Algebraic Geometry for Computer Vision

by

Joseph David Kileel

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Mathematics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Bernd Sturmfels, Chair
Professor Laurent El Ghaoui
Assistant Professor Nikhil Srivastava
Professor David Eisenbud

Spring 2017

Algebraic Geometry for Computer Vision

Abstract

Algebraic Geometry for Computer Vision

by

Joseph David Kileel

Doctor of Philosophy in Mathematics

University of California, Berkeley

Professor Bernd Sturmfels, Chair

This thesis uses tools from algebraic geometry to solve problems about three-dimensional scene reconstruction. 3D reconstruction is a fundamental task in multi-view geometry, a eld of computer vision. Given images of a world scene, taken by cameras in unknown positions, how can we best build a 3D model for the scene? Novel results are obtained for various challenging minimal problems, which are important algorithmic routines in Random Sampling Consensus pipelines for reconstruction. These routines reduce over tting when outliers are present in image data.

Our approach throughout is to formulate inverse problems as structured systems of polynomial equations, and then to exploit underlying geometry. We apply numerical algebraic geometry, commutative algebra and tropical geometry, and we derive new mathematical results in these elds. We present simulations on image data as well as an implementation of general-purpose homotopy-continuation software for implicitization in computational algebraic geometry.

Chapter 1 introduces some relevant computer vision. Chapters 2 and 3 are devoted to the recovery of camera positions from images. We resolve an open problem concerning two calibrated cameras raised by Sameer Agarwal, a vision expert at Google Research, by using the algebraic theory of Ulrich sheaves. This gives a robust test for identifying outliers in terms of spectral gaps. Next, we quantify the algebraic complexity for notorious poorly understood cases for three calibrated cameras. This is achieved by formulating in terms of structured linear sections of an explicit moduli space and then computing via homotopy-continuation. In Chapter 4, a new framework for modeling image distortion is proposed, based on lifting algebraic varieties in projective space to varieties in other toric varieties. We check that our formulation leads to faster and more stable solvers than the state of the art. Lastly, Chapter 5 concludes by studying possible pictures of simple objects, as

varieties inside products of projective planes. In particular, this dissertation exhibits that algebro-geometric methods can actually be useful in practical settings.

To Angela

Contents

List of Figures

List of Tables

Acknowledgments

I have been fortunate to have Bernd Sturmfels as an advisor. Thank you for suggesting this topic, for your amazing mentorship, and for keeping me somewhat on track (despite my best efforts). Prof. Sturmfels has had a fantastic impact on me.

Appreciation goes to my other committee members as well. Prof. El Ghaoui taught me convex optimization, Prof. Srivastava introduced me to applications of real-rootedness and Prof. Eisenbud's weekly seminar has been unmatched.

I am lucky to have benefitted from numerous collaborations at an early stage. Thanks to Cristiano Bocci, Enrico Carlini, Justin Chen, Gunnar Fløystad, Michael Joswig, Zuzana Kukelova, Giorgio Ottaviani, Tomas Pajdla, Bernd Sturmfels and André Wagner. I learned a lot from our work together, and had fun in the process.

Cheers to friends and classmates from Berkeley, starting with Justin, for helping with commutative algebra so many times, and extending to Anna, Benson, Bo, Chris, Elina, Franco, Frank, Grace, Jacob, Jose, Kieren, Mengyuan and Steven.

I am deeply gracious to my family for their support throughout my studies. I acknowledge generous funding from the Berkeley Fellowship, Chateaubriand Fellowship, Einstein Berlin Foundation, Max Planck Society and National Science Foundation. I got to travel across the world during my PhD, to the significant benefit of my research, and that would have been impossible in the absence of this support. In particular, I thank Didier Henrion for hosting my extended visit to LAAS-CNRS.

Additionally, acknowledgements are in order for the following:

- In Chapter 2, thanks to Anton Fonarev for pointing out the connection to Littlewood complexes in Remark 2.14, Frank Schreyer for useful conversations about [34] and Steven Sam for help with PieriMaps.
- In Chapter 3, thanks to Jonathan Hauenstein for teaching me how to use Bertini, Luke Oeding for many conversations about trifocal tensors, Kristian Ranestad for an inspiring visit to Oslo, and Frederick Kahl for his interest. Thanks also to Anton Leykin for being very supportive of our Macaulay2 package NumericalImplicitization, and David Eisenbud for conversations.
- In Chapter 4, thanks to Sameer Agarwal, Max Lieblick and Rekha Thomas, for organizing the May 2016 AIM workshop on *Algebraic Vision*, where this work began, and to Andrew Fitzgibbon for ideas about our CVPR paper.
- In Chapters 2-5, thanks, of course, to the authors of [5] for initiating the bridge between the applied algebraic geometry community and computer vision.

Lastly, I am so happy to thank Angela. Here is what I've been working on all this while. This thesis is dedicated to you.

Chapter 1

Motivation

As humans, we may take it for granted that three-dimensional structure can be inferred from two-dimensional images. Our visual perception systems do this naturally. While the neural processes behind this are fantastically complex, it is worth noting that retinal motion makes the reconstruction possible in the rst place [55]. To go from 2D to 3D, our brains use multiple images provided by eye movement.

In computer vision, estimating a 3D scene from multiple 2D images has been a fundamental task. Nowadays, Structure-from-Motion (SfM) algorithms support autonomously-driving cars [40] and large-scale photo tourism [2].

Figure 1.1: 3D model with 819,242 points of the Colosseum from 2106 Flickr images.

Such algorithms have diverse ingredients under the hood: band-pass filters, non-linear least squares optimization, sparse linear algebra, text retrieval ideas, distributed computing, and . . . algebraic geometry! In fact, projective geometry is the language used for formulating 3D reconstruction problems, as explained in the next subsection. The subfield of vision that studies connections with projective geometry is known as multiview geometry. The book [48] by Hartley and Zisserman is the standard introduction to this field.

In addition, SfM repeatedly solves zero-dimensional systems of polynomial equations [64]. Polynomial solvers are subroutines in Random Sampling Consensus (RANSAC) methods for robust estimation, i.e. regression in the presence of outliers. To deliver in real-time, minimal solvers are required to perform accurate, super-fast calculation (μs or ms scale).

In this dissertation, novel vision results are obtained by means of applying tools from algebraic geometry that are not traditionally used in multiview geometry or the design of minimal solvers.

1.1 Setup

According to the pinhole camera model [48, Chapter 6], a camera is simply a surjective linear projection $A : \mathbb{P}^3 \dashrightarrow \mathbb{P}^2$, where $\mathbb{P}^3$ represents the world and $\mathbb{P}^2$ represents the image plane. Thus, A is represented by a full rank real 3×4 matrix up to nonzero scale, also denoted by A. On affine charts, note that the map A is fractional affine-linear. The base locus $\ker(A) \in \mathbb{P}^3$ is interpreted as the camera center or focal point. For a camera A with center off the plane at infinity, RQ factorization applied to the left 3×3 submatrix of A induces a unique factorization $A = K[R|t]$, where $K \in \mathbb{R}^{3 \times 3}$ is upper triangular, with entries $K_{11} > 0, K_{22} > 0, K_{33} = 1$, where $R \in SO(3)$ is orthogonal and where $t \in \mathbb{R}^3$. Following [48, Section 6.2.4], K stores the internal parameters of A (focal lengths, principal point, skew) while $[R|t]$ stores the external parameters (center, orientation). In cases where K is known, left multiplication by K^{-1} normalizes $A = [R|t]$. In that case, the 3×4 matrix A is *calibrated*; calibration information is often available from image EXIF tags.

Standard Structure-from-Motion algorithms [85] perform detailed local reconstructions *first*. Afterwards, these are stitched together and refined via global optimization. Here, a local reconstruction accepts a small number of overlapping images (typically, two or three). The aim is to estimate the configuration in $\mathbb{P}^3$ of the two or three cameras that captured the images, as well as the coordinates of a large collection of 3D points visible in the images. In particular, the cameras' relative positions are deduced from the images, and this is an engine through all of SfM.

Recovery of camera configurations starts by matching features across images (for example, corner points and edges) according to neighborhood intensity patterns; see [73] details. The image matches impose constraints on the possible relative position of the cameras, e.g. [46]. At this point, an appropriately chosen loss function could be defined (see [48, Section 4.2] for so-called algebraic or geometric loss functions). Given the image matches, the camera configuration with least loss could be sought. However, in practice, this delivers poor results, because a non-neglible fraction of the putative image correspondences are wrongly matched. Thus, SfM must cope with outliers (mismatches) among image data.

To that end, Random Sampling Consensus is a method for parameter estimation in the presence of outliers. Invented in 1981 originally for vision applications [37], RANSAC randomly samples a *minimal* amount of data. Minimal means that the sample *exactly* determines only a finite (positive) number of possible parameter values. Those parameters are computed, and then treated as *competing hypotheses*. Each is tested against the rest of the data set. A hypothesis is accepted if it is approximately consistent with a sufficiently high fraction of the full data set (and more than any other hypothesis). Otherwise, a new minimal sample is drawn. RANSAC outputs a parameter estimate uninfluenced by outliers. Remarkably, it can process data sets with as high as 50% outliers. See Figure 1.2 for an illustration.

Thus, to recover camera configurations from image correspondences (containing some mismatches), SfM employs a RANSAC scheme. See [48, Section 4.7] for implementation details, including how thresholds are set adaptively. As an upshot, computing the finitely many parameters consistent with a minimal sample is a vital workhorse in SfM – repeated thousands of times in large-scale reconstructions. These calculations are called *minimal problems*. There is an industry in computer vision dedicated to building efficient solvers for minimal problems, e.g. [17, 38, 57, 64, 65, 94].

Like the matrix camera model above, minimal problems are *algebraic*. They are expressible as systems of multivariate polynomial equations with coefficients depending polynomially on image data. Moreover, a *geometric* formulation is often available. Frequently, a (fixed) algebraic variety $X \subseteq \mathbb{P}^n$ may be defined whose points are in bijection with camera configurations. Here X is an explicit moduli space, embedded in convenient coordinates. Image data defines (varying) linear subspaces $L \subseteq \mathbb{P}^n$. Then, minimal problems amount to computing the intersection $L \cap X$. As a noted example, Nistér's minimal problem solver [82] for recovering the relative position of two calibrated cameras from five image point pair matches fits into this framework; by now, the Gröbner basis script is *hardcoded* into most smartphones [66].

The relation of projective geometry and polynomial equations to 3D reconstruction is this dissertation's point de départ. Mixing classical algebraic geometry with

Figure 1.2: Fitted line from RANSAC. Outliers do not degrade the estimate.

modern computational tools, we answer concrete questions about computer vision and derive new math.

1.2 Main contributions

The main contributions of this dissertation are the following:

We obtain a matrix formula characterizing which six image point pairs are exactly consistent with two calibrated cameras (Theorem 2.1). This resolves a question raised in [1] by Sameer Agarwal, a vision expert apart of Google Research. Numerical experiments indicate the formula is robust to noise (Empirical Fact 2.27), thus it might be used for screening wrongly matched point pairs. Mathematically, the work is an instantiation of the theory of Ulrich sheaves, introduced in algebraic geometry by Eisenbud and Schreyer in [34]. A new determinantal description of the essential variety (Proposition 2.7) a ords a group action making Eisenbud-Schreyer's theory e ective in this case, by help from the representation theory of GL(4).

We quantify the algebraic complexity for the recovery of three calibrated cameras, given various sorts of image correspondences (Theorem 3.6). This helps clarify decades of partial progress on the three camera case (e.g. see [83] for nice complementary work). We build on the theory of trifocal tensors [46], and rely on powerful computational techniques from numerical algebraic geometry [11].

We contribute general-purpose homotopy-continuation software for implicitization in computational algebraic geometry (Section 3.7). This allows for the computation of invariants of an algebraic variety from a parametrization, when de ning equations are inaccessible.

We develop a new framework for modeling image distortion (Chapter 4), unifying existing models. The theory is based on lifting algebraic varieties in projective space to other ambient toric varieties, and it is of independent mathematical interest. We determine degrees in terms of the Chow polytope as well as de ning equations (Theorems 4.8 and 4.16). Tropical geometry [74] o ers a perspective on higher-dimensional distortions (Theorem 4.22).

We verify that our algebro-geometric theory of distortion leads to minimal solvers in vision that are competitive with, or superior to, the state of the art, as tested on synthetic data sets (Section 4.5).

We explore the space of possible pictures of simple objects, such as edges. The formulation is in terms of combinatorial commutative algebra, and we nd equations cutting the space out (Theorem 5.6). This works extends the in uential [5] to new settings of practical interest. It could form the basis for a polynomial/semide nite optimization [69] triangulation scheme, as in [3].

Chapter 2

Two Cameras

This chapter studies the recovery of the relative position of two calibrated cameras from image data. In particular, we are interested in the over-determined case. We characterize which super-minimal samples of image data are exactly consistent with a camera configuration. This connects to the classical theory of resultants [43]. To obtain an explicit result, we need the technology developed in [34]. This is joint work with Gunnar Fløystad and Giorgio Ottaviani [39] and it is to be published in the Journal of Symbolic Computation.

2.1 Introduction

The essential variety E is the variety of 3×3 real matrices with two equal singular values, and the third one equal to zero ($\sigma_1 = \sigma_2$, $\sigma_3 = 0$). It was introduced in the setting of computer vision; see [48, Section 9.6]. Its elements, the so-called essential matrices, have the form TR, where T is real skew-symmetric and R is real orthogonal. The essential variety is a cone of codimension 3 and degree 10 in the space of 3×3-matrices, defined by homogeneous cubic equations, that we recall in (2.2). The complex solutions of these cubic equations define the complexification $E_{\mathbb{C}}$ of the essential variety. This lives in the 8-dimensional complex projective space $\mathbb{P}^8_{\mathbb{C}}$. While the real essential variety is smooth, its complexification has a singular locus that we describe precisely in Section 2.2.

The Chow form of a codimension-c projective variety $X \subset \mathbb{P}^n$ is the equation $Ch(X)$ of the divisor in the Grassmannian $Gr(\mathbb{P}^{c-1}, \mathbb{P}^n)$ given by those linear subspaces of dimension $c-1$ which meet X. It is a basic and classical tool that allows one to recover much geometric information about X; for its main properties we refer to [43, Section 4]. In [1, Section 4], the problem of computing the Chow form of the essential variety was posed, while the analogous problem for the fundamental variety was solved, another important variety in computer vision.

The main goal of this chapter is to explicitly find the Chow form of the essential variety. This provides an important tool for the problem of detecting if a set of image point correspondences $\{(x^{(i)}, y^{(i)}) \in \mathbb{R}^2 \times \mathbb{R}^2 \mid i = 1, \ldots, m\}$ comes from m world points in $\mathbb{R}^3$ and two calibrated cameras. It furnishes an exact solution for $m = 6$ and it behaves well given noisy input, as we will see in Section 2.4. Mathematically, we can consider the system of equations:

$$\begin{cases} A\tilde{X}^{(i)} \sim \tilde{x}^{(i)} \\ B\tilde{X}^{(i)} \sim \tilde{y}^{(i)}. \end{cases} \tag{2.1}$$

Here $\tilde{x}^{(i)} = (x_1^{(i)} : x_2^{(i)} : 1)^T \in \mathbb{P}^2$ and $\tilde{y}^{(i)} = (y_1^{(i)} : y_2^{(i)} : 1)^T \in \mathbb{P}^2$ are the given image points. The unknowns are two 3×4 matrices A, B with rotations in their left 3×3 blocks and $m = 6$ points $\tilde{X}^{(i)} \in \mathbb{P}^3$. These represent calibrated cameras and world points, respectively. A calibrated camera has normalized image coordinates, as explained in [48, Section 8.5]. Here $\sim$ denotes equality up to nonzero scale. From our calculation of $\mathrm{Ch}(\mathcal{E}_{\mathbb{C}})$, we deduce:

Theorem 2.1. *There exists an explicit 20×20 skew-symmetric matrix $M(x, y)$ of degree $(6, 6)$ polynomials over $\mathbb{Z}$ in the coordinates of $(x^{(i)}, y^{(i)})$ with the following properties. If (2.1) admits a complex solution then $M(x^{(i)}, y^{(i)})$ is rank-deficient. Conversely, the variety of point correspondences $(x^{(i)}, y^{(i)})$ such that $M(x^{(i)}, y^{(i)})$ is rank-deficient contains a dense open subset for which (2.1) admits a complex solution.*

In fact, we will produce two such matrices. Both of them, along with related formulas we derive, are available in ancillary files accompanying the arXiv version of this work, and we have posted them at http://math.berkeley.edu/~jkileel/ChowFormulas.html.

Our construction of the Chow form uses the technique of *Ulrich sheaves* introduced in [34]. We construct rank 2 Ulrich sheaves on the essential variety $\mathcal{E}_{\mathbb{C}}$. For an analogous construction of the Chow form of K3 surfaces, see [7].

From the point of view of computer vision, this chapter contributes a complete characterization for an 'almost-minimal' problem. Here the motivation is *3D reconstruction*. Given multiple images of a world scene, taken by cameras in an unknown configuration, we want to estimate the camera configuration and a 3D model of the world scene. Algorithms for this are complex, and successful. See [2] for a reconstruction from 150,000 images.

By contrast, the system (2.1) encodes a tiny reconstruction problem. Suppose we are given six point correspondences in two calibrated pictures (the right-hand sides in (2.1)). We wish to reconstruct both the two cameras and the six world points (the left-hand sides in (2.1)). If an exact solution exists then it is typically unique, modulo the natural symmetries. However, an exact solution does not always exist. In order for this to happen, a giant polynomial of degree 120 in the 24 variables on the right-hand sides has to vanish. Theorem 2.1 gives an explicit matrix formula for that polynomial.

As explained in Chapter 1, the link between minimal or almost-minimal reconstructions and large-scale reconstructions is surprisingly strong. Algorithms for the latter use the former reconstructions repeatedly as core subroutines. In particular, solving the system (2.1) given $m = 5$ point pairs, instead of $m = 6$, is a subroutine in [2]. This solver is optimized in [82]. It is used to generate hypotheses inside **Random Sampling Consensus** (RANSAC) [37] schemes for robust reconstruction from pairs of calibrated images. See [48] for more vision background.

The rest of this chapter is organized as follows. In Section 2.2, we prove that E_C is a hyperplane section of the variety $PX^s_{4;2}$ of 4 4 symmetric matrices of rank 2. This implies a determinantal description of E_C; see Proposition 2.7. A side result of the construction is that E_C is the secant variety of its singular locus, which corresponds to pairs of isotropic vectors in C^3.

In Section 2.3, we construct two Ulrich sheaves on the variety of 44 symmetric matrices of rank 2. One of the constructions we propose is new, according to the best of our knowledge. Both sheaves are GL(4)-equivariant, and they admit \Pieri resolutions" in the sense of [92]. We carefully analyze the resolutions using representation theory, and in particular show that their middle di erentials may be represented by symmetric matrices; see Propositions 2.16 and 2.19.

In Section 2.4, we combine the results of the previous sections and we construct the Chow form of the essential variety. The construction from [34] starts with our rank 2 Ulrich sheaves and allows to de ne two 20 20 matrices in the Plücker coordinates of $Gr(P^2; P^8)$ each of which drops rank exactly when the corresponding subspace P^2 meets the essential variety E_C. It requires some technical e ort to put these matrices in skew-symmetric form, and here our analysis from Section 2.3 pays o . We conclude this work with numerical experiments demonstrating the robustness to noise that our matrix formulas in Theorem 2.1 enjoy.

2.2 The essential variety is determinantal

Intrinsic description

Let $\mathcal{E} \subset \mathbb{R}^{3\times 3}$ be the essential variety, which is defined by the following conditions on the three singular values of a 3×3 matrix:

$$\mathcal{E} := \{M \in \mathbb{R}^{3\times 3} \mid \sigma_1(M) = \sigma_2(M), \; \sigma_3(M) = 0\}.$$

The polynomial equations of $\mathcal{E}$ are (see [35, Section 4]) as follows:

$$\mathcal{E} = \{M \in \mathbb{R}^{3\times 3} \mid \det(M) = 0, \; 2(MM^T)M - \operatorname{tr}\left(MM^T\right)M = 0\}. \qquad (2.2)$$

These 10 cubics minimally generate the *real radical ideal* [13, p. 85] of the essential variety $\mathcal{E}$, and that ideal is prime. Indeed, the real radical property follows from our Proposition 2.2(i) and [75, Theorem 12.6.1]. We denote by $\mathcal{E}_{\mathbb{C}}$ the projective variety in $\mathbb{P}^8_{\mathbb{C}}$ given by the complex solutions of (2.2). The essential variety $\mathcal{E}_{\mathbb{C}}$ has codimension 3 and degree 10 (see [77, Theorem 5.16]). In this section, we will prove that it is isomorphic to a hyperplane section of the variety $\mathbb{P}X^s_{4,2}$ of complex symmetric 4×4 matrices of rank ≤ 2. The first step towards this is Proposition 2.2 below, and that relies on the group symmetries of $\mathcal{E}_{\mathbb{C}}$, which we now explain.

Consider $\mathbb{R}^3$ with the standard inner product Q, and the corresponding action of $SO(3, \mathbb{R})$ on $\mathbb{R}^3$. Complexify $\mathbb{R}^3$ and consider $\mathbb{C}^3$ with the action of $SO(3, \mathbb{C})$, which has universal cover $SL(2, \mathbb{C})$. It is technically simpler to work with the action of $SL(2, \mathbb{C})$. Denoting by U the irreducible 2-dimensional representation of $SL(2, \mathbb{C})$, we have the equivariant isomorphism $\mathbb{C}^3 \cong S_2U$. Writing Q also for the complexification of the Euclidean product, the projective space $\mathbb{P}(S_2U)$ divides into two $SL(2, \mathbb{C})$-orbits, namely the isotropic quadric with equation $Q(u) = 0$ and its complement. Let V be another complex vector space of dimension 2. The essential variety $\mathcal{E}_{\mathbb{C}}$ is embedded into the projective space of 3×3-matrices $\mathbb{P}(S_2U \otimes S_2V)$. Since the singular value conditions defining $\mathcal{E}$ are $SO(3, \mathbb{R}) \times SO(3, \mathbb{R})$-invariant, it follows that $\mathcal{E}_{\mathbb{C}}$ is $SL(U) \times SL(V)$-invariant using [29, Theorem 2.2].

The following is a new geometric description of the essential variety. From the computer vision application, we start with the set of real points $\mathcal{E}$. However, below we see that the surface $\operatorname{Sing}(\mathcal{E}_{\mathbb{C}})$ inside $\mathcal{E}_{\mathbb{C}}$, which has no real points, 'determines' the algebraic geometry. Part (i) of Proposition 2.2 is proved also in [77, Proposition 5.9].

Proposition 2.2. (i) *The singular locus of* $\mathcal{E}_{\mathbb{C}}$ *is the projective surface given by:*

$$\operatorname{Sing}(\mathcal{E}_{\mathbb{C}}) = \left\{ab^T \in \mathbb{P}(\mathbb{C}^{3\times 3}) \mid Q(a) = Q(b) = 0\right\}.$$

(ii) *The second secant variety of* $\operatorname{Sing}(\mathcal{E}_{\mathbb{C}})$ *equals* $\mathcal{E}_{\mathbb{C}}$.

Proof. Denote by S the variety $\{ab^T \in \mathbb{P}(\mathbb{C}^{3\times 3}) \mid Q(a) = Q(b) = 0\}$, and let $\hat{S}$ be the affine cone over it. The line secant variety $\sigma_2(\hat{S})$ consists of elements of the form $M = a_1b_1^T + a_2b_2^T \in \mathbb{C}^{3\times 3}$ such that $Q(a_i) = a_i^T a_i = Q(b_i) = b_i^T b_i = 0$ for $i = 1, 2$. We compute that $MM^T = a_1b_1^T b_2a_2^T + a_2b_2^T b_1a_1^T$ so that $\mathrm{tr}(MM^T) = 2(b_1^T b_2)(a_1^T a_2)$. Moreover $MM^T M = a_1b_1^T b_2a_2^T a_1b_1^T + a_2b_2^T b_1a_1^T a_2b_2^T = (b_1^T b_2)(a_1^T a_2)M$. Hence the equations (2.2) of $\mathcal{E}_{\mathbb{C}}$ are satisfied by M. This proves that $\sigma_2(S) \subseteq \mathcal{E}_{\mathbb{C}}$. S is a surface and $\sigma_2(S)$ has dimension 5 (see [22, Theorem 1.3]). Since $\sigma_2(S)$ and $\mathcal{E}_{\mathbb{C}}$ are both of codimension 3 and $\mathcal{E}_{\mathbb{C}}$ is irreducible, the equality $\sigma_2(S) = \mathcal{E}_{\mathbb{C}}$ follows. It remains to prove (i). Denote by $[a_i]$ the line generated by a_i. Every element $a_1b_1^T + a_2b_2^T$ with $[a_1] \neq [a_2]$, $[b_1] \neq [b_2]$ and $Q(a_i) = Q(b_i) = 0$ for $i = 1, 2$ can be taken by $SL(U) \times SL(V)$ to a scalar multiple of any other element of the same form. This is the open orbit of the action of $SL(U) \times SL(V)$ on $\mathcal{E}_{\mathbb{C}}$. The remaining orbits are the following:

1. the surface S, with set-theoretic equations $MM^T = M^T M = 0$.
2. $T_1 \setminus S$, where $T_1 = \{ab^T \in \mathbb{P}(\mathbb{C}^{3\times 3}) \mid Q(a) = 0\}$ is a threefold, with set-theoretic equations $M^T M = 0$.
3. $T_2 \setminus S$, where $T_2 = \{ab^T \in \mathbb{P}(\mathbb{C}^{3\times 3}) \mid Q(b) = 0\}$ is a threefold, with set-theoretic equations $MM^T = 0$.
4. $\mathrm{Tan}(S) \setminus (T_1 \cup T_2)$, where the *tangential variety* $\mathrm{Tan}(S)$ is the fourfold union of all tangent spaces to S, with set-theoretic equations $\mathrm{tr}(MM^T) = 0$, $MM^T M = 0$.

It is easy to check they are orbits, in a similar way than in [43, Example 14.4.5].

One can compute explicitly that the Jacobian matrix of $\mathcal{E}_{\mathbb{C}}$ at $\begin{pmatrix} 1 & 0 & 0 \\ \sqrt{-1} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \in T_1 \setminus S$ has rank 3. The following code in *Macaulay2* ([44]) does that computation:

```
R = QQ[m_(1,1)..m_(3,3)]
M = transpose(genericMatrix(R,3,3))
I = ideal(det(M))+minors(1,2*M*transpose(M)*M - trace(M*transpose(M))*M)
Jac = transpose jacobian I
S = QQ[q]/(1+q^2)
specializedJac = (map(S,R,{1,0,0,q,0,0,0,0,0}))(Jac)
minors(3,specializedJac)
```

Hence the points in $T_1 \setminus S$ are smooth points of $\mathcal{E}_{\mathbb{C}}$. By symmetry, also the points in $T_2 \setminus S$ are smooth. By semicontinuity, the points in $\mathrm{Tan}(S) \setminus (T_1 \cup T_2)$ are smooth.

Since points in S are singular for the secant variety $\sigma_2(S)$, this finishes the proof of (i). □

Remark 2.3. From Proposition 2.2, the essential variety is isomorphic to the variety of $2 \times 2 \times 2 \times 2$ tensors of rank ≤ 2 invariant under the permutations $S_2 \times S_2 \times S_4$. Hence, by the study of tensor decomposition, the parametric description in Proposition 2.2 is identifiable, meaning that, from the matrix $a_1 b_1^T + a_2 b_2^T$, all a_i, b_i are determined up to scalar multiple. That shows that real essential matrices have the form $a^T b + \bar{a}^T \bar{b}$ with $a, b \in \mathbb{C}^3$ and $Q(a) = Q(b) = 0$. This may be written in the alternative form $(u^2)^T v^2 + (\bar{u}^2)^T \bar{v}^2 \in S_2(U) \otimes S_2(V)$ with $u \in U$, $v \in V$. This may help in computing real essential matrices. Note that the four non-open orbits listed in the proof of Proposition 2.2 are contained in the isotropic quadric $\mathrm{tr}(MM^T) = 0$, hence they have no real points.

Remark 2.4. The surface $\mathrm{Sing}(E_\mathbb{C})$ is more familiar with the embedding by $\mathcal{O}(1,1)$, when it is the smooth quadric surface, doubly ruled by lines. In the embedding by $\mathcal{O}(2,2)$, the two rulings are given by conics. These observations suggest expressing $E_\mathbb{C}$ as a determinantal variety, as we do next in Proposition 2.5. Indeed, note that the smooth quadric surface embedded by $\mathcal{O}(2,2)$ is isomorphic to a linear section of the second Veronese embedding of $\mathbb{P}^3$, which is the variety of 4×4 symmetric matrices of rank 1.

In the following note that $S_2(U \otimes V)$ is 10-dimensional and identifies as the space of symmetric 4×4-matrices.

Proposition 2.5. The essential variety $E_\mathbb{C}$ is isomorphic to a hyperplane section of the variety of rank ≤ 2 elements in $\mathbb{P}(S_2(U \otimes V))$. Concretely, this latter variety identifies as the projective variety of 4×4 symmetric matrices of rank ≤ 2 (see also Subsection 2.3), and the section consists of traceless 4×4 symmetric matrices of rank ≤ 2.

Proof. The embedding of $\mathbb{P}(U) \times \mathbb{P}(V)$ in $\mathbb{P}(S_2(U) \otimes S_2(V))$ is given by $(u, v) \mapsto u^2 \otimes v^2$. Recall that Cauchy's formula states $S_2(U \otimes V) = (S_2(U) \otimes S_2(V)) \oplus (\wedge^2 U \otimes \wedge^2 V)$, where $\dim(U \otimes V) = 4$. Hence, $\mathbb{P}(S_2(U) \otimes S_2(V))$ is equivariantly embedded as a codimension one subspace in $\mathbb{P}(S_2(U \otimes V))$. The image is the subspace of traceless elements (since that is dimension 8 and invariant), and this map sends $u^2 \otimes v^2 \mapsto (u \otimes v)^2$. By Proposition 2.2, we have shown that $\mathrm{Sing}(E_\mathbb{C})$ embeds into a hyperplane section of the variety of rank 1 elements in $\mathbb{P}(S_2(U \otimes V))$. So, $E_\mathbb{C} = \sigma_2(\mathrm{Sing}(E_\mathbb{C}))$ embeds into that hyperplane section of the variety of rank ≤ 2 elements. This last variety has degree 10 by Segre formula [47, Proposition 12 (b)]. Comparing dimensions and degrees, the result follows. □

Remark 2.6. In light of the description in Proposition 2.5, it follows by Example 3.2 and Corollary 6.4 of [28] that the Euclidean distance degree is $\mathrm{EDdegree}(\mathcal{E}_{\mathbb{C}}) = 6$. This result has been proved also in [30], where the computation of EDdegree was performed in the more general setting of orthogonally invariant varieties. This quantity measures the algebraic complexity of finding the nearest point on $\mathcal{E}$ to a given noisy data point in $\mathbb{R}^{3\times 3}$.

Coordinate description

We now make the determinantal description of $\mathcal{E}_{\mathbb{C}}$ in Proposition 2.5 explicit in coordinates. For this, denote $a = (a_1; a_2; a_3)^T \in \mathbb{C}^3$. We have $Q(a) = a_1^2 + a_2^2 + a_3^2$. The $SL(2;\mathbb{C})$-orbit $Q(a) = 0$ is parametrized by $\left(u_1^2 - u_2^2; 2u_1u_2; \sqrt{-1}(u_1^2 + u_2^2)\right)^T$ where $(u_1; u_2)^T \in \mathbb{C}^2$. Let:

$$M = \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{pmatrix} \in \mathbb{C}^{3\times 3};$$

and define the 4×4 traceless symmetric matrix $s(M)$ (depending linearly on M):

$$s(M) := \frac{1}{2}\begin{pmatrix} m_{11} - m_{22} - m_{33} & m_{13} + m_{31} & m_{12} + m_{21} & m_{23} - m_{32} \\ m_{13} + m_{31} & -m_{11} - m_{22} + m_{33} & m_{23} + m_{32} & m_{12} - m_{21} \\ m_{12} + m_{21} & m_{23} + m_{32} & -m_{11} + m_{22} - m_{33} & -m_{13} + m_{31} \\ m_{23} - m_{32} & m_{12} - m_{21} & -m_{13} + m_{31} & m_{11} + m_{22} + m_{33} \end{pmatrix}. \tag{2.3}$$

This construction furnishes a new view on the essential variety $\mathcal{E}$, as described in Proposition 2.7.

Proposition 2.7. The linear map s in (2.3) is a real isometry from the space of 3×3 real matrices to the the space of traceless symmetric 4×4 real matrices. We have that:

$$M \in \mathcal{E} \iff \mathrm{rk}(s(M)) \leq 2.$$

The complexification of s, denoted again by s, satisfies for any $M \in \mathbb{C}^{3\times 3}$:

$$M \in \mathrm{Sing}(\mathcal{E}_{\mathbb{C}}) \iff \mathrm{rk}(s(M)) \leq 1;$$

$$M \in \mathcal{E}_{\mathbb{C}} \iff \mathrm{rk}(s(M)) \leq 2.$$

Proof. We construct the correspondence over $\mathbb{C}$ at the level of $\mathrm{Sing}(\mathcal{E}_\mathbb{C})$ and then we extend it by linearity. Choose coordinates (u_1, u_2) in U and coordinates (v_1, v_2) in V. Consider the following parametrization of matrices $M \in \mathrm{Sing}(\mathcal{E}_\mathbb{C})$:

$$M = \begin{pmatrix} u_1^2 - u_2^2 \\ 2u_1u_2 \\ \sqrt{-1}(u_1^2+u_2^2) \end{pmatrix} \otimes \left(v_1^2 - v_2^2,\ 2v_1v_2,\ \sqrt{-1}(v_1^2+v_2^2) \right). \tag{2.4}$$

Consider also the following parametrization of the Euclidean quadric in $U \otimes V$:

$$k = \left(u_2v_2 - u_1v_1,\ \sqrt{-1}(u_1v_1+u_2v_2),\ (u_1v_2+u_2v_1),\ \sqrt{-1}(u_1v_2 - u_2v_1) \right).$$

The variety of rank 1 traceless 4×4 symmetric matrices is accordingly parametrized by $k^T k$. Substituting (2.4) into the right-hand side below, a computation verifies that:

$$k^T k = s(M).$$

This proves the second equivalence in the statement above and explains the definition of $s(M)$, namely that it is the equivariant embedding from Proposition 2.5 in coordinates. The third equivalence follows because $\mathcal{E}_\mathbb{C} = \sigma_2(\mathrm{Sing}(\mathcal{E}_\mathbb{C}))$, by Proposition 2.2(ii). For the first equivalence, we note that s is defined over $\mathbb{R}$ and now a direct computation verifies that $\mathrm{tr}\left(s(M)s(M)^T\right) = \mathrm{tr}\left(MM^T\right)$ for $M \in \mathbb{R}^{3\times 3}$. □

Note that the ideal of 3-minors of $s(M)$ is indeed generated by the ten cubics in (2.2).

Remark 2.8. The critical points of the distance function from any data point $M \in \mathbb{R}^{3\times 3}$ to $\mathcal{E}$ can be computed by means of the SVD of $s(M)$, as in [28, Example 2.3].

2.3 Ulrich sheaves on the variety of symmetric 4×4 matrices of rank ≤ 2

Our goal is to construct the Chow form of the essential variety. By the theory of [34], this can be done provided one has an Ulrich sheaf on this variety. The notions of Ulrich sheaf, Chow forms and the construction of [34] will be explained below.

As shown in Section 2.2, the essential variety $\mathcal{E}_\mathbb{C}$ is a linear section of the projective variety of symmetric 4×4 matrices of rank ≤ 2, which we denote as $\mathbb{P}X^s_{4,2}$. If we construct an Ulrich sheaf on $\mathbb{P}X^s_{4,2}$, then a quotient of this sheaf by a linear form is an Ulrich sheaf on $\mathcal{E}_\mathbb{C}$ provided that linear form is regular for the Ulrich sheaf on $\mathbb{P}X^s_{4,2}$. We will achieve this twice, in Section 2.3 and Section 2.3.

Definition of Ulrich modules and sheaves

Definition 2.9. A graded module M over a polynomial ring $A = \mathbb{C}[x_0, \dots, x_n]$ is an Ulrich module provided:

1. It is generated in degree 0 and has a linear minimal free resolution:

$$0 \leftarrow M \leftarrow A^{\beta_0} \leftarrow A(-1)^{\beta_1} \leftarrow A(-2)^{\beta_2} \xleftarrow{d_2} \cdots \leftarrow A(-c)^{\beta_c} \leftarrow 0. \quad (2.5)$$

2. The length of the resolution c equals the codimension of the support of the module M.

2'. The Betti numbers are $\beta_i = \binom{c}{i}\beta_0$ for $i = 0, \dots, c$.

One can use either (1) and (2), or equivalently, (1) and (2)' as the definition.

A sheaf $\mathcal{F}$ on a projective space $\mathbb{P}^n$ with support of dimension ≥ 1 is an Ulrich sheaf provided it is the sheafication of an Ulrich module. Equivalently, the module of twisted global sections $M = \bigoplus_{d \in \mathbb{Z}} H^0(\mathbb{P}^n, \mathcal{F}(d))$ is an Ulrich module over the polynomial ring A.

Fact 2.10. If the support of an Ulrich sheaf $\mathcal{F}$ is a variety X of degree d, then β_0 is a multiple of d, say rd. This corresponds to $\mathcal{F}$ being a sheaf of rank r on X.

Since there is a one-to-one correspondence between Ulrich modules over A and Ulrich sheaves on $\mathbb{P}^n$, we interchangably speak of both. But in our constructions we focus on Ulrich modules. A prominent conjecture of [34, p.543] states that on any variety X in a projective space, there is an Ulrich sheaf whose support is X.

The variety of symmetric 4 × 4 matrices

We fix notation. Let X_4^s be the space of symmetric 4×4 matrices over the field $\mathbb{C}$. This identifies as $\mathbb{C}^{10}$. Let $x_{ij} = x_{ji}$ be the coordinate functions on X_4^s where $1 \leq i \leq j \leq 4$, so the coordinate ring of X_4^s is:

$$A = \mathbb{C}[x_{ij}]_{1 \leq i \leq j \leq 4}.$$

For $0 \leq r \leq 4$, denote by $X_{4,r}^s$ the affine subvariety of X_4^s consisting of matrices of rank $\leq r$. The ideal of $X_{4,r}^s$ is generated by the $(r+1) \times (r+1)$-minors of the generic 4×4 symmetric matrix (x_{ij}). This is in fact a prime ideal, by [104, Theorem

6.3.1]. The rank subvarieties have the following degrees and codimensions by [47, Proposition 12 (b)]:

variety	degree	codimension
$X^s_{4;4}$	1	0
$X^s_{4;3}$	4	1
$X^s_{4;2}$	10	3
$X^s_{4;1}$	8	6
$X^s_{4;0}$	1	10

Since the varieties $X^s_{4;r}$ are defined by homogeneous ideals, they give rise to projective varieties $PX^s_{4;r}$ in the projective space P^9. However, in Section 2.3 and Section 2.3 it will be convenient to work with affine varieties, and general (instead of special) linear group actions.

The group $GL(4;C)$ acts on X^s_4. If $M \in GL(4;C)$ and $X \in X^s_4$, the action is as follows:

$$M X = M \cdot X \cdot M^T:$$

Since any complex symmetric matrix can be diagonalized by a coordinate change, there are five orbits of the action of $GL(4;C)$ on X^s_4, one per rank of the symmetric matrix. Let:

$$E = C^4$$

be a four-dimensional complex vector space. The coordinate ring of X^s_4 identifies as $A = \mathrm{Sym}(S_2(E))$. The space of symmetric matrices X^s_4 may then be identified with the dual space $S_2(E)^*$, so again we see that $GL(E) = GL(4;C)$ acts on $S_2(E)^*$.

Representations and Pieri's rule

We shall recall some basic representation theory of the general linear group $GL(W)$, where W is a n-dimensional complex vector space. The irreducible representations of $GL(W)$ are given by Schur modules $S_\lambda(W)$ where λ is a generalized partition: a sequence of integers $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$. When $\lambda = d;0;\dots;0$, then $S_\lambda(W)$ is the d^{th} symmetric power $S_d(W)$. When $\lambda = 1;\dots;1;0;\dots;0$, with d 1's, then $S_\lambda(W)$ is the exterior wedge $\wedge^d W$. For all partitions λ there are isomorphisms of $GL(W)$-representations:

$$S_\lambda(W)^* \cong S_{-\lambda_n;\dots;-\lambda_1}(W) \quad \text{and} \quad S_\lambda(W) \otimes (\wedge^n W)^{\otimes r} \cong S_{\lambda+r\mathbf{1}}(W)$$

where $\mathbf{1} = 1;1;\dots;1$. Here $\wedge^n W$ is the one-dimensional representation C of $GL(W)$ where a linear map ϕ acts by its determinant.

Denote by $| \ |$:= $_1$ + + $_n$. Assume $_n$; $_n$ 0. The tensor product of two Schur modules S (W) S (W) splits into irreducibles as a direct sum of Schur modules:

$$\bigoplus u(; ;) S (W)$$

where the sum is over partitions with $| \ | = | \ | + | \ |$. The multiplicities u(; ;) 2 Z_0 are determined by the Littlewood-Richardson rule [41, Appendix A]. In one case, that will be important to us below, there is a particularly nice form of this rule. Given two partitions ^{0} and , we say that ^{0}= is a horizontal strip if ^0_i $_i$ $^0_{i+1}$.

Fact 2.11 (Pieri's rule). As GL(W)-representations, we have the rule:

$$S (W) \quad S_d(W) \quad = \bigoplus_{\substack{| ^0| = | \ | + d \\ ^0 = \text{ is a horizontal strip}}} S_{^0}(W):$$

The rst Ulrich sheaf

We are now ready to describe our rst Ulrich sheaf on the projective variety $P X^s_{4;2}$. We construct it as an Ulrich module supported on the variety $X^s_{4;2}$. We use notation from Section 2.3, so E is 4-dimensional. Consider $S_3(E)$ $S_2(E)$. By Pieri's rule this decomposes as:

$$S_5(E) \quad S_{4;1}(E) \quad S_{3;2}(E):$$

We therefore get a GL(E)-inclusion $S_{3;2}(E)$! $S_3(E)$ $S_2(E)$ unique up to nonzero scale. Since $A_1 = S_2(E)$ from Section 2.3, this extends uniquely to an A-module map:

$$S_3(E) \quad A \qquad S_{3;2}(E) \quad A(\ 1):$$

This map can easily be programmed using Macaulay2 and the package PieriMaps [90]:

```
R=QQ[a..d]
needsPackage "PieriMaps"
f=pieri({3,2},{2,2},R)
S=QQ[a..d,y_0..y_9]
a2=symmetricPower(2,matrix{{a..d}})
alpha=sum(10,i->contract(a2_(0,i),sub(f,S))*y_i)
```

We can then compute the resolution of the cokernel of in Macaulay2. It has the form:

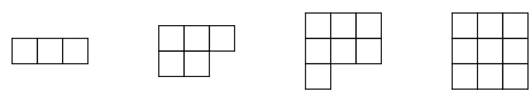
$$A^{20} \qquad A(\ 1)^{60} \qquad A(\ 2)^{60} \qquad A(\ 3)^{20}:$$

Thus the cokernel of is an Ulrich module by (1) and (2)' in De nition 2.9. An important point is that the res command in Macaulay2 computes di erential matrices in unenlightening bases. We completely and intrinsically describe the GL(E)-resolution below:

Proposition 2.12. The cokernel of is an Ulrich module M of rank 2 supported on the variety $X^{s}_{4;2}$. The resolution of M is $GL(E)$-equivariant and it is:

$$F : S_3(E) \quad A \qquad S_{3;2}(E) \quad A(\quad 1) \qquad S_{3;3;1}(E) \quad A(\quad 2) \qquad (2.6)$$
$$S_{3;3;3}(E) \quad A(\quad 3)$$

with ranks $20; 60; 60; 20$, and where all di erential maps are induced by Pieri's rule. The dual complex of this resolution is also a resolution, and these two resolutions are isomorphic up to twist. As in [92], we can visualize the resolution by:

$0 \qquad M$  $0:$

Proof. Since M is the cokernel of a $GL(E)$-map, it is $GL(E)$-equivariant. So, the support of M is a union of orbits. By De nition 2.9(2), M is supported in codimension 3. Since the only orbit of codimension 3 is $X^{s}_{4;2} n X^{s}_{4;3}$, the support of M is the closure of this orbit, which is $X^{s}_{4;2}$. It can also easily be checked with Macaulay2, by restricting to diagonal matrices of rank r for $r = 0; : : : ; 4$, that M is supported on the strata $X^{s}_{4;r}$ where $r \quad 2$. Also, the statement that the rank of M equals 2 is now immediate from Fact 2.10.

