrightarrow S^N$, the K_j's are the curvature coefficients and the $J_{N,j}$'s are the universal functions in (2.7). The integral version of the formula (see (2.8)) in our case reads as

$$\mathrm{Vol}(\mathcal{U}(V_{n,d}, \varepsilon)) = \int_{p \in V_{n,d}} \int_{t=0}^{\tan \varepsilon} \int_{\eta \in S(N_p V_{n,d})} \frac{t^{c-1} \det\left[I_n - t L_\eta\right]}{(1+t^2)^{\frac{N+1}{2}}} \, \mathrm{vol}_M \, d\eta \, dt, \tag{1.7}$$

where $S(N_pV_{n,d})$ is the unit sphere in the normal space to $V_{n,d}$ at p, I_n is the $n \times n$ identity matrix and L_η is the Weingarten operator of $V_{n,d} \hookrightarrow S^N$ at p in the normal direction η (see Section 2.1; this is just a different way of encoding the second fundamental form).

Weyl's tube formula holds whenever the radius of the neighborhood is smaller than a quantity called the *reach* of the embedded submanifold. The knowledge of the reach is therefore required if we want to meaningfully apply this technique. Moreover, if we want to evaluate explicitly (1.7) we need to understand its second fundamental form at every point. These are the two tasks we need to solve now: computing the reach and describing the second fundamental form of the embedding $V_{n,d} \hookrightarrow S^N$.

The reach of the Veronese variety

The reach of an embedded submanifold $M \hookrightarrow \overline{M}$ gives a bound to the radii of tubular neighborhoods that can be parametrized via the normal exponential map (Definition 2.5); we denote it by $\rho(M)$. As Weyl's tube formula is entirely based on this parametrization, it follows that it is only valid for radii smaller than the reach. See Proposition 2.6 for another characterization of the reach in terms of the *medial axis* of the submanifold.

Theorem 2.7 characterizes the reach as the minimum between the *bottleneck width* and the reciprocal of the *maximal curvature* of the submanifold. A *bottleneck* of M is a pair of points (x, y) on the submanifold such that the ambient geodesic connecting them is orthogonal to the submanifold at both endpoints; in other words the couple (x, y) is a critical point for the function computing the distance between pairs of points on M. The *width* of a bottleneck (x, y) is half the distance between the points x and y; the bottleneck width of the submanifold M is then the minimum of the width of a bottleneck. We denote it by $\mathrm{bw}(M)$. The maximal curvature, denoted by $\kappa_{\max}(M)$ is the largest principal curvature of M over all its points. Its reciprocal corresponds to the distance between the submanifold M and its closest *focal point*; these are the critical values of the normal exponential map. See Section 2.3 for more details.

The direct computation of these quantities is often unfeasible in the absence of some symmetries or special structures, as it requires the study of normal geodesics and curvature at every point. In our case, recalling the invariance property of $V_{n,d}$ under the action of the orthogonal group, we can show that it is enough to fix a point in the variety and perform the computations considering only geodesics originating from the chosen point (Lemma 3.15). This drastically reduces the complexity of the computation and allows us to obtain our first relevant result (see Theorem 3.17).

Theorem A (The reach of the spherical Veronese variety)**.** *The inverse of the maximal curvature and the bottleneck width of* $V_{n,d} \hookrightarrow S^N$ *are given by*

$$\kappa_{\max}^{-1}(V_{n,d}) = \frac{1}{\sqrt{2}}\sqrt{1 + \frac{1}{d-1}}$$
$$\mathrm{bw}(V_{n,d}) = \frac{\pi}{4}.$$

Therefore the reach of $V_{n,d} \hookrightarrow S^N$ *is given by*

$$\rho(V_{n,d}) = \min\{\kappa_{\max}^{-1}(V_{n,d}),\ \mathrm{bw}(V_{n,d})\} = \begin{cases} \frac{\pi}{4} & 2 \le d \le 5 \\ \frac{1}{\sqrt{2}}\sqrt{1 + \frac{1}{d-1}} & d \ge 6 \end{cases}.$$

Remark 1.3 (Bottleneck width of the Veronese variety)**.** From Theorem A it follows that the bottleneck width is the same for all Veronese varieties, as it shows no dependence on either n or d. Bottlenecks of algebraic varieties play an important role in optimization problems. For a general algebraic variety, the number of bottlenecks can be estimated using classical invariants such as Chern classes and polar classes [31]; see [31, 30] for more details.

Recalling the interpretation of $V_{n,d}$ as norm–1 rank–one symmetric tensors (up to sign), this theorem has an interesting consequence. By [40], the best rank–one approximation of a symmetric tensor is still symmetric. Then, as a corollary of Theorem A, we obtain that every real symmetric tensor sufficiently close to rank–one tensors admits a unique best rank–one approximation (Corollary 3.19). We here identify tensors with homogeneous polynomials via (1.3).

Corollary B (Uniqueness of best rank–one approximation)**.** *Every symmetric tensor p at distance less than $\sin(\rho(V_{n,d}))\|p\|_{\mathrm{BW}}$ from rank–one admits a unique best rank–one approximation.*

In particular, the normalized volume of a neighborhood of the Veronese variety of radius $\rho(V_{n,d})$ will provide a lower bound for the probability that a symmetric tensor has a unique best rank–one approximation. This can be explicitly computed by plugging in $\varepsilon = \rho(V_{n,d}) = \arcsin(\delta)$ in Theorem F below.

The curvature coefficients of the Veronese variety

The other ingredient we need in (1.7) is the curvature of the embedding $V_{n,d} \hookrightarrow S^N$, in particular the Weingarten operator along normal directions. As we already saw in the computation of the reach, also here the invariance property of the extrinsic geometry of $V_{n,d}$ under the action of the orthogonal group allows to simplify the computation. It is indeed enough to fix any point in $V_{n,d}$ and compute the Weingarten operator at that point; the description at every other point will then be the same. We perform the computation at the point $x_0^d \in V_{n,d}$ in Theorem 3.20 and extend it to the whole variety as follows (Corollary 3.21).

Theorem C (The normal bundle splitting)**.** *Let $p \in V_{n,d}$ and denote by L_η the symmetric matrix representing the second fundamental form of $V_{n,d} \hookrightarrow S^N$ along the normal direction $\eta \in N_p V_{n,d}$. There exists an orthogonal decomposition $N_p V_{n,d} = W_p \oplus P_p$ such that the following statements hold:*

1. *$L_\eta = 0$ for every $\eta \in P_p$;*
2. *W_p with its induced Bombieri–Weyl metric is isometric to $\mathrm{Sym}(n)$ with the Frobenius one. Moreover, if we pick $\eta \in W_p$ Gaussian, then $L_\eta \sim \sqrt{2}\left(\frac{d-1}{d}\right)^{\frac{1}{2}} \mathrm{GOE}(n)$, i.e. L_η is distributed, as a random matrix, as a multiple of a $\mathrm{GOE}(n)$ matrix (see Eq. (3.8)).*

This theorem could find applications going beyond the scope of this work. It gives a full description of the second fundamental form in terms of $\mathrm{GOE}(n)$ matrices. The use of this description in (1.7) has the effect of transforming the computation of the integral over the normal bundle into the computation of the expectation of a determinant involving $\mathrm{GOE}(n)$ matrices. This boils down to a purely combinatorial computation (3.44). We are thus able to evaluate (1.7) and obtain the volume of the tubular neighborhood (Theorem 3.22).

Theorem D (The volume of a neighborhood of the spherical Veronese variety)**.** *Let $V_{n,d} \hookrightarrow S^N$ be the spherical Veronese variety. For $\varepsilon < \rho(V_{n,d})$ the following formula holds:*

$$\mathrm{Vol}(\mathcal{U}(V_{n,d},\varepsilon)) = \sum_{\substack{0\le j\le n\\ j \text{ even}}} (-1)^{\frac{j}{2}} d^{\frac{n}{2}} \left(\frac{d-1}{d}\right)^{\frac{j}{2}} \frac{2^{n+2-j}\pi^{\frac{N}{2}}\Gamma\left(\frac{n}{2}+1\right)}{\Gamma\left(\frac{j}{2}+1\right)\Gamma(n+1-j)\Gamma\left(\frac{N+j-n}{2}\right)} J_{N,N-n+j}(\varepsilon),$$

where the functions J's are given by (2.7).

As a consequence we obtain the following explicit expressions for the curvature coefficients in Weyl's tube formula (Corollary 3.23).

Theorem E (The curvature coefficients of the spherical Veronese)**.** *The curvature coefficients of the Veronese variety $V_{n,d} \hookrightarrow S^N$ are as follows:*

$$K_{N-n+j}(V_{n,d}) = (-1)^{\frac{j}{2}} d^{\frac{n}{2}} \left(\frac{d-1}{d}\right)^{\frac{j}{2}} \frac{2^{n+2-j}\pi^{\frac{N}{2}}\Gamma\left(\frac{n}{2}+1\right)}{\Gamma\left(\frac{j}{2}+1\right)\Gamma(n+1-j)\Gamma\left(\frac{N+j-n}{2}\right)}$$

for $0 \le j \le n$ and j even, and $K_{N-n+j}(V_{n,d}) = 0$ otherwise.

As another consequence of Theorem D, we are finally able to give a complete answer to the question we started from: how likely is a symmetric tensor to be close to rank–one? We once again identify symmetric tensors with homogeneous polynomials. The following corresponds to Theorem 3.25.

Theorem F (The probability of being close to rank–one)**.** *Let p be a random Bombieri–Weyl symmetric tensor of order d on $\mathbb{R}^{n+1}$ and $\mathcal{V}_{n,d} \subset \mathbb{R}[x_0,\dots,x_n]_{(d)} \simeq \mathbb{R}^{N+1}$ be the Veronese variety of rank–one tensors. For every δ such that $0 \le \arcsin(\delta) < \frac{1}{\sqrt{2}}\sqrt{1+\frac{1}{d-1}}$ we have*

$$\mathbb{P}\left(\left\{d^{\mathrm{BW}}(p,\mathcal{V}_{n,d}) \le \delta\|p\|_{\mathrm{BW}}\right\}\right) = \sum_{\substack{0\le j\le n\\ j \text{ even}}} (-1)^{\frac{j}{2}} d^{\frac{n}{2}} \left(\frac{d-1}{d}\right)^{\frac{j}{2}} 2^{n-j+1}\pi^{-\frac{1}{2}} \cdot$$
$$\cdot \frac{\Gamma\left(\frac{n}{2}+1\right)\Gamma\left(\frac{N+1}{2}\right)}{\Gamma\left(\frac{j}{2}+1\right)\Gamma(n+1-j)\Gamma\left(\frac{N+j-n}{2}\right)} \int_0^{\frac{\delta}{\sqrt{1-\delta^2}}} \frac{t^{N-n+j-1}}{(1+t^2)^{\frac{N+1}{2}}}\mathrm{dt}.$$

We remark that the above expression, even if unpleasant, gives an exact formula for our probability. In general this has an exponential decay in the codimension of $V_{n,d} \hookrightarrow S^N$. In the case of rational normal curves, there is only one summand. It can be evaluated in a closed form, giving

$$\mathbb{P}\left(d^{\mathrm{BW}}(p,\mathcal{V}_{1,d}) \le \delta\|p\|_{\mathrm{BW}}\right) = \sqrt{d}\,\delta^{d-1}, \quad \text{for } \arcsin(\delta) \le \rho(V_{1,d}). \tag{1.8}$$

In fact in this case $N = d$ and, as one can easily verify,

$$\int_0^{\frac{\delta}{\sqrt{1-\delta^2}}} \frac{t^{d-2}}{(1+t^2)^{\frac{d+1}{2}}}\,\mathrm{dt} = \frac{\delta^{d-1}}{d-1}.$$

Once substituted in the above expression this immediately gives (1.8).

One could generalize this question to higher ranks by looking at secant varieties to the Veronese. The geometry of these varieties has been intensively studied, see [22] for a survey. Nevertheless, some technical difficulties, such as their non–smoothness, makes it hard, if not impossible, to apply the same techniques we used here. Therefore that task seems to be way harder than the rank–one case; we propose to investigate this in future works.

1.2 Typical ranks of order–three tensors

Determining the rank of a tensor is of interest for a broad variety of applications across many fields, such as psychometrics and data science; see [58] for a survey. While matrix rank is well understood and shows many useful properties, going up from matrices to higher order tensors things gets more complicated and almost all the properties of matrix rank are not preserved. For example, the maximal rank of an $m \times n$ matrix over either $\mathbb{C}$ or $\mathbb{R}$ is $\min\{m, n\}$; moreover the set of matrices with maximal rank is open and dense in the space of matrices of that given size. In the case of tensors, computing the maximal rank for a given format is instead a very hard task and very little is known (see for example [38]); moreover, the set of tensors of maximal rank is usually not even full-dimensional.

In the second problem we consider in this dissertation we focus on what are known as *typical ranks*. A positive integer r is called a typical rank of a given format if the set of tensors of that format that have rank equal to r is a full–dimensional semialgebraic set. Over the field $\mathbb{C}$ of complex numbers there is always only one typical rank, known as the *generic rank* of that format; this actually holds over any algebraically closed field. Notice that if the order of the tensor is at least 3, the generic rank is not the maximal one, marking a difference with the matrix (order–2) case. Over the field $\mathbb{R}$ of real numbers, while most formats still exhibit only one typical rank, there are many others for which multiple typical ranks are known to exist. In several other cases, the question whether or not there exist multiple typical ranks is still open.

From now on we only focus on the case of real order–3 tensors of format $m \times n \times \ell$. Since the rank of a tensor is invariant under permutations of the factors, we will always assume that $2 \leq m \leq n \leq \ell$. The study of typical ranks of order–3 tensors has been an active area of research since the late '90s and typical ranks have been computed in several cases. A first general result, valid for all formats, is that typical ranks are always consecutive: if two integers are typical ranks for a format, so are all the integers between them. Moreover, the smallest typical rank over $\mathbb{R}$ always corresponds to the generic rank over $\mathbb{C}$ [38]. In 1999 ten Berge and Kiers showed that $2 \times n \times \ell$ tensors have a unique typical rank, equal to $\min\{2n, \ell\}$, whenever $n \neq \ell$; if instead $n = \ell$, they admit two typical ranks, namely n and $n + 1$ [90]. For example, $2 \times 4 \times 4$ tensors have typical ranks 4 and 5; nevertheless the maximal rank of this format is 6. Moving on to larger formats, it is known that $3 \times 3 \times 3$ and $3 \times 3 \times 4$ formats have only one typical rank equal to 5 [91],[1]. In [89] ten Berge shows that instead $3 \times 3 \times 5$ tensors admit as typical ranks both 5 and 6. Sumi, Miyazaki and Sakata linked the presence of multiple typical ranks to the existence of a non–singular bilinear map $\mathbb{R}^m \times \mathbb{R}^n \longrightarrow \mathbb{R}^{mn-\ell}$, in the range where $(m-1)(n-1)+1 < \ell \leq mn$ [86, Theorem 1.1]; remark that if $\ell > mn$ the only typical rank is trivially mn. In the smaller range $(m-1)n < \ell < mn$, which ten Berge calls the range of *tall tensors*, there is a unique typical rank equal to ℓ [88],[73]. For a summary of known typical ranks for small formats see Table 1.1; for a more comprehensive survey see [38].

The perspective through which we approach typical ranks is that of probability. We will always consider tensors as multi–dimensional arrays, as explained at the beginning of Chapter 3. We use the notation $T = (t_{ijk}) \in \mathbb{R}^{m \times n \times \ell}$ for $i = 1, \ldots, m$, $j = 1, \ldots, n$ and $k = 1, \ldots, \ell$. We give a probabilistic definition of typical ranks as follows.

Definition 1.4 (Typical ranks, probabilistic viewpoint)**.** A random tensor $T \in \mathbb{R}^{m \times n \times \ell}$ is called *Gaussian* if its entries are i.i.d. standard Gaussians $t_{ijk} \sim \mathcal{N}(0, 1)$. A positive integer r is called a typical rank of the format $m \times n \times \ell$ if for a Gaussian tensor $T \in \mathbb{R}^{m \times n \times \ell}$ we

	$m=2$				$m=3$		
	$n=2$	$n=3$	$n=4$		$n=3$	$n=4$	$n=5$
$\ell=2$	{2,3}	3	4	$\ell=5$	{5,6}	?	?
$\ell=3$	3	{3,4}	4	$\ell=6$	6	?	?
$\ell=4$	4	4	{4,5}	$\ell=7$	7	{7,8}	?
$\ell=5$	4	5	5	$\ell=8$	8	{8,9}	?
$\ell=6$	4	6	6	$\ell=9$	9	9	{9,10}
$\ell=7$	4	6	7	$\ell=10$	9	10	10
$\ell=8$	4	6	8	$\ell=11$	9	11	11
$\ell=9$	4	6	8	$\ell=12$	9	12	12

Table 1.1: Typical ranks of real tensors of format $m \times n \times \ell$ based on results from [89, 73, 87]. The question marks "?" indicate that typical ranks of that format are not yet known.

have

$$\mathbb{P}\left(\{\operatorname{rank}(T) = r\}\right) > 0. \tag{1.9}$$

This definition is equivalent to the previous one. The probabilistic point of view sheds light on the importance of typical ranks in applications. Indeed, if the tensor represents real world data, the only possible ranks that you might witness are the typical ones, due to the presence of noise in the data.

For those formats for which multiple typical ranks are known to exist, an interesting question is what is the probability in (1.9) for each of the typical ranks. In this case, almost nothing is known. The only results we are aware of are due to Bergqvist and Forrester. In [11] Bergqvist gives an exact computation of the rank probabilities for $2 \times 2 \times 2$ and $2 \times 3 \times 3$ tensors. He finds that a Gaussian $2 \times 2 \times 2$ tensor has rank 2 with probability $\frac{\pi}{4}$, while a Gaussian $2 \times 3 \times 3$ tensor has rank 3 with probability $\frac{1}{2}$, which is therefore the same as having rank 4. His technique is based on the study of the average number of real solutions for the generalized eigenvalues problem for random Gaussian matrices. His result for $2 \times 2 \times 2$ tensors confirms a previous numerical simulation run by Kruskal [61]. Together with Forrester, they generalize these results and compute exact rank probabilities for all Gaussian tensors of format $2 \times n \times n$ in terms of Gamma and Barnes G-functions [12]. No other results are known to the author of this dissertation.

Most of the previous results about whether a format admits multiple typical ranks are obtained through purely algebraic means. In many instances, the proofs develop through a sequence of algebraic manipulations on the tensor and offer little geometric insight. We instead adopt a more geometric approach: even though we de not find new results on which formats admit multiple typical ranks, through our method we obtain new proofs that are often simpler than the original ones. Moreover, they offer a geometric viewpoint on tensor rank that is often lacking in the original works where the results were first proven. Precise references will be given every time that a result has already been discovered before.

In this work we move in two directions. In the first part we develop a general framework to study the typical ranks of some families of tensors. In the last part we instead turn our attention to the evaluation of rank probabilities of Gaussian tensors. The results we describe in this part have been obtained as joint work with Paul Breiding and Sarah Eggleston and were published in [16].

1.2.1 Overview of main contributions and sections

Our starting point is a result by Friedland ([38, Theorem 2.4]) that characterizes the rank of an $m \times n \times \ell$ tensor in terms of the minimum number of rank–one $m \times n$ matrices that span a subspace containing the slices of the tensor. This leads us to the question of when the span of the slices is itself spanned by rank–one matrices. We address this in Section 3.2.1. After introducing the projective Segre variety of rank–one matrices and the notion of a random uniform subspace, we prove a trichotomy (depending on the dimension of the subspace) showing that whether such a subspace is spanned or not by rank–one matrices depends on the way it slices the *real* Segre variety. We end that section by showing that the linear span of the slices of a Gaussian tensor is indeed a uniform subspace.

In Section 3.2.2 we prove a probabilistic counterpart of Friedland's theorem: we show that, with probability 1, the rank of an $m \times n \times \ell$ Gaussian tensor is bounded by ℓ if and only if the linear span of its slices projectively intersects the real Segre variety in at least ℓ distinct points. This leads to our first main result: in the range $(m-1)(n-1)+1 \leq \ell \leq mn$, the typical ranks of the format $m \times n \times \ell$ are always contained in $\{\ell, \ell+1\}$ and ℓ is always typical. The first part of this result was already proved by Sumi, Miyazaki and Sakata [86, Theorem 8.1] through algebraic means: our proof is independent and entirely built on our geometric framework. For the second part, it follows from a probabilistic version of the previously proved tricothomy. As a corollary we show that in the range $(m-1)(n-1)+1 < \ell \leq mn$, if the degree of the Segre variety of $m \times n$ matrices is odd, then we only have one typical rank equal to ℓ. In Section 3.2.3 we give a different proof of ten Berge's result that tall tensors admit a unique typical rank [88].

Section 3.2.4 is entirely focused on understanding typical ranks of $3 \times 3 \times 5$ Gaussian tensors. This is done via an unexpected connection with cubic surfaces. It is known that the typical ranks of this format are 5 and 6 [89]. Recall from classical algebraic geometry that every smooth cubic surface in $\mathbb{C}\mathrm{P}^3$ arises as the blow up of $\mathbb{C}\mathrm{P}^2$ in six points in general position and that every such surface contains 27 complex lines. A theorem from [29] associates to a general linear subspace of $\mathbb{C}\mathrm{P}^8 = \mathbb{P}(\mathbb{C}^{3\times 3})$ a determinantal smooth cubic surface in $\mathbb{C}\mathrm{P}^3$. We adapt this result to the case of a general subspace of $\mathbb{R}\mathrm{P}^8$; the associated surface will now be a real cubic surface. Using this fact and a classification of real cubic surfaces due to Polo–Blanco and Top [71], we prove that the probability of having a given number of real intersection points between the linear span of the slices of a $3 \times 3 \times 5$ Gaussian tensor and the Segre variety corresponds to the probability of having a given number of real lines on a determinantal real cubic surface. In particular, we prove that the probability that a $3 \times 3 \times 5$ Gaussian tensor has rank 5 is the same as the probability that on a determinantal real cubic surface all the 27 lines are real. As a by–product, we also obtain some bounds on the expected number of real lines on such surfaces.

At last, in Section 3.2.5 we use techniques from integral geometry to obtain the expected number of intersection points between the real Segre variety and a complementary dimension uniform subspace. Comparing the asymptotics of this expectation with the number of real intersection points required for a $m \times n \times (m-1)(n-1)+1$ tensor to have the smallest typical rank, which is equal to $(m-1)(n-1)+1$ by our previous results, we obtain some heuristics on how the rank probabilities behave. This is matched, for small enough formats, by numerical simulations.

1.2.2 Statement of the results

We fix the following notations. Given a tensor $T = (t_{ijk}) \in \mathbb{R}^{m\times n\times \ell}$, we define its *slices* to be the $m \times n$ matrices $T_1, \dots, T_\ell$ given by $(T_k)_{ij} = t_{ijk}$ for $k = 1, \dots, \ell$. Friedland proves that the rank of T is the minimum number r such that there exist rank–one matrices $A_1, \dots, A_r \in \mathbb{R}^{m\times n}$ with the property that

$$\langle T_1, \dots, T_\ell \rangle \subseteq \langle A_1, \dots, A_r \rangle, \tag{1.10}$$

where $\langle v_1, \dots, v_s \rangle$ denotes the linear subspace spanned by the vectors $v_1, \dots, v_s$ ([38, Theorem 2.4], here as Theorem 3.29). Since the rank ok a tensor is invariant under scaling, we can work in the projective setting; we will denoted by $\langle v_1, \dots, v_s \rangle_{\mathbb{P}}$ the projectivization of the linear span of $v_1, \dots, v_s$. The first question arising from (1.10) is whether we can say something about the cases where the inclusion is actually an equality. We therefore investigate when a linear subspace in $\mathbb{R}^{m\times n}$ is spanned by rank–one matrices.

Slicing the Segre variety

We introduce the *Segre variety* $\Sigma_{m,n}$ of projective rank–one matrices in $\mathbb{P}(\mathbb{R}^{m\times n}) = \mathbb{RP}^{mn-1}$. Its complex Zariski closure is the *complex Segre variety* of rank–one matrices in $\mathbb{CP}^{mn-1}$, denoted by $\Sigma^{\mathbb{C}}_{m,n}$. The dimension of $\Sigma_{m,n}$ is $m + n - 2$ and its codimension is $(m - 1)(n - 1)$. The degree of the complex variety $\Sigma^{\mathbb{C}}_{m,n}$ is $\binom{m+n-2}{m-1}$; recall that this means that a general linear subspace of dimension $(m - 1)(n - 1)$ in $\mathbb{RP}^{mn-1}$ will intersect $\Sigma^{\mathbb{C}}_{m,n}$ in that number of complex points. Recall also the definition of the Grassmannian $G(d, mn)$, that is the set of linear subspace of dimension d in $\mathbb{R}^{m\times n}$ (see Section 2.7). On $G(d, mn)$ there is a unique probability distribution that is invariant under the action of the orthogonal group $O(mn)$ (Theorem 2.23). We call this the *uniform distribution*; a random subspace with such distribution will be called a *uniform subspace*. Our first result is the following (Proposition 3.31). Its proof uses the *General Position Theorem* from algebraic geometry ([6, p.109]).

Theorem G (Subspaces spanned by rank–one)**.** *Let $0 \le d \le mn - 1$ and $L \subseteq \mathbb{RP}^{mn-1}$ be the projectivization of a uniform subspace in $G(d + 1, mn)$. With probability one the following trichotomy holds.*

1. *In the case $d < (m - 1)(n - 1)$: L is not spanned by rank–one matrices.*
2. *In the case $d = (m - 1)(n - 1)$: $\#(L \cap \Sigma_{m,n}) \ge d + 1$ implies that L is spanned by rank–one matrices.*
3. *In the case $d > (m - 1)(n - 1)$: $L \cap \Sigma_{m,n} \neq \emptyset$ implies that L is spanned by rank–one matrices.*

In particular, for any $0 \le d \le mn-1$ we have that, with probability one, $\#(L\cap\Sigma_{m,n}) \ge d+1$ implies that L is spanned by rank–one matrices.

Theorem G thus relates the property that a subspace is spanned by rank–one matrices to the number of real intersection points of its projectivization with the Segre variety. Notice that the dimension that discriminates the cases is the codimension of the Segre variety.

The connection to Gaussian tensors is realized by Lemma 3.33, which states that the linear span of the slices of a Gaussian $m \times n \times \ell$ tensor is uniform in $G(\ell, mn)$. As a

consequence we obtain a probabilistic version of Friedland's result. Indeed the rank of a Gaussian $m \times n \times \ell$ tensor T will not be lower than ℓ, with probability one, by the condition in (1.10). In order to have rank equal to ℓ, we need an equality in (1.10), i.e. that the span of the slices is spanned by rank–one. Applying Theorem G we see that this happens if and only if such subspace projectively intersects $\Sigma_{m,n}$ in at least ℓ points (Proposition 3.34).

Applications to typical ranks

We can now apply the geometric framework we developed in Theorem G and Proposition 3.34 to the study of typical ranks. Our main result in this direction is the following (Theorem 3.35).

Theorem H (Bounds on typical ranks)**.** *Let $(m-1)(n-1)+1 \leq \ell \leq mn$. Then the typical ranks of the format $m \times n \times \ell$ are always contained in $\{\ell, \ell+1\}$. Moreover, ℓ is always a typical rank.*

The first part of Theorem H was already proved in [86, Theorem 8.1] via algebraic means for $m \geq 3$. Our proof is instead entirely based on Theorem G and on the observation that given a d–dimensional general subspace in $\mathbb{R}\mathrm{P}^{mn-1}$, the linear span of it and a rank–one matrix not contained in it produces a $(d+1)$–dimensional general subspace in $\mathbb{R}\mathrm{P}^{mn-1}$. The second part of the theorem instead follows from the result below, which can be seen as the translation of Theorem G in terms of rank probabilities (Theorem 3.36).

Theorem I (Rank probabilities via slicing)**.** *Let $T \in \mathbb{R}^{m\times n\times \ell}$ be a Gaussian tensor and let $L \subseteq \mathbb{R}\mathrm{P}^{mn-1}$ be the projectivization of a uniform subspace in $G(\ell, mn)$. Then for $\ell = (m-1)(n-1)+1$ we have*

$$\mathbb{P}\left(\{\mathrm{rank}(T) = \ell\}\right) = \mathbb{P}\left(\{\#(L \cap \Sigma_{m,n}) \geq \ell)\}\right).$$

For $(m-1)(n-1)+1 < \ell \leq mn$ we have

$$\mathbb{P}\left(\{\mathrm{rank}(T) = \ell\}\right) = \mathbb{P}\left(\{L \cap \Sigma_{m,n} \neq \emptyset)\}\right).$$

The second part of Theorem H is then proved by showing that the two probabilities on the right are both non–zero; hence in both cases ℓ is a typical rank. An immediate corollary of Theorem I is that when $\ell > (m-1)(n-1)+1$ and the degree of the complex Segre variety is odd, then ℓ is the only typical rank, as we always have a non–empty intersection on the right (Corollary 3.38). As another consequence of Theorem H we are able to give a new proof that *tall* tensors, i.e. $m \times n \times \ell$ tensors with $(m-1)n < \ell \leq mn$, admit ℓ as their unique typical rank (Theorem 3.40, first proof by ten Berge in [88]).

$3 \times 3 \times 5$ tensors and cubic surfaces

In this part of the work, we aim at characterizing rank probabilities for $3 \times 3 \times 5$ Gaussian tensors, whose typical ranks are known to be 5 and 6 [89]. This will be achieved via a surprising connection with cubic surfaces. Recall that given 6 points in general position in $\mathbb{C}\mathrm{P}^2$, the blow up in those points produces a smooth cubic surface; viceversa every smooth cubic surface in $\mathbb{C}\mathrm{P}^3$ arises that way. The complex Segre variety $\Sigma^{\mathbb{C}}_{3,3}$ has degre 6 and codimension 4. It follows that a general 4–dimensional linear subspace in $\mathbb{R}\mathrm{P}^8 = \mathbb{P}(\mathbb{R}^{3\times 3})$ intersects $\Sigma^{\mathbb{C}}_{3,3}$ in 6 complex points. In [29, Theorem 6.9] they prove that a 4–dimensional linear subspace of $\mathbb{C}\mathrm{P}^8$ intersects the Segre variety $\Sigma^{\mathbb{C}}_{3,3}$ in the six rank–one matrices $x_1 \otimes y_1, \ldots, x_6 \otimes y_6$ if and only if there exists a cubic surface S that is both the blow up

of $\mathbb{C}\mathrm{P}^2$ in $x_1, \dots, x_6$ and the blow up of $\mathbb{C}\mathrm{P}^2$ in $y_1, \dots, y_6$. Moreover, this surface has the determinantal representation

$$S = \left\{ [z_0 : \cdots : z_3] \in \mathbb{C}\mathrm{P}^3 \mid \det(z_0 M_0 + \cdots + z_3 M_3) = 0 \right\}, \tag{1.11}$$

where $M_0, \dots, M_3$ span the orthogonal complement to L. In Theorem 3.41 we adapt this to the case of a general 4–dimensional linear subspace in $\mathbb{R}\mathrm{P}^8$; in this case the corresponding cubic surface S is real and admits the same determinantal representation as in (1.11). This way, given a $3 \times 3 \times 5$ Gaussian tensor T, we can associate to it a random determinantal cubic surface S as follows. Consider the projective 4–dimensional linear subspace given by $L := \langle T_1, \dots, T_5 \rangle_{\mathbb{P}}$. We can apply Theorem 3.41 to L, obtaining the random cubic surface S via a blow up procedure, as described in the theorem. Remark that since L is the projectivization of a uniform subspace in $G(5, 9)$, in this case the matrices $M_0, \dots, M_3$ in (1.11) are Gaussian matrices (Lemma 3.30). Therefore to a $3 \times 3 \times 5$ Gaussian tensor correspond the random determinantal cubic surface in (1.11) given by 3×3 Gaussian matrices $M_0, \dots, M_3$.

It is a classical result of enumerative geometry, dating back to Calyey and Salmon in 1849, that every smooth cubic surface contains exactly 27 lines [24]. Schläfli proved in 1858 that a real cubic surface can only contain either 3, 7, 15 or 27 real lines [77]. In [71] real cubic surfaces are explicitly classified by relating the number of real lines they contain to the number of real points in the blow up of $\mathbb{C}\mathrm{P}^2$ through which they can be built. Exploiting this classification and the association between Gaussian tensors and determinantal cubic surfaces, we prove the following (Theorem 3.46).

Theorem J (Rank probabilities of $3 \times 3 \times 5$ Gaussian tensors)**.** *Let S be the random cubic surface described by (1.11) with $M_0, \dots, M_3 \in \mathbb{R}^{3\times 3}$ Gaussian matrices and let $T \in \mathbb{R}^{3\times 3\times 5}$ be a Gaussian tensor. Then the following probabilities are equal*

$$\mathbb{P}\left(\{\mathrm{rank}(T) = 5\}\right) = \mathbb{P}\left(\{\textit{all the } 27 \textit{ lines on } S \textit{ are real}\}\right).$$

Theorem J follows from the rank of the tensor being 5 if and only if the projectivization of the span of the slices meets the Segre variety in at least 5 real points. Since we are intersecting real varieties, this is equivalent to having all the 6 intersection points real. By the classification of real cubic surfaces, the associated determinantal cubic will contain only real lines. As a by–product of our proofs, we obtain that the expected number of real lines on such a determinantal cubic is between 11 and 15 (Corollary 3.48): this is comparably large with respect to other known models; see [2, 4, 7].

Asymptotics and heuristics

Theorem I shows that for $\ell = (m-1)(n-1)+1$, the probability that a random Gaussian $m \times n \times \ell$ tensor has rank ℓ corresponds to the probability that a uniform slice of the Segre variety contains at least ℓ real points. Using a projective version of Howard's kinematic formula, we prove that the expected number of real points in such intersection is given by

$$\mathbb{E}[\#(L \cap \Sigma_{m,n})] = \sqrt{\pi} \cdot \frac{\Gamma(\frac{m+n-1}{2})}{\Gamma(\frac{m}{2})\Gamma(\frac{n}{2})}. \tag{1.12}$$

In Proposition 3.51 we show that, for m fixed and $n \to +\infty$, the asymptotic expansion of the expectation in (1.12) has principal order $n^{\frac{m-1}{2}}$.

For $m = 3$, the needed number of real points in the intersection to have the smaller rank is $\ell = 2n - 1$, while (1.12) is equal to n. Therefore the number of points we need to be real to have rank $\ell = 2n - 1$ grows faster, even though with the same order, with respect to their average number. We thus expect the probability of having rank $2n - 1$ to go to 0 as $n \to +\infty$. For $m \geq 4$, while ℓ grows linearly in n, the expectation (1.12) grows as a higher power of n, meaning that it is more and more likely to have enough real points in the intersection. Therefore the probability of having rank ℓ in this case approaches 1 as $n \to +\infty$. These are not formal proofs, but should rather be intended as heuristics and geometric considerations that aim to explain what is witnessed in numerical simulations. Formal proofs would require the study of the higher order moments of the random variable $\#(L \cap \Sigma_{m,n})$.

1.3 Nearest point problems in the Grassmannian

A widespread class of problems arising in applications are the so called *nearest point problems*. These are distance optimization tasks. In their most general setting, we are given a point outside of a subset of a metric space and the task is to compute the point of the subset closest to the given one; in other words we want to minimize the distance between the given point and the subset. We will often refer to the external point as the *data point* of the problem. We will always assume that a solution actually exists (maybe not unique); for example this is the case if X is a closed set.

Nearest point problems come in very different shapes, accounting for the variety of settings where they can be found. A classical example are classification tasks. In those problems, the metric space is usually the space of data endowed with a metric that captures the features of the required task, while the subset is given by a finite number of points corresponding to the different classes. The classification task asks, given a data point, to assign it to a class, which is equivalent to finding the closest point, in the given metric, among the ones representing the classes. Nearest point problems like this one where the subset is given by a finite number of points are sometimes called *nearest neighbor search*. In addition to classification problems, they also appear in pattern recognition for image processing and clustering. See [78] for a survey of the main applications and techniques used to solve them.

When X is not a finite set, the general strategy is usually to first compute the critical points of the function giving the distance from the data point restricted to the subset; then we look for the minimizer among them. This method is useful whenever the number of critical points is finite. As already observed in the general introduction, the models arising from applications in the sciences are often described by polynomial equations. They are often realized as *real algebraic varieties* embedded in some Euclidean space representing the space of parameters. In [34] they introduce for the first time a general framework for nearest point problems in this setting, i.e. when the metric space is a Euclidean one and the subset is a real algebraic variety. They show that the equations defining critical points are algebraic and that their solution set gives a finite variety. They define the *Euclidean Distance Degree* (EDD) of a real algebraic variety to be the number of complex solutions to the critical equations for a generic data point. As this gives a bound to the number of real critical points, it serves as a measure of the algebraic complexity of the nearest point problem of fitting data to the algebraic variety. Many well–known instances of optimization problems involving algebraic varieties can be cast in the framework of [34]. This is the case for example of *matrix low–rank approximation*, that we describe in Section 2.6, but also of

problems coming from algebraic statistics, computer vision or control theory.

The algebraic Euclidean case exhibits many desirable properties. First of all, for the generic data point the number of real critical points is always finite and uniformly bounded by the EDD of the variety. Moreover, in the instances where it is not possible to compute the EDD (mainly due to computational power constraints), explicit upper bounds for it are available. It is therefore natural to wonder whether these properties extend to other settings and if a notion similar to the EDD can be defined there. Here we will focus on the case where the metric space is the Grassmannian $G(k, n)$ of k–planes in $\mathbb{R}^n$ endowed with the Riemannian distance induced by the orthogonally invariant metric. The subsets we consider are smooth submanifolds. Under the assumption that the submanifold is also subanalytic[1], we propose the notion of *Grassmann Distance Complexity* (GDC), which we interpret as a bound to the complexity of the nearest point problem in this setting.

Our approach differs from the Euclidean case because the Riemannian distance we consider on the Grassmannian is neither smooth nor semialgebraic, preventing the direct application of techniques from smooth analysis and algebraic geometry. To deal with the non–smoothness, we employ Clarke's notion of critical points for locally Lipschitz functions and we analyze the subgradient of the distance function. To deal with the non–semialgebraicity, we use techniques from o–minimal geometry; this is why we use the term "complexity" rather than "degree". In the very last part of this work, we apply our framework to a class of simple Schubert varieties. For these we completely solve the nearest point problem, finding an explicit algorithm that given a generic data point computes the minimizer.

All the results that we present here have been obtained as joint work with Antonio Lerario. They can be found in the preprint [66]; this is currently under peer review.

1.3.1 Overview of main contributions and sections

In Section 4.1 we discuss the main features of the theory of *Euclidean Distance Degree* developed in [34]: our guiding question will be which of these properties we retain in the Grassmannian setting. The geometry of the Grassmannian and the properties of its Riemannian distance are discussed in detail in the background material (see Section 2.7).

We open Section 4.2 with the precise statement of the nearest point problem for smooth submanifolds of the Grassmannian endowed with the Grassmann metric. As already pointed out, the Grassmann distance function is not smooth; we therefore need a new definition of critical points. This is done in Section 4.2.1 through the notion of Clarke's subgradient and critical points for locally Lipschitz functions. We identify the set of non–differentiability points of the distance function with the *cut locus* of the data point. After a careful description of the cut locus, we compute the subgradient of the distance function at each of its points and characterize its dimension and structure. In Section 4.2.2 we show that for a generic data point, the subgradient of the restriction to a smooth submanifold of the distance function from the data point is the projection onto the tangent space to the submanifold of the subgradient of the unrestricted function. This leads to a characterization of critical points which can be seen as an extension of the *Lagrange*

[1] In this paper the world "subanalytic" means definable in the o–minimal structure of globally subanalytic sets.

multipliers rule to our non–smooth setting.

A major issue in this framework is that there exist submanifolds that admit infinitely many critical points, even for the generic data point; this, a priori, would make our strategy useless. Nevertheless, we prove in Section 4.2.3 one of the key results of our work: for a generic data point, the solution to the nearest point problem cannot not be found on the cut locus of the data point. Section 4.2.4 goes on to prove that, as soon as we exclude critical points on the cut locus, the nearest point problem for a generic data point admits finitely many critical points; this requires the submanifold to be either compact or subanalytic. In the subanalytic case, this follows from definability of the principal angles between k–planes.

In Section 4.2.5 we prove the second key result of our work, namely that for subanalytic submanifolds and generic data points, there exists a uniform bound to the number of critical points outside of the cut locus of the data point. This leads to our definition of the *Grassmann Distance Complexity* (GDC) of a smooth subanalytic submanifold as the maximum number of critical points for a generic data point. It is thus evident that this is a measure of the complexity of solving the nearest point problem in the Grassmannian setting. In Section 4.2.6 we show through an example that this problem is genuinely non–algebraic: this justifies the need for the use of techniques coming from o–minimal geometry and the given name of "complexity"rather than "degree".

In Section 4.2.7 we address the computability of bounds for the GDC of a submanifold. We show that for algebraic subvarities the critical points can be described via a system of *Pfaffian* equations and inequalities. Through the theory of Pfaffian functions, we obtain an explicit bound to the GDC of an algebraic hypersurfaces in terms of the degree of its defining equation. This result can be generalized to any algebraic subvariety for which equations are given.

The last part of the work deals with the nearest point problem for a specific class of Schubert varieties. These are simple Schubert varieties given by those k–planes that intersect another fixed k–plane in (at least) a given dimension; we introduce them in Section 4.3.1. We describe them in terms of varieties of matrices with rank constraints via the Riemannian exponential. Thanks to this description, we are able to show in Section 4.3.2 a surprising correspondence between our nearest point problem and the low–rank approximation task for matrices. In particular, we prove that critical points of the matrix approximation problem give critical points of the nearest point problem via the Riemannian exponential. In Section 4.3.3 we show, through a metric convergence result, that in some situations these critical points are the only ones in a neighborhood of the data point.

Finally, Section 4.3.4 gives the solution of the nearest point problem for these simple Schubert varieties. We indeed show that the minimizer of the low–rank approximation problem for matrices corresponds, through the construction of Section 4.3.2, to the global minimizer we were looking for. Moreover, this is generically unique and is explicitly computable in an algorithmic way. Section 4.3.5 deals with the opposite situation, i.e. we now want to maximize the distance from the data point. It shows, using a variation of the technique employed in Section 4.3.4, that we have a whole Grassmannian of global maxima, all of which are in the cut locus of the data point. This serves as an example of a class of varieties for which the nearest point problem admits infinitely many critical points.

1.3.2 Statement of the results

We begin by recalling some classical background on the Riemannian geometry of the Grassmannian $G(k,n)$ of k–planes in $\mathbb{R}^n$; it will be needed for the setup of the problem. For more details, see Section 2.7. We will always assume that $k \leq n-k$.

We endow the Grassmannian with the unique (up to multiples) orthogonally invariant Riemannian metric (Definition 2.26). Given two k–planes $\mathbf{E}, \mathbf{L} \in G(k,n)$, their Riemannian distance has the expression

$$\delta(\mathbf{E},\mathbf{L}) = \left(\sum_{i=1}^{k} \theta_i(\mathbf{E},\mathbf{L})\right)^{\frac{1}{2}},$$

where $0 \leq \theta_1(\mathbf{E},\mathbf{L}) \leq \cdots \leq \theta_k(\mathbf{E},\mathbf{L}) \leq \frac{\pi}{2}$ are the principal angles between $\mathbf{E}$ and $\mathbf{L}$ (Definition 2.28). We will refer to $\delta : G(k,n) \times G(k,n) \longrightarrow \mathbb{R}_{\geq 0}$ as the *Grassmann distance*.

Throughout this work, it will be useful to identify tangent vectors to $G(k,n)$ with $(n-k) \times k$ matrices as follows. Given $\mathbf{E} \in G(k,n)$, fix orthonormal basis for $\mathbf{E}$ and $\mathbf{E}^{\perp}$. Then any matrix $A \in \mathbb{R}^{(n-k)\times k}$ represents a linear map $\mathbf{E} \longrightarrow \mathbf{E}^{\perp}$ with respect to these basis. It follows that we can associate to any matrix $A \in \mathbb{R}^{(n-k)\times k}$ a tangent vector $v_A \in T_{\mathbf{E}}G(k,n)$; this gives indeed an isomorphism

$$\mathbb{R}^{(n-k)\times k} \cong T_{\mathbf{E}}G(k,n).$$

The orthogonally invariant metric on $G(k,n)$ corresponds through this isomorphism to the Frobenius scalar product for matrices (2.9).

Remark 1.5. The Riemannian exponential $\exp_{\mathbf{E}} : T_{\mathbf{E}}G(k,n) \longrightarrow G(k,n)$ maps the set of matrices with singular values bounded above by $\frac{\pi}{2}$ surjectively onto $G(k,n)$. For such a matrix, the principal angles between $\mathbf{E}$ and $\exp_{\mathbf{E}}(v_A)$ coincide with the singular values of A (see Proposition 2.30).

Given a point $\mathbf{L} \in G(k,n)$, we denote by $\delta_{\mathbf{L}} : G(k,n) \longrightarrow \mathbb{R}_{\geq 0}$ the function computing the Grassmann distance from $\mathbf{L}$, that is $\delta_{\mathbf{L}}(\mathbf{E}) = \delta(\mathbf{E},\mathbf{L})$ for every $\mathbf{E} \in G(k,n)$. We can now state the nearest point problem we aim to solve (Problem 4.9).

Problem 1.6 (The Grassmann nearest point problem)**.** Let $(G(k,n),\delta)$ be the Grassmannian endowed with the Grassmann distance and let $X \subseteq G(k,n)$ be a submanifold. For a generic $\mathbf{L} \notin X$, find the point $\mathbf{L}^* \in X$ that minimizes $\delta_{\mathbf{L}}$, i.e.

$$\delta_{\mathbf{L}}(\mathbf{L}^*) = \min_{\mathbf{E}\in X} \delta_{\mathbf{L}}(\mathbf{E}).$$

We denote the restriction of the distance function to the submanifold by $\delta_{\mathbf{L}}|_X$.

Critical points and the subgradient of the Grassmann distance function

If the distance function were smooth, one would find the critical points by setting $\nabla\delta_{\mathbf{L}}|_X = 0$, where ∇ denotes the usual gradient. However, $\delta_{\mathbf{L}}$ is not smooth everywhere; it is not even semialgebraic, but only real analytic (Proposition 2.32). Nevertheless, $\delta_{\mathbf{L}}$ is locally Lipschitz and hence differentiable almost everywhere by Rademacher's Theorem. We can

therefore apply Clarke's theory of *subgradient*. The subgradient of the locally Lipschitz function $\delta_{\mathbf{L}}$ at the point $\mathbf{E} \in G(k,n)$ is defined as

$$\partial_{\mathbf{E}}(\delta_{\mathbf{L}}) := \operatorname{co}\left\{ \lim_{\mathbf{E}_n \to \mathbf{E}} \nabla\delta_{\mathbf{L}}(\mathbf{E}_n) \;\middle|\; (\mathbf{E}_n)_n \subseteq \omega(\delta_{\mathbf{L}}) \text{ and the limit exists} \right\} \subseteq T_{\mathbf{E}}G(k,n),$$

where co is the convex hull, $\omega(\delta_{\mathbf{L}})$ is the set of differentiability points of $\delta_{\mathbf{L}}$ and $\nabla\delta_{\mathbf{L}}$ is the usual gradient at a differentiability point of $\delta_{\mathbf{L}}$ (Definition 4.10). It is a convex set in the tangent space to $G(k,n)$ at $\mathbf{E}$. At a differentiability point it reduces to a single point, i.e. the standard gradient. For locally Lipschitz functions, critical points are defined through the notion of subgradient: a point is critical if and only if the zero vector belongs to the subgradient at the point (Definition 4.12). Armed with this new definition of what a critical point is, we are interested in finding the critical points of the restriction of the Grassmann distance $\delta_{\mathbf{L}}|_X$. Remark that minima of a locally Lipschitz function are necessarily critical points in the subgradient sense we just described.

The *cut locus* of a point in a Riemannian manifold is the closure of the set of points that are connected to it by more than one shortest geodesic. For a point p we denote it by $\operatorname{cut}(p)$. It is known that the distance function from p induced by a Riemannian metric is not smooth at the point p itself and on the cut locus $\operatorname{cut}(p)$, while it is smooth everywhere else. Applying this to the Grassmannian, the function $\delta_{\mathbf{L}}$ is smooth on $G(k,n) \setminus (\mathbf{L} \cup \operatorname{cut}(\mathbf{L}))$, with gradient at a point $\mathbf{E}$ given by the tangent vector to the unique length minimizing gedoesic from $\mathbf{L}$ to $\mathbf{E}$ (Proposition 4.15). As for non–differentiability points, the cut locus of $\mathbf{L}$ consists of those planes with at least one principal angle with respect to $\mathbf{L}$ equal to $\frac{\pi}{2}$. It is a codimension one Schubert variety (Lemma 4.16) and it can be stratified as

$$\operatorname{cut}(\mathbf{L}) = \bigsqcup_{j=1}^{k} \Omega(j),$$

where $\Omega(j)$ consists of the k–planes for which exactly j principal angles with $\mathbf{L}$ are equal to $\frac{\pi}{2}$. We analyze the structure of the subgradient $\partial_{\mathbf{E}}\delta_{\mathbf{L}}$ for $\mathbf{E} \in \operatorname{cut}(\mathbf{L})$; this is carried out via the computation of the tangent vectors to all the length minimizing geodesics connecting $\mathbf{L}$ to $\mathbf{E}$. This leads to our first result (Corollary 4.20).

Proposition K (The subgradient of the Grassmann distance function)**.** *For every $\mathbf{L} \in G(k,n)$ and $\mathbf{E} \in \Omega(j)$ with $j = 1, \dots, k$, the subgradient of the distance function $\delta_{\mathbf{L}}$ at $\mathbf{E}$ is a convex set linearly isomorphic to the convex hull of the orthogonal group $O(j)$. In particular*

$$\dim(\partial_{\mathbf{E}}\delta_{\mathbf{L}}) = j^2.$$

See Theorem 4.19 for a more detailed statement with the explicit description of the vectors belonging to the subgradient.

We have thus described the subgradient of $\delta_{\mathbf{L}}$ at every point. We now have to understand how this changes when we restrict the distance function to a smooth submanifold $X \subseteq G(k,n)$. For a smooth function with a smooth constraint, the gradient of the restriction is just the projection of the gradient onto the tangent space to the constraint. We show that if the submanifold X intersects the cut locus of $\mathbf{L}$ transversely, condition that is satisfied for a generic $\mathbf{L}$, then this is the case also for the subgradient of $\delta_{\mathbf{L}}$ (Lemma 4.23).

Theorem L (The subgradient of the constrained Grassmann distance)**.** *Let $X \subseteq G(k,n)$ be a smooth submanifold. For a generic $\mathbf{L} \in G(k,n)$ and every $\mathbf{E} \in X$ we have*

$$\partial_{\mathbf{E}}(\delta_{\mathbf{L}}|_X) = \operatorname{proj}_{T_{\mathbf{E}}X}(\partial_{\mathbf{E}}\delta_{\mathbf{L}}),$$

where $\operatorname{proj}_{T_{\mathbf{E}}X}$ denotes the orthogonal projection onto the tangent space $T_{\mathbf{E}}X$.

By definition, a point $\mathbf{E} \in X$ is critical for $\delta_{\mathbf{L}}|_X$ if and only if the zero vector belongs to $\partial_{\mathbf{E}}(\delta_{\mathbf{L}}|_X)$. As an immediate consequence of Theorem L we obtain that this is equivalent to $\partial_{\mathbf{E}}\delta_{\mathbf{L}} \cap N_{\mathbf{E}}X \neq \emptyset$, where $N_{\mathbf{E}}X$ is the normal space to X at the point $\mathbf{E}$ (Corollary 4.24). In case that $\mathbf{E}$ is a smooth point for $\delta_{\mathbf{L}}$, this condition becomes that the gradient of $\delta_{\mathbf{L}}$ is orthogonal to the submanifold: this is the usual Lagrange multipliers rule. Corollary 4.24 can therefore be seen as a direct generalization of the Lagrange multipliers rule to this non–smooth setting.

The journey towards Grassmann Distance Complexity

A notable difference with the algebraic Euclidean case is that critical points may not be isolated; we give an example of this phenomenon in Section 4.3, where we exhibit a class of simple Schubert varieties with positive–dimensional families of critical points. This potentially invalidates our approach to Problem 1.6; indeed the strategy based on finding the critical points and looking for the minima among them is based on the assumption that we only have finitely many points to check. Nevertheless, in those examples the non–isolated critical points all lie in the cut locus of the data point, suggesting that the non–smoothness of the distance function is what might cause this issue. One of the key conceptual results of this work is that, under a mild transversality assumption, the points in the cut locus can never be minima (Theorem 4.25).

Theorem M (Ignoring the cut locus)**.** *Let $X \subseteq G(k,n)$ be a smooth submanifold. If X is transverse to $\mathrm{cut}(\mathbf{L})$, then no point in the intersection $X \cap \mathrm{cut}(\mathbf{L})$ can be a (local) minimum for $\delta_{\mathbf{L}}|_X$.*

From Theorem M it follows that for a generic data point $\mathbf{L}$ the minimizer in Problem 1.6 cannot be found on the cut locus of $\mathbf{L}$. When solving Problem 1.6, we can therefore ignore all the critical points coming from $\mathrm{cut}(\mathbf{L})$.

Since $\delta_{\mathbf{L}}$ is smooth on points outside of the cut locus, ignoring $\mathrm{cut}(\mathbf{L})$ we obtain again a smooth constrained optimization problem. A critical point $\mathbf{S}$ is therefore characterized by the minimizing geodesic connecting it to $\mathbf{L}$ being orthogonal to the submanifold X (Corollary 4.28). Thanks to the smoothness of the normal exponential map, we can prove that critical points not in the cut locus are discrete (Lemma 4.29) and hence obtain the following finiteness statement (Theorem 4.30).

Theorem N (Finiteness of critical points)**.** *Let $X \subseteq G(k,n)$ be a smooth submanifold, either compact or subanalytic. For a generic data point $\mathbf{L} \in G(k,n)$ the set of critical points of $\delta_{\mathbf{L}}|_X$ in $G(k,n) \setminus \mathrm{cut}(\mathbf{L})$ is finite.*

The next natural question is whether the number of critical points in uniformly bounded. Indeed, Theorem N says that for a generic data point we obtain finitely many critical points, but as the data point varies, this number might grow to infinity. In Proposition 2.32 we have shown that the Grassmann distance is definable in the o–minimal structure of subanalytic sets. Using techniques from o–minimality theory, in particular the *triviality lemma* for definable maps, we can prove that if X is also subanalytic the number of critical points is uniformly bounded (Proposition 4.31).

Theorem O. *Let $X \subseteq G(k,n)$ be a smooth subanalytic submanifold. There exists a constant $C_X \in \mathbb{N}$ such that for a generic $\mathbf{L} \in G(k,n)$ the following bound holds*

$$\#\{\mathbf{E} \in X \setminus \mathrm{cut}(\mathbf{L}) \mid \mathbf{E} \text{ is critical for } \delta_{\mathbf{L}}|_X\} \leq C_X. \tag{1.13}$$

We define the *Grassmann Distance Complexity* (GDC) of a smooth subanlytic submanifold X as the maximum of the left hand side of (1.13) for a generic data point $\mathbf{L} \in G(k,n)$. Theorem O guarantees that it is well defined.

Remark 1.7. If X is a stratified submanifold, we will only consider critical points in its smooth stratum. We can therefore extend the definition of GDC to stratified subanalytic submanifolds by defining it to be the GDC of the smooth stratum.

We stress that even if the submanifold X is real algebraic, the set of solutions to the critical equations is not semialgebraic (see Section 4.2.6 for an example). Nevertheless, if X is a real algebraic variety, we show that the critical points are definable in a Pfaffian o–minimal structure (see Section 2.5). In particular, a bound for the GDC is effectively computable. We carry out the explicit computation for hypersurfaces of degree d (Theorem 4.35).

Theorem P (Pfaffian bound for hypersurfaces)**.** *For every $0 \leq k \leq n$, there exist constants $c_1(k,n)$, $c_2(k,n) > 0$ such that for every smooth hypersurface $X \hookrightarrow G(k,n)$ of degree d, the following bound holds*

$$\mathrm{GDC}(X) \leq c_1(k,n)\, d^{c_2(k,n)}.$$

We remark that our method can be extended to any real algebraic variety for which the degrees of the defining equations are known and to pfaffian varieties.

The nearest point problem for simple Schubert varieties

In the last part of the work we apply our framework to the case when X is taken from the following family of simple Schubert varieties. Fix a k–plane $\mathbf{W} \in G(k,n)$ and define for $s = 1, \ldots, k-1$ the variety

$$\Omega_s = \{\mathbf{E} \in G(k,n) \mid \dim(\mathbf{E} \cap \mathbf{W}) \geq s\}.$$

In the same way we did for the cut locus, these varieties can be stratified according to the dimension of the intersection $\mathbf{E} \cap \mathbf{W}$. In particular, the smooth locus of Ω_s is given by those planes such that $\dim(\mathbf{E} \cap \mathbf{W}) = s$.

The main reason we focus on these varieties is that they admit a description in terms of matrices of bounded rank. Denote by $\mathcal{M}_{\leq r}$ the set of $(n-k) \times k$ matrices with rank less or equal than r and by $\mathcal{M}_r$ the set of those with rank exactly r. Let also $R_{\frac{\pi}{2}}$ be the set of $(n-k) \times k$ matrices with singular values less or equal than $\frac{\pi}{2}$. Recall that the dimension of the intersection between two k–planes correspond to the number of 0 principal angles between them. Then by Remark 1.5 we have

$$\Omega_s = \exp_{\mathbf{W}} \left(\mathcal{M}_{\leq k-s} \cap R_{\frac{\pi}{2}} \right), \tag{1.14}$$

where $\exp_{\mathbf{W}} : T_{\mathbf{W}} G(k,n) \longrightarrow G(k,n)$ is the Riemannian exponential of $G(k,n)$ at $\mathbf{W}$. The smooth stratum corresponds to the image of $\mathcal{M}_{k-s}$ in (1.14).

For a generic data point $\mathbf{L} \in G(k,n)$, its intersection with $\mathbf{W}$ is trivial, i.e. no principal angle between them is 0. Moreover, also its intersection with the orthogonal complement $\mathbf{W}^{\perp}$ is trivial, meaning that no angle is $\frac{\pi}{2}$. Thus the preimage of $\mathbf{L}$ through $\exp_{\mathbf{W}}$ inside $R_{\frac{\pi}{2}}$ is given by a unique matrix $A_{\mathbf{L}}$ of full rank. The idea will therefore be the following:

1. given a generic $\mathbf{L} \in G(k,n)$, compute the preimage matrix $A_{\mathbf{L}}$ discussed above; this will be a full rank matrix;
2. use *Eckart–Young Theorem* (Theorem 2.19) to find the critical points of the low–rank approximation problem for $A_{\mathbf{L}}$ with respect to $\mathcal{M}_{\leq k-s}$;
3. map back the so found critical points to Ω_s via $\exp_{\mathbf{W}}$.

Remark that Theorem 2.19 produces critical points of rank exactly $k-s$, so the points we obtain via this procedure are contained in the smooth locus of Ω_s. Applying the exponential map in the third step of the strategy outlined above, one should expect some distortion effect in the metric. Nevertheless, this procedure actually produces critical points for the nearest point problem of Ω_s (Theorem 4.43).

Theorem Q (Critical points via Eckart–Young)**.** *Let $\mathbf{L} \in G(k,n)$ be generic and $A_{\mathbf{L}}$ be the unique matrix in $R_{\frac{\pi}{2}}$ such that $\exp_{\mathbf{W}}(A_{\mathbf{L}}) = \mathbf{L}$. Then the critical points of the low–rank approximation problem for $A_{\mathbf{L}}$ with respect to $\mathcal{M}_{\leq k-s}$ with the Frobenius distance correspond via $\exp_{\mathbf{W}}$ to critical points of $\delta_{\mathbf{L}}|_{\Omega_s}$.*

The points provided by Theorem Q are explicitly computable. Remark that these are not all of the critical points of $\delta_{\mathbf{L}}|_{\Omega_s}$; for example local maxima will never be found among them. Nevertheless, this gives the lower bound

$$\mathrm{GDC}(\Omega_s) \geq \binom{k}{s}.$$

Let now $\widetilde{B} = B(0,\varepsilon)$ be the ball of radius $\varepsilon > 0$ centered at the origin of $\mathbb{R}^{(n-k)\times k} \cong T_{\mathbf{W}}G(k,n)$ (in the Frobenius norm) and $B = \exp_{\mathbf{W}}(\widetilde{B})$ be the corresponding geodesic ball in $G(k,n)$. For $\varepsilon > 0$ small enough, from (1.14) we have a diffeomorphism of pairs

$$\left(\widetilde{B}, \mathcal{M}_{\leq k-s} \cap \widetilde{B}\right) \overset{\exp_{\mathbf{W}}}{\longrightarrow} \left(B, \Omega_s \cap B\right).$$

We are able to prove that the Frobenius metric $\langle \cdot, \cdot \rangle_F$ and the pullback metric $\exp_{\mathbf{W}}^* g_{G(k,n)}$, where $g_{G(k,n)}$ is the Grassmann metric, converge to each other on $\widetilde{B}$ as $\varepsilon \to 0$ (Lemma 4.47). As a consequence, if the data point is close to $\mathbf{W}$, then all the critical points of $\delta_{\mathbf{L}}|_{\Omega_s}$ contained in a small enough ball around $\mathbf{W}$ are given by Theorem Q; see Proposition 4.48 for the full statement.

The closing result of this part of the work is that we are actually able to explicitly solve Problem 1.6 for the simple Schubert variety Ω_s through Theorem Q. Indeed, using a result from linear algebra relating the singular values of a matrix to those of its submatrices, we can bound the principal angles between any point $\mathbf{E} \in \Omega_s$ and the data point $\mathbf{L} \in G(k,n)$, in terms of the angles between $\mathbf{L}$ and $\mathbf{W}$. This way we obtain the following (Theorem 4.51).

Theorem R (The minimizer for Ω_s)**.** *For a generic $\mathbf{L} \in G(k,n)$, the restriction of $\delta_{\mathbf{L}}$ to Ω_s admits a unique global minimizer, corresponding to the minimum given by Eckart–Young Theorem through Theorem Q. This minimizer is generically unique.*

See Theorem 4.51 for an explicit expression of the minimizer in terms of the principal vectors between $\mathbf{L}$ and $\mathbf{W}$ and for the minimum distance in terms of principal angles. This makes the computation of the minimizer algorithmic, as it just boils down to the computation of principal vectors. Via a slight modification of the technique used in the proof of Theorem R we are also able to find the maxima (Theorem 4.54).

Theorem S (The maximizers for Ω_s)**.** *For a generic* $\mathbf{L} \in G(k,n)$*, the restriction of* $\delta_{\mathbf{L}}$ *to* Ω_s *admits a Grassmannian* $G(k-s, n-k-s)$ *of global maximizers.*

See Theorem 4.54 for an explicit expression of the maximum distance and its proof for an explicit description of such maxima points; these are smooth points of Ω_s but they all lie in $\mathrm{cut}(\mathbf{L})$. Theorem S therefore shows that the Ω_s' are examples of varieties that admit infinitely many critical points for a generic data point; this justifies the need for Theorem M.

Chapter 2

Background material

The aim of this chapter is to collect most of the basic tools and background results that will be needed in Chapter 3 and Chapter 4. The material is therefore quite heterogeneous. The presentation of these topics is not meant to be comprehensive by any means, but it has the specific goal of helping the reader getting familiar with the tools that will be used throughout this thesis.

The first four sections present the main tools from differential geometry that we will employ in Section 3.1. We start in Section 2.1 by briefly recalling the definition of second fundamental form and *Weingarten operator* of an embedded submanifold. Section 2.2 then introduces the notion of tubular neighborhood and Section 2.3 discuss that of reach of a submanifold through different points of view. Section 2.4 goes on to describe the celebrated Weyl's tube formula, that allows to compute volumes of tubular neighborhoods of embedded submanifolds.

In Section 2.5 we introduce the theory of Pfaffian functions. They are a useful tool to produce upper bounds to the topology of the set of real solutions to some special systems of equations. We will make use of them in Section 4.2.7 in the context of Grassmann Distance Complexity.

Section 2.6 deals with Eckart–Young theorem, which provides a full solution to the problem of low–rank matrix approximation. We will extensively use the ideas and results from this section throughout all of Chapter 4.

Section 2.7 gives a detailed overview of the geometry of Grassmannians, starting from their construction as homogeneous spaces. As part of our discussion, we introduce a natural orthogonally invariant Riemannian metric on them, which we then study through the notion of principal angles between linear subspaces. This allow to describe geodesics on Grassmannians and define a natural distance, which we call the *Grassmann distance*. We conclude with a paragraph where we easily show that principal angles can be defined in an o–minimal structure; this lemma is original work. This section serves as a solid basis on which all of Chapter 4 will rely.

2.1 The second fundamental form and Weingarten operator

Let $(\overline{M}, g)$ be a smooth Riemannian manifold and M be an embedded n-dimensional smooth submanifold of $\overline{M}$. Denote the embedding as $\iota : M \hookrightarrow \overline{M}$. We can endow M

with a Riemannian metric via the pullback through ι of the metric g; we denote it by $\iota^* g$. We then say that the Riemannian manifold $(M, \iota^* g)$ is an *isometrically embedded submanifold* of $(\overline{M}, g)$. We denote by T_pM, N_pM the tangent and normal spaces to M at $p \in M$. Denote also by $\overline{\nabla}$, ∇ the Riemannian (or *Levi–Civita*) connections of $\overline{M}$ and M respectively. For smooth vector fields X, Y on M, let $\overline{X}$,$\overline{Y}$ be their local extensions to smooth vector fields on $\overline{M}$. Then we have

$$\overline{\nabla}_{\overline{X}} \overline{Y} = \nabla_X Y + B(X, Y),$$

where $\nabla_X Y$ is the tangential component to M and $B(X, Y)$ is the normal one. By the well–known properties of the Riemannian connection, at every point $p \in M$ we can regard B as a symmetric bilinear map $B : T_pM \times T_pM \longrightarrow N_pM$. We call this the *second fundamental form* of the embedding $M \overset{\iota}{\hookrightarrow} \overline{M}$ at $p \in M$; when the embedding is clear from the context we will simply refer to it as the second fundamental form of M. Given a normal direction $\eta \in N_pM$ we also define the *second fundamental form along* η, denoted by $H_\eta : T_pM \times T_pM \longrightarrow \mathbb{R}$, as the projection of the form B along η, i.e. for every $v, w \in T_pM$

$$H_\eta(v, w) = g(B(v, w), \eta).$$

Geometrically, for v,$w \in T_pM$ and $\eta \in N_pM$ the quantity $H_\eta(v, w)$ describes the component along the normal direction η of the variation of the tangent vector w along a curve in M tangent to v. To this bilinear form we can associate a self–adjoint operator, called the *Weingarten operator along* η and denoted by $L_\eta : T_pM \longrightarrow T_pM$, such that for every $v, w \in T_pM$

$$g(L_\eta(v), w) = H_\eta(v, w) = g(B(v, w), \eta).$$

Given $p \in M$, consider a basis $\{e_1, \ldots, e_n\}$ of the tangent space T_pM. Set $H_{\eta,ij} = H_\eta(e_i, e_j)$ and denote by $L_\eta = (L_{\eta,ij})$ the matrix representing the Weingarten operator with respect to the given basis and the scalar product on T_pM defined by g. It is straightforward from the definitions that

$$H_{\eta,ij} = H_\eta(e_i, e_j) = g(L_\eta(e_i), e_j) = L_{\eta,ij}. \tag{2.1}$$

Therefore, computing the matrix representing the Weingarten operator with respect to a given basis is equivalent to the computation of the second fundamental form on the elements of that basis. This shows that the Weingarten operator is just a different way to encode the curvature information contained in the second fundamental form.

Being the Weingarten operator self–adjoint, its eigenvalues are all real. For $\eta \in N_pM$ we call the eigenvalues of L_η the *principal curvatures* of M at p in the normal direction η. We define the *maximal curvature* of M to be the maximum of the absolute value of the principal curvatures over all points of M and all normal directions. We denote it by $\kappa_{\max}(M)$. This will be a key quantity when we will discuss the notion of reach in Section 2.3.

Remark 2.1 (*Second fundamental form of submanifolds of Euclidean spaces and round spheres*)**.** Let M be an n-dimensional isometrically embedded submanifold of a Euclidean space with its standard metric and assume that $\varphi : \mathbb{R}^n \longrightarrow M$ is a given parametrization of M around a point $p \in M$, with coordinates denoted as $a_1, \ldots, a_n$. A basis for T_pM is given by the vectors

$$\frac{\partial \varphi}{\partial a_i} := D_{\varphi^{-1}(p)} \varphi \left(\frac{\partial}{\partial a_i} \right),$$

where $D_{\varphi^{-1}(p)} \varphi$ denotes the differential of the parametrization φ at the point $\varphi^{-1}(p)$. Recall that the Christoffel symbols of the Riemannian connection of the Euclidean space

are all null. It follows that to compute the second fundamental form of M at p along a normal direction η is enough to compute the second derivatives of the parametrization

$$H_\eta\left(\frac{\partial\varphi}{\partial a_i},\frac{\partial\varphi}{\partial a_j}\right)=\left\langle\nabla_{\frac{\partial\varphi}{\partial a_i}}\frac{\partial\varphi}{\partial a_j},\eta\right\rangle=\left\langle\frac{\partial^2\varphi}{\partial a_i\partial a_j},\eta\right\rangle,$$

where $\langle\cdot,\cdot\rangle$ is the Euclidean inner product. Let now M be an isometrically embedded submanifold of the *round sphere* (S^l,g_{S^l}), i.e. the unit sphere S^l with the Riemannian metric obtained by isometrically embedding it in $\mathbb{R}^{l+1}$ with the standard metric. Let $\eta\in N_pM$ be a normal vector to M at $p\in M$. Remark that we can also consider M as an isometrically embedded submanifold of $\mathbb{R}^{l+1}$ by composing the isometric embedding into S^l with that of the round sphere into $\mathbb{R}^{l+1}$. Then the second fundamental form of M along η is the same, whether we consider M as a submanifold of S^l or of $\mathbb{R}^{l+1}$. This means that above formula holds also for submanifolds of round spheres.

We will see that the Weingarten operator will be one of the main ingredients in Weyl's tube formula (see Section 2.4).

2.2 Tubular neighborhoods

Let $M\hookrightarrow\overline{M}$ be an isometrically embedded submanifold of a Riemannian manifold $(\overline{M},g)$. The metric g allows to define the norm of a tangent vector v to $\overline{M}$ as $\|v\|=g(v,v)^{\frac{1}{2}}$. We can therefore define the *ε–small normal bundle* of $M\hookrightarrow\overline{M}$, denoted as $N^\varepsilon M$, as the subset of the normal bundle NM consisting of those vectors with norm less than a given $\varepsilon>0$. Recall the definition of Riemannian exponential map of $\overline{M}$, $\exp:T\overline{M}\longrightarrow\overline{M}$, given for every $p\in\overline{M}$ and $v\in T_p\overline{M}$ by $\exp(p,v)=\gamma_{(p,v)}(1)$, where $\gamma_{(p,v)}$ is the unique geodesic in $\overline{M}$ such that $\gamma_{(p,v)}(0)=p$ and $\dot\gamma_{(p,v)}(0)=v$; we assume for simplicity that the exponential is well defined on the whole tangent bundle $T\overline{M}$. Its restriction to the normal bundle

$$\exp|_{NM}:NM\longrightarrow\overline{M}$$

is called the *normal exponential map*. Its further restriction to $N^\varepsilon M$ is called the *ε–small normal exponential map*.

Definition 2.2 (*Tubular neighborhood*)**.** If there exists an $\varepsilon>0$ such that the ε–small normal exponential

$$\exp|_{N^\varepsilon M}:N^\varepsilon M\longrightarrow\overline{M}$$

is a diffeomorphism on its image, we call $N^\varepsilon M$ a *tubular neighborhood of M in $\overline{M}$*.

The reason for this name will soon be clarified. Recall that the distance of a point $x\in\overline{M}$ from the submanifold M with respect to the metric g is given by

$$d^g(x,M):=\inf\{d^g(x,y)\mid y\in M\},$$

where $d^g(\cdot,\cdot)$ is the Riemannian distance induced by g on $\overline{M}$. We introduce the following set

$$\mathcal{U}(M,\varepsilon):=\{x\in\overline{M}\mid d^g(x,M)<\varepsilon\},\tag{2.2}$$

which consists of those points in $\overline{M}$ at distance less than $\varepsilon>0$ from M. Let us try to understand and visualize this set when M is a closed surface or a closed curve in $\mathbb{R}^3$ with the Euclidean metric. In this case, the distance of a point from the given surface or curve will always be given by the length of a segment starting from the point and meeting the submanifold orthogonally; this follows from Lagrange multipliers rule and the fact that

segments are length minimizing geodesics. The set $\mathcal{U}(M,\varepsilon)$ can therefore be seen as the set of points that you can reach if you move away from the submanifold along straight lines (i.e. geodesics) in all normal directions for a length less than ε, creating a small "tube"around your surface or curve; in other words, they are the points in the image of the ε–small normal exponential map.

The situation we just described can be generalized, as the following theorem shows. Even though this result is well known, we were not able to find a full explicit reference for the general case. We provide here a full proof, filling in the details of the outline given in [21, Theorem 6.6].

Theorem 2.3 (**Tubular neighborhood theorem**). *Let M be a compact isometrically embedded submanifold of a Riemannian manifold $(\overline{M},g)$. Then there exists an $\varepsilon>0$ small enough such that $\exp|_{N^\varepsilon M}: N^\varepsilon M \longrightarrow \overline{M}$ is a diffeomorphism on its image and $\exp(N^\varepsilon M)=\mathcal{U}(M,\varepsilon)$.*

Proof. Throughout the proof, we will identify M with the zero section of the normal bundle NM. We start by computing the differential of the normal exponential map along the zero section $(x,0)\in NM$ for any $x\in M$, i.e.

$$D_{(x,0)}(\exp|_{NM}) : T_{(x,0)}(NM) \longrightarrow T_x\overline{M}.$$

Notice that $\dim(T_{(x,0)}(NM)) = \dim(T_x\overline{M})$, hence the surjectivity of the differential already implies that is a linear isomorphism. As before denote by $\gamma_{(x,v)}$ the unique geodesic on $\overline{M}$ such that $\gamma_{(x,v)}(0)=x$ and $\dot\gamma_{(x,v)}(0)=v$. Let $y\in T_x\overline{M}$. Since the tangent bundle to $\overline{M}$ splits as the orthogonal direct sum $T\overline{M}=TM\oplus NM$, we can decompose y as $y=y_1+y_2$ with $y_1\in T_xM$ and $y_2\in N_xM$. Since y_1 is a tangent vector, there exists a curve $\sigma_1:(-\delta,\delta)\longrightarrow M$ adapted to it, i.e. $\sigma_1(0)=x$ and $\dot\sigma_1(0)=y_1$. Using it, we define a curve on the normal bundle $\sigma:(-\delta,\delta)\longrightarrow NM$ by $\sigma(t)=(\sigma_1(t),0)\in NM$. Since $\sigma(0)=(x,0)$ and $\dot\sigma(0)=(y_1,0)$, it follows that σ is an adapted curve to $(y_1,0)\in T_{(x,0)}(NM)$ and

$$D_{(x,0)}(\exp|_{NM})(y_1,0) = \frac{d}{dt}\exp(\sigma(t))\Big|_{t=0} = y_1,$$

showing that the tangent space T_xM is contained in the image of the differential.
Now consider $y_2\in N_xM$ and define a curve $\alpha:(-\delta,\delta)\longrightarrow NM$ by $\alpha(t)=(x,ty_2)$. Then we have that $\alpha(0)=(x,0)$ and $\dot\alpha(0)=(0,y_2)$, hence α is an adapted curve to $(0,y_2)\in T_{(x,0)}(NM)$. We can compute the differential

$$D_{(x,0)}(\exp|_{NM})(0,y_2)) = \frac{d}{dt}\exp(\alpha(t))\Big|_{t=0} = y_2,$$

therefore also N_xM is contained in the image of the differential. By its linearity it follows that $D_{(x,0)}(\exp|_{NM})$ is surjective and hence an isomorphism.

As a consequence of the differential being an isomorphism, for every $x\in M$ there exists an open neighborhood W_x of $(x,0)$ in NM such that the rank of the differential $D_{(q,v)}(\exp|_{NM})$ is maximal for every $(q,v)\in W_x$. Up to shrinking the neighborhood, we can assume that $W_x=(U_x\times B(0,\varepsilon_x))\cap NM$, where U_x is an open neighborhood of $x\in M$, $B(0,\varepsilon_x)$ is the ball of radius ε_x centered at the origin in $T_x\overline{M}$, and $\exp|_{W_x}$ is an embedding. By compactness we can find a finite set $\{x_1,\dots,x_r\}$ such that $\{U_{x_1},\dots,U_{x_r}\}$ is a finite covering of M. Choosing $\varepsilon:=\min\{\varepsilon_{x_1},\dots,\varepsilon_{x_r}\}$ we obtain that

$$\exp|_{N^\varepsilon M} : N^\varepsilon M \longrightarrow \overline{M}$$

is an immersion and a local embedding. Remark that these properties hold for every $\tilde{\varepsilon} \leq \varepsilon$. We claim that there exists an $\tilde{\varepsilon} < \varepsilon$ such that the restriction of exp to $N^{\tilde{\varepsilon}}M$ is also globally injective. If this is the case, then the restriction to the closure of the $\frac{\tilde{\varepsilon}}{2}$–small normal bundle is an embedding, since injective immersions with compact domain are embeddings. It follows that any number less than $\frac{\tilde{\varepsilon}}{2}$ satisfies the statement of the theorem.

We are only left to prove the claim. We argue by contradiction: suppose that for every $n \in \mathbb{N}$ there exist (x_n, v_n), $(y_n, w_n) \in N^{\frac{1}{n}}M$ such that $\exp(x_n, v_n) = \exp(y_n, w_n)$. By compactness of M, up to restricting to subsequences we can assume that x_n converges to $\overline{x} \in M$ and y_n converges to $\overline{y} \in M$, while v_n and w_n both converge to 0 since they belong to $B(0, \frac{1}{n})$ for every $n \in \mathbb{N}$. Moreover, again by compactness, there exists $r > 0$ such that for every $p \in M$ the map $\exp_p : B(0, r) \subset T_p\overline{M} \longrightarrow \overline{M}$ is a diffeomorphism on its image, where $\exp_p(z) = \exp(p, z)$. It follows that since $v_n \longrightarrow 0$, for n large enough $\gamma_{(x_n, v_n)}$ will be the unique geodesic joining x_n with $\exp_{x_n}(v_n) = \exp(x_n, v_n) = \gamma_{(x_n, v_n)}(1)$ and $d^g(x_n, \exp(x_n, v_n)) = \|v_n\|$. Analogously, for n large enough we will also have $d^g(y_n, \exp(y_n, w_n)) = \|w_n\|$, where we stress that the uniformity of r is crucial. By hypothesis $\exp(x_n, v_n) = \exp(y_n, w_n)$, hence by the triangular inequality we obtain

$$d^g(x_n, y_n) \leq d^g(x_n, \exp(x_n, v_n)) + d^g(y_n, \exp(y_n, w_n)) = \|v_n\| + \|w_n\| \longrightarrow 0,$$

and this forces $\overline{x} = \overline{y}$. Then, for n sufficiently large, we have that (x_n, v_n), $(y_n, w_n) \in W_p$ for some $p \in M$. Since on every W_p we have a local embedding, this is a contradiction. The proof is concluded. □

Theorem 2.3 can be seen in two different ways. On one hand, it can be considered as an existence result for tubular neighborhoods of compact submanifolds, as it states the existence of a small enough $\varepsilon > 0$ such that the ε–small normal exponential map is a diffeomorphism on its image. On the other hand, it provides a characterization of the set $\mathcal{U}(M, \varepsilon)$ for ε small enough. For this reason, in the following, we will refer also to $\mathcal{U}(M, \varepsilon)$ as a tubular neighborhood of M.

Remark 2.4. The compactness assumption in Theorem 2.3 is crucial. Without it, we can only prove the existence of a smooth function $\varepsilon(\cdot) : M \longrightarrow \mathbb{R}_{>0}$ such that the restriction of the exponential map to $N^{\varepsilon(\cdot)}M$ is an embedding, where $N^{\varepsilon(\cdot)}M = \{v \in N_xM \mid x \in M,\ \|v\| < \varepsilon(x)\}$, i.e. the radius ε is not uniform anymore but it depends on the point. In this case the existence of a tubular neighborhood is not guaranteed anymore. For example, in the non–compact curve in Fig. 2.1, to avoid any overlap the function $\varepsilon(\cdot)$ should go to zero as the point moves to the left along one of the two parts of the curve; in this case no tubular neighborhood can exist.

2.3 The reach of a submanifold

Let $M \hookrightarrow \overline{M}$ be an isometrically embedded submanifold of a Riemannian manifold $(\overline{M}, g)$ as in the previous sections. Recalling Definition 2.2 we can give the following definition.

Definition 2.5 (*Reach of a submanifold*)**.** The quantity

$$\rho(M) = \sup\{\varepsilon \geq 0 \mid N^{\varepsilon}M \text{ is a tubular neighborhood of } M\}$$

is called the *reach* of M.

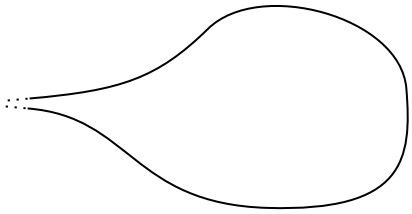


Figure 2.1: Example of a curve with no tubular neighborhood, i.e. with zero reach. The problem here is that the two parts of the curve on the left come infinitely close to each other.

It follows straight from above definition that Theorem 2.3 can be restated by saying that the reach of a compact submanifold is always positive. On the other hand, it is not positive in general, as the submanifold in Fig. 2.1 shows; no tubular neighborhood of radius greater than 0 exists in that case. Remark that even if for brevity we just write $\rho(M)$ and talk about the reach of a submanifold, it is not an intrinsic property of M but it strongly depends on the specific embedding $M \hookrightarrow \overline{M}$. It is therefore a property of the embedding rather than the submanifold.

We now introduce an intuitive geometric object that will allow us to get a more visual understanding of what the reach represents. The main reference for the rest of this section is [17, Chapter 7] and the references therein. Even though in [17] they only consider smooth subvarieties of Euclidean spaces, all the definitions and proofs in the part of our interest straightforwardly generalize to any smooth submanifold embedded in a Riemannian ambient manifold; we will avoid to stress this in what follows.

Given a subset X of a Riemannian manifold $(\overline{M}, g)$, we define the *medial axis* of X, denoted as $\mathrm{Med}(X)$, as the set of points $p \in \overline{M}$ such that there exists more than one closest point to p on $\overline{M}$, i.e. there is more than one point realizing the distance between X and p. The first observation is that by Definition 2.5 every point at distance less than $\rho(M)$ from M admits a unique closest point on M and it therefore cannot be contained in the medial axis of M. It follows that the distance between M and its medial axis $\mathrm{Med}(M)$ is at least the reach of M. Indeed, it is easily shown that equality holds.

Proposition 2.6 ([17], Proposition 7.2)**.** *The distance from a smooth submanifold M to its medial axis* $\mathrm{Med}(M)$ *is the reach of M.*

Remark that Proposition 2.6 can be used as an alternative definition of reach that works also in singular cases. Now that we have seen a different, more intuitive way of looking at the reach, we want to understand how to compute it. In order to do this, we go back to Definition 2.5: we need to understand when the ε–small normal exponential map of M ceases to be a diffeomorphism, as $\varepsilon > 0$ increases.

If $\exp|_{N^\varepsilon M}$ is not a diffeomorphism, then it is either not injective or not an immersion. Assume that $(p, \eta) \in NM$ is such that the normal exponential at (p, η) is not an immersion, i.e. the differential at that point is not injective. The corresponding critical value $\exp(p, \eta)$ is called a *focal point* of M. If this is the case, it can be shown, via computing the determinant of the differential, that $\|\eta\|^{-1}$ is a principal curvature of M at p in the normal direction η [17, (6.15)]. It follows that the supremum of the values of ε for which the ε–small normal exponential is an immersion coincides with the reciprocal of the maximal

curvature $\kappa_{\max}(M)$.

Let now $p, q \in M$ and let $\gamma : [0,1] \longrightarrow \overline{M}$ be an ambient geodesic such that $\gamma(0) = p$ and $\gamma(1) = q$. If such geodesic is orthogonal to M at its endpoints, i.e. $\dot{\gamma}(0) \in N_pM$ and $\dot{\gamma}(1) \in N_qM$, then we say that the couple $\{p, q\}$ is a *bottleneck* of M. We call half of the length of γ the *bottleneck width of* $\{p, q\}$ and we denote it by $\mathrm{bw}(p, q)$, while we call the midpoint of γ the *center* of the bottleneck. Remark that if γ was also length minimizing, then $\mathrm{bw}(p, q)$ corresponds to half the distance between p and q. We define the quantity

$$\mathrm{bw}(M) = \min_{\substack{\{p,q\} \\ \text{bottleneck of } M}} \mathrm{bw}(p, q)$$

the *bottleneck width of* M. It is evident that if $\{p, q\}$ is a bottleneck of M and $\varepsilon := \mathrm{bw}(p, q)$, then the normal exponential is not injective for radii greater than ε, as the center of the bottleneck will have more than one preimage. Therefore the bottleneck width of M bounds from above the reach.

Let now $\varepsilon := \rho(M)$ and assume that there exists a point $u \in \overline{M}$ that has at least two preimages through the ε–small normal exponential, say (x_1, η_1) and (x_2, η_2). If u is not a focal point of M, then it can be shown that $\{x_1, x_2\}$ form a bottleneck of width ε [17, Theorem 7.8]. All the previous discussion is summarized in the following result.

Theorem 2.7 ([17], Theorem 7.8)**.** *Let $M \hookrightarrow \overline{M}$ be an isometrically embedded submanifold of a Riemannian manifold $(\overline{M}, g)$. Then we have*

$$\rho(M) = \min\left\{\mathrm{bw}(M), \frac{1}{\kappa_{\max}(M)}\right\}.$$

Theorem 2.7 shows that the computation of the reach of an embedding $M \hookrightarrow \overline{M}$ can be decoupled into the computation of the bottleneck width and the maximal curvature of M. For what concerns the bottleneck width, it is clear from the definition that

$$\mathrm{bw}(M) = \frac{1}{2}\min\big\{l(\gamma) \mid \gamma : [0,1] \longrightarrow \overline{M} \text{ geodesic s.t. } \gamma(0), \gamma(1) \in M, \\ \dot{\gamma}(0) \in N_{\gamma(0)}M, \dot{\gamma}(1) \in N_{\gamma(1)}M\big\}, \tag{2.3}$$

where $l(\gamma)$ denotes the length of the geodesic γ. As for the maximal curvature, for submanifolds of round spheres we have the following expression

$$\kappa_{\max}(M) = \sup_{x \in M} \sup\big\{\|\ddot{\gamma}(0)\| \mid \gamma : (-\delta, \delta) \longrightarrow M \text{ arclength geodesic in } M \\ \text{with } \gamma(0) = x\big\}. \tag{2.4}$$

As one can imagine from (2.3) and (2.4), the evaluation of such expressions can be extremely computationally heavy, as they in general require to check every possible ambient geodesic between points of M that starts orthogonally to M for (2.3) and every geodesic on the submanifold M for (2.4); it goes without saying that these computations might be unfeasible, unless the geometry of M helps us in reducing the complexity of our task. This is exactly what will happen in Section 3.1.4 when we will compute the reach of the spherical Veronese variety. We will be able to greatly simplify (2.3) and (2.4) exploiting the symmetries of the variety, thus making the computation easier and actually feasible.

Remark 2.8. Bottlenecks of submanifolds/varieties play an important role in optimization problems. In the algebraic case, for a general variety the number of bottlenecks can be estimated using classical invariants such as Chern classes and polar classes [31]. For a general discussion on this subject in the context of algebraic varieties see [31, 30].

In the next section we will see the importance that the knowledge of the reach of an embedding plays when computing the volume of tubular neighborhoods.

2.4 Weyl's tube formula

In [94] Weyl presented a fundamental work, answering a question posed by Harold Hotelling: how can we compute the volume of a "tube", i.e. a tubular neighborhood of fixed radius in the sense of (2.2), around a closed n–dimensional submanifold in $\mathbb{R}^N$ or S^N? Hotelling himself had answered the case of curves ($n = 1$), both in the Euclidean and in the spherical setting. Weyl extended these results to any dimension n as follows.

Theorem 2.9 (**Weyl's tube formula**). *Let M be a smooth, n–dimensional, compact submanifold isometrically embedded in the Euclidean space $\mathbb{R}^N$ with its standard metric or in a round sphere S^N. Then, for $\varepsilon < \rho(M)$, the following formula holds:*

$$\mathrm{Vol}\left(\mathcal{U}(M,\varepsilon)\right) = \sum_{\substack{0\le e\le n \\ e \ even}} K_{c+e}(M) J_{N,c+e}(\varepsilon), \tag{2.5}$$

where $c := N - n$ is the codimension of M and $J_{N,c+e}$ are linearly independent functions of ε only. In the Euclidean case, the universal functions J have the following form:

$$J_{N,k}(\varepsilon) := \varepsilon^k, \tag{2.6}$$

while in the spherical one, they are given by

$$J_{N,k}(\varepsilon) := \int_0^{\varepsilon} (\sin\theta)^{k-1} (\cos\theta)^{N-k}\, d\theta = \int_0^{\tan\varepsilon} \frac{t^{k-1}}{(1+t^2)^{\frac{N+1}{2}}}\, dt. \tag{2.7}$$

Moreover, the coefficients $K_j(M)$ are isometric invariants of M.

Definition 2.10. The coefficients $K_j(M)$ in (2.5) are called the *curvature coefficients* of M.

The motivation for the name in Definition 2.10 accounts for the fact that such coefficients are integrals of functions depending on the second fundamental form of the embedding of M. Notice that in the original paper [94] Weyl uses different normalization constants for the universal functions (2.6) and (2.7). We choose to follow the normalization introduced by Nijenhuis in [70], which proves to be handier for applications, see for instance [19].

Weyl's tube formula presents some remarkable aspects. Firstly, the functions $J_{N,k}$ do not depend on the specific submanifold M but only on its dimension and codimension and on the kind of ambient space (Euclidean or spherical). Moreover they are easily computable, as shown by their explicit expressions (2.6) and (2.7). The second aspect to highlight is that the curvature coefficients are isometric invariants of M. This means that different, but isometric, embeddings of M will still give the same curvature coefficients.

From what we just observed, it might seem that tubular neighborhoods of different isometric embeddings of the same manifold M have the same volumes, so that the specific embedding plays no role in the formula; this is not the case. The idea behind Weyl's tube formula is relatively simple. Since we want to compute the volume of a set in the ambient space, we just integrate the volume form of the ambient space over that set. The whole point is then to parametrize the neighborhood in a clever way; this is done through the normal exponential map of M. The tubular neighborhood theorem (Theorem 2.3) ensures that, since we are assuming that M is compact, if we choose a radius ε small enough, namely less than the reach of M which is strictly positive, than the ε–small normal

exponential gives a good parametrization of the tubular neighborhood $\mathcal{U}(M,\varepsilon)$. Computing the Jacobian of the parametrization, together with some algebraic manipulations, then completes the proof of (2.5). The embedding thus determines up to which radius of the neighborhood Weyl's tube formula holds, since the reach is strongly dependent on the specific embedding.

While the proof of the existence of formula (2.5) is relatively straightforward as we outlined above, the most difficult (and maybe interesting) part of Weyl's paper is the proof of the isometric invariance of the curvature coefficients $K_k(M)$. Since we will not need such invariance property later on, we do not discuss it here and refer the reader to [94] for the original proof, while for a more accurate and modern treatment of the topic we refer to [45] and [10, Chapter 6]. Moreover, we remark that in [51] Howard contextualizes Weyl's result in the framework of *integral geometry*. There he considers more general integrals of polynomials on the second fundamental form of submanifolds of homogeneous spaces.

As a by–product of the proof of (2.5), there are explicit integral versions of formula (2.5) for both the Euclidean and spherical cases. For the latter, which is the only case we will need in Section 3.1, this reads as

$$\mathrm{Vol}\big(\mathcal{U}(M,\varepsilon)\big) = \int_{p\in M}\int_{t=0}^{\tan\varepsilon}\int_{S(N_pM)} \frac{t^{c-1}\det\left[I_n - tL_\eta\right]}{(1+t^2)^{\frac{N+1}{2}}}\,\mathrm{vol}_M\, d\eta\, dt, \tag{2.8}$$

where $S(N_pM)$ denotes the unit sphere in N_pM, I_n is the $n\times n$ identity matrix, L_η is the Weingarten operator of $M\hookrightarrow S^N$ along the unit normal vector η, and $d\eta$ is a short notation for the volume form on $S(N_pM)$. If one explicitly develops the determinant, it is easy to get back formula (2.5).

Remark 2.11. We cannot have a formula with the properties described in Theorem 2.9 for any ambient Riemannian manifold $\overline{M}$. In particular, it is easy to see that a necessary condition for such a formula to exist is that $\overline{M}$ has constant scalar curvature. Indeed, given $\overline{M}$, consider its simplest possible submanifold: a single point $M=\{p\}\hookrightarrow\overline{M}$. A tubular ε–neighborhood of M in $\overline{M}$ is given by a geodesic ball of radius ε in $\overline{M}$ centered at p. Then, for ε small enough, we have that the volume of such a neighborhood is given by

$$\mathrm{Vol}\big(\mathcal{U}(\{p\},\varepsilon)\big) = \mathrm{Vol}\big(B_{\mathbb{R}^N}(\varepsilon)\big)\left(1-\frac{K(p)}{12}\varepsilon^2+\mathcal{O}(\varepsilon^4)\right),$$

where $B_{\mathbb{R}^N}(\varepsilon)$ is a ball of radius ε in $\mathbb{R}^N$ with $N=\dim(\overline{M})$ and $K(p)$ is the scalar curvature of $\overline{M}$ at p. It follows that if the ambient manifold $\overline{M}$ does not have constant scalar curvature, the coefficients in above formula cannot be isometric invariants of the submanifold $M=\{p\}$.

2.5 Pfaffian functions

In this section we will introduce the basic notions on Pfaffian functions and briefly discuss the main results of the theory of Pfaffian systems. These are systems whose equations have a peculiar structure that allows to obtain bounds on the topology of the solution set. In particular, when the system is 0–dimensional, the theory provides upper bounds on the number of non–degenerate solutions. This setting was first introduced by Khovanskii ([56], [57]) in the first half of the '80s and later further developed and generalized by Gabrielov, Vorobjov, Zell and others. Our main reference will be [43].

A *Pfaffian chain* is an ordered sequence of analytic functions satisfying a triangular system of polynomial PDEs in a given domain. More precisely, a Pfaffian chain of *order* $r \geq 0$ and *degree* $\alpha \geq 1$ is an ordered sequence of analytic functions $\{f_1, \ldots, f_r\}$ on an open domain $\mathcal{G} \subseteq \mathbb{R}^n$ such that their partial derivatives satisfy the conditions

$$\frac{\partial f_i}{\partial x_j}(x) = P_{ij}(x, f_1(x), \ldots, f_i(x))$$

for every $i = 1, \ldots, r$ and $j = 1, \ldots, n$, where $P_{ij}(x_1, \ldots, x_n, y_1, \ldots, y_i)$ are polynomials with degrees bounded by α.

A *Pfaffian function* of *order* $r \geq 0$ and *degrees* (α, β) on an open domain $\mathcal{G} \subseteq \mathbb{R}^n$ is a function f such that

$$f(x) = P(x, f_1(x), \ldots, f_r(x)),$$

where $\{f_1, \ldots, f_r\}$ is a Pfaffian chain of order r and degree α on the domain $\mathcal{G}$ and $P(x_1, \ldots, x_n, y_1, \ldots, y_r)$ is a polynomial in $n + r$ variables with degree bounded by β. The triple (r, α, β) is called the *format* of the Pfaffian function f.

Remark 2.12. Every function contained in a Pfaffian chain of order r and degree α is also by definition a Pfaffian function of format $(r, \alpha, 1)$.

Example 2.13 (*The Pfaffian chain of sine and cosine*)**.** Let $\mathcal{G} = \{x \in \mathbb{R} \mid x \neq \pi + 2k\pi \ \forall k \in \mathbb{Z}\}$ and define the ordered sequence of functions on $\mathcal{G}$ given by $\{f_1, f_2, f_3, f_4\}$ where $f_1(x) := \tan\left(\frac{x}{2}\right)$, $f_2(x) := \cos^2\left(\frac{x}{2}\right)$, $f_3(x) := \cos(x)$ and $f_4(x) := \sin(x)$. We have the following identities involving the derivatives of $f_1, \ldots, f_4$:

$$\begin{aligned} f_1'(x) &= \frac{1}{2}\left(1 + f_1^2(x)\right) \\ f_2'(x) &= -f_1(x) f_2(x) \\ f_3'(x) &= -2 f_1(x) f_2(x) \\ f_4'(x) &= f_3(x). \end{aligned}$$

It follows that (f_1, f_2, f_3, f_4) is a Pfaffian chain of order 4 and degree 2 on $\mathcal{G}$.

Given a Pfaffian chain $\mathcal{C} = \{f_1, \ldots, f_r\}$, the family of all Pfaffian functions on the chain $\mathcal{C}$ is closed under partial differentiation. The following lemma, whose proof is immediate, describes the format of partial derivatives.

Lemma 2.14 ([43], Lemma 2.5)**.** *A partial derivative of a Pfaffian function of format* (r, α, β) *is a Pfaffian function of format* $(r, \alpha, \alpha + \beta - 1)$ *on the same defining chain.*

A *basic semi–Pfaffian* set $S \subset \mathbb{R}^n$ in a domain $\mathcal{G} \subseteq \mathbb{R}^n$ is a set defined by sign conditions on a family of Pfaffian functions defined through a common Pfaffian chain in $\mathcal{G}$. A *semi–Pfaffian set* is a finite union of basic semi-Pfaffian sets. Remark that the condition of having a common defining Pfaffian chain is not restrictive, as the concatenation of Pfaffian chains still gives a Pfaffian chain, even though it will have a higher order. The topology of a semi–Pfaffian set can be explicitly bounded in terms of how many Pfaffian functions are involved in its definition and their formats. Several different bounds can be found in the literature (see also [97]); we will use the following one.

Theorem 2.15 ([43], Theorem 3.4)**.** *Let* $S \subseteq \mathcal{G} \subseteq \mathbb{R}^n$ *be a semi-Pfaffian set defined by a conjunction of equations and strict inequalities involving* s *different Pfaffian functions having a common Pfaffian chain. Let* (r, α, β) *be the format of the Pfaffian functions. Then the sum of the Betti numbers of* S*, denoted by* $b(S)$*, is bounded by*

$$b(S) \leq s^n 2^{\frac{r(r-1)}{2}} \mathcal{O}\left(n\beta + \min\{n, r\}\alpha\right)^{n+r}.$$

We will make use of this result in Section 4.2.7 to exhibit a computable upper bound to the Grassmann Distance Complexity of an algebraic hypersurface of a Grassmannian.

Remark 2.16. Here we need to make a disclaimer. While the existence of upper bounds to the topology of a semi-Pfaffian set or to the number of solutions of a 0–dimensional system of Pfaffian functions in terms of their formats is extremely interesting and important from a theoretical point of view, we should specify that the bounds obtained are often too large to be of any practical use. The technique through which these bounds are obtained is via a sequence of modifications of the defining system in order to the reconduce it to a polynomial one exploiting the Pfaffian structure of the equations. In order to do this, the size of the system is doubled at each reduction step and the number of variables involved, as well as the degrees of the final polynomial equations, gets larger and larger, resulting in final bounds which are off by several orders of magnitude and of little (if any) practical utility.

2.6 Low–rank matrix approximation: the Eckart–Young theorem

A recurring problem in applications is to find a suitable way of approximating a given matrix with another one of smaller rank. Instances of this appear in a wide variety of different settings, going from machine learning and data analysis, where it has applications to dimensionality reduction and denoising of data sets, to image processing, signal processing and tomography, just to name a few. The approximation has to be done in such a way that the loss of information, once we fix the rank of the approximating matrix, is the least possible. Geometrically, this corresponds to minimizing the distance between the original matrix and its approximation with respect to some meaningful metric. Even though the applications we mentioned are quite recent, the problem is very natural and classical. A solution for the approximation problem in the setting that we will describe below was indeed first developed in a paper by Eckart and Young dating back to 1937 [35], where they needed such approximation in the context of statistics.

A natural metric on the space of $n \times m$ matrices is the one induced by the *Frobenius scalar product*. This is defined as follows: for $A = (a_{ij}) \in \mathbb{R}^{n\times m}$ and $B = (b_{ij}) \in \mathbb{R}^{n\times m}$ the Frobenius scalar product between A and B is given by

$$\langle A, B\rangle_F := \mathrm{tr}(A^\top B) = \sum_{i,j=1}^{n,m} a_{ij}b_{ij}, \tag{2.9}$$

where tr denotes the trace of a square matrix. Remark that this is just the standard Euclidean scalar product on $\mathbb{R}^{nm}$ when you vectorize matrices and see them as a vectors. The corresponding induced distance on $\mathbb{R}^{n\times m}$ is

$$\widehat{\delta}(A, B) := \|A - B\|_F = (\langle A - B, A - B\rangle_F)^{\frac{1}{2}} = \left(\sum_{i,j=1}^{n,m} (a_{ij} - b_{ij})^2\right)^{\frac{1}{2}}, \tag{2.10}$$

where $\|\cdot\|_F$ denotes the norm induced by the Frobenius scalar product. In order to rigorously state the approximation problem, we introduce the following notation for sets of matrices with rank constraints.

Definition 2.17. Let $r, m, n \in \mathbb{N}$ with $r \leq \min\{n, m\}$. We define:

$$\begin{aligned} \mathcal{M}_{\leq r} &:= \{A \in \mathbb{R}^{n\times m} \mid \mathrm{rank}(A) \leq r\}, \\ \mathcal{M}_r &:= \{A \in \mathbb{R}^{n\times m} \mid \mathrm{rank}(A) = r\}. \end{aligned} \tag{2.11}$$

Recall that the set $\mathcal{M}_{\leq r}$ is defined by the vanishing of all $(r+1) \times (r+1)$ minors of a matrix and is therefore a real algebraic variety; the set of its smooth points is exactly given by $\mathcal{M}_r$.

Using the just introduced notation, the low–rank approximation problem becomes the following: given a matrix $A \in \mathbb{R}^{n \times m}$ and $r < \operatorname{rank}(A)$, find a matrix of rank at most r that minimizes the Frobenius distance from A. In other words, we are asked to minimize the function

$$\widehat{\delta}_A : \mathbb{R}^{n \times m} \to \mathbb{R}, \quad \widehat{\delta}_A(X) := \widehat{\delta}(A, X) = \|X - A\|_F,$$

constrained to the set $\mathcal{M}_{\leq r}$.

A complete answer to this problem is given by the *Eckart–Young Theorem* [35]. Their result is based on the *Singular Values Decomposition* (SVD) of a matrix. Assuming without loss of generality that $m \leq n$, an SVD of the matrix $A \in \mathbb{R}^{m \times n}$ is a factorization $A = U\Sigma V^\top$ where $U \in O(n)$ and $V \in O(m)$ are orthogonal matrices, while $\Sigma = \operatorname{diag}(\sigma_1, \ldots, \sigma_m) \in \mathbb{R}^{n \times m}$ is a diagonal rectangular matrix with nonnegative entries, i.e. $\sigma_i \geq 0$ for all $i = 1, \ldots, m$. It is a basic fact in linear algebra that any matrix admits such a decomposition. The elements $\sigma_1, \ldots, \sigma_m$ are called *singular values* of A; they correspond to the square roots of the eigenvalues of $A^\top A$, which are all real and nonnegative given the symmetry of $A^\top A$. It then follows from the definition of Frobenius norm and from the properties of the trace of a square matrix that one has

$$\|A\|_F = \left(\operatorname{tr}(A^\top A)\right) = \left(\sum_{i=1}^{m} \sigma_i^2\right)^{\frac{1}{2}}.$$

Remark 2.18 (*Non-uniqueness of the SVD*)**.** The SVD of a matrix $A \in \mathbb{R}^{n \times m}$ is not unique. A first source for the lack of uniqueness comes from the different possible orderings of the singular values and consequent reordering of the corresponding columns of U and V. Even if we fix an ordering of the singular values, non-uniqueness still arises in the case of repeated eigenvalues. Lastly, if $m < n$, then the last $n - m$ columns of the orthogonal matrix $U \in O(n)$ give no contribution in the decomposition $A = U\Sigma V^\top$. Therefore different orthogonal completions of the first m columns of U will still give the same matrix A.

We are now ready to state the Eckart–Young Theorem, which provides a solution to the rank approximation problem for matrices, under some mild assumptions on the singular values of the matrix we want to approximate. The idea of the theorem is that critical points can be found by setting some of singular values of the matrix to 0, in order to obtain the wanted rank.

Theorem 2.19 ([35])**.** *Let $A \in \mathbb{R}^{n \times m}$ be a full rank matrix with distinct singular values. Let $A = U\Sigma V^\top$ be a singular values decomposition with singular values ordered as $\sigma_1 > \cdots > \sigma_m > 0$ (we assume $m \leq n$). For any subset $I \subseteq \{1, \ldots, m\}$ with $|I| = r$ define Σ_I as the rectangular $n \times m$ diagonal matrix with entries $(\Sigma_I)_{ii} = \sigma_i$ if $i \in I$ and $(\Sigma_I)_{ii} = 0$ if $i \notin I$. Define also $A_I = U\Sigma_I V^\top$, that is a rank–r matrix. Then:*

(i) the critical points of the Frobenius distance function $\widehat{\delta}_A$ constrained to the smooth manifold $\mathcal{M}_r$ are all and only the matrices A_I with $|I| = r$. In particular, restricted to this manifold, the function $\widehat{\delta}_A$ has $\binom{m}{r}$ critical points.

(ii) for every index set I with $|I| = r$ we have

$$\|A - A_I\|_F^2 = \sum_{i \notin I} \sigma_i^2. \tag{2.12}$$

In particular $A_{\{1,\dots,r\}}$ *is a global minimum of the restriction of* $\widehat{\delta}_A$ *to* $\mathcal{M}_{\leq r}$ *and*

$$\widehat{\delta}(A, \mathcal{M}_{\leq r}) = \|A - A_{\{1,\dots,r\}}\| = \sqrt{\sum_{i=r+1}^{m} \sigma_i^2}.$$

A slightly modified version of Theorem 2.19 covers the case when the singular values of A are not distinct; in that case the minimum might or might not be unique. Nevertheless, the idea of the approximation and formula (2.12) are still valid.

Remark 2.20. In Section 4.1 we will introduce what are usually called *nearest point problems* (see Problem 4.1). This class of problems requires, given a closed set X in a metric space and a point u outside of it, to find a point in X that is closest to u. The framework we will describe in Section 4.1 deals with the case where the metric space is just the standard Euclidean space, the set X is an algebraic variety and the point $u \notin X$ is generic. Since $\mathcal{M}_{\leq r}$ is an algebraic variety that framework applies to the rank approximation problem for matrices. Given that a generic matrix A satisfies its hypothesis, a consequence of Theorem 2.19 is that the *Euclidean Distance Degree* of the algebraic variety $\mathcal{M}_{\leq r}$ is $\binom{m}{r}$, assuming $m \leq n$ (see Definition 4.4).

In Section 4.3 we will connect the rank approximation problem for matrices with the nearest point problem to a Schubert variety in the real Grassmannian. In particular, we will use the Eckart–Young Theorem to find some of the critical solutions and, among them, the optimal ones.

2.7 The geometry of Grassmannians

We denote by $G(k, n)$ the real Grassmannian of k–planes in $\mathbb{R}^n$,

$$G(k, n) := \{\mathbf{E} \mid \mathbf{E} \text{ is a } k\text{–dimensional linear subspace of } \mathbb{R}^n\}.$$

For simplicity, from now on we will always assume that $k \leq n - k$.

The aim of this section is to give an overview of the geometry of the real Grassmannian, introduce a natural Riemannian metric on it and study the corresponding induced Riemannian distance. There at least two equivalent ways to do this: (1) viewing $G(k, n)$ as a Riemannian homogeneous space, that is as the quotient of a Lie group; (2) restricting the Fubini–Study metric to the image of the Plücker embedding. Here we will only develop the first approach; we refer the reader to [59] and [65, Section 4.8] for more details on comparing these two approaches.

2.7.1 Grassmannians as homogeneous spaces

Given orthonormal vectors $v_1, \dots, v_k \in \mathbb{R}^n$, we denote by $\langle v_1, \dots, v_k \rangle$ the k–plane spanned by $v_1, \dots, v_k$. The orthogonal group $O(n)$ acts transitively on $G(k, n)$ by rotations; explicitly, if $R \in O(n)$ and $\mathbf{E} = \langle v_1, \dots, v_k \rangle \in G(k, n)$, the orthogonal matrix R acts as

$$R \cdot \mathbf{E} = \langle Rv_1, \dots, Rv_k \rangle.$$

Given $\mathbf{E}_0 = \langle e_1, \dots, e_k \rangle$, where $\{e_1, \dots, e_n\}$ is the standard basis of $\mathbb{R}^n$, it is an easy computation to see that the isotropy subgroup of $\mathbf{E}_0$, i.e. the set of orthogonal matrices that fixes $\mathbf{E}_0$ with respect to above action, is

$$\left\{ \left(\begin{array}{c|c} A_1 & 0 \\ \hline 0 & A_2 \end{array} \right) \;\middle|\; A_1 \in O(k),\ A_2 \in O(n-k) \right\} \cong O(k) \times O(n-k).$$

By [64, Theorem 21.20], the Grassmannian $G(k,n)$ has a unique smooth manifold structure making the action of $O(n)$ smooth. Its dimension is given by

$$\dim(G(k,n)) = \dim(O(n)) - \dim(O(k)\times O(n-k)) = k(n-k).$$

The action induces a quotient map

$$\begin{aligned} \pi : O(n) &\longrightarrow G(k,n) \\ R &\longmapsto R\cdot \mathbf{E}_0, \end{aligned} \tag{2.13}$$

which projects an orthogonal matrix R to the span of its first k columns. This map is a smooth submersion with fibers isomorphic to $O(k)\times O(n-k)$. It follows that $G(k,n)$ has the structure of a homogeneous space, given by the quotient

$$G(k,n) \cong \frac{O(n)}{O(k)\times O(n-k)}.$$

2.7.2 Isomorphisms of tangent spaces

In this section we describe some useful isomorphisms between the tangent space $T_{\mathbf{E}}G(k,n)$, the space of matrices $\mathbb{R}^{(n-k)\times k}$ and $\mathrm{Hom}(\mathbf{E},\mathbf{E}^\perp)$. These isomorphisms give three different ways of looking at tangent vectors. In the following we will use all these different representations, depending on the context.

To start with, we endow $O(n)$ with the restriction of the rescaled Frobenius metric. We consider on the tangent spaces to $O(n)$ the rescaled version of the Frobenius scalar product introduced in (2.9), i.e. for $A,B\in T_RO(n)$ we have

$$\langle A,B\rangle_{\widetilde{F}} := \tfrac{1}{2}\mathrm{tr}(A^\top B). \tag{2.14}$$

This Riemannian metric on $O(n)$ is both left and right invariant ([65, Proposition 4.45]); the reason for the "$\frac{1}{2}$"scaling factor will be clear below.

For every $\mathbf{E}\in G(k,n)$, we denote by $\widehat{E}\in\pi^{-1}(\mathbf{E})$ a fixed representative in $O(n)$. As π is a submersion, the differential

$$D_{\widehat{E}}\pi : T_{\widehat{E}}O(n) \longrightarrow T_{\mathbf{E}}G(k,n)$$

is surjective at every point. In particular it defines a linear isomorphism between the orthogonal complement of its kernel and the image $T_{\mathbf{E}}G(k,n)$. Recall that the tangent spaces to $O(n)$ are just left translations of the Lie algebra $\mathfrak{so}(n)$, that is to say $T_{\widehat{E}}O(n) = \widehat{E}\cdot\mathfrak{so}(n)$. Using this observation and the definition of the metric, it is a straightforward computation to see that

$$\begin{aligned} \mathrm{Ker}(D_{\widehat{E}}\pi) &= \left\{ \widehat{E}\left(\begin{array}{c|c} M_1 & 0 \\ \hline 0 & M_2 \end{array}\right) \;\middle|\; M_1\in\mathfrak{so}(k),\ M_2\in\mathfrak{so}(n-k)\right\} \cong T_{\widehat{E}}\pi^{-1}(\mathbf{E}), \\ (\mathrm{Ker}(D_{\widehat{E}}\pi))^\perp &= \left\{ \widehat{E}\left(\begin{array}{c|c} 0 & -A^\top \\ \hline A & 0 \end{array}\right) \;\middle|\; A\in\mathbb{R}^{(n-k)\times k}\right\} \simeq T_{\mathbf{E}}G(k,n). \end{aligned} \tag{2.15}$$

For every matrix $A\in\mathbb{R}^{(n-k)\times k}$ we introduce the shorthand notation

$$M_A := \left(\begin{array}{c|c} 0 & -A^\top \\ \hline A & 0 \end{array}\right). \tag{2.16}$$

Then the following map

$$\begin{aligned} \mathbb{R}^{(n-k)\times k} &\longrightarrow (\mathrm{Ker}(D_{\widehat{E}}\pi))^{\perp} \\ A &\longmapsto \widehat{E}M_A \end{aligned}$$

is a linear isomorphism. It is then straightforward to check that this gives an isometry of Euclidean spaces

$$\left(\mathbb{R}^{(n-k)\times k}, \langle\cdot,\cdot\rangle_{\mathrm{F}}\right) \simeq \left((\mathrm{Ker}(D_{\widehat{E}}\pi))^{\perp}, \langle\cdot,\cdot\rangle_{\widetilde{\mathrm{F}}}\right). \tag{2.17}$$

Remark that this wouldn't be true without the rescaling factor in (2.14). We use above isomorphism to define the following one, which allows us to identify tangent vectors to the Grassmannian with matrices.

Definition 2.21 (*The isomorphism $A \mapsto v_A$*)**.** Let $\mathbf{E} \in G(k,n)$ and fix a representative $\widehat{E} \in \pi^{-1}(\mathbf{E})$. Then, we define the linear isomorphism

$$\begin{aligned} \mathbb{R}^{(n-k)\times k} &\xrightarrow{\simeq} T_{\mathbf{E}}G(k,n) \\ A &\longmapsto v_A := D_{\widehat{E}}\pi(\widehat{E}M_A). \end{aligned} \tag{2.18}$$

We stress that this isomorphism is not canonical, as it depends on the fixed representative $\widehat{E}$.

We can also establish another correspondence between $T_{\mathbf{E}}G(k,n)$ and the set of linear maps $\mathrm{Hom}(\mathbf{E}, \mathbf{E}^{\perp})$ as follows.

Definition 2.22 (*The isomoprhism $A \mapsto \varphi_A$*)**.** Let $\mathbf{E} \in G(k,n)$ and fix a representative $\widehat{E} \in \pi^{-1}(\mathbf{E})$. Then, we define the linear isomorphism

$$\begin{aligned} \mathbb{R}^{(n-k)\times k} &\xrightarrow{\simeq} \mathrm{Hom}(\mathbf{E}, \mathbf{E}^{\perp}) \\ A &\longmapsto \varphi_A \end{aligned}$$

defined by identifying a matrix A with the linear map $\varphi_A : \mathbf{E} \longrightarrow \mathbf{E}^{\perp}$ represented by A with respect to basis of $\mathbf{E}$ corresponding to the first k columns of $\widehat{E}$ and the basis of $\mathbf{E}^{\perp}$ corresponding to the last $n-k$ ones.

Summarizing, once we fix a representative $\widehat{E} \in \pi^{-1}(\mathbf{E})$, for any $A \in \mathbb{R}^{(n-k)\times k}$ we have isomorphic representations of tangent vectors to $G(k,n)$ at $\mathbf{E}$:

$$\mathrm{Hom}(\mathbf{E}, \mathbf{E}^{\perp}) \simeq \mathbb{R}^{(n-k)\times k} \simeq T_{\mathbf{E}}G(k,n), \quad \varphi_A \longleftrightarrow A \longleftrightarrow v_A.$$

2.7.3 The Grassmann metric and its geodesics

After having described the tangent spaces to a Grassmannian in Section 2.7.2, we now introduce an orthogonally invariant Riemannian metric on $G(k,n)$. In fact, this is the only $O(n)$–invariant metric up to multiples.

Theorem 2.23. *For every $k, n \in \mathbb{N}$ there exists a unique (up to multiples) Riemannian metric on $G(k,n)$ that is invariant under the action of the orthogonal group $O(n)$.*

Proof. See [65, Corollary 5.34] and [65, Remark 5.35]. □

Remark 2.24. If we only require the Riemannian metric on $G(k,n)$ to be invariant under the action of $SO(n)$, then this is still unique (up to multiples), except for the case $(k,n) = (2,4)$. In this case there is a one dimensional family of non–equivalent invariant metrics, see [65, Example 5.20 and Corollary 5.34].

Remark 2.25. There are other distances on the Grassmannian that are invariant for the natural orthogonal action, but they are not induced by a Riemannian metric. For a comprehensive overview of these distances, see [96, Section 2]. The distances listed there are all functions of the principal angles (see Section 2.7.4). This is not by accident, as every distance invariant under rotations in $O(n)$ has to be a function of the principal angles (see [96, Theorem 2] and [95, Theorem 3]).

The desired Riemannian metric is defined as follows.

Definition 2.26 (*The Grassmann metric*)**.** We define the Riemannian metric on $G(k,n)$ as the unique one such that $\pi : O(n) \to G(k,n)$ becomes a Riemannian submersion, i.e. such that for every $\mathbf{E} \in G(k,n)$ the map

$$D_{\widehat{E}}\pi : (\mathrm{Ker}(D_{\widehat{E}}\pi))^{\perp} \longrightarrow T_{\mathbf{E}}G(k,n)$$

is a linear isometry (this definition is well posed, see [65, Proposition 4.50]). We will refer to this metric as the *Grassmann metric* and denote it as $g_{G(k,n)}$.

More explicitly, choose a representative $\widehat{E} \in \pi^{-1}(\mathbf{E})$, and denote by $A_1, A_2 \in \mathbb{R}^{(n-k)\times k}$ matrices representing the vectors $v_1, v_2 \in T_{\mathbf{E}}G(k,n)$ under the isomorphism from Definition 2.21. Then, using (2.17), one easily sees that the Riemannian structure form the previous definition is the scalar product defined on $T_{\mathbf{E}}G(k,n)$ by

$$g_{G(k,n)}(v_1, v_2) := \langle A_1, A_2 \rangle_{\mathrm{F}}.$$

In other words, the space $T_{\mathbf{E}}G(k,n)$ with the Grassmann metric is isometiric to $\mathbb{R}^{(n-k)\times k}$ with the Frobenius scalar product under the previous identification.

We will now describe the Riemannian geodesics on $G(k,n)$ of the Grassmann metric. Observe that the projection map $\pi : O(n) \longrightarrow G(k,n)$ does not preserve geodesics in general, meaning that geodesics on $O(n)$ are not mapped to geodesics on $G(k,n)$. Nevertheless, if a geodesic $\gamma(t)$ on $O(n)$ is horizontal, i.e. for every t for which it is defined $\dot{\gamma}(t) \in (\mathrm{Ker}(D_{\gamma(t)}\pi))^{\perp}$, then the projected curve $\pi(\gamma(t))$ is a geodesic in $G(k,n)$. Viceversa, every geodesic in $G(k,n)$ admits a horizontal lift. Therefore we can exploit the knowledge of the geodesics on $O(n)$ with respect to the Frobenius metric to compute the geodesics and the exponential map on $G(k,n)$ with respect to the Grassmann metric.

To this end recall that the Riemannian exponential map of $O(n)$ at the identity coincides with the usual matrix exponential ([32, Exercise 3, Chapter 3]), which is defined for any square matrix X as

$$e^X := \sum_{i=0}^{+\infty} \frac{X^i}{i!}.$$

Given $\mathbf{E} \in G(k,n)$, the Riemannian exponential map of $G(k,n)$ at $\mathbf{E}$ is by definition

$$\begin{aligned} \exp_{\mathbf{E}} : T_{\mathbf{E}}G(k,n) &\longrightarrow G(k,n) \\ v &\longmapsto \gamma_{(\mathbf{E},v)}(1), \end{aligned}$$

where $\gamma_{(\mathbf{E},v)}$ is the geodesic starting at $\mathbf{E}$ with velocity v. We want to give an explicit description of this map. Remark that since $G(k,n)$ is a compact manifold, the domain of $\exp_{\mathbf{E}}$ is the whole tangent space for every $\mathbf{E} \in G(k,n)$.

For any $R \in O(n)$ we denote by $\exp_R^{O(n)}$ the Riemannian exponential map of $O(n)$ at the point R. Let us fix a representative $\widehat{E} \in \pi^{-1}(\mathbf{E})$ and let us identify the tangent space $T_{\mathbf{E}}G(k,n)$ with $\mathbb{R}^{(n-k)\times k}$ as in Definition 2.21. Then we have

$$\exp_{\mathbf{E}}(v_A) = \exp_{\mathbf{E}}\left(D_{\widehat{E}}\pi(\widehat{E}M_A)\right) = \pi\left(\exp_{\widehat{E}}^{O(n)}(\widehat{E}M_A)\right) = \pi\left(\widehat{E}\exp_{\mathbb{1}}^{O(n)}(M_A)\right) = \widehat{E}e^{M_A}\cdot\mathbf{E}_0,$$

where $\mathbb{1}$ is the identity matrix. Let now $A = U\Sigma V^\top$ be a SVD of A with $\Sigma = \mathrm{diag}(\mu_1,\dots,\mu_k) \in \mathbb{R}^{(n-k)\times k}$. Then we can decompose M_A as the product

$$M_A = \left(\begin{array}{c|c} V & 0 \\ \hline 0 & U \end{array}\right)\left(\begin{array}{c|c} 0 & -\Sigma^\top \\ \hline \Sigma & 0 \end{array}\right)\left(\begin{array}{c|c} V & 0 \\ \hline 0 & U \end{array}\right)^\top =: R_A M_\Sigma {R_A}^\top. \tag{2.19}$$

Remark that if $R \in O(n)$, then for every $B \in \mathbb{R}^{n\times n}$ we have $e^{RBR^\top} = Re^BR^\top$. From (2.19) it follows

$$e^{M_A} = R_A e^{M_\Sigma} {R_A}^\top. \tag{2.20}$$

By a straightforward computation, using the Taylor expansions of the sine and cosine functions, we see that the exponential of M_Σ is given by

$$e^{M_\Sigma} = \left(\begin{array}{ccc|ccc|c} \cos(\mu_1) & & & -\sin(\mu_1) & & & \\ & \ddots & & & \ddots & & 0 \\ & & \cos(\mu_k) & & & -\sin(\mu_k) & \\ \hline \sin(\mu_1) & & & \cos(\mu_1) & & & \\ & \ddots & & & \ddots & & 0 \\ & & \sin(\mu_k) & & & \cos(\mu_k) & \\ \hline & 0 & & & 0 & & 0 \end{array}\right). \tag{2.21}$$

From our computation above we therefore obtain a full description of the geodesics and the Riemannian exponential map on $G(k,n)$ for the Grassmann metric. Before stating the result, for practical reasons we introduce a different notation for elements of $G(k,n)$: given a full–rank matrix $M \in \mathbb{R}^{n\times k}$, we denote by $[M]$ the k–dimensional subspace spanned by the columns of M. We will also use the shorthand notation γ_A for the geodesic $\gamma_{(\mathbf{E},v_A)}$, whenever the association between the matrix $A \in \mathbb{R}^{(n-k)\times k}$ and the tangent vector $v_A \in T_{\mathbf{E}}G(k,n)$ is clear from the context.

Proposition 2.27. *Let $\mathbf{E} \in G(k,n)$ and choose a representative $\widehat{E} \in \pi^{-1}(\mathbf{E})$. Let also $v_A \in T_{\mathbf{E}}G(k,n) \simeq \mathbb{R}^{(n-k)\times k}$ be the tangent vector associated to a matrix $A \in \mathbb{R}^{(n-k)\times k}$ given by the isomorphism Definition 2.21. Let $A = U\Sigma V^\top$ be a SVD of A, with $U = (u_1 \dots u_{n-k}) \in O(n-k)$, $V = (v_1 \dots v_k) \in O(k)$ and $\Sigma = \mathrm{diag}(\mu_1,\dots,\mu_k) \in \mathbb{R}^{(n-k)\times k}$. Denote by $\gamma_A : (-\infty,+\infty) \longrightarrow G(k,n)$ the geodesic such that $\gamma_A(0) = \mathbf{E}$ and $\dot\gamma_A(0) = v_A$. Then*

$$\gamma_A(t) = \left[\widehat{E}\begin{pmatrix} \cos(t\mu_1)v_1 & \dots & \cos(t\mu_k)v_k \\ \sin(t\mu_1)u_1 & \dots & \sin(t\mu_k)u_k \end{pmatrix}\right], \tag{2.22}$$

In particular,

$$\exp_{\mathbf{E}}(v_A) = \left[\widehat{E}\begin{pmatrix} \cos(\mu_1)v_1 & \dots & \cos(\mu_k)v_k \\ \sin(\mu_1)u_1 & \dots & \sin(\mu_k)u_k \end{pmatrix}\right].$$

Proof. As we already observed, geodesics in $G(k,n)$ coincide with projections of horizontal geodesics under $\pi : O(n) \to G(k,n)$, i.e. of curves of the form given by (2.20) and (2.21). □

2.7.4 Principal angles, length minimizing geodesics and the Grassmann distance

We begin this section by recalling the notion of *principal vectors* and *principal angles*, following the algorithmic definition from [96]. This will allow us to give a different description of the geodesics of the Grassmann metric, find the lenght minimizing one and describe explicitly the Grassmann distance in terms of principal angles.

Definition 2.28 (*Principal vectors and principal angles*)**.** Let $\mathbf{E}_1$, $\mathbf{E}_2 \in G(k,n)$. The principal vectors $(p_1,q_1),\dots,(p_k,q_k) \in \mathbf{E}_1 \times \mathbf{E}_2$ and the principal angles $0 \leq \theta_1(\mathbf{E}_1,\mathbf{E}_2) \leq \cdots \leq \theta_k(\mathbf{E}_1,\mathbf{E}_2) \leq \frac{\pi}{2}$ between $\mathbf{E_1}$ and $\mathbf{E}_2$ are defined by the following recursive algorithm. The couple (p_1,q_1) is defined as a solution to the optimization problem:

$$\begin{aligned} \text{maximize} \quad & \langle p,q\rangle \\ \text{subject to} \quad & p \in \mathbf{E}_1,\ \|p\| = 1 \\ & q \in \mathbf{E}_2,\ \|q\| = 1, \end{aligned}$$

where $\langle\cdot,\cdot\rangle$ is the Euclidean scalar product of $\mathbb{R}^n$. Inductively, for $i = 2,\dots,k$, the couple (p_i,q_i) is defined as a solution to the optimization problem

$$\begin{aligned} \text{maximize} \quad & \langle p,q\rangle \\ \text{subject to} \quad & p \in \mathbf{E}_1,\ p \perp \{p_1,\dots,p_{i-1}\},\ \|p\| = 1 \\ & q \in \mathbf{E}_2,\ q \perp \{q_1,\dots,q_{i-1}\},\ \|q\| = 1, \end{aligned}$$

The obtained couples $(p_1,q_1),\dots,(p_k,q_k)$ are called *principal vectors* between $\mathbf{E}_1$ and $\mathbf{E}_2$. Recall that maximizing the scalar product between unit vectors is the same as minimizing the angle between them. It follows that principal vectors provide orthonormal basis for $\mathbf{E}_1$ and $\mathbf{E}_2$ such that the angles between corresponding vectors are the least possible. These angles are denoted by

$$\theta_i(\mathbf{E}_1,\mathbf{E}_2) := \arccos(\langle p_i,q_i\rangle), \quad i = 1,\dots,k,$$

and they are called the *principal angles* between $\mathbf{E}_1$ and $\mathbf{E}_2$.

It is clear that the principal angles satisfy $0 \leq \theta_1(\mathbf{E}_1,\mathbf{E}_2) \leq \cdots \leq \theta_k(\mathbf{E}_1,\mathbf{E}_2) \leq \frac{\pi}{2}$. Moreover we can easily observe that the following properties hold:

- if there are exactly ℓ principal angles that are 0, i.e. $\theta_1(\mathbf{E}_1,\mathbf{E}_2) = \cdots = \theta_\ell(\mathbf{E}_1,\mathbf{E}_2) = 0$ but $\theta_{\ell+1}(\mathbf{E}_1,\mathbf{E}_2) > 0$, then $\dim(\mathbf{E}_1 \cap \mathbf{E}_2) = \ell$;
- if there are exactly ℓ principal angles that are $\frac{\pi}{2}$, i.e. $\theta_{k-\ell+1}(\mathbf{E}_1,\mathbf{E}_2) = \cdots = \theta_k(\mathbf{E}_1,\mathbf{E}_2) = \frac{\pi}{2}$ but $\theta_{k-\ell}(\mathbf{E}_1,\mathbf{E}_2) < \frac{\pi}{2}$, then $\dim(\mathbf{E}_1 \cap \mathbf{E}_2^{\perp}) = \ell$;
- the planes $\mathfrak{P}_i = \operatorname{span}\{p_i,q_i\}$ are pairwise orthogonal for $i = 1,\dots,k$; we call them the *principal planes* between $\mathbf{E}_1$ and $\mathbf{E}_2$. Remark that in case of a 0 i–th angle, then $p_i = q_i$ and the i–th principal plane $\mathfrak{P}_i$ is just a line; nevertheless we will still use the term "plane" to indicate it.

To actually compute principal vectors and principal angles between two k–planes, we use once again the singular values decomposition. Let A, $B \in \mathbb{R}^{(n-k)\times k}$ be matrices with orthonormal columns such that $\mathbf{E}_1 = [A]$ and $\mathbf{E}_2 = [B]$. Consider a SVD of the product $A^\top B \in \mathbb{R}^{k\times k}$, i.e.

$$A^\top B = U\Sigma V^\top \tag{2.23}$$

with $U,V \in O(k)$ and $\Sigma = \operatorname{diag}(\sigma_1,\dots,\sigma_k) \in \mathbb{R}^{k\times k}$ with singular values ordered as $\sigma_1 \geq \cdots \geq \sigma_k \geq 0$. Then one can show that [13]

- $AU = (p_1 \dots p_k) \in \mathbb{R}^{n\times k}$ contains as columns the principal vectors for $\mathbf{E}_1$;
- $BV = (q_1 \dots q_k) \in \mathbb{R}^{n\times k}$ contains as columns the principal vectors for $\mathbf{E}_2$;
- $\sigma_i = \cos(\theta_i(\mathbf{E}_1, \mathbf{E}_2))$ for $i = 1, \dots, k$, where $0 \leq \theta_1(\mathbf{E}_1, \mathbf{E}_2) \leq \dots \leq \theta_k(\mathbf{E}_1, \mathbf{E}_2) \leq \frac{\pi}{2}$ are the principal angles between $\mathbf{E}_1$ and $\mathbf{E}_2$.

Therefore given two k-planes $\mathbf{E}_1$ and $\mathbf{E}_2 \in G(k, n)$, to compute the principal vectors and angles between them it is enough to pick two specific orthonormal basis, stack them into matrices representing the planes, take the product between them as in (2.23) and run an SVD to find the changes of basis needed to get principal vectors and to find the cosines of the principal angles.

Now that we have introduced principal vectors and principal angles, it becomes easy to describe the length minimizing geodesic connecting two points in $G(k, n)$ and hence the function giving the distance between them. To start with, we give an answer to the following question: given $\mathbf{E}_1$, $\mathbf{E}_2 \in G(k, n)$, how can we compute a tangent vector $v_A \in T_{\mathbf{E}_1} G(k, n)$ such that $\exp_{\mathbf{E}_1}(v_A) = \mathbf{E}_2$?