Now we prove that the $GL(E)$-equivariant minimal free resolution of M is F as above. By Pieri's rule there is a $GL(E)$-map unique up to nonzero scalar:

$$S_{3;2}(E) \quad S_2(E) \qquad S_{3;3;1}(E)$$

and a $GL(E)$-map unique up to nonzero scalar:

$$S_{3;3;1}(E) \quad S_2(E) \qquad S_{3;3;3}(E):$$

These are the maps and in F respectively. The composition maps $S_{3;3;1}(E)$ to a submodule of $S_3(E) \quad S_2(S_2(E))$. By [104, Proposition 2.3.8] the latter double symmetric power equals $S_4(E) \quad S_{2;2}(E)$, and so this tensor product decomposes as:

$$S_3(E) \quad S_4(E) \quad \overset{M}{} \quad S_3(E) \quad S_{2;2}(E):$$

By Pieri's rule, none of these summands contains $S_{3;3;1}(E)$. Hence is zero by Schur's lemma. The same type of argument shows that is zero. Thus F is a complex.

By our Macaulay2 computation of Betti numbers before the Proposition, ker() is generated in degree 2 by 60 minimal generators. In F these must be the image of $S_{3;3;1}(E)$, since that is 60-dimensional by the hook content formula and it maps injectively to F_1. So F is exact at F_1. Now again by the Macaulay2 computation, it follows that ker is generated in degree 3 by 20 generators. These must be the image of $S_{3;3;3}(E)$ since that is 20-dimensional and maps injectively to F_2. So F is exact at F_2. Finally, the computation implies that is injective, and F is the $GL(E)$-equivariant minimal free resolution of M.

For the statement about the dual, recall that since F is a resolution of a Cohen-Macaulay module, the dual complex, obtained by applying $Hom_A(; !_A)$ with $!_A = A(10)$, is also a resolution. If we twist this dual resolution with $(\wedge^4 E)^{3} \quad A(7)$, the terms will be as in the original resolution. Since the nonzero $GL(E)$-map is uniquely determined up to scale, it follows that F and its dual are isomorphic up to twist. □

Remark 2.13. The $GL(E)$-representations in this resolution could also have been computed using the Macaulay2 package HighestWeights [42].

Remark 2.14. The dual of this resolution is:

$$S_{3;3;3}(E) \quad A \quad S_{3;3;1}(E) \quad A(1) \quad S_{3;2}(E) \quad A(2) \quad S_3(E) \quad A(3): \tag{2.7}$$

A symmetric form q in $S_2(E)$ corresponds to a point in Spec(A) and a homomorphism $A\ !\ \mathbb{C}$. The ber of this complex over the point q is then an $SO(E; q)$-complex:

$$S_{3;3;3}(E) \quad S_{3;3;1}(E) \quad S_{3;2}(E) \quad S_3(E): \tag{2.8}$$

When q is a nondegenerate form, this is the Littlewood complex $L^{3;3;3}$ as de ned in [91, Section 4.2]. (The terms of $L^{3;3;3}$ can be computed using the plethysm in Section 4.6 of loc.cit.) This partition $= (3; 3; 3)$ is not admissible since $3 + 3 > 4$, see [91, Section 4.1]. The cohomology of (2.8) is then given by [91, Theorem 4.4] and it vanishes (since here $i_4( ) = 1$), as it should in agreement with Proposition 2.12. The dual resolution (2.7) of the Ulrich sheaf can then be thought of as a \universal" Littlewood complex for the parition $= (3; 3; 3)$. In other cases when Littlewood complexes are exact, it would be an interesting future research topic to investigate the sheaf that is resolved by the \universal Littlewood complex".

To obtain nicer formulas for the Chow form of the essential variety E_C in Section 2.4, we now prove that the middle map in the resolution (2.6) is symmetric, in the following appropriate sense. In general, suppose that we are given a linear map $W \to W \; L$ where L is a nite dimensional vector space. Dualizing, we get a map $W^{T} \; W \; L$ which in turn gives a map $W \; L \; W$. By de nition, the map is symmetric if = and skew-symmetric if = . If is symmetric and is represented as a matrix with entries in L with respect to dual bases of W and W, then that matrix is symmetric, and analogously when is skew-symmetric. Note that the map also induces a map $L \to W \; W$.

Fact 2.15. The map is symmetric if the image of is in the subspace $S_2(W)$ $W \; W$ and it is skew-symmetric if the image is in the subspace $\wedge^2 W \; W \; W$.

Proposition 2.16. The middle map in the resolution (2.6) is symmetric.

Proof. Consider the map in degree 3. It is:

$$S_{3;2}(E) \quad S_2(E) \qquad S_{3;3;1}(E) = S_{3;2}(E) \quad (\wedge^4 E)^{3}$$

and it induces the map:

$$S_{3;2}(E) \quad S_{3;2}(E) \qquad S_2(E) \quad (\wedge^4 E)^{3} = S_{3;3;3;1}(E):$$

By the Littlewood-Richardson rule, the right representation above occurs with multiplicity 1 in the left side. Now one can check that $S_{3;3;3;1}(E)$ occurs in $S_2(S_{3;2}(E))$. This follows by Corollary 5.5 in [19] or one can use the package SchurRings [95] in Macaulay2:

```
needsPackage "SchurRings"
S = schurRing(s,4,GroupActing=>"GL")
plethysm(s_2,s_{3,2})
```

Due to Fact 2.15, we can conclude that the map is symmetric. □

The second Ulrich sheaf

We construct another Ulrich sheaf on $P X^{s}_{4;2}$ and analyze it similarly to as above. This will lead to a second formula for Ch(E_C) in Section 2.4. Consider $S_{2;2;1}(E) \quad S_2(E)$. By Pieri's rule:

$$S_{2;2;1}(E) \quad S_2(E) = S_{4;2;1}(E) \quad S_{3;2;2}(E) \quad S_{3;2;1;1}(E) \quad S_{2;2;2;1}(E):$$

Thus there is a GL(E)-map, with nonzero degree 1 components unique up to scale:

$$S_{2;2;1}(E) \quad A \qquad (S_{3;2;2}(E) \quad S_{3;2;1;1}(E) \quad S_{2;2;2;1}(E)) \quad A(\ \ 1):$$

This map can be programmed in Macaulay2 using PieriMaps as follows:

```
R=QQ[a..d]
needsPackage "PieriMaps"
f1= transpose pieri({3,2,2,0},{1,3},R)
f2=transpose pieri({3,2,1,1},{1,4},R)
f3=transpose pieri({2,2,2,1},{3,4},R)
f = transpose (f1||f2||f3)
S=QQ[a..d,y_0..y_9]
a2=symmetricPower(2,matrix{{a..d}})
alpha=sum(10,i->contract(a2_(0,i),sub(f,S))*y_i)
```

We can then compute the resolution of coker() in Macaulay2. It has the form:

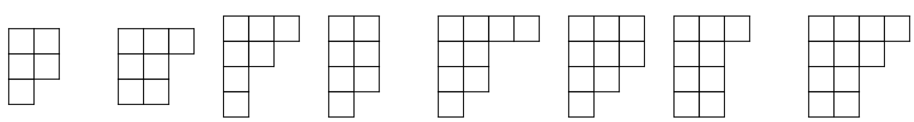
$$A^{20} \qquad A(\ \ 1)^{60} \qquad A(\ \ 2)^{60} \qquad A(\ \ 3)^{20}:$$

Thus the cokernel of is an Ulrich module, and moreover we have:

Proposition 2.17. *The cokernel of is an Ulrich module M of rank 2 supported on the variety $X^{s}_{4;2}$. The resolution of M is GL(E)-equivariant and it is:*

$$\begin{aligned} F \ : S_{2;2;1}(E) \quad A \qquad & (S_{3;2;2}(E) \quad S_{3;2;1;1}(E) \quad S_{2;2;2;1}(E)) \quad A(\ \ 1) \\ & (S_{4;2;2;1}(E) \quad S_{3;3;2;1}(E) \quad S_{3;2;2;2}(E)) \quad A(\ \ 2) \qquad (2.9) \\ & S_{4;3;2;2}(E) \quad A(\ \ 3) \end{aligned}$$

with ranks $20;60;60;20$. The dual complex of this resolution is also a resolution and these two resolutions are isomorphic up to twist. We can visualize the resolution by:

$0 \quad M$  $0:$

Proof. The argument concerning the support of M is exactly as in Proposition 2.12.

Now we prove that the minimal free resolution of M is of the form above, differently than in Proposition 2.12. To start, note that the module $S_{4;2;2;1}(E)$ occurs by Pieri once in each of:

$$S_{3;2;2}(E) \quad S_2(E); \quad S_{3;2;1;1}(E) \quad S_2(E); \quad S_{2;2;2;1}(E) \quad S_2(E):$$

On the other hand, it occurs in:

$$S_{2;2;1}(E) \quad S_2(S_2(E)) = S_{2;2;1}(E) \quad S_4(E) \quad S_{2;2;1}(E) \quad S_{2;2}(E)$$

only twice, as seen using Pieri's rule and the Littlewood-Richardson rule. Thus $S_{4;2;2;1}(E)$ occurs at least once in the degree 2 part of ker( ). Similarly we see that each of $S_{3;3;2;1}(E)$ and $S_{3;2;2;2}(E)$ occurs at least once in ker( ) in degree 2. But by the Macaulay2 computation before this Proposition, we know that ker( ) is a module with 60 generators in degree 2. And the sum of the dimensions of these three representations is 60. Hence each of them occurs exactly once in ker( ) in degree 2, and they generate ker( ).

Now let C be the 20-dimensional vector space generating ker( ). Since the resolution of M has length equal to codim(M), the module M is Cohen-Macaulay and the dual of its resolution, obtained by applying $Hom_A(\; ; !_A)$ where $!_A = A(\; 4)$, is again a resolution of $Ext^3_A(M; !_A)$. Thus the map from C $\quad$ A($\;$ 3) to each of:

$$S_{4;2;2;1}(E) \quad A(\; 2); \quad S_{3;3;2;1}(E) \quad A(\; 2); \quad S_{3;2;2;2}(E) \quad A(\; 2)$$

is nonzero. In particular C maps nontrivially to:

$$S_{3;2;2;2}(E) \quad S_2(E) = S_{5;2;2;2}(E) \quad S_{4;3;2;2}(E):$$

Each of the right-hand side representations have dimension 20, so one of them equals C. However only the last one occurs in $S_{3;3;2;1}(E) \quad S_2(E)$, and so $C = S_{4;3;2;2}(E)$. We have proven that the GL(E)-equivariant minimal free resolution of M indeed has the form F .

For the statement about the dual, recall that each of the three components of in degree 1 are nonzero. Also, as the dual complex is a resolution, here obtained by applying $Hom_A(\; ; !_A)$ with $!_A = A(\; 10)$, all three degree 1 components of are nonzero. If we twist this dual resolution with $(\wedge^4 E)^{4} \quad A(7)$, the terms will be as in the original resolution. Because each of the three nonzero components of the map are uniquely determined up to scale, the resolution F and its dual are isomorphic up to twist. □

Remark 2.18. Again the GL(E)-representations in this resolution could have been computed using the Macaulay2 package HighestWeights.

Proposition 2.19. The middle map in the resolution (2.9) is symmetric.

Proof. We rst show that the three `diagonal' components of in (2.9) are symmetric:

$$S_{3;2;2}(E) \quad S_2(E) \quad {}^{1} \quad S_{4;2;2;1}(E)$$
$$S_{3;2;1;1}(E) \quad S_2(E) \quad {}^{2} \quad S_{3;3;2;1}(E)$$
$$S_{2;2;2;1}(E) \quad S_2(E) \quad {}^{3} \quad S_{3;2;2;2}(E):$$

Twisting the third component $_3$ with $(\wedge^4 E\ )^{\ 2}$, it identi es as:

$$E \quad S_2(E) \quad E$$

and so $_3$ is obviously symmetric. Twisting the second map $_2$ with $\wedge^4 E$ it identi es as:

$$S_{2;1}(E) \quad S_2(E) \quad S_{2;2;1}(E) = (\ S_{2;1}(E)\ ) \quad (\wedge^4 E)^{\ 2};$$

which induces the map:

$$S_{2;1}(E) \quad S_{2;1}(E) \quad S_2(E) \quad (\wedge^4 E)^{\ 2} = S_{2;2;2}(E):$$

By the Littlewood-Richardson rule, the left tensor product contains $S_{2;2;2}(E)$ with multiplicity 1. By Corollary 5.5 in [19] or SchurRings in Macaulay2, this is in $S_2(S_{2;1}(E))$:

```
needsPackage "SchurRings"
S = schurRing(s,4,GroupActing=>"GL")
plethysm(s_2,s_{2,1})
```

So by Fact 2.15, the component $_2$ is symmetric. The rst map $_1$ may be identi ed as:

$$S_{3;2;2}(E) \quad S_2(E) \quad (S_{3;2;2}(E)) \quad (\wedge^4 E)^{\ 4};$$

which induces the map:

$$S_{3;2;2}(E) \quad S_{3;2;2}(E) \quad S_2(E) \quad (\wedge^4 E)^{\ 4} = S_{4;4;4;2}(E):$$

Again by Littlewood-Richardson, $S_{4;4;4;2}(E)$ is contained with multiplicity 1 in the left side. By Corollary 5.5 in [19] or the package SchurRings in Macaulay2, this is in $S_2(S_{3;2;2}(E))$:

```
needsPackage "SchurRings"
S = schurRing(s,4,GroupActing=>"GL")
plethysm(s_2,s_{3,2,2})
```

It is now convenient to tensor the resolution (2.9) by $(\wedge^4 E)^{-2}$, and to let:

$$T_1 = S_{1;0;0;-2}(E); \quad T_2 = S_{1;0;-1;-1}(E); \quad T_3 = S_{0;0;0;-1}(E).$$

We can then write the middle map as:

$$T_1 \otimes A(-1) \oplus T_2 \otimes A(-1) \oplus T_3 \otimes A(-1) \xrightarrow{\begin{pmatrix} \phi_1 & \alpha_2 & \beta_2 \\ \alpha_1 & \phi_2 & 0 \\ \beta_1 & 0 & \phi_3 \end{pmatrix}} T_1 \otimes A(-2) \oplus T_2 \otimes A(-2) \oplus T_3 \otimes A(-2) \tag{2.10}$$

Note indeed that the component:

$$S_{1;0;-1;-1}(E) \otimes S_2(E) = T_2 \otimes S_2(E) \to T_3 = S_1(E)$$

must be zero, since the left tensor product does not contain $S_1(E)$ by Pieri's rule. Similarly the map $T_3 \otimes S_2(E) \to T_2$ is zero.

We know the maps ϕ_1, ϕ_2 and ϕ_3 are symmetric. Consider:

$$T_2 \otimes A(-1) \xrightarrow{\alpha_1} T_1 \otimes A(-2); \quad T_1 \otimes A(-1) \xrightarrow{\alpha_2} T_2 \otimes A(-2).$$

Since the resolution (2.9) is isomorphic to its dual, either both α_1 and α_2 are nonzero, or they are both zero. Suppose both are nonzero. The dual of α_2 is (up to twist) $T_2 \otimes A(-1) \xrightarrow{\alpha_2^T} T_1 \otimes A(-2)$. But such a $GL(E)$-map is unique up to scalar, as is easily seen by Pieri's rule. Thus whatever the case we can say that $\alpha_1 = c\,\alpha_2^T$ for some nonzero scalar c. Similarly we get $\beta_1 = c\,\beta_2^T$. Composing the map (2.10) with the automorphism on its right given by the block matrix:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & c & 0 \\ 0 & 0 & c \end{pmatrix};$$

we get a middle map:

$$T_1 \otimes A(-1) \oplus T_2 \otimes A(-1) \oplus T_3 \otimes A(-1) \xrightarrow{\begin{pmatrix} \phi_1 & \alpha'_2 & \beta'_2 \\ \alpha_1 & \phi'_2 & 0 \\ \beta_1 & 0 & \phi'_3 \end{pmatrix}} T_1 \otimes A(-2) \oplus T_2 \otimes A(-2) \oplus T_3 \otimes A(-2)$$

where the diagonal maps are still symmetric, and $\alpha_1 = (\alpha'_2)^T$ and $\beta_1 = (\beta'_2)^T$. So we get a symmetric map, and the result about follows. □

This second Ulrich module constructed above in Proposition 2.17 is a particular instance of a general construction of Ulrich modules on the variety of symmetric $n \times n$ matrices of rank $\leq r$; see [104], Section 6.3 and Exercise 34 in Section 6. We briefly recall the general construction. Let $W = \mathbb{C}^n$ and G be the Grassmannian $Gr(n-r; W)$ of $(n-r)$-dimensional subspaces of W. There is a tautological exact sequence of algebraic vector bundles on G:

$$0 \to K \to W \otimes \mathcal{O}_G \to Q \to 0,$$

where r is the rank of Q. Let $X = X_n^s$ be the affine space of symmetric $n \times n$ matrices, and define Z to be the incidence subvariety of $X \times G$ given by:

$$Z = \{((W \xrightarrow{\phi} W), (\mathbb{C}^{n-r} \xrightarrow{i} W)) \in X \times G \mid \phi \circ i = 0\}.$$

The variety Z is the affine geometric bundle $V_G(S_2(Q))$ of the locally free sheaf $S_2(Q)$ on the Grassmannian G. There is a commutative diagram:

$$\begin{array}{ccc} Z & \longrightarrow & X \times G \\ \big\downarrow & & \big\downarrow \\ X_{n,r}^s & \longrightarrow & X \end{array}$$

in which Z is a desingularization of $X_{n,r}^s$. For any locally free sheaf E, the Schur functor S_λ applies to give a new locally free sheaf $S_\lambda(E)$. Consider then the locally free sheaf:

$$E(n;r) = S_{(n-r)^r}(Q) \otimes S_{n-r-1,n-r-2,\ldots,1,0}(K)$$

on the Grassmannian $Gr(n-r; W)$. Note that $S_{(n-r)^r}(Q) = (\det(Q))^{n-r}$ is a line bundle and $E(n;r)$ is a locally free sheaf of rank $2^{\binom{n-r}{2}}$. Let $Z \xrightarrow{p} G$ be the projection map. By pullback we get the locally free sheaf $p^*(E(n;r))$ on Z. The pushforward of this locally free sheaf down to $X_{n,r}^s$ is an Ulrich sheaf on this variety. Since $X_{n,r}^s$ is affine this corresponds to the module of global sections $H^0(Z; p^*E)$. The Ulrich module in Proposition 2.17 is that module when $n = 4$ and $r = 2$. For our computational purposes realized in Section 2.4, we worked out the equivariant minimal free resolution as above. Interestingly, we do not know yet whether the `simpler' Ulrich sheaf presented in Section 2.3, which is new to our knowledge, generalizes to a construction for other varieties.

2.4 The Chow form of the essential variety

Grassmannians and Chow divisors

The Grassmannian variety $Gr(c; n+1) = Gr(\mathbb{P}^{c-1}; \mathbb{P}^n)$ parametrizes the linear subspaces of dimension $c-1$ in $\mathbb{P}^n$, i.e the $\mathbb{P}^{c-1}$'s in $\mathbb{P}^n$. Such a linear subspace may be given as the rowspace of a $c \times (n+1)$ matrix. The tuple of maximal minors of this matrix is uniquely determined by the linear subspace up to scale. The number of such minors is $\binom{n+1}{c}$. Hence we get a well-defined point in the projective space $\mathbb{P}^{\binom{n+1}{c}-1}$. This defines an embedding of the Grassmannian $Gr(c; n+1)$ into that projective space, called the Plücker embedding. Somewhat more algebraically, let W be a vector space of dimension $n+1$ and let $\mathbb{P}(W)$ be the space of lines in W through the origin. Then a linear subspace V of dimension c in W defines a line $\wedge^c V$ in $\wedge^c W$, and so it defines a point in $\mathbb{P}(\wedge^c W) = \mathbb{P}^{\binom{n+1}{c}-1}$. Thus the Grassmannian $Gr(c; W)$ embeds into $\mathbb{P}(\wedge^c W)$.

If X is a variety of codimension c in a projective space $\mathbb{P}^n$, then a linear subspace of dimension $c-1$ will typically not intersect X. The set of points in the Grassmannian $Gr(c; n+1)$ that do have nonempty intersection with X forms a divisor in $Gr(c; n+1)$, called the *Chow divisor*. This is seen by counting dimensions in the incidence diagram:

$$X \leftarrow \mathcal{X} = \{(x, L) \in X \times Gr(\mathbb{P}^{c-1}; \mathbb{P}^n) \mid x \in L\} \rightarrow Gr(\mathbb{P}^{c-1}; \mathbb{P}^n).$$

In detail, the fibers of the left projection are isomorphic to $Gr(\mathbb{P}^{c-2}; \mathbb{P}^{n-1})$, so they have dimension $(c-1)(n-c+1)$. We conclude that

$$\dim(\mathcal{X}) = (c-1)(n-c+1) + (n-c) = c(n+1-c) - 1.$$

Since the right arrow is degree 1 onto its image, that image has dimension $\dim(\mathcal{X})$, which is 1 less than $\dim(Gr(c; n+1))$. Next recall that the divisor class group of $Gr(c; n+1)$ is isomorphic to $\mathbb{Z}$. Considering the Plücker embedding $Gr(c; n+1) \to \mathbb{P}^{\binom{n+1}{c}-1}$, any hyperplane in the latter projective space intersects the Grassmannian in a divisor which generates the divisor class group of $Gr(c; n+1)$. This follows from an application of [49, Chapter II, Proposition 6.5(c)]. The homogeneous coordinate ring of this projective space $\mathbb{P}^{\binom{n+1}{c}-1} = \mathbb{P}(\wedge^c W)$ is $\mathrm{Sym}(\wedge^c W^*)$. Note that here $\wedge^c W^*$ are the linear forms, i.e. the elements of degree 1. If X has degree d, then its Chow divisor is cut out by a single form $Ch(X)$ of degree d unique up to nonzero scale, called the *Chow form*, in the coordinate ring of the Grassmannian $\mathrm{Sym}(\wedge^c W^*)/I_{Gr(c;n+1)}$.

As the parameters n, c, d increase, Chow forms become unwieldy to even store on a computer file. Arguably, the most efficient (and useful) representations of Chow forms are as determinants or Pfaffians of a matrix with entries in $\wedge^c W$. As we explain next, Ulrich sheaves can give such formulas.

Construction of Chow forms

We now explain how to obtain the Chow form $Ch(X)$ of a variety X from an Ulrich sheaf F whose support is X. The reference for this is [34, p. 552-553]. Let $M = \oplus_{d \in Z} H^0(P^n, F(d))$ be the graded module of twisted global sections over the polynomial ring $A = C[x_0, \ldots, x_n]$. We write W for the vector space generated by the variables $x_0, \ldots, x_n$. Consider the minimal free resolution (2.5) of M. The map d_i may be represented by a matrix D_i of size $\beta_i \times \beta_{i+1}$, with entries in the linear space W. Since (2.5) is a complex the product of two successive matrices $D_{i-1} D_i$ is the zero matrix. Note that when we multiply the entries of these matrices, we are multiplying elements in the ring $A = Sym(W) = C[x_0, \ldots, x_n]$.

Now comes the shift of view: Let $B = \oplus_{i=0}^{n} \wedge^i W$ be the exterior algebra on the vector space W. We now consider the entries in the D_i (which are all degree one forms in $A_1 = W = B_1$) to be in the ring B instead. We then multiply together all the matrices D_i corresponding to the maps d_i. The multiplications of the entries are performed in the skew-commutative ring B. We then get a product:

$$D = D_0 \cdot D_1 \cdots D_{c-1},$$

where c is the codimension of the variety X which supports F. If F has rank r and the degree of X is d, the matrix D is a nonzero $rd \times rd$ matrix. The entries in the product D now lie in $\wedge^c W$. Now comes the second shift of view: We consider the entries of D to be linear forms in the polynomial ring $Sym(\wedge^c W)$. Then we take the determinant of D, computed in this polynomial ring, and get a form of degree rd in $Sym(\wedge^c W)$. When considered in the coordinate ring of the Grassmannian $Sym(\wedge^c W)/I_G$, then $\det(D)$ equals the r^{th} power of the Chow form of X. For more information on the fascinating links between the symmetric and exterior algebras, the reader can start with the Bernstein-Gel'fand-Gel'fand correspondence as treated in [32].

Skew-symmetry of the matrices computing the Chow form of $PX^s_{4,2}$

In Section 2.3 we constructed two different Ulrich modules of rank 2 on the variety $PX^s_{4,2}$ of symmetric 4×4 matrices of rank ≤ 2. That variety has degree 10. The matrix D thus in both cases is 20 20, and its determinant is a square in $\mathrm{Sym}(\wedge^c W)$ as we now show. In fact, and here our analysis of the equivariant resolutions pays o , the matrix D in both cases is skew-symmetric when we use the bases distinguished by representation theory for the di erential matrices:

Lemma 2.20. Let $A; B; C$ be matrices of linear forms in the exterior algebra. Their products behave as follows under transposition:

1. $(A \quad B)^T = \quad B^T \quad A^T$
2. $(A \quad B \quad C)^T = \quad C^T \quad B^T \quad A^T$.

Proof. Part (1) is because $uv = \quad vu$ when u and v are linear forms in the exterior algebra. Part (2) is because $uvw = \quad wvu$ for linear forms in the exterior algebra. □

The resolutions (2.6) and (2.9) of our two Ulrich sheaves, have the form:

$$F \qquad G \qquad G \qquad F : \tag{2.11}$$

Dualizing and twisting we get the resolution:

$$F \quad ^T \quad G \quad ^T \quad G \quad ^T \quad F :$$

Since $= \quad ^T$, both and ^{T} map isomorphically onto the same image. We can therefore replace the map in (2.11) with ^{T}, and get the $GL(E)$-equivariant resolution:

$$F \qquad G \qquad G \quad ^T \quad F :$$

Let $\_; \_$ and $\_^T$ be the maps in the resolution above, but now considered to live over the exterior algebra. The Chow form associated to the two Ulrich sheaves is then the Pfa an of the matrix:

$$\_ \_ \_^T :$$

Proposition 2.21. The Chow form $Ch(P X^s_{4;2})$ constructed from the Ulrich sheaf is, in each case, the Pfa an of a 20 20 skew-symmetric matrix.

Proof. The Chow form squared is the determinant of $\_ \_ \_^T$ and we have:

$$\_ \_ \_^T \; ^T = \quad (\_^T)^T \; \_^T \; \_^T = \quad \_ \_ \_^T :$$

□

Explicit matrices computing the Chow form of $\mathbb{P}X^{s}_{4,2}$

Even though our primary aim is to compute the Chow form of the essential variety, we get explicit matrix formulas for the Chow form of $\mathbb{P}X^{s}_{4,2}$ as a by-product of our method. We carried out the computation in Proposition 2.21 in Macaulay2 for both Ulrich modules on $\mathbb{P}X^{s}_{4,2}$. We used the package PieriMaps to make matrices D_1 and D_2 representing and with respect to the built-in choice of bases parametrized by semistandard tableaux. We had to multiply D_2 on the right by a change of basis matrix to get a matrix representative with respect to dual bases, i.e. symmetric. For example in the case of the first Ulrich module (2.6) this change of basis matrix computes the perfect pairing $S_{3,2}(E) \otimes S_{3,3,1}(E) \to (\wedge^4 E)^{\otimes 3}$. Let us describe the transposed inverse matrix that represents the dual pairing. Columns are labeled by the semistandard Young tableaux S of shape $(3,2)$, and rows are labeled by the semistandard Young tableaux T of shape $(3,3,1)$. The (S,T)-entry in the matrix is obtained by fitting together the tableau S and the tableau T rotated by 180° into a tableau of shape $(3,3,3,3)$, straightening, and then taking the coefficient of $\begin{array}{|c|c|c|}\hline 0&0&0\\\hline 1&1&1\\\hline 2&2&2\\\hline 3&3&3\\\hline\end{array}$.

To finish for each Ulrich module, we took the product $D_1 D_2 D_1^T$ over the exterior algebra.

The two resulting explicit 20 × 20 skew-symmetric matrices are available as arXiv ancillary files or at this chapter's webpage[1]. Their Pfaffians equal the Chow form of $\mathbb{P}X^{s}_{4,2}$, which is an element in the homogeneous coordinate of the $Gr(3,10) = Gr(\mathbb{P}^2, \mathbb{P}^9)$. To get a feel for the 'size' of this Chow form, note that this ring is a quotient of the polynomial ring $\mathrm{Sym}(\wedge^3 \mathrm{Sym}_2(E))$ in 120 Plücker variables, denoted $\mathbb{Q}[p_{\{11,12,13\}}, \ldots, p_{\{33,34,44\}}]$ on our website, by the ideal minimally generated by 2310 Plücker quadrics. We can compute that the degree 10 piece where $Ch(\mathbb{P}X^{s}_{4,2})$ lives is a 108,284,013,552-dimensional vector space.

Both 20 × 20 matrices afford extremely compact formulas for this special element. Their entries are linear forms in $p_{\{11,12,13\}}, \ldots, p_{\{33,34,44\}}$ with one- and two-digit relatively prime integer coefficients. No more than 5 of the p-variables appear in any entry. In the first matrix, 96 off-diagonal entries equal 0. The matrices give new expressions for one of the two irreducible factors of a discriminant studied since 1879 by [89] and as recently as 2011 [87], as we see next in Remark 2.22.

Remark 2.22. From the subject of plane curves, it is classical that every ternary quartic form $f \in \mathbb{C}[x,y,z]_4$ can be written as $f = \det(xA + yB + zC)$ for some 4 × 4 symmetric matrices A, B, C. Geometrically, this expresses $V(f)$ inside the net of

[1] http://math.berkeley.edu/~jkileel/ChowFormulas.html

plane quadrics $\langle A, B, C\rangle$ as the locus of singular quadrics. By Theorem 7.5 of [87], that plane quartic curve $V(f)$ is singular if and only if the Vinnikov discriminant:

$$\Delta(A, B, C) = M(A, B, C)P(A, B, C)^2$$

evaluates to 0. Here M is a degree $(16, 16, 16)$ polynomial known as the tact invariant and P is a degree $(10, 10, 10)$ polynomial. The factor P equals the Chow form $\mathrm{Ch}(PX^s_{4,2})$ after substituting Plücker coordinates for Stiefel coordinates:

$$p_{\{i_1j_1, i_2j_2, i_3j_3\}} = \det \begin{pmatrix} a_{i_1j_1} & a_{i_2j_2} & a_{i_3j_3} \\ b_{i_1j_1} & b_{i_2j_2} & b_{i_3j_3} \\ c_{i_1j_1} & c_{i_2j_2} & c_{i_3j_3} \end{pmatrix}.$$

Explicit matrices computing the Chow form of E_C

We now can put everything together and solve the problem raised by Agarwal, Lee, Sturmfels and Thomas in [1] of computing the Chow form of the essential variety. In Proposition 2.7, we constructed a linear embedding $s : \mathbb{P}^8 \to \mathbb{P}^9$ that restricts to an embedding $E_C \to PX^s_{4,2}$. Both of our Ulrich sheaves supported on $PX^s_{4,2}$ pull back to Ulrich sheaves supported on E_C, and their minimal free resolutions pull back to minimal free resolutions:

$$s^*F \longrightarrow s^*G \longrightarrow s^*G \longrightarrow s^*F.$$

Here we verified in Macaulay2 that s^* quotients by a linear form that is a nonzero divisor for the two Ulrich modules. So, to get the Chow form $\mathrm{Ch}(E_C)$ from Propositions 2.12 and 2.17, we took matrices D_1 and D_2 symmetrized from above, and applied s^*. That amounts to substituting $x_{ij} = s(M)_{ij}$, where $s(M)$ is from Section 2.2. We then multiplied $D_1 D_2 D_1^T$, which is a product of a 20×60, a 60×60 and a 60×20 matrix, over the exterior algebra.

The two resulting explicit 20×20 skew-symmetric matrices are available at the chapter's webpage. Their Pfaffians equal the Chow form of E_C, which is an element in the homogeneous coordinate of $\mathrm{Gr}(\mathbb{P}^2, \mathbb{P}^8)$. We denote that ring as the polynomial ring in 84 (dual) Plücker variables $\mathbb{Q}[q_{\{11,12,13\}}, \ldots, q_{\{31,32,33\}}]$ modulo 1050 Plücker quadrics. Here $\mathrm{Ch}(E_C)$ lives in the 9,386,849,472-dimensional subspace of degree 10 elements.

Both matrices are excellent representations of $\mathrm{Ch}(E_C)$. Their entries are linear forms in $q_{\{11,12,13\}}, \ldots, q_{\{31,32,33\}}$ with relatively prime integer coefficients less than 216 in absolute value. In the first matrix, 96 off-diagonal entries vanish, and no entries have full support.

Bringing this back to computer vision, we can now prove our main result stated in Section 2.1:

Proof of Theorem 2.1. We first construct $M(x^{(i)}, y^{(i)})$, and then we prove that it has the desired properties. For the construction, let Z denote the 6×9 matrix:

$$\begin{pmatrix}
y_1^{(1)}x_1^{(1)} & y_1^{(1)}x_2^{(1)} & y_1^{(1)} & y_2^{(1)}x_1^{(1)} & y_2^{(1)}x_2^{(1)} & y_2^{(1)} & x_1^{(1)} & x_2^{(1)} & 1 \\
y_1^{(2)}x_1^{(2)} & y_1^{(2)}x_2^{(2)} & y_1^{(2)} & y_2^{(2)}x_1^{(2)} & y_2^{(2)}x_2^{(2)} & y_2^{(2)} & x_1^{(2)} & x_2^{(2)} & 1 \\
y_1^{(3)}x_1^{(3)} & y_1^{(3)}x_2^{(3)} & y_1^{(3)} & y_2^{(3)}x_1^{(3)} & y_2^{(3)}x_2^{(3)} & y_2^{(3)} & x_1^{(3)} & x_2^{(3)} & 1 \\
y_1^{(4)}x_1^{(4)} & y_1^{(4)}x_2^{(4)} & y_1^{(4)} & y_2^{(4)}x_1^{(4)} & y_2^{(4)}x_2^{(4)} & y_2^{(4)} & x_1^{(4)} & x_2^{(4)} & 1 \\
y_1^{(5)}x_1^{(5)} & y_1^{(5)}x_2^{(5)} & y_1^{(5)} & y_2^{(5)}x_1^{(5)} & y_2^{(5)}x_2^{(5)} & y_2^{(5)} & x_1^{(5)} & x_2^{(5)} & 1 \\
y_1^{(6)}x_1^{(6)} & y_1^{(6)}x_2^{(6)} & y_1^{(6)} & y_2^{(6)}x_1^{(6)} & y_2^{(6)}x_2^{(6)} & y_2^{(6)} & x_1^{(6)} & x_2^{(6)} & 1
\end{pmatrix}$$

where the columns of Z are labeled by $11, 12, \ldots, 33$ respectively. Now set q_{ijk} to be the determinant of Z with columns i, j, k removed, and substitute into either of 20×20 skew-symmetric matrices above that compute the Chow form Ch_{E_C}. This constructs $M(x^{(i)}, y^{(i)})$.

Observe that $M(x^{(i)}, y^{(i)})$ drops rank if and only if the subspace $\ker(Z) \subset \mathbb{P}^8$ intersects E_C. This follows by definition of the Chow form, since the Plücker coordinates of $\ker(Z)$ equal the maximal minors of Z when Z is full-rank [43, p. 94]. Said differently:

$$M(x^{(i)}, y^{(i)}) \text{ drops rank} \iff \exists\, M \in E_C \text{ such that } \forall i = 1, \ldots, 6 \quad \begin{pmatrix} y_1^{(i)} & y_2^{(i)} & 1 \end{pmatrix} M \begin{pmatrix} x_1^{(i)} \\ x_2^{(i)} \\ 1 \end{pmatrix} = 0. \tag{2.12}$$

Indeed, (2.12) is a linear system for $M \in E_C$, while Z is the coefficient matrix of that system.

In the rest of the proof of Theorem 2.1, we relate solutions of (2.1) to solutions of (2.12). As goes computer vision parlance, we will move between cameras and world points to relative poses, and then back. In the first direction, given $\{(x^{(i)}, y^{(i)})\}$, suppose that we have a solution $A, B, X^{(1)}, \ldots, X^{(6)}$ to (2.1). Note that the group:

$$G := \{ g \in GL(4, \mathbb{C}) \mid (g_{ij})_{1 \le i,j \le 3} \in SO(3, \mathbb{C}) \text{ and } g_{41} = g_{42} = g_{43} = 0 \}$$

equals the stabilizer of the set of calibrated camera matrices inside $\mathbb{C}^{3\times 4}$, with respect to right multiplication. We now make two simplifying assumptions about our solution to (2.1).

- Without loss of generality, $A = [\mathrm{id}_{3\times 3} \mid 0]$. For otherwise, select $g \in G$ so that $Ag = [\mathrm{id}_{3\times 3} \mid 0]$, and then $Ag, Bg, g^{-1}\tilde{X}^{(1)}, \ldots, g^{-1}\tilde{X}^{(6)}$ is also a solution to (2.1).
- Denoting $B = [R \mid t]$ for $R \in SO(3, \mathbb{C})$ and $t \in \mathbb{C}^3$, then without loss of generality, $t \neq 0$. For otherwise, we may zero out the last coordinate of each $\tilde{X}^{(i)}$ and replace B by $[R \mid t']$ for any $t' \in \mathbb{C}^3$, and then we still have a solution to the system (2.1).

Denote $[t]_\times := \begin{pmatrix} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ -t_2 & t_1 & 0 \end{pmatrix}$. Set $M = [t]_\times R$. Then $M \in \mathcal{E}_{\mathbb{C}}$ and M is a solution to (2.12), as for each $i = 1, \ldots, 6$ we have:

$$
\begin{aligned}
\begin{pmatrix} y_1^{(i)} & y_2^{(i)} & 1 \end{pmatrix} M \begin{pmatrix} x_1^{(i)} \\ x_2^{(i)} \\ 1 \end{pmatrix} &\sim (B\tilde{X}^{(i)})^T M (A\tilde{X}^{(i)}) \\
&= \tilde{X}^{(i)T} \left([R \mid t]^T [t]_\times R [\mathrm{id}_{3\times 3} \mid 0]\right) \tilde{X}^{(i)} \\
&= \tilde{X}^{(i)T} \left([R \mid 0]^T [t]_\times [R \mid 0]\right) \tilde{X}^{(i)} \\
&= 0.
\end{aligned}
$$

Here the second-to-last equality is because $t^T[t]_\times = 0$, and the last equality is because the matrix in parentheses is skew-symmetric. We have shown the second sentence in Theorem 2.1.

Conversely, given $\{(x^{(i)}, y^{(i)})\}$, let us start with a solution $M \in \mathcal{E}_{\mathbb{C}}$ to system (2.12). From this, we will produce a solution to (2.1) provided that M is sufficiently nice. More precisely, assume:

1. M may be factored as a skew-symmetric matrix times a rotation matrix, i.e. $M = [t]_\times R$ where $t \in \mathbb{C}^3$ and $R \in SO(3, \mathbb{C})$.
2. For $i = 1, \ldots, 6$, we have $\begin{pmatrix} y_1^{(i)} & y_2^{(i)} & 1 \end{pmatrix} M \neq 0$ and $M \begin{pmatrix} x_1^{(i)} & x_2^{(i)} & 1 \end{pmatrix}^T \neq 0$.

For readers of [48, Section 9.2.4], condition 2 means that $\{(x^{(i)}, y^{(i)})\}$ avoids the epipoles of M. Supposing conditions 1 and 2 hold, we set $A = [\mathrm{id}_{3\times 3} \mid 0]$ and $B = [R \mid t]$. Then there exists $\tilde{X}^{(i)} \in \mathbb{P}^3$ such that $A\tilde{X}^{(i)} \sim \tilde{x}^{(i)}$ and $B\tilde{X}^{(i)} \sim \tilde{y}^{(i)}$. Indeed (dropping i for convenience), we take $\tilde{X} = \begin{pmatrix} x_1 & x_2 & 1 & \lambda \end{pmatrix}^T$ where $\lambda \in \mathbb{C}$ satisfies $R\tilde{x} + \lambda t \sim \tilde{y}$. To find such λ, we solve $\begin{pmatrix} R\tilde{x} & t & \tilde{y} \end{pmatrix}_{3\times 3} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = 0$ for $\alpha, \beta, \gamma \in \mathbb{C}$ with $\alpha \neq 0$. These equations are soluble since:

- $0 = \tilde{y}^T M \tilde{x} = \tilde{y}^T [t]_\times R\tilde{x} = \det\begin{pmatrix} R\tilde{x} & t & \tilde{y} \end{pmatrix}$ by Laplace expansion along the third column.
- $0 \neq M\tilde{x} = t \times R\tilde{x} \Rightarrow$ columns t and $R\tilde{x}$ are linearly independent. Likewise, $0 \neq \tilde{y}^T M = \tilde{y}^T [t]_\times R \Rightarrow$ columns $\tilde{y}$ and t are linearly independent.

This produces cameras and world points $A, B, \tilde{X}^{(1)}, \dots, \tilde{X}^{(6)}$ satisfying (2.1), given an essential matrix M satisfying the epipolar constraints (2.12) as well as the two regularity conditions above. To complete the proof of Theorem 2.1, it is enough to show that in the variety of point correspondences $\{(x^{(i)}, y^{(i)})\}$ where (2.12) is soluble, there is a dense, open subset of point correspondences where all solutions $M \in \mathcal{E}_{\mathbb{C}}$ satisfy conditions 1 and 2 above.

To this end, consider the diagram below, with projections π_1 and π_2 respectively:

$$(\mathbb{C}^2 \times \mathbb{C}^2)^6 \leftarrow (\mathbb{C}^2 \times \mathbb{C}^2)^6 \times \mathcal{E}_{\mathbb{C}} \rightarrow \mathcal{E}_{\mathbb{C}}.$$

Inside $(\mathbb{C}^2 \times \mathbb{C}^2)^6 \times \mathcal{E}_{\mathbb{C}}$ with coordinates (x, y, M), we consider three incidence varieties:

$$I_0 := \left\{ (x, y, M) \;\middle|\; \left(\tilde{y}^{(i)}\right)^T M \, \tilde{x}^{(i)} = 0 \text{ for each } i = 1, \dots, 6 \right\}$$

$$I_1 := \left\{ (x, y, M) \in I_0 \;\middle|\; M \text{ does } \underline{\text{not}} \text{ factor as } M = [t]_\times R \text{ for any } t \in \mathbb{C}^3, R \in SO(3, \mathbb{C}) \right\}$$

$$I_2 := \left\{ (x, y, M) \in I_0 \;\middle|\; M \tilde{x}^{(i)} = 0 \text{ or } \left(\tilde{y}^{(i)}\right)^T M = 0 \text{ for some } i = 1, \dots, 6 \right\}.$$

So, $\pi_1(I_0)$ is the variety of point correspondences where (2.12) is soluble, i.e. the hypersurface $V(\det(M(x^{(i)}, y^{(i)}))) \subset (\mathbb{C}^2 \times \mathbb{C}^2)^6$, while $\pi_1(I_0) \setminus \pi_1(I_1) \cup \pi_1(I_2)$ consists of those point correspondences where (2.12) is soluble and all solutions to

(2.12) satisfy conditions 1 and 2 above. We will show that $_1(I_1)$ and $_1(I_2)$ are closed subvarieties with dimension $<$ 23.

For I_1, note that $_2(I_1)$ E_C is a closed subvariety, the complement of the open orbit from the proof of Proposition 2.2, i.e. the 4-dimensional Tan(S). Also each ber of $_2|_{I_1}$ is 18-dimensional. It follows that I_1 is closed and 22-dimensional. So $_1(I_1)$ is closed and has dimension 22.

Next for I_2, note that $_2|_{I_2}$ surjects onto E_C and has general bers that are 17-dimensional. So I_2 is closed and 22-dimensional, implying that $_1(I_1)$ is closed and has dimension 22.

At this point, we have shown the converse in Theorem 2.1, and this completes the proof. □

We illustrate the main theorem with two examples. Note that since the rst example is a `positive', it is a strong (and reassuring) check of correctness for our formulas.

Example 2.23. Consider the image data of 6 point correspondences $\{(x^{(i)}, y^{(i)}) \in \mathbb{P}^2 \times \mathbb{P}^2 \mid i = 1, \ldots, m\}$ given by the corresponding rows of the two matrices:

$$[x^{(i)}] = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & \frac{1}{2} & 1 \\ 3 & 0 & 1 \\ \frac{3}{2} & \frac{5}{2} & 1 \\ 1 & \frac{1}{7} & 1 \end{pmatrix} \qquad [y^{(i)}] = \begin{pmatrix} \frac{8}{11} & \frac{16}{11} & 1 \\ \frac{7}{22} & \frac{5}{22} & 1 \\ \frac{8}{29} & \frac{34}{29} & 1 \\ \frac{17}{20} & 1 & 1 \\ \frac{1}{7} & \frac{1}{7} & 1 \\ \frac{9}{4} & \frac{3}{4} & 1 \end{pmatrix}:$$

In this example, they do come from world points $X^{(i)} \in \mathbb{P}^3$ and calibrated cameras $A; B$:

$$X^{(i)} = \begin{pmatrix} 0 & 0 & 2 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 2 & 4 & 1 \\ 3 & 0 & 1 & 1 \\ 3 & 5 & 2 & 1 \\ 7 & 1 & 7 & 1 \end{pmatrix}; \qquad A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}; \qquad B = \begin{pmatrix} \frac{7}{9} & \frac{4}{9} & \frac{4}{9} & 0 \\ \frac{4}{9} & \frac{1}{9} & \frac{8}{9} & 0 \\ \frac{4}{9} & \frac{8}{9} & \frac{1}{9} & 1 \end{pmatrix}:$$

To detect this, we form the 6×9 matrix Z from the proof of Theorem 2.1:

$$Z = \begin{pmatrix} 0 & 0 & \frac{8}{11} & 0 & 0 & \frac{16}{11} & 0 & 0 & 1 \\ \frac{7}{22} & \frac{7}{22} & \frac{7}{22} & \frac{5}{22} & \frac{5}{22} & \frac{5}{22} & 1 & 1 & 1 \\ 0 & \frac{4}{29} & \frac{8}{29} & 0 & \frac{17}{29} & \frac{34}{29} & 0 & \frac{1}{2} & 1 \\ \frac{51}{20} & 0 & \frac{17}{20} & 3 & 0 & 1 & 3 & 0 & 1 \\ \frac{3}{14} & \frac{5}{14} & \frac{1}{7} & \frac{3}{14} & \frac{5}{14} & \frac{1}{7} & \frac{3}{2} & \frac{5}{2} & 1 \\ \frac{9}{4} & \frac{9}{28} & \frac{9}{4} & \frac{3}{4} & \frac{3}{28} & \frac{3}{4} & 1 & \frac{1}{7} & 1 \end{pmatrix}.$$

We substitute the maximal minors of Z into the matrices computing $\mathrm{Ch}(E_C)$ in Macaulay2. The determinant command then outputs 0. This computation recovers the fact that the point correspondences are images of 6 world points under a pair of calibrated cameras.