This can be done via the following procedure. Given the two k–planes $\mathbf{E}_1 = [(a_1 \cdots a_k)]$ and $\mathbf{E}_2 = [(b_1 \cdots b_k)]$, fix a representative $\widehat{E}_1 = (a_1, \dots, a_k, w_1, \dots, w_{n-k}) \in \pi^{-1}(\mathbf{E}_1) \subseteq O(n)$ for $\mathbf{E}_1$, and take a decomposition as in (2.23). Define the $(n-k) \times k$ matrix $\Theta := \operatorname{diag}(\theta_1(\mathbf{E}_1, \mathbf{E}_2), \dots, \theta_k(\mathbf{E}_1, \mathbf{E}_2)) \in \mathbb{R}^{(n-k)\times k}$ containing the principal angles between $\mathbf{E}_1$ and $\mathbf{E}_2$ on the diagonal. Recall that the matrix U from (2.23) gives the change of basis that one needs to apply to the orthonormal basis $\{a_1, \dots, a_k\}$ of $\mathbf{E}_1$ to obtain the principal vectors $\{p_1, \dots, p_k\}$. Assume that $\theta_1(\mathbf{E}_1, \mathbf{E}_2) = \dots = \theta_\ell(\mathbf{E}_1, \mathbf{E}_2) = 0$, so that $p_1 = q_1, \dots, p_\ell = q_\ell$ and $\mathbf{E}_1 \cap \mathbf{E}_2$ has orthonormal basis $\{p_1, \dots, p_\ell\}$. For every $i = \ell + 1, \dots, k$ define $n_i \in \mathfrak{P}_i = \operatorname{span}\{p_i, q_i\}$ such that $\|n_i\| = 1$, $n_i \perp p_i$ and $\langle n_i, q_i \rangle > 0$. Then by definition of principal angles we can see that

$$q_i = p_i \cos(\theta_i(\mathbf{E}_1, \mathbf{E}_2)) + n_i \sin(\theta_i(\mathbf{E}_1, \mathbf{E}_2)). \tag{2.24}$$

Moreover, by the orthogonality relations between principal vectors, we also have $n_i \in \mathbf{E}_1^\perp$ for every $i = \ell + 1, \dots, k$. Complete the set $\{n_{\ell+1}, \dots, n_k\}$ to an orthonormal basis $\{n_1, \dots, n_{n-k}\}$ of $\mathbf{E}_1^\perp$ and define the orthogonal matrix $S \in O(n-k)$ giving the change of basis on $\mathbf{E}_1^\perp$ from $\{w_1, \dots, w_{n-k}\}$, that was a basis of $\mathbf{E}_1^\perp$ by construction, to $\{n_1, \dots, n_{n-k}\}$. We define

$$A := S\Theta U^\top.$$

Using (2.22) and (2.24), we can see that

$$\exp_{\mathbf{E}_1}(v_A) = \mathbf{E}_2. \tag{2.25}$$

Geometrically, the geodesic $\gamma_A(t) := \exp_{\mathbf{E}_1}(tv_A)$ just consists of simultaneous rotations from the principal vectors $p_i \in \mathbf{E}_1$ to the corresponding principal vectors $q_i \in \mathbf{E}_2$, performed inside the principal planes $\mathfrak{P}_i$ for $i = \ell + 1, \dots, k$ (nothing is done for $p_1, \dots, p_\ell$ which are in $\mathbf{E}_1 \cap \mathbf{E}_2$) with speed given by the angles $\theta_i(\mathbf{E}_1, \mathbf{E}_2)$. Remark that, if in the definition of Θ we replace an angle $\theta_i(\mathbf{E}_1, \mathbf{E}_2)$ with $\theta_i(\mathbf{E}_1, \mathbf{E}_2) + m\pi$ for any $m \in \mathbb{Z}$, equation (2.25) still holds true, but the corresponding geodesic will have a greater length for any $m \neq 0$.

One can show that the geodesic $\gamma_A(t)$ constructed via above procedure provides the length minimizing geodesic between $\mathbf{E}_1$ and $\mathbf{E}_2$. It follows that the Riemannian distance between $\mathbf{E}_1$ and $\mathbf{E}_2$ is given by the norm of the tangent vector v_A, which is the same as the Frobenius norm of A; this can be expressed in terms of the principal angles between them.

Definition 2.29 (*Grassmann distance*)**.** We denote by $\delta : G(k,n) \times G(k,n) \to \mathbb{R}$ the Riemannian distance function induced by the Grassmann metric (Definition 2.26) and we call this function the *Grassmann distance*. More explicitly, if $\theta_1(\mathbf{E}_1, \mathbf{E}_2), \dots, \theta_k(\mathbf{E}_1, \mathbf{E}_2)$ are the principal angles between the two k–planes $\mathbf{E}_1$ and $\mathbf{E}_2$, then

$$\delta(\mathbf{E}_1, \mathbf{E}_2) = \left(\sum_{i=1}^{k} \theta_i^2(\mathbf{E}_1, \mathbf{E}_2) \right)^{\frac{1}{2}}. \tag{2.26}$$

We introduce the notation

$$R_{\frac{\pi}{2}} = \left\{ A \in \mathbb{R}^{(n-k)\times k} \mid \text{the singular values of } A \text{ are bounded by } \frac{\pi}{2} \right\} \subseteq \mathbb{R}^{(n-k)\times k} \tag{2.27}$$

for the set of $(n-k) \times k$ matrices with bounded singular values. Identifying $R_{\frac{\pi}{2}}$ with a subset of $T_{\mathbf{E}}G(k,n)$ via the isomorphism in Definition 2.21, we can summarize the above discussion in the following result.

Proposition 2.30. *The image of $R_{\frac{\pi}{2}}$ through the exponential map $\exp_{\mathbf{E}} : T_{\mathbf{E}}G(k,n) \longrightarrow G(k,n)$ is the whole Grassmannian. Moreover, for every $A \in R_{\frac{\pi}{2}}$, the principal angles between $\mathbf{E}$ and $\exp_{\mathbf{E}}(v_A)$ coincide with the singular values of A.*

Remark 2.31. The content of the previous proposition can be visualized using the following commutative diagram that, for a fixed $\mathbf{E}$, relates the Frobenius distance on the tangent space and the Grassmann distance:

$$\begin{array}{ccc} R_{\frac{\pi}{2}} & \xrightarrow{\exp_{\mathbf{E}}} & G(k,n) \\ & {\scriptstyle \|\cdot\|_{\mathrm{F}}} \searrow \quad \swarrow {\scriptstyle \delta(\mathbf{E},\cdot)} & \\ & \mathbb{R} & \end{array} \quad , \qquad \delta(\mathbf{E}, \exp_{\mathbf{E}}(v_A)) = g_{G(k,n)}(v_A, v_A)^{\frac{1}{2}} = \|A\|_{\mathrm{F}}.$$

2.7.5 O–minimality of principal angles

The real Grassmannian $G(k,n)$ can be embedded in $\mathbb{R}\mathrm{P}^L = \mathbb{P}(\Lambda^k \mathbb{R}^N)$, where $L = \binom{n}{k} - 1$, via the well-known Plücker embedding. If a k–plane $\mathbf{E} \in G(k,n)$ has a basis $\{v_1, \dots, v_k\}$, its image via the Plücker embedding is the projective class of the wedge $[v_1 \wedge \cdots \wedge v_k]$. The image of this embedding is a real algebraic set (in particular semialgebraic), corresponding to the set of simple k–vectors. Notice that the principal angles *are not* semialgebraic functions, since they involve taking arccosines of the singular values (we will show in Section 4.2.6 that the whole nearest point problem to an algebraic set of $G(k,n)$ for the Grassmann metric is, in general, not semialgebraic). However, the aim of this section is to show that the principal angles between two k–planes are definable in the o–minimal structure of globally subanalytic sets. We will not discuss o–minimal structures or globally subanalytics sets; we refer the reader to [93] for the basics of o–minimality theory.

Proposition 2.32 (*o–minimality of the Grassmann distance*)**.** *The principal angles are definable in the o–minimal structure of globally subanalaytic sets and so is the Grassmann distance function.*

Proof. By definition of globally subanalytic function, it is enough to prove that the graph of the function $\theta_i(\mathbf{E}, \cdot)$, denoted as $\mathrm{graph}(\theta_i(\mathbf{E}, \cdot))$, is globally subanalytic for every $\mathbf{E} \in G(k,n)$, every $i = 1, \dots, k$ and every $1 \leq k \leq n$. By the orthogonal invariance of principal angles, we can fix a specific k–plane $\mathbf{E}_0 \in G(k,n)$ and assume without loss of generality that $\mathbf{E} = \mathbf{E}_0 = \langle e_1, \dots, e_k \rangle$ as defined before.

Given a set of k vectors $a_1, \ldots, a_k \in \mathbb{R}^n$, denote by A the matrix having those vectors as columns

$$A = (a_1, \ldots, a_k) = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} \in \mathbb{R}^{n \times k},$$

where A_1 is a $k \times k$ matrix and A_2 is a $(n-k) \times k$ matrix. Then the graph of $\theta_i(\mathbf{E}_0, \cdot)$ can be described as

$$\begin{aligned} \operatorname{graph}(\theta_i(\mathbf{E}_0, \cdot)) = \{(\mathbf{L}, t) \in G(k,n) \times \mathbb{R} \mid \exists\, a_1, \ldots, a_k \in \mathbb{R}^n,\ \mathbf{L} = [(a_1 \ldots a_k)], \\ (A_1^\top A_1 + A_2^\top A_2) = \mathbb{1},\ t = \arccos((\lambda_i(A_1 A_1^\top))^{\frac{1}{2}})\} \\ = \{(\mathbf{L}, t) \in G(k,n) \times \mathbb{R} \mid \exists\, a_1, \ldots, a_k \in \mathbb{R}^n,\ \mathbf{L} = [(a_1 \ldots a_k)], \\ (A_1^\top A_1 + A_2^\top A_2) = \mathbb{1},\ \cos t = \lambda_i(A_1 A_1^\top)^{\frac{1}{2}})\}, \end{aligned}$$

where $\lambda_i(\cdot)$ denotes the function giving the i–th eigenvalue of a $k \times k$ square matrix, which is a semialgebraic function. Remark that by the orthogonality condition on the columns of A, we only need the restriction of the cosine function to the interval $[0, \frac{\pi}{2}]$. Therefore the functions $\theta_i(\mathbf{E}, \cdot)$ are globally subanalytic.

The fact that the Grassmann distance is globally subanalytic follows now from (2.26) and the fact that sums, products and compositions of definable functions in an o–minimal structure are still definable. □

Chapter 3

Real random tensors

Tensors can be seen as multi–dimensional matrices, in a sense that will be made precise later, and as such they are often used as a compact way to store and represent data depending on many parameters. This is why their study encountered a renovated interest in the last decades; see Section 1.1 and the references therein for more details.

We begin this chapter with a brief account of the basic notions about tensors and recall the key definition of *tensor rank*. We then explore two questions related to tensor rank in the context of *real random tensors*. In Section 3.1 we ask ourselves what happens when we approximate a real symmetric tensor with one of rank one; in particular we focus on how far these two will be from a probabilistic point of view. In Section 3.2 we will then investigate the notion of *typical ranks*, with a focus on order–three tensors.

We recall that the research presented in this chapter is joint work with different people. In particular the results in Section 3.1 are a joint work with A. Lerario and have been published in [25]. The results in Section 3.2 are a joint work with P. Breiding and S. Eggleston and have been published in [16].

Basic notions on tensors

Let $\mathbb{F}$ be a field and V, W be vector spaces over $\mathbb{F}$. Denote by $\mathbb{F}(V \times W)$ the free vector space on $V \times W$, that is the $\mathbb{F}$–vector space with basis given by all the elements in $V \times W$. This is the set of formal finite linear combinations of elements in $V \times W$ with coefficients in the field $\mathbb{F}$

$$\mathbb{F}(V \times W) = \left\{ \sum_{i \in I} a_i (v_i, w_i) \;\middle|\; a_i \in \mathbb{F},\, v_i \in V,\, w_i \in W,\, \forall i \in I,\, |I| < \infty \right\}.$$

Consider the vector subspace of $\mathbb{F}(V \times W)$ generated by the elements of the following types:

- $(v, w) + (v', w) - (v + v', w)$;
- $(v, w) + (v, w') - (v, w + w')$;
- $(\mu v, w) - \mu(v, w)$;
- $(v, \mu w) - \mu(v, w)$,

for $v, v' \in V$, $w, w' \in W$, $\mu \in \mathbb{F}$. Denote such subspace by U.

Definition 3.1. The *tensor product* of V and W, denoted by $V \otimes W$, is the $\mathbb{F}$–vector space obtained as the quotient of $\mathbb{F}$–vector spaces

$$V \otimes W := \frac{\mathbb{F}(V \times W)}{U}.$$

We denote the class of $(v, w) \in V \times W$ by $v \otimes w$. The effect of taking the quotient by U is to introduce the following bilinear relations at the level of tensors

- $v \otimes w + v' \otimes w = (v + v') \otimes w$;
- $v \otimes w + v \otimes w' = v \otimes (w + w')$;
- $(\mu v) \otimes w = \mu(v \otimes w)$;
- $v \otimes (\mu w) = \mu(v \otimes w)$.

Using the bilinearity described above it is easy to prove the following.

Proposition 3.2. *Given $\mathbb{F}$–vector spaces V and W with basis $\{v_1, \dots, v_m\}$ and $\{w_1, \dots, w_n\}$ respectively, then a basis for $V \otimes W$ is given by*

$$\{v_i \otimes w_j \mid i = 1, \dots, m,\ j = 1, \dots, n\}.$$

Therefore $\dim(V \otimes W) = \dim(V)\dim(W)$.

Tensor product is associative: given three $\mathbb{F}$–vector spaces V, W, Z, we have

$$(V \otimes W) \otimes Z \cong V \otimes (W \otimes Z).$$

It follows that the tensor product $V \otimes W \otimes Z$ is well–defined. We can therefore extend Definition 3.1 and create the tensor product of finitely many vectors spaces: given $\mathbb{F}$–vector spaces $V_1, \dots, V_d$ of dimensions $n_1, \dots, n_d$ respectively, we denote the tensor product of $V_1, \dots, V_d$ by

$$V_1 \otimes \cdots \otimes V_d;$$

all the notations extend from the case $d = 2$ to a general $d \in \mathbb{N}$. The result of Proposition 3.2 also extends to finitely many vector spaces. If $\{v_j^{(i)} \mid j = 1, \dots, n_i\}$ is a basis for V_i, then

$$\left\{ v_{j_1}^{(1)} \otimes \cdots \otimes v_{j_d}^{(d)} \mid j_1 \in \{1, \dots, n_1\}, \dots, j_d \in \{1, \dots, n_d\} \right\} \tag{3.1}$$

is a basis of the tensor product $V_1 \otimes \cdots \otimes V_d$, which has dimension $n_1 \dots n_d$. This also shows that there is an isomorphism of $\mathbb{F}$–vector spaces

$$V_1 \otimes \cdots \otimes V_d \cong \mathbb{F}^{n_1} \otimes \cdots \otimes \mathbb{F}^{n_d} \cong \mathbb{F}^{n_1 \times \cdots \times n_d}.$$

Indeed, given a tensor $T \in V_1 \otimes \cdots \otimes V_d$ we can consider its coordinates with respect to the basis in (3.1). In other words, we identify a tensor T with its coordinates $(t_{i_1,\dots,i_d}) \in \mathbb{F}^{n_1 \times \cdots \times n_d}$ where

$$T = \sum_{i_1,\dots,i_d} t_{i_1,\dots,i_d}\, v_{i_1}^{(1)} \otimes \cdots \otimes v_{i_d}^{(d)}.$$

In this sense, a tensor is a multi–dimensional array, given the choice of a basis of $V_1, \dots, V_d$. In what follows, we will often use this isomorphism and consider tensors as elements of $\mathbb{F}^{n_1 \times \cdots \times n_d}$. We say that a tensor $T \in \mathbb{F}^{n_1 \times \cdots \times n_d}$ has *order* d (i.e. the number of factors)

and *format* $n_1 \times \cdots \times n_d$ (i.e. the dimensions of the factors).

Given an $\mathbb{F}$–vector space V, we can consider the tensor product of d copies of V

$$V^{\otimes d} := \underbrace{V \otimes \cdots \otimes V}_{d \text{ times}}.$$

We consider the vector subspace $\widetilde{U}$ of $V^{\otimes d}$ generated by the elements

$$v_1 \otimes \cdots \otimes v_d - v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(d)}$$

for $v_1, \ldots, v_d \in V$ and $\sigma \in S_d$, where S_d denotes the group of permutations on $\{1, \ldots, d\}$.

Definition 3.3. Given an $\mathbb{F}$–vector space V, its *d–th symmetric power* is defined as the quotient of $\mathbb{F}$–vector spaces

$$\mathcal{S}^d(V) := \frac{V^{\otimes d}}{\widetilde{U}}.$$

The elements of $\mathcal{S}^d(V)$ are called *symmetric tensors* of order d on V.

A symmetric tensor $v_1 \otimes \cdots \otimes v_d \in \mathcal{S}^d(V)$ can also be written as $v_1 \ldots v_d$ and the order of the factors has no importance. Indeed, taking the quotient by $\widetilde{U}$ introduces the symmetry relations

$$v_1 \ldots v_d = v_{\sigma(1)} \ldots v_{\sigma(d)}$$

for every $\sigma \in S_d$. Remark that given a basis $\{v_1, \ldots, v_n\}$ of V, a basis of $\mathcal{S}^d(V)$ is given by the monomials of degree d in $v_1, \ldots, v_n$. This is at the heart of the correspondence between symmetric tensors and homogeneous polynomials that is discussed in details in Section 3.1.1.

In the isomorphism $V^{\otimes d} \cong \mathbb{F}^{n \times \cdots \times n}$ that identifies tensors in $V^{\otimes d}$ with d–dimensional arrays of size $n \times \cdots \times n$, symmetric tensors correspond to arrays $(t_{i_1,\ldots,i_d}) \in \mathbb{F}^{n \times \cdots \times n}$ such that swapping indices give the same entry; from this point of view this is therefore an extension of the definition of symmetric matrix.

Tensor rank

The notion of *rank* of an object in linear and multilinear algebra is ubiquitous and can have several different meanings, depending on the type of objects involved and the property that we are interested in studying. Nevertheless, the idea behind the notion of rank is often the same: we define some objects, that we interpret as elementary ones, to have rank equal to one, and define the rank of an element to be the minimum number of rank–one objects that you need to consider to obtain the given one. We specialize this construction to the case of tensors as follows.

Definition 3.4 (Rank–one tensor)**.** Given $\mathbb{F}$–vector spaces $V_1, \ldots, V_d$ we say that a tensor $T \in V_1 \otimes \cdots \otimes V_d$ is a *rank–one* tensor if there exist $v_1 \in V_1, \ldots, v_d \in V_d$ such that

$$T = v_1 \otimes \cdots \otimes v_d.$$

In other words, a rank–one tensor corresponds to one of the generating elements in the very definition of tensor product of vector spaces. As a consequence, any tensor can be written as a sum of rank–one tensors. This motivates the following definition.

Definition 3.5 (Tensor rank)**.** Given a tensor $T \in V_1 \otimes \cdots \otimes V_d$ we define the *rank* of T as the minimum number $r \in \mathbb{N}_{\geq 0}$ such that there exist r rank–one tensors $T^{(1)}, \ldots, T^{(r)} \in V_1 \otimes \cdots \otimes V_d$ such that

$$T = \sum_{i=1}^{r} T^{(i)}. \tag{3.2}$$

Rank–one tensors are therefore to be considered as the building blocks of the space of tensors and the rank can itself be thought of as a measure of the complexity of a tensor. The varieties of rank–one tensors of a given format are called *Segre varieties*. These are among the most studied algebraic varieties but, since we will not need them in the rest of this chapter, we do not discuss them here. The only Segre variety that we will need is the one of tensors of order 2, i.e. matrices. The relevant properties will be given when needed in Section 3.2, where this variety will have a key role.

The definition of tensor rank given in Definition 3.5 extends also to symmetric tensors; indeed symmetric tensors are just tensors with an additional condition and we can therefore talk about the rank of a symmetric tensor. Another reasonable choice when considering a symmetric tensor would be to ask for the rank–one tensors in (3.2) to be themselves symmetric, i.e. to decompose the tensor by only looking at the space of symmetric tensors. This leads to the notion of *symmetric tensor rank*. Determining the relationship between tensor rank and symmetric tensor rank is an active area of research within the tensor community; we will not investigate it further here.

Remark that for $T \in \mathcal{S}^d(V)$, the symmetry condition forces all the vectors v_i in Definition 3.4 to be the same, i.e. a tensor $T \in \mathcal{S}^d(V)$ has rank–one if and only if there exists $v \in V$ such that

$$T = v^{\otimes d}.$$

The variety of such tensors is called the *Veronese variety*. It is one of the protagonists of Section 3.1. Its geometry will be described in details in Section 3.1.2, where it will also be shown that it can be identified with d–th powers of linear forms.

In Section 3.1 we will refer to the *best rank–one approximation* of a tensor $T \in \mathbb{F}^{n_1 \times \cdots \times n_d}$. In general, a *best rank approximation* of T is another tensor $\widetilde{T} \in \mathbb{F}^{n_1 \times \cdots \times n_d}$ with a prescribed rank, lower than the one of T, such that $\widetilde{T}$ is the closest to T among the tensors of the prescribed rank with respect to some chosen metric on $\mathbb{F}^{n_1 \times \cdots \times n_d}$. We have already seen an example of this in section Section 2.6, when talking about rank approximations of matrices. This kind of problems fall into the category of *nearest point problems* that we will describe at the beginning of Chapter 4.

3.1 Symmetric tensors close to rank–one

Rank–one approximations and tensor decomposition are crucial tools in many real–world applications involving tensors, such as signal processing, image analysis, phylogenetics and quantum computing (see for example Section 3.1.6 and the references in Section 1.1). In some of these applicative settings, it is not only important to approximate the tensor data, but also to be able to estimate how far the approximation is from the original tensor, in a suitable metric. Motivated by this, we address the following question:

"What is the probability for a real symmetric tensor to be "close" to rank–one?"

We will now specifiy what we mean by *probability* and *close*. We consider real symmetric tensors on $\mathbb{R}^n$ of order d, which we recall are the elements of $(\mathbb{R}^n)^{\otimes d}$ that are invariant under the action of the symmetric group acting by permuting the factors (see Definition 3.3); this set will be denoted by $\mathcal{S}^d(n) := \mathcal{S}^d(\mathbb{R}^n)$. We endow the space of tensors $(\mathbb{R}^n)^{\otimes d}$ with the *Frobenius scalar product*, denoted by $\langle \cdot, \cdot \rangle_F$. Given the standard basis $\{e_1, \dots, e_n\}$ of $\mathbb{R}^n$, this is defined by declaring the basis of $(\mathbb{R}^n)^{\otimes d}$ given by $\{e_{i_1} \otimes \cdots \otimes e_{i_d} \,|\, 1 \leq i_1, \dots, i_d \leq n\}$ to be orthonormal; it is a direct generalization of definition (2.9), where square matrices correspond to the case $d = 2$. The explicit formula we obtain is indeed the same as in (2.9): given $T = (T_{i_1,\dots,i_d})$, $U = (U_{i_1,\dots,i_d}) \in (\mathbb{R}^n)^{\otimes d}$, their Frobenius scalar product is given by

$$\langle T, U \rangle_F = \sum_{i_1,\dots,i_d=1}^{n} T_{i_1,\dots,i_d} \, U_{i_1,\dots,i_d}. \tag{3.3}$$

We still denote by $\langle \cdot, \cdot \rangle_F$ the restriction of the Frobenius scalar product to the space of symmetric tensors and we use the notation $\| \cdot \|_F$ for the associated norm. To be consistent with the notations adopted in Section 2.6, we denote the distance induced by the Frobenius product on the space of tensors as

$$\widehat{\delta}(T, U) := \|T - U\|_F = \left(\sum_{i_1,\dots,i_d=1}^{n} (T_{i_1,\dots,i_d} - U_{i_1,\dots,i_d})^2 \right)^{\frac{1}{2}}.$$

We therefore endow the space $\mathcal{S}^d(n)$ with the restriction of $\widehat{\delta}$ to that set and we will always refer to this distance when talking about how *close* symmetric tensors are.

Recall that any non–degenerate scalar product on a finite dimensional vector space V induces a non–degenerate centered Gaussian measure on V by assigning to every Borel subset $A \subseteq V$ the probability

$$\mathbb{P}(A) = \frac{1}{c} \int_A e^{-\frac{\|x\|^2}{2}} dx,$$

where $\| \cdot \|$ is the norm induced by the scalar product and c is a normalization constant. Viceversa, every non–degenerate centered Gaussian measure on V induces a norm and this correspondence is $1 : 1$. Therefore our choice of the Frobenius scalar product on $\mathcal{S}^d(n)$ naturally leads to the definition of the corresponding *Gaussian probability space* obtained by defining for every Borel set $U \subseteq \mathcal{S}^d(n)$

$$\mathbb{P}\Big\{ T \in U \Big\} := \frac{\displaystyle\int_U \exp\left(-\frac{1}{2} \|T\|_F^2 \right) \mathrm{d}\mu}{\displaystyle\int_{\mathcal{S}^d(n)} \exp\left(-\frac{1}{2} \|T\|_F^2 \right) \mathrm{d}\mu}, \tag{3.4}$$

where $\mathrm{d}\mu$ denotes the integration with respect to the Lebesgue measure on $\mathcal{S}^d(n)$. Thus, when we will talk about the *probability* of a certain event in the space of symmetric tensors, we will always refer to the probability distribution defined in (3.4). Remark that as the Frobenius scalar product has a natural invariance under the action of the orthogonal group on the space of tensors, the resulting Gaussian probability distribution in (3.4) inherits the same invariance property.

Now that we have properly defined what we mean by *probability* and by *close*, we can phrase our question more precisely. Recalling that the distance between a point and a

closed set is defined as the minimum distance bewteen the point and an element of the closed set, for a given threshold $\delta > 0$, we want to compute the following probability

$$\mathbb{P}\left\{T \in \mathcal{S}^d(n) \;\middle|\; \widehat{\delta}(T, \mathcal{S}_1^d(n)) \leq \delta\right\},$$

where $\mathcal{S}_1^d(n)$ is the set of rank–one symmetric tensors. Nevertheless, there is still one missing detail in this formulation. In applications and numerical analysis, including numerical algebraic geometry (see for example [20]), one often does not care about the *absolute error*, which in our context would be the distance between a tensor and its rank–one approximation, but is rather interested in the *relative error*, which is the distance *relative* to the size of the object you are trying to approximate. Therefore the correct quantity to look at is

$$\mathbb{P}\left\{T \in \mathcal{S}^d(n) \;\middle|\; \widehat{\delta}(T, \mathcal{S}_1^d(n)) \leq \delta\|T\|_F\right\}, \tag{3.5}$$

where the size of the tensor $\|T\|_F$ is also taken into account and δ now represents the relative error in going from the tensor T to its (best) rank–one approximation. From the computational point of view, this has also the effect of turning the problem into a conic one, which will make our life easier when doing the actual computations.

In Section 3.1.1 and Section 3.1.2 we will discuss some isomorphisms and geometric descriptions of the objects appearing in (3.5), namely $\mathcal{S}^d(n)$ and $\mathcal{S}_1^d(n)$, that will allow us to restate our question in a way that is easier for computations.

3.1.1 Symmetric tensors and homogeneous polynomials

The aim of this section is to describe a well known correspondence between symmetric tensors of order d and homogeneous polynomials of degree d. In addition to that, we will also relate the Frobenius scalar product on tensors with the Bombieri–Weyl one on polynomials, showing that the correspondence is indeed an isometry between these spaces.

We start by describing the Bombieri–Weyl structure for homogeneous polynomials. Let $\mathbb{C}[x_1, \ldots, x_n]_{(d)}$ be the space of complex homogeneous polynomials of degree d in n variables. The group of complex unitary $n \times n$ matrices $U(n)$ acts on this space by change of variables: for every $R \in U(n)$ and $p \in \mathbb{C}[x_1, \ldots, x_n]_{(d)}$ the matrix R acts on p by

$$(R \cdot p)(x) = p(R^{-1}x).$$

We denote this action by $\beta^{\mathbb{C}} : U(n) \longrightarrow \mathrm{GL}(\mathbb{C}[x_1, \ldots, x_n]_{(d)})$. In general, proving that a given representation is irreducible is not an easy task and it can be related to character theory. For representations of unitary groups another way of dealing with irreducibility is by means of the Schur–Weyl duality. We will not investigate these topics further and refer to [42] for a detailed explanation. Using those tools, one can prove that the representation $\beta^{\mathbb{C}}$ is a finite dimensional representation of $U(n)$ ([54]). As a consequence of Schur's lemma and the compactness of $U(n)$, there exists a unique (up to multiples) Hermitian structure on $\mathbb{C}[x_1, \ldots, x_n]_{(d)}$ which is $\beta^{\mathbb{C}}$–invariant, i.e. such that

$$\langle R \cdot p_1, R \cdot p_2 \rangle = \langle p_1, q_1 \rangle,$$

for every $R \in U(n)$ and every $p_1, p_2 \in \mathbb{C}[x_1, \ldots, x_n]_{(d)}$. It is easy to check that this unique complex Hermitian product is defined by

$$\langle p_1, p_2 \rangle_{\mathrm{BW}} := \frac{1}{\pi^{n+1}} \int_{\mathbb{C}^{n+1}} p_1(z)\overline{p_2(z)} e^{-\|z\|^2} \mathrm{d}z,$$

where $\mathrm{d}z := (i/2)^n \mathrm{d}z_0 \mathrm{d}\overline{z_0} \dots \mathrm{d}z_n \mathrm{d}\overline{z_n}$ is the Lebesgue measure. We call this the *Bombieri–Weyl* or *Kostlan* Hermitian structure on $\mathbb{C}[x_1, \dots, x_n]_{(d)}$. One can show that the following set

$$\left\{ \binom{d}{\alpha}^{\frac{1}{2}} x^\alpha \right\}_{\alpha=(\alpha_1,\dots,\alpha_n)\in\mathbb{Z}^n_{\geq 0},\ \alpha_1+\dots+\alpha_n=d} \tag{3.6}$$

gives an orthonormal basis for the Bombieri–Weyl structure, where we used the notations $\binom{d}{\alpha} = \frac{d!}{\alpha_1!\dots\alpha_n!}$ and $x^\alpha = x_1^{\alpha_1} \dots x_n^{\alpha_n}$. Remark that the elements of this basis are just the rescaled monomials of degree d.

Restricting now to real homogeneous polynomials $\mathbb{R}[x_1, \dots, x_n]_{(d)}$, we define an inner product by restricting the real part of the Hermitian Bombieri–Weyl product to real polynomials, and we call this the *Bombieri–Weyl* scalar product. Since the Hermitian structure was invariant under the complex action of $U(n)$, one can easily show that the Bombieri–Weyl scalar product is invariant under the action $\beta : O(n) \longrightarrow \mathrm{GL}(\mathbb{R}[x_1, \dots, x_n]_{(d)})$ of the orthogonal group on the real vector space of real homogeneous polynomials. We observe that the orthonormal basis in (3.6) is also a basis for the space of real homogeneous polynomials, as its element have real coefficients. Therefore (3.6) gives an orthonormal basis for the Bombieri–Weyl scalar product on $\mathbb{R}[x_1, \dots, x_n]_{(d)}$.

Remark 3.6. The action β of the orthogonal group on $\mathbb{R}[x_1, \dots, x_n]_{(d)}$, which is the restriction of the action $\beta^{\mathbb{C}}$ of $U(n)$, is not irreducible anymore: for example, the subspace of harmonic polynomials is a non–trivial invariant subspace for β, as it can be proved by a straightforward computation. It can be shown that the Bombieri–Weyl scalar product is not the unique (up to multiples) β–invariant scalar product on $\mathbb{R}[x_1, \dots, x_n]_{(d)}$, but a whole family of invariant ones can be constructed.

In the same way we turned $S^d(n)$ into a Gaussian probability space by considering the centered Gaussian measure induced by the Frobenius scalar product, we consider on $\mathbb{R}[x_1, \dots, x_n]_{(d)}$ the centered Gaussian measure induced by the Bombieri–Weyl product, which we call the *Bombieri–Weyl* probability distribution. Explicitly, it is the distribution of the random polynomial

$$P(x) = \sum_{\substack{\alpha\in\mathbb{Z}^n_{\geq 0} \\ \alpha_1+\dots+\alpha_n=d}} \xi_\alpha \binom{d}{\alpha}^{\frac{1}{2}} x^\alpha,$$

where ξ_α are i.i.d. standard Gaussians $\mathcal{N}(0, 1)$.

Before describing the general case of the correspondence between symmetric d–tensors and degree d homogeneous polynomials, we will start by investigating the case $d = 2$, which should be more familiar to the reader: this is indeed the well–known correspondence between symmetric matrices and homogeneous quadratic polynomials.

Let $\mathcal{S}^2(n) = \mathrm{Sym}(n, \mathbb{R})$ be the space of symmetric $n \times n$ matrices with real entries and denote by E_{ij} the elementary matrix having all entries 0 except for the ij-th one being 1. An orthonormal basis for the Frobenius scalar product (3.3) is then given by

$$\{E_{ii} \mid i = 1, \dots, n\} \cup \left\{ \frac{E_{ij} + E_{ji}}{2} \mid 1 \leq i < j \leq n \right\}. \tag{3.7}$$

It follows that the Gaussian measure induced by the Frobenius scalar product is explicitly given by the distribution of the random matrix defined by

$$Q = \sum_{i=1}^{n} \xi_{ii} E_{ii} + \sum_{i<j=1}^{n} \xi_{ij} \frac{E_{ij} + E_{ji}}{\sqrt{2}}, \tag{3.8}$$

where $\xi_{ij} \sim \mathcal{N}(0,1)$ are i.i.d. standard Gaussian variables for every $1 \leq i \leq j \leq n$. Such a random matrix Q has random Gaussian entries distributed as $\mathcal{N}(0,1)$ on the diagonal and $\mathcal{N}(0, \frac{1}{2})$ off-diagonal, independent except for the obvious symmetry condition. The set $\mathrm{Sym}(n, \mathbb{R})$ together with the distribution we just described it called the *Gaussian Orthogonal Ensemble* and it is denoted by $\mathrm{GOE}(n)$. In this case (3.4) has the simpler expression

$$\mathbb{P}\{Q \in U\} = \frac{1}{(2\pi)^{\frac{n(n+1)}{4}}} \int_U e^{-\frac{1}{2}\mathrm{tr}(A^2)} dA, \tag{3.9}$$

Recall that the orthogonal group $O(n)$ acts on $\mathrm{Sym}(n, \mathbb{R})$ by congruence. We denote this action by $\psi : O(n) \longrightarrow \mathrm{GL}(\mathrm{Sym}(n, \mathbb{R}))$: for every $R \in O(n)$ and $A \in \mathrm{Sym}(n, \mathbb{R})$, $\psi(R)(A) = R^t A R$, where R^t denotes the transpose of R. It is clear that the Frobenius scalar product, and therefore the $\mathrm{GOE}(n)$ distribution, are ψ–invariant. The well–known map

$$\begin{aligned} \phi : \mathrm{Sym}(n, \mathbb{R}) &\longrightarrow \mathbb{R}[x_1, \ldots, x_n]_{(2)} \\ Q &\longmapsto p_Q = x^t Q x \end{aligned} \tag{3.10}$$

defines an isomorphism of $\mathrm{Sym}(n, \mathbb{R})$ with $\mathbb{R}[x_1, \ldots, x_n]_{(2)}$. One can indeed easily check that the orthonormal basis of $\mathrm{Sym}(n, \mathbb{R})$ given by (3.7) is mapped to the orthonormal basis of $\mathbb{R}[x_1, \ldots, x_n]_(2)$ given by (3.6), which for $d = 2$ is given by $\{x_i^2, \sqrt{2} x_i x_j \mid 1 \leq i \leq j \leq n\}$. It follows that ϕ defines an isometry of $\mathrm{Sym}(n, \mathbb{R})$ endowed with the Frobenius inner product with $\mathbb{R}[x_1, \ldots, x_n]_{(2)}$ endowed with the Bombieri–Weyl one and thus we can identify the corresponding standard Gaussian probability distributions. Even more is true: ϕ is an isomorphism of representations between $\psi : O(n) \longrightarrow \mathrm{GL}(\mathrm{Sym}(n, \mathbb{R}))$ and $\beta : O(n) \longrightarrow \mathrm{GL}(\mathbb{R}[x_1, \ldots, x_n]_{(2)})$, i.e. the following diagram

$$\begin{array}{ccc} \mathrm{Sym}(n, \mathbb{R}) & \xrightarrow{\psi(R)} & \mathrm{Sym}(n, \mathbb{R}) \\ \phi \downarrow & & \downarrow \phi \\ \mathbb{R}[x_1, \ldots, x_n]_{(2)} & \xrightarrow[\beta(R)]{} & \mathbb{R}[x_1, \ldots, x_n]_{(2)} \end{array}$$

commutes for every $R \in O(n)$, as one can easily check by a straightforward computation.

Thanks to the isometry ϕ, we can therefore identify $\mathcal{S}^2(n)$ with $\mathbb{R}[x_1, \ldots, x_n]_{(2)}$, not just as vector spaces, but also as metric spaces and as probability spaces, where on the tensor side we consider the Frobenius product and its induced distance and probability distribution, while on the polynomial side we consider the Bombieri–Weyl product and its induced distance and probability distribution. We now show that we can extend this correspondence to any $d \geq 2$. Observe that the polynomial p_Q in (3.10) can be equivalently written as $p_Q = \langle Q, x^{\otimes 2} \rangle_F$. The generalization of (3.10) is then immediate and yields the following result.

Proposition 3.7. *The map ϕ_d given by*

$$\begin{aligned} \phi_d : \mathcal{S}^d(n) &\longrightarrow \mathbb{R}[x_1, \ldots, x_n]_{(d)} \\ T &\longmapsto p_T = \langle T, x^{\otimes d} \rangle_F \end{aligned} \tag{3.11}$$

gives an isometry of Euclidean spaces

$$(\mathcal{S}^d(n), \langle\cdot,\cdot\rangle_F) \xrightarrow{\simeq} (\mathbb{R}[x_1,\dots,x_n]_{(d)}, \langle\cdot,\cdot\rangle_{\mathrm{BW}}).$$

Proof. We only have to check that ϕ_d maps an orthonormal basis for $\left(\mathcal{S}^d(n), \langle\cdot,\cdot\rangle_{\mathrm{F}}\right)$ to an orthonormal basis for $\left(\mathbb{R}[x_1,\dots,x_n]_{(d)}, \langle\cdot,\cdot\rangle_{\mathrm{BW}}\right)$. For a multi–index $\underline{i} = (i_1,\dots,i_d)$ with $1 \le i_1,\dots,i_d \le n$, we say that $\underline{i}$ has type $\alpha = (\alpha_1,\dots,\alpha_n)$ if $\alpha_j = \#\{i_k = j\}$, i.e. for every $j = 1,\dots,n$, α_j is the number of occurences of j in the multi–index $\underline{i}$. It is straightforward to show that the following is an orthonormal basis for the Frobenius scalar product on $\mathcal{S}^d(n)$:

$$\left\{ \left(\frac{1}{\alpha_1!\dots\alpha_n!\ d!}\right)^{\frac{1}{2}} \sum_{\sigma\in S_d} e_{i_{\sigma(1)}}\otimes\dots\otimes e_{i_{\sigma(d)}} \;\middle|\; (i_1,\dots,i_d) \text{ type } \alpha = (\alpha_1,\dots,\alpha_n),\ 1 \le i_1,\dots,i_d \le n \right\},$$

where S_d is the symmetric group on $\{1,\dots,d\}$. Remark that different multi–indices $(i_1,\dots,i_d)$ and $(j_1,\dots,j_d)$ will give the same tensor in the above basis if and only if they have the same type. It follows that we can index the tensors of the basis according to multi–indices $\alpha = (\alpha_1,\dots,\alpha_n) \in \mathbb{Z}^n_{\ge 0}$ with $\alpha_1+\dots+\alpha_n = d$ and we set the notation v_α for the tensor above corresponding to such α.

We claim that ϕ_d sends v_α to the polynomial in (3.6) corresponding to multi–index α. If we prove the claim, the result follows. This is a straightforward computation. Indeed $\phi_d(v_\alpha) = p_{v_\alpha}$ where by definition

$$\begin{aligned}
p_{v_\alpha}(x) =& \langle v_\alpha, x^{\otimes d}\rangle_F \\
=& \langle \left(\frac{1}{\alpha_1!\dots\alpha_n!\ d!}\right)^{\frac{1}{2}} \sum_{\sigma\in S_d} e_{i_{\sigma_1}}\otimes\dots\otimes e_{i_{\sigma(d)}}, \left(\sum_{j_1=1}^n x_{j_1}e_{j_1}\right)\otimes\dots\otimes\left(\sum_{j_d=1}^n x_{j_d}e_{j_d}\right)\rangle_F \\
=& \left(\frac{1}{\alpha_1!\dots\alpha_n!\ d!}\right)^{\frac{1}{2}} \sum_{\sigma\in S_d}\sum_{j_1,\dots,j_d=1}^n \langle e_{i_{\sigma_1}}\otimes\dots\otimes e_{i_{\sigma(d)}}, x_{j_1}\dots x_{j_d}e_{j_1}\otimes\dots\otimes e_{j_d}\rangle_F \\
=& \left(\frac{1}{\alpha_1!\dots\alpha_n!\ d!}\right)^{\frac{1}{2}} \sum_{\sigma\in S_d} x_{i_{\sigma(1)}}\dots x_{i_{\sigma(d)}} = \left(\frac{1}{\alpha_1!\dots\alpha_n!\ d!}\right)^{\frac{1}{2}} d!\ x_1^{\alpha_1}\dots x_n^{\alpha_n} = \\
=& \left(\frac{d!}{\alpha_1!\dots\alpha_n!}\right)^{\frac{1}{2}} x_1^{\alpha_1}\dots x_n^{\alpha_n},
\end{aligned}$$

where the second to last equation follows from $x_{i_{\sigma(1)}}\dots x_{i_{\sigma(d)}} = x_1^{\alpha_1}\dots x_n^{\alpha_n}$ for every $\sigma\in S_d$, since permutations of the indices do not change the type of the multi–index, which is always $\alpha = (\alpha_1,\dots,\alpha_n)$. The proof is concluded. □

Proposition 3.7 shows that in our formulation of (3.5) we could substitute $\mathcal{S}^d(n)$ with the degree d homogeneous polynomials, the Frobenius distance with the Bombieri–Weyl one and the probability measure with the one induced by Bombieri–Weyl. But in the description of symmetric tensors as homogeneous polynomials, what plays the role of rank–one tensors? This is the question we address in the next section. We will show that rank–one symmetric tensors have a neat description as powers of linear forms.

3.1.2 The Veronese variety

For notational reasons, from now on we will consider the space of real homogeneous polynomials of degree d in $n+1$ variables $\mathbb{R}[x_0,\dots,x_n]_{(d)}$, endowed with the corresponding Bombieri–Weyl product with basis given by (3.6), where we now switch from n to $n+1$. In the same way, we will from now on consider $\mathcal{S}^d(n+1)$, that is order d real symmetric tensors on $\mathbb{R}^{n+1}$. We begin with the following definition.

Definition 3.8. For $n \geq 1$ and $d \geq 1$, the *real Bombieri–Weyl Veronese embedding* is the map

$$\begin{aligned} \nu_{n,d} : \mathbb{R}\mathrm{P}^n &\longrightarrow \mathbb{R}\mathrm{P}^N \\ [a] &\longmapsto \left[\binom{d}{\alpha}^{1/2} a^\alpha \right], \end{aligned}$$

where $\mathbb{R}\mathrm{P}^N \simeq \mathbb{P}(\mathbb{R}[x_0,\dots,x_n]_{(d)})$; the coordinates on the right are to be considered with respect to the Bombieri–Weyl basis. It is the Veronese projective embedding associated to the Bombieri–Weyl basis. We therefore call the *Bombieri–Weyl projective Veronese variety* the image of this embedding and we denote it by $\mathcal{P}V_{n,d} := \mathrm{im}(\nu_{n,d})$. The corresponding cone over $\mathcal{P}V_{n,d}$ is denoted by $\mathcal{V}_{n,d}$ and called the *Bombieri–Weyl Veronese variety*.

Remark 3.9. The construction in Definition 3.8 is a general one: given any basis of $\mathbb{R}[x_0,\dots,x_n]_{(d)}$, one can define an embedding by sending a point $[a] \in \mathbb{R}\mathrm{P}^n$ to the point of $\mathbb{R}\mathrm{P}^N$ whose coordinates with respect to the given basis are the evaluations of the polynomials of the basis at the point $[a]$; this is well–defined because the polynomials are homogeneous of the same degree. The *standard* Veronese variety is the one corresponding to the monomial basis. Moreover the construction works also over $\mathbb{C}$: complex Veronese varieties are indeed very classical objects in algebraic geometry and have been intensively studied. Notice that the "metric"properties of a Veronese variety depend dramatically on the chosen basis and therefore the chosen embedding. Here we will only work with the Bombieri–Weyl one because of the invariance properties this basis enjoy, as highlighted in the previous section. For this reason, in the following we might refer to this variety simply as the Veronese variety, omitting the reference to Bombieri–Weyl.

The main object we will consider in this section is the spherical counterpart of $\mathcal{P}V_{n,d}$, given as follows.

Definition 3.10. The *spherical (Bombieri–Weyl) Veronese map* is the map

$$\begin{aligned} \widehat{\nu}_{n,d} : S^n &\longrightarrow S^N \\ a &\longmapsto \left(\binom{d}{\alpha}^{1/2} a^\alpha \right), \end{aligned}$$

where S^n is the sphere with respect to the standard Euclidean product in $\mathbb{R}^{n+1}$ while S^N is the sphere with respect to the Bombieri–Weyl product in $\mathbb{R}[x_0,\dots,x_n]_{(d)}$. One can check via a straightforward computation that this map is well–defined. We call the image of this map the *spherical (Bombieri–Weyl) Veronese surface* and we denote it by $\mathcal{S}V_{n,d} := \mathrm{im}(\widehat{\nu}_{n,d})$.

The objects we just introduced have a nice description in terms of linear forms. Recall that to each vector $b = (b_0,\dots,b_n) \in \mathbb{R}^{n+1}$ we can associate the linear form on $\mathbb{R}^{n+1}$ given

by $l_b(x_0,\dots,x_n) = b_0x_0+\cdots+b_nx_n$. It is easy to show that $\mathcal{P}V_{n,d}$ parametrizes projective classes of $d-$th powers of linear forms

$$\begin{aligned}\mathcal{P}V_{n,d} &= \left\{[d\text{–th powers of linear forms on } \mathbb{R}^{n+1}]\right\} \\ &= \left\{ [a_\alpha] \in \mathbb{R}\mathrm{P}^N \mid \exists\, b=(b_0,\dots,b_n)\in\mathbb{R}^{n+1} \text{ s.t. } a_{\alpha_0,\dots,\alpha_n} = \binom{d}{\alpha}^{1/2} b_0^{\alpha_0}\dots b_n^{\alpha_n} \right\}\end{aligned} \tag{3.12}$$

as one can prove by showing that $\nu_{n,d}([b_0,\dots,b_n]) = [(b_0x_0+\cdots+b_nx_n)^d]$. Therefore the Veronese variety $\mathcal{V}_{n,d}$ parametrizes signed d–th powers of linear forms on $\mathbb{R}^{n+1}$. A similar description holds for the spherical Veronese surface, for which we have to add the condition that the linear forms have norm 1, that is to say that the associated vector has norm 1

$$\begin{aligned}\mathcal{S}V_{n,d} &= \left\{d\text{–th powers of norm–1 linear forms on } \mathbb{R}^{n+1}\right\} \\ &= \left\{ (a_\alpha) \in S^N \mid \exists\, b=(b_0,\dots,b_n)\in S^n \text{ s.t. } a_{\alpha_0,\dots,\alpha_n} = \binom{d}{\alpha}^{1/2} b_0^{\alpha_0}\dots b_n^{\alpha_n} \right\}.\end{aligned} \tag{3.13}$$

These descriptions of $\mathcal{V}_{n,d}$ and $\mathcal{S}V_{n,d}$ as powers of linear forms have two crucial consequences. The first one is related to the isomorphism between symmetric tensors and homogeneous polynomials described in Proposition 3.7. It is straightforward to show that given a symmetric rank–one tensor $T=\pm v^{\otimes d}$ with $v=(v_0,\dots,v_n)\in\mathbb{R}^{n+1}$, its image through the isomorphism (3.11) is given by the polynomial

$$p_T(x) = \langle \pm v^{\otimes d}, x^{\otimes d}\rangle_F = \pm(\langle v,x\rangle)^d = \pm(v_0x_0+\cdots+v_nx_n)^d,$$

that is a (signed) d–th power of a linear form. Viceversa, it is clear that any (signed) d–th power of a linear form can be obtained this way. It follows that the set of symmetric rank–one tensors $\mathcal{S}_1^d(n+1)$ can be identified via the isomorphism ϕ_d (3.11) with the Veronese variety $\mathcal{V}_{n,d}$.

The second consequence of (3.12) and (3.13) is the following result, which is also the main reason we work with this specific basis.

Proposition 3.11. *$\mathcal{P}V_{n,d}$ is an orbit for the action of the orthogonal group $O(n+1)$ induced by β on $\mathbb{R}\mathrm{P}^N = \mathbb{P}(\mathbb{R}[x_0,\dots,x_n]_{(d)})$ by change of variables. Similarly $\mathcal{S}V_{n,d}$ is an orbit for the induced action of the orthogonal group $O(n+1)$ on $S^N = S(\mathbb{R}[x_0,\dots,x_n]_{(d)})$.*

Proof. Consider the point $[x_0^d]\in\mathcal{P}V_{n,d}$ and another point $p=[(b_0x_0+\cdots+b_nx_n)^d]\in\mathcal{P}V_{n,d}$. Since we are working in a projective space, we can assume without loss of generality that the chosen representative is such that $b_0^2+\cdots+b_n^2=1$. Since the orthogonal group $O(n+1)$ acts transitively on the sphere S^n, there exists an orthogonal matrix R that maps the point $(1,0,\dots,0)\in S^n$ to the point $(b_0,\dots,b_n)\in S^n$. It follows that the map $\beta(R)$ sends the polynomial x_0^d to the polynomial $(b_0x_0+\cdots+b_nx_n)^d$, i.e. $\beta(R)([x_0^d]) = [(b_0x_0+\cdots+b_nx_n)^d]$ and $\mathcal{P}V_{n,d}$ is an orbit. The proof for $\mathcal{S}V_{n,d}$ is the same. □

We define the *spherical Veronese variety* $V_{n,d}$ to be the intersection of the affine Veronese variety with the unit sphere S^N:

$$V_{n,d} := \mathcal{V}_{n,d}\cap S^N.$$

Recall the usual two–fold covering map $\pi_N : S^N \longrightarrow \mathbb{R}\mathrm{P}^N$ given by the identification of antipodal points. Its restriction to the spherical Veronese surface gives a covering map

$$\widehat{\pi}_{n,d} : \mathcal{S}V_{n,d} \longrightarrow \mathcal{P}V_{n,d}$$

whose degree depends on the parity of d: if d is odd, $\widehat{\pi}_{n,d}$ is $2:1$ covering, while if d is even it is $1:1$, since in this case for $b \in S^n$ we have $\widehat{\nu}_{n,d}(b) = \widehat{\nu}_{n,d}(-b)$. Indeed we have that

$$V_{n,d} = \mathcal{S}V_{n,d} \cup -\mathcal{S}V_{n,d}, \tag{3.14}$$

where for d odd we have $\mathcal{S}V_{n,d} = -\mathcal{S}V_{n,d}$, $V_{n,d}$ is connected and coincides with $\mathcal{S}V_{n,d}$, while for d even $\mathcal{S}V_{n,d} \cap (-\mathcal{S}V_{n,d}) = \emptyset$, $V_{n,d}$ has two different connected components and it is a double copy of $\mathcal{S}V_{n,d}$. This is the reason we use different terms to describe $\mathcal{S}V_{n,d}$ and $V_{n,d}$, highlighting also the fact that $V_{n,d}$ is an algebraic variety. In the projective setting these differences cease to exist, since

$$\mathcal{P}V_{n,d} = \mathbb{P}(\mathcal{S}V_{n,d}) = \mathbb{P}(V_{n,d}) \subseteq \mathbb{R}\mathrm{P}^N.$$

We now turn to metric properties of the various Veronese objects we introduced. As usual, we consider the standard Euclidean metric on $\mathbb{R}^{n+1}$ and the Bombieri–Weyl one on $\mathbb{R}^{N+1} = \mathbb{R}[x_0, \dots, x_n]_{(d)}$. The metrics induced on S^n and S^N respectively are invariant under the antipodal map and therefore induce metrics on the corresponding projective spaces $\mathbb{R}\mathrm{P}^n$ and $\mathbb{R}\mathrm{P}^N$. We denote the round metrics on the spheres by g_{S^n} and g_{S^N} and those on the projective spaces by $g_{\mathbb{R}\mathrm{P}^n}$ and $g_{\mathbb{R}\mathrm{P}^N}$. Since the covering maps π_n, π_N are Riemannian coverings with respect to these metrics, any relation between g_{S^n} and g_{S^N} will also hold between $g_{\mathbb{R}\mathrm{P}^n}$ and $g_{\mathbb{R}\mathrm{P}^N}$ and viceversa. Through a direct computation, one can prove the following result.

Proposition 3.12. *Pulling back the Bombieri–Weyl metric on $\mathbb{R}\mathrm{P}^N$ through the Veronese embedding $\nu_{n,d} : \mathbb{R}\mathrm{P}^n \longrightarrow \mathbb{R}\mathrm{P}^N$, for any $n \geq 1$, $d \geq 1$ we have*

$$\nu_{n,d}^* \, g_{\mathbb{R}\mathrm{P}^N} = \sqrt{d} \, g_{\mathbb{R}\mathrm{P}^n}.$$

Sketch of proof. We will only prove the toy case $n = 1$, i.e. $\nu_{1,d} : \mathbb{R}\mathrm{P}^1 \longrightarrow \mathbb{R}\mathrm{P}^d$; the generalization to any $n \geq 1$ will follow the same steps. Since any relation between $g_{\mathbb{R}\mathrm{P}^1}$ and $g_{\mathbb{R}\mathrm{P}^d}$ will be the same between g_{S^1} and g_{S^d}, we look at the spherical Veronese embedding $\widehat{\nu}_{1,d} : S^1 \longrightarrow S^d$. Fix the point $p = (1, 0) \in S^1$. The tangent space of S^1 at $(1, 0)$ is spanned by the vector $v = (0, 1)$, that has norm equal to 1 with respect to the round metric g_{S^1}. Its norm with respect to the pullback metric is given by

$$\|v\|^2_{\widehat{\nu}^*_{1,d} g_{S^d}} = \widehat{\nu}^*_{1,d} g_{S^d}(v, v) = g_{S^d}(D_p\widehat{\nu}_{1,d}\,(v), D_p\widehat{\nu}_{1,d}\,(v)).$$

In order to compute the differential we consider the curve $\sigma : (-\varepsilon, \varepsilon) \longrightarrow S^1$ given by $\sigma(t) = (\cos t, \sin t)$, which is an adapted curve to v. Then we have

$$D_p\widehat{\nu}_{1,d}\,(v) = \frac{d}{dt}\widehat{\nu}_{1,d}\,(\sigma(t))\,\big|_{t=0} = (0, \sqrt{d}, 0, \dots, 0).$$

Recalling that the coordinates on $\mathbb{R}^d$ are expressed with respect to an orthonormal basis, it follows that

$$\|v\|^2_{\widehat{\nu}^*_{1,d} g_{S^d}} = d.$$

Since standard rotations are isometries for g_{S^1} and the corresponding induced trasnformations on S^d via $\widehat{\nu}_{1,d}$ are isometries for the Bombieri–Weyl metric, the computation above actually holds for any point in S^1. It follows that $\widehat{\nu}^*_{1,d}\, g_{S^d} = \sqrt{d}\, g_{S^1}$. □

The following corollary expresses the volume of the spherical Veronese surface $\mathcal{S}V_{n,d}$ in terms of the volume of the round sphere S^n.

Corollary 3.13. *For every $n, d \in \mathbb{N}$ and any smooth n–dimensional submanifold $C \hookrightarrow \mathcal{SV}_{n,d}$ we have*

$$\mathrm{Vol}_{n,\mathrm{BW}}(C) = \begin{cases} \frac{1}{2} d^{\frac{n}{2}} \mathrm{Vol}_n\left(\widehat{\nu}_{n,d}^{-1}(C)\right) & \text{for } d \text{ even} \\ \\ d^{\frac{n}{2}} \mathrm{Vol}_n\left(\widehat{\nu}_{n,d}^{-1}(C)\right) & \text{for } d \text{ odd} \end{cases},$$

where Vol_n is the n–dimensional volume with respect to g_{S^n} and $\mathrm{Vol}_{n,\mathrm{BW}}$ is the n–dimensional volume with respect to the metric induced by g_{S^N} on $\mathcal{SV}_{n,d}$. In particular

$$\mathrm{Vol}_{n,BW}(\mathcal{SV}_{n,d}) = \begin{cases} d^{\frac{n}{2}} \frac{\pi^{\frac{n+1}{2}}}{\Gamma(\frac{n+1}{2})} & \text{for } d \text{ even} \\ \\ 2d^{\frac{n}{2}} \frac{\pi^{\frac{n+1}{2}}}{\Gamma(\frac{n+1}{2})} & \text{for } d \text{ odd} \end{cases}. \tag{3.15}$$

In section 3.1.5 we will need the explicit expression of the volume of the spherical Veronese variety $V_{n,d}$. This immediately follows from formula (3.15), (3.14) and the discussion thereafter. Indeed we obtain

$$\mathrm{Vol}_{n,\mathrm{BW}}(V_{n,d}) = 2d^{\frac{n}{2}} \frac{\pi^{\frac{n+1}{2}}}{\Gamma(\frac{n+1}{2})} = d^{\frac{n}{2}} \mathrm{Vol}(S^n),$$

independent of the parity of d, where $\mathrm{Vol}(S^n)$ is the standard volume of the sphere.

Remark 3.14. For every orthogonal matrix $R \in O(n+1)$ we have the following commutative diagram

$$\begin{array}{ccc} (\mathbb{R}\mathrm{P}^n, g_{\mathbb{R}\mathrm{P}^n}) & \xrightarrow{R} & (\mathbb{R}\mathrm{P}^n, g_{\mathbb{R}\mathrm{P}^n}) \\ \downarrow{\scriptstyle \nu_{n,d}} & & \downarrow{\scriptstyle \nu_{n,d}} \\ \mathcal{P}V_{n,d} & \xrightarrow{\beta(R)|_{\mathcal{P}V_{n,d}}} & \mathcal{P}V_{n,d} \end{array},$$

since by Proposition 3.11 $\mathcal{P}V_{n,d}$ is an orbit for the action β and it is therefore preserved under $\beta(R)$. Moreover, by the invariance of the Bombieri–Weyl scalar product, $\beta(R)$ is an isometry of $(\mathbb{R}\mathrm{P}^N, g_{\mathbb{R}\mathrm{P}^N})$, and its restriction to $\mathcal{P}V_{n,d}$ defines an isometry of $\mathcal{P}V_{n,d}$. Therefore $\mathcal{P}V_{n,d}$ is an orbit for an isometric action of $O(n+1)$ over $\mathbb{R}[x_0, \ldots, x_n]_{(d)}$ and these isometries of $\mathcal{P}V_{n,d}$ are induced by isometries of the ambient space. The same property also holds for $\mathcal{SV}_{n,d}$ considering the induced action on the sphere S^N. This simple observation, which essentially follows from Proposition 3.11, is of fundamental importance, as it will allow us to drastically simplify the computations we will carry out in Section 3.1.4 and Section 3.1.5. Indeed, the consequence of this is that the extrinsic geometry of the isometric embedding $\mathcal{SV}_{n,d} \hookrightarrow S^N$ is exactly the same at every point of $\mathcal{SV}_{n,d}$; the same conclusion clearly holds for $V_{n,d} \hookrightarrow S^N$.

3.1.3 A restatement with polynomials

We are now ready to restate (3.5) in terms of homogeneous polynomials, exploiting the isomorphism ϕ_d given in Proposition 3.7. As we already stressed, symmetric tensors of order d on $\mathbb{R}^{n+1}$ correspond via ϕ_d to homogeneous polynomials of degree d in $n+1$ variables. We have seen in this section how the set of rank–one tensors in this correspondence is identified with the affine Veronese variety $\mathcal{V}_{n,d}$. Moreover the Bombieri–Weyl distance for

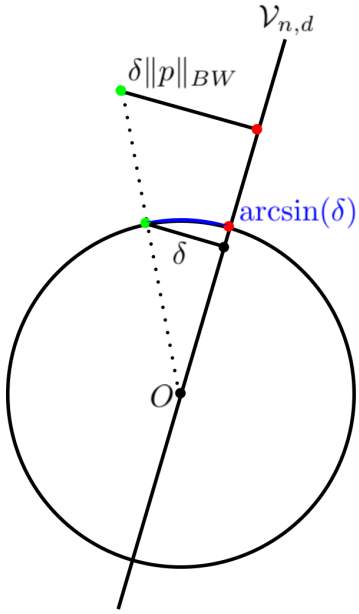


Figure 3.1: If a point is at distance $\delta\|p\|_{BW}$ from $\mathcal{V}_{n,d}$, rescaling it to have unit norm gives a point at distance δ from $\mathcal{V}_{n,d}$. The closest point to it on the spherical Veronese is on the same ray as the closest point on $\mathcal{V}_{n,d}$, but at spherical distance $\arcsin(\delta)$.

polynomials corresponds to the Frobenius one for tensors and the same happens at the level of induced probability measures. Therefore (3.5) is equivalent to

$$\mathbb{P}\left\{p \in \mathbb{R}[x_0,\dots,x_n]_d \;\middle|\; d^{\mathrm{BW}}(p,\mathcal{V}_{n,d}) \le \delta\|p\|_{\mathrm{BW}}\right\}, \tag{3.16}$$

where d^{BW} denotes the Bombieri–Weyl distance on $\mathbb{R}[x_0,\dots,x_n]_{(d)}$ and $\|\cdot\|_{\mathrm{BW}}$ is the corresponding norm. We now observe that if $p \in \mathbb{R}[x_0,\dots,x_n]_{(d)}$ is such that its distance from $\mathcal{V}_{n,d}$ is less than $\delta\|p\|_{\mathrm{BW}}$, then if we rescale everything by $\|p\|_{\mathrm{BW}}$ we obtain a point in S^N at distance less than δ from $\mathcal{V}_{n,d}$. Viceversa, given a point in S^N at distance less than δ from $\mathcal{V}_{n,d}$, every point in the affine ray over it will be at distance less than δ times its norm from the affine Veronese variety. Since a centered non–degenerate Gaussian measure induces the uniform distribution on the sphere, this implies that the quantity in (3.47) is equivalent to the volume of the intersection of a tubular neighborhood of radius δ around $\mathcal{V}_{n,d}$ with the unit sphere $S^N \subseteq \mathbb{R}[x_0,\dots,x_n]_{(d)}$, normalized by the volume of S^N. It is an exercise of elementary geometry to show that if we consider the spherical Veronese variety $V_{n,d}$, a tubular neighborhood of $\mathcal{V}_{n,d}$ of radius δ in the Bombieri–Weyl metric corresponds to a (spherical) tubular neighborhood of $V_{n,d}$ of radius $\varepsilon := \arcsin(\delta)$ in the induced metric g_{S^N} (see Fig. 3.1). We thus finally arrive to the equality

$$\mathbb{P}\left\{T \in \mathcal{S}^d(n) \;\middle|\; \widehat{\delta}(T,\mathcal{S}_1^d(n)) \le \delta\|T\|_F\right\} = \frac{\mathrm{Vol}\left(\mathcal{U}(V_{n,d},\varepsilon)\right)}{\mathrm{Vol}(S^N)}. \tag{3.17}$$

For practical reasons, we will work with $\mathcal{S}V_{n,d}$ instead of the full $V_{n,d}$. For this reason there will be a factor of 2 to be taken into account when switching from the Veronese surface $\mathcal{S}V_{n,d}$ to the rank–one variety $V_{n,d}$, depending on the parity of d.

Our aim in the next sections will be to exploit Weyl's tube formula to compute the volume of the tubular neighborhood $\mathcal{U}\left(\mathcal{S}V_{n,d},\varepsilon\right)$, from the result for $V_{n,d}$ will follow. As we discussed in Section 2.4, this formula holds only when the radius of the neighborhood is

smaller than the reach of the submanifold we introduced in Section 2.3. We therefore need to understand the reach of $\mathcal{S}V_{n,d}$ to be able to meaningfully use Weyl's result. We will in the next section this that this often unfeasible task becomes doable for our case thanks to the invariance properties discussed in Remark 3.14. We will then move on to compute the curvature coefficients of the spherical Veronese surface.

3.1.4 The reach of the spherical Veronese variety

In this section we will provide an explicit computation for the reach of $\mathcal{S}V_{n,d} \hookrightarrow S^N$. As theorem 2.9 shows, this quantity will serve as a lower bound on the radii for which Weyl's tube formula holds. Remark that, since $\mathcal{S}V_{n,d}$ is compact, by theorem 2.3 we have $\rho(\mathcal{S}V_{n,d}) > 0$.

Recall that by Theorem 2.7 we can express the reach as

$$\rho(\mathcal{S}V_{n,d}) = \min\left\{\mathrm{bw}(\mathcal{S}V_{n,d}),\ \frac{1}{\kappa_{\max}(\mathcal{S}V_{n,d})}\right\},$$

where the bottleneck width is given by formula (2.3) while the maximal curvature by formula (2.4). Using the observations in Remark 3.14 we can drastically simplify those expressions.

Lemma 3.15. *The formula in (2.4) simplifies to*

$$\kappa_{\max}(\mathcal{S}V_{n,d}) = \sup\left\{\, \|\ddot{\gamma}(0)\|_{\mathrm{BW}} \;\middle|\; \gamma : (-\delta,\delta) \longrightarrow \mathcal{S}V_{n,d} \text{ arclength geodesic in } \mathcal{S}V_{n,d} \text{ with } \gamma(0) = x_0^d \,\right\}, \tag{3.18}$$

that is to say the inner supremum in (2.4) does not depend on $x \in \mathcal{S}V_{n,d}$. Moreover (2.4) does not depend on the direction of $\dot{\gamma}(0)$. Similarly, the formula in (2.3) simplifies to

$$\mathrm{bw}(\mathcal{S}V_{n,d}) = \frac{1}{2}\ \inf\{l(\gamma) \mid \gamma : [a,b] \longrightarrow S^N \text{ geodesic s.t. } \gamma(a) = x_0^d,\ \gamma(b) \in \mathcal{S}V_{n,d},\ \dot{\gamma}(a) \in N_{x_0^d}\mathcal{S}V_{n,d},\ \dot{\gamma}(b) \in N_{\gamma(b)}\mathcal{S}V_{n,d}\}. \tag{3.19}$$

Proof. Consider an arclength geodesic $\gamma : (-\delta,\delta) \longrightarrow \mathcal{S}V_{n,d}$ with $\gamma(0) = p_1$ and pick another point $p_2 \in \mathcal{S}V_{n,d}$. By remark 3.14 there exists $R \in O(n+1)$ such that $\beta(R)(p_1) = p_2$. Recall that the image of a geodesic through an isometry is still a geodesic, hence the curve $\tilde{\gamma} := \beta(R) \circ \gamma$ is an arclength geodesic with $\tilde{\gamma}(0) = \beta(R)(\gamma(0)) = \beta(R)(p_1) = p_2$. Since $\beta(R)$ is not only an isometry of $\mathcal{S}V_{n,d}$ but also an isometry of the ambient space S^N, we have $\|\ddot{\gamma}(0)\|_{\mathrm{BW}} = \|\ddot{\tilde{\gamma}}(0)\|_{\mathrm{BW}}$. We just proved that given any two points in $\mathcal{S}V_{n,d}$, using the isometries $\beta(R)$ for $R \in O(n+1)$ we can transport any arclength geodesic passing through the first point into another arclength geodesic passing through the second point, preserving the norm of second derivatives. It follows that the expression in (2.4) is independent of the specific point $p \in \mathcal{S}V_{n,d}$. Now observe that given any arclength geodesic γ with $\gamma(0) = x_0^d$ fixed and tangent vector $\dot{\gamma}(0) = v$, we can change the direction of $\dot{\gamma}(0)$ through $\beta(R)$ for some $R \in O(n+1)$ with x_0^d being a fixed point (it is sufficient to choose a rotation R such that $(1,0,\ldots,0) \in S^n$ is in the axis of rotation), obtaining any other possible direction in $T_{x_0^d}\mathcal{S}V_{n,d}$, again without changing $\|\ddot{\gamma}(0)\|_{\mathrm{BW}}$. It follows that (3.18) does not depend on the specific direction of $\dot{\gamma}(0)$.
Since isometries preserve orthogonality and lengths, the second part of the lemma also follows in a similar way. □

The choice of the specific point x_0^d in formulas (3.18) and (3.19) is motivated by convenience for computations only; we could have chosen any other point.

The first step to compute (3.18) and (3.19) is to understand tangent and normal spaces of $\mathcal{SV}_{n,d} \longrightarrow S^N$. The following lemma describes tangent spaces. Normal spaces will then be the orthogonal complements of (3.20) in the tangent spaces of the ambient manifold S^N.

Lemma 3.16. *For $p \in \mathcal{SV}_{n,d}$ with $p = l^d$, where l is a norm-1 linear form, we have*

$$T_p\mathcal{SV}_{n,d} = \big\langle \{l^{d-1}\lambda \mid \lambda \text{ is a linear form orthogonal to } l\} \big\rangle, \tag{3.20}$$

where the brackets indicate the linear span.

Proof. Write $l(x) = a_0x_0 + \cdots + a_nx_n$ and set $a = (a_0, \ldots, a_n) \in S^n$. Recalling that $\mathcal{SV}_{n,d} = \mathrm{im}(\widehat{\nu}_{n,d})$, a curve on $\mathcal{SV}_{n,d}$ can be expressed as the image of a curve on S^n through $\widehat{\nu}_{n,d}$. Consider $b = (b_0, \ldots, b_n) \in S^n$ such that b is orthogonal to a. Then $\gamma(t) = (\cos t(a_0x_0 + \cdots + a_nx_n) + \sin t(b_0x_0 + \cdots + b_nx_n))^d$ is a curve in $\mathcal{SV}_{n,d}$ with $\gamma(0) = p$. Remark that the orthogonality condition is needed to ensure we are taking the d–th power of a norm–1 form. A computation shows that

$$\frac{d}{dt}\gamma(t)\big|_{t=0} = d\ l^{d-1}(b_0x_0 + \cdots + b_nx_n),$$

therefore $\big\langle \{l^{d-1}\lambda \mid \lambda \text{ is a linear form orthogonal to } l\} \big\rangle \subseteq T_p\mathcal{SV}_{n,d}$. By dimension count, equality follows. □

We are now ready to perform the computation of the reach $\rho(\mathcal{SV}_{n,d})$.

Theorem 3.17. *The inverse of the maximal curvature and the bottleneck width of $\mathcal{SV}_{n,d} \hookrightarrow S^N$ are*

$$\kappa_{\max}^{-1}(\mathcal{SV}_{n,d}) = \frac{1}{\sqrt{2}}\sqrt{1 + \frac{1}{d-1}} = \frac{1}{\sqrt{2}} + \frac{1}{2(d-1)\sqrt{2}} + \mathcal{O}\left(\frac{1}{d^2}\right),$$

$$\mathrm{bw}(\mathcal{SV}_{n,d}) = \frac{\pi}{4}.$$

Therefore the reach of the embedding $\mathcal{SV}_{n,d} \hookrightarrow S^N$ is given by

$$\rho(\mathcal{SV}_{n,d}) = \min\{\kappa_{\max}^{-1}(\mathcal{SV}_{n,d}),\ \mathrm{bw}(\mathcal{SV}_{n,d})\} = \begin{cases} \frac{\pi}{4} & 2 \le d \le 5 \\ \frac{1}{\sqrt{2}}\sqrt{1 + \frac{1}{d-1}} & d \ge 6 \end{cases}.$$

Proof. We begin with the computation of $\kappa_{\max}^{-1}(\mathcal{SV}_{n,d})$. By Proposition 3.12 geodesics in $\mathcal{SV}_{n,d}$ can be realized as images through $\widehat{\nu}_{n,d}$ of geodesics in S^n. Moreover, thanks to Lemma 3.15, it is enough to consider geodesics passing through $x_0^d = (1, 0, \ldots, 0) = \widehat{\nu}_{n,d}((1, 0, \ldots, 0))$ at time 0. As their direction plays no role, it is enough to consider the image of the single geodesic in S^n given by $\alpha(t) = x_0\cos(td^{-\frac{1}{2}}) + x_1\sin(td^{-\frac{1}{2}})$ with x_0 being the point of coordinates $(1, 0, \ldots, 0)$ and x_1 that of coordinates $(0, 1, 0, \ldots, 0)$. Explicitly we have $\alpha(t) = (\cos(td^{-\frac{1}{2}}), \sin(td^{-\frac{1}{2}}), 0, \ldots, 0)$ and the corresponding geodesic in $\mathcal{SV}_{n,d}$ is given by

$$\gamma(t) := (\widehat{\nu}_{n,d} \circ \alpha)(t) = \big(\cos^d(td^{-\frac{1}{2}}), \sqrt{d}\ \cos^{d-1}(td^{-\frac{1}{2}})\ \sin(td^{-\frac{1}{2}}), \ldots, \sin^d(td^{-\frac{1}{2}})\big),$$

with $\gamma(0) = x_0^d = (1, 0, \dots, 0)$ and $\|\dot\gamma(0)\|_{\mathrm{BW}} = 1$, so that γ is an arclength geodesic through x_0^d. Notice that the only components of $\gamma(t)$ which are not constantly zero are those corresponding to multi–indices $(\beta_0, \beta_1, 0, \dots, 0)$ with $\beta_0 + \beta_1 = d$. To compute $\|\ddot\gamma(0)\|$, we need the second derivatives of the non–constantly zero components of $\gamma(t)$. These components have the following expression for $k = 0, \dots, d$:

$$\binom{d}{k}^{\frac{1}{2}} \cos^k(td^{-\frac{1}{2}}) \sin^{d-k}(td^{-\frac{1}{2}}).$$

Computing second derivatives and evaluating at $t = 0$ we find

$$\ddot\gamma(0) = \left(-1, 0, \sqrt{2}\frac{\sqrt{d(d-1)}}{d}, 0, \dots, 0\right).$$

Since we are looking at $\mathcal{S}V_{n,d}$ as a submanifold of S^N, we first need to project this vector to the tangent space $T_{x_0^d}S^N$ and then compute its norm. This way we obtain

$$\begin{aligned}\|\ddot\gamma(0)\|_{\mathrm{BW}} =& \|\mathrm{proj}_{T_{x_0^d}S^N}\big(-1, 0, \sqrt{2}\frac{\sqrt{d(d-1)}}{d}, 0, \dots, 0\big)\|_{\mathrm{BW}} = \|\big(0, \sqrt{2}\frac{\sqrt{d(d-1)}}{d}, 0 \dots, 0\big)\|_{\mathrm{BW}} \\ =& \sqrt{2}\sqrt{\frac{d-1}{d}}.\end{aligned}$$

As a consequence we find that

$$\kappa_{\max}^{-1}(\mathcal{S}V_{n,d}) = \frac{1}{\|\ddot\gamma(0)\|_{\mathrm{BW}}} = \frac{1}{\sqrt{2}}\sqrt{\frac{d}{d-1}} = \frac{1}{\sqrt{2}}\sqrt{1 + \frac{1}{d-1}} = \frac{1}{\sqrt{2}} + \frac{1}{2(d-1)\sqrt{2}} + \mathcal{O}\left(\frac{1}{d^2}\right),$$

where in the last step we use the Taylor-MacLaurin expansion $\sqrt{1+x} = 1 + \frac{x}{2} + \mathcal{O}(x^2)$ with center $x = \frac{1}{d-1}$.

Remark 3.18. Notice that, since $\kappa_{\max}^{-1}(\mathcal{S}V_{n,d}) = \frac{1}{\sqrt{2}}\sqrt{1 + \frac{1}{d-1}}$, then $\kappa_{\max}^{-1}(\mathcal{S}V_{n,d}) > \frac{1}{\sqrt{2}}$ for all n, d.

For the computation of $\mathrm{bw}(\mathcal{S}V_{n,d})$, by Lemma 3.15 it is enough to consider geodesics $\gamma(\theta)$ in S^N starting from x_0^d at time $\theta = 0$. Recalling that the normal space at a point $p \in \mathcal{S}V_{n,d}$ is the orthogonal complement of $T_p\mathcal{S}V_{n,d}$ inside the tangent space T_pS^N, by Lemma 3.20 we have

$$N_{x_0^d}\mathcal{S}V_{n,d} = \left\langle \left\{ \binom{d}{\alpha}^{\frac{1}{2}} x_0^{\alpha_0} \dots x_n^{\alpha_n} \mid \alpha_0 < d-1 \right\} \right\rangle. \tag{3.21}$$

Pick a vector $w \in N_{x_0^d}\mathcal{S}V_{n,d}$ and let $\gamma_w(\theta)$ be the geodesic in S^N with $\gamma_w(0) = x_0^d$ and $\dot\gamma_w(0) = w$, i.e.

$$\gamma_w(\theta) = x_0^d \ \cos\big(\theta\|w\|\big) + \frac{w}{\|w\|} \ \sin\big(\theta\|w\|\big),$$

where all norms from now on as to be understood as Bombieri–Weyl ones. The goal now is to understand under which conditions on w such a geodesic γ_w meets again $\mathcal{S}V_{n,d}$ orthogonally. The first step is to find for which $b = (b_0, \dots, b_n) \in S^n$ and θ we at least have a solution to the equation

$$(b_0x_0 + \dots + b_nx_n)^d = x_0^d \ \cos\big(\theta\|w\|\big) + \frac{w}{\|w\|} \ \sin\big(\theta\|w\|\big).$$

On the right hand side, we can expand w as $w = \sum_{\alpha_0<d-1} w_\alpha \binom{d}{\alpha}^{\frac{1}{2}} x^\alpha$, while we expand the left hand side as $\sum_\alpha \binom{d}{\alpha} b^\alpha x^\alpha$ with multi–indices $\alpha = (\alpha_0, \dots, \alpha_n) \in \mathbb{Z}_{\geq 0}^{n+1}$ such that $\alpha_0 + \cdots + \alpha_n = d$. Hence, expanding it further in the Bombieri–Weyl basis, we get the equation

$$b_0^d x_0^d + \sqrt{d}\, x_0^{d-1}\left(\sum_{i=1}^n \sqrt{d}\, b_0^{d-1} b_i x_i\right) + \sum_{\alpha_0<d-1} \binom{d}{\alpha} b^\alpha x^\alpha = \tag{3.22}$$
$$x_0^d\, \cos\big(\theta\|w\|\big) + \sin\big(\theta\|w\|\big) \sum_{\alpha_0<d-1} \binom{d}{\alpha}^{\frac{1}{2}} \frac{w_\alpha}{\|w\|} x^\alpha.$$

Equating corresponding coefficients we get

$$\begin{cases} b_0^d = \cos\big(\theta\|w\|\big) \\ b_0^{d-1} b_i = 0 \qquad \forall\ i = 1, \dots, n \end{cases}.$$

Now two different cases can occur:

- if $b_0 \neq 0$, the second equation above implies that $b_i = 0$ for $i = 1, \dots, n$ and $b = (b_0, 0, \dots, 0)$. Since $b \in S^n$ it follows that $b_0 = \pm 1$. If $b_0 = 1$, then the meeting point is again x_0^d and γ_w comes back to it for $\theta = \frac{2\pi}{\|w\|}$, and the same happens if $b_0 = -1$ and d is even. If $b_0 = -1$ and d is odd, then the meeting point corresponds to $-x_0^d$, and the meeting time is $\theta = \frac{\pi}{\|w\|}$.
- if $b_0 = 0$ we get $\cos(\theta\|w\|) = 0$ and γ_w may meet $\mathcal{S}V_{n,d}$ at $(b_0 x_0 + \cdots + b_n x_n)^d$ for $\theta = \frac{\pi}{2\|w\|}$ or $\theta = \frac{3\pi}{2\|w\|}$.

Now that we know when the curve might intersect again $\mathcal{S}V_{n,d}$, we need to understand when it does that orthogonally. Notice that up to now we have used only some of the equations arising from (3.22). We will now use the others to impose the orthogonality condition.
We look at the case $b_0 = 0$ and $\theta = \frac{\pi}{2\|w\|}$. If we could find some $b = (0, b_1, \dots, b_n) \in S^n$ and $w \in N_{x_0^d}\mathcal{S}V_{n,d}$ such that γ_w meets $\mathcal{S}V_{n,d}$ orthogonally at $(b_0 x_0 + \cdots + b_n x_n)^d$ for $\theta = \frac{\pi}{2\|w\|}$, then by the previous computation no other curve satisfying the conditions in (3.19) could have length less than this one. Fix the following notations: $\tilde{b} := (b_1, \dots, b_n)$, $\tilde{x} := (x_1, \dots, x_n)$ and $\tilde{\alpha} := (\alpha_1, \dots, \alpha_n)$, i.e. we neglect the first entry of the corresponding vectors. We have $\|\tilde{b}\| = 1$ and for $\theta = \frac{\pi}{2\|w\|}$ equation (3.22) becomes

$$\sum_{\alpha_1+\cdots+\alpha_n=d} \binom{d}{\tilde{\alpha}} \tilde{b}^{\tilde{\alpha}} \tilde{x}^{\tilde{\alpha}} = \sum_{\alpha_0<d-1} \binom{d}{\alpha}^{\frac{1}{2}} \frac{w_\alpha}{\|w\|} x^\alpha. \tag{3.23}$$

Equating corresponding coefficients, we get $w_\alpha = 0$ for each $\alpha = (\alpha_0, \dots, \alpha_n)$ such that $\alpha_0 \neq 0$, while for the other multi–indices we get $\binom{d}{\tilde{\alpha}}^{\frac{1}{2}} \tilde{b}^{\tilde{\alpha}} = \frac{w_{\tilde{\alpha}}}{\|w\|}$. These equations admit a solution and therefore a curve γ_w with the properties described above exists.
Since we are assuming that $\gamma_w\big(\frac{\pi}{2\|w\|}\big) = (b_1 x_1 + \cdots + b_n x_n)^d$, by Lemma 3.20 we have

$$T_{\gamma_w\left(\frac{\pi}{2\|w\|}\right)} \mathcal{S}V_{n,d} = \left\langle \tilde{b}^{d-1}\lambda \mid \lambda \text{ is a linear form orthogonal to } \tilde{b} \right\rangle.$$

The tangent vector to the curve γ_w at the meeting point is given by

$$\dot{\gamma}_w\left(\frac{\pi}{2\|w\|}\right) = -x_0^d\, \|w\|.$$

Since the monomials of the Bombieri–Weyl basis are mutually orthogonal and x_0^d is not among those spanning the tangent space at the meeting point, γ_w comes back to $\mathcal{S}V_{n,d}$ orthogonally for any choice of $w \in N_{x_0^d}\mathcal{S}V_{n,d}$ and $b \in S^n$ satisfying the conditions given by (3.23). Moreover, the length of such a curve is independent of the specific w and always equal to $\frac{\pi}{2}$. Hence we get that $\mathrm{bw}(\mathcal{S}V_{n,d}) = \frac{\pi}{4}$.
Noticing that $\frac{1}{\sqrt{2}}\sqrt{1+\frac{1}{d-1}} < \frac{\pi}{4}$ as soon as $d \geq 6$, the proof is complete. □

Remark that the same results hold true for $V_{n,d} \hookrightarrow S^N$. Before moving on, let us comment on the meaning of the quantity $\mathrm{bw}(V_{n,d})$ in our context. First of all, [40] ensures that the best rank–one approximation of a symmetric tensor is still symmetric. Recall once again our interpretation of $V_{n,d} \hookrightarrow S^N$ as the set of symmetric rank–one tensors among norm–1 ones. Since $V_{n,d}$ is compact and in particular closed, for every point in S^N there will be at least one point in $V_{n,d}$ minimizing the distance between the given point and $V_{n,d}$. The point realizing the minimum need not be unique and indeed in general it is not. From the tensor perspective, for a fixed p, every distance minimizing point in $V_{n,d}$ provides a best rank–one approximation of the tensor represented by p. Since S^N is compact, it is geodesically complete, and therefore there exists distance minimizing geodesics joining p with each of its best rank–one approximations and each of the geodesics meet $V_{n,d}$ orthogonally. From this observation we obtain that a symmetric norm–1 tensor admits more than one best rank–one approximation if and only if the corresponding point in S^N admits more than one distance minimizing geodesic orthogonal to $V_{n,d}$. It follows that the injectivity of the normal exponential map on the ε–small normal bundle $N^{\varepsilon}V_{n,d}$ ensures that every tensor represented by a point in the image admits a unique best rank–one approximation, as it will be joined to $V_{n,d}$ by a unique distance minimizing orthogonal geodesic. Given this, we can restate the result of Theorem 3.17 about $\mathrm{bw}(V_{n,d})$ in the following way.

Corollary 3.19. *Every symmetric tensor p at distance less than $\sin(\rho(V_{n,d}))\|p\|_{\mathrm{BW}}$ from rank–one admits a unique best rank–one approximation.*

A consequence of this result is that the probability that a symmetric tensor admits a unique best rank–one approximation is bounded below by the normalized volume of the tubular neighborhood of radius $\rho_{n,d} := \rho(V_{n,d})$ on the sphere, that is to say,

$$\mathbb{P}\big(\text{symmetric norm–1 tensor s.t. } \exists! \text{ best rank–one approximation}\big) \geq \frac{\mathrm{Vol}\big(\mathcal{U}(V_{n,d}, \rho_{n,d})\big)}{\mathrm{Vol}(S^N)}.$$

We stress again that the reach computed in Theorem 3.17 gives a lower bound to the radii of validity of the formula for the volume of a tubular neighborhood of $V_{n,d}$ that we will compute in Section 3.1.5: we are guaranteed that it gives the correct result for any radius less than $\rho(\mathcal{S}V_{n,d})$.

3.1.5 The volume of tubular neighborhoods

The second fundamental form of the spherical Veronese surface

In this section we will carry out the computation of the second fundamental form of $\mathcal{S}V_{n,d}$, which is one of the ingredients needed to apply Weyl's tube formula.

By Remark 3.14, since we have an isometric transitive action of $O(n+1)$ on $\mathcal{S}V_{n,d}$ by restrictions of isometries of S^N, the extrinsic geometry of $\mathcal{S}V_{n,d} \hookrightarrow S^N$ is invariant under this action. Therefore it is enough compute the second fundamental form at a specific

point of $\mathcal{SV}_{n,d}$ to automatically know it at every other point. Hence we carry out the computation using the point $x_0^d \in \mathcal{SV}_{n,d}$ for simplicity.

By Remark 2.1 to compute the second fundamental form of $\mathcal{SV}_{n,d}$ at x_0^d along a normal direction $\eta \in N_{x_0^d}\mathcal{SV}_{n,d}$ it is enough to choose a local parametrization around x_0^d, compute its second derivatives and take their (Bombieri–Weyl) scalar product in $\mathbb{R}^{N+1}$ with η. This way we will obtain the matrix representing the Weingarten operator at x_0^d with respect to the basis of $T_{x_0^d}\mathcal{SV}_{n,d}$ given by the derivatives of the chosen parametrization. In order to compute such a parametrization, consider the projection on S^n from the tangent plane at $(1, 0 \dots, 0)$, giving a parametrization of the upper hemisphere of S^n. Composing this projection with the Veronese map $\widehat{\nu}_{n,d}$ we obtain a parametrization $\varphi_{n,d}$ of the part of $\mathcal{SV}_{n,d}$ contained in the upper hemisphere of S^N. This is given explicitly by

$$\begin{aligned} \varphi_{n.d} : \ \mathbb{R}^n &\longrightarrow U \subseteq \mathcal{SV}_{n,d} \\ a = (a_1, \dots, a_n) &\longmapsto \left(\frac{x_0 + a_1 x_1 + \cdots + a_n x_n}{(1 + \|a\|^2)^{\frac{1}{2}}} \right)^d , \end{aligned} \tag{3.24}$$

where $\|a\|$ is the Euclidean norm of a in $\mathbb{R}^n$. Since $\varphi_{n,d}^{-1}(x_0^d) = (0, \dots, 0)$, we have to compute the first and second derivatives of $\varphi_{n,d}$ at the origin. We obtain the following expressions

$$\begin{aligned} \frac{\partial \varphi_{n,d}}{\partial a_i}(a)\bigg|_{a=0} &= d\, x_0^{d-1} x_i, \\ \frac{\partial^2 \varphi_{n,d}}{\partial a_i \partial a_j}(a)\bigg|_{a=0} &= -\delta_{ij}(d\, x_0^d) + d(d-1)\, x_0^{d-2} x_i x_j, \end{aligned} \tag{3.25}$$

where $\delta_{ij} = 1$ if $i = j$ and 0 otherwise.

For $i = 1, \dots, n$ denote by $e_i = \sqrt{d}\, x_0^{d-1} x_i$ the orthonormal vectors in the Bombieri–Weyl basis with exponent $d-1$ on x_0. By (3.25) the basis of $T_{x_0^d}\mathcal{SV}_{n,d}$ given by the first derivatives of the parametrization is $\{\sqrt{d}\, e_i \mid i = 1, \dots, n\}$. Instead of using this basis, we compute the matrix representing the Weingarten operator with respect to the Bombieri–Weyl orthonormal basis $\{e_i\}_{i=1,\dots,n}$ along a normal direction $\eta \in N_{x_0^d}\mathcal{SV}_{n,d}$. Denoting by $L_\eta = (L_{\eta,ij})_{i,j=1,\dots,n}$ this matrix, by (2.1) we have

$$\begin{aligned} L_{\eta,ij} = H_\eta(e_i, e_j) \ &= \ \frac{1}{d}\, H_\eta\left(\sqrt{d}\, e_i, \sqrt{d}\, e_j\right) = \frac{1}{d}\, H_\eta\left(\frac{\partial \varphi_{n,d}}{\partial a_i}(a)\bigg|_{a=0}, \frac{\partial \varphi_{n,d}}{\partial a_j}(a)\bigg|_{a=0} \right) \\ &= \ \frac{1}{d} \left\langle \frac{\partial^2 \varphi_{n,d}}{\partial a_i \partial a_j}(a)\bigg|_{a=0}, \eta \right\rangle_{\mathbb{R}^{N+1}} = \frac{1}{d} \left\langle -\delta_{ij}(d\, x_0^d) + d(d-1)\, x_0^{d-2} x_i x_j, \eta \right\rangle_{\mathbb{R}^{N+1}} . \end{aligned} \tag{3.26}$$

Recalling that the normal space $N_{x_0^d}\mathcal{SV}_{n,d}$ is given by (3.21), we expand the normal direction as

$$\eta = \sum_{\alpha_0 \le d-2} \eta_\alpha \binom{d}{\alpha}^{\frac{1}{2}} x^\alpha .$$

Since everything is expressed in terms of an orthonormal basis, from (3.26),we obtain

$$L_{\eta,ii} = \sqrt{2\left(\frac{d-1}{d}\right)}\, \eta_{d-2,0,\dots,2,\dots,0} \qquad \text{for } i = 1, \dots, n, \tag{3.27}$$

$$L_{\eta,ij} = \sqrt{\frac{d-1}{d}}\, \eta_{d-2,\dots,1,\dots,1,\dots,0} \qquad \text{for } i, j = 1, \dots, n,\ i \neq j. \tag{3.28}$$

Consider the following decomposition of the normal space $N_{x_0^d}\mathcal{S}V_{n,d}$ as the orthogonal direct sum

$$N_{x_0^d}\mathcal{S}V_{n,d} = \left\langle \left\{ \binom{d}{\alpha}^{\frac{1}{2}} x_0^{d-2} x_i x_j \mid i,j = 1,\dots,n \right\} \right\rangle \oplus \left\langle \left\{ \binom{d}{\alpha}^{\frac{1}{2}} x^\alpha \mid \alpha_0 < d-2 \right\} \right\rangle =: W \oplus P, \tag{3.29}$$

according to the exponent corresponding to x_0. We define a map from the subspace W to the space of homogeneous quadratic polynomials in n variables $\mathbb{R}[x_1,\dots,x_n]_{(2)}$ by defining it on a basis of W as

$$\binom{d}{\alpha}^{\frac{1}{2}} x_0^{d-2} x_i x_j \longmapsto \binom{2}{(\alpha_i,\alpha_j)}^{\frac{1}{2}} x_i x_j$$

and extending by linearity. Since an orthonormal basis for the induced Bombieri–Weyl product on W is mapped to an orthonormal basis of $\mathbb{R}[x_1,\dots,x_n]_{(2)}$ with its own Bombieri–Weyl product, this defines a linear isometry between those spaces. Composing this with the inverse ϕ^{-1} of the isomorphism we described in (3.10), we can define a linear isometry of W with the space of symmetric matrices $\mathrm{Sym}(n,\mathbb{R})$ with the Frobenius scalar product. This isometry also yields a correspondence between the associated Gaussian probability distributions. A direct consequence of this construction, the description of $\mathrm{GOE}(n)$ matrices given in Section 3.1.1 and formulas (3.27) and (3.28), is the following theorem.

Theorem 3.20. *Consider the decomposition of the normal space $N_{x_0^d}\mathcal{S}V_{n,d} = W \oplus P$ given in (3.29). Then the following statements hold:*

1. *$L_\eta = 0$ for every $\eta \in P$;*
2. *if we pick a random vector $\eta \in W$, Gaussian w.r.t the Bombieri–Weyl metric, then the Weingarten operator at x_0^d along η is distributed, as a random matrix, as $L_\eta \sim \sqrt{2}\left(\frac{d-1}{d}\right)^{\frac{1}{2}} \mathrm{GOE}(n)$, i.e. as a multiple of a $\mathrm{GOE}(n)$ matrix.*

Corollary 3.21 (The normal bundle splitting)**.** *Let $p \in \mathcal{S}V_{n,d}$ and denote by L_η the symmetric matrix representing second fundamental form of $\mathcal{S}V_{n,d} \hookrightarrow S^N$ along a normal direction $\eta \in N_p\mathcal{S}V_{n,d}$. There exists an orthogonal decomposition $N_p\mathcal{S}V_{n,d} = W_p \oplus P_p$ such that the following statements hold:*

1. *$L_\eta = 0$ for every $\eta \in P_p$;*
2. *W_p with its induced Bombieri–Weyl metric is isometric to $\mathrm{Sym}(n)$ with the Frobenius one. Moreover, if we pick $\eta \in W_p$ Gaussian w.r.t the Bombieri–Weyl metric, then the Weingarten operator at p along η is distributed, as a random matrix, as $L_\eta \sim \sqrt{2}\left(\frac{d-1}{d}\right)^{\frac{1}{2}} \mathrm{GOE}(n)$, i.e. as a multiple of a $\mathrm{GOE}(n)$ matrix.*

Proof. The corollary follows immediately from Theorem 3.20 using the fact that for every $p \in \mathcal{S}V_{n,d}$ there exists an isometry of S^N preserving $\mathcal{S}V_{n,d}$ and sending p to x_0^d. □

Remark that Corollary 3.21 holds trivially also for $V_{n,d}$. Theorem 3.20 and Corollary 3.21 thus give a full description of the extrinsic geometry of $\mathcal{S}V_{n,d} \hookrightarrow S^N$ in terms of random matrices. This is crucial from the computational point of view, as it allows reducing integrals on the normal bundle of $\mathcal{S}V_{n,d}$ of quantities related to the second fundamental form to expected values of quantities related to $\mathrm{GOE}(n)$ matrices. In the next section, we will use this fact to explicitly compute the integrals appearing in Weyl's tube formula (2.8). This will lead to explicit expressions of the curvature coefficients of the embedding $V_{n,d} \hookrightarrow S^N$.

The curvature coefficients

All this section will be dedicated to proving the following result.

Theorem 3.22 (The volume of a neighborhood)**.** *Let $V_{n,d} \hookrightarrow S^N$ be the spherical Veronese variety and $\mathcal{U}(V_{n,d}, \varepsilon)$ be its tubular neighborhood of radius ε as described in (2.2). If $\varepsilon < \rho(V_{n,d})$, the following formula holds:*

$$\mathrm{Vol}(\mathcal{U}(V_{n,d},\varepsilon)) = \sum_{\substack{0\le j\le n\\ j \text{ even}}} (-1)^{\frac{j}{2}} d^{\frac{n}{2}} \left(\frac{d-1}{d}\right)^{\frac{j}{2}} \frac{2^{n+2-j}\pi^{\frac{N}{2}}\Gamma\left(\frac{n}{2}+1\right)}{\Gamma\left(\frac{j}{2}+1\right)\Gamma(n+1-j)\Gamma\left(\frac{N+j-n}{2}\right)} J_{N,N-n+j}(\varepsilon), \tag{3.30}$$

where for $0 \le j \le n$, j even, the functions $J_{N,N-n+j}$ are the universal functions given by (2.7) and Γ is the standard Gamma function.

Comparing the expression (3.30) to Weyl's tube formula (2.5), we immediately obtain the following corollary.

Corollary 3.23 (The curvature coefficients)**.** *The curvature coefficients of the spherical Veronese variety $V_{n,d} \hookrightarrow S^N$ are as follows:*

$$K_{N-n+j}(V_{n,d}) = (-1)^{\frac{j}{2}} d^{\frac{n}{2}} \left(\frac{d-1}{d}\right)^{\frac{j}{2}} \frac{2^{n+2-j}\pi^{\frac{N}{2}}\Gamma\left(\frac{n}{2}+1\right)}{\Gamma\left(\frac{j}{2}+1\right)\Gamma(n+1-j)\Gamma\left(\frac{N+j-n}{2}\right)},$$

for $0 \le j \le n$, j even, and $K_{N-n+j}(V_{n,d}) = 0$ otherwise.

We will prove Theorem 3.22 using Weyl's tube formula (2.8) applied to the spherical Veronese surface $\mathcal{S}V_{n,d} \hookrightarrow S^N$. Before diving into the computation, we will prove a useful lemma about integration on spheres. It will be needed to exploit Theorem 3.20 when dealing with the integrals on the normal bundle of $\mathcal{S}V_{n,d}$ arising from the tube formula.

Consider a sphere S^m for some $m \ge 2$ and fix some $k \in \{1, \dots, m-1\}$. Denote by $\iota : S^k \hookrightarrow \mathbb{R}^{k+1}$ the standard embedding and consider the map

$$\begin{array}{ccc} S^k \times \overset{\circ}{D}{}^{m-k} & \xrightarrow{\varphi} & S^m \subseteq \mathbb{R}^{m+1} = \mathbb{R}^{k+1} \times \mathbb{R}^{m-k} \\ (\sigma, z) & \longmapsto & (\sqrt{1-\|z\|^2}\,\iota(\sigma),\ z), \end{array} \tag{3.31}$$

where $\overset{\circ}{D}{}^{m-k}$ is the open unit disk in $\mathbb{R}^{m-k}$. This map gives a smooth parametrization of $S^m \setminus \left\{\{0\} \times S^{m-k-1}\right\} \subseteq \mathbb{R}^{k+1} \times \mathbb{R}^{m-k}$. For every $l \in \mathbb{N}$, consider $\mathbb{R}^l$ endowed with a non–degenerate scalar product and coordinate functions $x_1, \dots, x_l$ with respect to an orthonormal basis. Then we have the standard volume form $\mathrm{vol}_{\mathbb{R}^l} = dx_1 \wedge \cdots \wedge dx_l$ which induces a volume form $\mathrm{vol}_{\overset{\circ}{D}{}^l}$ on the open norm–1 disc $\overset{\circ}{D}{}^l$ and also a volume form $\mathrm{vol}_{S^{l-1}}$ on S^{l-1} through the pullback of the inclusion (these need not necessarily be the Euclidean ones). Notice that with respect to the volume form vol_{S^m}, the part of S^m not parametrized by φ in (3.31) has measure 0. We have the following result about integration with respect to these volume forms.

Lemma 3.24. *With the same notations above, the pullback of the volume form vol_{S^m} through φ is given by*

$$\varphi^*(\mathrm{vol}_{S^m}) = \left(1-\|z\|^2\right)^{\frac{k-1}{2}}\ \mathrm{vol}_{S^k} \wedge \mathrm{vol}_{\overset{\circ}{D}{}^{m-k}}.$$

In particular, this implies that for any measurable function f on S^m the following formula holds

$$\int_{S^m} f(p)\ \mathrm{vol}_{S^m} = \int_{S^k}\int_{\overset{\circ}{D}{}^{m-k}} f\big(\sqrt{1-\|z\|^2}\ \iota(\sigma),\ z\big)\big(1-\|z\|^2\big)^{\frac{k-1}{2}}\ \mathrm{vol}_{S^k}\wedge \mathrm{vol}_{\overset{\circ}{D}{}^{m-k}}.$$

Proof. By the usual formula for the pullback of a differential form through a diffeomorphism, this can be computed via

$$\varphi^*(\mathrm{vol}_{S^m}) = |\det(J\varphi^\top\cdot J\varphi)|^{\frac12}\ \mathrm{vol}_{\overset{\circ}{D}{}^k}\wedge \mathrm{vol}_{\overset{\circ}{D}{}^{m-k}},$$

where $J\varphi$ denotes the $(m+1)\times m$ Jacobian matrix of φ and $J\varphi^\top$ is its transpose. Denote also by $J\iota$ the Jacobian matrix of the inclusion $\iota: S^k\hookrightarrow\mathbb{R}^{k+1}$. Then the Jacobian $J\varphi$ at a point (σ,z) has the block structure

$$(J\varphi)(\sigma,z) = \left(\begin{array}{c|c} \sqrt{1-\|z\|^2}\,(J\iota)(\sigma) & \left(\frac{-z_j}{\sqrt{1-\|z\|^2}}\,\iota(\sigma)\right) \\ \hline 0 & I_{m-k}\end{array}\right),$$

where I_{m-k} denotes the $(m-k)\times(m-k)$ identity matrix. Since $(J\iota)(\sigma)^\top\cdot\iota(\sigma) = \iota(\sigma)\cdot(J\iota)(\sigma)=0$ and $\iota(\sigma)^\top\cdot\iota(\sigma)=1$, we obtain the block diagonal matrix

$$(J\varphi^\top\cdot J\varphi)(\sigma,z) = \left(\begin{array}{c|c} (1-\|z\|^2)\,(J\iota^\top\cdot J\iota)(\sigma) & 0 \\ \hline 0 & \left(I_{m-k}+\frac{z\cdot z^\top}{1-\|z\|^2}\right)\end{array}\right).$$

For every $z\in\mathbb{R}^{m-k}$ consider an orthogonal martrix $R\in O(m-k)$ such that $z = R\,e_1|z|$, where $e_1=(1,0,\dots,0)$ is the first standard basis vector. Then we can compute the determinant of the lower right block in the above matrix as

$$\det\left(I_{m-k}+\frac{z\cdot z^\top}{1-\|z\|^2}\right) = \det\left(R\left(I_{m-k}+\frac{\|z\|^2}{1-\|z\|^2}E_{11}\right)R^\top\right) \ = \ \frac{1}{1-\|z\|^2},$$

where $E_{11}=e_1\,e_1^\top$ is the matrix with all zero entries except for the $(1,1)$–th one, which is equal to 1. Recalling that the determinant of a diagonal block matrix is given by the product of the determinants of its blocks, we find the following expression

$$\begin{aligned}|\det(J\varphi^\top\cdot J\varphi)(\sigma,z))| \ &= \ \big(1-\|z\|^2\big)^k\ |\det((J\iota^\top\cdot J\iota)(\sigma,z))|\ \frac{1}{1-\|z\|^2}\\ &= \big(1-\|z\|^2\big)^{k-1}\ |\det((J\iota^\top\cdot J\iota)(\sigma,z))|.\end{aligned}$$

Finally, applying again the formula for the pullback of a differential form, we can conclude that

$$\begin{aligned}\varphi^*(\mathrm{vol}_{S^m}) &= \big(1-\|z\|^2\big)^{\frac{k-1}{2}}\ |\det((J\iota^t\cdot J\iota)(\sigma,z))|^{\frac12}\ \mathrm{vol}_{\overset{\circ}{D}{}^k}\wedge\mathrm{vol}_{\overset{\circ}{D}{}^{m-k}}\\ &= \big(1-\|z\|^2\big)^{\frac{k-1}{2}}\ \mathrm{vol}_{S^k}\wedge\mathrm{vol}_{\overset{\circ}{D}{}^{m-k}}.\end{aligned}$$

□

Proof of Theorem 3.22. We start from Weyl's tube formula (2.8) applied to $\mathcal{SV}_{n,d}\hookrightarrow S^N$. As we already noticed, Remark 3.14 implies that the Weingarten operator looks the same at every point. It follows that in this case the integrand in Weyl's tube formula (2.8) does not depend on the point $p\in\mathcal{SV}_{n,d}$. Thus the outer integral will simplify to give the volume:

$$\mathrm{Vol}\left(\mathcal{U}(\mathcal{SV}_{n,d},\varepsilon)\right) = \mathrm{Vol}(\mathcal{SV}_{n,d})\int_{t=0}^{\tan\varepsilon}\int_{\eta\in S(N_{x_0^d}\mathcal{SV}_{n,d})}\frac{t^{N-n-1}\det(I_n-tL_\eta)}{(1+t^2)^{\frac{N+1}{2}}}\,d\eta\,dt, \quad (3.32)$$

where $N-n$ is the codimension of $\mathcal{S}V_{n,d} \hookrightarrow S^N$, $S(N_{x_0^d}\mathcal{S}V_{n,d})$ is the unit sphere in the normal space to $\mathcal{S}V_{n,d}$ at x_0^d and I_n is the $n \times n$ identity matrix. The quantity $\mathrm{Vol}(\mathcal{S}V_{n,d})$ is given by (3.15).

Given the decomposition of $N_{x_0^d}\mathcal{S}V_{n,d}$ described in Theorem 3.20, we have $S(N_{x_0^d}\mathcal{S}V_{n,d}) = S(W \oplus P)$. Observe that $\dim(S(N_{x_0^d}\mathcal{S}V_{n,d})) = N-n-1$ and $\dim(W) = \dim(\mathrm{Sym}(n,\mathbb{R})) = \frac{n(n+1)}{2}$. Notice that if $d=2$, there is no actual decomposition, since in that case we have $N_{x_0^d}\mathcal{S}V_{n,d} = W$ and $P = 0$. If instead $d > 2$ we parametrize $S(W \oplus P)$ as in (3.31), where the dimension parameter here are $m = N-n-1$ and $k = \frac{n(n+1)}{2} - 1$. With the same notation of section Lemma 3.24, if $\varphi(\sigma, z) = \eta \in S(W \oplus P)$ for some $\sigma \in S(W)$ and $z \in \mathring{D}(P)$, we have that $\sqrt{1-\|z\|^2}\,\iota(\sigma)$ will be the component of η along W, while z itself will be the component of η along P. Moreover, we use the linear isometry between W and $\mathrm{Sym}(n,\mathbb{R})$ to change variable from $\sigma \in S(W)$ to $Q \in S(\mathrm{Sym}(n,\mathbb{R})) = S^{\frac{n(n+1)}{2}-1}$.

It is clear by its very definition that the Weingarten operator is linear in the normal vector argument: given an isometric embedding $M \hookrightarrow \overline{M}$, for every $p \in M$, η, $\xi \in N_pM$ and a, $b \in \mathbb{R}$, we have

$$L_{a\eta+b\xi} = aL_\eta + bL_\xi.$$

It follows that given $\eta = \varphi(Q,z) \in S(W \oplus P)$, the Weingarten operator L_η behaves as

$$L_\eta = L_{\sqrt{1-\|z\|^2}\,Q+z} = \sqrt{1-\|z\|^2}\,L_Q + L_z = \sqrt{1-\|z\|^2}\,L_Q. \tag{3.33}$$

Applying Lemma 3.24 to (3.32) and using (3.33) the integral becomes

$$\begin{aligned}\mathrm{Vol}\left(\mathcal{U}(\mathcal{S}V_{n,d},\varepsilon)\right) =& \mathrm{Vol}(\mathcal{S}V_{n,d}) \int_{t=0}^{\tan\varepsilon} \int_{S^{\frac{n(n+1)}{2}-1}} \int_{D^{N-n-\frac{n(n+1)}{2}}} \left[\frac{t^{N-n-1}}{(1+t^2)^{\frac{N+1}{2}}} \times\right. \\ & \left. \times \det(I_n - t\sqrt{1-\|z\|^2}\,L_Q)(1-\|z\|^2)^{\frac{n(n+1)}{4}-1}\right] dz\; dS(Q)\; dt,\end{aligned} \tag{3.34}$$

where dz is a short notation for $\mathrm{vol}_{D^{N-n-\frac{n(n+1)}{2}}}$ and $dS(Q)$ is a short notation for $\mathrm{vol}_{S^{\frac{n(n+1)}{2}-1}}$, with the convention that for $d=2$ the integral over D^0 is set to 1.

The only non-explicit term in (3.34) is the one involving the determinant. Recall that by Theorem 3.20, if $Q \in \mathrm{Sym}(n,\mathbb{R})$ is a random $\mathrm{GOE}(n)$ matrix, then L_Q is a random matrix distributed as $\sqrt{2}\left(\frac{d-1}{d}\right)^{\frac{1}{2}}\mathrm{GOE}(n)$. Set now $\tau := t\sqrt{2}\left(\frac{d-1}{d}\right)^{\frac{1}{2}}$. We can expand the determinant as

$$\det\left(I_n - \tau\sqrt{1-\|z\|^2}Q\right) = \sum_{j=0}^{n}(-1)^j\,\tau^j\,(1-\|z\|^2)^{\frac{j}{2}}\,g_j(Q), \tag{3.35}$$

where $g_j(Q)$ are homogeneous polynomials of degree j in the coefficients of Q for $j = 1,\dots,n$ and $g_0(Q) = 1$. Substituting the expansion (3.35) into (3.34), we can split the integral into a product of three terms and obtain

$$\begin{aligned}\mathrm{Vol}\left(\mathcal{U}(\mathcal{S}V_{n,d},\varepsilon)\right) =& \mathrm{Vol}(\mathcal{S}V_{n,d}) \sum_{j=0}^{n}(-1)^j\,2^{\frac{j}{2}}\left(\frac{d-1}{d}\right)^{\frac{j}{2}}\left(\int_0^{\tan\varepsilon}\frac{t^{N-n-1+j}}{(1+t^2)^{\frac{N+1}{2}}}\,dt\right)\times \\ & \times\left(\int_{D^{N-n-\frac{n(n+1)}{2}}}(1-\|z\|^2)^{\frac{n(n+1)}{4}-1+\frac{j}{2}}\,dz\right)\times \\ & \times\left(\int_{S^{\frac{n(n+1)}{2}-1}} g_j(Q)\,dS(Q)\right).\end{aligned} \tag{3.36}$$

The first of the three integrals is the integral of a rational function in t, while the second one is a "polynomial" in $\|z\|$. It remains to study the third one.

Since g_j are homogeneous polynomials, we have $g_j(Q) = \|Q\|^j\, g_j\left(\frac{Q}{\|Q\|}\right)$. Recalling the density (3.9) of a GOE(n) matrix we have

$$\begin{aligned} \mathop{\mathbb{E}}_{Q\in\mathrm{GOE}(n)} g_j(Q) &= \frac{1}{(2\pi)^{\frac{n(n+1)}{4}}} \int_{\mathrm{Sym}(n,\mathbb{R})} \|Q\|^j\, g_j\left(\frac{Q}{\|Q\|}\right) e^{-\frac{\|Q\|^2}{2}}\, dQ \;= \\ &= \frac{1}{(2\pi)^{\frac{n(n+1)}{4}}} \left(\int_0^{+\infty} \rho^{\frac{n(n+1)}{2}-1+j} e^{-\frac{\rho^2}{2}}\, d\rho\right) \left(\int_{S^{\frac{n(n+1)}{2}-1}} g_j(\tilde{Q})\; dS(\tilde{Q})\right), \end{aligned} \tag{3.37}$$

where $\mathbb{E}$ denotes the expectation of a random variable and in the last equality we switched to polar coordinates in order to decouple the integral further. Inverting (3.37) we obtain

$$\int_{S^{\frac{n(n+1)}{2}-1}} g_j(Q)\; dS(Q) \;=\; \frac{\mathop{\mathbb{E}}_{Q\in\mathrm{GOE}(n)}[g_j(Q)]\; (2\pi)^{\frac{n(n+1)}{4}}}{\int_0^{+\infty} \rho^{\frac{n(n+1)}{2}-1+j} e^{-\frac{\rho^2}{2}}\, d\rho}. \tag{3.38}$$

By the linearity of the expectation and the expansion $\det(I_n - \lambda Q) = \sum_{j=0}^{n} (-1)^j\, \lambda^j\, g_j(Q)$, to compute the expectation of $g_j(Q)$ it is enough to compute that of $\det(I_n - \lambda Q)$ for $Q \in \mathrm{GOE}(n)$ and look at the homogeneous part of degree j in λ. This procedure gives us the explicit expression for (3.38)

$$\int_{S^{\frac{n(n+1)}{2}-1}} g_j(Q)\; dS(Q) = \frac{(-1)^{\frac{j}{2}} (2\pi)^{\frac{n(n+1)}{4}}\, \frac{j!}{(\frac{j}{2})!} \binom{n}{j}}{2^j \int_0^{+\infty} \rho^{\frac{n(n+1)}{2}-1+j} e^{-\frac{\rho^2}{2}}\, d\rho} \quad \text{if } 0 \le j \le n,\; j \text{ even} \tag{3.39}$$

and 0 otherwise, see below for a proof of this result. By standard computations involving Gamma and Beta functions, one can show that the following identities hold

$$\int_0^{+\infty} \rho^{\frac{n(n+1)}{2}+j-1} e^{-\frac{\rho^2}{2}}\, d\rho \;=\; 2^{\frac{n(n+1)}{4}+\frac{j}{2}-1}\, \Gamma\left(\frac{1}{4}(n^2+n+2j)\right), \tag{3.40}$$

$$\int_{D^{N-n-\frac{n(n+1)}{2}}} \left(1-\|z\|^2\right)^{\frac{n(n+1)}{4}-1+\frac{j}{2}}\; dz \;=\; \pi^{\frac{2N-n^2-3n}{4}}\, \frac{\Gamma\left(\frac{1}{4}(n^2+n+2j)\right)}{\Gamma\left(\frac{1}{2}(N-n+j)\right)}, \tag{3.41}$$

where we notice that for $d = 2$ (3.41) gives 1, agreeing with our convention for the integral over D^0. Substituting (3.39), (3.40) and (3.41) into (3.36) and using the Legendre duplication formula for the Gamma function, we obtain the explicit expression of $\mathrm{Vol}\left(\mathcal{U}(\mathcal{S}V_{n,d}, \varepsilon)\right)$. Finally, recalling that $V_{n,d} = \mathcal{S}V_{n,d} \cup -\mathcal{S}V_{n,d}$ and using formula (3.15) to express $\mathrm{Vol}(\mathcal{S}V_{n,d})$, the proof of Theorem 3.22 is complete. □

Proof of (3.39). By the linearity of the expectation, in order to prove formula (3.39) all we have to do is computing

$$\mathop{\mathbb{E}}_{Q\in\mathrm{GOE}(n)}[\det(I_n - \lambda Q)] = \sum_{j=0}^{n} (-1)^j \lambda^j \mathop{\mathbb{E}}_{Q\in\mathrm{GOE}(n)}[g_j(Q)],$$

since the expectation of $g_j(Q)$ can then be deduced by looking at the degree j coefficient in above polynomial expression in λ. This can be computed immediately using [69, Eq. (2.2.38)], but we also give a simple derivation of this computation for the sake of completeness.

First, we write the determinant according to its very definition as

$$\det(I_n - \lambda Q) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} \left(\delta_{i\sigma(i)} - \lambda Q_{i\sigma(i)}\right),$$

where S_n is the group of permutations on $\{1, \dots, n\}$, δ_{ij} is the Kronecker delta and $\operatorname{sgn}(\sigma)$ is the signature of a permutation $\sigma \in S_n$, that is $\operatorname{sgn}(\sigma) = 1$ if the permutation is even, while it is -1 if odd.

Recall that for $Q \in \mathrm{GOE}(n)$ we have $Q_{ii} \sim \mathcal{N}(0,1)$ and $Q_{ij} \sim \mathcal{N}(0, \frac{1}{2})$ for $i \neq j$ and, apart from the obvious symmetry conditions, the entries are independent. By linearity

$$\mathop{\mathbb{E}}_{Q \in \mathrm{GOE}(n)}[\det(I_n - \lambda Q)] \;=\; \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \mathop{\mathbb{E}}_{Q \in \mathrm{GOE}(n)}\left[\prod_{i=1}^{n} \left(\delta_{i\sigma(i)} - \lambda Q_{i\sigma(i)}\right)\right]. \tag{3.42}$$

We now analyze the contributions of each permutation. Given $\sigma \in S_n$, suppose that it contains a cycle of length at least 3, i.e. there exists $i \in \{1, \dots, n\}$ such that $\sigma(i) \neq i$ and $\sigma^2(i) \neq i$. Then, by the independence of the entries, in the term of (3.42) corresponding to such σ, we can split the expectation into a product of expectations, separating the term corresponding to such i. Since $\delta_{i\sigma(i)} = 0$ and $Q_{i\sigma(i)}$ is a centered, this expectation is 0 and σ gives no contribution to the sum in (3.42). It follows that the only permutations contributing to (3.42) are those formed by disjoint transpositions and fixed points only.

Suppose now we have a $\sigma \in S_n$ containing only disjoint transpositions and fixed points. Denote by $\operatorname{fix}(\sigma) = \{i \in \{1, \dots, n\} \mid \sigma(i) = i\}$ the set of fixed points of σ and by $s(\sigma)$ the number of disjoint transpositions in σ. Then we have

$$\begin{aligned}
\mathop{\mathbb{E}}_{Q \in \mathrm{GOE}(n)}\left[\prod_{i=1}^{n} \left(\delta_{i\sigma(i)} - \lambda Q_{i\sigma(i)}\right)\right] \;&=\; \mathbb{E}\left[\prod_{i \in \operatorname{fix}(\sigma)} (1 - \lambda \xi_i)\right] \cdot \mathbb{E}\left[\prod_{k=1}^{s(\sigma)} \frac{1}{2}\lambda^2 \gamma_k^2\right] \;= \\
&= \left(\prod_{i \in \operatorname{fix}(\sigma)} \mathbb{E}\big[(1 - \lambda \xi_i)\big]\right) \cdot \left(\prod_{k=1}^{s(\sigma)} \mathbb{E}\left[\frac{1}{2}\lambda^2 \gamma_k^2\right]\right),
\end{aligned}$$

where $\xi_i \sim \mathcal{N}(0,1)$ and $\gamma_k \sim \mathcal{N}(0,1)$. For these terms we have

$$\begin{aligned}
&\mathbb{E}\big[(1 - \lambda \xi_i)\big] = 1, \\
&\mathbb{E}\left[\frac{1}{2}\lambda^2 \gamma_k^2\right] = \frac{1}{2}\lambda^2 \mathbb{E}\big[\gamma_k^2\big] = \frac{1}{2}\lambda^2.
\end{aligned}$$

The contribution of such $\sigma \in S_n$ in the sum in (3.42) is therefore given by

$$\operatorname{sgn}(\sigma) \mathop{\mathbb{E}}_{Q \in \mathrm{GOE(n)}}\left[\prod_{i=1}^{n} \left(\delta_{i\sigma(i)} - \lambda Q_{i\sigma(i)}\right)\right] \;=\; \operatorname{sgn}(\sigma) \left(\frac{1}{2}\lambda^2\right)^{s(\sigma)}. \tag{3.43}$$

Notice that the contribution only depends on the number of transpositions $s(\sigma)$. To conclude the computation we, therefore, have to count how many permutations in S_n are given by exactly k disjoint transpositions for every $k = 0, \dots, \lfloor \frac{n}{2} \rfloor$. Denote this number by $N(k)$. To construct a permutation with exactly k disjoint transpositions we proceed as follows: choose two elements in $\{1, \dots, n\}$ forming the first transposition, then choose another 2 among the remaining ones to form the second transposition and so on until the k–th and last transposition is formed. Since the supports of the transpositions are disjoint,

the order in which they are picked is not relevant. It follows that the number $N(k)$ can be combinatorially computed as

$$N(k) = \frac{\binom{n}{2}\binom{n-2}{2}\cdots\binom{n-2k+2}{2}}{k!} = \frac{n!}{2^k(n-2k)!\,k!}.$$

Using this expression and (3.43) in the sum (3.42) we obtain for the expectation of the determinant the formula

$$\mathop{\mathbb{E}}_{Q\in\mathrm{GOE}(n)}[\det(I_n - \tau Q)] = \sum_{k=0}^{\lfloor \frac{n}{2}\rfloor} (-1)^k N(k) \left(\frac{1}{2}\lambda^2\right)^k = \sum_{k=0}^{\lfloor \frac{n}{2}\rfloor} (-1)^k \lambda^{2k} \frac{(2k)!}{2^{2k}k!}\binom{n}{2k}. \tag{3.44}$$

Since the expectation of $g_j(Q)$ is given by the degree j term in (3.44) multiplied by $(-1)^j$, we obtain

$$\mathop{\mathbb{E}}_{Q\in\mathrm{GOE}(n)}[g_j(Q)] = \begin{cases} 0 & \text{if } j \text{ odd} \\ \frac{(-1)^{\frac{j}{2}}}{2^j}\frac{j!}{(\frac{j}{2})!}\binom{n}{j} & \text{if } 0 \le j \le n,\ j \text{ even} \end{cases}. \tag{3.45}$$

Plugging (3.45) into (3.38), we get (3.39). The proof is concluded. □

Theorem 3.22 gives the exact value of the volume of the tubular neighborhood of the spherical Veronese variety, given that the radius is small enough. In Section 3.1.3 we have shown how our original question (3.5) could be restated in geometric terms, leading to (3.17), that expressed the probability of being close to rank–one as a normalized volume. We are therefore now able to give an answer to the original question: how likely is a symmetric tensor to be close to rank–one?

Theorem 3.25 (The probability of being close to rank–one)**.** *Let p be a random Bombieri–Weyl symmetric tensor of order d on $\mathbb{R}^{n+1}$ and $\mathcal{V}_{n,d} \subset \mathbb{R}[x_0,\dots,x_n]_{(d)} \simeq \mathbb{R}^{N+1}$ be the Veronese variety of rank–one tensors. For every δ such that $0 \le \arcsin(\delta) < \frac{1}{\sqrt{2}}\sqrt{1+\frac{1}{d-1}}$ we have*

$$\mathbb{P}\left\{p \in \mathbb{R}[x_0,\dots,x_n]_d \;\middle|\; d^{\mathrm{BW}}(p,\mathcal{V}_{n,d}) \le \delta\|p\|_{\mathrm{BW}}\right\} = \sum_{\substack{0\le j\le n \\ j \text{ even}}} (-1)^{\frac{j}{2}} d^{\frac{n}{2}} \left(\frac{d-1}{d}\right)^{\frac{j}{2}} 2^{n-j+1}\pi^{-\frac{1}{2}} \cdot \frac{\Gamma\left(\frac{n}{2}+1\right)\Gamma\left(\frac{N+1}{2}\right)}{\Gamma\left(\frac{j}{2}+1\right)\Gamma(n+1-j)\Gamma\left(\frac{N+j-n}{2}\right)} \cdot \int_0^{\frac{\delta}{\sqrt{1-\delta^2}}} \frac{t^{N-n+j-1}}{(1+t^2)^{\frac{N+1}{2}}}\mathrm{dt}.$$

We remark that the above expression, even if unpleasant, gives an exact formula for our wanted probability. In general, this probability shows an exponential decay in the codimension of $V_{n,d} \hookrightarrow S^N$ for any n, d.

In the case of rational normal curves, corresponding to $n = 1$, there is only one summand in the above theorem. This can be evaluated exactly, giving

$$\mathbb{P}\left\{p \in \mathbb{R}[x_0,x_1]_d \;\middle|\; d^{\mathrm{BW}}(p,\mathcal{V}_{1,d}) \le \delta\|p\|_{\mathrm{BW}}\right\} = \sqrt{d}\,\delta^{d-1}, \tag{3.46}$$

whenever $\arcsin(\delta) \leq \rho(V_{1,d})$. In fact in this case one has $N = d$ and, as it can be easily verified,

$$\int_0^{\frac{\delta}{\sqrt{1-\delta^2}}} \frac{t^{d-2}}{(1+t^2)^{\frac{d+1}{2}}}\,\mathrm{dt} = \frac{\delta^{d-1}}{d-1}.$$

Once substituted in Theorem 3.25 this immediately gives (3.46).

3.1.6 Applications

This last section is devoted to briefly discussing two settings in which some of the results from the earlier parts of Section 3.1 have interesting applications. In the first example we focus on random matrix theory, in particular on the study of the eigenvalues of random symmetric matrices. The second example will instead have a completely different flavour as we enter the quantum physics world. We will indeed see how our random symmetric tensors and the notion of tensor rank have close ties with the theory of quantum entanglement.

Eigenvalues of random symmetric matrices

When $d = 2$ the space of symmetric tensors $\mathcal{S}^2(n)$ can be identified, as we have already shown, with the space of $n \times n$ real symmetric matrices. The Bombieri–Weyl distribution corresponds, in this identification, with the Frobenius scalar product. In this case, by the Spectral Theorem, the set $\mathcal{S}_1^2(n)$ of rank–one tensors coincides with the set of rank–one matrices, i.e. matrices of the form $A = \pm vv^\top$, for some $v \in \mathbb{R}^n$.

Given a symmetric matrix Q, if we order the singular values of Q increasingly, $\sigma_1(Q) \leq \cdots \leq \sigma_n(Q)$, by the Eckart–Young Theorem 2.19, the distance in the Frobenius norm between Q and $\mathcal{S}_1^2(n)$ is given by

$$\widehat{\delta}(Q, \mathcal{S}_1^2(n)) = \left(\sum_{k=1}^{n-1} \sigma_k(Q)^2\right)^{\frac{1}{2}}.$$

In this case, using the fact that $\|Q\|_F^2 = \sigma_1(Q)^2 + \cdots + \sigma_n(Q)^2$, the probability in (3.5) equals

$$\mathbb{P}\left\{(1-\delta^2)\|Q\|_F^2 \leq \sigma_n(Q)^2\right\}, \tag{3.47}$$

therefore our question becomes of interest from the point of view of Random Matrix Theory. Indeed, we can apply Theorem 3.25 and obtain a closed formula for the probability in (3.47), for $\delta \leq 1/\sqrt{2}$. For instance:

$$\begin{aligned}\mathbb{P}\left\{Q \in \mathcal{S}^2(3) \,\middle|\, (1-\delta^2)\|Q\|_F^2 \leq \sigma_3(Q)^2\right\} &= \frac{1}{3\pi}\left(2\delta\sqrt{1-\delta^2}(-3+14\delta^2) + 6\arctan\left(\frac{\delta}{\sqrt{1-\delta^2}}\right)\right) \\ &= \frac{32\delta^3}{3\pi} - \frac{64\delta^5}{15\pi} + \mathcal{O}(\delta^6). \end{aligned} \tag{3.48}$$

In the context of Random Matrix Theory, which is concerned with the study of the distribution of the eigenvalues of random symmetric matrices, expressions like (3.48) are especially interesting. In fact, most of the theory is focused on asymptotic results with the size of the matrices that goes to infinity (see, for instance, [5, Chapter 3]) and very little is known on the distribution of the eigenvalues of random matrices of fixed (small) size. Still for symmetric matrices, Corollary 3.19 says that every symmetric matrix $Q \in \mathcal{S}^2(n)$ with

$$\|Q\|_F^2 < 2\sigma_n(Q)^2$$

has a unique best rank–one approximation.

Geometric measure of entanglement

Quantum entanglement is one of the main topics of research in quantum information theory, where it provides an important resource for quantum computing. Recall that a d-partite pure (symmetric) state of a quantum system can be regarded as a normalized (symmetric) tensor in the tensor product of d Hilbert spaces $\mathcal{H}_1, \dots, \mathcal{H}_d$

$$\mathcal{H} = \otimes_{k=1}^{d} \mathcal{H}_k \,.$$

A quantum state $|\phi\rangle \in \mathcal{H}$ is called *separable* if it is a product state, i.e.

$$|\phi\rangle = \otimes_{k=1}^{d} |\phi^{(k)}\rangle \,,$$

with $\||\phi^{(k)}\rangle\|_{\mathcal{H}_k} = 1$ for $k = 1, \dots, d$; it is called *entangled* if it is not separable. In the language of tensors, separable pure states correspond to rank–one tensors, while entangled states to tensors of higher rank.

A way to quantify the entanglement of a pure state, i.e. to measure how distant that state is from being separable, is the *geometric measure of entanglement*, first introduced in [79] (see also [53] for the specific case of symmetric states). For a pure symmetric state $|\psi\rangle$ this is defined as [52]

$$E_G(\,|\psi\rangle\,) := 1 - \max_{\substack{\Phi = |\phi\rangle^{\otimes d} \in \mathcal{H} \\ \||\phi\rangle\| = 1}} |\langle \psi | \Phi \rangle|. \tag{3.49}$$

When the tensor space is $\mathcal{H} = (\mathbb{R}^n)^{\otimes d}$ endowed with the Euclidean norm, pure symmetric states are represented by a tensor $T \in S^d(n)$ with Frobenius norm 1 and (3.49) becomes

$$E_G(T) = 1 - \max_{x \in S^{n-1}} |\langle T, x^{\otimes d} \rangle_F| \,. \tag{3.50}$$

In the tensor community, given a tensor $T \in S^d(n)$ and a unit vector $x \in S^{n-1}$ the quantity

$$Tx^d := \langle T, x^{\otimes d} \rangle_F$$

is known as the *Generalized Rayleigh Quotient (GQR)* [98]. The problem of maximizing the absolute value of the GRQ over S^{n-1} for a tensor T is equivalent to finding the best rank–one approximation of T. Indeed, given a maximizer $x_* \in S^{n-1}$ for the absolute value of the GQR, i.e.

$$|Tx_*^d| = \max_{x \in S^{n-1}} |Tx^d|,$$

one can show that $(Tx_*^d)\, x_*^{\otimes d}$ is a best rank–one approximation of T and

$$\widehat{\delta}^{\,2}(T, S_1^d(n)) = \|T\|_F^2 - (Tx_*^d)^2,$$

see [72, Theorem 2.19]. When T represents a pure symmetric state as above, by definition $\|T\|_F = 1$ and (3.50) becomes

$$E_G(T) = 1 - \sqrt{1 - \widehat{\delta}^{\,2}(T, S_1^d(n))} \,.$$

A bound on the distance of T from rank-one symmetric tensors therefore implies a bound on the geometric measure of entanglement of the pure symmetric state represented by T. In particular, we see that

$$\mathbb{P}\left\{T \in \mathcal{S}^d(n),\ \|T\|_F = 1 \,\middle|\, \widehat{\delta}(T, \mathcal{S}_1^d(n)) \leq \delta\right\} = \mathbb{P}\left\{E_G(T) \leq 1 - \sqrt{1-\delta^2}\right\},$$

where the left-hand side is equivalent to (3.5).

An alternative formulation of the geometric measure of entanglement sometimes used in quantum computing [46] is given as

$$\widetilde{E}_G(|\psi\rangle) := -\log_2 \left(\max_{\substack{\Phi = |\phi\rangle^{\otimes d} \in \mathcal{H} \\ \||\phi\rangle\| = 1}} |\langle \psi \,|\, \Phi \rangle|^2 \right). \tag{3.51}$$

With the same reasoning as before we obtain

$$\mathbb{P}\left\{T \in \mathcal{S}^d(n),\ \|T\|_F = 1 \,\middle|\, \operatorname{dist}_F(T, \mathcal{S}_1^d(n)) \leq \delta\right\} = \mathbb{P}\left\{\tilde{E}_G(T) \leq -\log_2(1-\delta^2)\right\}. \tag{3.52}$$

We refer the reader to [46] for a detailed discussion on how the knowledge of bounds such as (3.52) might be relevant for quantum computing techniques and algorithms.

3.2 Typical ranks of order–three tensors

The second part of this chapter will be focused on typical ranks of real tensors, with a specific attention to the case of order–three ones. In all this section the base field $\mathbb{F}$ is either $\mathbb{R}$ or $\mathbb{C}$; it is specified when a distinction is needed. Here it is enough for us to consider tensors as multi–dimensional arrays.

We begin this section by highlighting some major differences between matrix rank and its generalization to higher order tensor rank. Matrix rank, that is the minimum number of rank–one matrices that add up to the given one, has several nice and well understood properties. The maximal rank of a $m \times n$ matrix is always $\min\{m, n\}$ and the set of matrices with maximal rank is an open dense set of $\mathbb{F}^{m\times n}$. The rank can be computed by looking at the minors of the matrix; this way we also obtain the equations defining the sets of matrices with bounded rank, which are always Zariski closed sets. Moreover, the rank of a matrix with real entries is the same whether you consider it over $\mathbb{R}$ or over $\mathbb{C}$; for this reason all the above properties hold regardless of whether $\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$.

When we move on to consider the rank of a higher order tensor, that we recall is the minimum number of rank–one tensors that add up to the given one (see Definition 3.5), all the nice properties that we pointed out for matrix rank fail. First of all, the sets of tensors with bounded rank are not Zariski closed anymore; to obtain closed sets one should look at a slightly more sofisticated notion called *border rank*. Even in this case, the algebraic equations defining these sets are not known in general. The maximal rank for a given format, already in the case of order–three tensors, is very hard to compute and not known apart from a few cases (see [38]).