Example 2.24. Random data $\{(x^{(i)}; y^{(i)}) \in \mathbb{R}^2 \times \mathbb{R}^2 \mid i = 1; \ldots; 6\}$ is expected to land outside the Chow divisor of E_C. We made an instance using the random(QQ) command in Macaulay2 for each coordinate of image point. The coordinates ranged from $\frac{1}{8}$ to 5 in absolute value. We carried out the substitution from Example 2.23, and got two full-rank skew-symmetric matrices with Pfaffians $\approx 5.5 \times 10^{25}$ and 1.3×10^{22}, respectively. These matrices certified that the system (2.1) admits no solutions for that random input.

The following proposition is based on general properties of Chow forms, collectively known as the U-resultant method to solve zero-dimensional polynomial systems. In our situation, it gives a connection with the 'five-point algorithm' for computing essential matrices. The proposition is computationally inefficient as-is for that purpose, but see [80] for a more efficient algorithm that would exploit our matrix formulas for $\mathrm{Ch}(E_C)$. Implementing the algorithms in [80] for our matrices is one avenue for future work.

Proposition 2.25. *Given a generic 5-tuple $\{(x^{(i)}; y^{(i)}) \in \mathbb{R}^2 \times \mathbb{R}^2 \mid i = 1; \ldots; 5\}$, if we make the substitution from the proof of Theorem 2.1, then the Chow form $\mathrm{Ch}(E_C)$*

specializes to a polynomial in $\mathbb{R}[x_1^{(6)}; x_2^{(6)}; y_1^{(6)}; y_2^{(6)}]$. Over $\mathbb{C}$, this specialization completely splits as:

$$\prod_{i=1}^{10} \begin{pmatrix} y_1^{(6)} & y_2^{(6)} & 1 \end{pmatrix} M^{(i)} \begin{pmatrix} x_1^{(6)} \\ x_2^{(6)} \\ 1 \end{pmatrix} :$$

Here $M^{(1)}; \ldots; M^{(10)} \in E_{\mathbb{C}}$ are the essential matrices determined by the given ve-tuple.

Proof. By the proof of Theorem 2.1, any zero of the above product is a zero of the specialization of $\mathrm{Ch}(E_{\mathbb{C}})$. By Hilbert's Nullstellensatz, this implies that the product divides the specialization. But both polynomials are inhomogeneous of degree 20, so they are . □

2.5 Numerical experiments with noisy point correspondences

In this section, we step back from the algebraic derivation above, and evaluate the output on noisy data. Commonly, a shortcoming in applications of algebra is that exact formulae cannot handle inexact data. In the present case, correctly matched point pairs come to the computer vision practitioner with noise, from the optical process in the camera itself as well as from pixelation. See Chapter 4 for a treatment of the related issue of image distortion.

Question 2.26. While in Theorem 2.1 the matrix $M(x; y)$ drops rank when there is an exact solution to (2.1), how can we tell if there is an approximate solution?

Since we have a matrix formula instead of a gigantic fully expanded polynomial formula, there is a positive answer to Question 2.26. We calculate the Singular Value Decomposition of the matrices $M(x; y)$ from Theorem 2.1, when a noisy six-tuple of image point correspondences is plugged in. An approximately rank-de cient SVD is expected when there exists an approximate solution to (2.1), as Singular Value Decomposition is numerically stable [26, Section 5.2]. In a slogan: given matrix formulas, we look at *spectral gaps* in the presence of noise.

Here is experimental evidence this works. For experiments, we assumed uniform noise from $\mathrm{unif}[-10^{-r}; 10^{-r}]$; this arises in image processing from pixelation [14, Section 4.5]. For each $r = 1; 1.5; 2; \ldots; 15$, we executed 500 of the following trials:

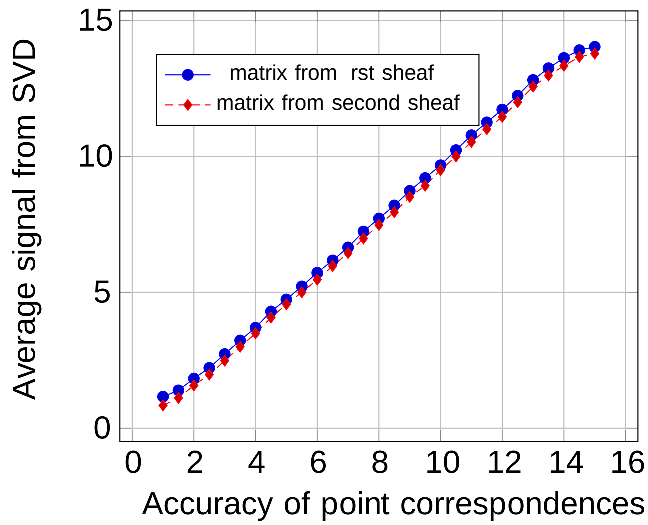


Figure 2.1: Both matrices from Theorem 2.1 detect approximately consistent point pairs.

- Pseudo-randomly generate an exact six-tuple of image point correspondences

$$\{(x^{(i)}, y^{(i)}) \in \mathbb{Q}^2 \times \mathbb{Q}^2 \mid i = 1, \ldots, 6\}$$

  with coordinates of size $O(1)$.

- Corrupt each image coordinate in the six-tuple by adding an independent and identically distributed sample from $\text{unif}[-10^{-r}, 10^{-r}]$.

- Compute the SVD's of both 20×20 matrices $M(x, y)$, derived from the first and second Ulrich sheaf respectively, with the above noisy image coordinates plugged in.

These experiments were performed in Macaulay2 using double precision for all floating-point arithmetic. Since it is a little subtle, we elaborate on our algorithm to pseudo-randomly generate exact correspondences in the first bullet. It breaks into three steps:

1. Generate calibrated cameras $A, B \in \mathbb{Q}^{3 \times 4}$. To do this, we sample twice from the Haar measure on $SO(3, \mathbb{R})$ and sample twice from the uniform measure on the radius 2 ball centered at the origin in $\mathbb{R}^3$. Then we concatenate nearby points in $SO(3, \mathbb{Q})$ and $\mathbb{Q}^3$ to obtain A and B. To find the nearby rotations, we pullback under $\mathbb{R}^3 \to S^3 \setminus \{N\} \to SO(3, \mathbb{R})$, we take nearby points in $\mathbb{Q}^3$, and then we pushforward.

2. Generate world points $X^{(i)} \in \mathbb{Q}^3$ $(i = 1, \ldots, 6)$. To do this, we sample six times from the uniform measure on the radius 6 ball centered at the origin in $\mathbb{R}^3$ (a choice fitting with some real-world data) and then we replace those by nearby points in $\mathbb{Q}^3$.

3. Set $x^{(i)} \leftarrow \{AX^{(i)}\}$ and $y^{(i)} \leftarrow \{BX^{(i)}\}$.

The most striking takeaway of our experiments is stated in the following result concerning the bottom spectral gaps we observed. Bear in mind that since $M(x;y)$ is skew-symmetric, its singular values occur with multiplicity two, so $\sigma_{19}(M(x;y)) = \sigma_{20}(M(x;y))$.

Empirical Fact 2.27. In the experiments described above, we observed for both matrices:

$$\frac{\sigma_{18}(M(x;y))}{\sigma_{20}(M(x;y))} = O(10^{-r}).$$

Here $M(x;y)$ has r-noisy image coordinates, and σ_i denotes the i^{th} largest singular value.

Figure 2.1 above plots $\mathrm{Log}_{10}\left(\frac{\sigma_{18}(M(x;y))}{\sigma_{20}(M(x;y))}\right)$ averaged over the 500 trials against r.

In this chapter, we resolved an open problem raised by Sameer Agarwal, vision expert at Google Research, by characterizing consistent point pairs across two calibrated views. Our output is an explicit matrix formula, robust to noisy measurements, which could be used for screening out wrongly matched point pairs inside RANSAC loops. Our derivation combined the algebraic theory of Ulrich sheaves with a geometric study based on secant varieties. In particular, we constructed a new low rank equivariant Ulrich sheaf supported on a determinantal variety.

Chapter 3

Three Cameras

This chapter is mostly based on my single-authored paper [61] about the recovery of 3 calibrated cameras from image data, to be published in the *SIAM Journal on Applied Algebra and Geometry*. The last section presents a general-purpose homotopy-continuation software for implicitization in computational algebraic geometry, joint with Justin Chen [21], currently submitted for publication and publicly released.

3.1 Introduction

As described in Chapter 1, 3D reconstruction is a fundamental task in computer vision, i.e. the recovery of three-dimensional scene geometry from two-dimensional images. In 1981, Fischler and Bolles proposed a methodology for 3D reconstruction that is robust to outliers in image data [37]. This is known as Random Sampling Consensus (RANSAC) and it is a paradigm in vision today [2]. RANSAC consists of three steps. Sketching the approach again, to compute a piece of the 3D scene:

Points, lines and other features that are images of the same source are detected in the photos. These matches are the *image data*.

A minimal sample of image data is randomly selected. *Minimal* means that only a positive finite number of 3D geometries are exactly consistent with the sample. Those 3D geometries are computed.

To each computed 3D geometry, the rest of the image data is compared. If one is approximately consistent with enough of the image data, it is kept. Else, the second step is repeated with a new sample.

Computing the finitely many 3D geometries is called a *minimal problem*. Typically, it is done by solving a corresponding zero-dimensional polynomial system, with coefficients that are functions of the sampled image data [64]. Since this step is carried out thousands of times in a full reconstruction, it is necessary to design efficient, specialized solvers. One of the most used *minimal solvers* in vision is Nistér's [82], based on Gröbner bases, to recover the relative position of two calibrated cameras. In Chapter 2 we considered a closely related problem about two calibrated cameras.

The concern of this chapter is the recovery of the relative position of *three* calibrated cameras from image data. To our knowledge, no satisfactory solution to this basic problem exists in the literature. Passing from two views to three views introduces a zoo of problems. Now feature lines, in addition to feature points, may be matched across images to recover camera positions. Our main result is the determination of the *algebraic degree* of 66 minimal problems for the recovery of three calibrated cameras; in other words, we find the generic number of complex solutions (see Theorem 3.6). Solution sets for particular random instances are available at:

https://math.berkeley.edu/~jkileel/CalibratedMinimalProblems.html.

As a by-product, we can derive minimal solvers for each case. Our techniques come from *numerical algebraic geometry* [92], and we rely on the homotopy continuation software Bertini [10]. This implies that our results are correct only with very high probability; in ideal arithmetic, with probability 1. Mathematically, the main object in this chapter is a particular projective algebraic variety $\mathcal{T}_{cal}$, which is a convenient moduli space for the relative position of three calibrated cameras. This variety is 11-dimensional, degree 4912 inside the projective space $\mathbb{P}^{26}$ of $3 \times 3 \times 3$ tensors (see Theorem 3.26). We call it the *calibrated trifocal variety*. Theorem 3.28 formulates our minimal problems as slicing $\mathcal{T}_{cal}$ by special linear subspaces of $\mathbb{P}^{26}$.

The rest of this chapter is organized as follows. In Section 3.2, we make our minimal problems mathematically precise and we state Theorem 3.6. In Section 3.3, we examine image correspondences using multiview varieties and then trifocal tensors [48, Chapter 15]. In Section 3.4, we prove that trifocal tensors and camera configurations are equivalent. In Section 3.5, we introduce the calibrated trifocal variety $\mathcal{T}_{cal}$ and prove several useful facts. Finally, in Section 3.6, we present a computational proof of the main result Theorem 3.6. In the last Section 3.7, we switch gears and present our Macaulay2 software package for implicitization in computational algebraic geometry. NumericalImplicitization is based on homotopy continuation, and my interest in writing general-purpose numerical algebraic geometry code grew out my approach to the minimal problems in Theorem 3.6.

3.2 Statement of main result

We begin by giving several definitions. Throughout this chapter, we work with the standard camera model of the projective camera [48, Section 6.2].

Definition 3.1. A (projective) *camera* is a full rank 3×4 matrix in $\mathbb{C}^{3 \times 4}$ defined up to multiplication by a nonzero scalar.

Thus, as noted in Section 1.1, a camera corresponds to a linear projection $\mathbb{P}^3 \dashrightarrow \mathbb{P}^2$. The *center* of a camera A is the point $\ker(A) \in \mathbb{P}^3$. A camera is *real* if $A \in \mathbb{R}^{3 \times 4}$.

Definition 3.2. A *calibrated camera* is a 3×4 matrix in $\mathbb{C}^{3 \times 4}$ whose left 3×3 submatrix is in the special orthogonal group $SO(3, \mathbb{C})$.

Real calibrated cameras have the interpretation of cameras with known and normalized internal parameters (e.g. focal length) [48, Subsection 6.2.4]. In practical situations, this information can be available during 3D reconstruction. Note that calibration of a camera is preserved by right multiplication by elements of the following subgroup of $GL(4, \mathbb{C})$:

$$G := \{ g \in \mathbb{C}^{4 \times 4} \mid (g_{ij})_{1 \leq i,j \leq 3} \in SO(3, \mathbb{C}),\ g_{41} = g_{42} = g_{43} = 0 \text{ and } g_{44} \neq 0 \}.$$

Elements in G act on $\mathbb{A}^3 \subset \mathbb{P}^3$ as composites of rotations, translations and central dilations. In the calibrated case of 3D reconstruction, one aims to recover camera positions (and afterwards the 3D scene) up to those motions, since recovery of absolute positions is not possible from image data alone.

Definition 3.3. A *configuration of three calibrated cameras* is an orbit of the action of the group G above on the set:

$$\{ (A, B, C) \mid A, B, C \text{ are calibrated cameras} \}$$

via simultaneous right multiplication.

By abuse of notation, we will call (A, B, C) a calibrated camera configuration, instead of always denoting the orbit containing (A, B, C).

As mentioned in Section 3.1, the image data used in 3D reconstruction typically are points and lines in the photos that match. This is made precise as follows. Call elements of $\mathbb{P}^2$ *image points*, and elements of the dual projective plane $(\mathbb{P}^2)^\vee$ *image lines*. An element of $(\mathbb{P}^2 \sqcup (\mathbb{P}^2)^\vee)^{\times 3}$ is a *point/line image correspondence*. For example, an element of $\mathbb{P}^2 \times \mathbb{P}^2 \times (\mathbb{P}^2)^\vee$ is called a point-point-line image correspondence, denoted PPL.

Definition 3.4. A calibrated camera configuration (A, B, C) is consistent with a given point/line image correspondence if there exist a point in $\mathbb{P}^3$ and a line in $\mathbb{P}^3$ containing it such that are such that (A, B, C) respectively map these to the given points and lines in $\mathbb{P}^2$.

For example, explicitly, a configuration (A, B, C) is consistent with a given point-point-line image correspondence $(x, x', \ell'') \in \mathbb{P}^2 \times \mathbb{P}^2 \times (\mathbb{P}^2)^\vee$ if there exist $(X, L) \in \mathbb{P}^3 \times \mathrm{Gr}(\mathbb{P}^1, \mathbb{P}^3)$ with $X \in L$ such that $AX = x, BX = x'$, and $CL = \ell''$. In particular, this implies that $X \notin \ker(A), \ker(B)$ and $\ker(C) \notin L$. We say that a configuration (A, B, C) is consistent with a set of point/line correspondences if it is consistent with each correspondence.

Example 3.5. Given the following set of real, random correspondences[1]:

$$\mathrm{PPP}: \begin{bmatrix} 0.6132 \\ 0.8549 \\ 0.5979 \end{bmatrix}, \begin{bmatrix} 0.4599 \\ 0.5713 \\ 0.1812 \end{bmatrix}, \begin{bmatrix} 0.6863 \\ 0.4508 \\ 0.1834 \end{bmatrix} \qquad \mathrm{PPL}: \begin{bmatrix} 0.6251 \\ 0.9248 \\ 0.9849 \end{bmatrix}, \begin{bmatrix} 0.3232 \\ 0.5453 \\ 0.6941 \end{bmatrix}, \begin{bmatrix} 0.3646 \\ 0.1497 \\ 0.1364 \end{bmatrix}$$

$$\mathrm{PPL}: \begin{bmatrix} 0.4970 \\ 0.6532 \\ 0.8429 \end{bmatrix}, \begin{bmatrix} 0.5405 \\ 0.8342 \\ 0.6734 \end{bmatrix}, \begin{bmatrix} 0.2692 \\ 0.8861 \\ 0.1333 \end{bmatrix} \qquad \mathrm{PPL}: \begin{bmatrix} 0.2896 \\ 0.6909 \\ 0.4914 \end{bmatrix}, \begin{bmatrix} 0.6898 \\ 0.9855 \\ 0.6777 \end{bmatrix}, \begin{bmatrix} 0.6519 \\ 0.8469 \\ 0.6855 \end{bmatrix}$$

$$\mathrm{PPL}: \begin{bmatrix} 0.8933 \\ 0.3375 \\ 0.1054 \end{bmatrix}, \begin{bmatrix} 0.7062 \\ 0.6669 \\ 0.7141 \end{bmatrix}, \begin{bmatrix} 0.3328 \\ 0.8228 \\ 0.6781 \end{bmatrix}.$$

In the notation of Theorem 3.6, this is a generic instance of the minimal problem '$1\,\mathrm{PPP} + 4\,\mathrm{PPL}$'. Up to the action of G, there are only a positive finite number of three calibrated cameras that are exactly consistent with this image data, namely 160 complex configurations. For this instance, it turns out that 18 of those configurations are real. For example, one is:

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, B = \begin{bmatrix} 0.22 & 0.95 & 0.18 & 1 \\ 0.96 & 0.24 & 0.08 & 1.44 \\ 0.12 & 0.15 & 0.97 & 0.97 \end{bmatrix}, C = \begin{bmatrix} 0.17 & 0.94 & 0.28 & 1.41 \\ 0.95 & 0.22 & 0.18 & 0.13 \\ 0.24 & 0.23 & 0.94 & 1.16 \end{bmatrix}.$$

In a RANSAC run for 3D reconstruction, the image data above is identified by feature detection software such as SIFT [73]. Also, only the real configurations are compared for agreement with further image data.

In Example 3.5 above, 160 is the *algebraic degree* of the minimal problem '$1\,\mathrm{PPP} + 4\,\mathrm{PPL}$'. This means that for correspondences in a nonempty Zariski open (hence measure 1) subset of $(\mathbb{P}^2 \times \mathbb{P}^2 \times \mathbb{P}^2) \times (\mathbb{P}^2 \times \mathbb{P}^2 \times (\mathbb{P}^2)^\vee)^{4}$, there are 160 consistent complex configurations. Given generic real correspondences, the number of real configurations varies, but 160 is an upper bound.

[1] For ease of presentation, double precision floating point numbers are truncated here.

Theorem 3.6. The rows of the following table display the algebraic degree for 66 minimal problems across three calibrated views. Given generic point/line image correspondences in the amount specified by the entries in the first five columns, then the number of calibrated camera configurations over $\mathbb{C}$ that are consistent with those correspondences equals the entry in the sixth column.

# PPP	# PPL	# PLP	# LLL	# PLL	#configurations
3	1	0	0	0	272
3	0	0	1	0	216
3	0	0	0	2	448
2	2	0	0	1	424
2	1	1	0	1	528
2	1	0	1	1	424
2	1	0	0	3	736
2	0	0	2	1	304
2	0	0	1	3	648
2	0	0	0	5	1072
1	4	0	0	0	160
1	3	1	0	0	520
1	3	0	1	0	360
1	3	0	0	2	520
1	2	2	0	0	672
1	2	1	1	0	552
1	2	1	0	2	912
1	2	0	2	0	408
1	2	0	1	2	704
1	2	0	0	4	1040
1	1	1	2	0	496
1	1	1	1	2	896
1	1	1	0	4	1344
1	1	0	3	0	368
1	1	0	2	2	736
1	1	0	1	4	1184
1	1	0	0	6	1672
1	0	0	4	0	360
1	0	0	3	2	696
1	0	0	2	4	1176
1	0	0	1	6	1680
1	0	0	0	8	2272
0	5	0	0	1	160
0	4	1	0	1	616
0	4	0	1	1	456
0	4	0	0	3	616
0	3	2	0	1	1152
0	3	1	1	1	880
0	3	1	0	3	1280
0	3	0	2	1	672
0	3	0	1	3	1008
0	3	0	0	5	1408
0	2	2	1	1	1168
0	2	2	0	3	1680
0	2	1	2	1	1032
0	2	1	1	3	1520
0	2	1	0	5	2072
0	2	0	3	1	800
0	2	0	2	3	1296
0	2	0	1	5	1848
0	2	0	0	7	2464
0	1	1	3	1	1016
0	1	1	2	3	1552
0	1	1	1	5	2144
0	1	1	0	7	2800
0	1	0	4	1	912
0	1	0	3	3	1456
0	1	0	2	5	2088
0	1	0	1	7	2808
0	1	0	0	9	3592
0	0	0	5	1	920
0	0	0	4	3	1464
0	0	0	3	5	2176
0	0	0	2	7	3024
0	0	0	1	9	3936
0	0	0	0	11	4912

Remark 3.7. A calibrated camera configuration (A, B, C) has 11 degrees of freedom (Theorem 3.26), and the first five columns in the table above represent conditions of codimension $3, 2, 2, 2, 1$; respectively (Theorem 3.28).

Remark 3.8. The algebraic degrees in Theorem 3.6 are intrinsic to the underlying camera geometry. However, our method of proof uses a device from multiview geometry called trifocal tensors, which breaks symmetry between (A, B, C). There are other minimal problems for three calibrated views involving image correspondences of type LPP , LPL , LLP . These also possess intrinsic algebraic degrees; but they are not covered by the non-symmetric proof technique used here.

3.3 Correspondences

In this section, we examine point/line image correspondences. In the first part, we use multiview varieties to describe correspondences. This approach furnishes exact polynomial systems for the minimal problems in Theorem 3.6. However, each parametrized system has a different structure (in terms of number and degrees of equations). This would force a direct analysis for Theorem 3.6 to proceed case-by-case, and moreover, each system so obtained is computationally unwieldy. In Subsection 3.3, we recall the construction of the trifocal tensor [48, Chapter 15]. This is a point $T_{A,B,C} \in \mathbb{C}^{3\times 3\times 3}$ associated to cameras (A, B, C). It encodes necessary conditions for (A, B, C) to be consistent with different types of correspondences. Tractable relaxations to the minimal problems in Theorem 3.6 are thus obtained, each with similar structure. We emphasize that everything in Section 3.3 applies equally to calibrated cameras (A, B, C) as well as to uncalibrated cameras.

multiview varieties

Let $A, B, C \in \mathbb{C}^{3\times 4}$ be three projective cameras, not necessarily calibrated. Denote by $: \mathbb{P}^3 \dashrightarrow \mathbb{P}^2_A$, $: \mathbb{P}^3 \dashrightarrow \mathbb{P}^2_B$, $: \mathbb{P}^3 \dashrightarrow \mathbb{P}^2_C$ the corresponding linear projections. We make:

Definition 3.9. Fix projective cameras A, B, C as above. Denote by $F\ell_{0,1}$ the incidence variety $(X, L) \in \mathbb{P}^3 \times Gr(\mathbb{P}^1, \mathbb{P}^3)$ $X \in L$. Then the:

PLL multiview variety denoted $X^{PLL}_{A,B,C}$ is the closure of the image of

$$F\ell_{0,1} \dashrightarrow \mathbb{P}^2_A \times (\mathbb{P}^2_B)^\vee \times (\mathbb{P}^2_C)^\vee, \quad (X, L) \mapsto \;\; (X), \;\; (L), \;\; (L)$$

LLL multiview variety denoted $X^{LLL}_{A,B,C}$ is the closure of the image of

$$\mathrm{Gr}(\mathbb{P}^1,\mathbb{P}^3) \dashrightarrow (\mathbb{P}^2_A)^\vee \times (\mathbb{P}^2_B)^\vee \times (\mathbb{P}^2_C)^\vee, \; L \mapsto (L), \; (L), \; (L)$$

PPL multiview variety denoted $X^{PPL}_{A,B,C}$ is the closure of the image of

$$F_{\ell_{0,1}} \dashrightarrow \mathbb{P}^2_A \times \mathbb{P}^2_B \times (\mathbb{P}^2_C)^\vee, \; (X,L) \mapsto (X), \; (X), \; (L)$$

PLP multiview variety denoted $X^{PLP}_{A,B,C}$ is the closure of the image of

$$F_{\ell_{0,1}} \dashrightarrow \mathbb{P}^2_A \times (\mathbb{P}^2_B)^\vee \times \mathbb{P}^2_C, \; (X,L) \mapsto (X), \; (L), \; (X)$$

PPP multiview variety denoted $X^{PPP}_{A,B,C}$ is the closure of the image of

$$\mathbb{P}^3 \dashrightarrow \mathbb{P}^2_A \times \mathbb{P}^2_B \times \mathbb{P}^2_C, \; X \mapsto (X), \; (X), \; (X) \, .$$

Next, we give the dimension and equations for these multiview varieties; the PPP case has appeared in [5]. In the following, we notate $x \in \mathbb{P}^2_A$, $x' \in \mathbb{P}^2_B$, $x'' \in \mathbb{P}^2_C$ for image points and $\ell \in (\mathbb{P}^2_A)^\vee$, $\ell' \in (\mathbb{P}^2_B)^\vee$, $\ell'' \in (\mathbb{P}^2_C)^\vee$ for image lines. Also, we postpone treatment of the PLL case to Subsection 3.3. In particular, the trilinear form $T_{A,B,C}(x,\ell',\ell'')$ will be defined there.

Theorem 3.10. Fix A, B, C. The multiview varieties from Definition 3.9 are irreducible. If A, B, C have linearly independent centers in $\mathbb{P}^3$, then the varieties have the following dimensions and multi-homogeneous prime ideals.

$\dim(X^{PLL}_{A,B,C}) = 5$ and $I(X^{PLL}_{A,B,C}) = \langle T_{A,B,C}(x,\ell',\ell'')\rangle \subseteq \mathbb{C}[x_i, \ell'_j, \ell''_k]$

$\dim(X^{LLL}_{A,B,C}) = 4$ and $I(X^{LLL}_{A,B,C}) \subseteq \mathbb{C}[\ell_i, \ell'_j, \ell''_k]$ is generated by the maximal minors of the matrix $\begin{bmatrix} A^T\ell & B^T\ell' & C^T\ell'' \end{bmatrix}_{4\times 3}$

$\dim(X^{PPL}_{A,B,C}) = 4$ and $I(X^{PPL}_{A,B,C}) \subseteq \mathbb{C}[x_i, x'_j, \ell''_k]$ is generated by the maximal minors of the matrix $\begin{pmatrix} A & x & 0 \\ B & 0 & x' \\ \ell''^T C & 0 & 0 \end{pmatrix}_{7\times 6}$

$\dim(X^{PLP}_{A,B,C}) = 4$ and $I(X^{PLP}_{A,B,C}) \subseteq \mathbb{C}[x_i, \ell'_j, x''_k]$ is generated by the maximal minors of the matrix $\begin{pmatrix} A & x & 0 \\ C & 0 & x'' \\ \ell'^T B & 0 & 0 \end{pmatrix}_{7\times 6}$

$\dim(X^{PPP}_{A,B,C}) = 3$ and $I(X^{PPP}_{A,B,C}) \subset \mathbb{C}[x_i, x'_j, x''_k]$ is generated by the maximal minors of the matrix $\begin{pmatrix} A & x & 0 & 0 \\ B & 0 & x' & 0 \\ C & 0 & 0 & x'' \end{pmatrix}_{9 \times 7}$ together with $\det \begin{pmatrix} A & x & 0 \\ B & 0 & x' \end{pmatrix}_{6 \times 6}$ and $\det \begin{pmatrix} A & x & 0 \\ C & 0 & x'' \end{pmatrix}_{6 \times 6}$ and $\det \begin{pmatrix} B & x' & 0 \\ C & 0 & x'' \end{pmatrix}_{6 \times 6}$

Proof. Irreducibility is clear from Definition 3.9. For the dimension and prime ideal statements, we may assume that:

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}; \quad B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}; \quad C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

This is without loss of generality in light of the following group symmetries. Let $g, g', g'' \in SL(3, \mathbb{C})$ and $h \in SL(4, \mathbb{C})$. To illustrate, consider the third case above, and let $J^{PPL}_{A,B,C} \subset \mathbb{C}[x_i, x'_j, \ell''_k]$ be the ideal generated by the maximal minors mentioned there. It is straightforward to check that:

$$I(X^{PPL}_{Ah,Bh,Ch}) = I(X^{PPL}_{A,B,C}) \quad \text{and} \quad J^{PPL}_{Ah,Bh,Ch} = J^{PPL}_{A,B,C}.$$

Also, we can check that:

$$I(X^{PPL}_{gA,\, g'B,\, g''C}) = (g, g', \wedge^2 g'') \cdot I(X^{PPL}_{A,B,C})$$
$$\text{and} \quad J^{PPL}_{gA,\, g'B,\, g''C} = (g, g', (g''^T)^{-1}) \cdot J^{PPL}_{A,B,C}.$$

Here the left, linear action of $SL(3,\mathbb{C}) \times SL(3,\mathbb{C}) \times SL(3,\mathbb{C})$ on $\mathbb{C}[x_i, x'_j, \ell''_k]$ is via $(g, g', g'') \cdot f(x, x', \ell'') = f(g^{-1}x, g'^{-1}x', g''^{-1}\ell'')$ for $f \in \mathbb{C}[x_i, x'_j, \ell''_k]$. Also, $\wedge^2 g'' = (g''^T)^{-1} \in \mathbb{C}^{3 \times 3}$. So, for the PPL case, I and J transform in the same way when (A, B, C) is replaced by $(gAh, g'Bh, g''Ch)$; in the other cases, this holds similarly. Assuming that A, B, C have linearly independent centers, we may choose g, g', g'', h to harmlessly move the cameras into the position above. Now using the computer algebra system Macaulay2 [44], we verify the dimension and prime ideal statements for this special position. □

Remark 3.11. In Theorem 3.10, if A, B, C do not have linearly independent centers, then the minors described still vanish on the multiview varieties, by continuity in (A, B, C).

Now, certainly a point/line correspondence that is consistent with (A, B, C) lies in the appropriate multiview variety; consistency means that the correspondence is a point in the set-theoretic image of the appropriate rational map in Definition 3.9. Since the multiview varieties are the Zariski closures of those set-theoretic images, care is needed to make a converse. We require:

Definition 3.12. Let A, B, C be three projective cameras with distinct centers. The *epipole* denoted $e_{1\leftarrow 2}$ is the point $A(\ker(B)) \in \mathbb{P}^2_A$. That is, $e_{1\leftarrow 2}$ is the image under A of the center of B. Epipoles $e_{1\leftarrow 3}, e_{2\leftarrow 1}, e_{2\leftarrow 3}, e_{3\leftarrow 1}, e_{3\leftarrow 2}$ are defined similarly.

Lemma 3.13. *Let A, B, C be three projective cameras with distinct centers. Let $\gamma \in (\mathbb{P}^2 \sqcup (\mathbb{P}^2)^{\vee})^{\times 3}$. Assume this point/line correspondence* avoids epipoles*. For example, if $\gamma = (x, x', \ell'') \in \mathbb{P}^2_A \times \mathbb{P}^2_B \times (\mathbb{P}^2_C)^{\vee}$, avoidance of epipoles means that $x \neq e_{1\leftarrow 2}, e_{1\leftarrow 3}$; $x' \neq e_{2\leftarrow 1}, e_{2\leftarrow 3}$; and $\ell'' \not\ni e_{3\leftarrow 1}, e_{3\leftarrow 2}$. Then γ is consistent with (A, B, C) if γ is in the suitable multiview variety.*

Proof. Assuming that γ is in the multiview variety, then γ satisfies the equations from Theorem 3.10. This is equivalent to containment conditions on the back-projections of γ, without any hypothesis on the centers of A, B, C.

We spell this out for the PPL case, where $\gamma = (x, x', \ell'') \in \mathbb{P}^2_A \times \mathbb{P}^2_B \times (\mathbb{P}^2_C)^{\vee}$. Here the back-projections are the lines $A^{-1}(x), B^{-1}(x') \subseteq \mathbb{P}^3$ and the plane $C^{-1}(\ell'') \subseteq \mathbb{P}^3$. The minors from Theorem 3.10 vanish if and only if there exists $(X, L) \in F\ell_{0,1}$ such that $X \in A^{-1}(x)$, $X \in B^{-1}(x')$ and $L \subseteq C^{-1}(\ell'')$. To see this, note that the minors vanish only if:

$$\begin{pmatrix} A & x & 0 \\ B & 0 & x' \\ \ell''^{T} C & 0 & 0 \end{pmatrix} \begin{pmatrix} X \\ \lambda \\ \lambda' \end{pmatrix} = 0 \ \text{ for some nonzero } \begin{pmatrix} X \\ \lambda \\ \lambda' \end{pmatrix} \in \mathbb{C}^6,$$

where $X \in \mathbb{C}^4$, $\lambda \in \mathbb{C}$ and $\lambda' \in \mathbb{C}$. Since $x, x' \in \mathbb{C}^3$ are nonzero, it follows that X is nonzero, and so defines a point $X \in \mathbb{P}^3$. From $AX = \lambda x$, the line $A^{-1}(x) \subseteq \mathbb{P}^3$ contains $X \in \mathbb{P}^3$. Similarly $AX = \lambda' x$ implies $X \in B^{-1}(x')$. Thirdly, $\ell''^{T} C X = 0$ says that X lies on the plane $C^{-1}(\ell'') \subseteq \mathbb{P}^3$. Now taking any line $L \subseteq \mathbb{P}^3$ with $X \in L \subseteq C^{-1}(\ell'')$ produces a satisfactory point $(X, L) \in F\ell_{0,1}$, and reversing the argument gives the converse.

Returning to the lemma, since γ avoids epipoles, the back-projections of γ avoid the centers of A, B, C. In the PPL case, this implies that (X, L) avoids the centers of A, B, C. Thus (X, L) witnesses consistency, because $A(X) = x$, $B(X) = x'$, $C(L) = \ell''$. The other cases are finished similarly. ☐

The results of this subsection have provided tight equational formulations for a camera configuration and a point/line image correspondence to be consistent. This leads to a parametrized system of polynomial equations for each minimal problem in Theorem 3.6. For instance, for the minimal problem '1 PPP + 4 PPL', the unknowns are the entries of A, B, C, up to the action of the group G. Due to Theorem 3.10, there are $\binom{9}{7} + 3 + 4\binom{7}{6} = 67$ quartic equations. Their coefficients are parametrized cubically and quadratically by the image data in $(\mathbb{P}^2)^{11} \times ((\mathbb{P}^2)^{\vee})^{4}$. Since this parameter space is irreducible, to find the generic number of solutions to the system, we may specialize to **one** random instance, such as in Example 3.5. Nonetheless, solving a single instance of this system – 'as is' – is computationally intractable, let alone solving systems for the other minimal problems present in Theorem 3.6.

The way out is to nontrivially replace the above systems with other systems, which enlarge the solution sets but amount to accessible computations. This key maneuver is based on *trifocal tensors* from multiview geometry. Before doing so, we justify calling the problems in Theorem 3.6 minimal.

Proposition 3.14. *For each problem in Theorem 3.6, given generic correspondence data, there is a finite number*[2] *of solutions, i.e. calibrated camera configurations (A, B, C). Moreover, solutions have linearly independent centers.*

Proof. For calibrated A, B, C, we may act by G so $A = \begin{bmatrix} I_{3\times 3} & 0 \end{bmatrix}$, $B = \begin{bmatrix} R_2 & t_2 \end{bmatrix}$ and $C = \begin{bmatrix} R_3 & t_3 \end{bmatrix}$ where $R_2, R_3 \in \mathrm{SO}(3,\mathbb{C})$ and $t_2, t_3 \in \mathbb{C}^3$. Furthermore, t_2 and t_3 may be jointly scaled. Thus, if A, B, C have non-identical centers, we get a point in $\mathrm{SO}(3,\mathbb{C})^{\times 2} \times \mathbb{P}^5$. This point is unique and configurations with non-identical centers are in bijection with $\mathrm{SO}(3,\mathbb{C})^{\times 2} \times \mathbb{P}^5$.

Now consider one of the minimal problems from Theorem 3.6, 'w_1PPP + w_2PPL + w_3PLP + w_4LLL + w_5PLL '. Notice that the problems in Theorem 3.6, are those for which the weights $(w_1, w_2, w_3, w_4, w_5) \in \mathbb{Z}_{\geq 0}$ satisfy $3w_1 + 2w_2 + 2w_3 + 2w_4 + w_5 = 11$ and $w_2 \geq w_3$. Image correspondence data is a point in the product $\mathcal{D}_w := (\mathbb{P}^2 \times \mathbb{P}^2 \times \mathbb{P}^2)^{\times w_1} \times \dots \times (\mathbb{P}^2 \times (\mathbb{P}^2)^{\vee} \times (\mathbb{P}^2)^{\vee})^{\times w_5}$.

Consider the incidence diagram:

$$\mathrm{SO}(3,\mathbb{C})^{\times 2} \times \mathbb{P}^5 \leftarrow \Gamma \rightarrow \mathcal{D}_w$$

where $\Gamma := \overline{\{((A,B,C), d) \in \mathrm{SO}(3,\mathbb{C})^{\times 2} \times \mathbb{P}^5 \times \mathcal{D}_w \mid (A,B,C) \text{ and } d \text{ are consistent}\}}$

and where the arrows are projections. The left map is surjective and a general fiber is a product of multiview varieties described by Theorem 3.10. In particular, the

[2] This number is shown to be positive in the proof of Theorem 3.6.

ber has dimension $3w_1 + 4w_2 + 4w_3 + 4w_4 + 5w_5$. Therefore, by [31, Corollary 13.5], has dimension $11 + 3w_1 + 4w_2 + 4w_3 + 4w_4 + 5w_5$, as $\dim(SO(3;\mathbb{C})^{2} \quad \mathbb{P}^5) = 11$. Now, the second arrow is a regular map between varieties of the same dimension, because $11 + 3w_1 + 4w_2 + 4w_3 + 4w_4 + 5w_5 = 6(w_1 + w_2 + w_3 + w_4 + w_5)$. So, if it is dominant, then again by [31, Corollary 13.5], a general ber has dimension 0; otherwise, a general ber is empty. However, note that points in a general ber of the second map correspond to solutions of a generic instance of the problem indexed by w from Theorem 3.6. This shows that those problems generically have nitely many solutions.

We can see that generically there are no solutions with non-identical but collinear centers, as follows. Let $C \quad SO(3;\mathbb{C})^{2} \quad \mathbb{P}^5$ be the closed variety of con gurations $(A;B;C)$ with non-identical but collinear centers. Consider:

$$C \qquad {}^{0}\!\rightarrow D_w$$

where the de nition of ^{0} is the de nition of with $SO(3;\mathbb{C})^{2} \quad \mathbb{P}^5$ replaced by C, and where the arrows are projections. Here $\dim(C) = 10$. The left arrow is surjective, and a general ber is a product of multiview varieties, with the same dimension as in the above case. This dimension statement is seen by calculating the multiview varieties as in the proof of Theorem 3.10, when $(A;B;C)$ have distinct, collinear centers. It follows that $\dim(\ ^{0}) = 10 + 3w_1 + 4w_2 + 4w_3 + 4w_4 + 5w_5 < 11 + 3w_1 + 4w_2 + 4w_3 + 4w_4 + 5w_5 = 6(w_1 + w_2 + w_3 + w_4 + w_5) = \dim(D_w)$ so that the right arrow is not dominant.

Finally, to see that generically there is no solution $(A;B;C)$ where the centers of $A;B;C$ are identical in $\mathbb{P}^3$, we may mimic the above argument with another dimension count. Calibrated con gurations with identical centers are in bijection with $SO(3;\mathbb{C})^{2}$, because each G-orbit has a unique representative of the form $A = I_{3\ 3}\ 0$, $B = R_2\ 0$, $C = R_3\ 0$ where $R_2;R_3 \in SO(3;\mathbb{C})$. So, analogously to before, we consider the diagram:

$$SO(3;\mathbb{C})^{2} \qquad {}^{00}\!\rightarrow D_w$$

where the de nition of ^{00} is the de nition of with $SO(3;\mathbb{C})^{2} \quad \mathbb{P}^5$ replaced by $SO(3;\mathbb{C})^{2}$, and where the arrows are projections. Again, the left arrow is surjective, and a general ber is a product of multiview varieties. Here, when $A;B;C$ have identical centers, a calculation as in the proof of Theorem 3.10 ver-i es that the dimensions of the multiview varieties drop, as follows: $\dim(X^{PLL}_{A;B;C}) = 3; \dim(X^{LLL}_{A;B;C}) = 2; \dim(X^{PPL}_{A;B;C}) = 3; \dim(X^{PLP}_{A;B;C}) = 3; \dim(X^{PPP}_{A;B;C}) = 2$. So the dimension of a general ber of the left arrow is $2w_1 + 3w_2 + 3w_3 + 2w_4 + 5w_3$. So $\dim(\ ^{00}) = 6 + 2w_1 + 3w_2 + 3w_3 + 2w_4 + 5w_3 < 11 + 3w_1 + 4w_2 + 4w_3 + 4w_4 + 5w_5 = 6(w_1 + w_2 + w_3 + w_4 + w_5) = \dim(D_w)$, whence the right arrow is not dominant. This completes the proof. □

Trifocal tensors

In this subsection, we re-derive the trifocal tensor $T_{A;B;C} \in \mathbb{C}^{3\times 3\times 3}$ associated to cameras $(A;B;C)$, following the projective geometry approach of Hartley [46]. This explains the notation in the $\mathcal{PLL}$ bullet of Theorem 3.10, and justifies the assertion made there. The trifocal tensor and its calibrated version are the analogs of the fundamental matrix and essential matrix from two-view geometry (see Chapter 2). We will review how $T_{A;B;C}$ encodes point/line correspondences besides $\mathcal{PLL}$ as well.

As in Subsection 3.3, let $A;B;C \in \mathbb{C}^{3\times 4}$ be three projective cameras, not necessarily calibrated, and denote by $\alpha : \mathbb{P}^3 \dashrightarrow \mathbb{P}^2_A$, $\beta : \mathbb{P}^3 \dashrightarrow \mathbb{P}^2_B$, $\gamma : \mathbb{P}^3 \dashrightarrow \mathbb{P}^2_C$ the corresponding linear projections. Let the point and lines $x \in \mathbb{P}^2_A; \ell' \in (\mathbb{P}^2_B)^\vee; \ell'' \in (\mathbb{P}^2_C)^\vee$ be given as column vectors. The pre-image $\alpha^{-1}(x)$ is a line in $\mathbb{P}^3$, while $\beta^{-1}(\ell')$ and $\gamma^{-1}(\ell'')$ are planes in $\mathbb{P}^3$. We can characterize when these three have non-empty intersection as follows.