The rank of a tensor with real entries is not the same whether you consider it over $\mathbb{R}$ or over $\mathbb{C}$: the complex rank is smaller or equal than the real rank, since a decomposition as a sum of real rank–one tensors is in particular a complex one, but the converse is not true. We now give a first definition of *typical ranks*.

Definition 3.26 (*Typical rank*)**.** Consider the space of order–d tensors of format $n_1\times\cdots\times n_d$, $\mathbb{F}^{n_1\times\cdots\times n_d}$. We say that an integer $r \in \mathbb{N}$ is a *typical rank* of the format $n_1 \times \cdots \times n_d$ if the set of tensors with rank equal to r

$$\left\{T \in \mathbb{F}^{n_1\times\cdots\times n_d} \mid \operatorname{rank}(T) = r\right\}$$

is a full dimensional semialgebraic set.

In the matrix case, since matrices with maximal rank are an open dense set, we have a unique typical rank, which is always the maximal one. For tensors over $\mathbb{C}$, it can be shown that there still exists a unique typical rank, which is called the *generic rank* of that format. Nevertheless, this rank is in general not the maximal one ([38]). When we instead consider the real rank of tensors with real entries, also uniqueness is not true anymore in general: there are formats which admit more than one typical rank.

Before going further into the study of real typical ranks, we want to introduce another point of view, with a more probabilistic flavour, which will be the main perspective from which we approach the study of typical ranks. We will therefore define what a *random* tensor will be for us in this section. Remark that the tensors we are considering here have no symmetry, hence the probability space we are about to introduce is not the same that we considered previously in Section 3.1.

Definition 3.27 (*Gaussian tensor*)**.** Given a tensor $T = (t_{i_1,\dots,i_d}) \in \mathbb{R}^{n_1\times\cdots\times n_d}$, we say that T is *Gaussian* is its entries are independent random variables distributed as standard Gaussians $\mathcal{N}(0,1)$.

Definition 3.28 (*Typical rank revisited*)**.** Consider a Gaussian tensor T of format $n_1 \times \cdots \times n_d$. We say that an integer $r \in \mathbb{N}$ is a *typical rank* of the format $n_1 \times \cdots \times n_d$ if the probability that T has rank equal to r is positive

$$\mathbb{P}\left(\{\text{rank}(T) = r\}\right) > 0.$$

It is easy to see that Definition 3.28 is equivalent to Definition 3.26 for real tensors. Indeed, the distribution of a Gaussian tensor is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^{n_1\times\cdots\times n_d}$ and non–degenerate, while being a full dimensional semialgebraic set implies containing an open Euclidean ball.

From now on we will only focus on the case of order–three real tensors. We will denote formats of these tensors by $m \times n \times \ell$, always assuming $m \leq n \leq \ell$. We here explicitly write the definition of tensor rank in the case of order–3 ones. A $m \times n \times \ell$ tensor T is said to have *rank one* if there exists $a = (a_i)_{i=1,\dots,m} \in \mathbb{F}^m$, $b = (b_j)_{j=1,\dots,n} \in \mathbb{F}^n$ and $c = (c_k)_{j=1,\dots,\ell} \in \mathbb{F}^\ell$ such that

$$T = a \otimes b \otimes c,$$

where the symbol $\otimes$ denotes the usual tensor product of vectors. In other words the (ijk)–th entry of T is given by the product of the corresponding entries of the vectors a,b and c (see Definition 3.4). The *rank* of a tensor $T \in \mathbb{F}^{m\times n\times \ell}$ is the minimum number of rank–one tensors that add up to T (see Definition 3.5).

The first thing one might ask is which are the order–3 formats that admit multiple typical ranks. Several results are known from the literature, see Section 1.2 for an account of the state of the art. Most of these results are obtained through linear algebra manipulations on the given tensors and thus exploit very little of the geometric structure of the space of tensors and of rank–one tensors.

The novelty of this work, which is based on a result by Friedland (Theorem 3.29), is to apply a more geometric, rather than algebraic, viewpoint on typical ranks. We approach the problem via studying how a linear subspace slices the Segre variety of rank–one matrices. Even though we do not find any new result on which formats admit multiple typical ranks, we are able to recover several results from the literature with our method. Our proofs are in most cases simpler and provide more geometrical insight than the original ones. We develop and apply our method in Section 3.2.1–Section 3.2.3.

The other question one might ask is which are the exact probabilities of having a certain rank for a Gaussian tensor, in the case where its format admits more than one typical rank. Almost nothing is known about this, apart from the works of Bergqvist and Forrester on $2 \times n \times \ell$ formats ([11], [12], see Section 1.2). In Section 3.2.4 we will focus on the $3 \times 3 \times 5$ case and connect it, via our geometric framework, with the probabilities that the lines on a random cubic surface are real rather than complex.

In Section 3.2.5 we study how a random linear subspace slices the Segre variety *on average*. We will then apply the result to show some heuristics explaining the behaviour of rank probabilities observed in numerical simulations.

3.2.1 A geometric perspective: (randomly) slicing the Segre variety

We start by setting the following notations. Given vectors $v_1, \dots, v_s \in \mathbb{R}^a$, we denote with angle brackets $\langle v_1, \dots, v_s \rangle$ the linear subspace generated by them. By $\langle v_1, \dots, v_s \rangle_{\mathbb{P}}$ we denote the projectivization of such subspace. Given a tensor $T = (t_{ijk}) \in \mathbb{R}^{m \times n \times \ell}$, we define its *slices* $T_1, \dots, T_\ell$ to be the $m \times n$ matrices obtained by fixing the third index of the tensor, i.e.

$$(T_k)_{ij} = t_{ijk}$$

for every $k = 1, \dots, \ell$ and every $i = 1, \dots, m$, $j = 1, \dots, n$. We say that we slice the tensor along its third direction; this will always be our convention when talking about slices.

The starting point of our geometric approach to typical ranks is the following result by Friedland, which will be instrumental in all our work.

Theorem 3.29 ([38, Theorem 2.4]). *Fix a tensor $T \in \mathbb{R}^{m \times n \times \ell}$ and let $T_1, \dots, T_\ell \in \mathbb{R}^{m \times n}$ be its slices. Then the rank r of T is the minimal number of rank–one matrices $A_1, \dots, A_r \in \mathbb{R}^{m \times n}$ such that they span a subspace containing the span of the slices*

$$\langle T_1, \dots, T_\ell \rangle \subseteq \langle A_1, \dots, A_r \rangle.$$

Proof. Remark that a tensor can be written in terms of its slices as $T = \sum_{i=1}^{\ell} T_i \otimes e_i$, where $\{e_1, \dots, e_\ell\}$ is the standard basis of $\mathbb{R}^\ell$. If the tensor admits a rank r decomposition as $T = \sum_{j=1}^{r} a_j \otimes b_j \otimes c_j$, we decompose the vectors $c_j \in \mathbb{R}^\ell$ in the standard basis as $c_j = \sum_{i=1}^{\ell} \lambda_{ij} e_i$ for some $\lambda_{ij} \in \mathbb{R}$. Then, using the multilinearity of tensor product, we have $T = \sum_{i=1}^{\ell} (\sum_{j=1}^{r} \lambda_{ij} a_j \otimes b_j) \otimes e_i$. Therefore every T_i is contained in the linear subspace spanned by r rank–one matrices $\langle a_1 \otimes b_1, \dots, a_r \otimes b_r \rangle$.
Conversely, if each T_i is contained in a subspace of the form $\langle a_1 \otimes b_1, \dots, a_r \otimes b_r \rangle$, then there exists coefficients $\lambda_{ij} \in \mathbb{R}$, such that for every $i = 1, \dots, \ell$ we have $T_i = \sum_{j=1}^{r} \lambda_{ij} a_j \otimes b_j$. It follows that

$$T = \sum_{i=1}^{\ell} T_i \otimes e_i = \sum_{i=1}^{\ell} \left(\sum_{j=1}^{r} \lambda_{ij} a_j \otimes b_j \right) \otimes e_i = \sum_{j=1}^{r} a_j \otimes b_j \otimes \left(\sum_{i=1}^{\ell} \lambda_{ij} e_i \right).$$

This is a rank r decomposition of T. □

Observe that from the matrices $A_1, \dots, A_r$ one can actually explicitly build a rank r decomposition of the tensor T.

In Theorem 3.29, to compute the rank of a tensor we thus need to determine the minimal number of rank-one matrices that are needed to span the linear subspace $\langle T_1, \dots, T_\ell \rangle$. Remark that since tensor rank, and in particular also matrix rank, are invariant under non–zero scaling, we can consider the projectivization of that linear space. A natural question is then whether or not the subspace $\langle T_1, \dots, T_\ell \rangle_{\mathbb{P}}$ is itself spanned by (projective) rank–one matrices. To approach this problem, we introduce the following notations.

The projectivization of the space of $m \times n$ matrices $\mathbb{R}^{m \times n}$ is $\mathbb{R}\mathrm{P}^{mn-1}$. Denote by

$$\Sigma_{m,n} := \{ M \in \mathbb{R}\mathrm{P}^{mn-1} \mid \mathrm{rank}(M) = 1 \}$$

the set of projective real rank–one matrices. This is a smooth algebraic variety in $\mathbb{RP}^{mn-1}$, called the *Segre variety*, defined by the vanishing of 2×2 minors of M. Its dimension and codimension in $\mathbb{RP}^{mn-1}$ are given by

$$\dim(\Sigma_{m,n}) = m + n - 2\,, \qquad \operatorname{codim}(\Sigma_{m,n}) = (m-1)(n-1).$$

The Segre variety is the image of the Segre embedding $\mathbb{RP}^{m-1} \times \mathbb{RP}^{n-1} \hookrightarrow \mathbb{RP}^{mn-1}$, given by taking tensor products $(x, y) \mapsto x \otimes y$. The complex Zariski closure of $\Sigma_{m,n}$ is the *complex Segre variety*

$$\Sigma^{\mathbb{C}}_{m,n} := \{M \in \mathbb{CP}^{mn-1} \mid \operatorname{rank}(M) = 1\}$$

of projective complex rank–one matrices. It is smooth and irreducible. By [48, Example 18.15], its degree is

$$\deg(\Sigma^{\mathbb{C}}_{m,n}) = \binom{m+n-2}{m-1}. \tag{3.53}$$

The latter means that a general complex linear subspace $L \subseteq \mathbb{CP}^{mn-1}$ of dimension $(m-1)(n-1)$ will intersect the variety $\Sigma^{\mathbb{C}}_{m,n}$ in $\binom{m+n-2}{m-1}$ many distinct complex points. Note that such an L corresponds to the projectivization of a general $((m-1)(n-1)+1)$-dimensional linear space in $\mathbb{C}^{m\times n}$. The word "general" here means that this holds for all L outside some lower dimensional semialgebraic set.

In Section 2.7 we introduced the Grassmannian of s–dimensional subspaces in $\mathbb{R}^a$, denoted by $G(s, a)$. Here we focus on Grassmannians $G(d, mn)$ for $1 \leq d \leq mn$, i.e.

$$G(d, mn) = \{L \subseteq \mathbb{R}^{m\times n} \mid L \text{ is a linear subspace of dimension } d\}.$$

Recall that Grassmannians are homogeneous spaces under the action of the orthogonal group $O(mn)$. As a consequence of Theorem 2.23, there is a unique probability distribution on $G(d, mn)$ that is invariant under the action of the orthogonal group $O(mn)$. We call this the *uniform distribution*, since due to its invariance, its density is the same on every point. A random subspace with distribution given by the uniform one will be called a *uniform subspace*.

Recall also that the we endowed the space of $m \times n$ matrices with the Frobenius scalar product

$$\langle A, B\rangle_F := \operatorname{tr}(A^\top B), \tag{3.54}$$

which coincides with the Euclidean scalar product on $\mathbb{R}^{m\times n}$. For a subspace $L \in G(d, mn)$ we denote its *orthogonal complement* by

$$L^\perp := \{A \in \mathbb{R}^{m\times n} \mid \langle A, B\rangle_F = 0 \text{ for all } B \in L\} \in G(mn-d, mn). \tag{3.55}$$

We have the following elementary lemma, which will be useful in later sections.

Lemma 3.30. *A random subspace $L \in G(d, mn)$ is uniform if and only if its orthogonal complement $L^\perp \in G(mn-d, mn)$ is uniform.*

Proof. A subspace $L \in G(d, mn)$ is $O(mn)$-invariant, if and only if its orthogonal complement $L^\perp \in G(mn-d, mn)$ is $O(mn)$-invariant. The result follows from the uniqueness of the uniform distribution on $G(d, mn)$ as the only invariant one under the action of $O(mn)$. □

We are now ready to study the matter whether a uniform subspace is spanned by rank–one matrices or not.

Proposition 3.31. *Let $0 \leq d \leq mn-1$ and $L \subseteq \mathbb{RP}^{mn-1}$ correspond to a uniform linear subspace in $G(d+1, mn)$. With probability one the following trichotomy holds.*

1. *In the case $d < (m-1)(n-1)$: L is not spanned by rank-one matrices.*
2. *In the case $d = (m-1)(n-1)$: $\#(L \cap \Sigma_{m,n}) \geq d+1$ implies that L is spanned by rank-one matrices.*
3. *In the case $d > (m-1)(n-1)$: $L \cap \Sigma_{m,n} \neq \emptyset$ implies that L is spanned by rank-one matrices.*

In particular, for any $0 \leq d \leq mn-1$ we have that, with probability one, $\#(L \cap \Sigma_{m,n}) \geq d+1$ implies that L is spanned by rank-one matrices.

Proof. With probability one, a uniform subspace in $G(d+1, mn)$ is a general subspace. Recall that the codimension of $\Sigma^{\mathbb{C}}_{m,n}$ in $\mathbb{CP}^{mn-1}$ is $(m-1)(n-1)$. For $L \subseteq \mathbb{RP}^{mn-1}$ corresponding to a linear subspace in $G(d+1, mn)$, let $L^{\mathbb{C}}$ denote the complex linear space spanned by L.

If $d < (m-1)(n-1)$, then with probability one we have that there is no intersection between L and the Segre variety, $L \cap \Sigma_{m,n} = \emptyset$. Therefore L is not spanned by rank-one matrices.

If $d = (m-1)(n-1)$, by definition of degree, with probability one the intersection $L^{\mathbb{C}} \cap \Sigma^{\mathbb{C}}_{m,n}$ consists of $\deg(\Sigma^{\mathbb{C}}_{m,n}) = \binom{m+n-2}{m-1}$ complex points. The *General Position Theorem* (see, e.g., [6, p. 109]) asserts that any subset of $d+1$ points in the intersection $L^{\mathbb{C}} \cap \Sigma^{\mathbb{C}}_{m,n}$ is linearly independent. Consequently, if there are at least $d+1$ real points among them, that is $L \cap \Sigma_{m,n}$ has cardinality at least $d+1$, then we can find $d+1$ linearly independent points in $L \cap \Sigma_{m,n}$. For dimensional reasons, their span must be the subspace L, which is therefore spanned by rank–one matrices.

Finally, suppose that $d > (m-1)(n-1)$. In this case, the intersection $L^{\mathbb{C}} \cap \Sigma^{\mathbb{C}}_{m,n}$ is a positive dimensional projective variety. Again by the *General Position Theorem*, a general choice of $d+1$ points in such intersection $L^{\mathbb{C}} \cap \Sigma^{\mathbb{C}}_{m,n}$ spans the complex subspace $L^{\mathbb{C}}$. If there is a real point of intersection, i.e. $L \cap \Sigma_{m,n} \neq \emptyset$, since a general linear section of $\Sigma^{\mathbb{C}}_{m,n}$ is smooth by Bertini's theorem, $L \cap \Sigma_{m,n}$ consists of smooth points and hence is Zariski-dense in the complex intersection $L^{\mathbb{C}} \cap \Sigma^{\mathbb{C}}_{m,n}$. Therefore, we can find $d+1$ points in $L \cap \Sigma_{m,n}$ that are general in $L^{\mathbb{C}} \cap \Sigma^{\mathbb{C}}_{m,n}$ and hence span the real subspace L.

□

Remark 3.32. In the statement of Proposition 3.31 we assume the subspace corresponds to a uniform one and we say that the trichotomy holds with probability one. We could have equivalently stated the result for the projectivization of a general linear subspace in $G(d+1, mn)$. We opted for the first option to make more evident the connection to Gaussian tensors (see Lemma 3.33 below).

We conclude this section by determining the distribution of the linear span of the slices of a Gaussian tensor. Let $T \in \mathbb{R}^{m \times n \times \ell}$ be a Gaussian tensor with slices $T_1, \ldots, T_\ell \in \mathbb{R}^{m \times n}$. With probability one, their linear span has dimension ℓ and hence defines a random point in the Grassmannian $G(\ell, mn)$. We characterize its distribution as follows.

Lemma 3.33. *Let $T \in \mathbb{R}^{m\times n\times \ell}$ be a Gaussian tensor with slices $T_1, \dots, T_\ell \in \mathbb{R}^{m\times n}$. Let $L := \langle T_1, \dots, T_\ell\rangle$ be the linear span of the slices; then the random linear subspace L is uniform in the Grassmannian $G(\ell, mn)$.*

Proof. Since matrices with independent normal Gaussians as variables are invariant under multiplication by orthogonal matrices, the distribution of the subspace $L = \langle T_1, \dots, T_\ell\rangle$ is $O(mn)$-invariant. Since there is a unique such distribution, L must be a uniform subspace. □

3.2.2 Probabilistic Friedland's theorem

The following result provides a probabilistic counterpart of Friedland's theorem (Theorem 3.29), using the notions we have introduced in the previous section.

Proposition 3.34. *Let $T \in \mathbb{R}^{m\times n\times \ell}$ be a Gaussian tensor, with slices $T_1, \dots, T_\ell \in \mathbb{R}^{m\times n}$. Let $L := \langle T_1, \dots, T_\ell\rangle_{\mathbb{P}}$ be the projectivization of the linear span of the slices. Then, with probability one,*

$$\operatorname{rank}(T) \le \ell \quad \iff \quad L \cap \Sigma_{m,n} \text{ contains at least } \ell \text{ points.}$$

Proof. Since the rank of T is always bounded from above by mn, the proposition is trivial if $\ell \ge mn$, since with probability one $L = \mathbb{R}\mathrm{P}^{mn-1}$. Thus we can assume $\ell < mn$.

Assume that $\operatorname{rank}(T) \le \ell$. Then by Theorem 3.29, there exist real rank–one $m \times n$ matrices $\{x_i \otimes y_i\}_{i=1,\dots,\ell}$ such that their linear span has dimension ℓ and contains the linear span of $T_1, \dots, T_\ell$. As already observed, since T has Gaussian independent entries, with probability one its slices span an ℓ-dimensional subspace. This implies that we have the equality of subspaces $L = \langle x_1 \otimes y_1, \dots, x_\ell \otimes y_\ell\rangle_{\mathbb{P}}$. Since the $\{x_i \otimes y_i\}_{i=1,\dots,\ell}$ are linearly independent, they correspond to different points in the projective space, proving that $L \cap \Sigma_{m,n}$ contains at least ℓ points.

Vice versa, assume that the intersection $L \cap \Sigma_{m,n}$ contains at least ℓ real points. Proposition 3.31 implies that, with probability one, L is spanned by rank–one matrices. Therefore, ℓ of the points in $L \cap \Sigma_{m,n}$, say $\{x_i \otimes y_i\}_{i=1,\dots,\ell}$, are linearly independent. Then, we have $L = \langle x_1 \otimes y_1, \dots, x_\ell \otimes y_\ell\rangle_{\mathbb{P}}$. By Theorem 3.29 it follows that $\operatorname{rank}(T) \le \ell$. □

We can now start looking at the applications of our geometric framework to the study of typical ranks. Our first main result is the following.

Theorem 3.35. *Suppose that $(m-1)(n-1)+1 \le \ell \le mn$. Then, the typical ranks of the format $m \times n \times \ell$ are always contained in $\{\ell, \ell+1\}$. Moreover, ℓ is always a typical rank.*

In [86, Theorem 8.1] the authors give an algebraic proof of the first part of Theorem 3.35 assuming $3 \le m$. Our proof instead only uses the geometric framework we just introduced and is essentially based on the *General Position Theorem* via Proposition 3.31.

Proof of Theorem 3.35 – First part. Let $T \in \mathbb{R}^{m\times n\times \ell}$ be a Gaussian tensor with slices $T_1, \dots, T_\ell \in \mathbb{R}^{m\times n}$ and let $L = \langle T_1, \dots, T_\ell\rangle$ be the linear span of the slices. By Lemma 3.33, L is a uniform subspace in $G(\ell, mn)$. In particular, with probability one we have that $\dim(L) = \ell$. It follows from Theorem 3.29 that the rank of T is at least ℓ. This shows that typical ranks are bounded from below by ℓ.

We now prove that typical ranks are bounded from above by $\ell+1$. We denote the projectivization of $L \in G(\ell, mn)$ also by L; it will be clear from the context whether we are referring to its projectivization or not. If $L \cap \Sigma_{m,n}$ contains at least ℓ real points, by Proposition 3.34 the rank of T is bounded by ℓ. Assume this is not the case and let $x \otimes y$ be a general rank–one matrix such that $x \otimes y \notin L$. Define $\widetilde{L} := L + \langle x \otimes y \rangle \in G(\ell+1, mn)$. In order to apply Proposition 3.31 to $\widetilde{L}$, we will now prove that $\widetilde{L}$ is general (see Remark 3.32).

Let $G_{\mathbb{C}}(d, N)$ denote the complex Grassmannian and recall that taking complex Zariski closures we have $\overline{G(d,N)} = G_{\mathbb{C}}(d,N)$ and $\overline{\Sigma_{m,n}} = \Sigma^{\mathbb{C}}_{m,n}$. Consider the rational map

$$\begin{aligned} \phi : G_{\mathbb{C}}(\ell, mn) \times \Sigma^{\mathbb{C}}_{m,n} &\dashrightarrow G_{\mathbb{C}}(\ell+1, mn) \\ (L, p) &\longmapsto L + \langle p \rangle \end{aligned} .$$

Denote by $A := \phi(G(\ell, mn) \times \Sigma_{m,n})$ the image of the real Grassmannian paired with the real Segre variety. We want to prove that elements in A are general in $G_{\mathbb{C}}(\ell+1, mn)$, that is, $\overline{A} = G_{\mathbb{C}}(\ell+1, mn)$. Note that $G_{\mathbb{C}}(\ell, mn) \times \Sigma^{\mathbb{C}}_{m,n}$ and $G_{\mathbb{C}}(\ell+1, mn)$ are irreducible varieties and $\overline{G(\ell, mn) \times \Sigma_{m,n}} = G_{\mathbb{C}}(\ell, mn) \times \Sigma^{\mathbb{C}}_{m,n}$. We have that the Zariski closure of A is the same of the Zariski closure of the image of ϕ (see, e.g., [14, Lemma 1.2]). By the last part of Proposition 3.31, a general point in $G_{\mathbb{C}}(\ell+1, mn)$ is spanned by rank–one matrices, which implies

$$\overline{A} = \overline{\phi(G_{\mathbb{C}}(\ell, mn) \times \Sigma^{\mathbb{C}}_{m,n})} = G_{\mathbb{C}}(\ell+1, mn).$$

This proves that $\widetilde{L}$ is general in $G_{\mathbb{C}}(\ell+1, mn)$. Since by construction we have a non–empty intersection $\widetilde{L} \cap \Sigma_{m,n} \neq \emptyset$, the last part of Proposition 3.31 implies that $\widetilde{L}$ is spanned by rank–one matrices. By Theorem 3.29 this implies that the rank of T is bounded by $\ell+1$. Thus, with probability one, the rank of T is always bounded by $\ell+1$. This concludes the proof of the first part of the theorem.

□

To prove the second part of Theorem 3.35 we need our second main result. It express that probability that a Gaussian $m \times n \times \ell$ tensor has rank given by ℓ in terms of properties of uniform random slices of the Segre variety.

Theorem 3.36. *Let $T \in \mathbb{R}^{m \times n \times \ell}$ be a tensor of format $m \times n \times \ell$. If $\ell = (m-1)(n-1)+1$, we have*

$$\mathop{\mathbb{P}}_{\substack{T \in \mathbb{R}^{m\times n\times \ell} \\ \text{Gaussian}}} (\{\operatorname{rank}(T) = \ell\}) = \mathop{\mathbb{P}}_{\substack{L \in G(\ell, mn) \\ \text{uniform}}} (\{\#(L \cap \Sigma_{m,n}) \geq \ell\}) .$$

For $(m-1)(n-1)+1 < \ell \leq mn$ we have

$$\mathop{\mathbb{P}}_{\substack{T \in \mathbb{R}^{m\times n\times \ell} \\ \text{Gaussian}}} (\{\operatorname{rank}(T) = \ell\}) = \mathop{\mathbb{P}}_{\substack{L \in G(\ell, mn) \\ \text{uniform}}} (\{L \cap \Sigma_{m,n} \neq \emptyset\}) .$$

Proof. By the first part of Theorem 3.35, no typical rank can be smaller than ℓ. Therefore, with probability one, it holds that $\operatorname{rank}(T) \geq \ell$. Together with Proposition 3.34, this shows that with probability one $\operatorname{rank}(T) = \ell$ if and only if $L \cap \Sigma_{m,n}$ contains at least ℓ points. The equality of probabilities follows using Lemma 3.33.

Furthermore, Proposition 3.31 implies that for $\ell > (m-1)(n-1)+1$, with probability one, a single point in $L \cap \Sigma_{m,n}$ suffices to show that L is spanned by rank-one matrices. Hence, $L \cap \Sigma_{m,n}$ contains at least $\dim(L) = \ell$ points. This proves the second equality. □

We can now proceed with the proof of the second part of Theorem 3.35.

Proof of Theorem 3.35 – Second part. We prove that ℓ is always a typical rank. In the case $\ell > (m-1)(n-1)+1$, the second part of Theorem 3.36 shows that ℓ is a typical rank if and only if the probability that a real uniform subspace in $G(\ell, mn)$ intersects the real Segre variety is positive. To show that this is the case, we can simply take a point $A \in \Sigma_{m,n}$ and a general linear space K that passes through A such that K is not tangential to $\Sigma_{m,n}$ at A. Then there exists a neighborhood of K such that for all L in this neighborhood, we have $L \cap \Sigma_{m,n} \neq \emptyset$; this follows from the stability under small perturbations of the number of non–degenerate real points in the intersection of two real varieties. It follows that $\mathbb{P}(\{L \cap \Sigma_{m,n} \neq \emptyset\}) > 0$ for a uniform subspace L.

The case $\ell = (m-1)(n-1)+1$ is more difficult. By Theorem 3.36, we have to prove that $\mathbb{P}(\{\#(L \cap \Sigma_{m,n}) \geq \ell\}) > 0$. To see this, denote by $k := mn - \ell = m + n - 2$ and consider general points $a_1, \ldots, a_k \in \mathbb{R}^m$ and $b_1 \ldots, b_k \in \mathbb{R}^n$. They define a linear subspace $\{M \in \mathbb{R}^{m \times n} \mid a_1^T M b_1 = \cdots = a_k^T M b_k = 0\}$ of dimension ℓ. For a rank–one matrix $x \otimes y = xy^\top \in \Sigma_{m,n}$ with $x \in \mathbb{R}^m$ and $y \in \mathbb{R}^n$, these equations become

$$(a_1^\top x)(b_1^\top y) = \cdots = (a_k^\top x)(b_k^\top y) = 0. \tag{3.56}$$

Since the a_i and b_i are general, at most $m-1$ of the $a_i^\top x$ and at most $n-1$ of the $b_j^\top y$ can be simultaneously zero. Consequently, if all of the products in (3.56) are supposed to be zero, we can choose x such that exactly $m-1$ of the $a_i^T x$ are zero and this then determines y. This shows that there are $\deg(\Sigma_{m,n}^{\mathbb{C}}) = \binom{m+n-2}{m-1} \geq \ell$ (see (3.53)) real solutions of (3.56). Since they are all regular solutions, they can be continued in an open neighborhood of the system (3.56). Hence, $\mathbb{P}(\{\#(L \cap \Sigma_{m,n}) \geq \ell\}) > 0$. This concludes the proof.

□

Remark 3.37. If in Eq. (3.56) we use pairs of complex conjugate rank–one matrices (and one additional real rank–one matrix if k is odd) instead of real ones, we get the following number of real solutions:

$$\begin{cases} 0, & \text{if } m \text{ and } n \text{ are even} \\ \binom{(m+n-3)/2}{(m-1)/2}, & \text{if } m \text{ is odd and } n \text{ is even} \\ \binom{(m+n-3)/2}{(m-2)/2}, & \text{if } m \text{ is even and } n \text{ is odd} \\ \binom{(m+n-2)/2}{(m-1)/2}, & \text{if } m \text{ and } n \text{ are odd} \end{cases}$$

This is the number $\alpha(m,n)$ from [87]. In particular, this shows that if m and n are even, we always have two typical ranks. In the case one between m and n is odd, if the quantity above is less than ℓ, then again we have two typical ranks. Notice that $\alpha(3,3) = 2$, but we know from Corollary 3.45 below that $\mathbb{P}\{\#(L \cap \Sigma_{3,3}) = 0\} > 0$. Therefore $\alpha(m,n)$ is not a lower bound for the number of points in $L \cap \Sigma_{m,n}$, and if the quantities above are greater than ℓ, this gives no information on the existence of more than one typical rank. We have thus found again, with a completely different approach, Lemma 5.4 and Theorem 5.5 from [87]. Our approach sheds more light on the geometric meaning of the quantity $\alpha(m,n)$.

A direct consquence of Theorem 3.36 is the following.

Corollary 3.38. *If $(m-1)(n-1)+1 < \ell \leq mn$ and the degree of the Segre variety $\Sigma_{m,n}$ is odd, then there is only one typical rank, equal to ℓ, of the format $m \times n \times \ell$.*

Proof. It follows from the second part of Theorem 3.36, since the intersection with a linear subspace L of complimentary dimension is always non empty if the degree is odd [39]. □

Another consequence of Proposition 3.34 is that, if $2 \leq m \leq n \leq \ell < (m-1)(n-1)+1$, then with probability one, the projectivization of the linear span of the $m \times n$ slices of a Gaussian tensor $T \in \mathbb{R}^{m \times n \times \ell}$ will have no intersection with the Segre variety $\Sigma_{m,n}$; its dimension is indeed lower than the codimension of $\Sigma_{m,n}$. Therefore the probability that the condition given in Proposition 3.34 to have $\operatorname{rank}(T) \leq \ell$ is satisfied is 0; in other words the probability that $\operatorname{rank}(T)$ is at least $\ell + 1$ is equal to 1. This is equivalent to the following.

Corollary 3.39. *If $2 \leq m \leq n \leq \ell < (m-1)(n-1)+1$, then all typical ranks of the format $m \times n \times \ell$ are strictly greater than ℓ.*

Deciding whether or not $\ell+1$ is a typical rank is more subtle. In view of Theorem 3.29, we can see that having rank $\ell+1$ is equivalent to the existence of an $(\ell+1)$-plane containing the linear span of the slices of T, such that its projectivization intersects $\Sigma_{m,n}$ in at least $\ell + 1$ linearly independent real points. The subtlety is that we are not asking for a general $(\ell + 1)$-plane containing the span of the slices to have this property, but we only require the existence of a single such plane. Therefore, dimension counting arguments, which only provide general results, cannot be used.

3.2.3 Typical ranks of tall tensors

We will now give another proof of a result of ten Berge. In [88] he defines an $m \times n \times \ell$ tensor to be *tall* if $(m-1)n < \ell < mn$, for $2 \leq m \leq n \leq \ell$. He shows that the only typical rank for such tensors is always ℓ. While he uses purely algebraic means to prove such result, we now will use Proposition 3.34 once again to give a different proof.

Theorem 3.40 ([88])**.** *Let $2 \leq m \leq n \leq \ell$ such that $(m-1)n < \ell < mn$. The only typical rank of the format $m \times n \times \ell$ is ℓ.*

Proof. Let $T \in \mathbb{R}^{m \times n \times \ell}$ be a tall Gaussian tensor and denote by $T_1, \ldots, T_\ell$ its $m \times n$ slices. Let $k := mn - \ell$. With probability one, the dimension of the linear span $\langle T_1, \ldots, T_\ell \rangle$ is ℓ and therefore there exist $A_1, \ldots, A_k \in \mathbb{R}^{m \times n}$ such that

$$\langle T_1, \ldots, T_\ell \rangle = \{M \in \mathbb{R}\mathrm{P}^{mn-1} \mid \operatorname{tr}(M^T A_1) = \cdots = \operatorname{tr}(M^T A_k) = 0\},$$

i.e. $A_1, \ldots, A_k$ are a basis for the orthogonal space to $\langle T_1, \ldots, T_k \rangle$ (see (3.54)). For a rank–one matrix $x \otimes y$ with $x \in \mathbb{R}^m$ and $y \in \mathbb{R}^n$, the condition that it belongs to the linear span of $T_1, \ldots, T_k$ is equivalent to the system of equations

$$x^T A_1\, y = \cdots = x^T A_k\, y = 0.$$

This is in turn equivalent to

$$x^T \cdot (A_1 y \ldots A_k y) = 0, \tag{3.57}$$

i.e. x^T is in the left kernel of the $m \times k$ matrix having columns $A_i y$ for $i = 1, \ldots, k$. For every $i = 1, \ldots, n$ we define the matrix $B_i \in \mathbb{R}^{m \times k}$ containing the i-th columns of the matrices $A_1, \ldots, A_k$, i.e. the j-th column of B_i is the i-th column of A_j. Then a simple computation shows that

$$(A_1 y \ldots A_k y) = (B_1 y_1 + \cdots + B_n y_n),$$

where $y = (y_1, \ldots, y_n) \in \mathbb{R}^n$. Equation (3.57) thus becomes

$$x^T \cdot (B_1 y_1 + \cdots + B_n y_n) = 0. \tag{3.58}$$

Since the tensor T is tall, we have $k < m$ and equation (3.58) always has non–trivial solutions. This means that for every $y \in \mathbb{R}^n$ we can find $x \in \mathbb{R}^m$ such that $x \otimes y$ belongs to $\langle T_1, \ldots, T_\ell \rangle$. By Proposition 3.34 it follows that $\operatorname{rank}(T) \leq \ell$ with probability 1, hence no typical rank can be higher than ℓ. By Theorem 3.35, ℓ is also the smallest typical rank, thus ℓ is the only typical rank. □

3.2.4 $3 \times 3 \times 5$ tensors and cubic surfaces

In this section, we focus on the special case of Gaussian tensors of format $3 \times 3 \times 5$. ten Berge [89] showed that tensors in $\mathbb{R}^{3\times3\times5}$ admit two typical ranks: 5 and 6. Here we find a beautiful connection between $3 \times 3 \times 5$ Gaussian tensors and random cubic surfaces. In particular, we express the probability that a Gaussian tensor T has rank 5 in terms of the probability that the 27 lines on a determinantal random cubic surface are all real.

We start by recalling some facts from classical algebraic geometry. Given six points in $\mathbb{CP}^2$ in general position, i.e. no three points are collinear and not all lie on a conic, blowing up $\mathbb{CP}^2$ in those points produces a smooth cubic surface in $\mathbb{CP}^3$; indeed every smooth cubic surface can be generated in this way [27].

Lemma 3.33 implies that a $3 \times 3 \times 5$ Gaussian tensor T induces a uniform subspace $L \subseteq \mathbb{RP}^8$ of dimension 4 via taking the linear span of its slices . By Theorem 3.36, with probability one, if T has rank 5, the subspace L intersects $\Sigma_{3,3}$ in 6 points $x_1 \otimes y_1, \ldots, x_6 \otimes y_6$. We can therefore associate to a Gaussian $3 \times 3 \times 5$ tensor T of rank 5 a random cubic surface that is the blow up of $\mathbb{RP}^2$ in $x_1, \ldots, x_6$ or $y_1, \ldots, y_6$.

Remark that $\operatorname{codim}(\Sigma^{\mathbb{C}}_{3,3}) = 4$ and $\deg(\Sigma^{\mathbb{C}}_{3,3}) = 6$, so a general 4–dimensional subspace of $\mathbb{CP}^8$ intersects $\Sigma^{\mathbb{C}}_{3,3}$ in 6 points. Our results are based on the following theorem.

Theorem 3.41. *Let L be a four-dimensional general linear subspace of $\mathbb{RP}^8$. Then L intersects $\Sigma^{\mathbb{C}}_{3,3}$ in the 6 points $x_1 \otimes y_1, \ldots, x_6 \otimes y_6 \in \mathbb{CP}^8$ if and only if there exists a cubic surface S that results both from blowing up $\mathbb{CP}^2$ in $x_1, \ldots, x_6$ as well as from blowing up $\mathbb{CP}^2$ in $y_1, \ldots, y_6$, such that S has the following determinantal representation*

$$S = \{[z_0 : \cdots : z_3] \in \mathbb{RP}^3 \mid \det(z_0 M_0 + \cdots + z_3 M_3) = 0\},$$

where $\langle M_0, \ldots, M_3 \rangle = L^\perp$ is the orthogonal complement of L (see (3.55)).

Proof. The theorem is proved in [29, Theorem 6.9] for a general linear subspace of $\mathbb{CP}^8$. Since the real Grassmannian is Zariski dense in the complex Grassmannian, it also holds for a general subspace in $\mathbb{RP}^8$. □

Remark 3.42. Remarkably, [29] is a paper that studies a problem in computer vision. The connection to ranks of random tensors is surprising and one of the most interesting aspects of our work.

Example 3.43. Consider the following real rank–one matrices $x_i \otimes y_i$ for $i = 1, \ldots, 6$ where

$$\begin{array}{lllll} x_1 = (1,0,0) & x_2 = (0,1,0) & x_3 = (0,0,1) & x_4 = (1,1,1) & x_5 = (3,5,1) \\ y_1 = (1,0,0) & y_2 = (0,1,0) & y_3 = (0,0,1) & y_4 = (1,1,1) & y_5 = (8,2,1) \end{array}$$

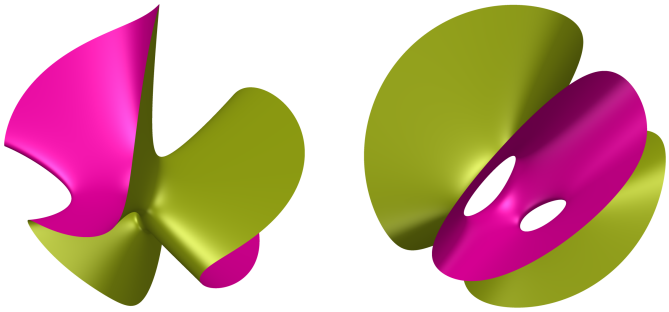


Figure 3.2: The cubic surface S in Eq. (3.60) intersected with the hyperplane $z_0 = 1$ seen from two different viewpoints. The pictures were created by S. Eggleston using `Surfer`.

and $x_6 = (-\frac{1}{3}, \frac{7}{5}, \frac{3}{17})$ and $y_6 = (-\frac{4}{3}, \frac{2}{5}, -\frac{1}{17})$. The first five rank–one matrices are linearly independent, but the sixth is linearly dependent on the first five. Therefore, the six rank–one matrices span a five–dimensional subspace $L \subseteq \mathbb{R}^{3\times 3}$. Its orthogonal complement is the 4–dimensional subspace

$$L^\perp = \left\{ \begin{bmatrix} 0 & 3z_0 & 3z_2 \\ -37z_0 - 2z_1 - 5z_2 + z_3 & 0 & 3z_3 \\ 34z_0 - z_1 + 2z_2 - 4z_3 & 3z_1 & 0 \end{bmatrix} \;\middle|\; z_0, \ldots, z_3 \in \mathbb{R} \right\}. \tag{3.59}$$

The four matrices $M_0, \ldots, M_3$ that give a basis for $L^\perp$ can be identified in (3.59) by separating the coefficients relative to the different parameters $z_0, \ldots, z_3$. The cubic surface S constructed in Theorem 3.41 is thus given in determinantal form by

$$S = \{z \in \mathbb{RP}^3 \mid z_0 z_3(34z_0 - z_1 + 2z_2 - 4z_3) + z_1 z_2(-37z_0 - 2z_1 - 5z_2 + z_3) = 0\}, \tag{3.60}$$

i.e. its equation is the determinant of the matrix in (3.59). The cubic surface S contains 27 real lines.

In 1849, Cayley and Salmon proved that every smooth cubic surface contains 27 lines [24]. Since their classical proof, many different proofs have been produced throughout the years. It was shown in 1858 by Schläfli [77] that a real cubic surface can only have 3, 7, 15, or 27 real lines, leading to a classification of cubic surfaces. We will need the following result due to Polo-Blanco and Top. Even if the proof is fairly easy and the methods do not differ from those used by Cayley and Schläfli in the 19th century, they observe that, to their surprise, they did not find this result anywhere in the literature.

Proposition 3.44. *[71, Proposition 2.1] Let S be a smooth real cubic surface. If S is obtained as the blow-up of*

1. *6 real points in the plane, it contains 27 real lines;*
2. *4 real points and a pair of complex conjugate points in the plane, it contains 15 real lines;*
3. *2 real points and 2 pairs of complex conjugate points in the plane, it contains 7 real lines*
4. *3 pairs of complex conjugate points in the plane, it contains 3 real lines.*

To be precise, there is also a fifth case in [71, Proposition 2.1]. Nevertheless, this case appears with probability zero and is not as simple as the others to write down. Therefore,

we do not include it in our description.

From this classification we derive the following corollary, establishing a connection between the probability of each of the cases in Proposition 3.44 and the probabilities of the different outcomes we might encounter in a uniform slice of $\Sigma_{3,3}$.

Corollary 3.45. *Let $M_0, M_1, M_2, M_3 \in \mathbb{R}^{3\times 3}$ be independent Gaussian matrices and consider the random cubic surface $S = \{[z \in \mathbb{RP}^3 \mid \det(z_0M_0 + \cdots + z_3M_3) = 0\}$. Let $L \in G(5,9)$ be a uniform subspace and denote also by L its projectivization. Denote by*

$$p_i := \mathbb{P}(\{\#(L \cap \Sigma_{3,3}) = i\}) \quad \text{and} \quad q_j := \mathbb{P}(\{S \text{ contains } j \text{ real lines}\}).$$

Then,

$$p_0 = q_3, \quad p_2 = q_7, \quad p_4 = q_{15}, \quad p_6 = q_{27}.$$

Proof. Since $\deg(\Sigma^{\mathbb{C}}_{3,3}) = 6$, a general real projective linear subspace L intersects $\Sigma_{3,3}$ in either $0, 2, 4$ or 6 real points, since the complex intersection points must always come in complex conjugate pairs. Therefore, $p_0 + p_2 + p_4 + p_6 = 1$. On the other hand, a smooth cubic surface contains either $3, 7, 15$ or 27 real lines. Since S is smooth with probability one, we obtain $q_3 + q_7 + q_{15} + q_{27} = 1$.

By Lemma 3.30, $L \in G(5,9)$ is uniform if and only if its orthogonal $L^\perp \in G(4,9)$ is uniform. Moreover, by Lemma 3.33 $\langle M_0, \ldots, M_3\rangle \in G(4,9)$ is also uniformly distributed. It follows from Theorem 3.41 that p_i is the probability that S is obtained as the blow up of i real points and $(6-i)/2$ pairs of complex conjugate points in the plane. Proposition 3.44 implies the inequalities $p_6 \leq q_{27}$, $p_4 \leq q_{15}$, $p_2 \leq q_7$ and $p_0 \leq q_3$. If one of those was a strict inequality, then we would have

$$q_3 + q_7 + q_{15} + q_{27} < p_0 + p_2 + p_4 + p_6 = 1,$$

which is a contradiction. Hence, all of them need to be equalities and the proof is concluded. □

The stage is now set to relate the probability that a random real $3 \times 3 \times 5$ tensor has rank 5 to the probability that all 27 lines on a determinantal random cubic surface are real.

Theorem 3.46. *Let $T \in \mathbb{R}^{3\times 3\times 5}$ be a Gaussian tensor and let $M_0, \ldots, M_3 \in \mathbb{R}^{3\times 3}$ be four independent Gaussian matrices. Define the random cubic surface*

$$S := \left\{[z_0, \ldots, z_3] \in \mathbb{RP}^3 \mid \det(z_0M_0 + \cdots + z_3M_3) = 0\right\}.$$

Then the following probabilities are equal

$$\mathbb{P}\left(\{\operatorname{rank}(T) = 5\}\right) = \mathbb{P}\left(\{\text{the } 27 \text{ lines on } S \text{ are all real}\}\right).$$

Proof. As before, we write $p_6 = \mathbb{P}(\{\#(L \cap \Sigma_{3,3}) = 6\})$. By Theorem 3.36, $\mathbb{P}(\{\operatorname{rank} T = 5\})$ is the probability that for a uniform subspace $L \in G(5,9)$, its projectivization L intersects $\Sigma_{3,3}$ in at least 5 points. Since L intersects $\Sigma_{3,3}$ in either 0, 2, 4 or 6 points with probability one, we have

$$\mathbb{P}(\{\operatorname{rank} T = 5\}) = p_6.$$

The statement now follows from Corollary 3.45. □

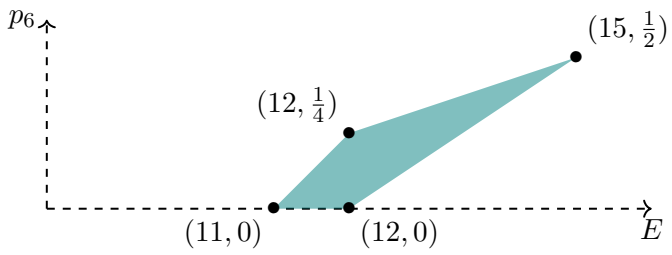


Figure 3.3: The polygon of possible values for (E, p_6), where E is the expected number of real lines on the random cubic, and p_6 the probability that it has 27 real lines.

Remark 3.47. Theorem 3.46 shows an extremely interesting connection between random tensors and random cubic surfaces, but yet does not produce a precise evaluation of the probability that a Gaussian $3 \times 3 \times 5$ tensor has rank 5 rather than 6. The reason why we were not able to explicitly evaluate the probability that the 27 lines on S are all real, is that taking the determinant results in the coefficients of the cubic surface now being product of three independent Gaussians. This is a complicated, but known, distribution. The real issue is that the same Gaussian now appears in multiple coefficients of the equation of the cubic, which are therefore not independent from each other. This makes almost all the methods that one could use to attack such computation, such as Kac-Rice formula, not applicable to this case.

We conclude this section by studying the expected number of real lines on our determinantal random cubic S. Denote by E the expectation

$$E := \mathbb{E}[\,\text{number of real lines on } S\,],$$

where $S = \{\det(z_0 M_0 + \cdots + z_3 M_3) = 0\}$ with M_i independent and Gaussian 3×3 matrices. We show in Example 3.50 below that the expected number of points in $L \cap \Sigma_{3,3}$ for uniform L is 3. Together with Corollary 3.45 we therefore obtain the following system of equalities and inequalities:

$$\begin{cases} 0 \leq p_0, p_2, p_4, p_6 \leq 1 \\ p_0 + p_2 + p_4 + p_6 = 1 \\ 2p_2 + 4p_4 + 6p_6 = 3 \\ 3p_0 + 7p_2 + 15p_4 + 27p_6 = E \end{cases}$$

Treating p_0, p_2, p_4, p_6 and also E as variables, the solution set of this system is a two–dimensional polytope. When projected to the (E, p_6)-plane, this polytope is the convex hull of the four points $(11, 0), (12, 0), (12, \frac{1}{4}), (15, \frac{1}{2})$ (see Fig. 3.3). From this we obtain the following nontrivial bound for the expectation.

Corollary 3.48. *For the expectation of the number of real lines on the determinantal cubic surface S we have the bounds $11 \leq E \leq 15$.*

The expected number of lines for other models of random cubic surfaces has been computed in [2, 4, 7]. Comparing Corollary 3.48 to these results, E seems to be comparably large. This model could therefore be useful in case one needs to sample cubic surfaces with many real lines.

3.2.5 Asymptotics and heuristics

In Theorem 3.36 we proved that the rank of a random tensor $T \in \mathbb{R}^{m\times n\times \ell}$ with $\ell := (m-1)(n-1)+1$ depends on whether the corresponding random linear subspace spanned by its slices projectively intersects the Segre variety in enough real points. In this section, we will compute the average number of real points in such an intersection and compute its asymptotic for m fixed as $n \to +\infty$. Comparing the growth of this expectation with the growth of the required number of real points, we obtain some insight into how the probability of having rank ℓ behaves.

We start by stating the real projective version of Howard's kinematic formula [51]. Recall that given a Riemannian manifold $(\overline{M}, g)$, an embedded submanifold $M \overset{\iota}{\hookrightarrow} \overline{M}$ inherits an induced Riemannian metric by pullback through ι, denoted as $\iota^* g$. By $\mathrm{Vol}(M)$ we denote the total volume of the submanifold M with respect to the induced metric (if $\dim(M) < \dim(\overline{M})$ its volume with respect to the ambient metric g is indeed 0). If M is compact, $\mathrm{Vol}(M)$ will always be finite.

As usual we denote by $O(n+1)$ the orthogonal group and by $d\lambda$ integration with respect to the uniform measure on $O(n+1)$. Recall that the orthogonal group acts transitively on projective spaces. Assume we are given two submanifolds $A, B \hookrightarrow \mathbb{R}\mathrm{P}^d$ embedded in some projective space. Howard's kinematic formula gives an expression for the average volume of the intersection between A and B, in the following sense.

Theorem 3.49 ([51]). *Let $A, B \hookrightarrow \mathbb{R}\mathrm{P}^d$ be compact submanifolds of a real projective space and let $a = \dim(A)$, $b = \dim(B)$ and $c = d-(d-a)-(d-b)$. For almost every $\lambda \in O(n+1)$ the intersection $A \cap \lambda B$ is transversal of dimension c. Moreover, the function $\lambda \longmapsto \mathrm{Vol}(A \cap \lambda B)$ is measurable, and the following formula holds:*

$$\int_{\lambda\in O(n+1)} \mathrm{Vol}(A\cap \lambda B)\, d\lambda = \mathrm{Vol}(\mathbb{R}\mathrm{P}^c)\, \frac{\mathrm{Vol}(A)}{\mathrm{Vol}(\mathbb{R}\mathrm{P}^a)}\, \frac{\mathrm{Vol}(B)}{\mathrm{Vol}(\mathbb{R}\mathrm{P}^b)}. \tag{3.61}$$

We wanto apply (3.61) in our context. The ambient projective space is the space of $m \times n$ matrices $\mathbb{R}\mathrm{P}^{mn-1}$. The first submanifold will be the real Segre variety $\Sigma_{m,n}$ and we are interested in the intersection between $\Sigma_{m,n}$ and the projectivization of a uniform linear subspace of complimentary dimension $\ell = (m-1)(n-1)+1$. To generate such a uniform subspace, it suffices to fix any ℓ–dimensional subspace $L \subseteq \mathbb{R}^{m\times n}$ and act on that via a uniform $\lambda \in O(n+1)$. Therefore our second submanifold will be the projectivization of the fixed subspace L, which we denote again by L. We can thus apply (3.61) with $d = mn-1$, $A = \Sigma_{m,n}$, $B = L$. The dimension of the intersection will generically be $c = 0$ by construction, and the 0–dimensional volume simply counts the number of points. Howard's kinematic formula in this setting thus yields

$$\mathop{\mathbb{E}}_{L \text{ uniform}} [\#(L\cap \Sigma_{m,n})] = \frac{\mathrm{Vol}(\Sigma_{m,n})}{\mathrm{Vol}(\mathbb{R}\mathrm{P}^{m+n-2})},$$

i.e. a formula for the expected number of intersection points in a uniform slice of the real Segre variety. Moreover, the volumes involved can be explicitly computed as $\mathrm{Vol}(\Sigma_{m,n}) = \mathrm{Vol}(\mathbb{R}\mathrm{P}^{m-1}) \cdot \mathrm{Vol}(\mathbb{R}\mathrm{P}^{n-1})$ and $\mathrm{Vol}(\mathbb{R}\mathrm{P}^{k-1}) = \frac{\pi^{\frac{k}{2}}}{\Gamma(\frac{k}{2})}$, leading to

$$\mathbb{E}[\#(L\cap \Sigma_{m,n})] = \sqrt{\pi} \cdot \frac{\Gamma(\frac{m+n-1}{2})}{\Gamma(\frac{m}{2})\Gamma(\frac{n}{2})}. \tag{3.62}$$

Example 3.50. For $m = 3$, Equation (3.62) becomes

$$\mathbb{E}[\#(L \cap \Sigma_{3,n})] = n.$$

Therefore, for a random tensor $T \in \mathbb{R}^{3\times n\times(2n-1)}$ with the usual notation, Theorem 3.36 combined with Markov's inequality yields

$$\mathbb{P}\left(\{\text{rank}(T) = 2n-1\}\right) = \mathbb{P}\left(\{\#(L \cap \Sigma_{3,n}) \geq 2n-1\}\right) \leq \frac{n}{2n-1} < 1.$$

This gives another proof of the fact that there are two typical ranks when $m = 3$ and $\ell = (m-1)(n-1)+1$, since it implies that $\mathbb{P}\left(\{\text{rank}(T) > 2n-1\}\right)$ is positive. This is also an alternative proof to [87, Theorem 5.5] for the case $m = 3$.

We are now interested in the asymptotics of (3.62) when m is fixed and $n \to +\infty$. In order to compute this, we consider two cases, depending on the parity of m.

Proposition 3.51. *Let L be a uniform projective $(m-1)(n-1)$-dimensional subspace. For m fixed, as $n \to +\infty$, we have the asymptotic expression*

$$\mathbb{E}[\#(L \cap \Sigma_{m,n})] = c \cdot n^{\frac{m-1}{2}}\left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right),$$

where $c = \frac{1}{(m-2)!!}$ for m odd and $c = \sqrt{\frac{\pi}{2}}\frac{1}{(m-2)!!}$ for m even.

Proof. We first consider the case when m is odd. Let $m = 2k+1$ for some $k \geq 1$. In this case, (3.62) reads

$$\mathbb{E}[\#(L \cap \Sigma_{2k+1,n})] = \sqrt{\pi} \cdot \frac{\Gamma(\frac{n}{2}+k)}{\Gamma(k+\frac{1}{2})\Gamma(\frac{n}{2})}.$$

By the multiplicative property of the Gamma function $\Gamma(x+1) = x\Gamma(x)$, we obtain

$$\Gamma\left(\frac{n}{2}+k\right) = \Gamma\left(\frac{n}{2}\right)\prod_{i=0}^{k-1}\left(i+\frac{n}{2}\right), \quad \text{and} \quad \Gamma\left(\frac{1}{2}+k\right) = \sqrt{\pi}\prod_{i=0}^{k-1}\left(i+\frac{1}{2}\right),$$

leading to

$$\mathbb{E}[\#(L \cap \Sigma_{2k+1,n})] = \prod_{i=0}^{k-1}\frac{\left(\frac{n}{2}+i\right)}{\left(\frac{1}{2}+i\right)}.$$

This is a polynomial of degree $k = \frac{m-1}{2}$ in n, with leading coefficient given by

$$\frac{(\frac{1}{2})^k}{\prod_{i=0}^{k-1}\frac{2i+1}{2}} = \frac{1}{\prod_{i=0}^{k-1}(1+2i)} = \frac{1}{(2k-1)!!},$$

where $(2k-1)!! = (2k-1)\cdot(2k-3)\cdot\cdots\cdot 3\cdot 1$ is the double factorial. Thus we obtain the asymptotic expression

$$\mathbb{E}[\#(L \cap \Sigma_{m,n})] = \frac{1}{(m-2)!!}\, n^{\frac{m-1}{2}}\left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right) \quad \text{for } m \text{ odd.}$$

We now consider the case when m is even. Let $m = 2k$ for some $k \geq 2$. In this case, (3.62) reads

$$\mathbb{E}[\#(L \cap \Sigma_{2k,n})] = \sqrt{\pi} \cdot \frac{\Gamma\left(\frac{n-1}{2}+k\right)}{\Gamma(k)\Gamma\left(\frac{n}{2}\right)}.$$

Again by the multiplicative property of the Gamma function we have

$$\Gamma\left(\frac{n-1}{2}+k\right)=\Gamma\left(\frac{n+1}{2}\right)\prod_{i=1}^{k-1}\left(\frac{n-1}{2}+i\right).$$

Using above identity and $\Gamma(k)=(k-1)!$, we obtain

$$\mathbb{E}[\#(L\cap\Sigma_{2k,n})]=\pi\cdot\frac{\prod_{i=1}^{k-1}\left(\frac{n-1}{2}+i\right)}{(k-1)!}\cdot\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{1}{2}\right)}=\pi\cdot\frac{\prod_{i=1}^{k-1}\left(\frac{n-1}{2}+i\right)}{(k-1)!}\cdot\frac{1}{B\left(\frac{1}{2},\frac{n}{2}\right)},$$

where $B(x,y)=\frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$ is the Beta function. Using the asymptotic $B\left(\frac{1}{2},\frac{n}{2}\right)\sim\sqrt{\frac{2\pi}{n}}$ (see [85, Equation 43:9:3]) and the identity $2^{k-1}(k-1)!=(2k-2)!!=(m-2)!!$, we have

$$\mathbb{E}[\#(L\cap\Sigma_{m,n})]=\sqrt{\frac{\pi}{2}}\frac{1}{(m-2)!!}\,n^{\frac{m-1}{2}}\left(1+\mathcal{O}\left(\frac{1}{n}\right)\right)\quad\text{for } m \text{ even}.$$

□

We can now compare the asymptotic growth in Proposition 3.51 with the threshold in Theorem 3.36. When we have $\ell=(m-1)(n-1)+1$, Theorem 3.36 states that for a random $m\times n\times\ell$ tensor having rank ℓ is, with probability one, equivalent to $\#(L\cap\Sigma_{m,n})\geq\ell$. While ℓ grows linearly in n for m fixed, the expectation in (3.62) grows as $n^{\frac{m-1}{2}}$.

For $m=3$, both quantities grow linearly, but while ℓ is given exactly by $2n-1$, the expectation of the number of intersection points is given by just n. Qualitatively, this means that the number of linearly independent points required to be real in order to have rank $\ell=2n-1$ grows faster, albeit with the same order, compared to the expected value of such number of points. This suggests that the probability that the rank is $2n-1$ should go to 0 as $n\to+\infty$.

For $m\geq 4$, while ℓ grows linearly, the expected value of real points grows as a higher power of n. Qualitatively, this tells us that as n increases, the number of points we require to be real becomes smaller and smaller relative to the expected value, i.e. it becomes more likely that the condition is satisfied. Based on this, we expect that the probability of having rank ℓ approaches 1 as $n\to+\infty$.

The above argument does not constitute a proof in the mathematical sense; in order to make it rigorous, one should study the higher order moments of the random variable given by the number of real intersection points. Nevertheless, it still gives a clear geometric explanation to what is observed in simulations. For example, we confirmed the behaviour described above for small values of m and n by randomly sampling uniform subspaces in $(m\times n\times(m-1)(n-1)+1)$ and computing numerically the intersection points between such subspace and the real Segre variety. The simulation was performed using *Julia* and its package *HomotopyContinuation.jl* [18].

Chapter 4

Nearest point problems in the Grassmannian

In this chapter we will focus on a class of optimization problems that is widespread in applications. This class is often referred to as *nearest point problems*. The general setting is the following.

Problem 4.1 (Nearest point problem)**.** Let (M, d) be a metric space and let $X \subseteq M$ be a subset. Given $u \notin X$, find a point $u^* \in X$ that is closest to u. In other words, denoting by $d_u : M \longrightarrow \mathbb{R}_{\geq 0}$ the function measuring the distance from u, that is $d_u(p) := d(u, p)$ for every $p \in M$, we have to find $u^* \in X$ such that

$$d_u(u^*) = \min_{x \in X} d_u(x).$$

We will often denote the restriction of a function f to a subset X as $f|_X$. With this notation, we are therefore looking for a minimizer of $d_u|_X$. We will always assume that the problem admits a solution; this is for example the case for every closed subset X.

This type of optimization problems appear naturally in many different settings. The metric space (M, d) and the subspace X involved depend on the kind of specific instance we need to solve. For example, a typical setting is when M represents the space of data (typically $\mathbb{R}^n$, but it can differ), d is any meaningful metric depending on the nature of the data we are looking at, and the subset X is a finite number of points. In this case, Problem 4.1 is usually called *nearest neighbor search*; it appears naturally in data analysis, pattern recognition in image processing, clustering and classification tasks ([78]). The method mostly employed for nearest neighbor search is *Divide and Conquer*, where you subdivide X in smaller subsets, solve the problem on each of them and then combine the results. There are several different implementations of this method, often exploiting the characteristics of the specific subset X. Another approach is via *Voronoi diagrams*, i.e. a decomposition of $\mathbb{R}^n$ in cells, one for each point of X, such that the cell corresponding to a point $x \in X$ comprises all of the points for which x is the closest point in X. This approach has the advantage that, once the diagram is built, finding the solution is very fast and exact. The disadvantage is that building the diagram becomes rapidly very expensive as the dimensionality of the data increases and, most importantly, the diagram can vary a lot when points are added or removed, which makes this approach highly impractical in dynamical situations.

We have already encountered another instance of Problem 4.1 in Section 2.6, where we considered the problem of approximating a $n \times m$ matrix with another one of lower rank.

In that case we had $M = \mathbb{R}^{n\times m}$ the space of $n \times m$ matrices with the Frobenius distance $\widehat{\delta} : \mathbb{R}^{n\times m} \times \mathbb{R}^{n\times m} \longrightarrow \mathbb{R}_{\geq 0}$ defined by (2.10). The set X was the set of matrices with rank bounded by some $r < \min(n, m)$, denoted by $\mathcal{M}_{\leq r}$, and the data point $A \in \mathbb{R}^{n\times m}$ was assumed to be generic; in particular it was full rank and with distinct singular values. Then Theorem 2.19 gave us the exact solution to Problem 4.1 in this context, via the singular values decomposition of A.

In general we are not as lucky as in Section 2.6 and we do not obtain a closed, exact expression for the minimizer in Problem 4.1. One strategy is usually to first find all the critical points of the function $d_u|_X$ and then look for the minima among them, if this is feasible. In Section 4.1 we will see how to use this approach when (M, d) is the Euclidean space $\mathbb{R}^n$ with its standard metric and the subset X is a real algebraic subvariety of $\mathbb{R}^n$. This setting was systematically studied for the first time in [34]. We will explain their main results and we will use the properties they highlight as our main guiding questions when studying Problem 4.1 in a different setting. Remark that the rank–approximation problem for matrices falls into this framework.

In Section 4.2 we will focus on Problem 4.1 in the Grassmannian setting. The ambient metric space will be given by the Grassmannian of k–planes in $\mathbb{R}^n$ with the distance induced by the Grassmann metric $g_{G(k,n)}$, i.e. $(M, d) = (G(k, n), \delta)$ (see Section 2.7). The subset X we consider is a smooth submanifold; when needed this will also be assumed to be subanalytic.

The results contained in Section 4.2 constitute the main project I carried out during my PhD studies. They are joint work with A. Lerario and can be found in the preprint [66].

4.1 The Euclidean Distance Degree

In the sciences, many models naturally arising in applications can be described by polynomial equations, that is as algebraic varieties in a given number of unknown variables that serve as parameters of the model. What typically happens is that you are given a data point, often coming from real–world measurements and observations, and you want to fit it to the model, that is to find the *best* point on the model that explains the given data. Instances of this kind of situation can be found in geometric modeling, n–view triangulation from computer vision, control theory, phylogenetics, algebraic statistics and many more.

All of these instances were originally studied case–by–case. In [34], J. Draisma, E. Horobet, G. Ottaviani, B. Sturmfels and R. R. Thomas gave the first unified treatment of these phenomena, describing a common setting for all of them in terms of a nearest point problem in the fashion of Problem 4.1. This section is intended as a gentle and very short introduction to the theory whose foundations were laid in [34], i.e. the theory of *Euclidean Distance Degree*. Even though we will not make use of the results of this section later, we include them as they constitute one of the main motivations for our work in [66]. Indeed, the many nice properties enjoyed by the Euclidean algebraic setting we are about to describe, served as guiding lights in our Grassmannian setting, where we try to understand which of the properties still hold and which ones do not. All the results we are going to present in this section can be found in [34].

We now describe the setting; as already pointed out, this will be phrased as an instance of a nearest point problem as in Problem 4.1. Let $X \subseteq \mathbb{R}^n$ be a real algebraic variety and

$u \in \mathbb{R}^n$ be a point not on X. The goal is to compute $u^* \in X$ that minimizes the Euclidean distance from u. Since minimizing a nonnegative function f or its square f^2 makes no difference, we look for $u^* \in X$ that minimizes the squared Euclidean distance function

$$d_u(x) := d(u,x) = \sum_{i=1}^{n} (u_i - x_i)^2 \tag{4.1}$$

restricted to the algebraic variety X, i.e.

$$d_u(u^*) := \min_{x \in X} d_u(x),$$

where $u = (u_1, \ldots, u_n) \in \mathbb{R}^n$ and $x = (x_1, \ldots, x_n) \in X$.

The strategy used to solve this problem relies on the algebraic nature of the Euclidean distance function (4.1). To exploit this algebraicity, we move to the complex world and consider the set of complex solutions to the polynomial equations defining X, i.e. we now look at $X \subseteq \mathbb{C}^n$; we use the same notation for X. The distance function d_u extends to a polynomial function on $\mathbb{C}^n$. Thus the value $d_u(x)$ for $x \in \mathbb{C}^n$ is a complex number and it can be 0 even if $u \neq x$. The idea is to first consider all the complex critical points of (4.1) and then look for the minimizer among them.

Remark 4.2. We are not using the standard Hermitian form on $\mathbb{C}^n$, but we are really just extending the Euclidean one to $\mathbb{C}^n$. The reason is that the operation of complex conjugation is not algebraic.