First, note that the plane $\beta^{-1}(\ell')$ is given by the column vector $B^T\ell'$ since $X \in \mathbb{P}^3$ satisfies $X \in \beta^{-1}(\ell')$ if and only if $0 = \ell'^T BX = (B^T\ell')^T X$. Similarly, the plane $\gamma^{-1}(\ell'')$ is given by $C^T\ell''$. For the line $\alpha^{-1}(x)$, note:

$$\alpha^{-1}(x) = \bigcap_{\substack{\ell \in (\mathbb{P}^2_A)^\vee \\ \ell^T x = 0}} \alpha^{-1}(\ell) \subseteq \alpha^{-1}\langle x; \begin{bmatrix}1 & 1 & 0\end{bmatrix}^T\rangle \cap \alpha^{-1}\langle x; \begin{bmatrix}1 & 0 & 1\end{bmatrix}^T\rangle:$$

Here $\langle\ \rangle$ denotes span, and auxiliary points $\begin{bmatrix}1 & 1 & 0\end{bmatrix}^T; \begin{bmatrix}1 & 0 & 1\end{bmatrix}^T \in \mathbb{P}^2_A$ are simply convenient choices for this calculation. Unless those two points and x are collinear, the inclusion above is an equality, and the intersectands in the RHS are the planes given by the column vectors $A^T[x]_\times \begin{bmatrix}1 & 1 & 0\end{bmatrix}^T$ and $A^T[x]_\times \begin{bmatrix}1 & 0 & 1\end{bmatrix}^T$. The notation means $[x]_\times = \begin{bmatrix}0 & -x_3 & x_2\\ x_3 & 0 & -x_1\\ -x_2 & x_1 & 0\end{bmatrix}$, and $[x]_\times y$ gives $\langle x;y\rangle$ for $x \neq y \in \mathbb{P}^2_A$. So, $\alpha^{-1}(x) \cap \beta^{-1}(\ell') \cap \gamma^{-1}(\ell'') \neq \emptyset$ only if:

$$\det\left(A^T[x]_\times\begin{bmatrix}1\\1\\0\end{bmatrix} \quad A^T[x]_\times\begin{bmatrix}1\\0\\1\end{bmatrix} \quad B^T\ell' \quad B^T\ell''\right)_{4\times 4} = 0: \tag{3.1}$$

This determinant is divisible by $(x_1 - x_2 - x_3)$, since that vanishes if and only if $x; \begin{bmatrix}1 & 1 & 0\end{bmatrix}^T; \begin{bmatrix}1 & 0 & 1\end{bmatrix}^T$ are collinear only if the first two columns above are linearly dependent. Hence, factoring out, we obtain a constraint that is trilinear in $x;\ell';\ell''$, i.e., we get for some tensor $T \in \mathbb{C}^{3\times 3\times 3}$:

$$\sum_{1 \le i,j,k \le 3} T_{ijk}\, x_i\, \ell'_j\, \ell''_k = 0.$$

The tensor entry T_{ijk} is computed by substituting into (3.1) the basis vectors $x = e_i, \ell' = e_j, \ell'' = e_k$. Breaking into cases according to i, this yields:

$$T_{1ij} = \tfrac{1}{(1\;0\;0)} \det\begin{bmatrix} a_3 & -a_2 & b_j & c_k \end{bmatrix} = \det\begin{bmatrix} a_2 & a_3 & b_j & c_k \end{bmatrix}$$

$$T_{2ij} = \tfrac{1}{(0\;1\;0)} \det\begin{bmatrix} -a_3 + a_1 & -a_3 & b_j & c_k \end{bmatrix} = -\det\begin{bmatrix} a_1 & a_3 & b_j & c_k \end{bmatrix}$$

$$T_{3ij} = \tfrac{1}{(0\;0\;1)} \det\begin{bmatrix} -a_1 + a_2 & -a_2 & b_j & c_k \end{bmatrix} = \det\begin{bmatrix} a_1 & a_2 & b_j & c_k \end{bmatrix}$$

where a_i denotes the transpose of the first row in A, and so on.

At this point, we have derived formula (17.12) from [48, p. 415]:

Definition 3.15. Let A, B, C be cameras. Their *trifocal tensor* $T_{A,B,C} \in \mathbb{C}^{3\times 3\times 3}$ is computed as follows. Form the 4×9 matrix $\begin{bmatrix} A^T & B^T & C^T \end{bmatrix}$. Then for $1 \le i,j,k \le 3$, the entry $(T_{A,B,C})_{ijk}$ is $(-1)^{i+1}$ times the determinant of the 4×4 submatrix gotten by omitting the i^{th} column from A^T, while keeping the j^{th} and k^{th} columns from B^T and C^T, respectively. If A, B, C are calibrated, then $T_{A,B,C}$ is said to be a *calibrated trifocal tensor* (first introduced by Weng et al. in [103] before [46]).

Remark 3.16. Since $A, B, C \in \mathbb{C}^{3\times 4}$ are each defined only up to multiplication by a nonzero scalar, the same is true of $T_{A,B,C} \in \mathbb{C}^{3\times 3\times 3}$.

Remark 3.17. By construction, $T_{A,B,C}(x, \ell', \ell'') := \sum_{1 \le i,j,k \le 3} T_{ijk}\, x_i\, \ell'_j\, \ell''_k = 0$ is equivalent to $A^{-1}(x) \cap B^{-1}(\ell') \cap C^{-1}(\ell'') \neq \emptyset$. In particular, $T_{A,B,C} = 0$ if and only if the centers of A, B, C are all the same. Moreover, the PLL cases in Theorem 3.10 and Lemma 3.13 postponed above are now immediate.

So far, we have constructed trifocal tensors so that they encode point-line-line image correspondences. Conveniently, the same tensors encode other point/line correspondences [46], up to extraneous components.

Proposition 3.18. Let A, B, C be projective cameras. Let $x \in \mathbb{P}^2_A, x' \in \mathbb{P}^2_B, x'' \in \mathbb{P}^2_C$ and $\ell \in (\mathbb{P}^2_A)^\vee, \ell' \in (\mathbb{P}^2_B)^\vee, \ell'' \in (\mathbb{P}^2_C)^\vee$. Putting $T = T_{A,B,C}$, then (A, B, C) is consistent with:

$(x;\ell';\ell'')$ only if $T(x;\ell';\ell'') = 0$ [PLL]

$(\ell;\ell';\ell'')$ only if $[\ell]\ \ T(\ ;\ell';\ell'') = 0$ [LLL]

$(x;\ell';x'')$ only if $[x'']\ \ T(x;\ell';\ \ ) = 0$ [PLP]

$(x;x';\ell'')$ only if $[x']\ \ T(x;\ \ ;\ell'') = 0$ [PPL]

$(x;x';x'')$ only if $[x'']\ \ T(x;\ \ ;\ \ )[x']\ \ = 0.$ [PPP]

In the middle bullets, each contraction of T with two vectors gives a column vector in $\mathbb{C}^3$. In the last bullet, $T(x;\ \ ;\ \ ) = \sum_{i=1}^{3} x_i (T_{ijk})_{1\ \ j;k\ \ 3} \in \mathbb{C}^{3\ \ 3}$.

Proof. This proposition matches Table 15.1 on [48, p. 372]. To be self-contained, we recall the proof. The first bullet is by construction of T.

For the second bullet, assume that $(\ell;\ell';\ell'')$ is consistent with $(A;B;C)$, i.e. there exists $L \in Gr(\mathbb{P}^1;\mathbb{P}^3)$ such that $\ (L) = \ell;\ \ (L) = \ell';\ \ (L) = \ell''$. Now let $y \in \ell$ be a point. So $\ ^{1}(x)$ is a line in the plane $\ ^{1}(\ell)$ and that plane contains the line L. This implies $\ ^{1}(x) \cap L \neq\ ;\ )\ \ \ ^{1}(x) \cap\ \ ^{1}(\ell') \cap\ \ ^{1}(\ell'') \neq\ ;\ ,\ \ T(y;\ell';\ell'') = 0$. It follows that for $y \in \mathbb{P}^2_A$, we have $y^T\ell = 0\ )\ \ y^T T(\ ;\ell';\ell'') = 0$. This means that ℓ and $T(\ ;\ell';\ell'')$ are linearly independent, i.e. $[\ell]\ \ T(\ ;\ell';\ell'') = 0$.

The third, fourth and fifth bullets are similar. They come from reasoning that the consistency implies, respectively:

$x'' \in k''\ )\ \ T(x;\ell';k'') = 0$

$x' \in k'\ )\ \ T(x;k';\ell'') = 0$

$x' \in k'$ and $x'' \in k''\ )\ \ T(x;k';k'') = 0,$

where $k' \in (\mathbb{P}^2_B)^-$ and $k'' \in (\mathbb{P}^2_C)^-$. □

Remark 3.19. The constraints in Proposition 3.18 are linear in T. We will exploit this in Section 3.6. Also, in fact, image correspondences of type LPL, LLP and LPP do not give linear constraints on $T_{A;B;C}$. This is the reason that these types are not considered in Theorem 3.6. To get linear constraints nonetheless, one could permute $A;B;C$ before forming the trifocal tensor.

In this subsection, we have presented a streamlined account of trifocal tensors, and the point/line image correspondences that they encode. Now, we sketch the relationship between the *tight* conditions in Theorem 3.10 and the *necessary* conditions in Proposition 3.18 for consistency.

Lemma 3.20. *Fix projective cameras A, B, C with linearly independent centers. Then the trilinearities in Proposition 3.18 cut out subschemes of three-factor products of $\mathbb{P}^2$ and $(\mathbb{P}^2)^\vee$. In all cases of Proposition 3.18, this subscheme is reduced and contains the corresponding multiview variety as a top-dimensional component.*

Proof. Without loss of generality, A, B, C are in the special position from the proof of Theorem 3.10. Then using Macaulay2, we form the ideal generated by the trilinearities of Proposition 3.18 and saturate with respect to the irrelevant ideal. This leaves a radical ideal; we compute its primary decomposition. □

For example, in the case of PPP, the trilinearities from Proposition 3.18 generate a radical ideal in $\mathbb{C}[x_i, x'_j, x''_k]$ that is the intersection of:

- the 3 irrelevant ideals for each factor of $\mathbb{P}^2$
- 2 linear ideals of codimension 4
- the multiview ideal $I(X^{PPP}_{A,B,C})$.

This discrepancy between the trifocal and multiview conditions for PPP correspondences was studied in [100]. To demonstrate our main result, in Section 3.6 we shall relax the tight multiview equations in Theorem 3.10 to the merely necessary trilinearities in Proposition 3.18. The ‘top-dimensional’ clause in Lemma 3.20, as well as Theorem 3.22 in Section 3.4 below, indicate that this gives ‘good’ approximations to the minimal problems in Theorem 3.6.

3.4 Configurations

In this section, it is proven that trifocal tensors, in both the uncalibrated and calibrated case, are in bijection with camera triples up to the appropriate group action, i.e. with camera configurations. Already, it is very well-known throughout the vision community that “trifocal tensors encode relative camera positions” (e.g. see the appendix of [46] or [54] for a proof for general uncalibrated camera triples). We contribute precise hypotheses under which the correspondence is valid, namely that the three camera centers are linearly independent. We also verify that the correspondence is one-to-one, instead of finite-to-one, for calibrated trifocal tensors and the subgroup of transformations G in Theorem 3.22 below. To our knowledge, this fact is new; subtly, the analog for two calibrated cameras is false [48, Result 9.19]. In terms of Theorem 3.6, Theorem 3.22 enables us to compute consistent calibrated trifocal tensors in exchange for consistent calibrated camera configurations.

Proposition 3.21. *Let A, B, C be three projective cameras, with linearly independent centers in $\mathbb{P}^3$. Let $\tilde{A}, \tilde{B}, \tilde{C}$ be another three projective cameras. Then $T_{A,B,C} = T_{\tilde{A},\tilde{B},\tilde{C}} \in \mathbb{P}(\mathbb{C}^{3\times 3\times 3})$ if and only if there exists $h \in SL(4,\mathbb{C})$ such that $Ah = \tilde{A}$, $Bh = \tilde{B}$, $Ch = \tilde{C} \in \mathbb{P}(\mathbb{C}^{3\times 4})$.*

Proof. As in the proof of Theorem 3.10, for $g, g', g'' \in SL(3,\mathbb{C})$, $h \in SL(4,\mathbb{C})$:

$$T_{gA,\, g'B,\, g''C} = (g, \wedge^2 g', \wedge^2 g'') \cdot T_{A,B,C} \quad \text{and} \quad T_{Ah,\, Bh,\, Ch} = T_{A,B,C}\,. \tag{3.2}$$

The second equality gives the 'if' direction. Conversely, for 'only if', for any $g, g', g'' \in SL(3,\mathbb{C})$, $h_1, h_2 \in SL(4,\mathbb{C})$, we are free to replace (A, B, C) by $(gAh_1, g'Bh_1, g''Ch_1)$ and to replace $(\tilde{A}, \tilde{B}, \tilde{C})$ by $(g\tilde{A}h_2, g'\tilde{B}h_2, g''\tilde{C}h_2)$, and then to exhibit an h as in the proposition. Hence we may assume that:

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \; B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \; C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\tilde{A} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \; \tilde{B} = \begin{bmatrix} * & * & * & * \\ * & * & * & * \\ 0 & 0 & 0 & 1 \end{bmatrix}, \; \tilde{C} = \begin{bmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{bmatrix}$$

where each '$*$' denotes an indeterminate. Now consider the nine equations:

$$(T_{A,B,C})_{i3k} = (T_{\tilde{A},\tilde{B},\tilde{C}})_{i3k}$$

where $1 \le i, k \le 3$. Under the above assumptions, these are linear and in the nine unknowns $\tilde{c}_{lm}$ for $1 \le l, m \le 3$. Here we have fixed the nonzero scale of $\tilde{C}$ so that these are indeed equalities, on the nose. It follows that:

$$\tilde{C} = \begin{bmatrix} 1 & 0 & 0 & * \\ 0 & 0 & 1 & * \\ 0 & 0 & 0 & * \end{bmatrix}.$$

At this point, we have reduced to solving 18 equations in 11 unknowns:

$$(T_{A,B,C})_{ijk} = (T_{\tilde{A},\tilde{B},\tilde{C}})_{ijk}$$

where $1 \le i,k \le 3$ and $1 \le j \le 2$. These equations are quadratic monomials and binomials. The system is simple to solve by hand or with Macaulay2:

$$\tilde{A} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}; \quad \tilde{B} = \begin{bmatrix} \lambda & 0 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}; \quad \tilde{C} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \lambda^{-1} \end{bmatrix}$$

for $\lambda \in \mathbb{C}^*$. Taking $h = \lambda^{-3/4}\operatorname{diag}(\lambda,\lambda,\lambda,1) \in SL(4,\mathbb{C})$ gives $Ah = \tilde{A}, Bh = \tilde{B}, Ch = \tilde{C} \in \mathbb{P}(\mathbb{C}^{3\times 4})$, as desired. This completes the proof. □

With a bit of work, we can promote Proposition 3.21 to the calibrated case. A little explanation may be helpful here. Only a subgroup of projective transformations acts on triples of calibrated cameras, namely G. The content of Theorem 3.22 is that h can be taken to lie in G instead of just $h \in SL(4,\mathbb{C})$. See [59] for related issues regarding critical configurations.

Theorem 3.22. Let A, B, C be three calibrated cameras, with linearly independent centers in $\mathbb{P}^3$. Let $\tilde{A}, \tilde{B}, \tilde{C}$ be another three calibrated cameras. Then $T_{A,B,C} = T_{\tilde{A},\tilde{B},\tilde{C}} \in \mathbb{P}(\mathbb{C}^{3\times 3\times 3})$ if and only if there exists $h \in G$ (where G is defined on page 2) such that $Ah = \tilde{A}, Bh = \tilde{B}, Ch = \tilde{C} \in \mathbb{P}(\mathbb{C}^{3\times 4})$.

Proof. The 'if' direction is from Proposition 3.21. For 'only if', here for any $g, g', g'' \in SO(3,\mathbb{C})$, $h_1, h_2 \in G$, we are free to replace (A, B, C) by $(gAh_1, g'Bh_1, g''Ch_1)$ and to replace $(\tilde{A}, \tilde{B}, \tilde{C})$ by $(g\tilde{A}h_2, g'\tilde{B}h_2, g''\tilde{C}h_2)$, and then to exhibit an $h \in G$ as above. In this way, we may assume that:

$$A = \begin{bmatrix} I_{3\times 3} & 0 \end{bmatrix}; \quad B = \begin{bmatrix} I_{3\times 3} & s_1 \end{bmatrix}; \quad C = \begin{bmatrix} I_{3\times 3} & s_2 \end{bmatrix}$$

$$\tilde{A} = \begin{bmatrix} I_{3\times 3} & 0 \end{bmatrix}; \quad \tilde{B} = \begin{bmatrix} R_1 & t_1 \end{bmatrix}; \quad \tilde{C} = \begin{bmatrix} R_2 & t_2 \end{bmatrix}$$

where $R_1, R_2 \in SO(3,\mathbb{C})$ and $s_1, s_2, t_1, t_2 \in \mathbb{C}^3$. Now from Proposition 3.21, there exists $h' \in SL(4,\mathbb{C})$ such that $Ah' = \tilde{A}, Bh' = \tilde{B}, Ch' = \tilde{C} \in \mathbb{P}(\mathbb{C}^{3\times 3})$. From the first equality, it follows that $h' = \begin{bmatrix} I_{3\times 3} & 0 \\ u^T & \lambda \end{bmatrix} \in \mathbb{P}(\mathbb{C}^{4\times 4})$ for some $u \in \mathbb{C}^3$, $\lambda \in \mathbb{C}^*$. It suffices to show that $u = 0$, so $h' \in G$. By way of contradiction, let us assume that $u \neq 0$. Substituting into $Bh' = \tilde{B}$ gives:

$$\begin{bmatrix} I_{3\times 3} & s_1 \end{bmatrix} \begin{bmatrix} I_{3\times 3} & 0 \\ u^T & \lambda \end{bmatrix} = \begin{bmatrix} I_{3\times 3} + s_1 u^T & \lambda s_1 \end{bmatrix} = \begin{bmatrix} R_1 & t_1 \end{bmatrix} \in \mathbb{P}(\mathbb{C}^{3\times 4}):$$

In particular, there is $\lambda_1 \in \mathbb{C}$ so that $\lambda_1(I_{3\times 3} + s_1u^T) = R_1$. In particular, $R_1 - \lambda_1 I_{3\times 3}$ is rank at most 1. Equivalently, λ_1 is an eigenvalue of the rotation $R_1 \in SO(3;\mathbb{C})$ of geometric multiplicity at least 2. The only possibilities are $\lambda_1 = 1$ or $\lambda_1 = -1$. If $\lambda_1 = 1$, then $R_1 = I$ and $s_1u^T = 0$. From $u \neq 0$, we get that $s_1 = 0$; but then $A = B$, contradicting linear independence of the centers of $A; B; C$. So in fact $\lambda_1 = -1$. Now R_1 is a 180 degree rotation. From $R_1 + I_{3\times 3} = s_1u^T \in \mathbb{C}^{3\times 3}$, it follows that the axis of rotation is the line through u, and $s_1 = \frac{2u}{u^Tu}$. The exact same analysis holds starting from $\bar{C}\bar{h}$. So in particular, $s_2 = \frac{2u}{u^Tu}$. But now $B = C$, contradicting linear independence of the centers of $A; B; C$. We conclude that $u = 0$. □

3.5 Varieties

So far in Subsection 3.3 and Section 3.4, we have worked with individual trifocal tensors, uncalibrated or calibrated. This is possible once a camera configuration $(A; B; C)$ is given. To determine an unknown camera configuration from image data, we need to work with the set of all trifocal tensors.

Definition 3.23. The *trifocal variety*, denoted $\mathcal{T} \subseteq \mathbb{P}(\mathbb{C}^{3\times 3\times 3})$, is defined to be the Zariski closure of the image of the following rational map

$$\mathbb{P}(\mathbb{C}^{3\times 4}) \times \mathbb{P}(\mathbb{C}^{3\times 4}) \times \mathbb{P}(\mathbb{C}^{3\times 4}) \dashrightarrow \mathbb{P}(\mathbb{C}^{3\times 3\times 3}); \quad (A; B; C) \mapsto T_{A;B;C}$$

$$\text{where } (T_{A;B;C})_{ijk} := (-1)^{i+1}\det\begin{bmatrix} \hat a_i \\ b_j \\ c_k \end{bmatrix}_{4\times 4} \text{ for } 1 \le i; j; k \le 3.$$

Here $\hat a_i$ is gotten from A by omitting the i^{th} row, and $b_j; c_k$ are the $j^{th}; k^{th}$ rows of $B; C$ respectively. So, $\mathcal{T}$ is the closure of the set of all trifocal tensors.

Definition 3.24. The *calibrated trifocal variety*, denoted $\mathcal{T}_{cal} \subseteq \mathbb{P}(\mathbb{C}^{3\times 3\times 3})$, is defined to be the Zariski closure of the image of the following rational map

$$(SO(3;\mathbb{C}) \times \mathbb{C}^3) \times (SO(3;\mathbb{C}) \times \mathbb{C}^3) \times (SO(3;\mathbb{C}) \times \mathbb{C}^3) \dashrightarrow \mathbb{P}(\mathbb{C}^{3\times 3\times 3});$$

$$(R_1; t_1); (R_2; t_2); (R_3; t_3) \mapsto T_{[R_1|t_1]; [R_2|t_2]; [R_3|t_3]}$$

where the formula for T is as in Definitions 3.15 and 3.23. So, $\mathcal{T}_{cal}$ is the closure of the set of all calibrated trifocal tensors.

In the remainder of this chapter, the calibrated trifocal variety $\mathcal{T}_{cal}$ is the main actor. It is the higher version of the essential variety $\mathcal{E}$ starring in Chapter 2 above.

The calibrated trifocal variety has recently been studied independently by Martyushev [75] and Matthews [76]. Both authors obtain implicit quartic equations for $\mathcal{T}_{cal}$. However, a full set of ideal generators for $I(\mathcal{T}_{cal}) \subseteq \mathbb{C}[T_{ijk}]$ is currently not known. We summarize the state of knowledge on implicit equations for $\mathcal{T}_{cal}$:

Proposition 3.25. *The prime ideal of the calibrated trifocal variety $I(\mathcal{T}_{cal}) \subseteq \mathbb{C}[T_{ijk}]$ contains the ideal of the trifocal variety $I(\mathcal{T})$, and $I(\mathcal{T})$ is minimally generated by 10 cubics, 81 quintics and 1980 sextics. Additionally, $I(\mathcal{T}_{cal})$ contains 15 linearly independent quartics that do not lie in $I(\mathcal{T})$.*

The ideal containment follows from $\mathcal{T}_{cal} \subseteq \mathcal{T}$, and the statement about minimal generators of $I(\mathcal{T})$ was proven by Aholt and Oeding [4]. For the additional quartics, see [75, Theorems 8, 11] and [76, Corollary 51].

In the rest of this chapter, using numerical algebraic geometry, we always interact with the calibrated trifocal variety $\mathcal{T}_{cal}$ directly via (a restriction of) its defining parametrization. Therefore, we do not need the ideal of implicit equations $I(\mathcal{T}_{cal})$, nor do we use the known equations from Proposition 3.25.

At this point, we discuss properties of the rational map in Definition 3.24. First, since the source $(SO(3, \mathbb{C}) \times \mathbb{C}^3)^{\times 3}$ is irreducible, the closure of the image $\mathcal{T}_{cal}$ is irreducible. Second, the base locus of the map consists of triples of calibrated cameras $([R_1|t_1], [R_2|t_2], [R_3|t_3])$ all with the same center in $\mathbb{P}^3$, by the remarks following Definition 3.15. Third, the two equations in (3.2), the second line of the proof of Proposition 3.21, mean that the rational map in Definition 3.24 satisfies group symmetries. Namely, the parametrization of $\mathcal{T}_{cal}$ is equivariant with respect to $SO(3, \mathbb{C})^{\times 3}$, and each of its fibers carry a G action. In vision, these two group actions are interpreted as changing image coordinates and changing world coordinates. Here, by the equivariance, it follows that $\mathcal{T}_{cal}$ is an $SO(3, \mathbb{C})^{\times 3}$-variety. Also, we can use the G action on fibers to pick out one point per fiber, and thus restrict the map in Definition 3.24 so that the restriction is generically injective and dominant onto $\mathcal{T}_{cal}$. Explicitly, we restrict to the domain where $[R_1 | t_1] = [I_{3 \times 3} \mid 0]$; $t_2 = [\ast \ \ast \ 1]^T$. This restriction $(SO(3; \mathbb{C}) \times \mathbb{C}^2) \times (SO(3; \mathbb{C}) \times \mathbb{C}^3) \dashrightarrow \mathcal{T}_{cal}$ is generically injective by Theorem 3.22. Generic injectivity makes the restricted map particularly amenable to numerical algebraic geometry, where computations regarding a parametrized variety are pulled back to the source of the parametrization. We now obtain the major theorem of this section using that technique:

Theorem 3.26. *The calibrated trifocal variety $\mathcal{T}_{cal} \subseteq \mathbb{P}(\mathbb{C}^{3 \times 3 \times 3})$ is irreducible, of dimension 11 and degree 4912. It equals the $SO(3, \mathbb{C})^{\times 3}$-orbit closure generated by the* following projective plane, parametrized by $\begin{bmatrix}\lambda_1 & \lambda_2 & \lambda_3\end{bmatrix}^T \in \mathbb{P}^2$:

$$T_1 = \begin{bmatrix} 0 & \lambda_1 & \lambda_2 \\ 0 & 0 & 0 \\ \lambda_1 & 0 & 0 \end{bmatrix}; \quad T_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \lambda_1 & \lambda_2 \\ 0 & \lambda_3 & 0 \end{bmatrix}; \quad T_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & \lambda_1 & \lambda_2 + \lambda_3 \end{bmatrix}:$$

Computational Proof. Dimension 11 follows from the generically injective parametrization given above. The $SO(3;\mathbb{C})^{\times 3}$ statement follows from (3.2). In more detail, given a calibrated camera configuration $(A; B; C)$ with linearly independent centers, we may act by G so that the centers of $A; B; C$ are:

$$\begin{bmatrix}0 & 0 & 0 & 1\end{bmatrix}^T; \quad \begin{bmatrix}0 & 0 & 1 & 1\end{bmatrix}^T; \quad \begin{bmatrix}0 & & & 1\end{bmatrix}^T;$$

respectively. Then we may act by $SO(3;\mathbb{C})^{\times 3}$ so that the left submatrices of $A; B; C$ equal $I_{3 \times 3}$. The calibrated trifocal tensor $T_{A;B;C}$ now lands in the stated $\mathbb{P}^2$. Hence, $\mathcal{T}_{cal}$ is that orbit closure due to transformation laws (3.2).

To compute the degree of $\mathcal{T}_{cal}$, we use the open-source homotopy continuation software Bertini. We fix a random linear subspace $\mathcal{L} \subset \mathbb{P}(\mathbb{C}^{3 \times 3 \times 3})$ of complementary dimension to $\mathcal{T}_{cal}$, i.e. $\dim(\mathcal{L}) = 15$. This is expressed in floating-point as the vanishing of 11 random linear forms $\ell_m(T_{ijk}) = 0$ (3.3), where $m = 1; \ldots; 11$. Our goal is to compute $\#(\mathcal{T}_{cal} \cap \mathcal{L})$. As homotopy continuation calculations are sensitive to the formulation used, we carefully explain our own formulation to calculate $\mathcal{T}_{cal} \cap \mathcal{L}$. Our formulation starts with the parametrization of $\mathcal{T}_{cal}$ above, and with its two copies of $SO(3;\mathbb{C})$.

Recall that unit norm quaternions double-cover $SO(3;\mathbb{R})$. Complexifying:

$$R_2 = \begin{pmatrix} a^2 + b^2 - c^2 - d^2 & 2(bc - ad) & 2(bd + ac) \\ 2(bc + ad) & a^2 + c^2 - b^2 - d^2 & 2(cd - ab) \\ 2(bd - ac) & 2(cd + ab) & a^2 + d^2 - b^2 - c^2 \end{pmatrix}$$

where $a; b; c; d \in \mathbb{C}$ and $a^2 + b^2 + c^2 + d^2 = 1$ (3.4). Similarly for R_3 with $e; f; g; h \in \mathbb{C}$ subject to $e^2 + f^2 + g^2 + h^2 = 1$ (3.5). For our purposes, it is computationally advantageous to replace (3.4) by a random patch $\lambda_1 a + \lambda_2 b + \lambda_3 c + \lambda_4 d = 1$ (3.6), where $\lambda_i \in \mathbb{C}$ are random floating-point numbers fixed once and for all. Similarly, we replace (3.5) by a random patch $\mu_1 e + \mu_2 f + \mu_3 g + \mu_4 h = 1$ (3.7). The patches (3.6) and (3.7) leave us with injective parameterizations of two subvarieties of $\mathbb{C}^{3 \times 3}$, that we denote by $SO(3;\mathbb{C})_\lambda; SO(3;\mathbb{C})_\mu$. These two varieties have the same closed affine cone as the closed affine cone of $SO(3;\mathbb{C})$. This affine cone is:

$$\widehat{SO}(3;\mathbb{C}) := \{R \in \mathbb{C}^{3 \times 3} : \exists \lambda \in \mathbb{C} \text{ s.t. } RR^T = R^T R = \lambda I_{3 \times 3}\}$$

and it is parametrized by $a; b; c; d$ as above, but with no restriction on $a; b; c; d$. In the definition of the cone $\widehat{SO}(3;\mathbb{C})$, note $\lambda = 0$ is possible; it corresponds to

$a^2+b^2+c^2+d^2=0$, or to $e^2+f^2+g^2+h^2=0$. By the first remark after Definition 3.15, we are free to scale cameras B and C so that their left 3×3 submatrices satisfy $R_2 \in SO(3,\mathbb{C})$ and $R_3 \in SO(3,\mathbb{C})$, and for our formulation here we do so. Finally, for $\mathbb{C}^5$ in the source of the parametrization of T_{cal}, write $t_2 = \begin{pmatrix} t_{2,1} & t_{2,2} & 1 \end{pmatrix}^T$ and $t_3 = \begin{pmatrix} t_{3,1} & t_{3,2} & t_{3,3} \end{pmatrix}^T$.

At this point, we have replaced the dominant, generically injective map

$$SO(3,\mathbb{C})^{\times 2} \times \mathbb{C}^5 \dashrightarrow T_{cal}$$

by the dominant, generically injective parametrization $SO(3,\mathbb{C}) \times SO(3,\mathbb{C}) \times \mathbb{C}^5 \dashrightarrow T_{cal}$. Also, we have injective, dominant maps $V(\lambda_1 a + \lambda_2 b + \lambda_3 c + \lambda_4 d - 1) \to SO(3,\mathbb{C})$ and $V(\lambda_1 e + \lambda_2 f + \lambda_3 g + \lambda_4 h - 1) \to SO(3,\mathbb{C})$. Composing gives the generically 1-to-1, dominant $V(\lambda_1 a + \lambda_2 b + \lambda_3 c + \lambda_4 d - 1) \times V(\lambda_1 e + \lambda_2 f + \lambda_3 g + \lambda_4 h - 1) \times \mathbb{C}^5 \dashrightarrow T_{cal}$. With exactly this parametrization of T_{cal}, it will be most convenient to perform numerical algebraic geometry calculations. Hence, here to compute $\deg(T_{cal}) = \#(T_{cal} \cap L)$, we consider the square polynomial system:

in 13 variables $a, b, c, d, e, f, g, h, t_{2,1}, t_{2,2}, t_{3,1}, t_{3,2}, t_{3,3} \in \mathbb{C}$;

with 13 equations the 11 cubics (3.3) and 2 linear equations (3.6), (3.7).

The solution set equals the preimage of $T_{cal} \cap L$. This system is expected to have $\deg(T_{cal})$ many solutions. We can solve zero-dimensional square systems of this size (in floating-point) using the UseRegeneration:1 setting in Bertini. That employs the *regeneration* solving technique from [53]. For the present system, overall, Bertini tracks 74,667 paths in 1.5 hours on a standard laptop computer to find 4912 solutions. Numerical path-tracking in Bertini is based on a *predictor-corrector* approach. Prediction by default is done by the Runge-Kutta 4^{th} order method; correction is by Newton steps. For more information, see [11, Section 2.2]. Here, this provides strong numerical evidence for the conclusion that $\deg(T_{cal}) = 4912$. Up to the numerical accuracy of Bertini and the reliability of our random number generator used to choose L, this computation is correct with probability 1. Practically speaking, 4912 is correct only with very high probability.

As a check for 4912, we apply the *trace test* from [50], [71] and [93]. A random linear form ℓ' on $\mathbb{P}(\mathbb{C}^{3 \times 3 \times 3})$ is fixed. For $s \in \mathbb{C}$, we set $L_s := V(\ell_1 + s\ell', \ldots, \ell_{11} + s\ell')$, so $L_0 = L$. Varying $s \in \mathbb{C}$, the intersection $T_{cal} \cap L_s$ consists of $\deg(T_{cal})$ many complex paths. Let $T_s \subseteq T_{cal} \cap L_s$ be a subset of paths. Then the trace test implies (for generic ℓ', ℓ_i) that $T_s = T_{cal} \cap L_s$ if and only if the centroid of T_s computed in a consistent affine chart $\mathbb{C}^{26}$, i.e.

$$\mathrm{cen}(T_s) := \frac{1}{\# T_s} \sum_{p_s \in T_s} p_s;$$

is an affine linear function of s. Here, we set T_0 to be the 4912 intersection points found above. Then we calculate T_1 with the UserHomotopy:1 setting in Bertini, where the variables are $a, \ldots, t_{3;3}$, and the start points are the preimages of T_0. After this homotopy in parameter space, T_1 is obtained by evaluating the endpoints of the track via TrackType:-4. Similarly, T_{-1} is computed. Then we calculate that the following quantity in $\mathbb{C}^{26}$:

$$(\mathrm{cen}(T_1) - \mathrm{cen}(T_0)) - (\mathrm{cen}(T_0) - \mathrm{cen}(T_{-1}))$$

is indeed numerically 0. This trace test is a further verification of 4912. □

Remark 3.27. In the proof of Theorem 3.26, when we select one point per fiber per member of $T_{cal} \cap L$, we obtain a pseudo-witness set W for T_{cal}. This is the fundamental data structure in numerical algebraic geometry for computing with parameterized varieties (see [52]). Precisely, here it is the quadruple:

- the parameter space $P \subset \mathbb{C}^{13}$, where $\mathbb{C}^{13}$ has coordinates $a, \ldots, t_{3;3}$ and $P = V(\lambda_1 a + \lambda_2 b + \lambda_3 c + \lambda_4 d - 1, \ \lambda_1 e + \lambda_2 f + \lambda_3 g + \lambda_4 h - 1)$
- the dominant map $\psi : P \dashrightarrow T_{cal}$ in the proof of Theorem 3.26, e.g. $\psi_{1;1;1} = -2bct_{2;1} - 2adt_{2;1} + a^2 t_{2;2} + b^2 t_{2;2} - c^2 t_{2;2} - d^2 t_{2;2}$
- the generic complimentary linear space $L = V(\ell_1, \ldots, \ell_{11}) \subset \mathbb{P}(\mathbb{C}^{3 \times 3 \times 3})$
- the finite set $W \subset P \subset \mathbb{C}^{13}$, mapping bijectively to $T_{cal} \cap L$.

We heavily use this representation of T_{cal} for the computations in Section 3.6.

Now, we re-visit Proposition 3.18. When $T_{A;B;C}$ is unknown but the point/line correspondence is known, the constraints there amount to special linear slices of T and of the subvariety T_{cal}. The next theorem may help the reader appreciate the specialness of these linear sections of T_{cal}; in general, the intersections are not irreducible, equidimensional, nor dimensionally transverse.

Theorem 3.28. Fix generic points $x, x', x'' \in \mathbb{P}^2$ and generic lines $\ell, \ell', \ell'' \in (\mathbb{P}^2)^{\vee}$. In the cases of Proposition 3.18, we have the following codimensions:

- [PLL] : $L = \{T \in \mathbb{P}(\mathbb{C}^{3 \times 3 \times 3}) : T(x, \ell', \ell'') = 0\}$ is a hyperplane and $T_{cal} \cap L$ consists of one irreducible component of codimension 1 in T_{cal}

[LLL] : $L = \{T \in \mathbb{P}(\mathbb{C}^{3\times 3\times 3}) : [\ell]_\times T(\cdot, \ell', \ell'') = 0\}$ is a codimension 2 subspace and $T_{cal} \cap L$ consists of two irreducible components both of codimension 2 in T_{cal}

[PLP] : $L = \{T \in \mathbb{P}(\mathbb{C}^{3\times 3\times 3}) : [x'']_\times T(x, \ell', \cdot) = 0\}$ is a codimension 2 subspace and $T_{cal} \cap L$ consists of two irreducible components both of codimension 2 in T_{cal}

[PPL] : $L = \{T \in \mathbb{P}(\mathbb{C}^{3\times 3\times 3}) : [x']_\times T(x, \cdot, \ell'') = 0\}$ is a codimension 2 subspace and $T_{cal} \cap L$ consists of two irreducible components both of codimension 2 in T_{cal}

[PPP] : $L = \{T \in \mathbb{P}(\mathbb{C}^{3\times 3\times 3}) : [x'']_\times T(x, \cdot, \cdot)[x']_\times = 0\}$ is a codimension 4 subspace and $T_{cal} \cap L$ consists of five irreducible components, one of codimension 3 and four of codimension 4 in T_{cal}.

Computational Proof. The statements about the subspaces may shown symbolically. In the case of LLL, e.g., work in the ring $\mathbb{Q}[\ell_0, \ldots, \ell''_2]$ with 8 variables, and write the constraint on $T \in \mathbb{P}(\mathbb{C}^{3\times 3\times 3})$ as the vanishing of a 3×27 matrix times a vectorization of T. Now we check that all of the 3×3 minors of that long matrix are identically 0, but not so for 2×2 minors.

For the statements about $T_{cal} \cap L$, we offer a probability 1, numerical argument. By [92, Theorem A.14.10] and the discussion on page 348 about generic irreducible decompositions, we can fix random floating-point coordinates for $x, x', x'', \ell, \ell', \ell''$. With the parametrization of T_{cal} from the proof of Theorem 3.26, the TrackType:1 setting in Bertini is used to compute a numerical irreducible decomposition for the preimage of $T_{cal} \cap L$ per each case. That outputs a witness set, i.e. general linear section, per irreducible component. Bertini's TrackType:1 is based on regeneration, monodromy and the trace test; see [92, Chapter 15] or [11, Chapter 8] for a description.

Here, the PPP case is most subtle since the subspace $L \subseteq \mathbb{P}(\mathbb{C}^{3\times 3\times 3})$ is codimension 4, but the linear section $T_{cal} \cap L \subseteq T_{cal}$ includes a codimension 3 component. The numerical irreducible decomposition above consists of five components of dimensions $8, 7, 7, 7, 7$ in $a, \ldots, t_{3,3}$-parameter space. Thus, it suffices to verify that the map to T_{cal} is generically injective restricted to the union of these components. For that, we take one general point on each component from the witness sets, and test whether that point satisfies $a^2 + b^2 + c^2 + d^2 \neq 0$ and $e^2 + f^2 + g^2 + h^2 \neq 0$. This indeed holds for all components. Then, we test using singular value decomposition (see [26, Theorem 3.2]) whether the point maps to a camera triple with linearly independent centers. Linear independence indeed holds for all components. From Theorem 3.22, the above parametrization is generically injective on this locus. Hence the image $T_{cal} \cap L$ consists of distinct components with the same dimensions, 7, 8, 7, 7, 7. This finishes PPP. The other cases are similar. □

Mimicking the proof of Proposition 3.14, and using the 'top-dimensional' clause in Lemma 3.20, we can establish the following finiteness result for T_{cal}:

Lemma 3.29. For each problem in Theorem 3.6, given generic image correspondence data, there are only finitely many tensors $T \in T_{cal}$ that satisfy all of the linear conditions from Proposition 3.18.

We have arrived at a relaxation for each minimal problem in Theorem 3.6, as promised. Namely, for a problem there we can fix a random instance of image data, and we seek those calibrated trifocal tensors that satisfy the – merely necessary – linear conditions in 3.18. Geometrically, this is equivalent intersect the special linear sections of T_{cal} from Theorem 3.28. In Section 3.6, we will use the pseudo-witness set representation $(P; \ ; L; W)$ of T_{cal} from Theorem 3.26 to compute these special slices of T_{cal} in Bertini. Conveniently, Bertini outputs a calibrated camera triple per calibrated trifocal tensor in the intersection; this is because all solving is done in the parameter space P, or in other words, camera space. To solve the original minimal problem, we then test these configurations against the tight conditions of Theorem 3.10.

3.6 Proof of main result

In this section, we put all the pieces together and we determine the algebraic degrees of the minimal problems in Theorem 3.6. Mathematically, these degrees represent interesting enumerative geometry problems; in vision, related work for three uncalibrated views appeared in [84]. The authors considered correspondences PPP and LLL and they determined 3 degrees for projective (uncalibrated) views, using the larger group actions present in that case. Here, all 66 degrees for calibrated views in Theorem 3.6 are new.

Now, recall from Proposition 3.14 that solutions $(A; B; C)$ to the problems in Theorem 3.6 in particular must have non-identical centers. So, by the second remark after Definition 3.15, they associate to nonzero tensors $T_{A;B;C}$, and thus to well-defined points in the projective variety T_{cal}. Conversely, however, there are special subloci of T_{cal} that are not physical. Points in these subvarieties (introduced next) are extraneous to Theorem 3.6, because they correspond to configurations with a

3×4 matrix whose left 3×3 submatrix R is not a rotation, but instead satisfies $RR^T = R^T R = 0$.

Definition/Proposition 3.30. Recall the parametrization of T_{cal} by $a, \ldots, t_{3,3}$ from Theorem 3.26. Let $T_{cal}^{0,1} \subseteq T_{cal}$ be the closure of the image of the rational map restricted to the locus $a^2 + b^2 + c^2 + d^2 = 0$. Let $T_{cal}^{1,0} \subseteq T_{cal}$ be the closure of the image of the rational map restricted to the locus $e^2 + f^2 + g^2 + h^2 = 0$. Let $T_{cal}^{0,0} \subseteq T_{cal}$ be the closure of the image of the rational map restricted to the locus $a^2 + b^2 + c^2 + d^2 = 0$ and $e^2 + f^2 + g^2 + h^2 = 0$. Then these subvarieties are irreducible with: $\dim(T_{cal}^{0,0}) = 9$ and $\deg(T_{cal}^{0,0}) = 1296$; $\dim(T_{cal}^{0,1}) = 10$ and $\deg(T_{cal}^{0,1}) = 2616$; $\dim(T_{cal}^{1,0}) = 10$ and $\deg(T_{cal}^{1,0}) = 2616$.

Computational Proof. The restricted parameter spaces:

$$P \cap V(a^2 + b^2 + c^2 + d^2), \; P \cap V(e^2 + f^2 + g^2 + h^2),$$
$$P \cap V(a^2 + b^2 + c^2 + d^2, e^2 + f^2 + g^2 + h^2) \subseteq \mathbb{C}^{13},$$

where $P = V(\lambda_1 a + \lambda_2 b + \lambda_3 c + \lambda_4 d - 1, \mu_1 e + \mu_2 f + \mu_3 g + \mu_4 h - 1)$, are irreducible, therefore their images $T_{cal}^{0,1}, T_{cal}^{1,0}, T_{cal}^{0,0} \subseteq \mathbb{P}(\mathbb{C}^{3 \times 3 \times 3})$ are irreducible. The dimension statements are verified by picking a random point in the restricted parameter spaces, and then by computing the rank of the derivative of the restricted rational map Φ at that point. This rank equals the dimension of the image with probability 1, by generic smoothness over $\mathbb{C}$ [46, III.10.5] and the preceding [46, III.10.4]. For the degree statements, the approach from Theorem 3.26 may be used. For $T_{cal}^{0,1}$ we fix a random linear subspace $M \subseteq \mathbb{P}(\mathbb{C}^{3 \times 3 \times 3})$ of complementary dimension, i.e $\dim(M) = 16$, so $\deg(T_{cal}^{0,1}) = \#(T_{cal}^{0,1} \cap M)$. We pull back to $P \cap V(a^2 + b^2 + c^2 + d^2) \cap \Phi^{-1}(M)$, and use the UseRegeneration:1 setting in Bertini to solve for this. This run outputs 2616 floating-point tuples in $a, \ldots, t_{3,3}$ coordinates. Then, we apply the parametrization Φ and check that the image of these are 2616 numerically distinct tensors, i.e. the restriction $\Phi|_{P \cap V(a^2 + b^2 + c^2 + d^2)}$ is generically injective. It follows that $\deg(T_{cal}^{0,1}) = 2616$, up to numerical accuracy and random choices. To verify this degree further, we apply the trace test as in Theorem 3.26, and this finishes the computation for $\deg(T_{cal}^{0,1})$. Since $T_{cal}^{0,1}$ and $T_{cal}^{0,1}$ are linearly isomorphic under the permutation $T_{ijk} \mapsto T_{ikj}$, this implies $\deg(T_{cal}^{1,0}) = 2616$. The computation for $\deg(T_{cal}^{0,0})$ is similar. □

Now, we come to the proof of Theorem 3.6, at last. The outline was given in the last paragraph of Section 3.5: for computations, solving the polynomial systems of multiview equations (see Theorem 3.10) is relaxed to taking a special linear section of the calibrated trifocal variety T_{cal} (see Theorem 3.28). Then, to take this slice, we use the numerical algebraic geometry technique of coefficient-parameter homotopy

[92, Theorems 7.1.1, A.13.1], i.e. a general linear section is moved in a homotopy to the special linear section.

Computational Proof of Theorem 3.6. Let weights $(w_1, w_2, w_3, w_4, w_5) \in \mathbb{Z}^5_{\geq 0}$ satisfy $3w_1 + 2w_2 + 2w_3 + 2w_4 + w_5 = 11$ and $w_2 \geq w_3$. Now consider the problem 'w_1PPP + w_2PPL + w_3PLP + w_4LLL + w_5PLL' in Theorem 3.6. Fix one general instance of this problem, by taking image data with random floating-point coordinates. Each point/line image correspondence in this instance defines a special linear subspace of $\mathbb{P}(\mathbb{C}^{3\times 3\times 3})$, as in Theorem 3.28. The intersection of these is one subspace $L_{special}$ expressed in floating-point; using singular value decomposition, we verify that its codimension in $\mathbb{P}(\mathbb{C}^{3\times 3\times 3})$ is the expected $4w_1 + 2w_2 + 2w_3 + 2w_4 + w_5 = 11 + w_1$. By Proposition 3.18, $L_{special}$ represents necessary conditions for consistency, so we seek $T_{cal} \cap L_{special}$. If $w_1 > 0$, then this intersection is not dimensionally transverse by the PPP clause of Theorem 3.28. To deal with a square polynomial system, we fix a general linear space $L'_{special} \supseteq L_{special}$ of codimension 11 in $\mathbb{P}(\mathbb{C}^{3\times 3\times 3})$ and now seek $T_{cal} \cap L'_{special}$. This step is known as *randomization* [92, Section 13.5] in numerical algebraic geometry, and it is needed to apply the parameter homotopy result [92, Theorem 7.1.1].

The linear section $T_{cal} \cap L'_{special}$ is found numerically by a degeneration. In the proof of Theorem 3.26, we computed a pseudo-witness set for T_{cal}. This includes a general complimentary linear section $T_{cal} \cap L$, and the preimage $\pi^{-1}(T_{cal} \cap L)$ of $\deg(T_{cal}) = 4912$ points in $a, \ldots, t_{3,3}$ space. Writing $L = V(\ell_1, \ldots, \ell_{11})$ and $L'_{special} = V(\ell'_1, \ldots, \ell'_{11})$ for linear forms ℓ_i and ℓ'_i on $\mathbb{P}(\mathbb{C}^{3\times 3\times 3})$, consider the following homotopy function $H : \mathbb{C}^{13} \times \mathbb{R} \to \mathbb{C}^{13}$:

$$H(a, \ldots, t_{3,3}, s) := \begin{bmatrix} s \cdot \ell_1(\pi(a, \ldots, t_{3,3})) + (1-s) \cdot \ell'_1(\pi(a, \ldots, t_{3,3})) \\ \vdots \\ s \cdot \ell_{11}(\pi(a, \ldots, t_{3,3})) + (1-s) \cdot \ell'_{11}(\pi(a, \ldots, t_{3,3})) \\ \lambda_1 a + \lambda_2 b + \lambda_3 c + \lambda_4 d - 1 \\ \lambda_1 e + \lambda_2 f + \lambda_3 g + \lambda_4 h - 1 \end{bmatrix}.$$

Here $s \in \mathbb{R}$ is the path variable. As s moves from 1 to 0, H defines a family of square polynomial systems in the 13 variables $a, \ldots, t_{3,3}$. The start system $H(a, \ldots, t_{3,3}, 1) = 0$ has solution set $\pi^{-1}(T_{cal} \cap L)$ and the target system $H(a, \ldots, t_{3,3}, 0) = 0$ has solution set $\pi^{-1}(T_{cal} \cap L'_{special})$. With the UserHomotopy:1 setting in Bertini, we track the 4912 solution paths from the start to target system. By genericity of L in the start system, these solution paths are smooth [92, Theorem 7.1.1(4), Lemma 7.1.2]. The finite endpoints of this track consist of solutions to the target system. By the principle of coefficient-parameter homotopy [92, Theorem A.13.1], every isolated point in $\pi^{-1}(T_{cal} \cap L^0_{special})$ is an endpoint, with probability 1. Note that in general, coefficient-parameter homotopy – i.e., the tracking of solutions of a *general* instance of a parametric system of equations to solutions of a *special* instance – may be used to find all isolated solutions to *square* polynomial systems. Here, by Lemma 3.29, $T_{cal} \cap L_{special}$ is a scheme with finitely many points. By Bertini's theorem [92, Theorem 13.5.1(1)], $T_{cal} \cap L^0_{special}$ also consists of finitely many points, using genericity of $L^0_{special}$. On the other hand, by Proposition 3.14, all solutions (A, B, C) to the instance of the original minimal problem indexed by $w \in \mathbb{Z}^5_{\geq 0}$ have linearly independent centers in $\mathbb{P}^3$. Moreover, a configuration (A, B, C) with linearly independent centers is an isolated point in $\pi^{-1}(T_{A,B,C})$, thanks to Theorem 3.22. Therefore, it follows that all solutions to the problem from Theorem 3.6 are among the isolated points in $\pi^{-1}(T_{cal} \cap L^0_{special})$, and so the endpoints of the above homotopy.

For each minimal problem in Theorem 3.6, after the above homotopy, Bertini returns 4912 finite endpoints in $a, \ldots, t_{3,3}$ space. We pick out which of these endpoints are solutions to the original minimal problem by performing a sequence of checks, as explained next. First of all, of these endpoints, let us keep only those that lie in $\pi^{-1}(T_{cal} \cap L_{special})$, as opposed to those that lie just in the squared-up target solution set $\pi^{-1}(T_{cal} \cap L^0_{special})$. Second, we remove points that satisfy $a^2 + b^2 + c^2 + d^2 \approx 0$ or $e^2 + f^2 + g^2 + h^2 \approx 0$, because they are non-physical (see Definition/Proposition 3.30). Third, we verify that, in fact, all remaining points correspond to camera configurations (A, B, C) with linearly independent centers. This means that the equations in Theorem 3.10 generate the multiview ideals (recall Definition 3.9). Fourth, we check which remaining points satisfy those tight multiview equations. To test this robustly in floating-point, note that the equations in Theorem 3.10 are equivalent to rank drops of the concatenated matrices there, hence we test for those rank drops using singular value decomposition. If the ratio of two consecutive singular values exceeds 10^5, then this is taken as an indication that all singular values below are numerically 0, thus the matrix drops rank. Fifth, and conversely, we verify that all remaining configurations (A, B, C) avoid epipoles (recall Definition 3.12) for the fixed random instance of image correspondence data, so the converse Lemma 3.13 applies to prove consistency. Lastly, we verify that all solutions are numerically distinct. Ultimately, the output of this procedure is a list of all calibrated camera configurations over $\mathbb{C}$ that are solutions to the fixed random instances of the minimal problems, where these solutions are expressed in floating-point and $a, \ldots, t_{3,3}$ coordinates. The numbers of solutions are the algebraic degrees from Theorem 3.6.

As a check for this numerical computation, we repeat the entire calculation for other random instances of correspondence data. For each minimal problem, we obtain the same algebraic degree each time. One instance per problem solved to high precision is provided on this chapter's webpage. □

Example 3.31. We illustrate the proof of Theorem 3.6 by returning to the instance of `1 PPP + 4 PPL ' in Example 3.5. Here $L_{special}$ $\mathbb{P}(\mathbb{C}^{3\ 3\ 3})$ formed by intersecting subspaces from Theorem 3.28 is codimension 12, hence $L^0_{special}$! $L_{special}$. Tracking deg(T_{cal}) many points in the pseudo-witness set ${}^1(T_{cal} \cap L\ )$ to the target ${}^1(T_{cal} \cap L^0_{special})$, we get 4912 nite endpoints. Testing membership in $L_{special}$, we get 2552 points in ${}^1(T_{cal} \cap L_{special})$. Among these, 888 points satisfy $a^2 + b^2 + c^2 + d^2$ 0, so they are non-physical (corresponding to 3 4 matrices with left submatrices that are not rotations). The remaining 1664 points turn out to correspond to calibrated camera con gurations with linearly independent centers. Checking satisfaction of the equations from Theorem 3.10, we end up with 160 solutions.

Remark 3.32. The proof of Theorem 3.6 is constructive. From the solved random instances, one may build solvers for each minimal problem, using coe cient-parameter homotopy. Here the start system is the solved instance of the minimal problem and the target system is another given instance. Such a solver is optimal in the sense that the number of paths tracked equals the true algebraic degree of the problem. Implementation is left to future work.

Remark 3.33. All degrees in Theorem 3.6 are divisible by 8. We would like to understand why. What are the Galois groups [51] for these minimal problems?

Remark 3.34. Practically speaking, given image correspondence data de ned over $\mathbb{R}$, only real solutions $(A; B; C)$ to the minimal problems in Theorem 3.6 are of interest to RANSAC-style 3D reconstruction algorithms. Does there exist image data such that all solutions are real? Also, for the image data observed in practice, what is the distribution of the number of real solutions?

3.7 Numerical implicitization

In this section, we switch gears from calibrated three-view geometry, and describe a stand-alone Macaulay2 software package [21] co-written with Justin Chen, for wide use in computational algebra. Our software NumericalImplicitization permits the user-friendly computation of invariants of the image of a polynomial map, such as dimension, degree and Hilbert function values. Like the computations performed already in this chapter, NumericalImplicitization relies on methods from numerical algebraic geometry, e.g. homotopy-continuation and monodromy. My own interest in writing general-purpose numerical algebraic geometry code grew out of my project on the calibrated trifocal variety.

Many varieties of interest in algebraic geometry and its applications are usefully described as images of polynomial maps, i.e. via a parametrization. For vision examples, see the third sentence of Section 2.1, Definition 3.23, Definition 3.24, Equation 4.10 and Equation 5.5. Implicitization is the process of converting a parametric description of a variety into an intrinsic – or implicit – description. Classically, implicitization refers to the procedure of computing the defining equations of a parametrized variety, and in theory this is accomplished by finding the kernel of a ring homomorphism, via Gröbner bases. In practice however, symbolic Gröbner basis computations are often time-consuming, even for medium-scale problems, and do not scale well with respect to the size of the input.

Despite this, one would often like to know basic information about a parametrized variety, even when symbolic methods are prohibitively expensive (in terms of computation time). The best examples of such information are discrete invariants such as the dimension, or degree and Hilbert function values if the variety is projective. Other examples include Boolean tests, e.g. whether or not a particular point lies on a parametrized variety. The goal of the present Macaulay2 [44] package is to provide such information – in other words, to *numerically implicitize* a parametrized variety – by using the methods of numerical algebraic geometry. NumericalImplicitization[3] builds on top of existing numerical algebraic geometry software, e.g. NAG4M2 [70], Bertini [10, 9] and PHCpack [102, 45]. Each of these can be used for path tracking and point sampling; by default, the native engine NAG4M2 is used.

Notation. The following notation will be used throughout the remainder of this section:

- $X \subseteq \mathbb{A}^n$ is a source variety, defined by an ideal $I = \langle g_1, \ldots, g_r \rangle$ in the polynomial ring $\mathbb{C}[x_1, \ldots, x_n]$
- $F = \{f_1, \ldots, f_m\}$, where $f_i \in \mathbb{C}[x_1, \ldots, x_n]$, is a list of polynomials specifying a map $\mathbb{A}^n \to \mathbb{A}^m$
- Y is the Zariski closure of the image $\overline{F(X)} = \overline{F(V(I))} \subseteq \mathbb{A}^m$; the target variety under consideration

[3]For up-to-date code and documentation, see https://github.com/Joe-Kileel/Numerical-Implicitization

- $\overline{Y} \subseteq \mathbb{P}^m$ is the projective closure of Y, with respect to the standard embedding $\mathbb{A}^m \subseteq \mathbb{P}^m$.

Currently, our code NumericalImplicitization is implemented for integral (i.e. reduced and irreducible) varieties X. Equivalently, the ideal I is prime. Since numerical methods are used, we always work over the complex numbers with floating-point arithmetic. Moreover, $\overline{Y}$ is internally represented by its affine cone. This is because it is easier for computers to work with points in affine space; at the same time, this suffices to find the invariants of $\overline{Y}$.

All the methods in this package rely crucially on the ability to sample general points on X. To this end, two methods are provided, numericalSourceSample and numericalImageSample, which allow the user to sample as many general points on X and Y as desired. numericalSourceSample will compute a witness set of X, unless $X = \mathbb{A}^n$, by taking a numerical irreducible decomposition of X. This time-consuming step cannot be avoided. Once a witness set is known, points on X can be sampled in negligible time. numericalImageSample works by sampling points in X via numericalSourceSample, and then applying the map F.

One way to view the difference in computation time between symbolic and numerical methods is that the upfront cost of computing a Gröbner basis is replaced with the upfront cost of computing a numerical irreducible decomposition, which is used to sample general points. However, if $X = \mathbb{A}^n$, then sampling is done by generating random tuples, and is essentially immediate. Thus, in this unrestricted parametrization case, the upfront cost of numerical methods becomes zero.

The most basic invariant of an algebraic variety is its dimension. To compute the dimension of the image of a variety numerically, we use the following theorem:

Theorem 3.35. *Let $F : X \to Y$ be a dominant morphism of irreducible varieties over $\mathbb{C}$. Then there is a Zariski open subset $U \subseteq X$ such that for all $x \in U$, the induced map on tangent spaces $dF_x : T_xX \to T_{F(x)}Y$ is surjective.*

Proof. This is an immediate corollary of *generic smoothness* [46, III.10.5] and the preceding [49, III.10.4]. □

In the setting above, since the singular locus Sing Y is a proper closed subset of Y, for general $y = F(x) \in Y$ we have that $\dim Y = \dim T_yY = \dim dF_x(T_xX) = \dim T_xX - \dim \ker dF_x$. Now T_xX is the kernel of the Jacobian matrix of I evaluated at x, given by $\mathrm{Jac}(I)(x) = ((\partial g_i/\partial x_j)(x))_{1 \le i \le r,\, 1 \le j \le n}$, and $\ker dF_x$ is the kernel of the Jacobian of F evaluated at x, intersected with T_xX. Explicitly, $\ker dF_x$ is the kernel of the $(r+m) \times n$ matrix:

$$\begin{bmatrix} \mathrm{Jac}(I)(x) \\ \mathrm{Jac}(F)(x) \end{bmatrix} = \begin{bmatrix} \frac{\partial g_1}{\partial x_1}(x) & \cdots & \frac{\partial g_1}{\partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial g_s}{\partial x_1}(x) & \cdots & \frac{\partial g_s}{\partial x_n}(x) \\ \frac{\partial F_1}{\partial x_1}(x) & \cdots & \frac{\partial F_1}{\partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial x_1}(x) & \cdots & \frac{\partial F_m}{\partial x_n}(x) \end{bmatrix}$$

We compute these kernel dimensions numerically, as explained prior to Chapter 3.38 below, to get $\dim Y$.

Example 3.36. Let $Y \subseteq \mathrm{Sym}^4(\mathbb{C}^5) = \mathbb{A}^{70}$ be the variety of $5 \times 5 \times 5 \times 5$ symmetric tensors of border rank ≤ 14. Equivalently, Y is the affine cone over $\sigma_{14}(\nu_4(\mathbb{P}^4))$, the 14^{th} secant variety of the fourth Veronese embedding of $\mathbb{P}^4$. Naively, one expects $\dim(Y) = 14 \cdot 4 + 13 + 1 = 70$. In fact, $\dim(Y) = 69$ as verified by the following code:

```
Macaulay2, version 1.9.2
i1 : needsPackage "NumericalImplicitization"
i2 : R = CC[s_(1,1)..s_(14,5)];
i3 : F = sum(1..14, i -> flatten entries basis(4, R, Variables =>
toList(s_(i,1)..s_(i,5))));
i4 : time numericalImageDim(F, ideal 0_R)
     -- used 0.106554 seconds
o4 = 69
```

This example is the largest exceptional case from the celebrated work [6]. Note the timing printed above.

We now turn to the problem of determining the Hilbert function of $\mathcal{Y}$. Recall that if $\mathcal{Y} \subseteq \mathbb{P}^m$ is a projective variety, given by a homogeneous ideal $J \subseteq \mathbb{C}[y_0, \ldots, y_m]$, then the Hilbert function of $\mathcal{Y}$ at an argument $d \in \mathbb{N}$ is by definition the vector space dimension of the d^{th} graded part of J, i.e. $H_{\mathcal{Y}}(d) := \dim J_d$. This counts the maximum number of linearly independent degree d hypersurfaces in $\mathbb{P}^m$ containing $\mathcal{Y}$.

To compute the Hilbert function of $\mathcal{Y}$ numerically, we use multivariate polynomial interpolation. For a fixed argument $d \in \mathbb{N}$, let $\{p_1, \ldots, p_N\}$ be a set of N general points on $\mathcal{V}$. For $1 \leq i \leq N$, consider an $i \times \binom{m+d}{d}$ interpolation matrix $A^{(i)}$ with rows indexed by points $\{p_1, \ldots, p_i\}$ and columns indexed by degree d monomials in $\mathbb{C}[y_0, \ldots, y_m]$, whose entries are the values of the monomials at the points. A vector in the kernel of $A^{(i)}$ corresponds to a choice of coefficients for a homogeneous degree d polynomial that vanishes on $p_1, \ldots, p_i$. If i is large, then one expects such a form to vanish on the entire variety $\mathcal{V}$. The following theorem makes this precise:

Theorem 3.37. *Let $\{p_1, \ldots, p_{s+1}\}$ be a set of general points on $\mathcal{V}$, and let $A^{(i)}$ be the interpolation matrix above. If $\dim \ker A^{(s)} = \dim \ker A^{(s+1)}$, then $\dim \ker A^{(s)} = \dim J_d$.*

Proof. Identifying a vector $v \in \ker A^{(i)}$ with the form in $\mathbb{C}[y_0, \ldots, y_m]$ of degree d having v as its coefficients, it suffices to show that $\ker A^{(s)} = J_d$. If $h \in J_d$, then h vanishes on all of $\mathcal{V}$, in particular on $\{p_1, \ldots, p_s\}$, so $h \in \ker A^{(s)}$. For the converse $\ker A^{(s)} \subseteq J_d$, we consider the universal interpolation matrices over $\mathbb{C}[y_{0,1}, y_{1,1}, \ldots, y_{m,i}]$

$$A^{(i)} := \begin{bmatrix} y_{0,1}^d & y_{0,1}^{d-1} y_{1,1} & \cdots & y_{m,1}^d \\ y_{0,2}^d & y_{0,2}^{d-1} y_{1,2} & \cdots & y_{m,2}^d \\ \vdots & \vdots & \ddots & \vdots \\ y_{0,i}^d & y_{0,i}^{d-1} y_{1,i} & \cdots & y_{m,i}^d \end{bmatrix}$$

Set $r_i := \min\{j \in \mathbb{Z}_{\geq 0} \mid$ every $(j+1) \times (j+1)$ minor of $A^{(i)}$ lies in the ideal of $\mathcal{V}^i \subseteq (\mathbb{P}^m)^i\}$. Then any specialization of $A^{(i)}$ to i points in $\mathcal{V}$ is a matrix over $\mathbb{C}$ of rank $\leq r_i$; moreover if the points are general, then the specialization has rank exactly r_i, since $\mathcal{V}$ is irreducible. In particular $\operatorname{rank}(A^s) = r_s$ and $\operatorname{rank}(A^{s+1}) = r_{s+1}$, so $\dim \ker A^{(s)} = \dim \ker A^{(s+1)}$ implies that $r_s = r_{s+1}$. It follows that specializing $A^{(s+1)}$ to $p_1, p_2, \ldots, p_s, q$ for any $q \in \mathcal{V}$ gives a rank r_s matrix. Hence, every degree d form in $\ker A^{(s)}$ evaluates to 0 at every $q \in \mathcal{V}$. Since $\mathcal{V}$ is reduced, we deduce that $\ker A^{(s)} \subseteq J_d$. □

It follows from Chapter 3.37 that the integers $\dim \ker A^{(1)}, \dim \ker A^{(2)}, \ldots$ decrease by exactly 1, until the first instance where they fail to decrease, at which point they stabilize: $\dim \ker A^{(i)} = \dim \ker A^{(s)}$ for $i \geq s$. This stable value is the value of the Hilbert function, $\dim \ker A^{(s)} = H_{\mathcal{V}}(d)$. In particular, it suffices to compute $\dim \ker A^{(N)}$ for $N = \binom{m+d}{d}$, i.e. one may assume the interpolation matrix is square. Although this may seem wasteful (as stabilization may have occurred with fewer rows), this is indeed what numericalHilbertFunction does, due to the algorithm used to compute kernel dimension numerically. To be precise, kernel dimension is found via a singular value decomposition (SVD) – namely, if a gap (= ratio of consecutive singular values) greater than the option SVDGapThreshold (with default value 200) is observed in the list of all singular values, then this is taken as an indication that all singular values past the greatest gap are numerically zero. On example problems, it was observed that taking only one more additional row than was needed often did not reveal a satisfactory gap in singular values. In addition, numerical stability is improved via preconditioning on the interpolation matrices – namely, each row is normalized in the Euclidean norm before computing the SVD.

Example 3.38. Let X be a random canonical curve of genus 4 in $\mathbb{P}^3$, so X is the complete intersection of a random quadric and cubic. Let $F : \mathbb{P}^3 \dashrightarrow \mathbb{P}^2$ be a projection by 3 random cubics. Then Y is a plane curve of degree $3^{\dim(Y)} \cdot \deg(X) = 3 \cdot 2 \cdot 3 = 18$, so the ideal of Y contains a single form of degree 18. We verify this as follows:

```
i5 : R = CC[w_0..w_3];
i6 : I = ideal(random(2,R), random(3,R));
i7 : F = toList(1..3)/(i -> random(3,R));
i8 : T = numericalHilbertFunction(F, I, 18)
Sampling image points ...
     -- used 4.76401 seconds
Creating interpolation matrix ...
     -- used 0.313925 seconds
Performing normalization preconditioning ...
     -- used 0.214475 seconds
Computing numerical kernel ...
     -- used 0.135864 seconds
Hilbert function value: 1
o8 = NumericalInterpolationTable
```

The output is a NumericalInterpolationTable, which is a HashTable storing the results of the interpolation computation described above. From this, one can obtain a floating-point approximation to a basis of J_d. This is done via the command extractImageEquations:

```
i9 : extractImageEquations T
o9 : | -.0000712719y_0^18+(.000317507-.000100639i)y_0^17y_1-(.0000906039-
```

```
-----------------------------------------------------------------------
.000616564i)y_0^16y_1^2-(.00197404+.00177936i)y_0^15y_1^3+(.0046344+
-----------------------------------------------------------------------
.00196825i)y_0^14y_1^4-(.00475536-.00157142i)y_0^13y_1^5+(.00550602-
-----------------------------------------------------------------------
.0100492i)y_0^12y_1^6-(.012252-.0188461i)y_0^11y_1^7+ ... |
```

An experimental feature to find equations over $\mathbb{Z}$ may be called with the option `attemptExact => true`.

After dimension, degree is the most basic invariant of a projective variety $Y \subseteq \mathbb{P}^m$. Set $k := \dim(Y)$. For a general linear space $L \in \mathrm{Gr}(\mathbb{P}^{m-k}, \mathbb{P}^m)$ of complementary dimension to Y, the intersection $L \cap Y$ is a finite set of reduced points. The degree of Y is by definition the cardinality of $L \cap Y$, which is independent of the general linear space L. Thus one approach to find $\deg(Y)$ is to fix a random L_0 and compute the set of points $L_0 \cap Y$.

`NumericalImplicitization` takes this tack, but the method used to find $L_0 \cap Y$ is not the most obvious. First and foremost, we do not know the equations of Y, so all solving must be done in X. Secondly, we do not compute $F^{-1}(L_0) \cap X$ from the equations of X and the equations of L_0 pulled back under F, because that has degree $\deg(F) \cdot \deg(Y)$ – potentially much bigger than $\deg(Y)$. Instead, monodromy is employed to find $L_0 \cap Y$.

To state the technique, we consider the map:

$$\Gamma := \{(L, y) \in \mathrm{Gr}(\mathbb{P}^{m-k}, \mathbb{P}^m) \times Y \mid y \in L\} \subseteq \mathrm{Gr}(\mathbb{P}^{m-k}, \mathbb{P}^m) \times Y \xrightarrow{\pi_1} \mathrm{Gr}(\mathbb{P}^{m-k}, \mathbb{P}^m)$$

where π_1 is projection onto the first factor. There is a nonempty Zariski open subset $U \subseteq \mathrm{Gr}(\mathbb{P}^{m-k}, \mathbb{P}^m)$ such that the restriction $\pi_1^{-1}(U) \to U$ is a $\deg(Y)$-to-1 covering map, namely U equals the complement of the Hurwitz divisor from [99]. Now fix a generic basepoint $L_0 \in U$. Then the fundamental group $\pi_1(U; L_0)$ acts on the fiber $\pi_1^{-1}(L_0) = L_0 \cap Y$. This action is known as monodromy. It is a key fact that the induced group homomorphism $\pi_1(U; L_0) \to \mathrm{Sym}(L_0 \cap Y) = \mathrm{Sym}_{\deg(Y)}$ is surjective, by irreducibility of Y. More explicitly:

Theorem 3.39. Let Y, U, L_0 be as above. Write $L_0 = V(\ell_0)$ for $\ell_0 \in (\mathbb{C}[y_0, \ldots, y_m]_1)^k$ a height k column vector of linear forms. Fix another generic point $L_1 = V(\ell_1) \in U$, where $\ell_1 \in (\mathbb{C}[y_0, \ldots, y_m]_1)^k$. For any $\gamma_0, \gamma_1 \in \mathbb{C}$, consider the following loop of linear subspaces of $\mathbb{P}^m$:

$$t \mapsto \begin{cases} V\left(\gamma_0 (1-2t)\ell_0 + \gamma_1 2t \ell_1\right) & \text{if } 0 \le t \le \frac{1}{2} \\ V\left(\gamma_1 (2-2t)\ell_1 + \gamma_0 (2t-1)\ell_0\right) & \text{if } \frac{1}{2} \le t \le 1. \end{cases}$$

For a nonempty Zariski open subset of $(\gamma_0, \gamma_1) \in \mathbb{C}^2$, this loop is contained in U. Moreover, the classes of these loops in $\pi_1(U, L_0)$ generate the full symmetric group $\mathrm{Sym}(L_0 \cap \Gamma)$.

Proof. Let $\mathcal{L}$ be the pencil of linear subspaces of $\mathbb{P}^m$ generated by ℓ_0 and ℓ_1. Via monodromy, $\pi_1(\mathcal{L} \cap U, L_0)$ maps surjectively onto $\mathrm{Sym}(L_0 \cap \Gamma)$, by [93, Corollary 3.5]. Here the topological space $\mathcal{L} \cap U$ is homeomorphic to the Riemann sphere $\mathbb{CP}^1$ minus a finite set of points, so $\pi_1(\mathcal{L} \cap U, L_0)$ is isomorphic to a free group on finitely many letters. The explicit loops in the theorem statement miss the finite set $\mathcal{L} \setminus (\mathcal{L} \cap U)$ for general γ_0, γ_1; moreover γ_0, γ_1 may be chosen so that the loop above encloses exactly one point in $\mathcal{L} \setminus (\mathcal{L} \cap U)$. Therefore, the classes of these loops generate $\pi_1(\mathcal{L} \cap U, L_0)$. To visualize these loops, the reader may consult the proof of [92, Lemma 7.1.3]. □

`numericalImageDegree` works by first sampling a general point $x \in X$, and manufacturing a general linear slice L_0 such that $F(x) \in L_0 \cap \Gamma$. Then, L_0 is moved around in a loop of the form described in Theorem 3.39. This loop pulls back to a homotopy in X, where we use the equations of X to track x. The endpoint of the track is a point $x' \in X$ such that $F(x') \in L_0 \cap \Gamma$. If $F(x)$ and $F(x')$ are numerically distinct, then the loop has *learned* a new point in $L_0 \cap \Gamma$; otherwise x' is discarded. We then repeat this process of tracking points in X over each known point in $L_0 \cap \Gamma$, according to loops in Theorem 3.39. Note that for random $\gamma_0, \gamma_1 \in \mathbb{C}$, each loop has a positive probability – bounded away from 0 – of learning new points in $L_0 \cap \Gamma$, up until all of $L_0 \cap \Gamma$ is known. Thus by carrying out many loops from Theorem 3.39, the probability of finding all points in $L_0 \cap \Gamma$ approaches 1. In practice, if several consecutive loops[4] do not learn new points in $L_0 \cap \Gamma$, then we suspect that all of $L_0 \cap \Gamma$ has been calculated. To verify this, we pass to the *trace test* (see [93, Corollary 2.2], [50, §5] or [71, §1]), which provides a characterization for when a subset of $L_0 \cap \Gamma$ equals $L_0 \cap \Gamma$. If the trace test is failed, then L_0 is replaced by a new random L_0' and preimages in X of known points of $L_0 \cap \Gamma$ are tracked to those preimages of points of $L_0' \cap \Gamma$. Afterwards, monodromy for $L_0' \cap \Gamma$ begins anew. If the trace test is failed `maxTraceTests` (= 10 by default) times in total, then `numericalImageDegree` exits with only a lower bound on $\deg(\Gamma)$.

[4] This is specified by the option `maxRepetitiveMonodromies` (with default value 4).

Example 3.40. Let $\mathcal{Y} = \sigma_2(\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1) \subset \mathbb{P}^{31}$. We find that $\deg(\mathcal{Y}) = 3256$, using the commands below:

```
i10 : R = CC[a_1..a_5, b_1..b_5, t_0, t_1];
i11 : F1 = terms product(apply(toList(1..5), i -> 1 + a_i));
i12 : F2 = terms product(apply(toList(1..5), i -> 1 + b_i));
i13 : F = apply(toList(0..<2^5), i -> t_0*F1#i + t_1*F2#i);
i14 : time numericalImageDegree(F, ideal 0_R, maxRepetitiveMonodromies=>2)
Sampling point in source ...
Tracking monodromy loops ...
Points found: 2
Points found: 4
Points found: 8
Points found: 16
Points found: 32
Points found: 62
Points found: 123
Points found: 239
Points found: 466
Points found: 860
Points found: 1492
Points found: 2314
Points found: 3007
Points found: 3229
Points found: 3256
Points found: 3256
Points found: 3256
Running trace test ...
Degree of image: 3256
     -- used 388.989 seconds
o14 = PseudoWitnessSet
```

In [88, Theorem 4.1], it is proven via representation theory and combinatorics that the prime ideal J of $\mathcal{Y}$ is generated by the 3×3 minors of all flattenings of $2^{\times 5}$ tensors, so we can confirm that $\deg(J) = 3256$. However, the naive attempt to compute the degree of $\mathcal{Y}$ symbolically by taking the kernel of a ring map – from a polynomial ring in 32 variables – has no hope of finishing in any reasonable amount of time.

The output o14 above is a PseudoWitnessSet, which is a Macaulay2 HashTable that stores the computation of $L_0 \cap \mathcal{Y}$. This numerical representation of parameterized varieties was introduced in [52].

Classically, given a variety $Y \subseteq \mathbb{A}^m$ and a point $y \in \mathbb{A}^m$, we determine whether or not $y \in Y$ by finding set-theoretic equations of Y (which generate the ideal of Y up to radical), and then testing if y satisfies these equations. If a PseudoWitnessSet for Y is available, then point membership in Y can instead be verified by parameter homotopy. More precisely, isOnImage determines if y lies in the constructible set $F(X) \subseteq Y$, as follows. We fix a general affine linear subspace $L_y \subseteq \mathbb{A}^m$ of complementary dimension $m - k$ passing through y. Then $y \in F(X)$ if and only if $y \in L_y \cap F(X)$, so it suffices to compute the set $L_y \cap F(X)$. Now, a PseudoWitnessSet for Y provides a general section $L \cap F(X)$, and preimages in X. We move L to L_y as in [92, Theorem 7.1.6]. This pulls back to a homotopy in X, where we use the equations of X to track those preimages. Applying F to the endpoints of the track gives all isolated points in $L_y \cap F(X)$ by [92, Theorem 7.1.6]. Since L_y was general, the proof of [31, Corollary 10.5] shows $L_y \cap F(X)$ is zero-dimensional, so this procedure computes the entire set $L_y \cap F(X)$.

Example 3.41. Let $Y \subseteq \mathbb{A}^{18}$ be defined by the resultant of three quadratic equations in three unknowns, i.e., Y consists of all $(c_1, \ldots, c_6, d_1, \ldots, d_6, e_1, \ldots, e_6) \in \mathbb{A}^{18}$ such that the system

$$\begin{aligned}
0 &= c_1x^2 + c_2xy + c_3xz + c_4y^2 + c_5yz + c_6z^2\\
0 &= d_1x^2 + d_2xy + d_3xz + d_4y^2 + d_5yz + d_6z^2\\
0 &= e_1x^2 + e_2xy + e_3xz + e_4y^2 + e_5yz + e_6z^2
\end{aligned}$$

admits a solution $(x : y : z) \in \mathbb{P}^2$. Here Y is a hypersurface, and a matrix formula for its defining equation was derived in [34], using Ulrich sheaf and exterior algebra methods, similarly to own approach in Chapter 2 above. Here, we can rapidly determine point membership in Y numerically as follows.

```
i15 : R = CC[c_1..c_6, d_1..d_6, e_1..e_6, x, y, z];
i16 : I = ideal(c_1*x^2+c_2*x*y+c_3*x*z+c_4*y^2+c_5*y*z+c_6*z^2,
                d_1*x^2+d_2*x*y+d_3*x*z+d_4*y^2+d_5*y*z+d_6*z^2,
                e_1*x^2+e_2*x*y+e_3*x*z+e_4*y^2+e_5*y*z+e_6*z^2);
i17 : F = toList(c_1..c_6 | d_1..d_6 | e_1..e_6);
i18 : W = numericalImageDegree(F, I, verboseOutput => false); -- Y has degree 12
i19 : p1 = numericalImageSample(F, I); p2 = point random(CC^1, CC^#F);
i21 : time (isOnImage(W, p1), isOnImage(W, p2))
     -- used 0.186637 seconds
o21 = (true, false)
```

In this chapter, we determined algebraic degrees for minimal problems in the recovery of three calibrated cameras. This recovery has resisted e orts from the vision community; our results quantify the complexity. Numerical algebraic geometry furnished a powerful toolkit. Additionally, we relaxed zero-dimensional polynomial systems to systems with more geometric structure, hence easier to solve. In the last section, a software package for numerical implicitization was presented.

Chapter 4

Image Distortion

This chapter develops an algebro-geometric framework for dealing with image distortion. To that end, we introduce a general construction for lifting varieties in projective space to other toric varieties. We prove exact formulas for degree and defining equations, and we draw a connection with tropical geometry. These results unify and extend an existing body of work in computer vision. Our formulations lead to minimal solvers that competitive with or superior to the state of the art. The chapter is mostly based on my work [62] joint with Zuzana Kukelova, Tomas Pajdla and Bernd Sturmfels accepted for journal publication in *Foundations of Computational Mathematics*. In addition, the last section led to our subsequent paper [67], accepted for presentation at the *2017 IEEE Conference on Computer Vision and Pattern Recognition* in Honolulu, Hawaii.

4.1 Introduction

This chapter introduces a construction in algebraic geometry that is motivated by multiview geometry in computer vision. As we have seen in Chapters 2 and 3, in that field, one thinks of a camera as a linear projection $\mathbb{P}^3 \dashrightarrow \mathbb{P}^2$, and a model is a projective variety $X \subseteq \mathbb{P}^n$ that represents the relative positions of two or more such cameras. The data are correspondences of image points in $\mathbb{P}^2$ (or, in the case of three or more cameras, image lines in $(\mathbb{P}^2)^\vee$). These correspondences define a linear subspace $L \subseteq \mathbb{P}^n$, and the task is to compute the real points in the intersection $L \cap X$ as fast and accurately as possible. That kind of formulation already played prominently in Chapters 2 and 3 above. See [48, Chapter 9] for a textbook introduction.

A model for cameras with image distortion allows for an additional unknown parameter λ. Each coordinate of X gets multiplied by a polynomial in λ whose coefficients also depend on the data. We seek to estimate both λ and the point in X, where the data now specify a subspace L' in a larger projective space $\mathbb{P}^N$. The distortion variety X' lives in that $\mathbb{P}^N$, it satisfies $\dim(X') = \dim(X) + 1$, and the task is to compute $L' \cap X'$ in $\mathbb{P}^N$ fast and accurately.

We illustrate the idea of distortion varieties for the basic scenario in two-view geometry.

Example 4.1. The relative position of two uncalibrated cameras is expressed by a 3×3-matrix $x = (x_{ij})$ of rank 2, known as the *fundamental matrix*. Let $n = 8$ and write F for the hypersurface in $\mathbb{P}^8$ defined by the 3×3-determinant. Seven (generic) image correspondences in two views determine a line L in $\mathbb{P}^8$, and one rapidly computes the three points in $L \cap F$.

The 8-point radial distortion problem [64, Section 7.1.3] is modeled as follows in our setting. We duplicate the coordinates in the last row and last column of x, and we set

$$(x_{11} : x_{12} : x_{13} : y_{13} : x_{21} : x_{22} : x_{23} : y_{23} : x_{31} : y_{31} : x_{32} : y_{32} : x_{33} : y_{33} : z_{33}) =$$
$$\big(x_{11} : x_{12} : x_{13} : \lambda x_{13} : x_{21} : x_{22} : x_{23} : \lambda x_{23} : x_{31} : \lambda x_{31} : x_{32} : \lambda x_{32} : x_{33} : \lambda x_{33} : \lambda^2 x_{33}\big). \tag{4.1}$$

Here $N = 14$. The distortion variety F' is the closure of the set of matrices (4.1) where $x \in F$ and $\lambda \in \mathbb{C}$. The variety F' has dimension 8 and degree 16 in $\mathbb{P}^{14}$, whereas F has dimension 7 and degree 3 in $\mathbb{P}^8$. To estimate both λ and the relative camera positions, we now need eight image correspondences. These data specify a linear space L' of dimension 6 in $\mathbb{P}^{14}$. The task in the computer vision application is to rapidly compute the 16 points in $L' \cap F'$.

The prime ideal of the distortion variety F' is minimally generated by 18 polynomials in the 15 variables. First, there are 15 quadratic binomials, namely the 2×2-minors of matrix

$$\begin{pmatrix} x_{13} & x_{23} & x_{31} & x_{32} & x_{33} & y_{33} \\ y_{13} & y_{23} & y_{31} & y_{32} & y_{33} & z_{33} \end{pmatrix}. \tag{4.2}$$

Note that this matrix has rank 1 under the substitution (4.1). Second, there are three cubics

$$\begin{aligned} &x_{11}x_{22}x_{33} - x_{11}x_{23}x_{32} - x_{12}x_{21}x_{33} + x_{12}x_{23}x_{31} + x_{13}x_{21}x_{32} - x_{13}x_{22}x_{31}, \\ &x_{13}x_{22}y_{31} - x_{12}x_{23}y_{31} - x_{13}x_{21}y_{32} + x_{11}x_{23}y_{32} + x_{12}x_{21}y_{33} - x_{11}x_{22}y_{33}, \\ &x_{22}y_{13}y_{31} - x_{12}y_{23}y_{31} - x_{21}y_{13}y_{32} + x_{11}y_{23}y_{32} + x_{12}x_{21}z_{33} - x_{11}x_{22}z_{33}. \end{aligned} \tag{4.3}$$

These three 3×3-determinants replicate the equation that defines the original model F. ◇

This chapter is organized as follows. Section 4.2 gives the relevant concepts and definitions from computer vision and algebraic geometry. We present camera models with image distortion, with focus on distortions with respect to a single parameter . The resulting distortion varieties $X_{[u]}$ live in the rational normal scroll S_u, where $u = (u_0, u_1, \ldots, u_n)$ is a vector of non-negative integers. This distortion vector indicates that the coordinate x_i on $\mathbb{P}^n$ is replicated u_i times when passing to $\mathbb{P}^N$. In Example 4.1 we have $u = (0, 0, 1, 0, 0, 1, 1, 1, 2)$ and S_u is the 9-dimensional rational normal scroll defined by the 2×2-minors of (4.2).

Our results on one-parameter distortions of arbitrary varieties are stated and proved in Section 4.3. Theorem 4.8 expresses the degree of $X_{[u]}$ in terms of the Chow polytope of X. Theorem 4.16 derives ideal generators for $X_{[u]}$ from a Gröbner basis of X. These results explain what we observed in Example 4.1, namely the degree 16 and the equations in (4.2)-(4.3).

Section 4.4 deals with multi-parameter distortions. We first derive various camera models that are useful for applications, and we then present the relevant algebraic geometry.

Section 4.5 is concerned with a concrete application to solving minimal problems in computer vision. We focus on the distortion variety f+E+ of degree 23 derived in Section 4.2.

4.2 One-parameter distortions

This section has three parts. First, we derive the relevant camera models from computer vision. Second, we introduce the distortion varieties $X_{[u]}$ of an arbitrary projective variety X. And, third, we study the distortion varieties for the camera models from the first part.

Multiview geometry with image distortion

A perspective camera in computer vision [48, p. 158] is a linear projection $\mathbb{P}^3 \dashrightarrow \mathbb{P}^2$. The 3×4-matrix that represents this map is written as $K[R \mid t]$ where $R \in SO(3)$, $t \in \mathbb{R}^3$, and K is an upper-triangular 3×3 matrix known as the calibration matrix. This transforms a point $X \in \mathbb{P}^3$ from the world Cartesian coordinate system to the camera Cartesian coordinate system. Here, we usually normalize homogeneous coordinates on $\mathbb{P}^3$ and $\mathbb{P}^2$ so that the last coordinate equals 1. With this, points in $\mathbb{R}^3$ map to $\mathbb{R}^2$ under the action of the camera.

The following camera model was introduced in [78, Equation 3] to deal with image distortions:

$$\begin{bmatrix} R \mid t \end{bmatrix} X = \mu \begin{bmatrix} h(\|AU + b\|)\,(AU + b) \\ g(\|AU + b\|) \end{bmatrix} \quad \text{for some } \mu \in \mathbb{R}\setminus\{0\}. \tag{4.4}$$

The two functions $h: \mathbb{R} \to \mathbb{R}$ and $g: \mathbb{R} \to \mathbb{R}$ represent the distortion. The invertible matrix $A \in \mathbb{R}^{2\times 2}$ and the vector $b \in \mathbb{R}^2$ are used to transform the image point $U \in \mathbb{R}^2$ into the image Cartesian coordinate system. The perspective camera in the previous paragraph is obtained by setting $h = g = 1$ and taking the calibration matrix K to be the inverse of $\begin{bmatrix} A & b \\ 0\,0 & 1 \end{bmatrix}$.

Micusik and Pajdla [78] studied applications to fish eye lenses as well as catadioptric cameras. In this context they found that it often suffices to fix $h = 1$ and to take a quadratic polynomial for g. For the following derivation we choose $g(t) = 1 + \gamma t^2$, where γ is an unknown parameter. We also assume that the calibration matrix has the diagonal form $K = \operatorname{diag}(f, f, 1)$. If we set $\lambda = \gamma/f^2$ then the model (4.4) simplifies to

$$\begin{bmatrix} R \mid t \end{bmatrix} X = \mu K^{-1} \begin{bmatrix} U \\ 1 + \lambda \|U\|^2 \end{bmatrix} \quad \text{for some } \mu \in \mathbb{R}\setminus\{0\}. \tag{4.5}$$

Let us now analyze two-view geometry for the model (4.5). The quantity $\lambda = \gamma/f^2$ is our distortion parameter. Throughout the discussion in Section 4.2 there is only one such parameter. Later, in Section 4.4, there will be two or more different distortion parameters.

Following [48, Section 9.6] we represent two camera matrices $[R_1 \mid t_1]$ and $[R_2 \mid t_2]$ by their essential matrix E. This 3×3-matrix has rank 2 and satisfies the Demazure equations. The equations were first derived in [25]; they take the matrix form $2EE^\top E - \operatorname{trace}(EE^\top)E = 0$. For a pair (U_1, U_2) of corresponding points in two images, the *epipolar constraint* now reads

$$0 = \begin{bmatrix} AU_2 \\ 1 + \gamma\|AU_2\|^2 \end{bmatrix}^\top E \begin{bmatrix} AU_1 \\ 1 + \gamma\|AU_1\|^2 \end{bmatrix} = \begin{bmatrix} U_2 \\ 1 + \lambda\|U_2\|^2 \end{bmatrix}^\top K^{-\top} E K^{-1} \begin{bmatrix} U_1 \\ 1 + \lambda\|U_1\|^2 \end{bmatrix}. \tag{4.6}$$

In this way, the essential matrix E expresses a necessary condition for two points U_1 and U_2 in the image planes to be pictures of the same world point. The *fundamental matrix* is obtained from the essential matrix and the calibration matrix:

$$F = \begin{pmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{pmatrix} = K^{-\top} E K^{-1}. \tag{4.7}$$

Using the coordinates of $U_1 = [u_1; v_1]^\top$ and $U_2 = [u_2; v_2]^\top$, the epipolar constraint (4.6) is

$$
\begin{aligned}
0 = {} & u_2u_1f_{11} + u_2v_1f_{12} + u_2f_{13} + u_2\|U_1\|^2 f_{13} + v_2u_1f_{21} + v_2v_1f_{22} + v_2f_{23} + v_2\|U_1\|^2 f_{23} + \\
& u_1f_{31} + u_1\|U_2\|^2 f_{31} + v_1f_{32} + v_1\|U_2\|^2 f_{32} + f_{33} + (\|U_1\|^2 + \|U_2\|^2) f_{33} + \|U_1\|^2\|U_2\|^2\ {}^2 f_{33}.
\end{aligned}
$$

This is a sum of 15 terms. The corresponding monomials in the unknowns form the vector

$$m^\top = \left(f_{11}; f_{12}; f_{13}; f_{13}\ ; f_{21}; f_{22}; f_{23}; f_{23}\ ; f_{31}; f_{31}\ ; f_{32}; f_{32}\ ; f_{33}; f_{33}\ ; f_{33}\ {}^2 \right). \qquad (4.8)$$

The 15 coefficients are real numbers given by the data. The coefficient vector c is equal to

$$\left(u_2u_1; u_2v_1; u_2; u_2\|U_1\|^2; v_2u_1; v_2v_1; v_2; v_2\|U_1\|^2; u_1; u_1\|U_2\|^2; v_1; v_1\|U_2\|^2; 1; \|U_1\|^2 + \|U_2\|^2; \|U_1\|^2\|U_2\|^2 \right)^\top.$$

With this notation, the epipolar constraint given by one point correspondence is simply

$$c^\top m = 0. \qquad (4.9)$$

At this stage we have derived the distortion variety in Example 4.1. Identifying f_{ij} with the variables x_{ij}, the vector (4.8) is precisely the same as that in (4.1). This is the parametrization of the rational normal scroll S_u in $\mathbb{P}^{14}$ where $u = (0; 0; 1; 0; 0; 1; 1; 1; 2)$. The set of fundamental matrices is dense in the hypersurface $X = \{\det(F) = 0\}$ in $\mathbb{P}^8$. Its distortion variety $X_{[u]}$ has dimension 8 and degree 16 in $\mathbb{P}^{14}$. Each point correspondence $(U_1; U_2)$ determines a vector c and hence a hyperplane in $\mathbb{P}^{14}$. The constraint (4.9) means intersecting $X_{[u]}$ with that hyperplane. Eight point correspondences determine a 6-dimensional linear space in $\mathbb{P}^{14}$. Intersecting $X_{[u]}$ with that linear subspace is the same as solving the 8-point radial distortion problem in [64, Section 7.1.3]. The expected number of complex solutions is 16.