In order to do study critical points, we restrict to the smooth locus of X, which we denote by X^{sm}. Since the function d_u is smooth everywhere except at $u \notin X$, critical points for the restriction $d_u|_X$ which are smooth for X can be computed via the usual Lagrange multipliers rule: a point $x \in X^{\text{sm}}$ is critical for $d_u|_X$ if and only if

$$\nabla d_u(x) \perp T_x X, \tag{4.2}$$

where $\nabla d_u(x)$ is the gradient of the distance function at x and $T_x X$ is the tangent space to X at the point x. The gradient of d_u at x is just given by twice the difference between u and x, i.e.

$$\nabla d_u(x) = 2(u - x). \tag{4.3}$$

Assume that $I_X = \langle f_1, \ldots, f_s \rangle$ is a radical ideal such that $X \subseteq \mathbb{C}^n$ is its variety, where the angle brackets here mean that $f_1, \ldots, f_s \in \mathbb{R}[x_1, \ldots, x_n]$ are the generators of I_X. We will also always assume that X is irreducible and I_X is prime. We set the notation $f = (f_1, \ldots, f_s)$ and denote by $J(f)$ the *Jacobian* of f, that is the $s \times n$ matrix of partial derivatives of $f_1, \ldots, f_s$. Assume also that the codimension of X is c. Then a point $x \in X$ is smooth if and only if the rank of the Jacobian $J(f)$ at the point x has rank c, while the rank drops at singular points. It follows that singular points are identified by the vanishing of the $c \times c$ minors of $J(f)$. We define the ideal

$$I_{X^{\text{sing}}} := I_X + \langle\, (c \times c) - \text{minors of } J(f) \,\rangle\,.$$

If $x \in X^{\text{sm}}$, the rowspan of $J(f)(x)$, i.e. the span of the gradients $\nabla f_1(x), \ldots, \nabla f_s(x)$, is the normal space to X at x and has dimension c. From (4.3) we therefore see that the condition in (4.2) is equivalent to the vector $u - x$ being in the rowspan of $J(f)(x)$. We can create the *augmented Jacobian*, which is the $(s+1) \times n$ matrix obtained by adding the

row $u - x$ to $J(f)$. At a smooth point, condition (4.2) is then equivalent to the vanishing of $(c+1) \times (c+1)$–minors of the augmented Jacobian; indeed when adding the new row $u - x$, the rank should not change. We define the *critical ideal* to be

$$I_{\text{crit}} = \left(I_X + \left\langle (c+1) \times (c+1) - \text{minors of } \begin{pmatrix} u - x \\ J(f) \end{pmatrix} \right\rangle \right) : (I_{X^{\text{sing}}})^{\infty}, \qquad (4.4)$$

where the last part denotes saturation by $I_{X^{\text{sing}}}$; geometrically this means that we are removing the singular points of X. As the name of the ideal suggests, we have the following.

Lemma 4.3 ([34], Lemma 2.1)**.** *For a general $u \in \mathbb{C}^n$, the variety $V(I_{\text{crit}})$ defined by the critical ideal in $\mathbb{C}^n$ is finite and consists of the critical points of the squared Euclidean distance d_u restricted to X^{sm}.*

For general $u \in \mathbb{C}^n$, the number of points in $V(I_{\text{crit}})$ is not only finite, but also constant. Therefore the following definition makes sense.

Definition 4.4 (Euclidean Distance Degree)**.** Let $X \subseteq \mathbb{R}^n$ be a real algebraic variety. We define the *Euclidean Distance Degree* of X, denoted by $\text{EDD}(X)$, to be the number of complex critical points of the squared Euclidean distance function d_u on X^{sm} for a generic $u \in \mathbb{C}^n$, where we look at X as a variety in $\mathbb{C}^n$.

The Euclidean Distance Degree of an algebraic variety $X \subseteq \mathbb{R}^n$ can therefore be thought of as a measure of the algebraic complexity of the nearest point problem for X with respect to the Euclidean distance. It is therefore useful to have an a priori bound on $\text{EDD}(X)$, depending on the degrees of the defining equations. This is given by the following result. Remark that also the codimension of the variety is important.

Proposition 4.5 ([34], Proposition 2.6)**.** *Let $X \subseteq \mathbb{C}^n$ be an algebraic variety of codimension c. Assume that the defining polynomials are given by $f_1, \dots, f_c, \dots, f_s$ ordered by decreasing degrees $d_1 \geq d_2 \geq \cdots \geq d_c \geq d_s$. Then the following bound holds*

$$\text{EDD}(X) \leq d_1 d_2 \dots d_c \sum_{i_1 + i_2 + \cdots + i_c \leq n - c} (d_1 - 1)^{i_1} (d_2 - 1)^{i_2} \dots (d_c - 1)^{i_c} .$$

If X is a complete intersection (i.e. $s = c$), then equality holds.

From a computational point of view, the algebraic description of critical points given by I_{crit} in (4.4) makes the Euclidean Distance Degree computable, in theory, via symbolic computer algebra systems like *Macaulay2*. In practice, this only works for small enough varieties, since the task becomes rapidly unfeasible for symbolic computations as the degrees and number of equations grow larger. One can then try to switch to numerical algebra systems like *Bertini* or *Julia*, and use homotopy continuation packages like *HomotopyContinuation.jl* [18]. Even if you can push the computations further with these softwares, when the variety gets too complicated they again fail to give an answer.

On the theoretical side, there are formulas relating the Euclidean Distance Degree of an algebraic variety $X \subseteq \mathbb{C}^n$ to well–known geometric objects that can be associated to X. In particular, the $\text{EDD}(X)$ has an expression in terms of *polar classes* [34, Theorem 5.4] and in terms of *Chern classes* [34, Theorem 5.8].

We remark that there is a version of this construction also for projective varieties. In that case, the Euclidean Distance Degree of a projective variety is defined to be the one of the affine cone over the variety; see [34, Lemma 2.8]. We also stress that $\text{EDD}(X)$ is

highly sensitive to the metric used; for example changing the coefficients in (4.1) can alter dramatically the number of complex critical points; see [34, Example 3.2].

When we deal with real–world applications, the data point u is usually a real–valued point that we seek to approximate with a real–valued point on the model X. Therefore it makes sense to consider only real–valued critical points for d_u on the smooth part of X. For a generic data point $u \in \mathbb{R}^n$, their number is not constant anymore and only upper bounded by $\mathrm{EDD}(X)$. Nevertheless, their number is constant when $u \in \mathbb{C}^n$ varies in a connected component of the complement of an algebraic hypersurface. The polynomial defining this hypersurface is called the *ED discriminant*. This is the locus of data points for which we have less than $\mathrm{EDD}(X)$ complex critical points, i.e. where one of the critical points have higher multiplicity. For plane curves the ED discriminant defines the *evolute* of the curve, which is a very classical object dating back to 19–th century. More recently, ED discriminants were studied under the name of *focal loci* by Catanese and Trifogli ([23], [92]).

Remark 4.6. Some algebraic varieties have the property that all the complex critical points of the function d_u on X^{sm} are actually real–valued. In this case the number of real critical points for generic $u \in \mathbb{R}^n$ is then constant and equal to the $\mathrm{EDD}(X)$. This happens becuase the set of real points of the hypersurface defined by the ED discriminant has actually codimension at least two, which implies that it cannot disconnect $\mathbb{R}^n$. One example of this is the variety of matrices with bounded rank $\mathcal{M}_{\geq r}$ encountered in Section 2.6. Theorem 2.19 indeed describes all the critical points for the corresponding Euclidean nearest point problem and shows that all the critical points are real, whenever the generic data point u is real–valued.

Since their number varies, it makes sense to ask *on average* how many real critical points we have for a given algebraic variety X. To do this we first equip the space $\mathbb{R}^n$ with a volume form ω such that its density over $\mathbb{R}^n$ integrates to 1; in other words we endow $\mathbb{R}^n$ with a probability measure admitting a density. Then we define the *average Euclidean Distance Degree* of (X, ω) as the expectation

$$\mathrm{aEDD}(X, \omega) = \int_{\mathbb{R}^n} \#\left\{\text{real critical points of } d_u \text{ on } X^{\mathrm{sm}}\right\}\ \omega(u).$$

This number strongly depends on the chosen volume form ω and is, in general, way lower than the EDD of the variety. One way of computing this number is to randomly sample data points $u \in \mathbb{R}^n$ according to the distribution given by ω, compute the number of real critical points for such data point and then average over all the samplings. Nevertheless, whenever a uni–rational parametrization of the variety X is available, there is an integral formula [34, Formula (4.3)] which behaves numerically and computationally way better than the sampling approach. This formula is based on the *ED correspondence*, which is the last object we introduce. It describes the couples consisting of a data point and a critical point correspond to the data.

Definition 4.7. Let $X \subseteq \mathbb{C}^n$ be an irreducible algebraic variety. The ED correspondence $\mathcal{E}_X \subseteq \mathbb{C}^n \times \mathbb{C}^n$ is the variety defined by $I_{\mathrm{crit}} \subseteq \mathbb{C}[x_1, \ldots, x_n, u_1, \ldots, u_n]$, i.e. when also the u is treated as a variable. Explicitly, this is the Zariski closure of the set

$$\{\, (x, u) \mid x \in X^{\mathrm{sm}},\ x \text{ is critical for } d_u|_X \,\} \subseteq \mathbb{C}^n \times \mathbb{C}^n.$$

The next result shows the geometric structure of the ED correspondence.

Theorem 4.8 ([34], Theorem 4.1)**.** *The* ED *correspondence* $\mathcal{E}_X$ *of an irreducible algebraic variety* $X \subseteq \mathbb{C}^n$ *of codimension* c *is an irreducible variety of dimension* n *in* $\mathbb{C}^n \times \mathbb{C}^n$*. The projection on the first factor* $\pi_1 : \mathcal{E}_X \longrightarrow X$ *is an affine vector bundle of rank* c *over* X^{sm}*. For a general data point* $u \in \mathbb{C}^n$*, the second projection* $\pi_2 : \mathcal{E}_X \longrightarrow \mathbb{C}^n$ *has finite fibers of cardinality equal to* $\text{EDD}(X)$*.*

Observe that the preimage $\pi_1^{-1}(x)$ for a point $x \in X^{\text{sm}}$ describes all the points u for which x is a critical point of $d_u|_X$, i.e. all the points for which $u - x \perp T_x X$. In can therefore be identified as the embedded normal space to X at x. The other way round, the preimage $\pi_2^{-1}(u)$ describes all the smooth critical points of d_u on X. In Section 4.2.5 we will construct a similar object for the Grassmannian setting in order to prove uniform boundedness on the number of (real) critical points; see the proof of Proposition 4.31.

Many more things could be said on Euclidean Distance Degree; it has been, and still is, an active area of research since the first paper [34] was published. For us, it is sufficient to have outlined here the most relevant aspects, for our purpose, of Problem 4.1 in the Euclidean setting:

1. for general data point $u \in \mathbb{C}^n$ we have a constant number of critical points, called the *Euclidean Distance Degree*; in particular, the number of real critical points is finite and uniformly bounded.
2. there exists an explicit bound to the number of critical points in terms of the structure of the defining polynomial equations of the constraint variety (Proposition 4.5).
3. the *ED correspondence* allows to simultaneously describe data points and their corresponding critical points.

In the next section we will change the setting for Problem 4.1 and move to the Grassmannian world. A large part of our effort will be devoted to understanding the analogues of (1) and (2) from the list above, in a world where algebraicity and smoothness of the distance function are not available anymore; in doing so we will also introduce the Grassmannian version of (3) under some additional hypothesis.

4.2 The Grassmann Distance Complexity

In Section 4.1 we have seen how nearest point problems behave in the Euclidean algebraic setting. This raises a natural question: can we extend those results to other ambient spaces (M, d) and other types of subsets X? In this section we will investigate Problem 4.1 in the following Grassmannian setting.

Let $(G(k,n), g_{G(k,n)})$ be the Grassmannian of k–planes in $\mathbb{R}^n$ endowed with the Grassmann metric (Definition 2.26). Let $\delta : G(k,n) \times G(k,n) \longrightarrow \mathbb{R}_{\geq 0}$ be the Grassmann distance induced by $g_{G(k,n)}$ (Definition 2.29); recall that this is given in terms of principal angles between k–planes by (2.26). Our version of Problem 4.1 is the following.

Problem 4.9 (Grassmannian nearest point problem)**.** Let $(G(k,n), \delta)$ be the Grassmannian endowed with its Grassmann distance and let $X \subseteq G(k,n)$ be a submanifold. Given $\mathbf{L} \notin X$, find the point $\mathbf{L}^* \in X$ that is closest to $\mathbf{L}$. Denoting by $\delta_{\mathbf{L}} : G(k,n) \longrightarrow \mathbb{R}_{\geq 0}$ the distance function from $\mathbf{L}$

$$\delta_{\mathbf{L}}(\mathbf{E}) = \delta(\mathbf{L}, \mathbf{E})$$

for every $\mathbf{E} \in G(k,n)$, we seek to find $\mathbf{L}^* \in X$ such that

$$\delta_{\mathbf{L}}(\mathbf{L}^*) = \min_{\mathbf{E} \in X} \delta_{\mathbf{L}}(\mathbf{E}).$$

The strategy that we adopt is the same as in other nearest point problems, including the Euclidean algebraic one: for a generic input data $\mathbf{L} \in G(k,n)$, we first look for the critical points of the function $\delta_{\mathbf{L}}$ on the smooth locus of X and then we look for the minimizer among those. Nevertheless, as soon as we start looking at our setting more carefully, we are immediately faced with two issues that complicate our study.

The Grassmann distance is defined in terms of principal angles between k–planes (2.26). In section Section 2.7.4 we have shown explicitly how to compute them and in Section 2.7.5 we discussed their non–semialgebraicity. This implies that also the Grassmann distance is not a semialgebraic function of the input variables; this can also be seen from the non–semialgebraicity of the geodesics of the Grassmann metric described in Section 2.7.4 (see also Section 4.2.6). This is the first issue we have to face: we cannot follow the same path as in Section 4.1, since we do not have the tools of algebraic geometry at our disposal anymore; we have to think of a different approach. For this reason, there is also no need to look at complex critical points, but we will always stick to the realm of real numbers. Instead of algebraic methods, we will make use of tools coming from real geometry and o–minimality theory, as the Grassmann distance function is definable in an appropriate o–minimal structure (Proposition 2.32).

The other issue is that the Grassmann distance from a point $\mathbf{L}$ is not a smooth function on $G(k,n) \setminus \{\mathbf{L}\}$. We will explain how non–smoothness is related to the *cut locus* of a point in $G(k,n)$ being not empty. Because of this, we have to change our usual notion of critical points. We will tackle this issue in Section 4.2.1, where we define critical points following Clarke's theory of subgradients.

Our goal in our study of Problem 4.9 is to understand which of the properties that hold for the Euclidean algebraic setting hold also in this one. In analogy to the list at the end of Section 4.1 our guiding questions will be the following:

1. for a general data point $\mathbf{L} \in G(k,n)$, do we have finitely many critical points of $\delta|_{\mathbf{L}}$ on the smooth part of X? If we do, is their number uniformly bounded?
2. if finite, is there an explicit bound to the number of critical points, in terms of the structure of the defining equations, for some specific class of submanifolds?
3. is there any meaningful analogue of the ED correspondence in this setting?

We will see that the answer to (1) is negative: there exist submanifolds, even algebraic ones, for which the number of critical points of $\delta_{\mathbf{L}}$ on X^{sm} is not finite for a generic $\mathbf{L}$; Section 4.3 gives an example of this. Nevertheless, in Section 4.2.3 we show that the most part of these critical points can be ignored for the solution of Problem 4.9 and in Section 4.2.4 we prove that the meaningful ones are actually finitely many. Moreover, under some assumptions on the nature of the submanifold, we can prove uniform boundedness and define an analogue of the Euclidean Distance Degree, which we call *Grassmann Distance complexity* (Section 4.2.5). As for (2) we will give an example of such a bound in Section 4.2.7.

4.2.1 The subgradient of the Grassmann distance function

As already pointed out in the previous section, when trying to solve the optimization problem Problem 4.9 we are faced with two main issues. On one hand, the Grassmann distance from a fixed point is not smooth; on the other it is not (semi)algebraic. The aim of this section is to address the first of these issues.

We will start by introducing a suitable notion of critical points for the Grassmann distance function from a point through the theory of Clarke's subgradient. We then analyze the locus of points for which the function is not differentiable. Lastly we compute a full description of the subgradient at non–smoothness points.

The subgradient of locally Lipschitz functions

Let (M, g) be a Riemannian manifold and let $f : M \longrightarrow \mathbb{R}$ be a locally Lipschitz function; remark that a Riemannian structure is needed in order to talk about "Lipschitzianity"and gradient of a function. By Rademacher's Theorem, every locally Lipschitz function is differentiable almost everywhere. Therefore the following object is well defined.

Definition 4.10 ([26])**.** The *subgradient* of a locally Lipschitz function $f : M \longrightarrow \mathbb{R}$ at a point p of its domain is defined as

$$\partial_p f := \mathrm{co}\left\{ \lim_{p_n \to p} \nabla f(p_n) \mid (p_n)_n \subseteq \omega(f) \text{ and the limit exists} \right\},$$

where co denotes the convex hull, $\omega(f)$ is the set differentiability points for f and ∇f is the gradient of f.

Note that, if $p \in M$ is a differentiability point for f, i.e. $p \in \omega(f)$, then $\partial_p f = \{\nabla f(p)\}$. It follows that the notion of subgradient directly extends that of gradient for points of non-differentiability.

Remark 4.11. Subgradients are defined by Clarke in greater generality for locally Lipschitz functions defined on Banach spaces [26, Chapter 2]. He proves in [26, Theorem 2.5.1] that when the Banach space is finite dimensional, the subgradient is characterized as in Definition 4.10.

With this new notion available, we can *define* what are critical points of a possibly non–smooth locally Lipschitz function.

Definition 4.12. A point $p \in M$ is a critical point for a locally Lipschitz function $f : M \longrightarrow \mathbb{R}$ if and only if the zero vector belongs to the subgradient

$$0 \in \partial_p f\,.$$

We denote the set of critical points of a locally Lipschitz function f by $\mathrm{Crit}(f)$.

From the observation above, if p is a differentiability point of f, then p is critical if and only if the zero vector belongs to the subgradient, which in this case coincides with the gradient. Therefore p is critical if and only if the gradient of f at p is zero, showing that Definition 4.12 is consistent with the usual notion of critical points for differentiable functions.

Example 4.13. Consider the locally Lipschitz function $f : \mathbb{R} \longrightarrow \mathbb{R}$ given by the absolut value $f(x) = |x|$. This function has a unique point of non–differentiability, namely $x = 0$. As it is evident, this point is a global minimizer of f and any reasonable generalization of the notion of critical point should therefore identify it as critical. It is immediate to see that

$$\partial_0 f = \mathrm{co}\{-1, 1\} = [-1, 1]\,.$$

Since $0 \in \partial_0 f$, the point $x = 0$ is a critical point of the absolute value function f. This example, though it is somewhat naive, should help the reader getting some intuition behind the notion of critical point. Indeed remark that what really matters in this example is that the two limits, 1 and -1, have different signs, meaning that on one side the function is increasing, while on the other it is decreasing. This different behaviour is captured exactly by 0 being in the segment which is the convex hull of the possible limits.

The cut locus of a point

Now that we have a notion of critical points for our non–smooth setting, we can turn to the following question: given a fixed $\mathbf{L} \in G(k, n)$, which are the non–differentiability points of the Grassmann distance function $\delta_{\mathbf{L}} : G(k, n) \longrightarrow \mathbb{R}$? In order to state the basic result concerning the differentiability properties of $\delta_{\mathbf{L}}$, we first recall the following definition.

Definition 4.14 (Cut locus)**.** Let (M, g) be a Riemannian manifold and $p \in M$. The *cut locus* of p, denoted by $\mathrm{cut}(p)$, is the closure of the set of points in M that are connected to p by two or more distinct shortest geodesics.

It is a standard fact in differential geometry that, in a Riemannian manifold (M, g), for a given $p \in M$, the Riemannian distance function from p is smooth away from $\{p\} \cup \mathrm{cut}(p)$. The gradient at a differentiability point x is given by the tangent vector to the unique length minimizing geodesic joining p with x. From this we get the following proposition.

Proposition 4.15. *The distance function $\delta_{\mathbf{L}}$ is smooth on $G(k, n) \setminus (\mathrm{cut}(\mathbf{L}) \cup \mathbf{L})$. Moreover, at a differentiability point $\mathbf{E}$ the gradient of $\delta_{\mathbf{L}}$ is given by the tangent vector to the unique length minimizing geodesic going from $\mathbf{L}$ to $\mathbf{E}$.*

To understand the differentiability locus of $\delta_{\mathbf{L}}$, we thus need to understand the cut locus of a point in $G(k, n)$ with respect to the Grassmann metric. It is known [95] that in this case the cut locus of $\mathbf{L} \in G(k, n)$ consists of the planes $\mathbf{E} \in G(k, n)$ having at least one principal angle with $\mathbf{L}$ equal to $\frac{\pi}{2}$:

$$\mathrm{cut}(\mathbf{L}) = \left\{ \mathbf{E} \in G(k, n) \;\middle|\; \theta_k(\mathbf{L}, \mathbf{E}) = \frac{\pi}{2} \right\}. \tag{4.5}$$

As the following lemma shows, the cut locus of $\mathbf{L}$ is in fact an algebraic variety.

Lemma 4.16. *For any point $\mathbf{L} \in G(k, n)$ the cut locus $\mathrm{cut}(\mathbf{L})$ is a Schubert variety of codimension one.*

Proof. By the properties of principal angles, the condition in (4.5) is equivalent to the fact $\dim(\mathbf{E} \cap \mathbf{L}^{\perp}) \geq 1$. Consider a complete flag where the element of dimension $n - k$ is given by $\mathbf{L}^{\perp}$. We can see that

$$\mathrm{cut}(\mathbf{L}) = \left\{ \mathbf{E} \in G(k, n) \mid \dim(\mathbf{E} \cap \mathbf{L}^{\perp}) \geq 1 \right\}$$

is the Schubert variety given the Young diagram consisting of just one box with respect to the chosen flag. Therefore it is of codimension one (see Section 4.3.1 and [41] for details on Schubert varieties and their representation via Young diagrams). □

The main advantage of writing the cut locus as a Schubert variety as in Lemma 4.16 is that we can use algebraic geometry to stratify it. Its strata are given by

$$\operatorname{cut}(\mathbf{L}) = \bigsqcup_{i=1}^{k} \Omega(j) \tag{4.6}$$

where

$$\Omega(j) = \Big\{ \mathbf{E} \in G(k,n) \mid \dim(\mathbf{E} \cap \mathbf{L}^{\perp}) = j \Big\}. \tag{4.7}$$

Each $\Omega(j)$ is itself the smooth part of a Schubert variety. It is the one associated to the Young diagram given by the partition $(j, \dots, j, 0, \dots, 0)$, where the first j entries are j, with respect to a complete flag containing $\mathbf{L}^{\perp}$ as the $(n-k)$-dimensional element. In particular $\Omega(1)$, i.e. the locus of the planes with intersection with $\mathbf{L}^{\perp}$ of dimension exactly 1, is the smooth locus of $\operatorname{cut}(\mathbf{L})$. Moreover, from the Young diagram description of the strata, we see that the codimension of each $\Omega(j)$ in $G(k,n)$ is j^2. In terms of principal angles, the strata can be described as

$$\Omega(j) = \Big\{ \mathbf{E} \in G(k,n) \mid \theta_k(\mathbf{L},\mathbf{E}) = \cdots = \theta_{k-j+1}(\mathbf{L},\mathbf{E}) = \frac{\pi}{2} > \theta_{k-j}(\mathbf{L},\mathbf{E}) \Big\}.$$

The reason for which the description of the strata as Schubert varieties is particularly convenient is that the tangent space at a smooth point of a Schubert variety can be explicitly described. More precisely, assume we have a simple Schubert variety $\Omega \subseteq G(k,n)$, i.e. a variety defined by the single condition

$$\Omega = \{ \mathbf{E} \in G(k,n) \mid \dim(\mathbf{E} \cap \mathbf{V}_i) \geq r \}$$

for some fixed i–dimensional subspace $\mathbf{V}_i$ and some $r \in \{0, \dots, \min\{k,i\}\}$. Then at a smooth point $\mathbf{E} \in \Omega$ (i.e. a point such that $\dim(\mathbf{E} \cap \mathbf{V}_i) = r$), using the description of tangent vectors as linear homomorphisms, we have [37, Proposition 4.3]

$$T_{\mathbf{E}}\Omega = \{ \varphi : \mathbf{E} \longrightarrow \mathbf{E}^{\perp} \mid \varphi(\mathbf{E} \cap \mathbf{V}_i) \subseteq \mathbf{E}^{\perp} \cap \mathbf{V}_i \}, \tag{4.8}$$

that is, under the isomorphism from Definition 2.22, those maps who preserve the intersection between $\mathbf{E}$ and $\mathbf{V}_i$ still inside $\mathbf{V}_i$. We will discuss this in greater generality in Section 4.3.1. We will need this description in the later sections.

The structure of the subgradient

The following fact is elementary and can be proved exactly as in [47, Proposition 2.1.2].

Proposition 4.17. *Let $\mathbf{L} \in G(k,n)$. For every point $\mathbf{S} \in G(k,n)$ the subgradient of the Grassmann distance function from $\mathbf{L}$ at $\mathbf{S}$ is given by*

$$\partial_{\mathbf{S}} \delta_{\mathbf{L}} = \operatorname{co}\left\{ D_v \exp_{\mathbf{L}}\left(\frac{v}{\|v\|}\right) \;\middle|\; v \in \exp_{\mathbf{L}}^{-1}(\mathbf{S}) \cap R_{\frac{\pi}{2}} \right\} \subseteq T_{\mathbf{S}} G(k,n)\,, \tag{4.9}$$

where $\exp_{\mathbf{L}}$ is the exponential map of $G(k,n)$ at $\mathbf{L}$ and $R_{\frac{\pi}{2}}$ is given by (2.27), identifying tangent vectors and matrices via the isomorphism (2.18). In particular, if $\mathbf{S} \notin (\operatorname{cut}(\mathbf{L}) \cup \mathbf{L})$, then $\partial_{\mathbf{S}} \delta_{\mathbf{L}} = \{ \nabla \delta_{\mathbf{L}}(\mathbf{S}) \}$.

The geometric meaning of Proposition 4.17 is that the subgradient of $\delta_{\mathbf{L}}$ at a point $\mathbf{S} \in G(k,n)$ can be computed as the convex hull of the directions, at $\mathbf{S}$, of the lenght–minimizing geodesics that connect $\mathbf{L}$ to $\mathbf{S}$. It follows that, in order to get a description of the subgradient, we need first to understand the set of preimages $\exp_{\mathbf{L}}^{-1}(\mathbf{S}) \cap R_{\frac{\pi}{2}}$. These are described by the following proposition. Recall that we denote by $\pi : O(n) \longrightarrow G(k,n)$ the quotient map defined in (2.13).

Proposition 4.18. *Let $\Omega(j)$ be the j–th stratum of the cut locus $\mathrm{cut}(\mathbf{L})$ for some $\mathbf{L} \in G(k,n)$ and let $\mathbf{S} \in \Omega(j)$. Fix a representative $\widehat{L} \in \pi^{-1}(\mathbf{L})$. Using the isomorphism given by Definition 2.21, fix an element $A_0 \in \exp_{\mathbf{L}}^{-1}(\mathbf{S}) \cap R_{\frac{\pi}{2}}$. Let $A_0 = U\Sigma V^\top$ be a SVD for A_0 with $\Sigma = \mathrm{diag}\,(\theta_1, \dots, \theta_{k-j}, \frac{\pi}{2}, \dots, \frac{\pi}{2}) \in \mathbb{R}^{(n-k)\times k}$. Then*

$$\exp_{\mathbf{L}}^{-1}(\mathbf{S}) \cap R_{\frac{\pi}{2}} = \left\{ A_W := U\Sigma \begin{pmatrix} \mathbb{1}_{k-j} & 0 \\ 0 & W \end{pmatrix} V^\top \,\middle|\, W \in O(j) \right\}.$$

Moreover, under the isomorphism of Definition 2.21, if $W_1 \neq W_2$ are two distinct orthogonal matrices in $O(j)$, the matrices A_{W_1} and A_{W_2} correspond to different geodesics connecting $\mathbf{L}$ to $\mathbf{S}$, i.e. for every $0 < t < 1$ we have $\exp_{\mathbf{L}}(tv_{A_{W_1}}) \neq \exp_{\mathbf{L}}(tv_{A_{W_2}})$.

Proof. Consider a matrix $A \in \exp_{\mathbf{L}}^{-1}(\mathbf{S})$ with $A \neq A_0$ and with SVD $A = \widetilde{U}\widetilde{\Sigma}\widetilde{V}^\top$. We want to understand the relation between A_0 and A.

Note that all the points in $\exp_{\mathbf{L}}^{-1}(\mathbf{S}) \cap R_{\frac{\pi}{2}}$ will actually be on the boundary $\partial R_{\frac{\pi}{2}}$ by Proposition 2.30.

By the results of Section 2.7.4, we know the following: the columns of $\widetilde{V}$ correspond to the coordinates of the principal vectors of $\mathbf{L}$ (with $\mathbf{S}$) with respect to the basis given by the fixed representative $\widehat{L}$; the matrix $\widetilde{U}$ contains the coordinates of the vectors orthogonal to the principal ones of $\mathbf{L}$ in the planes $\mathfrak{P}$ spanned by the couples of principal vectors (the vectors n_i in (2.24)); the matrix $\widetilde{\Sigma}$ contains the principal angles between $\mathbf{L}$ and $\mathbf{S}$. In particular we must have $\Sigma = \widetilde{\Sigma}$, while $\widetilde{V}$ and $\widetilde{U}$ must correspond to a different choice of principal vectors with respect to V and U. We then have to understand how much freedom do we have when choosing principal vectors and when different choices actually lead to different matrices. Depending on the multiplicity of principal angles, three cases can occur (compare with [9, Proposition B.1]):

- if a principal angle has multiplicity 1, the corresponding pair of principal vectors is uniquely determined and the corresponding columns in U and V will be the same as in $\widetilde{U}$ and $\widetilde{V}$;
- if a principal angle $\mu < \frac{\pi}{2}$ has multiplicity m, the corresponding principal vectors are determined up to an orthogonal transformation in $O(m)$, which has to be the same on both spaces. More explicitly, if $\{p_{i+1}, \dots, p_{i+m}\}$ and $\{q_{i+1}, \dots, q_{i+m}\}$ are principal vectors corresponding to the angle μ, any other choice of principal vectors for the same angle μ, say $\{\widetilde{p}_{i+1}, \dots, \widetilde{p}_{i+m}\}$ and $\{\widetilde{q}_{i+1}, \dots, \widetilde{q}_{i+m}\}$, must be of the form

$$(\widetilde{p}_{i+1} \dots \widetilde{p}_{i+m}) = (p_{i+1} \dots p_{i+m})\ Q$$
$$(\widetilde{q}_{i+1} \dots \widetilde{q}_{i+m}) = (q_{i+1} \dots q_{i+m})\ Q$$

  for some $Q \in O(m)$. The fact that the matrix is the same for both the p's and the q's implies that this different choice of principal vectors in the SVD has no effect, as the matrix Q cancels out when multiplying back $\widetilde{U}$ and $\widetilde{V}$ (in order to obtain the matrix A);

– if an angle of $\frac{\pi}{2}$ has multiplicity j and $\{p_{k-j+1},\dots,p_k\}$ and $\{q_{k-j+1},\dots,q_k\}$ are principal vectors corresponding to that angle, then any vector in $\mathrm{span}\{p_{k-j+1},\dots,p_k\}$ is orthogonal to any vector in $\mathrm{span}\{q_{k-j+1},\dots,q_k\}$. This means that if $\{\widetilde{p}_{k-j+1},\dots,\widetilde{p}_k\}$ and $\{\widetilde{q}_{k-j+1},\dots,\widetilde{q}_k\}$ are another choice of principal vectors corresponding to the angle of $\frac{\pi}{2}$, the orthogonal changes to apply to the original basis to get the new one need not be related one another. In other words

$$(\widetilde{p}_{k-j+1}\dots\widetilde{p}_k) = (p_{k-j+1}\dots p_k)\ W_1$$
$$(\widetilde{q}_{k-j+1}\dots\widetilde{q}_k) = (q_{k-j+1}\dots q_k)\ W_2$$

for some $W_1, W_2 \in O(j)$ arbitrary. This change now produces a matrix A which is in general different from the original A_0. Indeed, denoting by $\Theta = \mathrm{diag}(\theta_1,\dots,\theta_{k-j}) \in \mathbb{R}^{(k-j)\times(k-j)}$ the matrix of principal angles which are not $\frac{\pi}{2}$, we have

$$A = \widetilde{U}\Sigma\widetilde{V} = U\begin{pmatrix} \mathbb{1}_{k-j} & 0 & 0 \\ 0 & W_1 & 0 \\ 0 & 0 & \mathbb{1}_{n-2k} \end{pmatrix}\begin{pmatrix} \Theta & 0 \\ 0 & \frac{\pi}{2}\mathbb{1}_j \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \mathbb{1}_{k-j} & 0 \\ 0 & W_2^\top \end{pmatrix}V^\top =$$
$$= U\Sigma\begin{pmatrix} \mathbb{1}_{k-j} & 0 \\ 0 & W_1W_2^\top \end{pmatrix}V^\top.$$

Remark that above formula implies that what matters in order to have a different matrix $A \neq A_0$ are not the specific orthogonal changes W_1 and W_2, but rather the relative change in the principal vectors given by $W_1W_2^\top$.

It follows from this that any $A \in \partial R_{\frac{\pi}{2}}$ such that $\exp_{\mathbf{L}}(v_A) = \mathbf{S}$ must be of the form

$$A = U\Sigma\begin{pmatrix} \mathbb{1}_{k-j} & 0 \\ 0 & W \end{pmatrix}V^\top$$

for some $W \in O(j)$, where $A_0 = U\Sigma V^\top$ is a fixed matrix in $\partial R_{\frac{\pi}{2}}$ such that $\exp_{\mathbf{L}}(v_{A_0}) = \mathbf{S}$. Viceversa any choice $W \in O(j)$ gives a different matrix A.

The second part of the statement is immediate to verify by explicitly writing down the geodesics. We leave this computation to the reader. □

Now that we understand how to describe the vectors involve in (4.9), we proceed with the description of the subgradient. This is slightly more complicated than the formulation of the previous result, because we need to apply the differential of the exponential map.

Theorem 4.19. *Let $\Omega(j)$ be the j–th stratum of the cut locus $\mathrm{cut}(\mathbf{L})$ for some $\mathbf{L} \in G(k,n)$ and let $\mathbf{S} \in \Omega(j)$. Fix a representative $\widehat{S} \in \pi^{-1}(\mathbf{S})$. Using the isomorphism given in Definition 2.21, pick an element $B_0 \in \exp_{\mathbf{S}}^{-1}(\mathbf{L}) \cap R_{\frac{\pi}{2}}$. Let $U\Sigma V^\top$ be an SVD for B_0 and define the matrix*

$$C := \begin{pmatrix} V & 0 \\ 0 & U\Sigma \end{pmatrix},$$

and, for every $W \in \mathbb{R}^{j\times j}$, the skew–symmetric matrix

$$\overline{W} := \begin{pmatrix} 0 & -\begin{pmatrix} \mathbb{1}_{k-j} & 0 \\ 0 & W \end{pmatrix} \\ \begin{pmatrix} \mathbb{1}_{k-j} & 0 \\ 0 & W^\top \end{pmatrix} & 0 \end{pmatrix}.$$

Then the subgradient of the distance function from $\mathbf{L}$ at the point $\mathbf{S}$ is given by

$$\partial_{\mathbf{S}}\delta_{\mathbf{L}} = -\frac{1}{\delta(\mathbf{L},\mathbf{S})}\, D_{\widehat{S}}\pi\Big(\widehat{S}\,C\,\{\overline{W}\,|\,W \text{ has singular values} \leq 1\}\,C^\top\Big) \subseteq T_{\mathbf{S}}G(k,n)\,. \quad (4.10)$$

Proof. Proposition 4.18 gives a description of all the tangent vectors in $T_{\mathbf{L}}G(k,n)$ corresponding to length minimizing geodesics connecting $\mathbf{L}$ to $\mathbf{S}$ for any $\mathbf{S} \in \mathrm{cut}(\mathbf{L})$. The expression (4.9) for a point $\mathbf{S} \in \Omega(j)$ thus becomes

$$\partial_{\mathbf{S}}\delta_{\mathbf{L}} = \mathrm{co}\left\{ D_{v_{A_W}} \exp_{\mathbf{L}}\left(\frac{v_{A_W}}{\|v_{A_W}\|}\right) \;\middle|\; W \in O(j) \right\}, \tag{4.11}$$

where A_W is defined as in Proposition 4.18.

Denote the length minimizing geodesic corresponding to the tangent vector $v_{A_W} \in T_{\mathbf{L}}G(k,n)$ as γ_W, i.e. $\gamma_W(t) := \exp_{\mathbf{L}}(tv_{A_W})$, $\gamma_W(0) = \mathbf{L}$ and $\gamma_W(1) = \mathbf{S}$. Then we can compute the image vectors of the differential in (4.11) as follows. Define the curve $\sigma : (-\varepsilon, \varepsilon) \longrightarrow T_{\mathbf{L}}G(k,n)$ given by

$$\sigma(t) := v_{A_W} + t\frac{v_{A_w}}{\|v_{A_W}\|}.$$

Then $\sigma(0) = v_{A_W}$ and $\dot\sigma(0) = \frac{v_{A_W}}{\|v_{A_W}\|}$ so that σ is an adapted to curve to $\frac{v_{A_W}}{\|v_{A_W}\|} \in T_{v_{A_W}}(T_{\mathbf{L}}G(k,n)) \cong T_{\mathbf{L}}G(k,n)$. We can compute the differential using the adapted curve σ as

$$\begin{aligned}
D_{v_{A_W}} \exp_{\mathbf{L}}\left(\frac{v_{A_W}}{\|v_{A_W}\|}\right) &= \frac{d}{dt}\exp_{\mathbf{L}}(\sigma(t))\Big|_{t=0} = \frac{d}{dt}\exp_{\mathbf{L}}\left(v_{A_W} + t\frac{v_{A_w}}{\|v_{A_W}\|}\right)\Big|_{t=0} \\
&= \frac{d}{dt}\exp_{\mathbf{L}}\left(\left(1 + \frac{t}{\|v_{A_W}\|}\right)v_{A_W}\right)\Big|_{t=0} \\
&= \frac{d}{dt}\gamma_W\left(1 + \frac{t}{\|v_{A_W}\|}\right)\Big|_{t=0} \\
&= \frac{1}{\|v_{A_W}\|}\,\dot\gamma_W(1),
\end{aligned}$$

which is the tangent vector to the geodesic γ_W at the point $\mathbf{S}$, rescaled by the its constant speed. The subgradient (4.11) can therefore be written as

$$\partial_{\mathbf{S}}\delta_{\mathbf{L}} = \mathrm{co}\left\{ \frac{1}{\|v_{A_W}\|}\,\dot\gamma_W(1) \;\middle|\; W \in O(j) \right\}. \tag{4.12}$$

Now we invert the motion along the geodesics to go from $\mathbf{S}$ to $\mathbf{L}$. Define $\widetilde\gamma_W(t) := \gamma_W(1-t)$, so that $\widetilde\gamma_W(0) = \mathbf{S}$, $\widetilde\gamma_W(1) = \mathbf{L}$ and $\dot{\widetilde\gamma}_W(t) = -\dot\gamma_W(1-t)$, which implies that $\widetilde\gamma_W$ is a lenght minimizing geodesic connecting $\mathbf{S}$ to $\mathbf{L}$. Remark that the speed of $\widetilde\gamma_W$ is the same as that of γ_W. Therefore $\dot\gamma_W(1) = -\dot{\widetilde\gamma}_W(0)$ and the subdifferential (4.12) can be described as the convex hull of the tangent vectors in $T_{\mathbf{S}}G(k,n) \cap R_{\frac{\pi}{2}}$ corresponding to length minimizing geodesics from $\mathbf{S}$ to $\mathbf{L}$, rescaled by their norm and multiplied by -1.

The computation of such vectors is completely analogous to that we carried out in Proposition 4.18. For every $W \in O(j)$ define

$$B_W := U\Sigma\begin{pmatrix} \mathbb{1}_{k-j} & 0 \\ 0 & W^\top \end{pmatrix}V^\top.$$

By Proposition 4.18 we have

$$\exp_{\mathbf{S}}^{-1}(\mathbf{L}) \cap R_{\frac{\pi}{2}} = \{v_{B_W} \mid W \in O(j)\}.$$

Indeed it is immediate to see that if $\mathbf{S}$ is in the j-th stratum of $\mathrm{cut}(\mathbf{L})$, viceversa $\mathbf{L}$ is in the j-th stratum of $\mathrm{cut}(\mathbf{S})$ by symmetry of principal angles. Remark that

$$\|v_{A_W}\| = \delta(\mathbf{L}, \exp_{\mathbf{L}}(v_{A_W})) = \delta(\mathbf{L}, \mathbf{S}),$$

so that (4.12) becomes

$$\partial_{\mathbf{S}}\delta_{\mathbf{L}} = -\frac{1}{\delta(\mathbf{L},\mathbf{S})}\ \mathrm{co}\{v_{B_W} \mid W \in O(j)\}.$$

Recall that, by the construction of the isomorphism from Definition 2.21, we have

$$v_B = D_{\widehat{S}}\pi\left(\widehat{S}\left(\begin{array}{c|c} 0 & -B^\top \\ \hline B & 0 \end{array}\right)\right) = D_{\widehat{S}}\pi(\widehat{S}M_B),$$

where the matrix M_B is defined in (2.16). Using the linearity of the differential we obtain

$$\begin{aligned}\partial_{\mathbf{S}}\delta_{\mathbf{L}} &= -\frac{1}{\delta(\mathbf{L},\mathbf{S})}\ \mathrm{co}\Big\{D_{\widehat{S}}\pi(\widehat{S}M_{B_W}) \mid W \in O(j)\Big\} = \\ &= -\frac{1}{\delta(\mathbf{L},\mathbf{S})}D_{\widehat{S}}\pi\Big(\widehat{S}\ \mathrm{co}\{M_{B_W} \mid W \in O(j)\}\Big).\end{aligned}$$

Given the definition of B_W, we can check that

$$M_{B_W} = C\,\overline{W}\,C^\top.$$

Therefore, as the matrix C is fixed and does not depend on $W \in O(j)$, we obtain

$$\mathrm{co}\{M_{B_W} \mid W \in O(j)\} = C\ \mathrm{co}\big\{\overline{W} \mid W \in O(j)\big\}\ C^\top.$$

To understand the convex hull of the matrices $\overline{W}$, it is enough to look at the convex hull of the matrices

$$\begin{pmatrix} \mathbb{1}_{k-j} & 0 \\ 0 & W \end{pmatrix}, \quad W \in O(j).$$

Pick $\lambda_1, \ldots, \lambda_r \geq 0$ such that $\lambda_1 + \cdots + \lambda_r = 1$ and $W_1, \ldots, W_r \in O(j)$. Then

$$\sum_{i=1}^{r}\lambda_i\begin{pmatrix} \mathbb{1}_{k-j} & 0 \\ 0 & W_i \end{pmatrix} = \sum_{i=1}^{r}\begin{pmatrix} \lambda_i\mathbb{1}_{k-j} & 0 \\ 0 & \lambda_i W_i \end{pmatrix} = \begin{pmatrix} \mathbb{1}_{k-j} & 0 \\ 0 & \sum_{i=1}^{r}\lambda_i W_i \end{pmatrix},$$

from which we obtain

$$\mathrm{co}\left\{\begin{pmatrix} \mathbb{1}_{k-j} & 0 \\ 0 & W \end{pmatrix} \;\middle|\; W \in O(j)\right\} = \left\{\begin{pmatrix} \mathbb{1}_{k-j} & 0 \\ 0 & B \end{pmatrix} \;\middle|\; B \in \mathrm{co}(O(j))\right\}.$$

It is well–known that the convex hull of the orthogonal group is the set of matrices with operator norm smaller or equal than 1 (see [55, Section 3.4] and [62, Theorem 2.1]). This condition on operator norm is the same as asking for the singular values to be bounded above by 1. The proof is concluded. □

As a corollary, we deduce the following geometric description of the subgradient of $\delta_{\mathbf{L}}$, highlighting its most relevant aspects.

Corollary 4.20. *For every $\mathbf{L} \in G(k,n)$ and $\mathbf{S} \in \Omega(j) \subseteq \mathrm{cut}(\mathbf{L})$ with $j \in \{1,\ldots,k\}$, the subgradient of the distance function from $\mathbf{L}$ at $\mathbf{S}$ is a convex set linearly isomorphic to the convex hull of the orthogonal group $O(j)$. In particular it is of dimension j^2.*

Proof. We use the same notation as in Theorem 4.19. We define the matrix

$$Y_B := U\Sigma\begin{pmatrix} \mathbb{1}_{k-j} & 0 \\ 0 & B^\top \end{pmatrix}V^\top \in \mathbb{R}^{(n-k)\times k}$$

for any $B \in \mathrm{co}\{O(j)\}$. Then by the expression in (4.10) we obtain

$$\dim(\partial_{\mathbf{S}}\delta_{\mathbf{L}}) = \dim\left(D_{\widehat{S}}\pi\left(\left\{\widehat{S}\begin{pmatrix} 0 & -Y_B^\top \\ Y_B & 0 \end{pmatrix} \mid B \in \mathrm{co}\{O(j)\}\right\}\right)\right). \tag{4.13}$$

By (2.15) we have that the matrices in (4.13) to which we are applying the differential $D_{\widehat{S}}\pi$ lie in the orthogonal to complement to its kernel. It follows that $D_{\widehat{S}}\pi$ preserves the dimension of such set. Since the matrix $\widehat{S}$ is orthogonal, we obtain

$$\dim(\partial_{\mathbf{S}}\delta_{\mathbf{L}}) = \dim\left(\left\{\begin{pmatrix} 0 & -Y_B^\top \\ Y_B & 0 \end{pmatrix} \;\middle|\; B \in \mathrm{co}\{O(j)\}\right\}\right) = \dim(\{Y_B \mid B \in \mathrm{co}\{O(j)\}\}).$$

Since U and V are orthogonal transformations, we only need to check how the diagonal matrix $\Sigma = \mathrm{diag}(\theta_1, \ldots, \theta_{k-j}, \frac{\pi}{2}, \ldots, \frac{\pi}{2})$ acts by multiplication. It is easy to see that given the block structure of the product defining Y_B, the possible presence of 0 entries in Σ, corresponding to a possible non-trivial intersection between $\mathbf{S}$ and $\mathbf{L}$, does not affect the part of the product involving the matrix B. As a consequence

$$\dim(\partial_{\mathbf{S}}\delta_{\mathbf{L}}) = \dim(\{Y_B \mid B \in \mathrm{co}\{O(j)\}\}) = \dim(\mathrm{co}\{O(j)\}) = j^2.$$

□

4.2.2 Restricting the Grassmann distance to a submanifold

In the previous section we computed the subgradient of the Grassmann distance function from a point in $G(k,n)$. What we are now interested in is understanding how the subgradient behaves when we restrict the function $\delta_{\mathbf{L}}$ to a smooth submanifold $X \subseteq G(k,n)$. As we will see, in order to have a neat description of the subgradient of the restricted function, we will need to impose some nondegeneracy conditions on the intersection between the submanifold and the cut locus of $\mathbf{L}$.

Definition 4.21. We say that a smooth submanifold $X \subseteq G(k,n)$ is transverse to the cut locus $\mathrm{cut}(\mathbf{L})$ if it is transverse to each stratum $\Omega(j)$ of the stratification given in (4.6).

We begin our discussion with the following easy observation.

Lemma 4.22. *Let $X \subseteq G(k,n)$ be a smooth submanifold. Then, for a generic point $\mathbf{L} \in G(k,n)$, the manifold X is transverse to the cut locus* $\mathrm{cut}(\mathbf{L})$.

Proof. Let $g \in O(n)$ be such that $\mathbf{L} = g^{-1} \cdot \mathbf{L}_0$ for some $\mathbf{L}_0 \in G(k,n)$. Since the action of $O(n)$ on $G(k,n)$ is by diffeomorphisms which are also isometries for the Grassmann metric, we have that X is transverse to $\mathrm{cut}(\mathbf{L})$ if and only if $g \cdot X$ is transverse to $\mathrm{cut}(\mathbf{L}_0)$. It is thus sufficient to check the transversality condition for $g \cdot X$ with respect to $\mathrm{cut}(\mathbf{L}_0)$ for the generic $g \in O(n)$. Consider the map

$$\begin{aligned} F : O(n) \times X &\longrightarrow G(k,n) \\ (g, \mathbf{S}) &\longmapsto g \cdot \mathbf{S}. \end{aligned}$$

Since the projection map $O(n) \longrightarrow G(k,n)$ is a (Riemannian) submersion, the $O(n)$ part is already enough to give the surjectivity of the differential of F. It follows that the map F is transverse to $\mathrm{cut}(\mathbf{L}_0)$. By the Parametric Transversality Theorem [49, Chapter 3, Theorem 2.7], it follows that the map $F(g,\cdot) : X \longrightarrow G(k,n)$ is transverse to $\mathrm{cut}(\mathbf{L}_0)$ for almost every $g \in O(n)$, which is equivalent to $g \cdot X$ being transverse to $\mathrm{cut}(\mathbf{L}_0)$.

□

Next, we use the genericity assumptions on $\mathbf{L} \in G(k,n)$ to prove the following lemma, which is an analogue of [47, Proposition 2.1.10]. There the authors deal with the Euclidean distance function. We will follow along the steps of their proof and adapt it to the Grassmann distance case we are interested in.

Lemma 4.23. *Let $X \subseteq G(k,n)$ be a smooth submanifold. For a generic $\mathbf{L} \in G(k,n)$ and every $\mathbf{S} \in X$, we have*

$$\partial_{\mathbf{S}}(\delta_{\mathbf{L}}|_X) = \mathrm{proj}_{T_{\mathbf{S}}X}(\partial_{\mathbf{S}}\delta_{\mathbf{L}}),$$

where $\delta_{\mathbf{L}}|_X$ is the restriction of the distance function from $\mathbf{L}$ to X and $\mathrm{proj}_{T_{\mathbf{S}}X}$ is the orthogonal projection onto the tangent space to X at $\mathbf{S}$.

Proof. Remark that since $\mathbf{L}$ is just a point, X is trivially transverse to $\mathbf{L}$. By **??**, for the generic $\mathbf{L}$, the manifold X is also transverse to $\mathrm{cut}(\mathbf{L})$. The definition of subgradient (4.9) in this case reads as

$$\partial_{\mathbf{S}}(\delta_{\mathbf{L}}|_X) = \mathrm{co}\left\{ \lim_{\mathbf{S}_n \to \mathbf{S}} \nabla\delta_{\mathbf{L}}|_X(\mathbf{S}_n) \;\middle|\; (\mathbf{S}_n) \subseteq \omega(\delta_{\mathbf{L}}|_X) \text{ and the limit exists} \right\}. \tag{4.14}$$

By Proposition 4.15, we have $\omega(\delta_{\mathbf{L}}) = G(k,n) \setminus (\mathbf{L} \cup \mathrm{cut}(\mathbf{L}))$. Since $\mathrm{cut}(\mathbf{L})$ has codimension 1, by transversality the intersection $X \cap (\mathbf{L} \cup \mathrm{cut}(\mathbf{L}))$ has codimension at least 1 in X. Moreover $\omega(\delta_{\mathbf{L}}) \cap X \subseteq \omega(\delta_{\mathbf{L}}|_X)$ since, if $\delta_{\mathbf{L}}$ is differentiable at a point $\mathbf{E} \in X$, so is its restriciton to X.

These two facts together imply that any point in $\omega(\delta_{\mathbf{L}}|_X)$ can be approached by a sequence of points in $\omega(\delta_{\mathbf{L}}) \cap X$. As a consequence we obtain

$$(4.14) = \mathrm{co}\left\{ \lim_{\mathbf{S}_n \to \mathbf{S}} \nabla\delta_{\mathbf{L}}|_X(\mathbf{S}_n) \;\middle|\; (\mathbf{S}_n) \subseteq \omega(\delta_{\mathbf{L}}) \cap X \text{ and the limit exists} \right\}. \tag{4.15}$$

At any differentiability point $\mathbf{E} \in \omega(\delta_{\mathbf{L}}) \cap X$ we have that

$$\nabla\delta_{\mathbf{L}}|_X(\mathbf{E}) = \mathrm{proj}_{T_{\mathbf{E}}X}(\nabla\delta_{\mathbf{L}}(\mathbf{E})),$$

since the gradient of the restriction of a function to a submanifold is the projection of the gradient of the function on the tangent space to the submanifold. Since the projection map $(\mathbf{E}, v) \longmapsto \mathrm{proj}_{T_{\mathbf{E}}X}(v)$ is continuous, we obtain

$$\begin{aligned}
(4.15) &= \mathrm{co}\left\{ \lim_{\mathbf{S}_n \to \mathbf{S}} \mathrm{proj}_{T_{\mathbf{S}_n}X}(\nabla\delta_{\mathbf{L}}(\mathbf{S}_n)) \;\middle|\; (\mathbf{S}_n) \subseteq \omega(\delta_{\mathbf{L}}) \cap X \text{ and the limit exists} \right\} \\
&= \mathrm{co}\left\{ \lim_{\mathbf{S}_n \to \mathbf{S}} \mathrm{proj}_{T_{\mathbf{S}}X}(\nabla\delta_{\mathbf{L}}(\mathbf{S}_n)) \;\middle|\; (\mathbf{S}_n) \subseteq \omega(\delta_{\mathbf{L}}) \cap X \text{ and the limit exists} \right\} \\
&= \mathrm{proj}_{T_{\mathbf{S}}X}\left(\mathrm{co}\left\{ \lim_{\mathbf{S}_n \to \mathbf{S}} \nabla\delta_{\mathbf{L}}(\mathbf{S}_n) \;\middle|\; (\mathbf{S}_n) \subseteq \omega(\delta_{\mathbf{L}}) \cap X \text{ and the limit exists} \right\} \right),
\end{aligned} \tag{4.16}$$

where the last equality follows by linearity of the projection.

Now let $U_{\mathbf{S}} \subseteq G(k,n)$ be a small neighborhood of $\mathbf{S}$ and consider the disjoint union of connected components

$$U_{\mathbf{S}} \setminus (\mathbf{L} \cup \mathrm{cut}(\mathbf{L})) = \bigsqcup_{i=1}^{r} A_i.$$

The function $\delta_{\mathbf{L}}$ is smooth on each A_i and we can define

$$V_i := \left\{ \lim_{\mathbf{S}_n \to \mathbf{S}} \nabla \delta_{\mathbf{L}}(\mathbf{S}_n) \;\middle|\; (\mathbf{S}_n) \subseteq A_i \right\}$$

for $i = 1, \ldots, r$. For $U_{\mathbf{S}}$ small enough we have that

$$\partial_{\mathbf{S}} \delta_{\mathbf{L}} = \operatorname{co} \{V_1, \ldots, V_r\}.$$

By transversality, for every $i = 1, \ldots, r$ the intersection $X \cap A_i$ is non empty and there exists a sequence $(\mathbf{S}_n)_n \subseteq X \cap A_i$ that converges to $\mathbf{S}$. It follows that

$$\begin{aligned}
&\operatorname{co} \left\{ \lim_{\mathbf{S}_n \to \mathbf{S}} \nabla \delta_{\mathbf{L}}(\mathbf{S}_n) \;\middle|\; (\mathbf{S}_n) \subseteq \omega(\delta_{\mathbf{L}}) \cap X \text{ and the limit exists} \right\} \\
&= \operatorname{co} \left\{ \lim_{\mathbf{S}_n \to \mathbf{S}} \nabla \delta_{\mathbf{L}}(\mathbf{S}_n) \;\middle|\; (\mathbf{S}_n) \subseteq \bigsqcup_{i=1}^{r} A_i \cap X \text{ and the limit exists} \right\} \\
&= \operatorname{co} \left\{ \operatorname{proj}_{T_{\mathbf{S}} X}(V_1 \cup \cdots \cup V_r) \right\} = \operatorname{proj}_{T_{\mathbf{S}} X}(\operatorname{co} \{V_1, \ldots, V_r\}) \\
&= \operatorname{proj}_{T_{\mathbf{S}} X}(\partial_{\mathbf{S}} \delta_{\mathbf{L}}).
\end{aligned}$$

As a consequence, from (4.16) we obtain that the subgradient is given by

$$\partial_{\mathbf{S}}(\delta_{\mathbf{L}}|_X) = \operatorname{proj}_{T_{\mathbf{S}} X}(\partial_{\mathbf{S}} \delta_{\mathbf{L}}),$$

and the proof is finished. □

The geometric meaning of Lemma 4.23 is that to compute the subgradient of the restriction of $\delta_{\mathbf{L}}$ to a smooth submanifold, we first compute the subgradient of the non–restricted function and then we project this convex set to the tangent space to the submanifold. This generalizes the standard result for gradients of restrictions of smooth functions. Recalling our Definition 4.12 of critical points, from this result it immediately follows a characterization of the critical points of the restricted distance function.

Corollary 4.24. *Let X be a smooth submanifold of $G(k, n)$. For a generic $\mathbf{L} \in G(k, n)$, a point $\mathbf{S} \in X$ is critical for the restriction of the Grassmann distance function from $\mathbf{L}$ to X, $\delta_{\mathbf{L}}|_X$, if and only if $N_{\mathbf{S}} X \cap \partial_{\mathbf{S}} \delta_{\mathbf{L}} \neq \emptyset$.*

Proof. A point $\mathbf{S} \in X$ is critical for $\delta_{\mathbf{L}}|_X$ if and only if $0 \in \partial_{\mathbf{S}} \delta_{\mathbf{L}}|_X$, by definition. By Lemma 4.23 this condition means that the 0 vector has to belong to the projection of the subgradient $\partial_{\mathbf{S}} \delta_L$ to the tangent space to X at $\mathbf{S}$. Since the only vectors being mapped to 0 in this projection are the normal vectors in $N_{\mathbf{S}} X$, this condition is equivalent to having a non–empty intersection between the normal space $N_{\mathbf{S}} X$ and the subgradient $\partial_{\mathbf{S}} \delta_{\mathbf{L}}$. □

In the same way that Lemma 4.23 generalized the usual result for smooth functions, the characterization of critical points in Corollary 4.24 can be seen as a generalization of the usual Lagrange multipliers rule to our non–smooth setting.

4.2.3 Ignoring the cut locus

After having discussed how to deal with the non–smoothness of the Grassmann distance function from a point, let us go back to the nearest point problem Problem 4.9. The structure of the subgradient of $\delta_{\mathbf{L}}$ complicates the study of critical points of $\delta_{\mathbf{L}}|_X$, even for a generic $\mathbf{L}$ and a smooth submanifold $X \subseteq G(k, n)$. Indeed, Corollary 4.24 shows that finding critical points requires to find for which points the normal space to the submanifold

and the convex body given by the subgradient intersect non–trivially. However, in this section we show that, if we are truly interested in the minimization problem, this becomes again a standard minimization problem with smooth constraint (for instance, one can use the Lagrange multiplier's rule). This is accomplished by showing that certain critical points can be a priori excluded from consideration, as they do not correspond to (local) minima in the distance minimization problem.

More precisely, the next result shows that whenever the transversality condition holds, the minimum of the distance function $\delta_{\mathbf{L}}|_X$ cannot be located on the intersection $\mathrm{cut}(\mathbf{L}) \cap X$. As this is the locus where nondifferentiability points of $\delta_{\mathbf{L}}|_X$ can occur, this reduces the problem to the standard smooth case. Remark that this need not be true without the transversality hypothesis; this is the case for example of smooth submanifolds entirely contained in the cut locus.

Theorem 4.25. *Let X be a smooth submanifold of $G(k,n)$ and $\mathbf{L} \in G(k,n)$ be a point such that X is transverse to the cut locus $\mathrm{cut}(\mathbf{L})$. Then no point in the intersection $X \cap \mathrm{cut}(\mathbf{L})$ can be a (local) minimum for the function $\delta_{\mathbf{L}}|_X$.*

Proof. Let $\mathbf{E} \in X \cap \mathrm{cut}(\mathbf{L})$ and assume in particular that $\mathbf{E}$ is in the j–th stratum Ω_j of the cut locus of $\mathbf{E}$; remark that X need not intersect all the strata of $\mathrm{cut}(\mathbf{L})$. Denote by $\{p_1, \ldots, p_k\}$ and $\{q_1, \ldots, q_k\}$ some orthonormal basis of principal vectors for $\mathbf{L}$ and $\mathbf{E}$ respectively, with principal angles $0 \leq \theta_1(\mathbf{L},\mathbf{E}) \leq \cdots \leq \theta_{k-j}(\mathbf{L},\mathbf{E}) < \frac{\pi}{2}$ and the last j ones being $\frac{\pi}{2}$. Let $\gamma : [0,1] \longrightarrow G(k,n)$ be the lenght minimizing geodesic such that $\gamma(0) = \mathbf{E}$ and $\gamma(1) = \mathbf{L}$. For $i = 1, \ldots, k-j$ let $n_i \in \mathrm{span}\{q_i, p_i\}$ be the unit vector orthogonal to q_i such that $p_i = q_i \cos(\theta_i) + n_i \sin(\theta_i)$. Then one can check that, under the isomorphism from Definition 2.22, $\dot\gamma(0)$ corresponds to the linear map $\varphi : \mathbf{E} \longrightarrow \mathbf{E}^\perp$ which acts like

$$\begin{aligned}
\varphi(q_1) &= \theta_1\, n_1 \\
&\vdots \\
\varphi(q_{k-j}) &= \theta_{k-j}\, n_{k-j} \\
\varphi(q_{k-j+1}) &= \frac{\pi}{2}\, p_{k-j+1} \\
&\vdots \\
\varphi(q_k) &= \frac{\pi}{2}\, p_k.
\end{aligned}$$

From the description of tangent spaces in (4.8) and the definition in (4.7) and, we have that

$$T_{\mathbf{E}}\Omega(j) = \{\varphi : \mathbf{E} \longrightarrow \mathbf{E}^\perp \mid \varphi(\mathbf{E} \cap \mathbf{L}^\perp) \subseteq \mathbf{E}^\perp \cap \mathbf{L}^\perp\}.$$

By construction we have that $\mathbf{E} \cap \mathbf{L}^\perp = \mathrm{span}\{q_{k-j+1}, \ldots, q_k\}$ and this is mapped to $\mathrm{span}\{p_{k-j+1}, \ldots, p_k\} \subseteq \mathbf{L}$. Therefore $\dot\gamma(0) \notin T_{\mathbf{E}}\mathrm{cut}(\mathbf{L})$, as it does not preserve $\mathbf{L}^\perp$. Since by hypothesis X is transverse to $\Omega(j)$ there exists a vector $v \in T_{\mathbf{E}}X$ such that the scalar product $\langle v, \dot\gamma(0)\rangle > 0$. Let $\sigma : (-\varepsilon, \varepsilon) \longrightarrow X$ be an adapted curve to v, i.e. $\sigma(0) = \mathbf{E}$ and $\dot\sigma(0) = v$. Then for $t > 0$ small enough we have

$$\delta_{\mathbf{L}}(\mathbf{E}) > \delta_{\mathbf{L}}(\sigma(t)),$$

which implies that $\mathbf{E}$ cannot be a local minimum for $\delta_{\mathbf{L}}$ on X.

□

We showed in Lemma 4.22 that for the generic $\mathbf{L} \in G(k,n)$ the transversality condition in Theorem 4.25 holds. As a consequence, we obtain the following corollary.

Corollary 4.26. *Let X be a smooth submanifold of $G(k,n)$. Then for a generic $\mathbf{L} \in G(k,n)$ the minimizer of the distance function $\delta_{\mathbf{L}}|_X$ cannot be on the intersection $X \cap \mathrm{cut}(\mathbf{L})$.*

Remark 4.27. Corollary 4.26 is one of the conceptual key results of our analysis of the nearest point problem Problem 4.9. Indeed, unlike in the algebraic Euclidean setting, it is not true that the number of critical points of $\delta_{\mathbf{L}}|_X$ for the generic $\mathbf{L} \in G(k,n)$ is finite for every smooth submanifold $X \subseteq G(k,n)$. In Section 4.3 we will exhibit an example of a class of Schubert varieties admitting infinitely many global maxima (and therefore critical points) for the generic $\mathbf{L} \in G(k,n)$, such that each of these maxima is on the smooth stratum of the variety. Nevertheless, all these maxima will be on the cut locus of the point $\mathbf{L}$. We will see in Section 4.2.4 that this is not by chance: as soon as we exclude points lying on the cut locus, we generically obtain only a finite amount of critical points. Thanks to Corollary 4.26, we know that only these finitely many points outside of the cut locus will matter for the nearest point problem, as the minimum will be found among them.

4.2.4 Discreteness of critical points

Under the genericity hypothesis for $\mathbf{L} \in G(k,n)$, we have proved in Corollary 4.24 that a point $\mathbf{S} \in X$ is critical for $\delta_{\mathbf{L}}|_X$ if and only if the normal space to the submanifold X at $\mathbf{S}$ meets the subgradient $\partial_{\mathbf{S}}\delta_{\mathbf{L}}$ of the Grassmann distance function from $\mathbf{L}$. If the point $\mathbf{S}$ does not belong to the cut locus of $\mathbf{L}$ and does not coincide with $\mathbf{L}$ itself, then as we already remarked the function $\delta_{\mathbf{L}}$ is smooth at $\mathbf{S}$ and its subgradient coincides with the usual gradient. In this case, it is the tangent vector to the unique length minimizing geodesic connecting $\mathbf{L}$ to $\mathbf{S}$ at $\mathbf{S}$. This gives the following corollary.

Corollary 4.28. *For a smooth submanifold $X \subseteq G(k,n)$ and a generic $\mathbf{L} \in G(k,n)$, if $\mathbf{S} \notin \mathrm{cut}(\mathbf{L})$ and $\mathbf{S} \neq \mathbf{L}$, then $\mathbf{S}$ is critical for $\delta_{\mathbf{L}}|_X$ if and only if the unique length minimizing geodesic from $\mathbf{L}$ to $\mathbf{S}$ meets X orthogonally at $\mathbf{S}$.*

If instead the point $\mathbf{S}$ belongs to the cut locus of $\mathbf{L}$, being critical for the restricted function $\delta_{\mathbf{L}}|_X$ does not imply the existence of a geodesic from $\mathbf{L}$ to $\mathbf{S}$ being normal to the submanifold. Indeed, in this case the subgradient is just the convex hull of the tangent vectors to length minimizing geodesics between $\mathbf{L}$ and $\mathbf{S}$ (see Section 4.2.1). This means that it can intersect the normal space to X at $\mathbf{S}$ even if no geodesic is normal. For example, this phenomenon will appear in the specific case that we will describe in Section 4.3, when dealing with global maxima of the restricted distance function: no geodesic from the external $\mathbf{L}$ to a global maximum will be orthogonal to the submanifold, but all of such points are in the cut locus.

We will now prove that the set of critical points not on the cut locus is generically finite (Theorem 4.30 below). Let X be a smooth submanifold of $G(k,n)$. As critical points not in the cut locus are described by Corollary 4.28, we consider the normal bundle to X, denoted as NX. Consider the normal exponential map, which is the restriction of the exponential map to NX (see Section 2.2):

$$\exp|_{NX} : NX \longrightarrow G(k,n).$$

It is a smooth map, given the smoothness of the exponential map itself. By Sard's Lemma the set of critical values of $\exp|_{NX}$ has measure zero in $G(k,n)$: in other words, a generic $\mathbf{L} \in G(k,n)$ is a regular value of the normal exponential $\exp|_{NX}$. It is then immediate to prove the following.