Scrolls and distortions

This subsection introduces the algebro-geometric objects studied in this chapter. We fix a non-zero vector $u = (u_0; u_1; \ldots; u_n) \in \mathbb{N}^{n+1}$ of non-negative integers, we abbreviate $|u| = u_0 + u_1 + \cdots + u_n$, and we set $N = |u| + n$. The rational normal scroll S_u is a smooth projective variety of dimension $n + 1$ and degree $|u|$ in $\mathbb{P}^N$. It has the parametric representation

$$x_0 : x_0\ : x_0\ {}^2 : \quad : x_0\ {}^{u_0} : x_1 : x_1\ : x_1\ {}^2 : \quad : x_1\ {}^{u_1} : \quad : x_n : x_n\ : \quad : x_n\ {}^{u_n}. \qquad (4.10)$$

The coordinates are monomials, so the scroll S_u is also a toric variety [23]. Since degree$(S_u) = |u|$ equals codim$(S_u) + 1 = N - n + 1$, it is a variety of minimal degree [46, Example 1.14].

Restriction to the coordinates $(x_0 : x_1 : \cdots : x_n)$ defines a rational map $S_u \dashrightarrow \mathbb{P}^n$. This is a toric fibration [27]. Its fibers are curves parametrized by λ. The base locus is a coordinate subspace $\mathbb{P}^n \subset \mathbb{P}^N$. Its points have support on the last coordinate in each of the $n+1$ groups. For instance, in Example 4.2 the base locus is the $\mathbb{P}^2$ defined by $\langle a_0, b_0, b_1, c_0, c_1, c_2 \rangle$ in $\mathbb{P}^8$.

The prime ideal of the scroll S_u is generated by the 2×2-minors of a $2 \times |u|$-matrix of unknowns that is obtained by concatenating Hankel matrices on the blocks of unknowns; see [33, Lemma 2.1], [86], and Example 4.2 below. For a textbook reference see [46, Theorem 19.9].

We now consider an arbitrary projective variety X of dimension d in $\mathbb{P}^n$. This is the underlying model in some application, such as computer vision. We define the distortion variety of level u, denoted $X_{[u]}$, to be the closure of the preimage of X under the map $S_u \dashrightarrow \mathbb{P}^n$. The fibers of this map are curves. The distortion variety $X_{[u]}$ lives in $\mathbb{P}^N$. It has dimension $d+1$. Points on $X_{[u]}$ represent points on X whose coordinates have been distorted by an unknown parameter. The parametrization above is the rule for the distortion. In other words, $X_{[u]}$ is the closure of the image of the regular map $X \times \mathbb{C} \to \mathbb{P}^N$ given by (4.10).

Each distortion variety represents a *minimal problem* [64] in polynomial systems solving. Data points define linear constraints on $\mathbb{P}^N$, like (4.9). Our problem is to solve $d+1$ such linear equations on $X_{[u]}$. The number of complex solutions is the degree of $X_{[u]}$. A simple bound for that degree is stated in Proposition 4.7, and an exact formulas can be found in Theorem 4.8. Of course, in applications we are primarily interested in the real solutions.

We already saw one example of a distortion variety in Example 4.1. In the following example, we discuss some surfaces in $\mathbb{P}^N$ that arise as distortion varieties of plane curves.

Example 4.2. Let $n = 2$ and $u = (1, 2, 3)$. The rational normal scroll is a 3-dimensional smooth toric variety in $\mathbb{P}^8$. Its implicit equations are the 2×2-minors of the 2×6-matrix

$$\begin{pmatrix} a_0 & b_0 & b_1 & c_0 & c_1 & c_2 \\ a_1 & b_1 & b_2 & c_1 & c_2 & c_3 \end{pmatrix}. \tag{4.11}$$

This is the "concatenated Hankel matrix" mentioned above. Its pattern generalizes to all u.

Let X be a general curve of degree d in $\mathbb{P}^2$. The distortion variety $X_{[u]}$ is a surface of degree $5d$ in $\mathbb{P}^8$. Its prime ideal is generated by the 15 minors of (4.11) together with $d+1$ polynomials of degree d. These are obtained from the ternary form that defines X by the distortion process in Theorem 4.16. For special curves X, the degree of $X_{[u]}$ may drop below $5d$. For instance, given a line $X = V(a + b + c)$ in P^2, the distortion surface $X_{[u]}$ has degree 5 if $\neq 0$, it has degree 4 if $= 0$ but $\neq 0$, and it has degree 3 if $= = 0$. For any curve X, the property $\deg(X_{[u]}) = 5 \deg(X)$ holds after a coordinate change in P^2. If $X = \{p\}$ is a single point in P^2 then $X_{[u]}$ is a curve in P^8. It has degree 3 unless $p \in V(c)$. }

Back to two-view geometry

In this subsection we describe several variants of Example 4.1. These highlight the role of distortion varieties in two-view geometry. We fix $n = 8$, $N = 14$ and $u = (0;0;1;0;0;1;1;1;2)$ as above. The scroll S_u is the image of the map (4.1) and its ideal is generated by the 2 2-minors of (4.2). Each of the following varieties live in the space of 3 3-matrices $x = (x_{ij})$.

Example 4.3 (Essential Matrices). We now write E for the essential variety (see [25] or Chapter 2). It has dimension 5 and degree 10 in P^8. Its points x are the essential matrices in (4.6). The ideal of E is generated by ten cubics, namely $\det(x)$ and the nine entries of the matrix $2xx^T x - \mathrm{trace}(xx^T)x$. The distortion variety $E_{[u]}$ has dimension 6 and degree 52 in P^{14}. Its ideal is generated by 15 quadrics and 18 cubics, derived from the ten Demazure cubics. }

Example 4.4 (Essential Matrices plus Two Equal Focal Lengths). Fix a diagonal calibration matrix $k = \mathrm{diag}(f;f;1)$, where f is a new unknown. We define G to be the closure in P^8 of the set of 3 3-matrices x such that $kxk \in E$ for some f. To compute the ideal of the variety G, we use the following lines of code in the computer algebra system Macaulay2 [44]:

```
R=QQ[f,x11,x12,x13,x21,x22,x23,x31,x32,x33,y13,y23,y33,y31,y32,z33,t];
X=matrix {{x11,x12,x13},{x21,x22,x23},{x31,x32,x33}}
K=matrix {{f,0,0},{0,f,0},{0,0,1}};
P=K*X*K;
E=minors(1,2*P*transpose(P)*P-trace(P*transpose(P))*P)+ideal(det(P));
G=eliminate({f},saturate(E,ideal(f)))
codim G, degree G, betti mingens G
```

The output tells us that the variety G has dimension 6 and degree 15, and that G is the complete intersection of two hypersurfaces in P^8, namely the cubic $\det(x)$ and the quintic

$$\begin{aligned}
&x_{11}x_{13}^3x_{31} + x_{13}^2x_{21}x_{23}x_{31} + x_{11}x_{13}x_{23}^2x_{31} + x_{21}x_{23}^3x_{31} - x_{11}x_{13}x_{31}^3 - x_{21}x_{23}x_{31}^3 + \\
&x_{12}x_{13}^3x_{32} + x_{13}^2x_{22}x_{23}x_{32} + x_{12}x_{13}x_{23}^2x_{32} + x_{22}x_{23}^3x_{32} - x_{12}x_{13}x_{31}^2x_{32} - x_{12}^2x_{13}^2x_{33} \\
&- x_{11}x_{13}x_{31}x_{32}^2 - x_{21}x_{23}x_{31}x_{32}^2 - x_{12}x_{13}x_{32}^3 - x_{22}x_{23}x_{32}^3 - x_{11}^2x_{13}^2x_{33} - x_{22}x_{23}x_{31}^2x_{32} \\
&- 2x_{11}x_{13}x_{21}x_{23}x_{33} - 2x_{12}x_{13}x_{22}x_{23}x_{33} - x_{21}^2x_{23}^2x_{33} - x_{22}^2x_{23}^2x_{33} + x_{11}^2x_{31}^2x_{33} \\
&+ x_{21}^2x_{31}^2x_{33} + 2x_{11}x_{12}x_{31}x_{32}x_{33} + 2x_{21}x_{22}x_{31}x_{32}x_{33} + x_{12}^2x_{32}^2x_{33} + x_{22}^2x_{32}^2x_{33}.
\end{aligned} \tag{4.12}$$

The distortion variety $G_{[u]}$ is now computed by the following lines in Macaulay2:

```
Gu = eliminate({t}, G +
  ideal(y13-x13*t,y23-x23*t,y31-x31*t,y32-x32*t,y33-x33*t,z33-x33*t^2))
codim Gu, degree Gu, betti mingens Gu
```

We learn that $G_{[u]}$ has dimension 7 and degree 68 in $\mathbb{P}^{14}$. Modulo the 15 quadrics for S_u, its ideal is generated by three cubics, like those in (4.3), and five quintics, derived from (4.12). }

Example 4.5 (Essential Matrices plus One Focal Length Unknown). Let G^0 denote the 6-dimensional subvariety of $\mathbb{P}^8$ defined by the four maximal minors of the 3×4-matrix

$$\begin{pmatrix} x_{11} & x_{12} & x_{13} & x_{21}x_{31} + x_{22}x_{32} + x_{23}x_{33} \\ x_{21} & x_{22} & x_{23} & -x_{11}x_{31} - x_{12}x_{32} - x_{13}x_{33} \\ x_{31} & x_{32} & x_{33} & 0 \end{pmatrix}. \tag{4.13}$$

This variety has dimension 6 and degree 9 in $\mathbb{P}^8$. It is defined by one cubic and three quartics. The variety G^0 is similar to G in Example 4.4, but with the identity matrix as the calibration matrix for one of the two cameras. We can compute G^0 by running the Macaulay2 code above but with the line `P = K*X*K` replaced with the line `P = X*K`. This model was studied in [16].

The distortion variety $G^0_{[u]}$ has dimension 7 and degree 42 in $\mathbb{P}^{14}$. Modulo the 15 quadrics that define S_u, the ideal of $G^0_{[u]}$ is minimally generated by three cubics and 11 quartics. }

Table 4.1 summarizes the four models we discussed in Examples 4.1, 4.3, 4.4 and 4.5. The first column points to a reference in computer vision where this model has been studied. The last column shows the upper bound for $\deg(X_{[u]})$ given in Proposition 4.7. That bound is not tight in any of our examples. In the second half of the table we report the same data for the four models when only only one of the two cameras undergoes radial distortion.

$u =$ 0; 0; 1; 0; 0; 1; 1; 1; 2	Ref	dim(X)	deg(X)	$\dim(X_{[u]})$	$\deg(X_{[u]})$	Prop 4.7
F in Ex 4.1: +F+	[64]	7	3	8	16	18
E in Ex 4.3: +E+	[64]	5	10	6	52	60
G in Ex 4.4: f+E+f	[57]	6	15	7	68	90
G^0 in Ex 4.5: +E+f		6	9	7	42	54
$v =$ 0; 0; 1; 0; 0; 1; 0; 0; 1	Ref	dim(X)	deg(X)	$\dim(X_{[v]})$	$\deg(X_{[v]})$	Prop 4.7
F in Ex 4.6: F+	[63]	7	3	8	8	9
E in Ex 4.6: E+	[63]	5	10	6	26	30
G in Ex 4.6: f+E+f		6	15	7	37	45
G^0 in Ex 4.6: E+f	[63]	6	9	7	19	27
G^{00} in Ex 4.6: f+E+		6	9	7	23	27

Table 4.1: Dimensions and degrees of two-view models and their radial distortions.

Example 4.6. We revisit the four two-view models discussed above, but with distortion vector $v = (0;0;1;0;0;1;0;0;1)$. Now, $N = 11$ and only one camera is distorted. The rational normal scroll S_v has codimension 2 and degree 3 in P^{11}. Its parametric representation is

$$x_{11} : x_{12} : x_{13} : x_{13} \quad : x_{21} : x_{22} : x_{23} : x_{23} \quad : x_{31} : x_{32} : x_{33} : x_{33} \quad :$$

The distortion varieties $F_{[v]}$, $E_{[v]}$, $G_{[v]}$ and $G^0_{[v]}$ live in P^{11}. Their degrees are shown in the lower half of Table 4.1. For instance, consider the last two rows. The notation E+f means that the right camera has unknown focal length and it is also distorted.

The fth row refers to another variety G^{00}. This is the image of G^0 under the linear isomorphism that maps a 3 3-matrix to its transpose. Since v is not a symmetric matrix, unlike u, the variety $G^{00}_{[v]}$ is actually dierent from $G^0_{[v]}$. The descriptor f+E+ of $G^{00}_{[v]}$ expresses that the left camera has unknown focal length and the right camera is distorted. The variety $G^{00}_{[v]}$ has dimension 7 and degree 23 in P^{11}. In addition to the three quadrics $x_{3i} y_{3j} \quad x_{3j} y_{3i}$ that dene S_v, the ideal generators for $G^{00}_{[v]}$ are two cubics and ve quartics. The minimal problem [63, 64] for this distortion variety is studied in detail in Section 4.5. }

4.3 Equations and degrees

In this section we express the degree and equations of $X_{[u]}$ in terms of those of X. Throughout we assume that X is an irreducible variety of codimension c in P^n and the distortion vector u 2 N^{n+1} satises $u_0 \quad u_1 \qquad u_n$. We begin with a general upper bound for the degree.

Proposition 4.7. *Suppose $u_n \geq 1$. The degree of the distortion variety satisfies*

$$\deg(X_{[u]}) \leq \deg(X)\cdot(u_c + u_{c+1} + \cdots + u_n). \tag{4.14}$$

This holds with equality if the coordinates are chosen so that X is in general position in $\mathbb{P}^n$.

The upper bound in Proposition 4.7 is shown for our models in the last column of Table 4.1. This result will be strengthened in Theorem 4.8 below, where we give an exact degree formula that works for all X. It is instructive to begin with the two extreme cases. If $c = 0$ and $X = \mathbb{P}^n$ then we recover the fact that the scroll $X_{[u]} = S_u$ has degree $N - n = u_0 + \cdots + u_n$. If $c = n$ and X is a general point in $\mathbb{P}^n$ then $X_{[u]}$ is a rational normal curve of degree u_n.

The following proof, and the subsequent development in this section, assumes familiarity with two tools from computational algebraic geometry: the construction of initial ideals with respect to weight vectors, as in [97], and the *Chow form* of a projective variety [24, 39, 43, 60].

Proof of Proposition 4.7. Fix $\dim(X_{[u]}) = n - c + 1$ general linear forms on $\mathbb{P}^N$, denoted $\ell_0, \ell_1, \ldots, \ell_{n-c}$. We write their coefficients as the rows of the $(n-c+1)\times(N+1)$ matrix

$$\begin{bmatrix} \gamma_{0,0} & \gamma_{0,1} & \gamma_{0,2} & \cdots & \gamma_{0,N} \\ \gamma_{1,0} & \gamma_{1,1} & \gamma_{1,2} & \cdots & \gamma_{1,N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \gamma_{n-c,0} & \gamma_{n-c,1} & \gamma_{n-c,2} & \cdots & \gamma_{n-c,N} \end{bmatrix}. \tag{4.15}$$

Here $\gamma_{i,j} \in \mathbb{C}$. The degree of $X_{[u]}$ equals $\#\big(X_{[u]} \cap V(\ell_0, \ldots, \ell_{n-c})\big)$. We shall do this count. Recall that $X_{[u]}$ is the closure of the image of the injective map $X \times \mathbb{C} \to \mathbb{P}^N$ given in (4.10). The image of this map is dense in $X_{[u]}$. Its complement is the $\mathbb{P}^n$ consisting of all points whose coordinates in each the $n+1$ groups are zero except for the last one. Since the linear forms ℓ_i are generic, all points of $X_{[u]} \cap V(\ell_0, \ldots, \ell_{n-c})$ lie in this image. By injectivity of the map, $\deg(X_{[u]})$ is the number of pairs $(x, \lambda) \in X \times \mathbb{C}$ which map into $X_{[u]} \cap V(\ell_0, \ldots, \ell_{n-c})$.

We formulate this condition on (x, λ) as follows. Consider the $(n-c+1)\times(n+1)$ matrix

$$\begin{bmatrix} \gamma_{0,0} + \gamma_{0,1}\lambda + \cdots + \gamma_{0,u_0}\lambda^{u_0} & & & \gamma_{0,u_0+\cdots+u_{n-1}+1} + \cdots + \gamma_{0,N-n}\lambda^{u_n} \\ \gamma_{1,0} + \gamma_{1,1}\lambda + \cdots + \gamma_{1,u_0}\lambda^{u_0} & & \cdots & \gamma_{1,u_0+\cdots+u_{n-1}+1} + \cdots + \gamma_{1,N-n}\lambda^{u_n} \\ \vdots & \ddots & \ddots & \vdots \\ \gamma_{n-c,0} + \gamma_{n-c,1}\lambda + \cdots + \gamma_{n-c,u_0}\lambda^{u_0} & & & \gamma_{n-c,u_0+\cdots+u_{n-1}+1} + \cdots + \gamma_{n-c,N-n}\lambda^{u_n} \end{bmatrix}. \tag{4.16}$$

We want to count pairs $(x;\lambda) \in \mathbb{P}^n \times \mathbb{C}$ such that $x \in X$ and x lies in the kernel of this matrix. By genericity of ℓ_i, this matrix has rank $n - c + 1$ for all $\lambda \in \mathbb{C}$. So for each $\lambda \in \mathbb{C}$, the kernel of the matrix (4.16) is a linear subspace of dimension $c - 1$ in $\mathbb{P}^n$.

We conclude that (4.16) defines a rational curve in the Grassmannian $Gr(\mathbb{P}^{c-1};\mathbb{P}^n)$. Here the $\alpha_{i;j}$ are fixed generic complex numbers and λ is an unknown that parametrizes the curve. If we take the Grassmannian in its Plücker embedding then the degree of our curve is $u_c + u_{c+1} + \cdots + u_n$, which is the largest degree in λ of any maximal minor of (4.16).

At this point we use the Chow form Ch_X of the variety X. As in Chapter 2, following [24, 43], this is the defining equation of an irreducible hypersurface in the Grassmannian $Gr(\mathbb{P}^{c-1};\mathbb{P}^n)$. Its points are the subspaces that intersect X. The degree of Ch_X in Plücker coordinates is $\deg(X)$.

We now consider the intersection of our curve with the hypersurface defined by Ch_X. Equivalently, we substitute the maximal minors of (4.16) into Ch_X and we examine the resulting polynomial in λ. Since the matrix entries $\alpha_{i;j}$ in (4.15) are generic, the curve intersects the hypersurface of the Chow form Ch_X outside its singular locus. By Bezout's Theorem, the number of intersection points is bounded above by $\deg(X) \cdot (u_c + u_{c+1} + \cdots + u_n)$.

Each intersection point is non-singular on $V(Ch_X)$, and so the corresponding linear space intersects the variety X in a unique point x. We conclude that the number of desired pairs $(x;\lambda)$ is at most $\deg(X) \cdot (u_c + u_{c+1} + \cdots + u_n)$. This establishes the upper bound.

For the second assertion, we apply a general linear change of coordinates to X in $\mathbb{P}^n$. Consider the lexicographically last Plücker coordinate, denoted $p_{c;c+1;\dots;n}$. The monomial $p_{c;c+1;\dots;n}^{\deg(X)}$ appears with non-zero coefficient in the Chow form Ch_X. Substituting the maximal minors of (4.16) into Ch_X, we obtain a polynomial in λ of degree $\deg(X) \cdot (u_c + u_{c+1} + \cdots + u_n)$. By the genericity hypothesis on (4.15), this polynomial has distinct roots in $\mathbb{C}$. These represent distinct points in $X_{[u]} \setminus V(\ell_0;\dots;\ell_{n-c})$, and we conclude that the upper bound is attained. □

We will now refine the method in the proof above to derive an exact formula for the degree of $X_{[u]}$ that works in all cases. The Chow form Ch_X is expressed in primal Plücker coordinates $p_{i_0;i_1;\dots;i_{n-c}}$ on $Gr(\mathbb{P}^{c-1};\mathbb{P}^n)$. The *weight* of such a coordinate is the vector $e_{i_0} + e_{i_1} + \cdots + e_{i_{n-c}}$, and the weight of a monomial is the sum of the weights of its variables. The *Chow polytope* of X is the convex hull of the weights of all Plücker monomials appearing in Ch_X; see [60].

Theorem 4.8. *The degree of $X_{[u]}$ is the maximum value attained by the linear functional $w \mapsto u \cdot w$ on the Chow polytope of X. This positive integer can be computed by the formula*

$$\text{degree}(X_{[u]}) \;=\; \sum_{j=0}^{n} u_j \cdot \text{degree}\left(\text{in}_{-u}(X) : \langle x_j \rangle^{\infty} \right) , \qquad (4.17)$$

where $\text{in}_{-u}(X)$ is the initial monomial ideal of X with respect to a term order that refines $-u$.

Proof. Let M be a monomial ideal in $x_0, x_1, \ldots, x_n$ whose variety is pure of codimension c. Each of its irreducible components is a subspace $\text{span}(e_{i_0}, e_{i_1}, \ldots, e_{i_{n-c}})$ of P^n. We write $\mu_{i_0,i_1,\ldots,i_{n-c}}$ for the multiplicity of M along that coordinate subspace. By [60, Theorem 2.6], the Chow form of (the cycle given by) M is the Plücker monomial $\prod p_{i_0,i_1,\ldots,i_{n-c}}^{\mu_{i_0,i_1,\ldots,i_{n-c}}}$, and the Chow polytope of M is the point $\sum \mu_{i_0,i_1,\ldots,i_{n-c}}(e_{i_0} + e_{i_1} + \cdots + e_{i_{n-c}})$. The j-th coordinate of that point can be computed from M without performing a monomial primary decomposition. Namely, the j-th coordinate of the Chow point of M is the degree of the saturation $M : \langle x_j \rangle^{\infty}$. This follows from [60, Proposition 3.2] and the proof of [60, Theorem 3.3].

We now substitute each maximal minor of the matrix (4.16) for the corresponding Plücker coordinate $p_{i_0,i_1,\ldots,i_{n-c}}$. This results in a general polynomial of degree $u_{i_0} + u_{i_1} + \cdots + u_{i_{n-c}}$ in the one unknown τ. When carrying out this substitution in the Chow form Ch_X, the highest degree terms do not cancel, and we obtain a polynomial in τ whose degree is the largest u-weight among all Plücker monomials in Ch_X. Equivalently, this degree in τ is the maximum inner product of the vector u with any vertex of the Chow polytope of X.

One vertex that attains this maximum is the Chow point of the monomial ideal $M = \text{in}_{-u}(X)$ in the proof of Proposition 4.7. Note that we had chosen one particular term order to refine the partial order given by $-u$. If we vary that term order then we obtain all vertices on the face of the Chow polytope supported by u. The saturation formula for the Chow point of the monomial ideal M in the first paragraph of the proof completes our argument. □

We are now able to characterize when the upper bound in Proposition 4.7 is attained. Let c_- and c_+ be the smallest and largest index respectively such that $u_{c_-} = u_c = u_{c_+}$. We define a set L_u of $n - c + 1$ linear forms as follows. Start with the $n - c_+$ variables $x_{c_+ +1}$, $x_{c_+ +2}$, …, x_n and then take $c_+ - c + 1$ generic linear forms in the variables $x_{c_-}, x_{c_- +1}, \ldots, x_{c_+}$. In the case when u has distinct coordinates, $V(L_u)$ is simply the subspace spanned by $e_0, e_1, \ldots, e_{n-c}$.

Corollary 4.9. The degree of $X_{[u]}$ is the right hand side of (4.14) if and only if $V(L_u) \cap X = \emptyset$.

Proof. The quantity $\deg(X)\cdot(u_c + u_{c+1} + \cdots + u_n)$ is the maximal u-weight among Plücker monomials of degree equal to $\deg(X)$. The monomials that attain this maximal u-weight are products of $\deg(X)$ many Plücker coordinates of weight $u_c + u_{c+1} + \cdots + u_n$. These are precisely the Plücker coordinates $p_{i_0,i_1,\ldots,i_{c+\ -c},u_{c+\ +1},\ldots,u_n}$, where $c \le i_0 < i_1 < \cdots < i_{c+\ -c} \le c+$.

Such monomials are non-zero when evaluated at the subspace $V(L_u)$. All other monomials, namely those having smaller u-weight, evaluate to zero on $V(L_u)$. Hence the Chow form Ch_X has terms of degree $\deg(X)\cdot(u_c + u_{c+1} + \cdots + u_n)$ if and only if Ch_X evaluates to a non-zero constant on $V(L)$ if and only if the intersection of X with $V(L_u)$ is empty. □

We present two example to illustrate the exact degree formula in Theorem 4.8.

Example 4.10. Suppose X is a hypersurface in $\mathbb{P}^n$, defined by a homogeneous polynomial $f(x_0,\ldots,x_n)$ of degree d. Let φ be the tropicalization of f, with respect to min-plus algebra, as in [74]. Equivalently, φ is the support function of the Newton polytope of f. Then

$$\deg(X_{[u]}) = d\cdot|u| - \varphi(u_0,u_1,\ldots,u_n). \qquad (4.18)$$

For instance, let $n = 8, d = 3$ and f the determinant of a 3×3-matrix. Hence X is the variety of fundamental matrices, as in Example 4.1. The tropicalization of the 3×3-determinant is

$$\varphi = \min\left(u_{11}+u_{22}+u_{33},\, u_{11}+u_{23}+u_{32},\, u_{12}+u_{21}+u_{33},\, u_{12}+u_{23}+u_{31},\, u_{13}+u_{21}+u_{32},\, u_{13}+u_{22}+u_{31}\right).$$

The degree of the distortion variety $X_{[u]}$ equals $3\cdot\left(\sum u_{ij}\right) - \varphi$. This explains the degree 16 we had observed in Example 4.1 for the radial distortion of the fundamental matrices. }

Example 4.11. Let X be the variety of essential matrices with the same distortion vector u. In Example 4.3, we found that $\deg(X_{[u]}) = 52$. The following Macaulay2 code verifies this:

```
U = {0,0,1,0,0,1,1,1,2};
R = QQ[x11,x12,x13,x21,x22,x23,x31,x32,x33,Weights=>apply(U,i->10-i)];
P = matrix {{x11,x12,x13},{x21,x22,x23},{x31,x32,x33}}
X = minors(1,2*P*transpose(P)*P-trace(P*transpose(P))*P)+ideal(det(P));
M = ideal leadTerm X;
sum apply( 9, i -> U_i * degree(saturate(M,ideal((gens R)_i))) )
```

Here, M is the monomial ideal in $_u(X)$, and the last line is our saturation formula in (4.17). }

We next derive the equations that de ne the distortion variety $X_{[u]}$ from those that de ne the underlying variety X. Our point of departure is the ideal of the rational normal scroll S_u. It is generated by the $\binom{N\quad n}{2}$ minors of the concatenated Hankel matrix. The following lemma is well-known and easy to verify using Buchberger's S-pair criterion; see also [86].

Lemma 4.12. The $2 \quad 2$-minors that de ne the rational normal scroll S_u form a Gröbner basis with respect to the diagonal monomial order. The initial monomial ideal is squarefree.

For instance, in Example 4.2, when $n = 2$ and $u = (1;2;3)$, the initial monomial ideal is

$$h a_0b_1; a_0b_2; a_0c_1; a_0c_2; a_0c_3; b_0b_2; b_0c_1; b_0c_2; b_0c_3; b_1c_1; b_1c_2; b_1c_3; c_0c_2; c_0c_3; c_1c_3 i : \quad (4.19)$$

A monomial m is standard if it does not lie in this initial ideal. The weight of a monomial m is the sum of its indices. Equivalently, the weight of m is the degree in of the monomial in $N+1$ variables that arises from m when substituting in the parametrization of S_u.

Lemma 4.13. Consider any monomial $x = x_0^{\ 0} x_1^{\ 1} \quad x_n^{\ n}$ of degree j in the coordinates of P^n. For any nonnegative integer $i \quad u$ there exists a unique monomial m in the coordinates on P^N such that m is standard and maps to $x \quad ^i$ under the parametrization of the scroll S_u.

Proof. The polyhedral cone corresponding to the toric variety S_u consists of all pairs $(;i)\ 2\ R_0^{n+2}$ with $0 \quad i \quad u$. Its lattice points correspond to monomials $x\ t^i$ on S_u. Since the initial ideal in Lemma 4.12 is square-free, the associated regular triangulation of the polytope is unimodular, by [97, Corollary 8.9]. Each lattice point $(;i)$ has a unique representation as an N-linear combination of generators that span a cone in the triangulation. Equivalently, $x\ t^i$ has a unique representation as a standard monomial in the $N+1$ coordinates on P^N. □

We refer to the standard monomial m in Lemma 4.13 as the ith distortion of the given x .

Example 4.14. In Example 4.2 we have $n = 2$, $N = 8$, and S_u corresponds to the cone over a triangular prism. The lattice points in that cone are the monomials $x_0^{\nu_0} x_1^{\nu_1} x_2^{\nu_2} t^i$ with $0 \le i \le \nu_0 + 2\nu_1 + 3\nu_2$. Using the ambient coordinates on $\mathbb{P}^8$, each such monomial is written uniquely as $a_0^{\lambda_{00}} a_1^{\lambda_{01}} b_0^{\lambda_{10}} b_1^{\lambda_{11}} b_2^{\lambda_{12}} c_0^{\lambda_{20}} c_1^{\lambda_{21}} c_2^{\lambda_{22}} c_3^{\lambda_{23}}$ that is not in (4.19) and satisfies $\lambda_{00} + \lambda_{01} = \nu_0$; $\lambda_{10} + \lambda_{11} + \lambda_{12} = \nu_1$; $\lambda_{20} + \lambda_{21} + \lambda_{22} + \lambda_{23} = \nu_2$; $\lambda_{01} + \lambda_{11} + 2\lambda_{12} + \lambda_{21} + 2\lambda_{22} + 3\lambda_{23} = i$. For instance, if $x^\nu = x_0^3 x_1^2 x_2^2$ then its various distortions, for $0 \le i \le 13$, are the monomials

$$a_0^3 b_0^2 c_0^2;\ a_0^3 b_0^2 c_0 c_1;\ a_0^3 b_0^2 c_0 c_2;\ a_0^3 b_0^2 c_0 c_3;\ a_0^3 b_0^2 c_1 c_3;\ a_0^3 b_0^2 c_2 c_3;\ a_0^3 b_0^2 c_3^2;$$
$$a_0^3 b_0 b_1 c_3^2;\ a_0^3 b_0 b_2 c_3^2;\ a_0^3 b_1 b_2 c_3^2;\ a_0^3 b_2^2 c_3^2;\ a_0^2 a_1 b_2^2 c_3^2;\ a_0 a_1^2 b_2^2 c_3^2;\ a_1^3 b_2^2 c_3^2.$$

Given any homogeneous polynomial p in the unknowns $x_0, x_1, \ldots, x_n$, we write $p_{[i]}$ for the polynomial on $\mathbb{P}^N$ that is obtained by replacing each monomial in p by its ith distortion.

Example 4.15. For the scroll in Example 4.2, the distortions of the sextic $p = a^6 + a^2 b^2 c^2$ are

$$p_{[0]} = a_0^6 + a_0^2 b_0^2 c_0^2;\ p_{[1]} = a_0^5 a_1 + a_0 a_1 b_0^2 c_0^2;\ \ldots;\ p_{[5]} = a_0 a_1^5 + a_1^2 b_1 b_2 c_0^2;\ p_{[6]} = a_1^6 + a_1^2 b_2^2 c_0^2;\ \ldots$$

The following result shows how the equations of $X_{[u]}$ can be read off from those of X.

Theorem 4.16. The ideal of the distortion variety $X_{[u]}$ is generated by the $\binom{N-n}{2}$ quadrics that define S_u together with the distortions $p_{[i]}$ of the elements p in the reduced Gröbner basis of X for a term order that refines the weights u. Hence, the ideal is generated by polynomials whose degree is at most the maximal degree of any monomial generator of $M = \mathrm{in}_u(X)$.

Proof. Since $X_{[u]} \subseteq S_u$, the binomial quadrics that define S_u lie in the ideal $I(X_{[u]})$. Also, if p is a polynomial that vanishes on X then all of its distortions $p_{[i]}$ are in $I(X_{[u]})$ because

$$p_{[i]}\left(x_0, \lambda x_0, \ldots, \lambda^{u_0} x_0, x_1, \ldots, \lambda^{u_n} x_n\right) = \lambda^i p(x) = 0 \quad \text{for } \lambda \in \mathbb{C} \text{ and } x \in X.$$

Conversely, consider any homogeneous polynomial F in $I(X_{[u]})$. It must be shown that F is a polynomial linear combination of the specified quadrics and distortion polynomials. Without loss of generality, we may assume that F is standard with respect to the Gröbner basis in Lemma 4.12, and that each monomial in F has the same weight i. This implies

$$F\left(x_0, \lambda x_0, \ldots, \lambda^{u_0} x_0, x_1, \ldots, \lambda^{u_n} x_n\right) = \lambda^i f(x)$$

for some homogeneous $f \in \mathbb{C}[x_0, \dots, x_n]$. Since $F \in I(X_{[u]})$, we have $f \in I(X)$. We write

$$f = h_1 p_1 + h_2 p_2 + \cdots + h_k p_k,$$

where $p_1, p_2, \dots, p_k$ are in the reduced Gröbner basis of $I(X)$ with respect to a term order refining u, and the multipliers satisfy $\deg_u(f) \geq \deg_u(h_j p_j) = \deg_u(h_j) + \deg_u(p_j)$ for $j = 1, 2, \dots, k$. Since $F = f_{[i]}$, we have $\deg_u(f) \geq i$. Hence, for each j there exist nonnegative integers a_j and b_j such that $a_j + b_j = i$ and $\deg_u(h_j) \leq a_j$ and $\deg_u(p_j) \leq b_j$. The latter inequalities imply that the distortion polynomials $(h_j)_{[a_j]}$ and $(p_j)_{[b_j]}$ exist.

Now consider the following polynomial in the coordinates of $\mathbb{P}^N$:

$$\tilde{F} = (h_1)_{[a_1]} \cdot (p_1)_{[b_1]} + \cdots + (h_k)_{[a_k]} \cdot (p_k)_{[b_k]}.$$

By construction, $\tilde{F}$ and F both map to $\lambda^i f$ under the parameterization of the scroll S_u. Thus, $\tilde{F} - F \in I(S_u)$. This shows that F is a polynomial linear combination of generators of $I(S_u)$ and distortions of Gröbner basis elements $p_1, \dots, p_k$. This completes the proof. □

We illustrate this result with two examples.

Example 4.17. If X is a hypersurface of degree $d \geq 2$ then the ideal $I(X_{[u]})$ is generated by binomial quadrics and distortion polynomials of degree d. More generally, if the generators of $I(X)$ happen to be a Gröbner basis for u then the degree of the generators of $I(X_{[u]})$ does not go up. This happens for all the varieties from computer vision seen in Section 2. ♢

In general, however, the maximal degree among the generators of $I(X_{[u]})$ can be much larger than that same degree for $I(X)$. This happens for complete intersection curves in $\mathbb{P}^3$:

Example 4.18. Let X be the curve in $\mathbb{P}^3$ obtained as the intersection of two random surfaces of degree 4. We fix $u = (2, 3, 4, 4)$. The initial ideal $M = \mathrm{in}_u(X)$ has 51 monomial generators. The largest degree is 32. We now consider the distortion surface $X_{[u]}$ in $\mathbb{P}^{12}$. The ideal of $I(X_{[u]})$ is minimally generated by 133 polynomials. The largest degree is 32. ♢

4.4 Multi-parameter distortions

In this section we study multi-parameter distortions of a given projective variety $X \subset \mathbb{P}^n$. Now, $\lambda = (\lambda_1, \dots, \lambda_r)$ is a vector of r parameters, and $u = (u_0, \dots, u_n)$

where $u_i = \{u_{i,1}, u_{i,2}, \ldots, u_{i,s_i}\}$ is an arbitrary finite subset of $\mathbb{N}^r$. Each point $u_{i,j}$ represents a monomial in the r parameters, denoted $\lambda^{u_{i,j}}$. We set $|u| = \sum_{i=0}^{n} |u_i| = \sum_{i=0}^{n} s_i$ and $N = |u| - 1$. The role of the scroll is played by a toric variety C_u of dimension $n + r$ in $\mathbb{P}^N$ that is usually not smooth. Generalizing (4.10), we define the Cayley variety C_u in $\mathbb{P}^N$ by the parametrization

$$\left(x_0\lambda^{u_{0,1}} : x_0\lambda^{u_{0,2}} : \cdots : x_0\lambda^{u_{0,s_0}} : x_1\lambda^{u_{1,1}} : \cdots : x_1\lambda^{u_{1,s_1}} : \cdots : x_r\lambda^{u_{r,1}} : \cdots : x_r\lambda^{u_{r,s_r}}\right). \tag{4.20}$$

The name was chosen because C_u is the toric variety associated with the Cayley configuration of the configuration u. Its convex hull is the *Cayley polytope*; see [27, Section 3] and [74, Definition 4.6.1].

The *distortion variety* $X_{[u]}$ is defined as the closure of the set of all points (4.20) in $\mathbb{P}^N$ where $x \in X$ and $\lambda \in (\mathbb{C}^*)^r$. Hence $X_{[u]}$ is a subvariety of the Cayley variety C_u, typically of dimension $d + r$ where $d = \dim(X)$. Note that, even in the single-parameter setting ($r = 1$), we have generalized our construction, by permitting u_i to not be an initial segment of $\mathbb{N}$.

Example 4.19. Let $r = n = 2$, $u_0 = \{(0,0),(0,1)\}$, $u_1 = \{(0,0),(1,0)\}$, $u_2 = \{(2,2),(1,1)\}$. The Cayley variety C_u is the singular hypersurface in $\mathbb{P}^5$ defined by $a_0b_0c_0 - a_1b_1c_1$. Let X be the conic in $\mathbb{P}^2$ given by $x_0^2 + x_1^2 - x_2^2$. The distortion variety $X_{[u]}$ is a threefold of degree 10. Its ideal is $\langle a_0b_0c_0 - a_1b_1c_1,\ a_0^2c_0^2 + b_0^2c_0^2 - c_1^4,\ a_0^2a_1b_1c_0 + a_1b_0^2b_1c_0 - a_0b_0c_1^3,\ a_0^2a_1^2b_1^2 + a_1^2b_0^2b_1^2 - a_0^2b_0^2c_1^2\rangle$. $\diamond$

Two views with two or four distortion parameters

We now present some motivating examples from computer vision. Multi-dimensional distortions arise when several cameras have different unknown radial distortions, or when the distortion function $g(t) = 1 + \lambda t^2$ in (4.4)–(4.5) is replaced by a polynomial of higher degree.

We return to the setting of Section 4.2, and we introduce two distinct distortion parameters λ_1 and λ_2, one for each of the two cameras. The role of the equation (4.6) is played by

$$0 = \begin{pmatrix} U_2 \\ 1 + \lambda_2\|U_2\|^2 \end{pmatrix}^T \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix} \begin{pmatrix} U_1 \\ 1 + \lambda_1\|U_1\|^2 \end{pmatrix}. \tag{4.21}$$

Just like in (4.9), this translates into one linear equation $c^{>} m = 0$, where now $m^{>} =$ $[x_{11}; x_{12}; x_{13};\ _1x_{13}; x_{21}; x_{22}; x_{23};\ _1x_{23}; x_{31}; x_{31}\ _2; x_{32}; x_{32}\ _2; x_{33}; x_{33}\ _2; x_{33}\ _1; x_{33}\ _1\ _2]$ and $c^{>} =$ $u_2u_1; u_2v_1; u_2; u_2kU_1k^2; v_2u_1; v_2v_1; v_2; v_2kU_1k^2; u_1; u_1kU_2k^2; v_1; v_1kU_2k^2; 1; kU_1k^2; kU_2k^2; kU_1k^2kU_2k^2$.

Here c is a real vector of data, whereas $= (\ _1;\ _2)$ and $x = (x_{ij})$ comprise 11 unknowns. The vector m is a monomial parametrization of the form (4.20). The corresponding con guration u is given by $u_{11} = u_{12} = u_{21} = u_{22} = f(0;0)g; u_{13} = u_{23} = f(0;0);(1;0)g; u_{31} = u_{32} = f(0;0);(0;1)g; u_{33} = f(0;0);(1;0);(0;1);(1;1)g$. The Cayley variety C_u lives in P^{15}. It has dimension 10 and degree 10. Its toric ideal is generated by 11 quadratic binomials.

Let $X \quad P^8$ be one of the two-view models F, E, G, or G^0 in Subsection 4.2. The following table concerns the distortion varieties $X_{[u]}$ in P^{15}. It is an extension of Table 4.1.

	dim(X), deg(X)	dim($X_{[u]}$)	deg($X_{[u]}$)	Prop 4.7 iterated	# ideal gens of deg 2, 3, 4, 5
F in Ex 4.1: $_1$+F+ $_2$	7, 3	9	24	36	11, 4, 0, 0
E in Ex 4.3: $_1$+E+ $_2$	5, 10	7	76	120	11, 20, 0, 0
G in Ex 4.4: $_1$f+E+f $_2$	6, 15	8	104	180	11, 4, 0, 4
G^0 in Ex 4.5: $_1$+E+f $_2$	6, 9	8	56	108	11, 4, 15, 0

Table 4.2: Dimensions, degrees, mingens of two-view models and their two-parameter radial distortions.

On each $X_{[u]}$ we consider linear systems of equations $c^{>} m = 0$ that arise from point correspondences. For a minimal problem, the number of such epipolar constraints is dim($X_{[u]}$), and the expected number of its complex solutions is deg($X_{[u]}$) (though e.g. in three-view geometry, degree drops occur; see Theorem 3.6). The last column summarizes the number of minimal generators of the ideal of $X_{[u]}$. For instance, the variety $X_{[u]} = E_{[u]}$ for essential matrices is de ned by 11 quadrics (from C_u), 20 cubics, 0 quartics and 0 quintics. If we add 7 general linear equations to these then we have a system with 76 solutions in P^{15}. The penultimate column of Table 4.2 gives an upper bound on deg($X_{[u]}$) that is obtained by applying Proposition 4.7 twice, after decomposing u into two one-parameter distortions.

We next discuss four-parameter distortions for two cameras. These are based on the following model for epipolar constraints, which is a higher-order version of equation (4.21):

$$0 = \begin{bmatrix} U_2 \\ 1 + \ _2kU_2k^2 + \ _2kU_2k^4 \end{bmatrix}^{>} \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix} \begin{bmatrix} U_1 \\ 1 + \ _1kU_1k^2 + \ _1kU_1k^4 \end{bmatrix} : \quad (4.22)$$

As before, the 3 3-matrix $x = (x_{ij})$ belongs to a two-view camera model E, F, G or G'. We rewrite (4.22) as the inner product $c^{\top} m = 0$ of two vectors, where c records the data and m is a parametrization for the distortion variety. We now have $n = 9$, $r = 4$ and $|u| = 25$. The configurations in $\mathbb{N}^4$ that furnish the degrees for this four-parameter distortion are

$$u_{11} = u_{12} = u_{21} = u_{22} = \{0\};$$
$$u_{13} = u_{23} = \{0, (1,0,0,0), (0,0,1,0)\}, u_{31} = u_{32} = \{0, (0,1,0,0), (0,0,0,1)\},$$
$$u_{33} = \{0, (1,0,0,0), (0,1,0,0), (0,0,1,0), (0,0,0,1), (1,0,1,0), (1,0,0,1), (0,1,1,0), (0,1,0,1)\}.$$

Each of the resulting distortion varieties $X_{[u]}$ lives in $\mathbb{P}^{24}$ and satisfies $\dim(X_{[u]}) = \dim(X) + 4$. As before, we may compute the prime ideals for these distortion varieties by elimination, for instance in Macaulay2. From this, we obtain the information displayed in Table 4.3.

	dim	deg	quadrics	cubics	quartics	quintics
F in Ex 4.1: $_1\,_1 + F + \,_2\,_2$	11	115	51	9		
E in Ex 4.3: $_1\,_1 + E + \,_2\,_2$	9	354	51	34		
G in Ex 4.4: $_1\,_1 f + E + f\,_2\,_2$	10	245	51	9	42	
G' in Ex 4.5: $_1\,_1 + E + f\,_2\,_2$	10	475	51	9		9

Table 4.3: Dimension, degrees, number of minimal generators for four-parameter radial distortions.