Lemma 4.29. *Let $\mathbf{L} \in G(k,n)$ be generic. Then the preimage $(\exp|_{NX})^{-1}(\mathbf{L})$ is discrete.*

Proof. By Sard's Lemma $\mathbf{L}$ is a regular value of $\exp|_{NX}$. Observe that the vector bundle NX and the manifold $G(k,n)$ have the same dimension. It follows that $(\exp|_{NX})^{-1}(\mathbf{L})$ is a zero–dimensional submanifold and is thus discrete. □

Theorem 4.30. *Let X be a smooth compact submanifold of $G(k,n)$ and $\mathbf{L} \in G(k,n)$ be generic. Then the set of critical points of the restricted distance function $\delta_{\mathbf{L}}|_X$ on $X \setminus \mathrm{cut}(\mathbf{L})$ is finite. The same holds true if X is subanalytic (possibly non compact).*

Proof. By Corollary 4.28 a critical point $\mathbf{S} \in X \setminus (\mathrm{cut}(\mathbf{L}) \cup \mathbf{L})$ is characterized by the geodesic connecting it to $\mathbf{L}$ being normal to X. Denoting such geodesic as $\gamma_{\mathbf{S}} : [0,1] \longrightarrow G(k,n)$ with $\gamma_{\mathbf{S}}(0) = \mathbf{S}$ and $\gamma_{\mathbf{S}}(1) = \mathbf{L}$, this means that

$$\big(\mathbf{S}, \dot{\gamma}_{\mathbf{S}}(0)\big) \in \exp|^{-1}_{NX}(\mathbf{L})$$

for every such critical point $\mathbf{S}$. Remark that since $\exp_{\mathbf{E}}(R_{\frac{\pi}{2}}) = G(k,n)$ for every $\mathbf{E}$, we can restrict the normal exponential map to normal vectors in $R_{\frac{\pi}{2}}$, which is compact in the tangent space at each point. Therefore if also X is compact, the discreteness in Lemma 4.29 implies that the set of preimages of $\mathbf{L}$ in $\exp^{-1}|_{NX}(\mathbf{L}) \cap R_{\frac{\pi}{2}}$ is finite. Thus only finitely many $\mathbf{S}$ away from $\mathrm{cut}(\mathbf{L})$ can be critical for $\delta_{\mathbf{L}}|_X$.

If now X is not compact but subanalytic, finiteness follow from o–minimality. □

4.2.5 Uniform boundedness: the Grassmann Distance Complexity

In the previous section we have proved that for a generic point $\mathbf{L} \in G(k,n)$ we obtain finitely many critical points outside the cut locus of $\mathbf{L}$, whenever the constraining submanifold X is compact or subanalytic. Nevertheless, thus far we have no control on the number of such points; it might well be unbounded as the generic data point $\mathbf{L}$ varies. If we want to use the number of critical points as a measure of the complexity of the nearest point problem for a given submanifold X, in the same fashion as the Euclidean Distance Degree is conceived, we need to be able to give at least an upper bound to such number of points. The aim of this section is to settle this in the case of subanalytic submanifolds, that is proving that for this class of submanifolds the number of critical points for the generic $\mathbf{L} \in G(k,n)$ is uniformly bounded. This is achieved using the o–minimality of principal angles and the Grassmann distance function proved in Section 2.7.5.

Proposition 4.31. *Let X be a smooth subanalytic submanifold of $G(k,n)$. There exists a constant $C_X \in \mathbb{N}$ such that for a generic $\mathbf{L} \in G(k,n)$ we have*

$$\#\{\mathbf{S} \in X \setminus \mathrm{cut}(\mathbf{L}) \mid \textit{s.t. } \mathbf{S} \textit{ is critical for } \delta_{\mathbf{L}}|_X\} \leq C_X, \tag{4.17}$$

where $\delta_{\mathbf{L}}|_X$ is the restriction to X of the Grassmann distance function from $\mathbf{L}$.

Proof. Define the following correspondence set

$$A = \{(\mathbf{S}, \mathbf{L}) \in X \times G(k,n) \mid \mathbf{S} \notin \mathrm{cut}(\mathbf{L}),\ \mathbf{S} \in \mathrm{Crit}(\delta_{\mathbf{L}}|_X)\},$$

where $\mathrm{Crit}(f)$ denotes the critical points of a locally Lipschitz function f. By Proposition 2.32 and the definability of X, it follows that also the set A is definable in the o–minimal structure of globally subanalytics sets.

Consider the projection map on the second factor $\pi_2 : A \longrightarrow G(k,n)$. As this is continuous and definable, by the trivialization theorem [93, Chapter 9, Theorem 1.2], there exist a partition of the codomain $G(k,n)$ as a finite disjoint union of definable sets

$$G(k,n) = \bigsqcup_{i=1}^{l} M_i \,,$$

definable sets $F_1, \ldots, F_l$ and definable maps $\varphi_i : F_i \times M_i \longrightarrow \pi_2^{-1}(M_i)$ such that the following diagram commutes

$$\begin{array}{ccc} F_i \times M_i & \xrightarrow{\quad\varphi_i\quad} & \pi_2^{-1}(M_i) \\ & {\scriptstyle p_2}\searrow \quad \swarrow{\scriptstyle \pi_2} & \\ & M_i & \end{array} \,,$$

where $p_2 : F_i \times M_i \longrightarrow M_i$ is also the projection on the second factor. In particular for every $\mathbf{L} \in M_i$ we have $\pi_2^{-1}(\mathbf{L}) \cong F_i$, where the isomorphism is subanalytic. By the construction of A, the fiber $\pi_2^{-1}(\mathbf{L})$ describes exactly the points in $X \setminus \mathrm{cut}(\mathbf{L})$ that are critical for $\delta_{\mathbf{L}}|_X$. Therefore by Theorem 4.30 for the generic $\mathbf{L}$ such fiber consists of finitely many points. It follows that, if we define the union of the strata with 0–dimensional fibers

$$\widetilde{M} := \bigcup_{\dim(F_i)=0} M_i,$$

we have that the complement set $G(k,n) \setminus \widetilde{M}$ has codimension at least 1 in $G(k,n)$. For every i such that $\dim(F_i) = 0$ we have a finite cardinality $|F_i| = c_i < +\infty$. Define $C_X := \max\{c_i \mid \dim(F_i) = 0\}$, which is well–defined since we only have finitely many strata. Then for every $\mathbf{L} \in \widetilde{M}$ it holds that

$$\#\{\mathbf{S} \in X \setminus \mathrm{cut}(\mathbf{L}) \mid \mathbf{S} \text{ is critical for } \delta_{\mathbf{L}}|_X\} \leq C_X$$

and the proof is concluded. □

As we anticipated at the beginning of this section, the meaning of the constant C_X in Proposition 4.31 is the following. When we are looking for solutions to our problem of minimizing the function $\delta_{\mathbf{L}}$ for a generic $\mathbf{L}$ restricted to a submanifold $X \subseteq G(k,n)$, we already observed that one way to proceed is to first look for critical points and then find the minimum among them. Theorem 4.25 tells us that we do not need to care about critical points that lie on the cut locus of $\mathbf{L}$. Therefore the constant C_X serves as a bound on the number of points that we will have to check and compare in order to find the minimum. In other words, it bounds the *complexity* of the minimization problem for the submanifold X. In the same spirit that led to the definition of the Euclidean Distance Degree for algebraic subvarieties of $\mathbb{R}^n$, we introduce the following notion.

Definition 4.32 (Grassmann Distance Complexity)**.** Let X be a smooth subanalytic submanifold of $G(k,n)$. We define the *Grassmann Distance Complexity* of X to be the maximum of the left hand side in (4.17). We will denote this as $\mathrm{GDC}(X)$.

Remark 4.33. Suppose that X is a stratified submanifold of $G(k,n)$. As in the Euclidean setting, we will be interested in critical points lying in the smooth stratum of X. We can therefore extend Definition 4.32 to the stratified or singular case by defining $\mathrm{GDC}(X)$ as the Grassmann Distance Complexity of the smooth stratum.

Observe that proof that the Grassmann Distance Complexity is well–defined is not constructive, i.e. it does not produce an explicit bound to the number of critical points. However, in some instances we are able to produce a bound to the GDC of a submanifold using different methods. This is the case of Section 4.2.7, where we study the Grassmann Distance Complexity of smooth algebraic subvarieties in terms of their defining equations via the theory of Pfaffian functions introduced in Section 2.5.

4.2.6 Non-semialgebraicity of the critical equations

We discuss in this section an example showing that the nearest point to an algebraic variety in $G(k,n)$ is genuinely non–semialgebraic, i.e. that the critical points of the Grassmann distance to a point constrained to an algebraic set are, in general, solutions of a system of non–algebraic equations.

For this example, we will work in the *oriented Grassmannian* $\widetilde{G}(2,4)$, i.e. the Grassmannian of oriented 2–planes in $\mathbb{R}^4$. This is a double cover of $G(2,4)$, with the quotient map simply being the map that forgets the orientation. The oriented Grassmannian embeds, via the spherical Plücker embedding, in the unit sphere in $\mathbb{R}^6$ as the quadric

$$\widetilde{G}(2,4) = \left\{ (x,y) \in \mathbb{R}^3 \oplus \mathbb{R}^3 \;\middle|\; \|x\|^2 = \|y\|^2 = 1 \right\} \simeq S^2 \times S^2;$$

this can be realized via a linear change of coordinates with respect to the standard Plücker ones.

The above discussion extends verbatim to the oriented Grassmannian. Up to a multiple, the Grassmann metric is the product metric of the two sphere factors ([59]), therefore the (oriented) Grassmann distance is the function

$$\delta : \widetilde{G}(2,4) \times \widetilde{G}(2,4) \to \mathbb{R},$$

given for $(x,y),(z,w) \in \widetilde{G}(2,4) \simeq S^2 \times S^2$ by

$$\delta\left((x,y),(z,w)\right) = \sqrt{\left(\arccos\langle x,z\rangle\right)^2 + \left(\arccos\langle y,w\rangle\right)^2}\,. \tag{4.18}$$

Let now $e_0 = (1,0,0) \in \mathbb{R}^3$. Then by (4.18), the Riemmanian sphere S_R of radius $0 < R < \frac{\pi}{2}$ centered at $\mathbf{L} := (e_0, e_0)$ is given by

$$S_R = \left\{ (x,y) \in \mathbb{R}^3 \oplus \mathbb{R}^3 \;\middle|\; \|x\|^2 = \|y\|^2 = 1,\ (\arccos x_1)^2 + (\arccos y_1)^2 = R^2) \right\}.$$

(Notice that this *is not* an algebraic set in $\mathbb{R}^6$.) It is easily seen that the cut locus of $\mathbf{L}$ equals

$$\operatorname{cut}(\mathbf{L}) = \left(\{-e_0\} \times S^2\right) \cup \left(S^2 \times \{-e_0\}\right). \tag{4.19}$$

Consider now the semialgebraic family $\{X_\beta\}_{\beta>0} \subseteq \widetilde{G}(2,4)$ of algebraic sets defined by

$$X_\beta := \left\{ (x,y) \in \mathbb{R}^3 \oplus \mathbb{R}^3 \;\middle|\; \|x\|^2 = \|y\|^2 = 1,\ x_1 - \beta y_1 = 0 \right\}. \tag{4.20}$$

Each X_β is a hypersurface in $\widetilde{G}(2,4)$, obtained as a linear slice.

The critical points of the distance function $\delta_{\mathbf{L}}|_{X_\beta}$ on $X_\beta \setminus \mathrm{cut}(\mathbf{L})$ are those points where the gradient of $\delta_{\mathbf{L}}$ is a multiple of the gradient of defining equation of X_β in $S^2 \times S^2$. Denoting by $\alpha : [-1, 1] \to \mathbb{R}$ the function

$$\alpha(w) := \frac{\arccos w}{\sqrt{1 - w^2}},$$

we see that these are the points in X_β such that the following matrix has rank less than 4:

$$M := \begin{pmatrix} x_1 & 0 & \alpha(x_1) & 1 \\ x_2 & 0 & 0 & 0 \\ x_3 & 0 & 0 & 0 \\ 0 & y_1 & \alpha(y_1) & -\beta \\ 0 & y_2 & 0 & 0 \\ 0 & y_3 & 0 & 0 \end{pmatrix}.$$

The determinant of the 4×4 matrix $M^\top M$ vanishes exactly when one of the non–trivial singular values of M is zero, i.e. M has rank at most 3. This determinant equals $(x_2^2 + x_3^2)(y_2^2 + y_3^2)(\alpha(y_1) + \beta\alpha(x_1))^2$. Therefore, the critical points of $\delta_{\mathbf{L}}|_{X_\beta}$ not on the cut locus (4.19) are the points (x, y) in $\mathbb{R}^3 \oplus \mathbb{R}^3$ solving the following system of equations:

$$\|x\|^2 = \|y\|^2 = 1, \quad x_1 = \beta y_1, \quad \frac{\arccos(y_1)}{\sqrt{1 - y_1^2}} + \beta \frac{\arccos(\beta y_1)}{\sqrt{1 - \beta^2 y_1^2}} = 0.$$

From this, one could see that the family of systems that we obtain, depending on the semialgebraic parameter $\beta > 0$, is not a semialgebraic system. The formal proof that this is not semialgebraic would require more work and goes beyond the scope of this work.

4.2.7 Pfaffian bounds

We discuss in this section how to obtain upper bounds on the Grassmann Distance Complexity of a smooth algebraic subvariety, depending explicitly on the format of its defining equations. This will be achieved using known bounds on the topology of the solution set of a systems of Pfaffian equations and inequalities. For the basic results and definitions about Pfaffian systems and inequalities, see Section 2.5 and the references therein.

Recall that in Proposition 2.27 we have described the exponential map of the Grassmann metric in terms of tangent matrices, while in Proposition 2.30 the relation between principal angles and singular values of such matrices was highlighted. From these results and the Pfaffian description of the cosine function given in Example 2.13, it follows that for an algebraic submanifold $X \hookrightarrow G(k, n)$ and a generic point $\mathbf{L} \in G(k, n)$ the critical points of the restricted distance function $\delta_{\mathbf{L}}|_X$ admit a Pfaffian description. We will explain how to obtain such a description through the worked out example of a hypersurface. It will be evident how to generalize it to deal with any algebraic submanifold for which the degrees of the defining equations are known.

Recall the Plücker embedding of Grassmannians into a projective space

$$\iota_P : G(k, n) \hookrightarrow \mathbb{P}(\Lambda^k \mathbb{R}^n) \cong \mathbb{R}\mathrm{P}^L$$

sending a plane $\mathbf{E}$ with orthonormal basis $\{v_1, \ldots, v_k\}$ to the projective class of the wedge product $\iota_P(\mathbf{E}) = [v_1 \wedge \cdots \wedge v_k]$. If $\{e_1, \ldots, e_n\}$ is the standard basis of $\mathbb{R}^n$, for any $w_1 \wedge \cdots \wedge w_k \in \mathbb{R}^{L+1}$ we write

$$w_1 \wedge \cdots \wedge w_k = \sum_{1 \le i_1 < \cdots < i_k \le n} c_{i_1, \ldots, i_k} e_{i_1} \wedge \cdots \wedge e_{i_k}$$

and the collection $(c_{i_1,\dots,i_k})$ define a system of coordinates on $\mathbb{R}^{L+1}$. The Plücker coordinates of a plane $\mathbf{E} \in G(k,n)$ are the coordinates $[c_{i_1,\dots,i_k}]_{1\le i_1<\dots<i_k\le n}$ of its image in the Plücker embedding $\iota_P(\mathbf{E}) \in \mathbb{R}\mathrm{P}^L$. The standard coordinate charts of $\mathbb{R}\mathrm{P}^L$ are $(\mathcal{U}_{j_1\dots j_k}, \psi_{j_1,\dots,j_k})$ given by

$$\mathcal{U}_{j_1\dots j_k} := \{[c_{i_1,\dots,i_k}] \in \mathbb{R}\mathrm{P}^L \mid c_{j_1,\dots,j_k} \neq 0\} \xrightarrow{\psi_{j_1,\dots,j_k}} \mathbb{R}^L$$

where $\psi_{j_1,\dots,j_k}$ divides the coordinates by $c_{j_1,\dots,j_k}$ and forgets the $(j_1,\dots,j_k)$–th one.

Without loss of generality, we assume that $\mathbf{L} = \operatorname{span}\{e_1,\dots,e_k\}$. From Proposition 2.27, for any tangent vector $v_A \in T_{\mathbf{L}}G(k,n) \cong \mathbb{R}^{(n-k)\times k}$ the exponential map is given by

$$\exp_{\mathbf{L}}(v_A) = \begin{bmatrix} \cos(\mu_1)v_1 & \dots & \cos(\mu_k)v_k \\ \sin(\mu_1)u_1 & \dots & \sin(\mu_k)u_k \end{bmatrix}, \tag{4.21}$$

where $A = U\Sigma V^\top$ is a singular values decomposition of A, $U = (u_1 \dots u_{n-k}) \in O(n-k)$, $V = (v_1 \dots v_k) \in O(k)$ and $\Sigma = \operatorname{diag}(\mu_1 \dots \mu_k) \in \mathbb{R}^{(n-k)\times k}$. Remark that, in (4.21), only the first k columns of U actually matter, so we can use the reduced singular values decomposition and assume that U is in the Stiefel manifold of orthonormal k–frames in $\mathbb{R}^{n-k}$, denoted $\mathrm{St}(k,n-k)$.

From Proposition 2.30 we also have that

$$G(k,n) \setminus \operatorname{cut}(\mathbf{L}) = \exp_{\mathbf{L}}(\operatorname{int}(R_{\frac{\pi}{2}})),$$

where we recall that $\operatorname{int}(R_{\frac{\pi}{2}})$ is the set of matrices in the tangent space $T_{\mathbf{L}}G(k,n)$ with singular values strictly less than $\frac{\pi}{2}$. Indeed, if the singular values of the matrices we exponentiate are strictly less than $\frac{\pi}{2}$, so will be all the principal angles between $\mathbf{L}$, which is the center of the exponential map, and the images that we obtain. Hence we obtain all and only the points with no principal angle of $\frac{\pi}{2}$ with $\mathbf{L}$, that is exactly the complement of the cut locus of $\mathbf{L}$ by (4.5). Moreover the exponential $\exp_{\mathbf{L}}$ is a diffeomorophism onto its image when restricted to this set. In the following we will make use of this easy lemma.

Lemma 4.34. *The image in the Plücker embedding of the complement $G(k,n) \setminus \operatorname{cut}(\mathbf{L})$ is contained in the coordinate chart $\mathcal{U}_{1,\dots,k}$ for $\mathbf{L} = \operatorname{span}\{e_1,\dots,e_k\}$.*

Proof. Every point of $G(k,n) \setminus \operatorname{cut}(\mathbf{L})$ can be written as in (4.21) for some $A \in \operatorname{int}(R_{\frac{\pi}{2}})$; in particular $0 \le \mu_i < \frac{\pi}{2}$ for every $i = 1,\dots,k$. The Plücker coordinate with multi–index $(1,\dots,k)$ of (4.21) is given by $\cos(\mu_1)\dots\cos(\mu_k)\det(V)$. This is always non–zero, since $V \in O(k)$ and $\cos(\mu_i) \neq 0$ for every $i = 1,\dots,k$. This proves the statement. □

Let now $X \hookrightarrow G(k,n)$ be a smooth algebraic hypersurface of degree d, in the sense of [44, Ch. 3, Sect. A]. By [44, Prop. 2.1, Chapter 3] there exists a homogeneous polynomial p of degree d in the Plücker coordinates such that X is the vanishing locus of p,

$$X = \{p = 0\}.$$

We will now prove the following upper bound to the Grassmann Distance Complexity GDC of X.

Theorem 4.35. *For every $0 \le k \le n$ there exist constants $c_1(k,n), c_2(k,n) > 0$ depending on k and n only, such that for every smooth, degree d hypersurface $X \hookrightarrow G(k,n)$, the following bound holds:*

$$\mathrm{GDC}(X) \le c_1(k,n) d^{c_2(k,n)}.$$

The constants $c_1(k,n)$ and $c_2(k,n)$ above are given explicitly, respectively, by (4.32) and (4.30) from the following proof.

Proof. The proof consists of several steps. First we perform a preliminary reduction (a "change of variables") so that we reduce to study the set of critical points of a new function f defined on a new set Y, in such a way that these critical points admits a semi–Pfaffian description; then we bound the topology of this set of critical points, call it S, using Theorem 2.15; finally we show that each critical point of the restricted distance function $\delta_{\mathbf{L}}|_X$ from a generic $\mathbf{L} \in G(k,n)$ lying in $X \setminus \mathrm{cut}(\mathbf{L})$ corresponds to at most one connected component of S and, therefore, the bound we have obtained for the Betti numbers of S also serves as a bound for $\mathrm{GDC}(X)$.

First step: a preliminary reduction

We need to describe, given a generic $\mathbf{L} \notin X$, the critical points of the function $\delta_{\mathbf{L}}|_X$ lying in $X \setminus \mathrm{cut}(\mathbf{L})$. As above, we can assume without loss of generality that $\mathbf{L} = \mathrm{span}\{e_1, \ldots, e_k\}$. Let $X = \{p = 0\}$ for a homogeneous polynomial p of degree d in the Plücker coordinates. By Lemma 4.34, in order to find the critical points not in the cut locus of $\mathbf{L}$, we can work in the affine chart $(\mathcal{U}_{1,\ldots,k}, \psi_{1,\ldots,k})$. Here we dehomogeneize p, obtaining the corresponding polynomial function $\hat{p} : \mathbb{R}^L \longrightarrow \mathbb{R}$; thus $X \setminus \mathrm{cut}(\mathbf{L})$ is defined in these coordinates by $\hat{p} = 0$.

Consider now the set $\mathcal{G} := \mathrm{St}(k, n-k) \times O(k) \times \left[0, \frac{\pi}{2}\right)^k \subseteq \mathbb{R}^{(n-k)\times k} \times \mathbb{R}^{k\times k} \times \mathbb{R}^k \cong \mathbb{R}^{k(n+1)}$ and define the following map, where the variables are now the SVD components of the tangent vectors:

$$\begin{array}{rccc} \widetilde{\exp}_{\mathbf{L}} : & \mathcal{G} & \longrightarrow & G(k,n) \\ & (U,V,\mu) & \longmapsto & \begin{bmatrix} \cos(\mu_1)v_1 & \ldots & \cos(\mu_k)v_k \\ \sin(\mu_1)u_1 & \ldots & \sin(\mu_k)u_k \end{bmatrix}, \end{array} \tag{4.22}$$

with $U = (u_1, \ldots, u_k) \in \mathrm{St}(k, n-k)$, $V = (v_1, \ldots, v_k) \in O(k)$ and $\mu = (\mu_1, \ldots, \mu_k)$, and recall the notation $[M]$ to denote the k–dimensional span of a matrix $M \in \mathbb{R}^{(n-k)\times k}$ of full–rank. We express the polynomial function $\hat{p}$ in these new coordinates (i.e. the entries of U, V and μ) as $\widetilde{p} := \hat{p} \circ \iota_P \circ \widetilde{\exp}_{\mathbf{L}}$. The relation between the maps involved is depicted in the following commutative diagram:

$$\begin{array}{ccccc} G(k,n) \setminus \mathrm{cut}(\mathbf{L}) & \xrightarrow{\iota_P} & \mathcal{U}_{1\ldots k} & \xrightarrow{\widehat{p}} & \mathbb{R} \\ \uparrow{\scriptstyle \exp_{\mathbf{L}}} & & & \nearrow{\scriptstyle \widetilde{p}} & \\ \mathrm{int}\left(R_{\frac{\pi}{2}}\right) & & & & \\ \uparrow{\scriptstyle \widetilde{\exp}_{\mathbf{L}}} & & & & \\ \mathcal{G} & & & & \end{array} .$$

It follows that the preimage $Y := \widetilde{\exp}_{\mathbf{L}}^{-1}(X \setminus \mathrm{cut}(\mathbf{L})) \subseteq \mathcal{G}$ is defined in the coordinates of

$$\mathcal{U} := \mathbb{R}^{(n-k)\times k} \times \mathbb{R}^{k\times k} \times \left(0, \frac{\pi}{2}\right)^k$$

by the system of equations

$$\begin{cases} \widetilde{p}(U,V,\mu) = 0 \\ U^\top U = \mathbb{1}_k \\ V^\top V = \mathbb{1}_k \end{cases}, \tag{4.23}$$

where the last two sets of equations are the defining equations for $\mathrm{St}(k, n-k)$ and $O(k)$ respectively. In these coordinates the squared distance function from $\mathbf{L}$ is given by

$$f(U, V, \mu) := \delta^2_{\mathbf{L}}\left(\widetilde{\exp}_{\mathbf{L}}(U, V, \mu)\right) = \mu_1^2 + \cdots + \mu_k^2. \tag{4.24}$$

The main idea for the rest of the proof is now to find critical points of the restriction of this new function f to Y and then relate them to the critical points of the original function $\delta_{\mathbf{L}}|_X$ lying in $X \setminus \mathrm{cut}(\mathbf{L})$.

Second step: bound on the Betti numbers of the critical set of $f|_Y$.

The critical points of $f|_Y$ can be found through Lagrange multipliers rule, i.e. imposing that the gradient of (4.24) is a linear combination of the gradients of the equations of the defining equations of Y in (4.23). This gradient is given by the vector

$$\nabla f(U, V, \mu) = 2\,(0, 0, \mu).$$

To shorten notations, we introduce the function $F : \mathbb{R}^{(n-k)\times k} \times \mathbb{R}^{k\times k} \longrightarrow \mathbb{R}^{k(k+1)}$ given by

$$F(U, V) := ((\langle u_i, u_j\rangle - \delta_{ij})_{1\le i\le j\le k}, (\langle v_i, v_j\rangle - \delta_{ij})_{1\le i\le j\le k});$$

then $F(U, V) = 0$ are exactly the defining equations for $\mathrm{St}(k, n-k) \times O(k)$. This way (4.23) can be equivalently written as

$$\begin{cases} \widetilde{p}(U, V, \mu) = 0 \\ F(U, V) = 0 \end{cases}. \tag{4.25}$$

Since the defining equations for Stiefel manifolds and orthogonal groups are regular, at every point $(U, V) \in F^{-1}(0)$ the differential of F, denoted by $D_{(U,V)}F$, has maximal rank, equal to $k(k+1)$. Recall that the rows of the differential matrix $D_{(U,V)}F$ are given by the gradients of the components of F. Denote also by $\partial_{(U,V)}\widetilde{p}$ the vector of partial derivatives of $\widetilde{p}$ with respect to the variables in U and V, and by $\partial_\mu \widetilde{p}$ the vector of partial derivatives of $\widetilde{p}$ with respect to the variables $\mu_1, \ldots, \mu_k$. To visualize the criticality conditions better, we look at the block matrix containing as rows the gradients of the equations in (4.25) and ∇f

$$M := \begin{pmatrix} \partial_{(U,V)}\widetilde{p} & \partial_\mu \widetilde{p} \\ D_{(U,V)}F & 0 \\ 0 & \mu \end{pmatrix}; \tag{4.26}$$

the criticality conditions are that (4.25) is satisfied and the last row of (4.26) is linearly dependent from the other rows.

If $\partial_\mu \widetilde{p} = 0$, the criticality condition about the dependence of gradients would force $\mu = 0$. This solution would then correspond to the unique point $\mathbf{L}$, which doesn't lie on X. Therefore on the solution set we have $\mu \neq 0$ and $\partial_\mu \widetilde{p} \neq 0$ and the upper part of M consisting of the gradient of $\widetilde{p}$ and F has maximal rank, equal to $k(k+1)+1$. It follows that the criticality condition involving M is equivalent to $\partial_\mu \widetilde{p} \neq 0$ and all maximal minors of M vanish.

To enforce the condition $\partial_\mu \widetilde{p} \neq 0$, we subdivide the system into $2k$ systems of the following forms

$$\begin{cases} \widetilde{p}(U, V, \mu) = 0 \\ F(U, V) = 0 \\ \partial_{\mu_i}\widetilde{p} > 0 \\ \mathrm{rank}(M) = k(k+1)+1 \end{cases}, \qquad \begin{cases} \widetilde{p}(U, V, \mu) = 0 \\ F(U, V) = 0 \\ \partial_{\mu_i}\widetilde{p} < 0 \\ \mathrm{rank}(M) = k(k+1)+1 \end{cases}, \tag{4.27}$$

for $i = 1, \dots, k$. It is an easy linear algebra exercise to show that

$$\operatorname{rank}(M) = k(k+1)+1 \iff \operatorname{rank}\begin{pmatrix} \partial_\mu \widetilde{p} \\ \mu \end{pmatrix} = 1 \text{ and } \operatorname{rank}\begin{pmatrix} \partial_{(U,V)} \widetilde{p} \\ D_{(U,V)} F \end{pmatrix} = k(k+1), \quad (4.28)$$

so that in the systems (4.27) the last condition can be replaced by the right side of (4.28).

At this point, the crucial observation is that all the functions involved in the equations in (4.27) and (4.28) are Pfaffian functions in the domain $\mathcal{U} \subseteq \mathbb{R}^{k(n+1)}$ in a common Pfaffian chain. To see this, define the functions

$$\begin{aligned} f_{1,i}(U, V, \mu) &:= \tan\left(\frac{\mu_i}{2}\right), \\ f_{2,i}(U, V, \mu) &:= \cos^2\left(\frac{\mu_i}{2}\right), \\ f_{3,i}(U, V, \mu) &:= \cos(\mu_i), \\ f_{4,i}(U, V, \mu) &:= \sin(\mu_i). \end{aligned}$$

By Example 2.13 the ordered sequence $\mathfrak{C} = (f_{1,1}, \dots, f_{4,1}, f_{1,2}, \dots, f_{4,k})$ is a Pfaffian chain in $\mathcal{U} \subseteq \mathbb{R}^{k(n+1)}$ of order $4k$ and degree 2. We now examine all the functions in (4.27) and (4.28), computing their pfaffian formats with respect to the chain $\mathfrak{C}$ and the $k(n+1)$ variables (U, V, μ):

- $\widetilde{p}$ is a polynomial of degree at most d in the affine Plücker coordinates obtained taking the $k \times k$ minors of the matrix in (4.22) and dividing by the product $\cos(\mu_1) \dots \cos(\mu_k)$. Each of the entries of the matrix is a degree 2 polynomial in the functions of the Pfaffian chain above. Therefore $\widetilde{p}$ is a Pfaffian function of order $4k$ and degrees bounded by $(2, 2kd)$ on $\mathcal{U}$;
- $F(U, V) = 0$ corresponds to $k(k+1)$ equations, each being polynomial of degree 2;
- being the derivative of a Pfaffian function of format $(4k, 2, 2kd)$, for every $i = 1, \dots, k$ the function $\partial_{\mu_i} \widetilde{p}$ is Pfaffian of format $(4k, 2, 2kd+1)$ by Lemma 2.14;
- the condition of $\partial_\mu \widetilde{p}$ and μ forming a rank–1 matrix can be expressed through $k-1$ equations of Pfaffian format $(4k, 2, 2kd+2)$. Indeed depending on which of the systems we are looking at, we know that there exists $i \in \{1, \dots, k\}$ such that the column $(\partial_{\mu_i} \widetilde{p}, \mu_i)^\top$ is non-zero. It is then sufficient to impose that the other $k-1$ columns are proportional to this one;
- the condition that the matrix of partial derivatives of $\widetilde{p}$ and F with respect to U, V has rank $k(k+1)$ corresponds to all of its maximal minors being zero. As the matrix has size $(k(k+1)+1) \times nk$, this gives $\binom{nk}{k(k+1)+1}$ equations of Pfaffian format $(4k, 2, 2dk + k(k+1) + 1)$.

From these computations follows that (4.27) are systems involving $\binom{nk}{k(k+1)+1} + (k+1)^2$ Pfaffian functions of formats bounded by $(4k, 2, 2dk + k(k+1) + 1)$ in $k(n+1)$ variables.

Denote now by $S_i^+ \subseteq \mathcal{U}$ the solution set of the system on the left in (4.27) and by S_i^- that of the one on the right. Define the semi-Pfaffian set consisting of the union of all those solution sets

$$S := \bigcup_{\substack{i=1,\dots,k \\ * \in \{+,-\}}} S_i^*.$$

A direct application of the bound provided by Theorem 2.15 provides the following upper bound for the sum of the Betti numbers of S_i^* for every $i = 1, \dots, k$ and $* \in \{+, -\}$:

$$b(S_i^*) \leq \tilde{c}_1(k, n) d^{c_2(k,n)}, \quad (4.29)$$

where the $\tilde{c}_1(k,n)$ and $c_2(k,n)$ depend on k and n only and are given by

$$\begin{aligned}\tilde{c}_1(k,n) &:= c\,2^{8k^2-2k}\left(\binom{nk}{k(k+1)+1}+(k+1)^2\right)^{k(n+1)}(2k^2(n+1))^{k(n+5)},\\ c_2(k,n) &:= k(n+5),\end{aligned} \tag{4.30}$$

with c being a constant.

Third step: relate the critical points of $f|_Y$ to the critical points of $\delta_{\mathbf{L}}|_X$ lying in $X \setminus \mathrm{cut}(\mathbf{L})$.

Every $(U,V,\mu) \in S$ corresponds to a critical point of $\delta_{\mathbf{L}}|_X$ through $\widetilde{\exp}_{\mathbf{L}}$. Recall indeed that the variables (U,V,μ) were introduced to play the role of the singular values decomposition of tangent matrices in $T_{\mathbf{L}}G(k,n)$, to which apply the exponential at $\mathbf{L}$ to obtain the whole $G(k,n) \setminus \mathrm{cut}(\mathbf{L})$. Therefore the map that reconstructs critical points of $\delta_{\mathbf{L}}|_X$ is given by

$$(U,V,\mu) \in S \longmapsto \exp_{\mathbf{L}}(U\,\mathrm{diag}(\mu_1,\dots,\mu_k)\,V^\top) \in X \setminus \mathrm{cut}(\mathbf{L}).$$

Since this map is continuous, the discreteness of critical points outside of the cut locus of $\mathbf{L}$ given by Lemma 4.29 ensures that every connected component of S corresponds to exactly one critical point. Denoting by $b_0(S)$ the number of connected components of S, we therefore have the upper bound

$$\mathrm{GDC}(X) \le b_0(S) \le \sum_{\substack{i=1,\dots,k\\ *\in\{+,-\}}} b_0(S_i^*) \le \sum_{\substack{i=1,\dots,k\\ *\in\{+,-\}}} b(S_i^*). \tag{4.31}$$

The statement of the theorem follows from (4.31) and (4.29) with

$$c_1(k,n) := 2k\,\tilde{c}_1(k,n). \tag{4.32}$$

□

Remark 4.36. The bound provided by Theorem 4.35 is enormous: already in the case $k=2$, $n=4$ the constant $c_1(2,4)$ has order 10^{50}! This is a typical feature of the bounds obtained through the theory of Pfaffian functions (see Remark 2.16). The importance of Theorem 4.35 is not the value of the bound, but its existence and computability. Notice however that different choices of variables or ways of writing the critical equations might lead to improved bounds, even though obtaining reasonable bounds this way is hopeless.

Remark 4.37. An important feature of the strategy outlined in the proof of Theorem 4.35 is that it can be easily extended to any algebraic X for which explicit equations are given; the drawback is that the critical conditions might get more complicated if X is defined by more than one equation. Notice also that we could extend the procedure to *Pfaffian varieties*, i.e. varieties or submanifolds defined by Pfaffian equations. In that case we would need to keep into account also the format of the defining Pfaffian equations in the computation of the bound.

4.3 Nearest point problems to simple Schubert varieties

In this section we will apply the previous results to some specific subvarieties of $G(k,n)$. These varieties will be instances of simple Schubert varieties, chosen in such a way that we can exploit the results of Section 2.7 to describe them in terms of matrix varieties. For them we will be able to explicitly solve the nearest point problem Problem 4.9 for a generic input $\mathbf{L} \in G(k,n)$ and to also give a characterization of the global maxima.

4.3.1 Schubert varieties

Schubert varieties are an important class of subvarieties of Grassmannians, defined by requiring non-trivial intersection conditions between the k–planes of $G(k,n)$ and the elements of a fixed complete flag of linear spaces in $\mathbb{R}^n$. They arise naturally in a wide variety of different contexts and applications and have been therefore intensively studied. For example, we have already seen in Section 4.2.1 that the cut locus of a point in $G(k,n)$ can be described as a Schubert variety.

We will here briefly recall the definition of Schubert varieties through the correspondence with Young diagrams. The ambient space will be the Grassmannian $G(k,n)$ and we fix a complete flag

$$\{0\} = \mathbf{W}_0 \subsetneq \mathbf{W}_1 \subsetneq \cdots \subsetneq \mathbf{W}_n = \mathbb{R}^n,$$

where $\dim(\mathbf{W}_i) = i$ for every $i = 1, \dots, n$.

A *Young diagram* of format $k \times (n-k)$ is a collection of boxes chosen from a grid with k rows and $n-k$ columns such that, for every chosen box, all the boxes to its left and above it are also chosen. If we denote the Young diagram by λ and index the boxes in the grid from left to right and from top to bottom with indices (i,j), for $i \in \{1,\dots,k\}$ and $j \in \{1,\dots,n-k\}$, this is equivalent to $(i,j) \in \lambda$ implying that $(i-a,\, j-b) \in \lambda$ for every $a \in \{1,\dots,i-1\}$ and $b \in \{1,\dots,j-1\}$. We say that a box (i,j) is an *outer corner* of λ if $(i,j) \in \lambda$ but $(i+1,j) \notin \lambda$ and $(i,j+1) \notin \lambda$. It is immediate to see that a Young diagram is uniquely identified by its outer corners, since they completely determine which boxes belong to the diagram.

Definition 4.38. Given a Young diagram λ of format $k \times (n-k)$, the *Schubert variety associated to* λ is defined as

$$\Omega_\lambda := \{\mathbf{E} \in G(k,n) \mid \dim(\mathbf{E} \cap \mathbf{W}_{n-k+i-j}) \geq i \text{ for every } (i,j) \in \lambda\}. \tag{4.33}$$

Observe that if the condition in (4.33) relative to a box (i,j) is satisfied, then the strict inclusions between the subspaces in the fixed flag imply that also all the conditions relative to boxes to the left and above the box (i,j) are also satisfied, making the definition of a Schubert variety consistent with that of a Young diagram. It follows also that if the diagram λ has outer corners given by the boxes $(s_1,r_1),\dots,(s_a,r_a)$, then the only conditions actually needed in Definition 4.38 are those corresponding to those boxes, i.e

$$\Omega_\lambda := \{\mathbf{E} \in G(k,n) \mid \dim(\mathbf{E} \cap \mathbf{W}_{n-k+s_l-r_l}) \geq s_l \text{ for every } l = 1,\dots,a\}. \tag{4.34}$$

Schubert varieties can in general be defined without the use of Young diagrams, by requiring the intersections with some $\mathbf{W}_i$'s from the fixed flag to have at least some specific dimensions, in the same fashion as in (4.33). It is easy to see that any such non–trivial condition corresponds to some box in a Young diagram. This way, to every Young diagram we can associate a Schubert variety and viceversa every Schubert variety can be associated to a Young diagram.

There are a bunch of reasons why the representation via Young diagrams is particularly useful. One of them is that the codimension of the Schubert variety Ω_λ is given by the number of boxes in the defining Young diagram λ (see [41, Sec. 9.4]). Another one is that Schubert varieties over the complex numbers form a basis of the cohomology of the complex Grassmannian. Thus, intersection theory in the complex Grassmannian can be

reconduced to computing cup products between cohomology classes of Schubert varieties. The machinery behind these products is known as *Schubert calculus* and it works nicely with the Young diagrams representation.

Schubert varieties are not smooth in general. Using the same notation as in (4.34) the smooth locus, denoted by $\Omega_\lambda^{\text{sm}}$, is obtained replacing inequalities with equalities

$$\Omega_\lambda^{\text{sm}} := \{\mathbf{E} \in G(k,n) \mid \dim(\mathbf{E} \cap \mathbf{W}_{n-k+s_l-r_l}) = s_l \text{ for every } l = 1, \dots, a\},$$

while the singular part, denoted by $\Omega_\lambda^{\text{sing}}$, is obtained imposing strict inequalities for at least one of the defining conditions

$$\Omega_\lambda^{\text{sing}} := \{\mathbf{E} \in G(k,n) \mid \exists\, l \in \{1, \dots, a\} \text{ such that } \dim(\mathbf{E} \cap \mathbf{W}_{n-k+s_l-r_l}) > s_l\} \subseteq \Omega_\lambda\,.$$

The singular part admits a stratification in terms of how many of the inequalities in (4.34) are strict and how big is the difference between the actual dimension of intersection and the required one. One way of understanding this, is that every time that an inequality in (4.34) is strict for a plane $\mathbf{E} \in \Omega_\lambda$, such plane satisfies a condition coming from a box $(i,j) \notin \lambda$. In other words it satisfies the conditions for a higher codimension Schubert variety, contained in Ω_λ: these varieties are the singular strata of Ω_λ and correspond to Young diagrams containing the given λ. The stratification of the cut locus of a point $\mathbf{L} \in G(k,n)$ given by (4.7) is a particular instance of this construction.

Given a Schubert variety Ω_λ as in (4.34), it is not hard to describe the tangent space at a smooth point. This was done in [37, Proposition 4.3] for varieties defined by a diagram with a unique outer corner and generalized by Mathis in his PhD Thesis [67]. Recall the identification of the tangent space to $G(k,n)$ at a point $\mathbf{E}$ with the space of linear maps from $\mathbf{E}$ to $\mathbf{E}^\perp$.

Lemma 4.39. *Let λ be a Young diagram with outer corners $(s_1, r_1), \dots, (s_a, r_a)$ and let Ω_λ be the associated Schubert variety. The tangent space at a smooth point $\mathbf{E} \in \Omega_\lambda^{\text{sm}}$ is given by*

$$T_{\mathbf{E}}\Omega_\lambda^{\text{sm}} = \left\{\varphi : \mathbf{E} \longrightarrow \mathbf{E}^\perp \mid \varphi(\mathbf{E} \cap \mathbf{W}_{n-k+s_l-r_l}) \subseteq \mathbf{E}^\perp \cap \mathbf{W}_{n-k+s_l-r_l}\ \forall\, l = 1, \dots, a\right\}. \tag{4.35}$$

In other words, the tangent space is given by those linear maps that preserve the intersection between the plane $\mathbf{E}$ and the elements of the flag. The formula (4.8), that we used to describe the tangent space at a smooth point of the cut locus is a particular instance of the description in (4.35).

Definition 4.40. A Schubert variety is said to be *simple* if it is defined by a single condition of the type in (4.33) or, equivalently, if its associated Young diagram has a unique outer corner.

If the unique outer corner of the Young diagram λ is given by the box (s,r), the associated simple Schubert variety is thus given by

$$\Omega_\lambda := \{\mathbf{E} \in G(k,n) \mid \dim(\mathbf{E} \cap \mathbf{W}_{n-k+s-r}) \geq s\},$$

with its smooth locus corresponding to the equality case in above condition. In this case Ω_λ can be easily stratified in terms of the dimension of the intersection of a plane $\mathbf{E}$ with $\mathbf{W}_{n-k+s-r}$. We can thus write the stratification as

$$\Omega_\lambda = \Omega_\lambda^{\text{sm}} \sqcup \Omega_\lambda^{\text{sing}} = \Omega_\lambda^{\text{sm}} \sqcup \left(\bigsqcup_{j=1}^{k-s} \Omega_\lambda^{\text{sing},j} \right), \tag{4.36}$$

where

$$\Omega_\lambda^{\text{sm}} = \{\mathbf{E} \in G(k,n) \mid \dim(\mathbf{E} \cap \mathbf{W}_{n-k+s-r}) = s\}, \tag{4.37}$$

$$\Omega_\lambda^{\text{sing},j} = \{\mathbf{E} \in G(k,n) \mid \dim(\mathbf{E} \cap \mathbf{W}_{n-k+s-r}) = s+j\} \qquad j = 1, \dots, k-s, \tag{4.38}$$

where a stratum $\Omega_\lambda^{\text{sing},j}$ is empty if $j > n-k-r$.

A class of simple Schubert varieties

We will now turn our attention to a specific class of simple Schubert varieties. These will be the main focus of the rest of this chapter. The definition is as follows.

Definition 4.41 (Simple Schubert variety Ω_s)**.** Denote the k–dimensional element of the fixed complete flag as $\mathbf{W} := \mathbf{W}_k \in G(k,n)$. For any $s \in \{1, \dots, k-1\}$ we define the simple Schubert variety Ω_s as

$$\Omega_s := \{\mathbf{E} \in G(k,n) \mid \dim(\mathbf{E} \cap \mathbf{W}) \geq s\}. \tag{4.39}$$

This is the simple Schubert variety corresponding to the $k \times (n-k)$ Young diagram with unique outer corner given by $(s, n-2k+s)$, that is the outer corner lies on the diagonal going up from the bottom-right corner of the grid. The codimension of Ω_s in $G(k,n)$ is $s(n-2k+s)$. Observe that allowing $s = k$ in the definition would yield a variety with a single point $\Omega_k = \{\mathbf{W}\}$, making the nearest point problem in this case trivial.

The stratification of simple Schubert varieties described in (4.36) can be applied to our varieties Ω_s, giving

$$\Omega_s = \Omega_s^{\text{sm}} \sqcup \Omega_s^{\text{sing}} = \Omega_s^{\text{sm}} \sqcup \left(\bigsqcup_{j=1}^{k-s} \Omega_s^{\text{sing},j} \right),$$

with the strata described in (4.37) and (4.38) being given in this case by

$$\Omega_s^{\text{sm}} = \{\mathbf{E} \in G(k,n) \mid \dim(\mathbf{E} \cap \mathbf{W}) = s\},$$

$$\Omega_s^{\text{sing},j} = \{\mathbf{E} \in G(k,n) \mid \dim(\mathbf{E} \cap \mathbf{W}) = s+j\} \qquad j = 1, \dots, k-s.$$

All the singular strata are non empty, the deepest one being given by the single point $\mathbf{W}$ and the other ones being given by varieties Ω_t for some $t > s$. These varieties can therefore also be seen as a nested family of varieties; we will see that the nearest point problem respects this nested structure, in the sense that the global solution of the problem for Ω_s will always be in the smooth stratum of Ω_s and not in one of the smallest subvarieties Ω_t for $t > s$.

The main reason why we restrict to this specific class of simple Schubert varieties is that the element of the flag which is involved in the definition of Ω_s lives in the same

Grassmannian $G(k,n)$ as the points of the variety (it is indeed itself the most singular point of it). In the following we will sometimes use the notation $\theta_i(\mathbf{E}) := \theta_i(\mathbf{E},\mathbf{W})$ for the principal angles between a point $\mathbf{E} \in G(k,n)$ and the fixed $\mathbf{W}$, whenever it is clear which angles we are referring to. Recalling the properties of principal angles described in Section 2.7.4, Ω_s and its strata admit an equivalent description in terms of principal angles between planes $\mathbf{E} \in G(k,n)$ and $\mathbf{W}$ as

$$\begin{aligned}
\Omega_s &= \{\mathbf{E} \in G(k,n) \mid \theta_1(\mathbf{E}) = \cdots = \theta_s(\mathbf{E}) = 0\},\\
\Omega_s^{\mathrm{sm}} &= \{\mathbf{E} \in G(k,n) \mid \theta_1(\mathbf{E}) = \cdots = \theta_s(\mathbf{E}) = 0,\ \theta_{s+1}(\mathbf{E}) > 0\},\\
\Omega_s^{\mathrm{sing,j}} &= \{\mathbf{E} \in G(k,n) \mid \theta_1(\mathbf{E}) = \cdots = \theta_{s+j}(\mathbf{E}) = 0,\ \theta_{s+j+1}(\mathbf{E}) > 0\},
\end{aligned}$$

for $j = 1,\dots,k-s$.

Recall that in Proposition 2.30 we have shown how the singular values of a $(n-k)\times k$ matrix $A \in R_{\frac{\pi}{2}}$ correspond to the principal angles between $\mathbf{W}$ and the image of the tangent vector $v_A \in T_{\mathbf{W}}G(k,n)$ through the exponential map at $\mathbf{W}$ (we use the identification between $\mathbb{R}^{(n-k)\times k}$ and $T_{\mathbf{W}}G(k,n)$ in Definition 2.21, the choice of a representative $\widehat{W} \in O(n)$ is implicit). Since the rank of a rectangular matrix coincides with the number of its non–zero singular values, the rank of $A \in R_{\frac{\pi}{2}}$ corresponds to the number of non–zero principal angles between $\mathbf{W}$ and the exponential of $\exp_{\mathbf{W}}(v_A)$. Recalling the notation introduced in (2.11) for the set of matrices with rank constraints and the fact that the image of $R_{\frac{\pi}{2}}$ through the exponential $\exp_{\mathbf{W}}$ is the whole Grassmannian $G(k,n)$, we obtain the following corollary of Proposition 2.30.

Corollary 4.42. *Let* $\exp_{\mathbf{W}} : \mathbb{R}^{(n-k)\times k} \longrightarrow G(k,n)$ *be the exponential map of* $G(k,n)$ *at* $\mathbf{W}$ *and* Ω_s *and its strata be defined as above. Then*

$$\begin{aligned}
\Omega_s &= \exp_{\mathbf{W}}\left(\mathcal{M}_{\leq k-s} \cap R_{\frac{\pi}{2}}\right),\\
\Omega_s^{\mathrm{sm}} &= \exp_{\mathbf{W}}\left(\mathcal{M}_{k-s} \cap R_{\frac{\pi}{2}}\right),\\
\Omega_s^{\mathrm{sing},j} &= \exp_{\mathbf{W}}\left(\mathcal{M}_{k-s-j} \cap R_{\frac{\pi}{2}}\right),
\end{aligned} \tag{4.40}$$

for $j = 1,\dots,k-s$.

This corollary thus gives a full description of the variety Ω_s as the images of matrices with some fixed rank. This result allows us to use the techniques coming from matrix rank–approximation, and in particular Eckart–Young theorem Theorem 2.19, to find some critical points for the Grassmann distance function from a generic $\mathbf{L} \notin \Omega_s$ restricted to Ω_s.

4.3.2 Eckart–Young theorem and critical points

In this section we explore how to exploit the description given by Corollary 4.42 to find critical points for the distance function $\delta_{\mathbf{L}}$ restricted to the variety Ω_s for a generic $\mathbf{L} \in G(k,n)$. As anticipated, our method will rely on ideas coming from the Eckart–Young rank–approximation theorem for matrices (Theorem 2.19).

The first thing to observe that for a generic $\mathbf{L} \in G(k,n)$, the intersection $\mathbf{L} \cap \mathbf{W}$ is trivial, so that no principal angle between $\mathbf{L}$ and $\mathbf{W}$ is generically zero. The same is true for $\mathbf{L} \cap \mathbf{W}^{\perp}$, meaning that no principal angle will be generically $\frac{\pi}{2}$. Moreover the angles will all be distinct. In particular we have that the generic $\mathbf{L}$ is not in the cut locus of $\mathbf{W}$.

Thus the preimage through the exponential map at $\mathbf{W}$ inside the set $R_{\frac{\pi}{2}}$ consists of only one matrix with full rank. We denote this matrix by $A_{\mathbf{L}}$:

$$A_{\mathbf{L}} := \exp_{\mathbf{W}}^{-1}(\mathbf{L}) \cap R_{\frac{\pi}{2}}. \tag{4.41}$$

By (4.40) we have that the preimage of the whole variety Ω_s in $R_{\frac{\pi}{2}}$ corresponds to the matrices of bounded rank $\mathcal{M}_{\leq k-s} \cap R_{\frac{\pi}{2}}$. As we want to find critical points for the restriction of the distance function from $\mathbf{L}$ to Ω_s, we might wonder whether this problem can be translated to the matrix case in $T_{\mathbf{W}}G(k,n)$ via the inverse of the exponential map $\exp_{\mathbf{W}}^{-1}$, where $\mathbf{L}$ corresponds to the full-rank matrix $A_{\mathbf{L}}$ and Ω_s corresponds to matrices with rank bounded by $k-s$.

The (crazy at this point!) idea would be the following:

(i) given $\mathbf{L} \notin \Omega_s$ generic, find the unique full–rank matrix $A_{\mathbf{L}} \in R_{\frac{\pi}{2}}$ such that $\exp_{\mathbf{W}}(A_{\mathbf{L}}) = \mathbf{L}$;

(ii) use Eckart–Young (Theorem 2.19) to find the critical rank–approximations of $A_{\mathbf{L}}$ of rank $k-s$;

(iii) map back on $G(k,n)$ the rank–approximated matrices with the exponential $\exp_{\mathbf{W}}$.

Remark that this procedure gives smooth points of Ω_s, since the critical rank–approximation matrices given by the Eckart–Young theorem have exactly rank $k-s$ and therefore correspond to smooth points of Ω_s through the exponential map. A priori, there is no reason to believe that the points found via above strategy are critical points for the restricted distance function $\delta_{\mathbf{L}}|_{\Omega_s}$. One would indeed expect a distortion effect in the metric (and thus the critical point equations) given by applying the exponential map. Suprisingly, this procedure actually produces critical points, as shown in the following result.

Theorem 4.43. *Let $\mathbf{L} \notin \Omega_s$ be a generic point in $G(k,n)$ and $A_{\mathbf{L}} \in R_{\frac{\pi}{2}}$ be the unique full rank matrix such that $\exp_{\mathbf{W}}(A_{\mathbf{L}}) = \mathbf{L}$. For any subset $I \subseteq \{1,\dots,k\}$ of cardinality $|I| = k-s$ let $A_{\mathbf{L},I} \in \mathcal{M}_{k-s}$ be the rank–approximation obtained from $A_{\mathbf{L}}$ via Theorem 2.19. Define its image via the exponential map $\mathbf{L}_I := \exp_{\mathbf{W}}(A_{\mathbf{L},I}) \in \Omega_s^{\mathrm{sm}}$. Then $\mathbf{L}_I$ is a critical point for $\delta_{\mathbf{L}}|_{\Omega_s}$.*

Before proving the theorem, let us try to understand what is happening from a geometric point of view when we approximate $A_{\mathbf{L}}$ with some $A_{\mathbf{L},I}$, going from the corresponding point $\mathbf{L}$ to $\mathbf{L}_I$. Imagine that we have the basis of principal vectors of $\mathbf{W}$ and $\mathbf{L}$ relative to each other. When we apply the exponential map $\exp_{\mathbf{W}}$ to $A_{\mathbf{L}}$, we rotate the principal vectors of $\mathbf{W}$ onto those of $\mathbf{L}$ with the speed prescribed by the corresponding principal angles $0 < \theta_1(\mathbf{L}) < \cdots < \theta_k(\mathbf{L}) < \frac{\pi}{2}$, which in turn correspond to the singular values of $A_{\mathbf{L}}$. When we approximate $A_{\mathbf{L}}$ with $A_{\mathbf{L},I}$, the singular values, and therefore by Proposition 2.30 the principal angles, corresponding to indices not in I are set to 0. This means that applying $\exp_{\mathbf{W}}$ to $A_{\mathbf{L},I}$, while the principal vectors corresponding to indices in I still get rotated to the corresponding ones in $\mathbf{L}$, the ones corresponding to indices not in I will now stay still, as the corresponding angles of rotation have been set to 0. Remark that, starting from $\mathbf{W}$, rotating the principal vectors of $\mathbf{W}$ towards those of $\mathbf{L}$ descreases the distance from $\mathbf{L}$. Therefore, setting some singular values in $A_{\mathbf{L}}$ to zero "preserves" the starting distance between $\mathbf{W}$ and $\mathbf{L}$ corresponding to those principal angles, while keeping a singular value unchanged allows to perform the rotation and annihilate the contribution of that angle to the distance.

Proof of Theorem 4.43. The proof will be divided into several steps. First, given $\mathbf{W}$ and $\mathbf{L}$, we will compute the matrix $A_{\mathbf{L}}$. We will then switch to $A_{\mathbf{L},I}$ and compute $\mathbf{L}_I$. As a last step we will compute the length minimizing geodesic connecting $\mathbf{L}_I$ to $\mathbf{L}$ and prove that this geodesic starts orthogonal to Ω_s, showing that $\mathbf{L}_I$ is indeed a critical point for $\delta_{\mathbf{L}}|_{\Omega_s}$.

To compute $A_{\mathbf{L}}$ we proceed as in Section 2.7. Assume we fixed some representatives $\widehat{W} = (w_1 \dots w_k \; b_1 \dots b_{n-k}) \in \pi^{-1}(\mathbf{W})$ and $\widehat{L} \in \pi^{-1}(\mathbf{L})$, where recall that $\pi : O(n) \longrightarrow G(k,n)$ denotes the Riemannian submersion that gives the quotient structure of $G(k,n)$. Let $0 < \theta_1(\mathbf{L}) < \cdots < \theta_k(\mathbf{L}) < \frac{\pi}{2}$ be the principal angles bewteen $\mathbf{W}$ and $\mathbf{L}$. As we already observed, by the genericity of $\mathbf{L} \in G(k,n)$ we can assume they are all distinct and different from 0 and $\frac{\pi}{2}$. Let $\Sigma = \operatorname{diag}(\theta_1(\mathbf{L}), \dots, \theta_k(\mathbf{L})) \in \mathbb{R}^{(n-k)\times k}$ and $\{p_1, \dots, p_k\}$ and $\{q_1, \dots, q_k\}$ be corresponding basis of principal vectors for $\mathbf{W}$ and $\mathbf{L}$ respectively. As in Section 2.7 define $\mathfrak{P}_i := \operatorname{span}\{p_i, q_i\}$ and $n_i \in \mathfrak{P}_i$ as the unit vector such that

$$q_i = p_i \cos(\theta_i(\mathbf{L})) + n_i \sin(\theta_i(\mathbf{L})).$$

We can complete the $\{n_1, \dots, n_k\}$ to an orthonormal basis of $\mathbf{W}^{\perp}$, and denote it by $\{n_1, \dots, n_{n-k}\}$.

Define $V = (v_1 \dots v_k) \in O(k)$ to be the change of basis matrix from $\{w_1, \dots w_k\}$ to $\{p_1, \dots, p_k\}$, i.e. $(w_1 \dots w_k)V = (p_1 \dots p_k)$. Similarly define $U = (u_1 \dots u_{n-k}) \in O(n-k)$ as the change of basis matrix from $\{b_1, \dots, b_{n-k}\}$ to $\{n_1, \dots, n_{n-k}\}$. We know by Section 2.7 that the matrix $A_{\mathbf{L}}$ can then be constructed as

$$A_{\mathbf{L}} = U\Sigma V^{\top}, \qquad \exp_{\mathbf{W}}(A_{\mathbf{L}}) = \mathbf{L}.$$

For any subset of cardinality $k-s$, $I = \{i_1, \dots, i_{k-s}\} \subseteq \{1, \dots, k\}$, define the matrix $\widetilde{\Sigma} = \operatorname{diag}(\widetilde{\theta}_1(\mathbf{L}), \dots, \widetilde{\theta}_k(\mathbf{L}))$ such that $\widetilde{\theta}_i(\mathbf{L}) = \theta_i(\mathbf{L})$ if $i \in I$ and $\widetilde{\theta}_i(\mathbf{L}) = 0$ otherwise, with increasing ordering $i_1 < \cdots < i_{k-s}$. Define $A_{\mathbf{L},I} := U\widetilde{\Sigma}V^{\top}$ and $\mathbf{L}_I := \exp_{\mathbf{W}}(A_{\mathbf{L},I})$. By construction we have $\operatorname{rank}(A_{\mathbf{L},I}) = k - s$ and therefore the point $\mathbf{L}_I$ is in the smooth stratum of Ω_s.

By Proposition 2.27 we have

$$\mathbf{L}_I = \left[\widehat{W}\begin{pmatrix} \cos(\widetilde{\theta}_1(\mathbf{L}))\, v_1 & \dots & \cos(\widetilde{\theta}_k(\mathbf{L}))\, v_k \\ \sin(\widetilde{\theta}_1(\mathbf{L}))\, u_1 & \dots & \sin(\widetilde{\theta}_k(\mathbf{L}))\, u_k \end{pmatrix}\right].$$

Explicitly, an orthonormal basis for $\mathbf{L}_I$ is given by

$$\begin{aligned} i \in I: &\qquad \widehat{W}\begin{pmatrix} \cos(\theta_i(\mathbf{L}))\, v_i \\ \sin(\theta_i(\mathbf{L}))\, u_i \end{pmatrix} = \cos(\theta_i(\mathbf{L}))\, p_i + \sin(\theta_i(\mathbf{L}))\, n_i = q_i, \\ i \notin I: &\qquad \widehat{W}\begin{pmatrix} v_i \\ 0 \end{pmatrix} = p_i. \end{aligned}$$

We now want to describe the unique length minimizing geodesic connecting $\mathbf{L}_I$ to $\mathbf{L}$. We denote such curve as $\gamma_I : [0,1] \longrightarrow G(k,n)$ with $\gamma_I(0) = \mathbf{L}_I$ and $\gamma_I(1) = \mathbf{L}$. We use the procedure already described in Section 2.7. Denote by $\{1, \dots, k\} \setminus I = \{j_1, \dots, j_s\}$ the complement set of I, with increasing ordering $j_1 < \cdots < j_s$. As orthonormal bases for the subspaces we use $\beta_1 := \{q_{i_1}, \dots, q_{i_{k-s}}, p_{j_1}, \dots, p_{j_s}\}$ for $\mathbf{L}_I$ and $\{q_{i_1}, \dots, q_{i_{k-s}}, q_{j_1}, \dots, q_{j_s}\}$

for $\mathbf{L}$. Then by construction and the orthogonality properties of principal vectors we have

$$\begin{pmatrix} q_{i_1}^\top \\ \vdots \\ q_{i_{k-s}}^\top \\ \\ q_{j_1}^\top \\ \vdots \\ q_{j_s}^\top \end{pmatrix} \begin{pmatrix} q_{i_1} \dots q_{i_{k-s}}\ p_{j_1} \dots p_{j_s} \end{pmatrix} = \operatorname{diag}\Big(1, \dots, 1, \cos(\theta_{j_1}(\mathbf{L})), \dots, \cos(\theta_{j_s}(\mathbf{L}))\Big) \in \mathbb{R}^{k\times k},$$

which is already in a singular values decomposition form. This means that the basis above are already bases of principal vectors for $\mathbf{L}_I$ and $\mathbf{L}$, with the first $k-s$ principal angles being 0 and the remaining s being $0 < \theta_{j_1}(\mathbf{L}) < \cdots < \theta_{j_s}(\mathbf{L}) < \frac{\pi}{2}$.

We complete the orthonormal set $\{n_{j_1}, \dots, n_{j_s}\} \subseteq \mathbf{L}_I^\perp$ to an orthonormal basis of $\mathbf{L}_I^\perp$ denoted by $\beta_2 := \{f_1, \dots, f_{k-s}, n_{j_1}, \dots, n_{j_s}, f_{k+1}, \dots, f_{n-k}\}$. To write the geodesic γ_I we use the representative $\widehat{L}_I \in \pi^{-1}(\mathbf{L}_I)$ given by the basis β_1 and β_2, in order to fix the isomorphism $T_{\mathbf{L}_I}G(k,n) \cong \mathbb{R}^{(n-k)\times k}$. Then, the matrix $M_I := \operatorname{diag}(0, \dots, 0, \theta_{j_1}(\mathbf{L}), \dots, \theta_{j_s}(\mathbf{L})) \in \mathbb{R}^{(n-k)\times k}$ is such that $\exp_{\mathbf{L}_I}(v_{M_I}) = \mathbf{L}$. Thus, we are only left to check that v_{M_I} is in the normal space to Ω_s at $\mathbf{L}_I$, as v_{M_I} is the tangent vector to γ_I at the starting point $\mathbf{L}_I$. Using the identification of tangent vectors v_A as linear maps φ_A from $\mathbf{L}_I$ to $\mathbf{L}_I^\perp$, from (4.35) we obtain that the tangent space to Ω_s at the smooth point $\mathbf{L}_I$ is given by

$$T_{\mathbf{L}_I}\Omega_s = \{\varphi_A : \mathbf{L}_I \longmapsto \mathbf{L}_I^\perp \mid \varphi_A(\mathbf{L}_I \cap \mathbf{W}) \subseteq \mathbf{L}_I^\perp \cap \mathbf{W}\}.$$

We can then see that the normal space to Ω_s at $\mathbf{L}_I$, which is the orthogonal complement of the tangent space inside the tangent space to the Grassmannian $T_{\mathbf{L}_I}G(k,n)$, is given by

$$N_{\mathbf{L}_I}\Omega_s = \{\varphi_A : \mathbf{L}_I \longrightarrow \mathbf{L}_I^\perp \mid \varphi_A(\mathbf{L}_I \cap \mathbf{W}) \subseteq (\mathbf{L}_I^\perp \cap \mathbf{W})^\perp \cap \mathbf{L}_I^\perp \quad \text{and} \quad \varphi_A((\mathbf{L}_I \cap \mathbf{W})^\perp \cap \mathbf{L}_I) = \{0\}\}.$$

By construction, a basis for the subspace $\mathbf{L}_I \cap \mathbf{W}$ is given by $\{p_{j_1}, \dots, p_{j_s}\}$. The matrix M_I represents a linear map $\varphi_{M_I} : \mathbf{L}_I \longrightarrow \mathbf{L}_I^\perp$ with respect to the basis β_1 and β_2. Therefore we have

$$\begin{aligned} &\varphi_{M_I}(q_{i_1}) = 0, \\ &\qquad \vdots \\ &\varphi_{M_I}(q_{i_{k-s}}) = 0, \\ &\varphi_{M_I}(p_{j_1}) = \theta_{j_1}(\mathbf{L})\, n_{j_1}, \\ &\qquad \vdots \\ &\varphi_{M_I}(p_{j_s}) = \theta_{j_s}(\mathbf{L})\, n_{j_s}. \end{aligned}$$

Since $n_{j_1}, \dots, n_{j_s}$ are orthogonal to both $\mathbf{L}_I$ and $\mathbf{W}$, it follows that $M_I \in N_{\mathbf{L}_I}\Omega_s$. Therefore $\mathbf{L}_I$ is a critical point of the restricted distance function $\delta_{\mathbf{L}}|_{\Omega_s}$ by Proposition 4.15 and Corollary 4.24.