In each case, the 51 quadrics are binomials that define the ambient Cayley variety C_u in $\mathbb{P}^{24}$. The minimal problems are now more challenging than those in Tables 4.1 and 4.2. For instance, to recover the essential matrix along with four distortion parameters from 9 general point correspondences, we must solve a polynomial system that has 354 complex solutions.

Iterated distortions and their tropicalization

In what follows we take a few steps towards a geometric theory of multi-parameter distortions. We begin with the observation that multi-parameter distortions arising in practice, including those in Subsection 4.4, will often have an inductive structure. Such a structure allows us to decompose them as successive one-parameter distortions where the degrees form an initial segment of the non-negative integers $\mathbb{N}$. In that case the results of Section 4.2 can be applied iteratively. The following proposition characterizes when this is possible. For $u_i \subset \mathbb{N}^r$ and $k < r$, we write $u_i|_{\mathbb{N}^k} \subset \mathbb{N}^k$ for the projection of the set u_i onto the first k coordinates.

Proposition 4.20. Let $u = (u_0, \ldots, u_n)$ be a sequence of finite nonempty subsets of $\mathbb{N}^r$. The multi-parameter distortion with respect to u in $\lambda_1, \ldots, \lambda_r$ is a succession of one-parameter distortions by initial segments, in λ_1, then λ_2, and so on, if and only if each fiber of the maps $u_i|_{\mathbb{N}^k} \to u_i|_{\mathbb{N}^{k-1}}$ becomes an initial segment of $\mathbb{N}$ when projected onto the k^{th} coordinate. This condition holds when each u_i is an order ideal in the poset $\mathbb{N}^r$, with coordinate-wise order.

Proof. We show this for $r = 2$. The general case is similar but notationally more cumbersome. The two-parameter distortion given by a sequence u decomposes into two one-parameter distortions if and only if there exist vectors $v = (v_0, \ldots, v_n) \in \mathbb{N}^{n+1}$ and $w = (w_0, \ldots, w_n) \in \mathbb{N}^{v_0+1} \times \cdots \times \mathbb{N}^{v_n+1}$ such that $u_i = \{(s, t) : 0 \le s \le v_i$ and $0 \le t \le w_{is}\}$ for $i = 0, 1, \ldots, n$. This means that both the Cayley variety and any distortion subvariety decomposes as follows:

$$C_u = (S_v)_{[w]} \quad \text{and} \quad X_{[u]} = (X_{[v]})_{[w]}. \tag{4.23}$$

The segment $[0, v_i]$ in $\mathbb{N}$ is the unique fiber of the map $u_i|_{\mathbb{N}^1} \to u_i|_{\mathbb{N}^0} = \{0\}$. The fiber of $u_i|_{\mathbb{N}^2} \to u_i|_{\mathbb{N}^1} = [0, v_i]$ over an integers is the segment $[0, w_{is}]$ in $\mathbb{N}$. Thus the stated condition on fibers is equivalent to the existence of the non-negative integers v_i and w_{is}. For the second claim, we note that the set u_i is an order ideal in $\mathbb{N}^2$ precisely when $w_{i0} \ge w_{i1} \ge \cdots \ge w_{is}$. □

Proposition 4.20 applies to all models seen in Subsection 4.4 since the u_i are order ideals.

Example 4.21. Consider the two-parameter radial distortion model for two cameras derived in (4.21). The vectors in the above proof are $v = (0, 0, 1, 0, 0, 1, 0, 0, 1)$ and $w = \big(0, 0, (0, 0), 0, 0, (0, 0), 1, 1, (1, 1)\big)$. The decomposition (4.23) holds for all four models $X = E, F, G, G^0$. The penultimate column of Table 4.2 says that the degree of $(X_{[v]})_{[w]}$ is bounded above by $12 \deg(X)$. This follows directly from Proposition 4.7 because $12 = |v| \cdot |w|$. }

The exact degrees for $X_{[u]}$ shown in Tables 4.2 and 4.3 were found using Gröbner bases. This computation starts from the ideal of X and incorporates the structure in Proposition 4.20.

Tropical Geometry [74] furnishes tools for studying multi-parameter distortion varieties. In what follows, we identify any variety $X \subseteq \mathbb{P}^n$ with its reembedding into $\mathbb{P}^N$, where the i-th coordinate x_i has been duplicated $|u_i|$ times. Consider the distortion variety $1_{[u]}$ of the point $1 = (1 : 1 : \cdots : 1)$ in $\mathbb{P}^n$. This is the toric variety in $\mathbb{P}^N$ given by the parametrization

$$\lambda^{u_{0,1}} : \lambda^{u_{0,2}} : \cdots : \lambda^{u_{0,s_0}} : \lambda^{u_{1,1}} : \cdots : \lambda^{u_{1,s_1}} : \cdots : \lambda^{u_{r,1}} : \cdots : \lambda^{u_{r,s_r}} \quad \text{for} \quad \lambda \in (\mathbb{C}^*)^{r+1}.$$

Let $\tilde{u}$ denote the $(r+1) \times (N+1)$-matrix whose columns are vectors in the sets u_i for $i = 0, 1, \ldots, n$, augmented by an extra all-one row vector $(1, 1, \ldots, 1)$. This matrix represents the toric variety $1_{[u]}$. Recall that the Hadamard product $\star$ of two vectors in $\mathbb{C}^{n+1}$ is their coordinate-wise product. This operation extends to points in $\mathbb{P}^n$ and also to subvarieties.

Theorem 4.22. *Fix a projective variety $X \subseteq \mathbb{P}^n$ and any distortion system u, regarded as $r \times (N+1)$-matrix. The distortion variety is the Hadamard product of X with a toric variety:*

$$X_{[u]} = X \star 1_{[u]}$$

Its tropicalization is the Minkowski sum of the tropicalization of X with a linear space:

$$\mathrm{trop}(X_{[u]}) = \mathrm{trop}(X) + \mathrm{trop}(1_{[u]}) = \mathrm{trop}(X) + \mathrm{rowspace}(\tilde{u}). \tag{4.24}$$

Proof. This follows from equation (4.20) and [74, Section 5]. The toric variety $1_{[u]}$ in $\mathbb{P}^N$ is represented by the matrix $\tilde{u}$, in the sense of [97], so its tropicalization is the row space of $\tilde{u}$. Tropicalization takes Hadamard products into Minkowski sums, by [12, Proposition 5.1] or [74, Proposition 5.5.11]. □

Theorem 4.22 suggests the following method for computing degrees of multi-parameter distortion varieties. Let L be the standard tropical linear space of codimension $r + \dim(X)$ in $\mathbb{R}^{N+1}/\mathbb{R}\mathbf{1}$, as in [74, Corollary 3.6.16]. Fix a general point γ in $\mathbb{R}^{N+1}/\mathbb{R}\mathbf{1}$. Then $\deg(X_{[u]})$ is the number of points, counted with multiplicity, in the intersection of the tropical variety (4.24) with the tropical linear space $\gamma + L$. In practice, X is fixed and we precompute $\mathrm{trop}(X)$. That fan then gets intersected with $\gamma + L + \mathrm{rowspace}(\tilde{u})$ for various configurations u.

Corollary 4.23. *The degree of $X_{[u]}$ is a piecewise-linear function in the maximal minors of $\tilde{u}$.*

Proof. The maximal minors of $\tilde{u}$ are the Plücker coodinates of the row space of $\tilde{u}$. An argument as in [20, Section 4] leads to a polyhedral chamber decomposition of the relevant Grassmannian, according to which pairs of cones in $\mathrm{trop}(X)$ and in $\gamma + L + \mathrm{rowspace}(\tilde{u})$ actually intersect. Each such intersection is a point, and its multiplicity is one of the maximal minors of $\tilde{u}$. □

Using the software Gfan [56], we precomputed the tropical varieties $\mathrm{trop}(X)$ for our four basic two-view models, namely $X = E, F, G, G'$. The results are summarized in Table 4.4.

Variety X	dim	lineality	f-vector	multiplicities
F in Example 4.1	7	4	(9, 18, 15)	1_{15}
E in Example 4.3	5	0	(591, 4506, 12588, 15102, 6498)	$2_{6426}; 4_{72}$
G in Example 4.4	6	1	(32, 213, 603, 780, 390)	$1_{336}; 2_{54}$
G' in Example 4.5	6	1	(100, 746, 2158, 2800, 1380)	$1_{800}; 2_{572}; 4_{8}$

Table 4.4: The tropical varieties in $\mathbb{R}^9/\mathbb{R}1$ associated with the two-view models.

The lineality space corresponds to a torus action on X. Its dimension is given in column 2. Modulo this space, trop(X) is a pointed fan. Column 3 records the number of i-dimensional cones for $i = 1, 2, 3, \ldots$. Each maximal cone comes with an integer multiplicity [74, Section 3.4]. These multiplicities are 1, 2 or 4 for our examples. Column 4 indicates their distribution.

4.5 Application to minimal problems

This section offers a case study for one minimal problem which has not yet been treated in the computer vision literature. We build and test an efficient Gröbner basis solver for it. Our approach follows [65, 64, 67] and applies in principle to any zero-dimensional parameterized polynomial system. This illustrates how the theory in Sections 4.2, 4.3, 4.4 ties in with practice.

We fix the distortion variety f+E+λ in Table 4.1. This is the variety $G''_{[v]}$ which lives in $\mathbb{P}^{11}$ and has dimension 7 and degree 23. We represent its defining equations by the matrix

$$\begin{pmatrix} x_{11} & x_{12} & x_{21}x_{31} + x_{22}x_{32} + x_{23}x_{33} & x_{13} & y_{13} \\ x_{21} & x_{22} & x_{11}x_{31} \quad x_{12}x_{32} \quad x_{13}x_{33} & x_{23} & y_{23} \\ x_{31} & x_{32} & 0 & x_{33} & y_{33} \end{pmatrix}. \qquad (4.25)$$

This matrix is derived by augmenting (4.13) with the y-column. The prime ideal of $G''_{[v]}$ is generated by all 3×3-minors of (4.25) and the 2×2-minors in the last two columns. The real points on this projective variety represent the relative position of two cameras, one with an unknown focal length f, and the other with an unknown radial distortion parameter λ.

Each pair (U_1, U_2) of image points gives a constraint (4.6) which translates into a linear equation (4.9) on $G''_{[v]} \cap L' \subset \mathbb{P}^{11}$. Here:

$$m^{\top} = [x_{11}, x_{12}, x_{13}, y_{13}, x_{21}, x_{22}, x_{23}, y_{23}, x_{31}, x_{32}, x_{33}, y_{33}]$$

is the vector of unknowns. Using notation above, the coefficient vector of the equation $c^\top m = 0$ is $c^\top = \left(u_2u_1, u_2v_1, u_2, u_2\|U_1\|^2, v_2u_1, v_2v_1, v_2, v_2\|U_1\|^2, u_1, v_1, 1, \|U_1\|^2\right)$.

Seven pairs determine a linear system $C\,m = 0$ where the coefficient matrix C has format 7×12. For general data, the matrix C has full rank 7. The solution set is a 5-dimensional linear subspace in $\mathbb{R}^{12}$, or, equivalently, a 4-dimensional subspace L' in $\mathbb{P}^{11}$. The intersection $G''_{[v]} \cap L'$ consists of 23 points. Our aim is to compute these fast and accurately. This is what is meant by the *minimal problem* associated with the distortion variety $G''_{[v]}$.

First build elimination template, then solve instances very fast

We shall employ the method of *automatic generation of Gröbner solvers*. This has already been applied with considerable success to a wide range of camera geometry problems in computer vision; see e.g [65, 64]. We start by computing a suitable basis $\{n_1, n_2, n_3, n_4, n_5\}$ for the null space of C in $\mathbb{R}^{12}$. We then introduce four unknowns $\lambda_1, \ldots, \lambda_4$, and we substitute

$$m = \lambda_1 n_1 + \lambda_2 n_2 + \lambda_3 n_3 + \lambda_4 n_4 + n_5. \tag{4.26}$$

Our rank constraints on (4.25) translate into ten equations in $\lambda_1, \lambda_2, \lambda_3, \lambda_4$. This system has 23 solutions in $\mathbb{C}^4$. Our aim is to compute these within a few tens or hundreds of microseconds.

Efficient and stable Gröbner solvers are often based on *Stickelberger's Theorem* [98, Theorem 2.6], which expresses the solutions as the joint eigenvalues of its companion matrices. Let $I \subset \mathbb{R}[\lambda]$ be the ideal generated by our ten polynomials in $\lambda = (\lambda_1, \lambda_2, \lambda_3, \lambda_4)$. The quotient ring $\mathbb{R}[\lambda]/I$ is isomorphic to $\mathbb{R}^{23}$. An $\mathbb{R}$-vector space basis $\mathcal{B}$ is given by the standard monomials with respect to any Gröbner basis of I. The multiplication map $M_i : \mathbb{R}[\lambda]/I \to \mathbb{R}[\lambda]/I$, $f \mapsto f \cdot \lambda_i$ is $\mathbb{R}$-linear. Using the basis $\mathcal{B}$, this becomes a 23×23-matrix. The matrices M_1, M_2, M_3, M_4 commute pairwise. These are the *companion matrices*. As an $\mathbb{R}$-algebra, $\mathbb{R}[M_1, M_2, M_3, M_4] \simeq \mathbb{R}[\lambda]/I$. Since I is radical, there are 23 linearly independent joint eigenvectors x, satisfying $M_i x = \lambda_i x$. The vectors $(\lambda_1, \lambda_2, \lambda_3, \lambda_4) \in \mathbb{C}^4$ are the zeros of I.

In practice, it suffices to construct only one of the companion matrices M_i, since we can recover the zeros of I from eigenvectors x of M_i. Thus, our primary task is to compute either M_1, M_2, M_3 or M_4 from seven point correspondences (U_1, U_2) in a manner that is both very fast and numerically stable. For this purpose, the *automatic generator* of Gröbner solvers [65, 64] is used. We now explain this method and illustrate it for the $f{+}E{+}\lambda$ problem.

To achieve speed in computation, we exploit that, for generic data, Buchberger's algorithm always rewrites the input polynomials in the same way. The resulting Gröbner trace [101] is always the same. Therefore, we can construct a single trace for all generic systems by tracing the construction of a Gröbner basis of a single "generic" system. This is done only once in an *off-line* stage of solver generation. It produces an *elimination template*, which is then reused again and again for efficient on-line computations on generic data.

The off-line part of the solver generation is a variant of the Gröbner trace algorithm in [101]. Based on the F4 algorithm [36] for a particular generic system, it produces an elimination template for constructing a Gröbner basis of $\langle F \rangle$. The input polynomial system $F = \{f_1, \ldots, f_{10}\}$ is written in the form $A\,m = 0$, where A is the matrix of coefficients and m is the vectors of monomials of the system. Every Gröbner basis G of F can be constructed by Gauss-Jordan (G-J) elimination of a coefficient matrix A_d derived from F by multiplying each polynomial $f_i \in F$, by all monomials up to degree $\max\{0, d - d_i\}$, where $d_i = \deg(f_i)$.

To find an appropriate d, our solver generator starts with $d = \min\{d_i\}$, sets $m_d = m$, and G-J eliminates the matrix $A_{\min\{d_i\}} = A$. Then, it checks if a Gröbner basis G has been generated. If not, it increases d by one, builds the next A_d and m_d, and goes back to the check. This is repeated until a suitable d and a Gröbner basis G has been found. Often, we can remove some rows (polynomials) from A_d at this stage and form a smaller elimination template, denoted A_d^0. For this, another heuristic optimization procedure is employed, aimed at removing unnecessary polynomials and provide an efficient template leading from F to the reduced coefficient matrix A_d^0. For a detailed description see [65] and [64, Section 4.4.3].

In order to guide this process, we first precompute the reduced Gröbner basis of I, e.g. w.r.t. grevlex ordering in *Macaulay2* [44], and the associated monomial basis B of $R[\;]/I$. This has to be done in exact arithmetic over $\mathbb{Q}$, which is computationally very demanding, due to the coefficient growth [8]. We alleviate this problem by using modular arithmetic [36] or by computing directly in a finite field modulo a single "lucky prime number" [101]. For many practical problems [18, 82, 94], small primes like 30011 or 30013 are sufficient.

The output of this off-line algorithm is the elimination template for constructing A_d^0, i.e. the list of monomials multiplying each polynomial of F to produce A_d^0 and m_d^0. The template is encoded as manipulations of sparse coefficient matrices. After removing unnecessary rows and columns, the matrix A_d^0 has size $s \times (s + |B|)$ for some s. The left $s \times s$-block is invertible. Multiplying A_d^0 by that inverse and extracting appropriate rows, one obtains the $|B| \times |B|$ matrix M_1 that represents the linear map $R[\;]/I \rightarrow R[\;]/I;\ f \mapsto f\,{}_1$ in the basis B.

We applied this off-line algorithm to the f+E+ problem, with standard monomial basis

$$B = (1;\ {}_1;\ {}_1\ {}_3;\ {}_1\ {}_3\ {}_4;\ {}_1\ {}_4;\ {}_1\ {}_4^2;\ {}_2;\ {}_2\ {}_3;\ {}_2\ {}_3\ {}_4;\ {}_2\ {}_4;\ {}_2\ {}_4^2;\ {}_2\ {}_4^3;\ {}_3;\ {}_3^2;\ {}_3^3;\ {}_3^2\ {}_4;\ {}_3\ {}_4;\ {}_3\ {}_4^2;\ {}_3\ {}_4^3;\ {}_4;\ {}_4^2;\ {}_4^3;\ {}_4^4).$$

Note that $jBj = 23$. The matrix (4.25) gives the following ten ideal generators (with $d_1 = d_2 = d_3 = 2;\ d_4 = d_5 = 3;\ d_6 = \quad = d_{10} = 4$) for the variety $G^{00}_{[u]}$ encoding the f+E+ problem:

$$
\begin{aligned}
f_1 &= y_{23}x_{33} \quad x_{23}y_{33}\\
f_2 &= y_{13}x_{33} \quad x_{13}y_{33}\\
f_3 &= y_{13}x_{23} \quad x_{13}y_{23}\\
f_4 &= y_{13}x_{22}x_{31} \quad x_{12}y_{23}x_{31} \quad y_{13}x_{21}x_{32} + x_{11}y_{23}x_{32} + x_{12}x_{21}y_{33} \quad x_{11}x_{22}y_{33}\\
f_5 &= x_{13}x_{22}x_{31} \quad x_{12}x_{23}x_{31} \quad x_{13}x_{21}x_{32} + x_{11}x_{23}x_{32} + x_{12}x_{21}x_{33} \quad x_{11}x_{22}x_{33}\\
f_6 &= x_{11}y_{13}x_{31}x_{32} + x_{21}y_{23}x_{31}x_{32} + x_{12}y_{13}x_{32}^2 + x_{22}y_{23}x_{32}^2 \quad x_{11}x_{12}x_{31}y_{33} \quad x_{21}x_{22}x_{31}y_{33}\\
&\quad x_{12}^2x_{32}y_{33} + x_{13}^2x_{32}y_{33} \quad x_{22}^2x_{32}y_{33} + x_{23}^2x_{32}y_{33} \quad x_{12}x_{13}x_{33}y_{33} \quad x_{22}x_{23}x_{33}y_{33}\\
\\
f_{10} &= x_{11}x_{12}x_{31}^2 + x_{21}x_{22}x_{31}^2 \quad x_{11}^2x_{31}x_{32} + x_{12}^2x_{31}x_{32} \quad x_{21}^2x_{31}x_{32} + x_{22}^2x_{31}x_{32}\\
&\quad x_{11}x_{12}x_{32}^2 \quad x_{21}x_{22}x_{32}^2 + x_{12}x_{13}x_{31}x_{33} + x_{22}x_{23}x_{31}x_{33} \quad x_{11}x_{13}x_{32}x_{33} \quad x_{21}x_{23}x_{32}x_{33}
\end{aligned}
$$

Using (4.26), these are inhomogeneous polynomials in ${}_1;\ {}_2;\ {}_3;\ {}_4$. In the o -line algorithm, we multiply f_i by all monomials up to degree $5\ d_i$ in these four variables. Each of $f_1; f_2; f_3$ is multiplied by the 35 monomials of degree 3, each of $f_4; f_5$ is multiplied by the 15 monomials of degree 2, and each of $f_6; \ldots; f_{10}$ is multiplied by the 5 monomials of degree 1. The resulting $160 = 10 + 105 + 30 + 25$ polynomials are written as a matrix A_5 with 160 rows. Only 103 rows are needed to construct the matrix M_1. We conclude with an elimination template matrix A_5^0 of format 103 126. For any data C, the on-line solver performs G-J elimination on that matrix, and it computes the eigenvectors of a 23 23 matrix M_1.

To avoid coe cient growth in the on-line stage, exact computations over $\mathbb{Q}$ are replaced by approximate computations with oating point numbers in $\mathbb{R}$. In a naive implementation, expected cancellations may fail to occur due to rounding errors, thus leading to incorrect results. This is not a problem in our method because we follow the precomputed elimination template: we use only matrix entries that were non-zero in the o -line stage. Still, replacing the symbolic F4 algorithm with a numerical computation may lead to very unstable behavior.

It has been observed [15] that di erent formulations, term orderings, pair selection strategies, etc., can have a dramatic e ect on the stability and speed of the nal solver. It is hence crucial to validate every solver experimentally, by simulations as well as on real data.

Computational results

A complete solution, in the engineering sense, to a minimal problem is a solution that is: 1) fast and 2) numerically stable for most of the data that occur in practice. Moreover, for applications it is important to study the distribution of real solutions of the minimal solver.

Minimal solvers are often used inside RANSAC style loops [37]. They form parts of much larger systems, such as structure-from-motion and 3D reconstruction pipelines or localization systems. Maximizing the efficiency of these solvers is an essential task. Inside a RANSAC loop, all real zeros returned by the solver are seen as possible solutions to the problem. The consistency w.r.t. all measurements is tested for each of them. Since that test may be computationally expensive, the study of the distribution of real solutions is important.

In this section we present graphs and statistics that display properties of the complete solution we offer for the f+E+ problem. We studied the performance of our Gröbner solver on synthetically generated 3D scenes with known ground-truth parameters. We generated 500,000 different scenes with 3D points randomly distributed in a cube $[\ 10; 10]^3$ and cameras with random feasible poses. Each 3D point was projected by two cameras. The focal length f of the left camera was drawn uniformly from the interval $[0{:}5; 2{:}5]$ and the focal length of the right camera was set to 1. The orientations and positions of the cameras were selected at random so as to look at the scene from a random distance, varying from 20 to 40 from the center of the scene. Next, the image projections in the right camera were corrupted by random radial distortion, following the one-parameter division model in [38]. The radial distortion was drawn uniformly from the interval $[\ 0{:}7; 0]$. The aim was to investigate the behavior of the algorithms for large as well as small amounts of radial distortion.

Computation and its speed. The proposed f+E+ solver performs the following steps:

1. Fill the 103 126 elimination template matrix A_5^0 with coefficients derived from the input measurements.
2. Perform G-J elimination on the matrix A_5^0.
3. Extract the desired coefficients from the eliminated matrix.
4. Create the multiplication matrix from extracted coefficients.
5. Compute the eigenvectors of the multiplication matrix.
6. Extract 23 complex solutions $({}_1;\ {}_2;\ {}_3;\ {}_4)$ from the eigenvectors.

7. For each real solution $(_1; \ _2; \ _3; \ _4)$, recover the monomial vector m as in (4.26), the fundamental matrix F, the focal length f, and the radial distortion .

All seven steps were implemented e ciently. The nal f+E+ solver runs in less than 1ms.

(a) (b)

Figure 4.1: Numerical stability. (a) Log_{10} of the relative error of the estimated radial distortion. (b) Log_{10} of the relative error of the estimated focal length.

Numerical stability. We studied the behavior of our solver on noise-free data. Figure 4.1(a) shows the experimental frequency of the base 10 logarithm of the relative error of the radial distortion parameter estimated using the new f+E+ solver. These result were obtained by selecting the real roots closest to the ground truth values. The results suggest that the solver delivers correct solutions and its numerical stability is suitable for real word applications.

Figure 4.1(b) shows the distribution of Log_{10} of the relative error of the estimated focal length f. Again these result were obtained by selecting the real roots closest to the ground truth values. Note that the f+E+ solver does not directly compute the focal length f. Its output is the monomial vector in m (4.26), from which we extract and the fundamental matrix $F = (x_{ij})$. To obtain the unknown focal length from F, we use the following formula:

Lemma 4.24. Let $X = (x_{ij})_{1 \ i;j \ 3}$ be a generic point in the variety G^{00} from Example 4.6. Then there are exactly two pairs of essential matrix and focal length $(E; f)$ such that $X = \mathrm{diag}(f \ ^1; f \ ^1; 1)E$. If one of them is $(E; f)$ then the other is (diag(1; 1; 1)E; f). In particular, f is determined up to sign by X. A formula to recover f from X is as follows:

$$f^2 = \frac{x_{23}x_{31}^2 + x_{23}x_{32}^2 \quad 2x_{21}x_{31}x_{33} \quad 2x_{22}x_{32}x_{33} \quad x_{23}x_{33}^2}{2x_{11}x_{13}x_{21} + 2x_{12}x_{13}x_{22} \quad x_{11}^2x_{23} \quad x_{12}^2x_{23} + x_{13}^2x_{23} + x_{21}^2x_{23} + x_{22}^2x_{23} + x_{23}^3} : \quad (4.27)$$

Proof. Consider the map $E \quad C \ ! \quad P^8$, $(E; f)$ 7! diag($f^{1}; f^{1}; 1)E$. Let I $Q[e_{ij}; f; x_{ij}]$ be the ideal of the graph of this map. So I is generated by the ten Demazure cubics and the nine entries of X diag($f^{1}; f^{1}; 1)E$. We computed the elimination ideal $I \setminus Q[f; x_{ij}]$ in Macaulay2. The polynomial gotten by clearing the denominator and subtracting the RHS from the LHS in the formula (4.27) lies in this elimination ideal. This proves the lemma. □

(a) (b)

Figure 4.2: Number of real solutions for oating point computation with noise-free image data.

Counting real solutions. In the next experiment we studied the distribution of the number of real solutions (; F) and the number of real solutions for the focal length f.

Figure 4.2 (a) shows the histogram of the number of real solutions on the distortion variety $G^{00}_{[v]}$. All odd integers between 1 and 23 were observed. Most of the time we got an odd number of real solutions between 7 and 15. The empirical probabilities are in Table 4.5.

Figure 4.2 (b) shows the histogram of the number of solutions for the focal length f, computed from the distortion variety $G^{00}_{[v]}$ using the formula (4.27). Of the 46

real roots in $G^{00}_{[v]}$	1	3	5	7	9	11	13	15	17	19	21	23
%	0.003	0.276	2.47	9.50	21.0	28.0	22.8	11.5	3.60	0.681	0.078	0.003

Table 4.5: Percentage of the number of real solutions in the distortion variety $G^{00}_{[v]}$.

complex solutions, at most 23 could be real and positive. The largest number of positive real solutions f observed in in 500,000 runs was 16. The empirical probabilities from this experiment are in Table 4.6.

real f	0	1	2	3	4	5	6	7	8	9	10	11
%	0.003	0.397	3.16	7.93	14.5	18.8	19.9	15.5	10.5	5.54	2.52	0.894
real f	12	13	14	15	16							
%	0.295	0.075	0.023	0.005	0.001							

Table 4.6: Percentage of the number of positive real roots for the focal length f.

We performed the same experiment with image measurements corrupted by Gaussian noise with the standard deviation set to 2 pixels. The distribution of the real roots in the distortion variety $G^{00}_{[v]}$ was very similar to the distribution for noise-free data. The main difference between these result and those for noise-free data was in the number of real values for the focal length f. For a fundamental matrix corrupted by noise, the formula (4.27) results in no real solutions more often. See Tables 4.7 and 4.8 for the empirical probabilities.

real roots	1	3	5	7	9	11	13	15	17	19	21	23
%	0.021	0.509	3.23	11.2	22.4	27.7	21.1	10.1	3.07	0.566	0.062	0.004

Table 4.7: Percentage of the number of real solutions in the distortion variety $G^{00}_{[v]}$ for image measurements corrupted with Gaussian noise with $= 2$ pixels.

real f	0	1	2	3	4	5	6	7	8	9	10	11
%	0.243	1.30	4.92	10.2	16.1	19.0	18.5	13.7	8.79	4.33	1.96	0.689
real f	12	13	14	15	16							
%	0.217	0.048	0.015	0.002	0.001							

Table 4.8: Percentage of the number of real roots for the focal length f with data as in Table 4.7.

Finally, we performed the same experiments for a special camera motion. It is known [81, 96] that the focal length cannot be determined by the formula (4.27) from the fundamental matrix if the optical axes are parallel to each other, e.g. for a sideways motion of cameras. Therefore, we generated cameras undergoing "close-to-sideways motion". To model this scenario, 100 points were again placed in a 3D cube $[-10, 10]^3$. Then 500,000 different camera pairs were generated such that both cameras were first pointed in the same direction (optical axes were intersecting at infinity) and then translated laterally. Next, a small amount of rotational noise of 0.01 degrees was introduced into the camera poses by right-multiplying the projection matrices by respective rotation matrices. This multiplication slightly rotated the optical axes of cameras (as not to intersect at infinity) as well as simultaneously displaced the camera centers.

The results for noise-free data are displayed in Tables 4.9 and 4.10. For this special close-to-sideways motion, the formula (4.27) provides up to 20 real solutions for the focal length f.

real roots	1	3	5	7	9	11	13	15	17	19	21	23
%	0.007	0.544	5.14	16.83	26.2	24.9	16.2	7.37	2.30	0.475	0.061	0.006

Table 4.9: Real solutions in the distortion variety $G^{00}_{[v]}$ for the close-to-sideways motion scenario.

real f	0	1	2	3	4	5	6	7	8	9	10
%	0.006	0.755	3.08	10.2	12.9	20.9	16.2	16.0	8.73	6.17	2.61
real f	11	12	13	14	15	16	17	18	19	20	
%	1.58	0.556	0.253	0.086	0.033	0.011	0.0044	0.0016	0.0012	0.0002	

Table 4.10: Real solutions for the focal length f in the close-to-sideways motion scenario.

Example 4.25. In [67], Kukelova, Pajdla, Sturmfels and I apply a similar elimination strategy inspired by distortion varieties to derive new minimal solvers for problems with solvers already, for purposes of comparison. In particular, see [67, Section 3.3] for a new solver for the case E+f from Table 4.1, corresponding to the distortion variety G^0 in $\mathbb{P}^{11}$ with dimension 7 and degree 19. It is shown that our solver compares favorably to the state of the art (SOTA) solver due to Kuang et al. [63]. Our solver's elimination template has size 51×70, while SOTA's elimination template is 200×231. The smaller solver is faster and moreover it has competitive numerical stability properties. See [67, Figure 3] for details.

In this chapter, we presented a mathematical theory for describing distortion in images. It is based on lifting varieties in projective space to other toric varieties. The framework uni es existing models in vision, and leads to fast minimal solvers for cases with distortion. Our theorems about degree, de ning equations and tropicalization are of independent interest in combinatorial algebraic geometry.

Chapter 5

Modeling Spaces of Pictures

In this chapter, we model spaces of pictures of simple objects, such as edges. Here cameras are fixed, a world object varies in position and we are interested in its space of possible simultaneous pictures. Our understanding could enhance triangulation algorithms. Our approach is to use combinatorial commutative algebra; in the simplest case, we consider subvarieties of products of the projective plane. This chapter is based on my publication [58] in the *International Journal of Algebra and Computation* 26 (2016) joint with Michael Joswig, Bernd Sturmfels and André Wagner.

5.1 Introduction

The emerging field of Algebraic Vision is concerned with interactions between computer vision and algebraic geometry. A central role in this endeavor is played by projective varieties that arise in multiview geometry [48].

The set-up is as follows: A *camera* is a linear map from the three-dimensional projective space $\mathbb{P}^3$ to the projective plane $\mathbb{P}^2$, both over $\mathbb{R}$. We represent n cameras by matrices $A_1, A_2, \dots, A_n \in \mathbb{R}^{3\times 4}$ of rank 3. The kernel of A_j is the focal point $f_j \in \mathbb{P}^3$. Each image point $u_j \in \mathbb{P}^2$ of camera A_j has a line through f_j as its fiber in $\mathbb{P}^3$. This is the *back-projected line*.

We assume throughout that the focal points of the n cameras are in *general position*, i.e. all distinct, no three on a line, and no four on a plane. Let ℓ_{jk} denote the line in $\mathbb{P}^3$ spanned by the focal points f_j and f_k. This is the *baseline* of the camera pair A_j, A_k. The image of the focal point f_j in the image plane $\mathbb{P}^2$ of the camera A_k is the *epipole* $e_{k\leftarrow j}$. Note that the baseline ℓ_{jk} is the back-projected line of $e_{k\leftarrow j}$ with respect to A_j and also the back-projected line of $e_{j\leftarrow k}$ with respect to A_k. See Figure 5.1 for a sketch.

Figure 5.1: Two-view geometry (cf. Chapter 2).

Fix a point X in $\mathbb{P}^3$ which is not on the baseline β_{jk}, and let u_j and u_k be the images of X under A_j and A_k. Since X is not on the baseline, neither image point is the epipole for the other camera. The two back-projected lines of u_j and u_k meet in a unique point, which is X. This process of reconstructing X from two images u_j and u_k is called *triangulation* [48, §9.1].

The triangulation procedure amounts to solving the linear equations

$$B^{jk} \begin{bmatrix} X \\ -\lambda_j \\ -\lambda_k \end{bmatrix} = 0 \quad \text{where} \quad B^{jk} = \begin{bmatrix} A_j & u_j & 0 \\ A_k & 0 & u_k \end{bmatrix} \in \mathbb{R}^{6 \times 6}. \tag{5.1}$$

For general data we have $\mathrm{rank}(B^{jk}) = \mathrm{rank}(B^{jk}_1) = \cdots = \mathrm{rank}(B^{jk}_6) = 5$, where B^{jk}_i is obtained from B^{jk} by deleting the ith row. Cramer's Rule can be used to recover X. Let $\wedge_5 B^{jk}_i \in \mathbb{R}^6$ be the column vector formed by the signed maximal minors of B^{jk}_i. Write $\tilde{\wedge}_5 B^{jk}_i \in \mathbb{R}^4$ for the first four coordinates of $\wedge_5 B^{jk}_i$. These are bilinear functions of u_j and u_k. They yield

$$X = \tilde{\wedge}_5 B^{jk}_1 = \tilde{\wedge}_5 B^{jk}_2 = \cdots = \tilde{\wedge}_5 B^{jk}_6. \tag{5.2}$$

We note that, in most practical applications, the data $u_1, \ldots, u_n$ will be noisy, in which case triangulation requires techniques from optimization [3].

The multiview variety V_A of the camera configuration $A = (A_1, \ldots, A_n)$ was defined in [5] as the closure of the image of the rational map

$$\begin{aligned} \phi_A : \mathbb{P}^3 &\dashrightarrow \mathbb{P}^2 \times \mathbb{P}^2 \times \cdots \times \mathbb{P}^2, \\ X &\mapsto (A_1 X, A_2 X, \ldots, A_n X). \end{aligned} \tag{5.3}$$

The points $(u_1, u_2, \ldots, u_n) \in V_A$ are the consistent views in n cameras. The prime ideal I_A of V_A was determined in [5, Corollary 2.7]. It is generated by the $\binom{n}{2}$ bilinear polynomials $\det(B^{jk})$ plus $\binom{n}{3}$ further trilinear polynomials. See [72] for the natural generalization of this variety to higher dimensions.

The analysis in [5] was restricted to a single world point $X \in \mathbb{P}^3$ (cf. Definition 3.9). In this chapter we study the case of two world points $X, Y \in \mathbb{P}^3$ that are linked by a distance constraint. Consider the hypersurface $V(Q)$ in $\mathbb{P}^3 \times \mathbb{P}^3$ defined by

$$Q = (X_0Y_3 - Y_0X_3)^2 + (X_1Y_3 - Y_1X_3)^2 + (X_2Y_3 - Y_2X_3)^2 - X_3^2Y_3^2. \tag{5.4}$$

The affine variety $V_{\mathbb{R}}(Q) \cap \{X_3 = Y_3 = 1\}$ in $\mathbb{R}^3 \times \mathbb{R}^3$ consists of pairs of points whose Euclidean distance is 1. The *rigid multiview map* is the rational map

$$\begin{array}{llllll} \psi_A : & V(Q) & \hookrightarrow & \mathbb{P}^3 \times \mathbb{P}^3 & \dashrightarrow & (\mathbb{P}^2)^n \times (\mathbb{P}^2)^n, \\ & (X, Y) & & & \mapsto & \big((A_1X, \ldots A_nX), (A_1Y, \ldots A_nY)\big). \end{array} \tag{5.5}$$

The *rigid multiview variety* is the image of this map. This is a 5-dimensional subvariety of $(\mathbb{P}^2)^{2n}$. Its multihomogeneous prime ideal J_A lives in the polynomial ring $\mathbb{R}[u, v] = \mathbb{R}[u_{i0}, u_{i1}, u_{i2}, v_{i0}, v_{i1}, v_{i2} : i = 1, \ldots, n]$, where $(u_{i0}{:}u_{i1}{:}u_{i2})$ and $(v_{i0}{:}v_{i1}{:}v_{i2})$ are coordinates for the ith factor $\mathbb{P}^2$ on the left respectively right in $(\mathbb{P}^2)^n \times (\mathbb{P}^2)^n$. Our aim is to determine the ideal J_A. Knowing generators of J_A has the potential of being useful for designing optimization tools as in [3] for triangulation in the presence of distance constraints.

The choice of world and image coordinates for the camera configuration $A = (A_1, \ldots, A_n)$ gives our problem the following group symmetries. Let N be an element of the *Euclidean group of motions* $\mathrm{SE}(3, \mathbb{R})$, which is generated by rotations and translations. We may multiply the camera configuration on the right by N to obtain $AN = (A_1N, \ldots, A_nN)$. Then $J_A = J_{AN}$ since $V(Q)$ is invariant under $\mathrm{SE}(3, \mathbb{R})$. For $M_1, \ldots, M_n \in \mathrm{GL}(3, \mathbb{R})$, we may multiply A on the left to obtain $A' = (M_1A, \ldots, M_nA)$. Then $J_{A'} = (M_1 \otimes \ldots \otimes M_n)J_A$.

This chapter is organized as follows. In Section 5.2 we present the explicit computation of the rigid multiview ideal for $n = 2, 3, 4$. Our main result, to be stated and proved in Section 5.3, is a system of equations that cuts out the rigid multiview variety $V(J_A)$ for any n. Section 5.4 is devoted to generalizations. The general idea is to replace $V(Q)$ by arbitrary subvarieties of $(\mathbb{P}^3)^m$ that represent polynomial constraints on $m \geq 2$ world points. We focus on scenarios that are of interest in applications to computer vision.

Our results in Propositions 5.1, 5.2, 5.3 and Corollary 5.1 are proved by computations with `Macaulay2` [44]. Following standard practice in computational algebraic geometry, we carry out the computation on many samples in a Zariski dense set of parameters, and then conclude that it holds generically.

5.2 Two, three and four pictures

In this section we o er a detailed case study of the rigid multiview variety when the number n of cameras is small. We begin with the case $n = 2$. The prime ideal J_A lives in the polynomial ring $\mathbb{R}[u, v]$ in 12 variables. This is the homogeneous coordinate ring of $(\mathbb{P}^2)^4$, so it is naturally $\mathbb{Z}^4$-graded. The variables u_{10}, u_{11}, u_{12} have degree $(1, 0, 0, 0)$, the variables u_{20}, u_{21}, u_{22} have degree $(0, 1, 0, 0)$, the variables v_{10}, v_{11}, v_{12} have degree $(0, 0, 1, 0)$, and the variables v_{20}, v_{21}, v_{22} have degree $(0, 0, 0, 1)$. Our ideal J_A is $\mathbb{Z}^4$-homogeneous.

Throughout this section we shall assume that the camera con guration A is *generic* in the sense of algebraic geometry. This means that A lies in the complement of a certain (unknown) proper algebraic subvariety in the a ne space of all n-tuples of 3×4-matrices. All our results in Section 5.2 were obtained by symbolic computations with several random choices of A. Such choices of camera matrices are generic. They will be attained with with probability 1.

Proposition 5.1. *For $n = 2$, the rigid multiview ideal J_A is minimally generated by eleven $\mathbb{Z}^4$-homogeneous polynomials in twelve variables, one of degree $(1, 1, 0, 0)$, one of degree $(0, 0, 1, 1)$, and nine of degree $(2, 2, 2, 2)$.*

Let us look at the result in more detail. The rst two bilinear generators are the familiar 6×6-determinants

$$\det \begin{bmatrix} A_1 & u_1 & 0 \\ A_2 & 0 & u_2 \end{bmatrix} \qquad \text{and} \qquad \det \begin{bmatrix} A_1 & v_1 & 0 \\ A_2 & 0 & v_2 \end{bmatrix}. \tag{5.6}$$

These cut out two copies of the multiview threefold $V_A \subset (\mathbb{P}^2)^2$, in separate variables, for $X \mapsto u = (u_1, u_2)$ and $Y \mapsto v = (v_1, v_2)$. If we write the two bilinear forms in (5.6) as $u_1^\top F u_2$ and $v_1^\top F v_2$ then F is a real 3×3-matrix of rank 2, known as the *fundamental matrix* [48, Chapter 9] of the camera pair (A_1, A_2).

The rigid multiview variety $V(J_A)$ is a divisor in $V_A \times V_A \subset (\mathbb{P}^2)^2 \times (\mathbb{P}^2)^2$. The nine octics that cut out this divisor can be understood as follows. We write B and C for the 6×6-matrices in (5.6), and B_i and C_i for the matrices obtained by deleting their ith rows. The kernels of these 5×6-matrices are represented, via Cramer's Rule, by $\wedge_5 B_i$ and $\wedge_5 C_i$. We write $\widetilde{\wedge}_5 B_i$ and $\widetilde{\wedge}_5 C_i$ for the vectors given by their rst four entries. As in (5.2), these represent the two world points X and Y in $\mathbb{P}^3$.

Their coordinates are bilinear forms in (u_1, u_2) or (v_1, v_2), where each coe cient is a 3 × 3-minor of $\begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$. For instance, writing a_i^{jk} for the (j, k) entry of A_i, the rst coordinate of $\widetilde{\wedge}_5 B_1$ is

$$
\begin{aligned}
&-(a_1^{32}a_2^{23}a_2^{34} - a_1^{32}a_2^{24}a_2^{33} - a_1^{33}a_2^{22}a_2^{34} + a_1^{33}a_2^{24}a_2^{32} + a_1^{34}a_2^{22}a_2^{33} - a_1^{34}a_2^{23}a_2^{32})\, u_{11}u_{20} \\
&+(a_1^{32}a_2^{13}a_2^{34} - a_1^{32}a_2^{14}a_2^{33} - a_1^{33}a_2^{12}a_2^{34} + a_1^{33}a_2^{14}a_2^{32} + a_1^{34}a_2^{12}a_2^{33} - a_1^{34}a_2^{13}a_2^{32})\, u_{11}u_{21} \\
&-(a_1^{32}a_2^{13}a_2^{24} - a_1^{32}a_2^{14}a_2^{23} - a_1^{33}a_2^{12}a_2^{24} + a_1^{33}a_2^{14}a_2^{22} + a_1^{34}a_2^{12}a_2^{23} - a_1^{34}a_2^{13}a_2^{22})\, u_{11}u_{22} \\
&+(a_1^{22}a_2^{23}a_2^{34} - a_1^{22}a_2^{24}a_2^{33} - a_1^{23}a_2^{22}a_2^{34} + a_1^{23}a_2^{24}a_2^{32} + a_1^{24}a_2^{22}a_2^{33} - a_1^{24}a_2^{23}a_2^{32})\, u_{12}u_{20} \\
&-(a_1^{22}a_2^{13}a_2^{34} - a_1^{22}a_2^{14}a_2^{33} - a_1^{23}a_2^{12}a_2^{34} + a_1^{23}a_2^{14}a_2^{32} + a_1^{24}a_2^{12}a_2^{33} - a_1^{24}a_2^{13}a_2^{32})\, u_{12}u_{21} \\
&+(a_1^{22}a_2^{13}a_2^{24} - a_1^{22}a_2^{14}a_2^{23} - a_1^{23}a_2^{12}a_2^{24} + a_1^{23}a_2^{14}a_2^{22} + a_1^{24}a_2^{12}a_2^{23} - a_1^{24}a_2^{13}a_2^{22})\, u_{12}u_{22}.
\end{aligned}
$$

Recall that the two world points in $\mathbb{P}^3$ are linked by a distance constraint (5.4), expressed as a biquadratic polynomial Q. We set $Q(X, Y) = T(X, X, Y, Y)$, where $T(\bullet, \bullet, \bullet, \bullet)$ is a quadrilinear form. We regard T as a tensor of order 4. It lives in the subspace $\mathrm{Sym}_2(\mathbb{R}^4) \otimes \mathrm{Sym}_2(\mathbb{R}^4) \simeq \mathbb{R}^{100}$ of $(\mathbb{R}^4)^{\otimes 4} \simeq \mathbb{R}^{256}$. Here $\mathrm{Sym}_k(\,\cdot\,)$ denotes the space of symmetric tensors of order k.

We now substitute our Cramer's Rule formulas for X and Y into the quadrilinear form T. For any choice of indices $1 \le i \le j \le 6$ and $1 \le k \le l \le 6$,

$$
T\big(\, \widetilde{\wedge}_5 B_i \,,\, \widetilde{\wedge}_5 B_j \,,\, \widetilde{\wedge}_5 C_k \,,\, \widetilde{\wedge}_5 C_l \,\big) \tag{5.7}
$$

is a multihomogeneous polynomial in (u_1, u_2, v_1, v_2) of degree $(2, 2, 2, 2)$. This polynomial lies in J_A but not in the ideal $I_A(u) + I_A(v)$ of $V_A \times V_A$, so it can serve as one of the nine minimal generators described in Proposition 5.1.

The number of distinct polynomials appearing in (5.7) equals $\binom{7}{2}^2 = 441$. A computation veri es that these polynomials span a real vector space of dimension 126. The image of that vector space modulo the degree $(2, 2, 2, 2)$ component of the ideal $I_A(u) + I_A(v)$ has dimension 9.

We record three more features of the rigid multiview with $n = 2$ cameras. The rst is the *multidegree* [79, Section 8.5], or, equivalently, the cohomology class of $V(J_A)$ in $H^*\big((\mathbb{P}^2)^4, \mathbb{Z}\big) = \mathbb{Z}[u_1, u_2, v_1, v_2]/\langle u_1^3, u_2^3, v_1^3, v_2^3\rangle$. It equals

$$
\begin{aligned}
&2u_1^2v_1 + 2u_1u_2v_1 + 2u_2^2v_1 + 2u_1^2v_2 + 2u_1u_2v_2 + 2u_2^2v_2 \\
&+2u_1v_1^2 + 2u_1v_1v_2 + 2u_1v_2^2 + 2u_2v_1^2 + 2u_2v_1v_2 + 2u_2v_2^2.
\end{aligned}
$$

This is found with the built-in command `multidegree` in `Macaulay2`.

The second is the table of the Betti numbers of the minimal free resolution of J_A in the format of `Macaulay2` [44]. In that format, the columns correspond to the syzygy modules, while rows denote the degrees. For $n = 2$ we obtain

	0	1	2	3	4	5	6	7	8	9	10	11
total:	1	177	1432	5128	10584	13951	12315	7410	3018	801	126	9
0:	1	.	.	.	.	.	.	.	.	.	.	.
1:	.	6	.	.	.	.	.	.	.	.	.	.
2:	.	2	21	6	.	.	.	.	.	.	.	.
3:	.	.	6	36	18	.	.	.	.	.	.	.
4:	.	.	1	12	42	36	9	.	.	.	.	.
5:	.	1	.	.	.	.	.	.	.	.	.	.
6:	.	24	108	166	120	42	6	.	.	.	.	.
7:	.	144	1296	4908	10404	13873	12300	7410	3018	801	126	9

Table 5.1: Betti numbers for the rigid multiview ideal with $n = 3$.

	0	1	2	3	4	5
total:	1	11	25	22	8	1
0:	1	.	.	.	.	.
1:	.	2	.	.	.	.
2:	.	.	1	.	.	.
7:	.	9	24	22	8	1

The column labeled `1` lists the minimal generators from Proposition 5.1. Since the codimension of $V(J_A)$ is 3, the table shows that J_A is not Cohen-Macaulay. The unique 5th syzygy has degree $(3,3,3,3)$ in the $\mathbb{Z}^4$-grading.

The third point is an explicit choice for the nine generators of degree $(2,2,2,2)$ in Proposition 5.1. Namely, we take $i = j \leq 3$ and $k = l \leq 3$ in (5.7). The following corollary is also found by computation:

Corollary 5.1. *The rigid multiview ideal J_A for $n = 2$ is generated by $I_A(u) + I_A(v)$ together with the nine polynomials $Q\big(\widetilde{\wedge}_5 B_i, \widetilde{\wedge}_5 C_k\big)$ for $1 \leq i, k \leq 3$.*

We next come to the case of three cameras:

Proposition 5.2. *For $n = 3$, the rigid multiview ideal J_A is minimally generated by 177 polynomials in 18 variables. Its Betti table is given in Table 5.1.*

Proposition 5.2 is proved by computation. The 177 generators occur in eight symmetry classes of multidegrees. Their numbers in these classes are

(110000) : 1	(220111) : 3	(220220) : 9	(211211) : 1
(111000) : 1	(211111) : 1	(220211) : 3	(111111) : 1

For instance, there are nine generators in degree $(2,2,0,2,2,0)$, arising from Proposition 5.1 for the first two cameras. Using various pairs among the three cameras when forming the matrices B_i, B_j, C_k and C_l in (5.7), we can construct the generators of degree classes $(2,2,0,2,1,1)$ and $(2,1,1,2,1,1)$.

Table 5.1 shows the Betti table for J_A in `Macaulay2` format. The first two entries (6 and 2) in the 1-column refer to the eight minimal generators of $I_A(u) + I_A(v)$. These are six bilinear forms, representing the three fundamental matrices, and two trilinear forms, representing the *trifocal tensor* of the three cameras (cf. Chapter 3, [4], [48, Chapter 15]). The entry 1 in row 5 of column 1 marks the unique sextic generator of J_A, which has $\mathbb{Z}^6$-degree $(1,1,1,1,1,1)$.

For the case of four cameras we obtain the following result.

Proposition 5.3. *For $n = 4$, the rigid multiview ideal J_A is minimally generated by 1176 polynomials in 24 variables. All of them are induced from $n = 3$. Up to symmetry, the degrees of the generators in the $\mathbb{Z}^8$-grading are*

(11000000) : 1	(22001110) : 3	(22002200) : 9	(21102110) : 1
(11100000) : 1	(21101110) : 1	(22002110) : 3	(11101110) : 1

We next give a brief explanation of how the rigid multiview ideals J_A were computed with `Macaulay2` [44]. For the purpose of efficiency, we introduce projective coordinates for the image points and affine coordinates for the world points. We work in the corresponding polynomial ring

$$\mathbb{Q}[u,v][X_0,X_1,X_2,Y_0,Y_1,Y_2].$$

The rigid multiview map ψ_A is thus restricted to $\mathbb{R}^3 \times \mathbb{R}^3$. The prime ideal of its graph is generated by the following two classes of polynomials:

1. the 2×2 minors of the 3×2 matrices

$$\left[\ A_i \cdot (X_0,X_1,X_2,1)^\top \ \middle|\ u_i\ \right], \quad \left[\ A_i \cdot (Y_0,Y_1,Y_2,1)^\top \ \middle|\ v_i\ \right],$$

2. the dehomogenized distance constraint

$$Q\big((X_0,X_1,X_2,1)^\top,(Y_0,Y_1,Y_2,1)^\top\big).$$

From this ideal we eliminate the six world coordinates $\{X_0,X_1,X_2,Y_0,Y_1,Y_2\}$.

For a speed up, we exploit the group actions described in Section 5.1. We replace $A = (A_1,...,A_n)$ and $Q = Q(X,Y)$ by $A' = (M_1A_1N,...,M_nA_nN)$ and $Q' = Q(N^{-1}X,N^{-1}Y)$. Here $M_i \in \mathrm{GL}_3(\mathbb{R})$ and $N \in \mathrm{GL}_4(\mathbb{R})$ are chosen so that A'

is sparse. The modi cation to Q is needed since we generally use $N \notin \mathrm{SE}(3,\mathbb{R})$. The elimination above now computes the ideal $(M_1 \otimes \ldots \otimes M_n)J_A$, and it terminates much faster. For example, for $n = 4$, the computation took two minutes for sparse A' and more than one hour for non-sparse A. For $n = 5$, `Macaulay2` ran out of memory after 18 hours of CPU time for non-sparse A. The complete code used in this chapter can be accessed via `http://www3.math.tu-berlin.de/combi/dmg/data/rigidMulti/`.

One last question is whether the Grobner basis property in [5, Section 2] extends to the rigid case. This does not seem to be the case in general. Only in Proposition 5.1 can we choose minimal generators that form a Grobner basis.

Remark 5.4. Let $n = 2$. The reduced Grobner basis of J_A in the reverse lexicographic term order is a minimal generating set. For a generic choice of cameras the initial ideal equals

$$\begin{aligned}\mathrm{in}(J_A) \;=\; \langle\, & u_{10}u_{20},\ v_{10}v_{20},\ u_{10}^2u_{21}^2v_{10}^2v_{21}^2,\ u_{10}^2u_{21}^2v_{11}^2v_{20}v_{21},\ u_{10}^2u_{21}^2v_{11}^2v_{20}^2,\\ & u_{11}^2u_{20}^2v_{10}^2v_{21}^2,\ u_{11}^2u_{20}u_{21}v_{10}^2v_{21}^2,\ u_{11}^2u_{20}^2v_{11}^2v_{20}v_{21},\\ & u_{11}^2u_{20}^2v_{11}^2v_{20}^2,\ u_{11}^2u_{20}u_{21}v_{11}^2v_{20}v_{21},\ u_{11}^2u_{20}u_{21}v_{11}^2v_{20}^2\,\rangle.\end{aligned}$$

For special cameras the exact form of the initial ideal may change. However, up to symmetry the degrees of the generators in the $\mathbb{Z}^4$-grading stay the same. In general, a universal Grobner basis for the rigid multiview ideal J_A consists of octics of degree $(2,2,2,2)$ plus the two quadrics (5.6). This was veri ed using the `Gfan` [56] package in `Macaulay2`. Analogous statements do not hold for $n \geq 3$.

5.3 Equations for the rigid multiview variety

The computations presented in Section 2 suggest the following conjecture.

Conjecture 5.5. *The rigid multiview ideal J_A is minimally generated by $\frac{4}{9}n^6 - \frac{2}{3}n^5 + \frac{1}{36}n^4 + \frac{1}{2}n^3 + \frac{1}{36}n^2 - \frac{1}{3}n$ polynomials. These polynomials come from two triples of cameras, and their number per class of degrees is*

(110..000..) :	$1 \cdot 2\binom{n}{2}$	(220..111..) :	$3 \cdot 2\binom{n}{2}\binom{n}{3}$
(220..220..) :	$9 \cdot \binom{n}{2}^2$	(211..211..) :	$1 \cdot n^2\binom{n-1}{2}^2$
(111..000..) :	$1 \cdot 2\binom{n}{3}$	(211..111..) :	$1 \cdot 2n\binom{n-1}{2}\binom{n}{3}$
(220..211..) :	$3 \cdot 2n\binom{n}{2}\binom{n-1}{2}$	(111..111..) :	$1 \cdot \binom{n}{3}^2$

$n\backslash$degree	2	3	6	7	8	total	timing (s)
2	2				9	1	<1
3	6	2	1	24	144	177	14
4	12	8	16	240	900	1176	130
5	20	20	100	1200	3600	4940	24064

Table 5.2: The known minimal generators of the rigid multiview ideals, listed by total degree, for up to ve cameras. There are no minimal generators of degrees 4 or 5. Average timings (in seconds), using the speed up described above, are in the last column.

At the moment we have a computational proof only up to $n = 5$. Table 5.2 o ers a summary of the corresponding numbers of generators.

Conjecture 5.5 implies that $V(J_A)$ is set-theoretically de ned by the equations coming from triples of cameras. It turns out that, for the set-theoretic description, pairs of cameras su ce. The following is our main result:

Theorem 5.6. *Suppose that the n focal points of A are in general position in $\mathbb{P}^3$. The rigid multiview variety $V(J_A)$ is cut out as a subset of $V_A \times V_A$ by the $9\binom{n}{2}^2$ octic generators of degree class* (220..220..). *In other words, equations coming from any two pairs of cameras suffice set-theoretically.*

With notation as in the introduction, the relevant octic polynomials are

$$T\big(\,\widetilde{\wedge}_5 B_{i_1}^{j_1k_1}\,,\,\widetilde{\wedge}_5 B_{i_2}^{j_1k_1},\widetilde{\wedge}_5 C_{i_3}^{j_2k_2}\,,\,\widetilde{\wedge}_5 C_{i_4}^{j_2k_2}\,\big),$$

for all possible choices of indices. Let H_A denote the ideal generated by these polynomials in $\mathbb{R}[u,v]$, the polynomial ring in $6n$ variables. As before, we write $I_A(u)+I_A(v)$ for the prime ideal that de nes the 6-dimensional variety $V_A \times V_A$ in $(\mathbb{P}^2)^n \times (\mathbb{P}^2)^n$. It is generated by $2\binom{n}{2}$ bilinear forms and $2\binom{n}{3}$ trilinear forms, corresponding to fundamental matrices and trifocal tensors. In light of Hilbert's Nullstellensatz, Theorem 5.6 states that the radical of $H_A + I_A(u) + I_A(v)$ is equal to J_A. To prove this, we need a lemma.

A point u in the multiview variety $V_A \subset (\mathbb{P}^2)^n$ is *triangulable* if there exists a pair of indices (j,k) such that the matrix B^{jk} has rank 5. Equivalently, there exists a pair of cameras for which the unique world point X can be found by triangulation. Algebraically, this means $X = \widetilde{\wedge}_5 B_i^{jk}$ for some i.

Lemma 5.7. *All points in V_A are triangulable except for the pair of epipoles, denoted $(e_{1\leftarrow 2}, e_{2\leftarrow 1})$, in the case where $n = 2$. Here, the rigid multiview variety $V(J_A)$ contains the threefolds $V_A(u) \times (e_{1\leftarrow 2}, e_{2\leftarrow 1})$ and $(e_{1\leftarrow 2}, e_{2\leftarrow 1}) \times V_A(v)$.*

Proof. Let us rst consider the case of $n = 2$ cameras. The rst claim holds because the back-projected lines of the two camera images u_1 and u_2 always span a plane in $\mathbb{P}^3$ except when $u_1 = e_{1\leftarrow 2}$ and $u_2 = e_{2\leftarrow 1}$. In that case both back-projected lines agree with the common baseline β_{12}. Alternatively, we can check algebraically that the variety de ned by the 5×5-minors of the matrix B consists of the single point $(e_{1\leftarrow 2}, e_{2\leftarrow 1})$.

For the second claim, x a generic point X in $\mathbb{P}^3$ and consider the surface

$$X^Q \;=\; \big\{ Y \in \mathbb{P}^3 \,:\, Q(X, Y) = 0 \big\}. \tag{5.8}$$

Working over $\mathbb{C}$, the baseline β_{12} is either tangent to X^Q, or it meets that quadric in exactly two points. Our assumption on the genericity of X implies that no point in the intersection $\beta_{12} \cap X^Q$ is a focal point. This gives

$$(A_1X, A_2X, A_1Y_X, A_2Y_X) \;=\; (A_1X, A_2X, e_{1\leftarrow 2}, e_{2\leftarrow 1}). \tag{5.9}$$

The point (A_1X, A_2X) lies in the multiview variety $V_A(u)$. Each generic point in $V_A(u)$ has this form for some X. Hence (5.9) proves the desired inclusion $V_A(u) \times (e_{1\leftarrow 2}, e_{2\leftarrow 1}) \subset V(J_A)$. The other inclusion $(e_{1\leftarrow 2}, e_{2\leftarrow 1}) \times V_A(v) \subset V(J_A)$ follows by switching the roles of u and v.

If there are more than two cameras then for each world point X, due to general position of the cameras, there is a pair of cameras such that X avoids the pair's baseline. This shows that each point is triangulable if $n \geq 3$. □

Proof of Theorem 5.6. It follows immediately from the de nition of the ideals in question that the following inclusion of varieties holds in $(\mathbb{P}^2)^n \times (\mathbb{P}^2)^n$:

$$V(J_A) \;\subseteq\; V\big(I_A(u) + I_A(v) + H_A\big).$$

We prove the reverse inclusion. Let (u, v) be a point in the right hand side.

Suppose that u and v are both triangulable. Then u has a unique preimage X in $\mathbb{P}^3$, determined by a single camera pair $\{A_{j_1}, A_{k_1}\}$. Likewise, v has a unique preimage Y in $\mathbb{P}^3$, also determined by a single camera pair $\{A_{j_2}, A_{k_2}\}$. There exist indices $i_1, i_2 \in \{1, 2, 3, 4, 5, 6\}$ such that

$$X = \widetilde{\wedge}_5 B_{i_1}^{j_1 k_1} \quad \text{and} \quad Y = \widetilde{\wedge}_5 C_{i_2}^{j_2 k_2}.$$

Suppose that (u,v) is not in $V(J_A)$. Then $Q(X,Y) \neq 0$. This implies

$$Q(X,Y) = T(X,X,Y,Y) = T\big(\widetilde{\wedge}_5 B_{i_1}^{j_1k_1}, \widetilde{\wedge}_5 B_{i_1}^{j_1k_1}, \widetilde{\wedge}_5 C_{i_2}^{j_2k_2}, \widetilde{\wedge}_5 C_{i_2}^{j_2k_2}\big) \neq 0,$$

and hence $(u,v) \notin V(H_A)$. This is a contradiction to our choice of (u,v).

It remains to consider the case where v is not triangulable. By Lemma 5.7, we have $n = 2$, as well as $v = (e_{1\leftarrow 2}, e_{2\leftarrow 1})$ and $(u,v) \in V(J_A)$. The case where u is not triangulable is symmetric, and this proves the theorem. □

The equations in Theorem 5.6 are fairly robust, in the sense that they work as well for many special position scenarios. However, when the cameras $A_1, A_2, \ldots, A_n$ are generic then the number $9\binom{n}{2}^2$ of octics that cut out the divisor $V(J_A)$ inside $V_A \times V_A$ can be reduced dramatically, namely to 16.

Corollary 5.8. *As a subset of the* 6*-dimensional ambient space* $V_A \times V_A$*, the* 5*-dimensional rigid multiview variety* $V(J_A)$ *is cut out by* 16 *polynomials of degree class* (220..220..)*. One choice of such polynomials is given by*

$$\begin{array}{ll} Q\big(\widetilde{\wedge}_5 B_i^{12}, \widetilde{\wedge}_5 C_k^{12}\big), & Q\big(\widetilde{\wedge}_5 B_i^{12}, \widetilde{\wedge}_5 C_k^{13}\big) \\ Q\big(\widetilde{\wedge}_5 B_i^{13}, \widetilde{\wedge}_5 C_k^{12}\big), & Q\big(\widetilde{\wedge}_5 B_i^{13}, \widetilde{\wedge}_5 C_k^{13}\big) \end{array} \qquad \textit{for all } 1 \leq i,k \leq 2.$$

Proof. First we claim that for each triangulable point u at least one of the matrices B^{12} or B^{13} has rank 5, and the same for v with C^{12} or C^{13}. We prove this by contradiction. By symmetry between u and v, we can assume that $\mathrm{rk}(B^{12}) = \mathrm{rk}(B^{13}) = 4$. Then $u_3 = e_{3\leftarrow 1}$, $u_2 = e_{2\leftarrow 1}$, and $u_1 = e_{1\leftarrow 2} = e_{1\leftarrow 3}$. However, this last equality of the two epipoles is a contradiction to the hypothesis that the focal points of the cameras A_1, A_2, A_3 are not collinear.

Next we claim that if B^{12} has rank 5 then at least one of the submatrices B_1^{12} or B_2^{12} has rank 5, and the same for B^{13}, C^{12} and C^{13}. Note that the bottom 4×6 submatrix of B^{12} has rank 4, since the first four columns are linearly independent, by genericity of A_1 and A_2. The claim follows. □

5.4 Other constraints, more points, and no labels

In this section we discuss several extensions of our results. A first observation is that there was nothing special about the constraint Q in (5.4). For instance, fix positive integers d and e, and let $Q(X,Y)$ be any irreducible polynomial that is bihomogeneous of degree (d,e). Its variety $V(Q)$ is a hypersurface of degree (d,e) in $\mathbb{P}^3 \times \mathbb{P}^3$. The following analogue to Theorem 5.6 holds, if we define the map ψ_A as in (5.5).

Theorem 5.9. *The closure of the image of the map ψ_A is cut out in $V_A \times V_A$ by $9\binom{n}{2}^2$ polynomials of degree class $(d, d, 0, \ldots, e, e, 0, \ldots)$. In other words, the equations coming from any two pairs of cameras suffice set-theoretically.*

Proof. The tensor T that represents Q now lives in $\mathrm{Sym}_d(\mathbb{R}^4) \otimes \mathrm{Sym}_e(\mathbb{R}^4)$. The polynomial (5.7) vanishes on the image of ψ_A and has degree (d, d, e, e). The proof of Theorem 5.6 remains valid. The surface X^Q in (5.8) is irreducible of degree e in $\mathbb{P}^3$. These polynomials cut out that image inside $V_A \times V_A$. □

Remark 5.10. In the generic case, we can replace $9\binom{n}{2}^2$ by 16, as in Corollary 5.8.

Another natural generalization is to consider m world points $X_1, \ldots, X_m$ that are linked by one or several constraints in $(\mathbb{P}^3)^m$. Taking images with n cameras, we obtain a variety $V(J_A)$ which lives in $(\mathbb{P}^2)^{mn}$. For instance, if $m = 4$ and X_1, X_2, X_3, X_4 are constrained to lie on a plane in $\mathbb{P}^3$, then $Q = \det(X_1, X_2, X_3, X_4)$ and $V(J_A)$ is a variety of dimension 11 in $(\mathbb{P}^2)^{4n}$. Taking 6×6-matrices B, C, D, E as in (5.1) for the four points, we then form

$$\det\left(\widetilde{\wedge}_5 B_i, \widetilde{\wedge}_5 C_j, \widetilde{\wedge}_5 D_k, \widetilde{\wedge}_5 E_l\right) \qquad \text{for all } 1 \leq i, j, k, l \leq 6. \tag{5.10}$$

For $n = 2$ we veri ed with `Macaulay2` that the prime ideal J_A is generated by 16 of these determinants, along with the four bilinear forms for ${V_A}^4$.

Proposition 5.11. *The variety $V(J_A)$ is cut out in ${V_A}^4$ by the $16\binom{n}{2}^4$ polynomials from (5.10). In other words, the equations coming from any two pairs of cameras suffice set-theoretically.*

Proof. Each polynomial (5.10) is in J_A. The proof of Theorem 5.6 remains valid. The planes $(X_i, X_j, X_k)^Q$ intersect the baseline β_{12} in one point each. □

To continue the theme of rigidity, we may impose distance constraints on pairs of points. Fixing a nonzero distance d_{ij} between points i and j gives

$$Q_{ij} = (X_{i0}X_{j3} - X_{j0}X_{i3})^2 + (X_{i1}X_{j3} - X_{j1}X_{i3})^2 + (X_{i2}X_{j3} - X_{j2}X_{i3})^2 - d_{ij}^2 X_{i3}^2 X_{j3}^2.$$

We are interested in the image of the variety $\mathcal{V} = V(Q_{ij} : 1 \leq i < j \leq m)$ under the multiview map ψ_A that takes $(\mathbb{P}^3)^m$ to $(\mathbb{P}^2)^{mn}$. For instance, for $m = 3$, we consider the variety $\mathcal{V} = V(Q_{12}, Q_{13}, Q_{23})$ in $(\mathbb{P}^3)^3$, and we seek the equations for its image under the multiview map ψ_A into $(\mathbb{P}^2)^{3n}$. Note that $\mathcal{V}$ has dimension 6, unless we are in the collinear case. Algebraically,

$$(d_{12} + d_{13} + d_{23})(d_{12} + d_{13} - d_{23})(d_{12} - d_{13} + d_{23})(-d_{12} + d_{13} + d_{23}) \;=\; 0. \quad (5.11)$$

If this holds then $dim(\mathcal{V}) = 5$. The same argument as in Theorem 5.6 yields:

Corollary 5.12. *The rigid multiview variety $\overline{\psi_A(\mathcal{V})}$ has dimension six, unless (5.11) holds, in which case the dimension is five. It has real points if and only if d_{12}, d_{13}, d_{23} satisfy the triangle inequality. It is cut out in ${V_A}^3$ by $27\binom{n}{2}^2$ biquadratic equations, coming from the $9\binom{n}{2}^2$ equations for any two of the three points.*

In many computer vision applications, the m world points and their images in $\mathbb{P}^2$ will be unlabeled. To study such questions, we propose to work with the *unlabeled rigid multiview variety*. This is the image of the rigid multiview variety under the quotient map $\big((\mathbb{P}^2)^m\big)^n \to \big(\mathrm{Sym}_m(\mathbb{P}^2)\big)^n$.

Indeed, while labeled configurations in the plane are points in $(\mathbb{P}^2)^m$, unlabeled configurations are points in the *Chow variety* $\mathrm{Sym}_m(\mathbb{P}^2)$. This is the variety of ternary forms that are products of m linear forms (cf. [68, §8.6]). It is embedded in the space $\mathbb{P}^{\binom{m+2}{2}-1}$ of all ternary forms of degree m.

Example 5.13. Let $m = n = 2$. The Chow variety $\mathrm{Sym}_2(\mathbb{P}^2)$ is the hypersurface in $\mathbb{P}^5$ defined by the determinant of a symmetric 3×3-matrix (a_{ij}). The quotient map $(\mathbb{P}^2)^2 \to \mathrm{Sym}_2(\mathbb{P}^2) \subset \mathbb{P}^5$ is given by the formulas

$$\begin{array}{lll} a_{00} = 2u_{10}v_{10}, & a_{11} = 2u_{11}v_{11}, & a_{22} = 2u_{12}v_{12}, \\ a_{01} = u_{11}v_{10} + u_{10}v_{11}, & a_{02} = u_{12}v_{10} + u_{10}v_{12}, & a_{12} = u_{12}v_{11} + u_{11}v_{12}. \end{array}$$

Similarly, for the two unlabeled images under the second camera we use

$$\begin{array}{lll} b_{00} = 2u_{20}v_{20}, & b_{11} = 2u_{21}v_{21}, & b_{22} = 2u_{22}v_{22}, \\ b_{01} = u_{21}v_{20} + u_{20}v_{21}, & b_{02} = u_{22}v_{20} + u_{20}v_{22}, & b_{12} = u_{22}v_{21} + u_{21}v_{22}. \end{array}$$

The unlabeled rigid multiview variety is the image of $V(J_A) \subset V_A \times V_A$ under the quotient map that takes two copies of $(\mathbb{P}^2)^2$ to two copies of $\mathrm{Sym}_2(\mathbb{P})^2 \subset \mathbb{P}^5$. This quotient map is given by $(u_1, v_1) \mapsto a,\ (u_2, v_2) \mapsto b$.

We first compute the image of $V_A \times V_A$ in $\mathbb{P}^5 \times \mathbb{P}^5$, denoted $\mathrm{Sym}_2(V_A)$. Its ideal has seven minimal generators, three of degree $(1, 1)$, and one each in degrees

$(3,0),(2,1),(1,2),(0,3)$. The generators in degrees $(3,0)$ and $(0,3)$ are $\det(a_{ij})$ and $\det(b_{ij})$. The five others depend on the cameras A_1, A_2.

Now, to get equations for the unlabeled rigid multiview variety, we intersect the ideal J_A with the subring $\mathbb{R}[a,b]$ of bisymmetric homogeneous polynomials in $\mathbb{R}[u,v]$. This results in nine new generators which represent the distance constraint. One of them is a quartic of degree $(2,2)$ in (a,b). The other eight are quintics, four of degree $(2,3)$ and four of degree $(3,2)$.

Independently of the specific constraints considered in this chapter, it is of interest to characterize the pictures of m unlabeled points using n cameras. This gives rise to the *unlabeled multiview variety* $\mathrm{Sym}_m(V_A)$ in $\left(\mathbb{P}^{\binom{m+2}{2}-1}\right)^n$. It would be desirable to know the prime ideal of $\mathrm{Sym}_m(V_A)$ for any n and m.

In this chapter, we modeled spaces of pictures of simple objects, using subvarieties of products of the projective plane. We determined defining equations that cut these subvarieties out, and we proposed various scenarios of practical interest. Our results might be helpful in polynomial optimization schemes for triangulation, following [3].

Bibliography

[1] S. Agarwal, H.-L. Lee, B. Sturmfels, and R. Thomas. Certifying the existence of epipolar matrices. *Int. J. Comput. Vision*, 121:403{415, 2017.

[2] S. Agarwal, N. Snavely, I. Simon, S.M. Seitz, and R. Szeliski. Building Rome in a day. *Proc. Int. Conf. on Comput. Vision*, pages 72{79, 2009.

[3] C. Aholt, S. Agarwal, and R. Thomas. A QCQP approach to triangulation. *Proc. European Conf. Comput. Vision; Lecture Notes in Computer Science*, 7572:654{667, 2012.

[4] C. Aholt and L. Oeding. The ideal of the trifocal variety. *Math. Comput.*, 83:2553{2574, 2014.

[5] C. Aholt, B. Sturmfels, and R. Thomas. A Hilbert scheme in computer vision. *Can. J. Math.*, 65:961{988, 2013.

[6] J. Alexander and A. Hirschowitz. Polynomial interpolation in several variables. *J. Alg. Geom.*, 4(2):201{222, 1995.

[7] M. Aprodu, G. Farkas, and A. Ortega. Minimal resolutions, Chow forms and Ulrich bundles on K3 surfaces. *J. Reine Angew. Math.*, to appear.

[8] E. Arnold. Modular algorithms for computing Gröbner bases. *J. Symbolic Comput.*, 35:403{419, 2003.

[9] D.J. Bates, E. Gross, A Leykin, and J.I. Rodriguez. Bertini for Macaulay2. *arXiv:1310.3297v1*.

[10] D.J. Bates, J.D. Hauenstein, A.J. Sommese, and C.W. Wampler. Bertini: Software for numerical algebraic geometry. *Available at* `http://bertini.nd.edu`.

[11] D.J. Bates, J.D. Hauenstein, A.J. Sommese, and C.W. Wampler. *Numerically solving polynomial systems with Bertini; Software, environments, and tools, vol. 25*. Society for Industrial and Applied Mathematics, Philadelphia, 2013.

[12] C. Bocci, E. Carlini, and J. Kileel. Hadamard products of linear spaces. *J. Algebra*, 448:595{617, 2016.

[13] J. Bochnak, M. Coste, and M.-F. Roy. *Real algebraic geometry; A series of modern surveys in mathematics, vol. 36*. Springer-Verlag, Berlin, 1998.

[14] A. Bovik. *Handbook of image and video processing, 2nd ed.* Academic Press, San Diego, 2005.

[15] M. Bujnak. *Algebraic solutions to absolute pose problems*. Doctoral Thesis, Czech Technical University in Prague, 2012.

[16] M. Bujnak, Z. Kukelova, and T. Pajdla. 3D reconstruction from image collections with a single known focal length. *Proc. Int. Conf. on Comput. Vision*, pages 351{358, 2009.

[17] M. Bujnak, Z. Kukelova, and T. Pajdla. Making minimal solvers fast. *Proc. IEEE Conf. Comput. Vision Pattern Recog.*, pages 1506{1513, 2012.

[18] M. Byrod, Z. Kukelova, K. Josephson, T. Pajdla, and K. Astrom. Fast and robust numerical solutions to minimal problems for cameras with radial distortion. *Proc. IEEE Conf. Comput. Vision Pattern Recog.*, pages 1{8, 2008.

[19] C. Carre and B. Leclerc. Splitting the square of a Schur function into its symmetric and antisymmetric parts. *J. Algebraic Combin.*, 4(3):201{231, 1995.

[20] E. Cattani, M. A. Cueto, A. Dickenstein, S. Di Rocco, and B. Sturmfels. Mixed discriminants. *Math. Z.*, 274:761{778, 2013.

[21] J. Chen and J. Kileel. Numerical implicitization for Macaulay2. *arXiv:1610.03034v1*.

[22] L. Chiantini and C. Ciliberto. Weakly defective varieties. *Trans. Amer. Math. Soc.*, 354:151{178, 2002.

[23] D. Cox, J. Little, and H. Schenck. *Toric varieties; Graduate studies in mathematics, vol. 124*. American Mathematical Society, Providence, 2011.

[24] J. Dalbec and B. Sturmfels. *Introduction to Chow forms; Invariant Methods in Discrete and Computational Geometry (ed. N. White), pages 37-58*. Springer, New York, 1995.

[25] M. Demazure. *Sur deux problèmes de reconstruction; Technical Report 882*. INRIA, Rocquencourt, France, 1988.

[26] J. Demmel. *Applied Numerical Linear Algebra*. Society for Industrial and Applied Mathematics, Philadelphia, 1997.

[27] S. Di Rocco. *Linear toric fibrations; Combinatorial Algebraic Geometry, Lecture Notes in Mathematics, vol. 2108, pages 119-149*. Springer, 2014.

[28] J. Draisma, E. Horobet, G. Ottaviani, B. Sturmfels, and R. Thomas. The Euclidean distance degree of an algebraic variety. *Found. Comput. Math.*, 16(1):99{149, 2016.

[29] D. Drusvyatski, H.-L. Lee, G. Ottaviani, and R. Thomas. The Euclidean distance degree of orthogonally invariant matrix varieties. *Israel J. Math.*, to appear.

[30] D. Drusvyatski, H.-L. Lee, and R. Thomas. Counting real critical points of the distance to orthogonally invariant matrix sets. *SIAM J. Matrix Anal. Appl.*, 36(3):1360{1380, 2015.

[31] D. Eisenbud. *Commutative algebra: with a view toward algebraic geometry; Graduate texts in mathematics, vol. 150*. Springer-Verlag, New York, 1995.

[32] D. Eisenbud, G. Fl ystad, and F. Schreyer. Sheaf cohomology and free resolutions over exterior algebras. *Trans. Amer. Math. Soc.*, 355(11):4397{4426, 2003.

[33] D. Eisenbud and J. Harris. *On varieties of minimal degree (a centennial account); Algebraic Geometry, Bowdoin, Proc. Sympos. Pure Math., vol. 46, part 1, 3-13*. American Mathematical Society, Providence, 1987.

[34] D. Eisenbud, F. Schreyer, and J. Weyman. Resultants and Chow forms via exterior syzygies. *J. Amer. Math. Soc.*, 16(3):537{579, 2003.

[35] O.D. Faugeras and S. Maybank. Motion from point matches: multiplicity of solutions. *Int. J. Comput. Vision*, 4(3):225{246, 1990.

[36] J.-C. Faugere. A new e cient algorithm for computing Grobner bases (F4). *J. Pure. Appl. Algebra*, 139:61{88, 1999.

[37] M. Fischler and R. Bolles. Random sample consensus: a paradigm for model tting with application to image analysis and automated cartography. *Commun. Assoc. Comp. Mach.*, 24:381{395, 1981.

[38] A. Fitzgibbon. Simultaneous linear estimation of multiple view geometry and lens distortion. *Proc. IEEE Conf. Comput. Vision Pattern Recog.*, pages 125{132, 2001.

[39] G. Fl ystad, J. Kileel, and G. Ottaviani. The Chow form of the essential variety in computer vision. *J. Symbolic Comput.*, to appear.

[40] F. Fraundorfer, L. Heng, D. Honegger, G.H. Lee, L. Meier, P. Tanskanen, and M. Pollefeys. Vision-based autonomous mapping and exploration using a quadrotor MAV. *Proc. Int. Conf. Intell. Robots Systems*, pages 4557{4564, 2012.

[41] W. Fulton and J. Harris. *Representation theory: a first course; Graduate texts in mathematics, vol. 129*. Springer-Verlag, New York, 1991.

[42] F. Galetto. Free resolutions and modules with a semisimple Lie group action. *J. Softw. Algebra Geom.*, 7:17{29, 2015.

[43] I.M. Gelfand, M.M. Kapranov, and A.V. Zelevinsky. *Discriminants, resultants and multidimensional determinants; Mathematics: theory & applications*. Birkhauser, Boston, 1994.

[44] D. Grayson and M. Stillman. Macaulay2, a software system for research in algebraic geometry. *Available at* `http://www.math.uiuc.edu/Macaulay2/`.

[45] E. Gross, S. Petrovic, and J. Verschelde. Interfacing with PHCpack. *J. Softw. Algebra Geom.*, 5:20{25, 2013.

[46] J. Harris. *Algebraic Geometry: a First Course; Graduate Texts in Mathematics, vol. 133*. Springer-Verlag, New York, 1992.

[47] J. Harris and L. Tu. On symmetric and skew-symmetric determinantal varieties. *Topology*, 23:71{84, 1984.

[48] R.I. Hartley and A. Zisserman. *Multiple view geometry in computer vision, 2nd ed.* Cambridge University Press, Cambridge, 2004.

[49] R. Hartshorne. *Algebraic geometry; Graduate texts in mathematics, vol. 52*. Springer, New York, 1977.

[50] J.D. Hauenstein and J.I. Rodriguez. Numerical irreducible decomposition of multiprojective varieties. *arXiv:1507.07069v2*.

[51] J.D. Hauenstein, J.I. Rodriguez, and F. Sottile. Numerical computation of Galois groups. *Found. Comput. Math.*, to appear.

[52] J.D. Hauenstein and A.J. Sommese. Witness sets of projections. *Appl. Math. Comput.*, 217(7):3349{3354, 2010.

[53] J.D. Hauenstein, A.J. Sommese, and C.W. Wampler. Regeneration homotopies for solving systems of polynomials. *Math. Comp.*, 80(273):345{377, 2011.

[54] A. Heyden. Tensorial properties of multiple view constraints. *Math. Meth. Appl. Sci.*, 23:169{202, 2000.

[55] S. Iwaki, G. Bonmassar, and J.W. Belliveau. Brain activities during 3-D structure perception from 2-D motion as assessed by combined MEG/fMRI techniques. *Proc. Int. Conf. Comp. Medical Engin.*, pages 1394{1399, 2007.

[56] A. Jensen. Gfan, a software system for Grobner fans and tropical varieties. *Available at* `http://home.imf.au.dk/jensen/software/gfan/gfan.html`.

[57] F. Jiang, Y. Kuang, J.E. Solem, and K. Astrom. A minimal solution to relative pose with unknown focal length and radial distortion. *Proc. Asian Conf. Comput. Vision*, pages 443{456, 2014.

[58] M. Joswig, J. Kileel, B. Sturmfels, and A. Wagner. Rigid multiview varieties. *Int. J. Algebra Comput.*, 26:775{787, 2016.

[59] F. Kahl, B. Triggs, and K. Astrom. Critical motions for auto-calibration when some intrinsic parameters can vary. *J. Math. Imagining Vis.*, 13(2):131{146, 2004.

[60] M. Kapranov, B. Sturmfels, and A. Zelevinsky. Chow polytopes and general resultants. *Duke Math. J.*, 67:189{218, 1992.

[61] J. Kileel. Minimal problems for the calibrated trifocal variety. *SIAM Appl. Alg. Geom.*, to appear.

[62] J. Kileel, Z. Kukelova, T. Pajdla, and B. Sturmfels. Distortion varieties. *Found. Comput. Math.*, to appear.

[63] Y. Kuang, J.E. Solem, F. Kahl, and K. Astrøm. Minimal solvers for relative pose with a single unknown radial distortion. *Proc. Conf. Comput. Vision Pattern Recogn.*, 67:33{40, 2014.

[64] Z. Kukelova. *Algebraic methods in computer vision*. Doctoral Thesis, Czech Technical University in Prague, 2013.

[65] Z. Kukelova, M. Bujnak, and T. Pajdla. Automatic generator of minimal problem solvers. *European Conf. Comput. Vision; Lecture Notes in Computer Science, vol. 5304*, pages 302{315, 2008.

[66] Z. Kukelova, M. Bujnak, and T. Pajdla. Real-time solution to the absolute pose problem with unknown distortion and focal length. *Proc. Int. Conf. Comput. Vision*, 2013.

[67] Z. Kukelova, J. Kileel, T. Pajdla, and B. Sturmfels. A clever elimination strategy for e cient minimal solvers. *Proc. Conf. Comput. Vision Pattern Recog.*, to appear.

[68] J.M. Landberg. *Tensors: geometry and applications; Graduate studies in mathematics, vol. 128*. American Mathematical Society, Providence, 2012.

[69] J.B. Lasserre. *An Introduction to polynomial and semi-algebraic optimization; Cambridge texts in applied mathematics, vol. 52*. Cambridge University Press, Cambridge, 2015.

[70] A. Leykin. Numerical algebraic geometry. *J. Softw. Algebra Geom.*, 3:5{10, 2011.

[71] A. Leykin, J.I. Rodriguez, and F. Sottile. Trace test. *arxiv:1608.00540v2*.

[72] B. Li. Images of rational maps of projective spaces. *Int. Math. Res. Notices*, to appear.

[73] D.G. Lowe. Object recognition from local scale-invariant features. *Proc. Int. Conf. on Comput. Vision*, pages 1150{1157, 1999.

[74] D. Maclagan and B. Sturmfels. *Introduction to Tropical Geometry; Graduate Studies in Mathematics, vol. 161*. American Mathematical Society, Providence, 2015.

[75] M. Marshall. *Positive polynomials and sums of squares; Mathematical surveys and monographs, vol. 146*. American Mathematical Society, Providence, 2008.

[76] J. Matthews. Multi-focal tensors as invariant differential forms. *arXiv:1610.04294v1*.

[77] S. Maybank. *Theory of reconstruction from image motion*. Springer, Berlin, 1993.

[78] B. Micusik and T. Pajdla. Structure from motion with wide circular field of view cameras. *IEEE T. Pattern Anal.*, 28:1135–1149, 2006.

[79] E. Miller and B. Sturmfels. *Combinatorial Commutative Algebra; Graduate Texts in Mathematics*. Springer-Verlag, New York, 2004.

[80] H. Murao, H. Kobayashi, and T. Fujise. On factorizing the symbolic U-resultant – Application of the *ddet* operator. *J. Symbolic Comput.*, 15(2):123–142, 1993.

[81] G. Newsam, D. Q. Huynh, M. Brooks, and H. P. Pan. Recovering unknown focal lengths in self-calibration: an essentially linear algorithm and degenerate configurations. *ISPRS J. Photogramm.*, XXXI(B3):575–580, 1996.

[82] D. Nister. An efficient solution to the five-point relative pose problem. *IEEE T. Pattern Anal.*, 26(6):756–770, 2004.

[83] D. Nister and F. Schaffalitzky. Four points in two or three calibrated views: theory and practice. *Int. J. Comput. Vision*, 67(2):211–231, 2006.

[84] M. Oskarsson, A. Zisserman, and K. Aström. Minimal projective reconstruction for combinations of points and lines in three views. *Image Vision Comput.*, 22(10):777–785, 2004.

[85] O. Özyesil, V. Voroninski, R. Basri, and A. Singer. A survey on structure from motion. *Acta Numer.*, 26:305–364, 2017.

[86] S. Petrovic. On the universal Gröbner bases of varieties of minimal degree. *Math. Res. Lett.*, 15:1211–1221, 2008.

[87] D. Plaumann, B. Sturmfels, and C. Vinzant. Quartic curves and their bitangents. *J. Symbolic Comput.*, 46(6):712–733, 2011.

[88] C. Raicu. Secant varieties of Segre-Veronese varieties. *Algebra Number Theory*, 6(8):1817–1868, 2012.

[89] G. Salmon. *A treatise on the higher plane curves: Intended as a sequel to A treatise on conic sections, 3rd ed.* Dublin, 1879; reprinted by Chelsea Publ. Co., New York, 1960.

[90] S. Sam. Computing inclusions of Schur modules. *J. Softw. Algebra Geom.*, 1:5{10, 2009.

[91] S. Sam, A. Snowden, and J. Weyman. Homology of Littlewood complexes. *Selecta Math.*, 19(3):655{698, 2013.

[92] S. Sam and J. Weyman. Pieri resolutions for classical groups. *J. Algebra*, 329(1):222{259, 2011.

[93] A.J. Sommese, J. Verschelde, and C.W. Wampler. Symmetric functions applied to decomposing solution sets of polynomial systems. *SIAM J. Numer. Anal.*, 40(6):2012{2046, 2002.

[94] H. Stewenius, D. Nister, F. Kahl, and F. Scha alitzky. A minimal solution for relative pose with unknown focal length. *Proc. Int. Conf. on Comput. Vision*, pages 789{794, 2005.

[95] M. Stillman, H. Schenck, and C. Raicu. SchurRings, a package for Macaulay2. *Available at* `http://www.math.uiuc.edu/Macaulay2/doc/Macaulay2-1.8.2/share/doc/Macaulay2/SchurRings/html/`.

[96] P. Sturm. On focal length calibration from two views. *Proc. Int. Conf. Comput. Vision*, pages 145{150, 2001.

[97] B. Sturmfels. *Gröbner Bases and Convex Polytopes; Univ. Lectures Series, vol. 8*. American Mathematical Society, Providence, 1996.

[98] B. Sturmfels. *Solving Systems of Polynomial Equations; CBMS Regional Conferences Series, vol. 97*. American Mathematical Society, Providence, 2002.

[99] B. Sturmfels. The Hurwitz form of a projective variety. *J. Symb. Comput.*, 79:186{196, 2017.

[100] M. Trager, J. Ponce, and M. Hebert. Trinocular geometry revisited. *Int. J. Comput. Vision*, 120(2):134{152, 2016.

[101] C. Traverso. *Gröbner trace algorithms; Symbolic and Algebraic Computation, Lecture Notes in Computer Science, vol. 358, pages 125-138*. Springer, Providence, 2005.

[102] J. Verschelde. Algorithm 795: PHCPACK: A general-purpose solver for polynomial systems by homotopy continuation. *ACM Trans. Math. Software*, 25(2):251{276, 1999.

[103] J. Weng, T.S. Huang, and N. Ahuja. Motion and structure from line correspondences: closed-form solution, uniqueness, and optimization. *IEEE T. Pattern Anal.*, 14(3):318{336, 1992.

[104] J. Weyman. *Cohomology of vector bundles and syzygies; Cambridge tracts in mathematics, vol. 149*. Cambridge University Press, Cambridge, 2003.