□

Remark 4.44 (The Eckart–Young minimizer)**.** In the literature, the Eckart–Young theorem is stated with the singular values of the matrix being listed in decreasing order. We chose in our statement of Theorem 2.19 to follow the same notation. Nevertheless, remark that

in the proof of Theorem 4.43, when we write explicitly the SVD of the matrix $A_{\mathbf{L}}$, the order of the singular values, which coincide with the principal angles between $\mathbf{W}$ and $\mathbf{L}$, is reversed, as the angles are always listed in increasing order. Therefore the point with minimum distance from $\mathbf{L}$ among the points $\mathbf{L}_I$ built in Theorem 4.43 corresponds the choice of the last $k-s$ indices, instead of the first ones, that is $I = \{s+1, \dots, k\}$.

Remark 4.45. Eckart–Young theorem does not contemplate the possibility of "increasing" a singular value, because the variety $\mathcal{M}_{\leq k-s}$ is unbounded. In particular, maxima cannot be found by increasing singular values, because we could always increase them more and move further away. But in our setting, we actually restrict to the compact set $\mathcal{M}_{\leq k-s} \cap R_{\frac{\pi}{2}}$ and it makes sense to ask what happens if instead of modifying a singular value of $A_{\mathbf{L}}$ by sending it to 0, thus not performing any rotation, we modify it with the aim of producing a point of maximum. In order to achieve this, instead of rotating from the principal vectors for $\mathbf{W}$ to those for $\mathbf{L}$ or staying still, we have to rotate in the opposite direction until we create an angle of $\frac{\pi}{2}$ with respect to the original vector, which is the maximum achievable in that direction. In other words, if the original angle between $\mathbf{W}$ and $\mathbf{L}$ is $\theta_i(\mathbf{L})$, we have to rotate the i–th principal vector p_i of $\mathbf{W}$ in the opposite direction with respect to that of $\mathbf{L}$ q_i of an angle of $\frac{\pi}{2} - \theta_i(\mathbf{L})$. We believe that this new kind of modification of $A_{\mathbf{L}}$ again produces critical points for the restricted distance function $\delta_{\mathbf{L}}|_{\Omega_s}$, with the difference that now they can also be at a distance from $\mathbf{L}$ greater than that of $\mathbf{W}$ (which was not the case in Theorem 4.43; the distance was always decreasing). The two approaches, the one aimed at minimizing the distance in some directions, and the one that aims to maximize it, can also be mixed to produce new types of critical points. All these newly created points will be in the cut locus of $\mathbf{L}$, therefore we cannot use the characterization of critical points through normal geodesics. In order to show that these are critical points one should explicitly compute the subgradient and check that it intersects the normal space to Ω_s at those points. Since we already know that no such points are minimizers, we will not perform the computation here. Nevertheless, we will use this idea when computing global maximizers in Section 4.3.5.

Remark 4.46. Theorem 4.43 tells us that the points we obtain are critical points, but these *are not* all the critical points outside of the cut locus, in general. Nevertheless, it provides a lower bound to the Grassmann Distance Complexity of the algebraic subvarieties Ω_s. In particular, Theorem 4.43 gives a critical point for every choice of a subset of $k-s$ indices among $\{1, \dots, k\}$, which amounts to $\binom{k}{s}$ critical points. Thus we have

$$\mathrm{GDC}(\Omega_s) \geq \binom{k}{s}.$$

In the next section we will describe the local structure of the Schubert variety Ω_s around $\mathbf{W}$. When the generic point is chosen close enough to $\mathbf{W}$, this will allow us to characterize all the critical points close to $\mathbf{W}$: they will be *exactly* those provided by Theorem 4.43.

4.3.3 The local structure around the deepest singularity

We begin this section by describing a degenerate setting for our problem. Instead of picking a generic point $\mathbf{L} \notin \Omega_s$ as usual, we look at what happens when the point from which we compute the distance is specifically $\mathbf{W} \in \Omega_s$, i.e. the point corresponding to the deepest singularity of Ω_s. In this case we can show that the situation for the critical points of the restricted distance function $\delta_{\mathbf{W}}|_{\Omega_s}$ is the following:

– there is a unique global minimum, which is trivially $\mathbf{W}$ itself;

– there are no other local minima or saddle points;

– there is a $G(s,k) \times G(k-s, n-k)$ submanifold of global maximizers and no other local ones.

To see this, remark that for every point $\mathbf{E} \in \Omega_s$, the first s principal angles with $\mathbf{W}$ must be 0, so that the maximum distance possible from $\mathbf{W}$ is a priori given by $\sqrt{(k-s)}\,\frac{\pi}{2}$, corresponding to when the remaining $k-s$ principal angles are $\frac{\pi}{2}$. By the properties of principal angles, if $(k-s)$ principal angles are $\frac{\pi}{2}$, it means that $\mathbf{E} \cap \mathbf{W}^{\perp}$ has dimension $k-s$. Therefore any global maximizer $\mathbf{E}$ can be constructed as the orthogonal direct sum of an s–dimensional subspace of $\mathbf{W}$, accounting for the defining condition of Ω_s, and a $(k-s)$–dimensional subspace of $\mathbf{W}^{\perp}$. It follows that the set of global maximizers can be identified with a $G(s,k) \times G(k-s, n-k)$. To see that no other points can be critical, it is enough to observe that the geodesics connecting $\mathbf{W}$ and any other point $\mathbf{E} \in \Omega_s$ are entirely contained in Ω_s, so that no other point apart from the global maximizers and $\mathbf{W}$ itself can be critical.

Now we ask the following question: what happens if we choose a point $\mathbf{L} \notin \Omega_s$ sufficiently close to $\mathbf{W}$? To answer this question, we first investigate the local structure of the simple Schubert variety Ω_s in a neighborhood of its deepest singularity $\mathbf{W}$.

Let $D := D(0,1)$ be the unit disk in $T_{\mathbf{W}}G(k,n)$ and D_ε be the disk of radius ε (norms of tangent vectors have to be understood in the Grassmann metric). For every $\varepsilon > 0$ define the map

$$\phi_\varepsilon : D \xrightarrow{\cdot\varepsilon} D_\varepsilon \xrightarrow{\exp_{\mathbf{W}}} B(\mathbf{W}, \varepsilon), \tag{4.42}$$

where $B(\mathbf{W}, \varepsilon)$ is the geodesic ball of radius ε in $G(k,n)$ centered at $\mathbf{W}$. The family of maps ϕ_ε thus rescale the unit disk by a factor ε and then apply the exponential at $\mathbf{W}$ to map the rescaled vectors to $G(k,n)$. We can define a corresponding family of metrics $\{g_\varepsilon\}_{\varepsilon \geq 0}$ on D associated to this family of maps via a rescaled pullback as

$$g_\varepsilon := \frac{1}{\varepsilon^2}\phi_\varepsilon^* g, \tag{4.43}$$

where g is the Grassmann metric on the geodesic ball $B(\mathbf{W}, \varepsilon)$ and ϕ_ε^* denotes the pullback via ϕ_ε. The meaning of the new family of metrics $\{g_\varepsilon\}_{\varepsilon \geq 0}$ is that it gives information on the distortion effect on the metric on $T_{\mathbf{W}}G(k,n)$ close to the origin caused by the application of the exponential map. Denoting by g_F the Frobenius metric on $T_{\mathbf{W}}G(k,n)$ (we are now looking at it as a space of matrices), we have the following result.

Lemma 4.47. *The family of metrics g_ε converges C^∞ to the Frobenius metric g_F as $\varepsilon \longrightarrow 0$.*

Proof. Let $A \in D \subseteq T_{\mathbf{W}}G(k,n)$ and $B_1, B_2 \in T_A D \cong \mathbb{R}^{(n-k)\times k}$ be tangent vectors to D at A. Then

$$\begin{aligned}(g_\varepsilon)_A(B_1,B_2) = \left(\frac{1}{\varepsilon^2}\phi_\varepsilon^* g\right)_A (B_1,B_2) = \frac{1}{\varepsilon^2} g_{\phi_\varepsilon(A)}(D_A\phi_\varepsilon(B_1), D_A\phi_\varepsilon(B_2)) = \\ = \frac{1}{\varepsilon^2} g_{\phi_\varepsilon(A)}(\varepsilon D_{\varepsilon A}\exp_{\mathbf{W}}(B_1), \varepsilon D_{\varepsilon A}\exp_{\mathbf{W}}(B_2)) = \\ = g_{\phi_\varepsilon(A)}(D_{\varepsilon A}\exp_{\mathbf{W}}(B_1), D_{\varepsilon A}\exp_{\mathbf{W}}(B_2)),\end{aligned}$$

where $D_A\phi_\varepsilon$ denotes the differential of ϕ_ε at A and $D_{\varepsilon A}\exp_{\mathbf{W}}$ the differential of the exponential map $\exp_{\mathbf{W}}$ at εA. Since the differential of the exponential at the origin $D_0\exp_{\mathbf{W}}$ is the identity, we obtain the wanted C^∞ convergence. □

Proposition 4.48. *There exists $\varepsilon > 0$ such that for the generic $\mathbf{L} \in B(\mathbf{W}, \widetilde{\varepsilon})$, $\mathbf{L} \notin \Omega_s$ with $\widetilde{\varepsilon} < \varepsilon$ and $\mathbf{L} = \exp_{\mathbf{W}}(v_A)$ for some $v_A \in T_{\mathbf{W}}G(k,n)$, the critical points for the restricted distance function $\delta_{\mathbf{L}}|_{\Omega_s}$ contained in the geodesic ball $B(\mathbf{W}, \varepsilon)$ are approximations of the critical points for $\widehat{\delta}_A|_{\mathcal{M}_{k-s}}$ coming from Eckart–Young theorem.*

Remark 4.49. Given a matrix $A \in \mathbb{R}^{(n-k)\times k}$, all its rank–approximations coming from Eckart–Young theorem have norm smaller than $\|A\|_F$, as they are obtained by canceling out some of the singular values.

Proof of Proposition 4.48. We give a sketch of the proof. Observe that the distance between $\mathbf{L}$ and its cut locus is always $\frac{\pi}{2}$. Therefore for ε small enough, the ball $B(\mathbf{W}, \varepsilon)$ will not intersect the cut locus of any $\mathbf{L} \in B(\mathbf{W}, \widetilde{\varepsilon})$ for $\widetilde{\varepsilon} < \varepsilon$. The result then follows directly from Lemma 4.47 and the fact that critical points away from the cut locus are nondegenerate. □

Remark 4.50. By Proposition 4.48 we have that critical points close to $\mathbf{W}$ will be given by approximations of the Eckart–Young ones. But Theorem 4.43 tells us that the critical points coming from Eckart–Young are actually critical points also in our setting, suggesting that the exponential mapping has no distortion effect on critical points. The reason why this happens, admittedly, is still not clear to us and will be subject to further investigation.

As a corollary of Proposition 4.48 and Theorem 4.43 we have that for $\mathbf{L}$ close enough to $\mathbf{W}$, the minimum for $\delta_{\mathbf{L}}|_{\Omega_s}$ is given by the minimum in Eckart–Young theorem, as no other critical point closer to $\mathbf{L}$ than the ones coming from Theorem 4.43 can be found. In the next section we will see that this is not specific to data points close to $\mathbf{W}$, but it is a general fact: for the generic $\mathbf{L} \in G(k,n)$ the minimum for the restricted distance function $\delta_{\mathbf{L}}|_{\Omega_s}$ will always be the one provided by Theorem 4.43.

4.3.4 Global minimum

We will now prove a result that provides the global minimum for the restriction of $\delta_{\mathbf{L}}$ to Ω_s for a generic $\mathbf{L} \in G(k,n)$, thus solving the nearest point problem for this class of simple Schubert varieties.

The geometric idea is the following. Every point $\mathbf{E} \in \Omega_s$ has to to satisfy the condition that $\mathbf{E} \cap \mathbf{W}$ is of dimension at least s. In other words $\mathbf{E}$ has to contain an s–dimensional subspace of $\mathbf{W}$. Suppose that $0 < \theta_1(\mathbf{L}, \mathbf{W}) < \cdots < \theta_k(\mathbf{L}, \mathbf{W}) < \frac{\pi}{2}$ are the principal angles between $\mathbf{L}$ and $\mathbf{W}$. Then we claim that, since $\mathbf{E}$ contains an s–dimensional subspace of $\mathbf{W}$, the squared distance between $\mathbf{E}$ and $\mathbf{L}$ must be at least $\sum_{i=1}^{s} \theta_i(\mathbf{L}, \mathbf{W})^2$, that is to say some of the angles between $\mathbf{E}$ and $\mathbf{L}$ have to be at least the same as the smallest s angles between $\mathbf{W}$ and $\mathbf{L}$; this argument will be made rigorous in the following. As we can show that this quantity is actually achieved by the critical point corresponding to the Eckart–Young global minimizer via Theorem 4.43, we state the following theorem.

Theorem 4.51. *For the generic $\mathbf{L} \in G(k,n)$, the restriction of $\delta_{\mathbf{L}}$ to Ω_s admits a unique global minimizer, corresponding to the minimum given by the Eckart–Young Theorem through Theorem 4.43. Explicitly, if $0 < \theta_1(\mathbf{L}, \mathbf{W}) < \cdots < \theta_k(\mathbf{L}, \mathbf{W}) < \frac{\pi}{2}$ are the angles between $\mathbf{L}$ and $\mathbf{W}$ with corresponding principal vectors $\{q_1, \ldots, q_k\}$ and $\{p_1, \ldots, p_k\}$ for $\mathbf{L}$ and $\mathbf{W}$ respectively, we have*

$$\min \delta_{\mathbf{L}}|_{\Omega_s} = \left(\theta_1(\mathbf{L}, \mathbf{W})^2 + \cdots + \theta_s(\mathbf{L}, \mathbf{W})^2\right)^{\frac{1}{2}},$$

with the unique global minimizer being

$$\mathbf{E}_m = \operatorname{span}\{p_1, \ldots, p_s, q_{s+1}, \ldots, q_k\}. \tag{4.44}$$

The proof will be mainly based on the following proposition ([50, Corollary 3.1.3]) and the subsequent Corollary 4.53.

Proposition 4.52 ([50, Corollary 3.1.3])**.** *Let $B \in \mathbb{R}^{n\times k}$ and let B_r denote a submatrix of B obtained by deleting r columns from B. Ordering the singular values of a matrix in decreasing order, we have*

$$\sigma_i(B) \geq \sigma_i(B_r) \geq \sigma_{i+r}(B),$$

for every $i = 1, \dots, k$, where for a matrix $X \in \mathbb{R}^{p\times q}$ we set $\sigma_j(X) = 0$ if $j > \min\{p, q\}$.

Recall also that the algorithm used to define principal vectors and angles can be extended to define principal vectors and angles between subspaces of different dimensions, exactly in the same way as it was done in [96]. Here we need a modification of [96, Corollary 10], that we state using the same notation of [96].

Corollary 4.53. *Let $\mathbf{A} \in G(k, n)$, $\mathbf{B} \in G(k, n)$ and $\widetilde{\mathbf{B}} \in G(\widetilde{k}, n)$ such that $\widetilde{k} \leq k$ and $\widetilde{\mathbf{B}} \subseteq \mathbf{B}$. Then we have the following inequalities on principal angles*

$$\theta_i(\mathbf{A}, \mathbf{B}) \leq \theta_i(\mathbf{A}, \widetilde{\mathbf{B}}) \leq \theta_{i+k-\widetilde{k}}(\mathbf{A}, \mathbf{B}), \tag{4.45}$$

for every $i = 1, \dots, \widetilde{k}$.

Proof. Let $\{w_1, \dots, w_{\widetilde{k}}\}$ be an orthonormal basis for $\widetilde{\mathbf{B}}$ and complete it to an orthonormal basis of $\mathbf{B}$ $\{w_1, \dots, w_k\}$. Let $\widetilde{B} = (w_1 \dots w_{\widetilde{k}})$ and $B = (w_1 \dots w_k)$, so that the matrix $\widetilde{B}$ is obtained removing the last $k - \widetilde{k}$ columns from the matrix B. Let $A = (a_1 \dots a_k) \in \mathbb{R}^{n\times k}$ be a matrix whose columns are an orhtonormal basis of $\mathbf{A}$. Then we know that

$$\begin{aligned} \theta_i(\mathbf{A}, \mathbf{B}) &= \arccos(\sigma_i(A^\top B)) \quad i = 1, \dots, k, \\ \theta_i(\mathbf{A}, \widetilde{\mathbf{B}}) &= \arccos(\sigma_i(A^\top \widetilde{B})) \quad i = 1, \dots, \widetilde{k}. \end{aligned} \tag{4.46}$$

Since $A^\top \widetilde{B}$ is obtained from $A^\top B$ by deleting $k - \widetilde{k}$ columns, by Proposition 4.52 we have

$$\sigma_i(A^\top B) \geq \sigma_i(A^\top \widetilde{B}) \geq \sigma_{i+k-\widetilde{k}}(A^\top B), \tag{4.47}$$

for every $i = 1, \dots, \widetilde{k}$. Since all the singular values are in $[0, 1]$ and the function arccos is descreasing on that interval, from (4.46) and (4.47) we obtain (4.45).

□

Proof of Theorem 4.51. Let $\mathbf{E} \in \Omega_s$ and denote $\mathbf{V} := \mathbf{E} \cap \mathbf{W}$. Since $\mathbf{E} \in \Omega_s$, we know that $\dim(\mathbf{V}) = \widetilde{s} \geq s$. Let $0 \leq \theta_1(\mathbf{L}, \mathbf{V}) \leq \dots \leq \theta_{\widetilde{s}}(\mathbf{L}, \mathbf{V}) \leq \frac{\pi}{2}$ be the principal angles between $\mathbf{V}$ and $\mathbf{L}$ (remark that these are angles between subspaces of different dimensions). Recall that the distance function from $\mathbf{L}$ is given by

$$\delta_{\mathbf{L}}(\mathbf{E}) = \left(\theta_1(\mathbf{L}, \mathbf{E})^2 + \dots + \theta_k(\mathbf{L}, \mathbf{E})^2\right)^{\frac{1}{2}}.$$

Since $\mathbf{V}$ is a subspace of $\mathbf{W}$, the left side of (4.45) implies that

$$\theta_i(\mathbf{L}, \mathbf{W}) \leq \theta_i(\mathbf{L}, \mathbf{V}) \qquad i = 1, \dots, \widetilde{s}.$$

But $\mathbf{V}$ can also be regarded as a subspace of $\mathbf{E}$, hence the right hand side of (4.45) implies that

$$\theta_i(\mathbf{L}, \mathbf{V}) \leq \theta_{i+k-\widetilde{s}}(\mathbf{L}, \mathbf{E}) \qquad i = 1, \dots, \widetilde{s}.$$

The two inequalities above together give

$$\begin{aligned}\theta_1(\mathbf{L},\mathbf{W}) &\leq \theta_{k-\widetilde{s}+1}(\mathbf{L},\mathbf{E}),\\ &\vdots\\ \theta_{\widetilde{s}}(\mathbf{L},\mathbf{W}) &\leq \theta_k(\mathbf{L},\mathbf{E}).\end{aligned}$$

In turn, this gives the inequality

$$\begin{aligned}\delta_{\mathbf{L}}(\mathbf{E}) &= \left(\theta_1(\mathbf{L},\mathbf{E})^2+\cdots+\theta_k(\mathbf{L},\mathbf{E})^2\right)^{\frac{1}{2}}\\ &\geq \left(\theta_1(\mathbf{L},\mathbf{E})^2+\cdots+\theta_{k-\widetilde{s}}(\mathbf{L},\mathbf{E})^2+\theta_1(\mathbf{L},\mathbf{W})^2+\cdots+\theta_{\widetilde{s}}(\mathbf{L},\mathbf{W})^2\right)^{\frac{1}{2}}\\ &\geq \left(\theta_1(\mathbf{L},\mathbf{W})^2+\cdots+\theta_{\widetilde{s}}(\mathbf{L},\mathbf{W})^2\right)^{\frac{1}{2}}\\ &\geq \left(\theta_1(\mathbf{L},\mathbf{W})^2+\cdots+\theta_s(\mathbf{L},\mathbf{W})^2\right)^{\frac{1}{2}},\end{aligned} \tag{4.48}$$

where in the last inequality we used $\widetilde{s}\geq s$.

It is easy to check that the plane $\mathbf{E}_m$ defined in (4.44) achieves this lower bound by computing principal angles between $\mathbf{E}_m$ and $\mathbf{W}$ in the usual way. This shows that

$$\delta_{\mathbf{L}}(\mathbf{E}_m)=\min\delta_{\mathbf{L}}|_{\Omega_s}=\left(\theta_1(\mathbf{L},\mathbf{W})^2+\cdots+\theta_s(\mathbf{L},\mathbf{W})^2\right)^{\frac{1}{2}}.$$

Moreover, by construction, $\mathbf{E}_m$ is the point obtained in Theorem 4.43 by choosing the approximation corresponding to the minimizer provided by Theorem 2.19 (see Remark 4.44).

To prove that $\mathbf{E}_m$ is the unique minimizer, we start by observing that since all the angles $\theta_i(\mathbf{L},\mathbf{W})$ are positive by genericity of $\mathbf{L}$, if $\widetilde{s}>s$ the inequality (4.48) is strict and the minimum cannot be achieved. Therefore for every minimizer it holds that $\widetilde{s}=s$, i.e. minimizers are always smooth points of Ω_s.

Another necessary condition for the minimum to be achieved by a point $\mathbf{E}$ is that the first $k-s$ angles vanish, i.e.

$$\theta_1(\mathbf{L},\mathbf{E})=\cdots=\theta_{k-s}(\mathbf{L},\mathbf{E})=0,$$

which means that $\dim(\mathbf{E}\cap\mathbf{L})=k-s$, since we know from the proved inequalities that the $(k-s+1)$–th angle will be positive. Together with $\mathbf{E}\in\Omega_s$ and the intersection between $\mathbf{L}$ and $\mathbf{W}$ being trivial, this implies that $\mathbf{E}$ must be an orthogonal direct sum

$$\mathbf{E}=\mathbf{E}_{\mathbf{L}}\oplus\mathbf{E}_{\mathbf{W}},$$

where $\mathbf{E}_{\mathbf{L}}$ is a $(k-s)$–dimensional subspace of $\mathbf{L}$ and $\mathbf{E}_{\mathbf{W}}$ is an s–dimensional subspace of $\mathbf{W}$. By construction, for a minimizer we must have that

$$\begin{aligned}\theta_{k-s+1}(\mathbf{L},\mathbf{E})=\theta_1(\mathbf{L},\mathbf{E}_{\mathbf{W}})&=\theta_1(\mathbf{L},\mathbf{W}),\\ &\vdots\\ \theta_k(\mathbf{L},\mathbf{E})=\theta_s(\mathbf{L},\mathbf{E}_{\mathbf{W}})&=\theta_s(\mathbf{L},\mathbf{W}).\end{aligned}$$

Since $\mathbf{E}_{\mathbf{W}}\subseteq\mathbf{W}$ and every angle $\theta_i(\mathbf{L},\mathbf{W})$ has multiplicitly one by genericity, it follows that $\mathbf{E}_{\mathbf{W}}$ is uniquely determined as the span of the principal vectors of $\mathbf{W}$ corresponding to the angles $\theta_1(\mathbf{L},\mathbf{W}),\ldots,\theta_s(\mathbf{L},\mathbf{W})$ with respect to $\mathbf{L}$, that is

$$\mathbf{E}_{\mathbf{W}}=\operatorname{span}\{p_1,\ldots,p_s\}.$$

Then the $\mathbf{E}_\mathbf{L}$ component is uniquely determined as

$$\mathbf{E}_\mathbf{L} = \mathbf{L} \cap (\mathbf{E}_\mathbf{W})^\perp = \operatorname{span}\{q_{s+1}, \dots, q_k\}.$$

It follows that the minimizer $\mathbf{E}_m$ is unique. □

We have thus completely solved the nearest point problem for the varieties Ω_s for the generic data point $\mathbf{L} \in G(k,n)$. Our solution is algorithmic: given the k–plane $\mathbf{W}$ that defines the variety Ω_s and the data point $\mathbf{L}$, we are able to explicitly compute the principal angles and principal vectors as described in Section 2.7.4. Once we have them, it is then enough to select the vectors in (4.44) to obtain the generically unique minimizer.

4.3.5 Global maximum

We will now apply the same ideas we used in the computation of the minimizer to find the maximizers of the restricted distance function $\delta_\mathbf{L}|_{\Omega_s}$ for a generic $\mathbf{L} \in G(k,n)$. Using Corollary 4.53 in a different order than in above proof and the idea explained in Remark 4.45, we can indeed find explicitly the global maximizers. In particular we will see that there can be infinitely many of them. This is not in contrast with the finiteness of the Grassmann Distance Complexity, since they all belong to the cut locus of the data point $\mathbf{L}$. Moreover, one could check that no geodesic from $\mathbf{L}$ to a global maximum is normal to Ω_s, as we expect from Lemma 4.29.

Theorem 4.54. *For the generic $\mathbf{L} \in G(k,n)$, the restriction of $\delta_\mathbf{L}$ to Ω_s admits a Grassmannian $G(k-s, n-k-s)$ of global maximizers, which can be explicitly computed. If $0 < \theta_1(\mathbf{L},\mathbf{W}) < \cdots < \theta_k(\mathbf{L},\mathbf{W}) < \frac{\pi}{2}$ are the principal angles between $\mathbf{L}$ and $\mathbf{W}$, we have*

$$\max \delta_\mathbf{L}|_{\Omega_s} = \left(\theta_{k-s+1}(\mathbf{L},\mathbf{W})^2 + \cdots + \theta_k(\mathbf{L},\mathbf{W})^2 + (k-s)\frac{\pi}{4}^2\right)^{\frac{1}{2}}.$$

Proof. Let $\mathbf{E} \in \Omega_s$ and denote $\mathbf{V} := \mathbf{E} \cap \mathbf{W}$. Since $\mathbf{E} \in \Omega_s$, we know that $\dim(\mathbf{V}) = \widetilde{s} \geq s$. Since $\mathbf{V}$ is a subspace of $\mathbf{E}$, the left hand side of (4.45) implies that

$$\theta_i(\mathbf{L},\mathbf{E}) \leq \theta_i(\mathbf{L},\mathbf{V}) \quad i = 1, \dots, \widetilde{s}.$$

But $\mathbf{V}$ is also a subspace of $\mathbf{W}$, so that the right hand side of (4.45) gives

$$\theta_i(\mathbf{L},\mathbf{V}) \leq \theta_{i+k-\widetilde{s}}(\mathbf{L},\mathbf{W}) \quad i = 1, \dots, \widetilde{s}.$$

These inequalities together give

$$\begin{aligned} \theta_1(\mathbf{L},\mathbf{E}) &\leq \theta_{k-\widetilde{s}+1}(\mathbf{L},\mathbf{W}) < \frac{\pi}{2}, \\ &\vdots \\ \theta_{\widetilde{s}}(\mathbf{L},\mathbf{E}) &\leq \theta_k(\mathbf{L},\mathbf{E}) < \frac{\pi}{2}, \end{aligned}$$

and the trivial bounds for the remaining angles $\theta_j(\mathbf{L},\mathbf{E}) \leq \frac{\pi}{2}$ for $j = \widetilde{s}+1, \dots, k$. We thus have the following chain of inequalities

$$\begin{aligned} \delta^2_\mathbf{L}(\mathbf{E}) &= \theta_1(\mathbf{L},\mathbf{E})^2 + \cdots + \theta_k(\mathbf{L},\mathbf{E})^2 \\ &\leq \theta_{k-\widetilde{s}+1}(\mathbf{L},\mathbf{W})^2 + \cdots + \theta_k(\mathbf{L},\mathbf{W})^2 + \theta_{\widetilde{s}+1}(\mathbf{L},\mathbf{E})^2 + \cdots + \theta_k(\mathbf{L},\mathbf{E})^2 \\ &\leq \theta_{k-\widetilde{s}+1}(\mathbf{L},\mathbf{W})^2 + \cdots + \theta_k(\mathbf{L},\mathbf{W})^2 + (k-\widetilde{s})\,\frac{\pi}{4}^2 \\ &\leq \theta_{k-s+1}(\mathbf{L},\mathbf{W})^2 + \cdots + \theta_k(\mathbf{L},\mathbf{W})^2 + (k-s)\,\frac{\pi}{4}^2, \end{aligned}$$

where the last inequality follows from $s \leq \widetilde{s}$ and all angles being bounded by $\frac{\pi}{2}$; moreover, since by genericity $\theta_j(\mathbf{L}, \mathbf{W}) < \frac{\pi}{2}$ for every $j = 1, \ldots, k$, this inequality is strict if $s < \widetilde{s}$. Since these inequalities hold for every $\mathbf{E} \in \Omega_s$, it follows that

$$\max \delta_{\mathbf{L}}|_{\Omega_s} \leq \left(\theta_{k-s+1}(\mathbf{L}, \mathbf{W})^2 + \cdots + \theta_k(\mathbf{L}, \mathbf{W})^2 + (k-s)\frac{\pi}{4}^2 \right)^{\frac{1}{2}}. \tag{4.49}$$

We will now build a family of points in Ω_s achieving this bound, proving that it is the maximum and that all maximizers are in the smooth locus of Ω_s. As already pointed out, in order for the inequality to not be strict, we need to have $\widetilde{s} = s$, that is the point $\mathbf{E}$ has to be a smooth point of Ω_s. Moreover we need

$$\begin{aligned} \theta_1(\mathbf{L}, \mathbf{E}) &= \theta_{k-s+1}(\mathbf{L}, \mathbf{W}) \\ &\vdots \\ \theta_s(\mathbf{L}, \mathbf{E}) &= \theta_k(\mathbf{L}, \mathbf{W}) \\ \theta_{s+1}(\mathbf{L}, \mathbf{E}) = \cdots &= \theta_k(\mathbf{L}, \mathbf{E}) = \frac{\pi}{2}. \end{aligned}$$

In particular $\dim(\mathbf{E} \cap \mathbf{L}^\perp) = k - s$. This forces $\mathbf{E}$ to be a direct sum

$$\mathbf{E} = \mathbf{E}_{\mathbf{L}^\perp} \oplus \mathbf{E}_{\mathbf{W}},$$

where $\mathbf{E}_{\mathbf{L}^\perp}$ is a $(k-s)$–dimensional subspace of $\mathbf{L}^\perp$ and $\mathbf{E}_{\mathbf{W}}$ is a s–dimensional subspace of $\mathbf{W}$. For such an $\mathbf{E}$ we can see that

$$\theta_i(\mathbf{L}, \mathbf{E}) = \theta_i(\mathbf{L}, \mathbf{E}_{\mathbf{W}}) = \theta_{k-s+i}(\mathbf{L}, \mathbf{W})$$

for every $i = 1, \ldots, s$. Since every angle $\theta_j(\mathbf{L}, \mathbf{W})$ has multiplicity one by genericity and these are the maximal ones, $\mathbf{E}_{\mathbf{W}}$ is uniquely determined as the subspace corresponding to the last s principal vectors of $\mathbf{W}$ with respect to $\mathbf{L}$. Denote them as $\{p_{k-s+1}, \ldots, p_k\}$. To build the remaining subspace $\mathbf{E}_{\mathbf{L}^\perp}$ we have to select a $(k-s)$–dimensional subspace of

$$\mathbf{A} := \mathbf{L}^\perp \cap \langle p_{k-s+1}, \ldots, p_k \rangle^\perp.$$

The space $\mathbf{A}$ has generically dimension $n-k-s$. Let $G(k-s, \mathbf{A}) \cong G(k-s, n-k-s)$ be the Grassmannian of $(k-s)$–subspaces in $\mathbf{A}$ and pick any $\mathbf{B} \in G(k-s, \mathbf{A})$. Define $\mathbf{E}_{\mathbf{B}} := \mathbf{B} \oplus \mathbf{E}_{\mathbf{W}} = \mathbf{B} \oplus \langle p_{k-s+1}, \ldots, p_k \rangle$. It is easy to check via the usual algorithm that for every choice of $\mathbf{B} \in G(k-s, \mathbf{A})$ the plane $\mathbf{E}_{\mathbf{B}}$ achieves the bound (4.49). These are the unique ones achieving it, as their construction highlighted.

□

Remark 4.55. From the proof of Theorem 4.51 it follows that the minimum distance from $\mathbf{L}$ achieved on the strata of Ω_s increases strictly every time we go to a deeper singular stratum. In the same way, from the proof of Theorem 4.54 we see that the maximum distance strictly decreases. The geometric idea behind this phenomenon is that every time we move to a deeper stratum, we are increasing the constraints on the intersection between the points and $\mathbf{W}$, which reflects on more constraints on the principal angles between those points and $\mathbf{L}$. We can also explicitly describe what the minimizers and maximizers are on each stratum, either by just reading through the proofs above, or by using the fact that the strata of a given Ω_s are themselves the smooth parts of some Ω_t's for $t > s$.

This concludes our study of the nearest point problem for this class of simple Schubert varieties. We have shown how to algorithmically compute global minimizers and maximizers for the generic data point $\mathbf{L} \in G(k, n)$. If the point is not generic, in the sense that either some principal angles between $\mathbf{L}$ and the defining $\mathbf{W}$ are 0, $\frac{\pi}{2}$ or have higher multiplicity, some of the properties above, such as for example uniqueness of the minimizer, might break down. Nevertheless, it easy to see that the proofs and the construction still work also in the non–generic case, up to some slight modifications.

Another question is whether we can generalize this study to other simple Schubert varieties. An improved version of Corollary 4.53 would be needed, where we allow for the spaces $\mathbf{A}$ and $\mathbf{B}$ to have different dimensions, accounting for the defining subspace $\mathbf{W}$ not having fixed dimension k anymore. Nevertheless, the geometric idea behind our study is exactly the same: the defining condition of the simple Schubert variety on the intersection between the k–planes in $G(k, n)$ and the given $\mathbf{W}$ will impose restrictions on the principal angles between the data point $\mathbf{L} \in G(k, n)$ and the points of the variety in terms of the angles between $\mathbf{L}$ and $\mathbf{W}$. We propose to give a full statement and proof of this idea in a future work.

Table of notations

The order follows the order of appearance from Chapter 2 onward.

$(\overline{M}, g)$	Riemannian manifold
$\iota : M \hookrightarrow \overline{M}$	embedding of the submanifold M into $\overline{M}$
$\iota^* g$	pullback of a Riemannian metric g through an embedding ι
TM	tangent bundle of M
T_pM	tangent space to M at the point p
NM	normal bundle of a submanifold M
N_pM	normal space to a submanifold M at the point p
$\overline{\nabla}_X Y$, $\nabla_X Y$	Levi–Civita connections of $\overline{M}$ and M respectively
$B : T_pM \times T_pM \to N_pM$	second fundamental form of $M \hookrightarrow \overline{M}$
H_η	second fundamental form along a normal direction η
L_η	Weingarten operator along a normal direction η
D_pf	differential of a smooth map $f : M \to N$ at $p \in M$
S^n	unit Euclidean sphere in $\mathbb{R}^{n+1}$
$g_{S^n}, g_{\mathbb{R}\mathrm{P}^n}$	round metric on S^n and induced metric on $\mathbb{R}\mathrm{P}^n$
$N^\varepsilon M$	ε–small normal bundle
$\exp : T\overline{M} \to \overline{M}$	Riemannian exponential map of $\overline{M}$
$\exp\|_{NM} : NM \to \overline{M}$	normal exponential map of $M \hookrightarrow \overline{M}$
$d^g(p, q)$	distance induced by the Riemannian metric g
$\mathcal{U}(M, \varepsilon)$	tubular neighborhood of $M \hookrightarrow \overline{M}$ of radius $\varepsilon > 0$
$\gamma_{(p,v)}$	unique geodesic with $\gamma_{(p,v)}(0) = p$ and $\dot{\gamma}_{(p,v)}(0) = v$
$\rho(M)$	reach of $M \hookrightarrow \overline{M}$
$\mathrm{Med}(M)$	medial axis of $M \hookrightarrow \overline{M}$
$\kappa_{\max}(M)$	maximal curvature of $M \hookrightarrow \overline{M}$
$\mathrm{bw}(p, q)$	width of the bottleneck $\{p, q\}$
$\mathrm{bw}(M)$	bottleneck width of $M \hookrightarrow \overline{M}$
K_j	j–th curvature coefficient of $M \hookrightarrow \overline{M}$
J_j	universal functions in Weyl's formula ((2.6),(2.7))
(r, α, β)	format of a Pfaffian function
$\mathbb{R}^{m \times n}$	space of $m \times n$ real matrices
$\langle A, B \rangle_F$	Frobenius scalar product between two matrices A, $B \in \mathbb{R}^{m \times n}$ (2.9); extends to tensors (3.3)
$\|A\|_F$	Frobenius norm of the matrix $A \in \mathbb{R}^{m \times n}$; extends to tensors
$\widehat{\delta}(A, B)$	distance between two matrices A, $B \in \mathbb{R}^{m \times n}$ induced by the Frobenius scalar product (2.10); extends to tensors
$\mathcal{M}_{\leq r}$	variety of $m \times n$ matrices with rank bounded by r
$\mathcal{M}_r$	set of $m \times n$ matrices with rank equal to r
$\widehat{\delta}_A : \mathbb{R}^{m \times n} \to \mathbb{R}$	distance function from $A \in \mathbb{R}^{m \times n}$ defined by $\widehat{\delta}_A(B) = \widehat{\delta}(A, B)$

$O(n)$	orthogonal group of $n \times n$ matrices
$\sigma_i(A)$	i–th singular value of the matrix A
A_I	for an index set $I \subseteq \{1, \dots, m\}$ the matrix defined in Eckart–Young theorem Theorem 2.19
$G(k, n)$	real Grassmannian of k–planes in $\mathbb{R}^n$
$\langle v_1, \dots, v_k \rangle$	linear span of the vectors $v_1, \dots, v_k$
$\mathbf{E}, \mathbf{L}, \mathbf{W}, \mathbf{S}$	elements of $G(k, n)$
$\mathbf{E}_0$	k–plane given by the span of the first k vectors of the standard basis of $\mathbb{R}^n$
$\pi : O(n) \to G(k, n)$	quotient map from $O(n)$ to $G(k, n)$ (2.13)
$\mathrm{Hom}(V, W)$	set of linear maps $V \to W$
$\langle A, B \rangle_{\widetilde{F}}$	rescaled Frobenius product on $O(n)$ (2.14)
$\widehat{E}$	representative in $O(n)$ of the k–plane $\mathbf{E}$
$\mathfrak{so}(n)$	Lie algebra of skew–symemtric $n \times n$ matrices
$\mathrm{Ker}(f)$	Kernel of a linear map f
M_A	skew–symmetric $n \times n$ matrix associated to $A \in \mathbb{R}^{(n-k)\times k}$ defined in (2.16)
v_A	tangent vector to $G(k, n)$ corresponding to a matrix $A \in R^{(n-k)\times k}$ defined in Definition 2.21
φ_A	element of $\mathrm{Hom}(\mathbf{E}, \mathbf{E}^\perp)$ corresponding to a matrix $A \in \mathbb{R}^{(n-k)\times k}$ defined in Definition 2.22
$g_{G(k,n)}$	orthogonally invariant Riemannian *Grassmann* metric on $G(k, n)$
e^X	matrix exponential of the square matrix X
$\exp_{\mathbf{E}}$	Riemannian exponential on $G(k, n)$ at the point $\mathbf{E}$
$\exp_R^{O(n)}$	Riemannian exponential on $O(n)$ at the point R
R_A	block diagonal $n \times n$ matrix associated to $A \in \mathbb{R}^{(n-k)\times k}$ defined in (2.19)
$[M]$	linear subspace spanned by the columns of $M \in \mathbb{R}^{n\times k}$
γ_A	geodesic on $G(k, n)$ corresponding to the tangent vector v_A
$\theta_i(\mathbf{E}, \mathbf{L})$	i–th principal angle between $\mathbf{E}$ and $\mathbf{L}$
(p_i, q_i)	couple of i–th principal vectors between $\mathbf{E}$ and $\mathbf{L}$
$\mathfrak{P}_i$	principal plane spanned by the couple (p_i, q_i)
n_i	vector orthogonal to p_i in $\mathfrak{P}_i$ as in (2.24)
$\delta(\mathbf{E}, \mathbf{L})$	distance between two k–planes $\mathbf{E}, \mathbf{L}$ induced by the Grassmann metric (2.26)
$R_{\frac{\pi}{2}}$	set of $(n-k) \times k$ matrices with singular values bounded by $\frac{\pi}{2}$
$\mathbb{P}(\Lambda^k \mathbb{R}^n)$	projectivization of k–vector on $\mathbb{R}^n$
$\mathbb{F}(V \times W)$	free $\mathbb{F}$–vector space on $V \times W$
$V_1 \otimes \cdots \otimes V_d$	tensor product of d vector spaces $V_1, \dots V_d$
$\mathbb{F}^{n_1 \times \cdots \times n_d}$	tensor product of $\mathbb{F}^{n_1}, \dots, \mathbb{F}^{n_d}$, i.e. the space of d–tensors of format $n_1 \times \cdots \times n_d$ over $\mathbb{F}$
$v_1 \otimes \cdots \otimes v_d$	tensor product of the vectors $v_1, \dots, v_d$
$V^{\otimes d}$	d–th tensor product of V with itself
$\mathcal{S}^d(V)$	d–th symmetric power of V
$\mathcal{S}^d(n)$	symmetric d–tensors on $\mathbb{R}^n$
$v^{\otimes d}$	d–th tensor product of a vector with itself, i.e. a symmetric order–d tensor
$\mathcal{S}^d_1(n)$	symmetric rank–one order–d tensors on $\mathbb{R}^n$

$\mathbb{F}[x_1,\dots,x_n]_{(d)}$	space of homogeneous polynomials of degree d in n variables, $\mathbb{F}=\mathbb{R},\mathbb{C}$
$U(n)$	unitary group of $n\times n$ matrices
$\beta^{\mathbb{C}}$	action of $U(n)$ on $\mathbb{C}[x_1,\dots,x_n]_{(d)}$ by change of variables
β	action of $O(n)$ on $\mathbb{R}[x_1,\dots,x_n]_{(d)}$ by change of variables
$\langle p,q\rangle_{\mathrm{BW}}$	Bombieri–Weyl hermitian/scalar product of two complex/real polynomials p,q
$\mathcal{N}(0,1)$	standard Gaussian variable with 0 mean and variance 1
$\mathrm{Sym}(n,\mathbb{R})$	space of symmetric $n\times n$ real matrices
E_{ij}	elementary $n\times n$ matrix with (i,j)–th entry 1 and the others 0
$\mathrm{GOE}(n)$	Gaussian Orthogonal Ensemble of $n\times n$ matrices
ψ	action of $O(n)$ on $\mathrm{Sym}(n,\mathbb{R})$ by congruence
ϕ_d	isomorphism between $\mathcal{S}^d(n)$ and $\mathbb{R}[x_1,\dots,x_n]_{(d)}$ given by Proposition 3.7
p_T	homogeneous polynomial corresponding to the tensor T via ϕ_d
$\mathbb{R}\mathrm{P}^N$	projectivization $\mathbb{P}(\mathbb{R}[x_0,\dots,x_n]_{(d)})$
$\nu_{n,d}:\mathbb{R}\mathrm{P}^n\to\mathbb{R}\mathrm{P}^N$	real Bombieri–Weyl Veronese embedding
$\mathcal{P}V_{n,d}$	Bombieri–Weyl projective Veronese variety
$\mathcal{V}_{n,d}$	Bombieri–Weyl Veronese variety
$\widehat{\nu}_{n,d}:S^n\to S^N$	spherical Bombieri–Weyl Veronese map
$\mathcal{S}V_{n,d}$	spherical Veronese surface
$V_{n,d}$	spherical Veronese variety
l_b	linear form on $\mathbb{R}^{n+1}$ associated to $b\in\mathbb{R}^{n+1}$
$\pi_N:S^N\to\mathbb{R}\mathrm{P}^N$	covering map from S^N to $\mathbb{R}\mathrm{P}^N$
$\widehat{\pi}_{n,d}:\mathcal{S}V_{n,d}\to\mathcal{P}V_{n,d}$	covering map from $\mathcal{S}V_{n,d}$ to $\mathcal{P}V_{n,d}$
$\mathrm{Vol}_{n,\mathrm{BW}}$	n–dimensional volume induced by the Bombieri–Weyl metric
$\|p\|_{\mathrm{BW}}$	Bombieri–Weyl norm of a homogeneous polynomial
$d^{\mathrm{BW}}(p,q)$	distance between two homoegenous polynomials p,q induced by the Bombieri–Weyl scalar product
proj_W	orthogonal projection onto a linear subspace W
$\varphi_{n,d}$	parametrization of $\mathcal{S}V_{n,d}$ in (3.24)
W,P	subspaces of $N^d_{x_0}\mathcal{S}V_{n.d}$ in the decomposition (3.29)
W_p,P_p	subspaces of $N_p\mathcal{S}V_{n,d}$ in Corollary 3.21
D^n	unit disk in $\mathbb{R}^n$
$\mathring{D}^n$	open unit disk in $\mathbb{R}^n$
vol_M	volume form on a Riemannian manifold M
Jf	Jacobian of a smooth map $f:M\to N$
I_n	$n\times n$ identity matrix
$S(V)$	unit sphere in a normed vector space V
$g_j(Q)$	homogeneous polynomials of degree j defined in (3.35)
$\mathbb{E}[X]$	expectation of a random variable X
S_n	group of permutations on $\{1,\dots,n\}$
$\mathrm{sgn}(\sigma)$	sign on a permutation $\sigma\in S_n$
$\mathrm{fix}(\sigma)$	fixed points of a permutation $\sigma\in S_n$
$s(\sigma)$	number of disjoint transpositions in a permutation $\sigma\in S_n$
$N(k)$	number of permutations in S_n given by exactly k disjoint transpositions
$\mathcal{H}$	tensor product of d Hilbert spaces $\mathcal{H}_1,\dots\mathcal{H}_d$
$\|\phi\rangle\in\mathcal{H}$	quantum state in a Hilbert space

$E_G(\lvert\psi\rangle)$	geometric measure of entanglement as in (3.49)
$\widetilde{E}_G(\lvert\psi\rangle)$	geometric measure of entanglement as in (3.51)
$\operatorname{rank}(T)$	rank of a tensor T
$\langle v_1,\dots,v_s\rangle_{\mathbb{P}}$	projectivization of the linear span of the vectors $v_1,\dots,v_s$
$T_1,\dots,T_\ell$	$m\times n$ slices of a tensor $T\in\mathbb{R}^{m\times n\times\ell}$
$\Sigma_{m,n}$	Segre variety of projective rank–one real $m\times n$ matrices
$\Sigma^{\mathbb{C}}_{m,n}$	Segre variety of projective complex rank–one $m\times n$ matrices
$G_{\mathbb{C}}(k,n)$	complex Grassmannian of k–planes in $\mathbb{C}^n$
$S\subseteq\mathbb{C}\mathrm{P}^3$	determinantal cubic surface in Theorem 3.41
$f\lvert_X$	restriction of a function f to a subset X
$d_u:\mathbb{R}^n\to\mathbb{R}_{\geq 0}$	Euclidean distance function from a point u (4.1)
X^{sm}	smooth locus of an algebraic variety or stratified submanifold
X^{sing}	singular locus of an algebraic variety or stratified submanifold
∇f	gradient of a smooth function f
I_X	defining ideal of an algebraic variety X
$J(f)$	Jacobian matrix of $f=(f_1,\dots,f_s)$
I_{crit}	ideal of critical points in (4.4)
$\operatorname{EDD}(X)$	Euclidean Distance Degree of an algebraic variety X
$\operatorname{aEDD}(X,\omega)$	average Euclidean Distance Degree of an algebraic variety X with respect to the density ω
$\mathcal{E}_X$	ED corrspondence of an algebraic variety X
$\delta_{\mathbf{L}}:G(k,n)\to\mathbb{R}_{\geq 0}$	distance function from $\mathbf{L}$ defined by $\delta_{\mathbf{L}}(\mathbf{E})=\delta(\mathbf{L},\mathbf{E})$
$\partial_p f$	subgradient of a locally Lipschitz function f at the point p
$\omega(f)$	differentiability locus of a locally Lipschitz function f
$\operatorname{co}(A)$	convex hull of the set A
$\operatorname{Crit}(f)$	set of critical points of a locally Lipschitz function f
$\operatorname{cut}(p)$	cut locus of a point p in a Riemannian manifold (M,g)
$\Omega(j)$	j–th stratum of the cut locus defined by (4.7)
A_W	matrix associated to $W\in O(n)$ as defined in Proposition 4.18
C_X	constant in Proposition 4.31 for a smooth subanalytic X
$\operatorname{GDC}(X)$	Grassmann Distance Complexity of a smooth subanalytic X
$\widetilde{G}(2,4)$	oriented Grassmannian of 2–planes in $\mathbb{R}^4$
X_β	hypersurface in $\widetilde{G}(2,4)$ in (4.20)
ι_P	Plücker embedding of $G(k,n)$ in $\mathbb{P}(\Lambda^k\mathbb{R}^n)$
$(\mathcal{U}_{j_1\dots j_k},\psi_{j_1\dots j_k})$	charts for $\mathbb{P}(\Lambda^k\mathbb{R}^n)$
$\operatorname{St}(k,n-k)$	Stiefel manifold of orthonormal k–frames in $\mathbb{R}^{n-k}$
$\operatorname{int}(A)$	interior part of a set A
$c_1(k,n),c_2(k,n)$	pfaffian constants in Theorem 4.35
$\mathbf{W}_i$	i–th element of a complete flag in $\mathbb{R}^n$
Ω_λ	Schubert variety in $G(k,n)$ associated to the $(n-k)\times k$ Young diagram λ
Ω_s	simple Schubert variety defined in (4.39)
$A_{\mathbf{L}}$	matrix associated to a generic $\mathbf{L}$ defined by (4.41)
$\mathbf{L}_I$	element of Ω_s defined in Theorem 4.43
ϕ_ε	rescaling map defined by (4.42)
g_ε	zooming metric defined by (4.43)
$B(\mathbf{W},\varepsilon)$	geodesic ball in $G(k,n)$ centered at $\mathbf{W}$ of radius ε
$\mathbf{E}_m$	minimizer for the nearest point problem of Ω_s (4.44)

Acknowledgements

Acknowledgements will only be included in the final version.

Bibliography

[1] H. Abo, G. Ottaviani, and C. Peterson. Induction for secant varieties of segre varieties. *Trans. Amer. Math. Soc.*, 361:767–792, 02 2009.

[2] R. Ait El Manssour, M Belotti, and C. Meroni. Real lines on random cubic surfaces. *Arnold Math J.*, 7:541–559, 12 2021.

[3] Andrew Alexander, Jee Lee, Mariana Lazar, and Aaron Field. Diffusion tensor imaging of the brain. *Neurotherapeutics : the journal of the American Society for Experimental NeuroTherapeutics*, 4:316–29, 08 2007.

[4] D. Allcock, J.A. Carlson, and D. Toledo. Hyperbolic geometry and moduli of real cubic surfaces. *Ann. Sci. Éc. Norm. Supér. (4)*, 43:69–115, 2010.

[5] Greg W. Anderson, Alice Guionnet, and Ofer Zeitouni. *An introduction to random matrices*, volume 118 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010.

[6] E. Arbarello, M. Cornalba, P. Griffiths, and J.D. Harris. *Geometry of Algebraic Curves: Volume I.* Grundlehren der mathematischen Wissenschaften. Springer New York, 1985.

[7] S. Basu, A. Lerario, E. Lundberg, and C. Peterson. Random fields and the enumerative geometry of lines on real and complex hypersurfaces. *Math. Ann.*, 374:1773–1810, 08 2019.

[8] Saugata Basu and Antonio Lerario. Hausdorff approximations and volume of tubes of singular algebraic sets. *Mathematische Annalen*, August 2022.

[9] Thomas Bendokat, Ralf Zimmermann, and P.-A. Absil. A Grassmann manifold handbook: basic geometry and computational aspects. *Adv. Comput. Math.*, 50(1):Paper No. 6, 51, 2024.

[10] Marcel Berger and Bernard Gostiaux. *Differential geometry: manifolds, curves, and surfaces*, volume 115 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1988. Translated from the French by Silvio Levy.

[11] G. Bergqvist. Exact probabilities for typical ranks of $2 \times 2 \times 2$ and $3 \times 3 \times 2$ tensors. *Linear Algebra Appl.*, 438:663–667, 01 2013.

[12] G. Bergqvist and P. Forrester. Rank probabilities for real random $n \times n \times 2$ tensors. *Electron. Commun. Probab.*, 16:630–637, 2011.

[13] Åke Björck and Gene H. Golub. Numerical methods for computing angles between linear subspaces. *Math. Comp.*, 27:579–594, 1973.

[14] P. Breiding, F. Rydell, E. Shehu, and A. Torres. Line multiview varieties. *SIAM J. Appl. Algebra Geom.*, 7(2):470–504, 2023.

[15] Paul Breiding. How many eigenvalues of a random symmetric tensor are real? *Trans. Amer. Math. Soc.*, 372(11):7857–7887, 2019.

[16] Paul Breiding, Sarah Eggleston, and Andrea Rosana. Typical Ranks of Random Order-Three Tensors. *Int. Math. Res. Not. IMRN*, (4):rnaf018, 2025.

[17] Paul Breiding, Kathlén Kohn, and Bernd Sturmfels. *Metric algebraic geometry*, volume 53 of *Oberwolfach Seminars*. Birkhäuser/Springer, Cham, [2024] ©2024.

[18] Paul Breiding and Sascha Timme. Homotopycontinuation.jl: A package for homotopy continuation in julia. In James H. Davenport, Manuel Kauers, George Labahn, and Josef Urban, editors, *Mathematical Software – ICMS 2018*, pages 458–465, Cham, 2018. Springer International Publishing.

[19] Peter Buergisser. Average volume, curvatures, and euler characteristic of random real algebraic varieties, 2006.

[20] Peter Bürgisser and Felipe Cucker. *Condition*, volume 349 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Heidelberg, 2013. The geometry of numerical algorithms.

[21] Ana Cannas da Silva. *Lectures on symplectic geometry*, volume 1764 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2001.

[22] Enrico Carlini, Nathan Grieve, and Luke Oeding. Four lectures on secant varieties. In *Connections between algebra, combinatorics, and geometry*, volume 76 of *Springer Proc. Math. Stat.*, pages 101–146. Springer, New York, 2014.

[23] Fabrizio Catanese and Cecilia Trifogli. Focal loci of algebraic varieties. I. volume 28, pages 6017–6057. 2000. Special issue in honor of Robin Hartshorne.

[24] A. Cayley. On the triple tangent planes of surfaces of the third order. *Camb. Dublin Math. J.*, 4:118–138, 1849.

[25] Alberto Cazzaniga, Antonio Lerario, and Andrea Rosana. What is the probability that a random symmetric tensor is close to rank-one? *SIAM J. Appl. Algebra Geom.*, 8(2):227–258, 2024.

[26] F. H. Clarke. *Optimization and nonsmooth analysis*, volume 5 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, second edition, 1990.

[27] A. Clebsch. Die Geometrie auf den Flächen dritter Ordnung. *Journal für die reine und angewandte Mathematik*, 65:359–380, 1866.

[28] Pierre Comon. Ma identification using fourth order cumulants. *Signal Processing*, 26(3):381–388, 1992.

[29] E. Connelly, S. Agarwal, A. Ergur, and R.R. Thomas. The geometry of rank drop in a class of face-splitting matrix products: Part I. *Adv. Geom.*, 24:369–394, 2024.

[30] Sandra Di Rocco, David Eklund, and Oliver Gäfvert. Sampling and homology via bottlenecks. *Math. Comp.*, 91(338):2969–2995, 2022.

[31] Sandra Di Rocco, David Eklund, and Madeleine Weinstein. The bottleneck degree of algebraic varieties. *SIAM J. Appl. Algebra Geom.*, 4(1):227–253, 2020.

[32] Manfredo Perdigão do Carmo. *Riemannian geometry*. Mathematics: Theory & Applications. Birkhäuser Boston, Inc., Boston, MA, 1992. Translated from the second Portuguese edition by Francis Flaherty.

[33] Jan Draisma and Emil Horobeţ. The average number of critical rank-one approximations to a tensor. *Linear Multilinear Algebra*, 64(12):2498–2518, 2016.

[34] Jan Draisma, Emil Horobeţ, Giorgio Ottaviani, Bernd Sturmfels, and Rekha R. Thomas. The Euclidean distance degree of an algebraic variety. *Found. Comput. Math.*, 16(1):99–149, 2016.

[35] Carl Eckart and Gale Young. The approximation of one matrix by another of lower rank. *Psychometrika*, 1(3):211–218, 1936.

[36] Alan Edelman and Eric Kostlan. How many zeros of a random polynomial are real? *Bull. Amer. Math. Soc. (N.S.)*, 32(1):1–37, 1995.

[37] László M. Fehér and Ákos K. Matszangosz. Real solutions of a problem in enumerative geometry. *Period. Math. Hungar.*, 73(2):137–156, 2016.

[38] S. Friedland. On the generic and typical ranks of 3-tensors. *Linear Algebra Appl.*, 436:478–497, 02 2012.

[39] S. Friedland and A. Libgober. Generalizations of the odd degree theorem and applications. *Israel J. Math.*, 136:353–371, 12 2003.

[40] Shmuel Friedland. Best rank one approximation of real symmetric tensors can be chosen symmetric. *Front. Math. China*, 8(1):19–40, 2013.

[41] William Fulton. *Young tableaux*, volume 35 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 1997. With applications to representation theory and geometry.

[42] William Fulton and Joe Harris. *Representation theory*, volume 129 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1991. A first course, Readings in Mathematics.

[43] Andrei Gabrielov and Nicolai Vorobjov. Complexity of computations with Pfaffian and Noetherian functions. In *Normal forms, bifurcations and finiteness problems in differential equations*, volume 137 of *NATO Sci. Ser. II Math. Phys. Chem.*, pages 211–250. Kluwer Acad. Publ., Dordrecht, 2004.

[44] I. M. Gelfand, M. M. Kapranov, and A. V. Zelevinsky. *Discriminants, resultants and multidimensional determinants*. Modern Birkhäuser Classics. Birkhäuser Boston, Inc., Boston, MA, 2008. Reprint of the 1994 edition.

[45] Alfred Gray. *Tubes*, volume 221 of *Progress in Mathematics*. Birkhäuser Verlag, Basel, second edition, 2004. With a preface by Vicente Miquel.

[46] D. Gross, S. T. Flammia, and J. Eisert. Most quantum states are too entangled to be useful as computational resources. *Phys. Rev. Lett.*, 102(19):190501, 4, 2009.

[47] Andrea Guidolin, Antonio Lerario, Isaac Ren, and Martina Scolamiero. Morse theory of euclidean distance functions and applications to real algebraic geometry. *https://arxiv.org/abs/2402.08639*, 2024.

[48] J. Harris. Algebraic geometry: a first course. *Grad. Texts in Math.*, 133, 1992.

[49] Morris W. Hirsch. *Differential topology*, volume 33 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1994. Corrected reprint of the 1976 original.

[50] Roger A. Horn and Charles R. Johnson. *Topics in matrix analysis*. Cambridge University Press, Cambridge, 1991.

[51] Ralph Howard. The kinematic formula in Riemannian homogeneous spaces. *Mem. Amer. Math. Soc.*, 106(509):vi+69, 1993.

[52] Shenglong Hu, Liqun Qi, and Guofeng Zhang. Computing the geometric measure of entanglement of multipartite pure states by means of non-negative tensors. *Phys. Rev. A*, 93(1):012304, 7, 2016.

[53] Robert Hübener, Matthias Kleinmann, Tzu-Chieh Wei, Carlos González-Guillén, and Otfried Gühne. Geometric measure of entanglement for symmetric states. *Phys. Rev. A (3)*, 80(3):032324, 5, 2009.

[54] C. Itzykson and M. Nauenberg. Unitary groups: Representations and decompositions. *Rev. Modern Phys.*, 38:95–120, 1966.

[55] Michel Journée, Yurii Nesterov, Peter Richtárik, and Rodolphe Sepulchre. Generalized power method for sparse principal component analysis. *J. Mach. Learn. Res.*, 11:517–553, 2010.

[56] A. G. Khovanskiĭ. *Fewnomials*, volume 88 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1991. Translated from the Russian by Smilka Zdravkovska.

[57] A. G. Khovanskiĭ i. A class of systems of transcendental equations. *Dokl. Akad. Nauk SSSR*, 255(4):804–807, 1980.

[58] T. G. Kolda and B. W. Bader. Tensor decompositions and applications. *SIAM Rev.*, 51(3):455–500, 2009.

[59] S. E. Kozlov. Geometry of real Grassmannian manifolds. I, II, III. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, 246:84–107, 108–129, 197–198, 1997.

[60] Pieter M. Kroonenberg. *Applied multiway data analysis*. Wiley Series in Probability and Statistics. Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, 2008. With a foreword by Willem J. Heiser and Jarqueline Meulman.

[61] J. B. Kruskal. Rank, decomposition, and uniqueness for 3-way and N-way arrays. In *Multiway data analysis (Rome, 1988)*, pages 7–18. North-Holland, Amsterdam, 1989.

[62] P.-A. Absil Kyle A. Gallivan. Note on the convex hull of the stiefel manifold. preprint on webpage at https://www.math.fsu.edu/%7Ealuffi/archive/paper386.pdf.

[63] J. M. Landsberg. *Tensors: geometry and applications*, volume 128 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2012.

[64] John M. Lee. *Introduction to smooth manifolds*, volume 218 of *Graduate Texts in Mathematics*. Springer, New York, second edition, 2013.

[65] A. Lerario. Lectures on metric algebraic geometry. preprint on webpage at https://drive.google.com/file/d/1yTaym4-NUwdjHymy0FUewYuzSG_Tl8FA/view.

[66] Antonio Lerario and Andrea Rosana. The grassmann distance complexity. https://arxiv.org/abs/2411.16589, 2024.

[67] L. Mathis. The handbook of zonoid calculus. *PhD thesis*, 2022.

[68] Peter McCullagh. *Tensor methods in statistics*. Monographs on Statistics and Applied Probability. Chapman & Hall, London, 1987.

[69] Madan Lal Mehta. *Random matrices*. Academic Press, Boston, New York, San Diego, 1991.

[70] Albert Nijenhuis. On Chern's kinematic formula in integral geometry. *J. Differential Geometry*, 9:475–482, 1974.

[71] I. Polo-Blanco and J. Top. Explicit real cubic surfaces. *Canad. Math. Bull.*, 51:125–133, 03 2008.

[72] Liqun Qi and Ziyan Luo. *Tensor analysis*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 2017. Spectral theory and special tensors.

[73] T Sakata, T. Sumi, and M. Miyazaki. *Algebraic and Computational Aspects of Real Tensor Ranks*. Springer Briefs in Statistics. Springer, 2016.

[74] Toshio Sakata. *Applied Matrix and Tensor Variate Data Analysis*. 01 2016.

[75] George Salmon. *A Treatise on the Higher Plane Curves*. Hodges and Smith, 1852.

[76] George Salmon. *A Treatise on the Analytic Geometry of Three Dimensions*. Hodges, Smith, and Company, 1865.

[77] L. Schläfli. An attempt to determine the twenty-seven lines upon a surface of the third order, and to divide such surfaces into species in reference to the reality of the lines upon the surface. *Q. J. Pure Appl. Math.*, 2:110–120, 1858.

[78] G Shakhnarovich, T. Darrell, and P. Indyk. *Nearest-Neighbor Methods in Learning and Vision: Theory and Practice*. Neural Information Processing Series. The MIT Press, 2006.

[79] Abner Shimony. Degree of entanglement. In *Fundamental problems in quantum theory (Baltimore, MD, 1994)*, volume 755 of *Ann. New York Acad. Sci.*, pages 675–679. New York Acad. Sci., New York, 1995.

[80] M. Shub and S. Smale. Complexity of Bezout's theorem. II. Volumes and probabilities. In *Computational algebraic geometry (Nice, 1992)*, volume 109 of *Progr. Math.*, pages 267–285. Birkhäuser Boston, Boston, MA, 1993.

[81] Michael Shub and Steve Smale. Complexity of Bézout's theorem. I. Geometric aspects. *J. Amer. Math. Soc.*, 6(2):459–501, 1993.

[82] Michael Shub and Steve Smale. Complexity of Bezout's theorem. III. Condition number and packing. volume 9, pages 4–14. 1993. Festschrift for Joseph F. Traub, Part I.

[83] Nicholas D. Sidiropoulos, Lieven De Lathauwer, Xiao Fu, Kejun Huang, Evangelos E. Papalexakis, and Christos Faloutsos. Tensor decomposition for signal processing and machine learning. *IEEE Trans. Signal Process.*, 65(13):3551–3582, 2017.

[84] A. Smilde, Rasmus Bro, and P. Geladi. Multi way analysis — applications in chemical sciences. 01 2004.

[85] J. Spanier, K. B. Oldham, and J. Myland. *An atlas of functions*. Springer, second edition, 2009.

[86] T. Sumi, M. Miyazaki, and T. Sakata. Typical ranks for 3-tensors, nonsingular bilinear maps and determinantal ideals. *J. Algebra*, 471:409–453, 2017.

[87] T. Sumi, M. Miyazaki, and T. Sakata. Typical ranks of semi-tall real 3-tensors. *Linear Multilinear Algebra*, 69:1734–1746, 2021.

[88] J.M.F. ten Berge. The typical rank of tall three-way arrays. *Psychometrika*, 65:525–532, 12 2000.

[89] J.M.F. ten Berge. Partial uniqueness in CANDECOMP/PARAFAC. *J. Chemometrics*, 18:12–16, 05 2004.

[90] J.M.F. ten Berge and H.A.L. Kiers. Simplicity of core arrays in three-way principal component analysis and the typical rank of $p \times q \times 2$ arrays. *Linear Algebra Appl.*, 294:169–179, 06 1999.

[91] J.M.F. ten Berge and A. Stegeman. Symmetry transformations for square sliced three-way arrays, with applications to their typical rank. *Linear Algebra Appl.*, 418:215–224, 10 2006.

[92] C. Trifogli. Focal loci of algebraic hypersurfaces: a general theory. *Geom. Dedicata*, 70(1):1–26, 1998.

[93] Lou van den Dries. *Tame topology and o-minimal structures*, volume 248 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1998.

[94] Hermann Weyl. On the Volume of Tubes. *Amer. J. Math.*, 61(2):461–472, 1939.

[95] Yung-chow Wong. Differential geometry of Grassmann manifolds. *Proc. Nat. Acad. Sci. U.S.A.*, 57:589–594, 1967.

[96] Ke Ye and Lek-Heng Lim. Schubert varieties and distances between subspaces of different dimensions. *SIAM J. Matrix Anal. Appl.*, 37(3):1176–1197, 2016.

[97] Thierry Zell. Betti numbers of semi-Pfaffian sets. *J. Pure Appl. Algebra*, 139(1):323–338, 1999.

[98] Tong Zhang and Gene H. Golub. Rank-one approximation to high order tensors. *SIAM J. Matrix Anal. Appl.*, 23(2):534–550, 2001.