Canad. J. Math. Vol. **65** (5), 2013 pp. 961–988
http://dx.doi.org/10.4153/CJM-2012-023-2

A Hilbert Scheme in Computer Vision

Chris Aholt, Bernd Sturmfels, and Rekha Thomas

Abstract. Multiview geometry is the study of two-dimensional images of three-dimensional scenes, a foundational subject in computer vision. We determine a universal Gröbner basis for the multiview ideal of n generic cameras. As the cameras move, the multiview varieties vary in a family of dimension $11n - 15$. This family is the distinguished component of a multigraded Hilbert scheme with a unique Borel-fixed point. We present a combinatorial study of ideals lying on that Hilbert scheme.

1 Introduction

Computer vision is based on mathematical foundations known as *multiview geometry* [7, 9] or *epipolar geometry* [11, §9]. In that subject one studies the space of pictures of three-dimensional objects seen from $n \geq 2$ cameras. Each camera is represented by a 3×4-matrix A_i of rank 3. The matrix specifies a linear projection from $\mathbb{P}^3$ to $\mathbb{P}^2$, which is well defined on $\mathbb{P}^3 \backslash \{f_i\}$, where the focal point f_i is represented by a generator of the kernel of A_i.

The space of pictures from the n cameras is the image of the rational map

$$\phi_A : \mathbb{P}^3 \dashrightarrow (\mathbb{P}^2)^n, \quad \mathbf{x} \mapsto (A_1\mathbf{x}, A_2\mathbf{x}, \dots, A_n\mathbf{x}). \tag{1.1}$$

The closure of this image is an algebraic variety, denoted V_A and called the *multiview variety* of the given n-tuple of 3×4-matrices $A = (A_1, A_2, \dots, A_n)$. In geometric language, the multiview variety V_A is the blow-up of $\mathbb{P}^3$ at the cameras $f_1, \dots, f_n$, and we study this threefold as a subvariety of $(\mathbb{P}^2)^n$.

The *multiview ideal* J_A is the prime ideal of all polynomials that vanish on the multiview variety V_A. It lives in a polynomial ring $K[x, y, z]$ in $3n$ unknowns (x_i, y_i, z_i), $i = 1, 2, \dots, n$, that serve as coordinates on $(\mathbb{P}^2)^n$. In Section 2 we give a determinantal representation of J_A for generic A and identify a universal Gröbner basis consisting of multilinear polynomials of degree 2, 3, and 4. This extends previous results of Heyden and Åström [12].

The multiview ideal J_A has a distinguished initial monomial ideal M_n that is independent of A, provided the configuration A is generic. Section 3 gives an explicit description of M_n and shows that it is the unique Borel-fixed ideal with its $\mathbb{Z}^n$-graded Hilbert function. Following [3], we introduce the multigraded Hilbert scheme $\mathcal{H}_n$ that parametrizes $\mathbb{Z}^n$-homogeneous ideals in $K[x, y, z]$ with the same Hilbert function as M_n. We show in Section 6 that, for $n \geq 3$, $\mathcal{H}_n$ has a distinguished component

Received by the editors April 13, 2012.
Published electronically July 19, 2012.
All three authors were partially supported by the US National Science Foundation.
AMS subject classification: **14N**, 14Q, 68.
Keywords: multigraded Hilbert Scheme, computer vision, monomial ideal, Groebner basis, generic initial ideal.

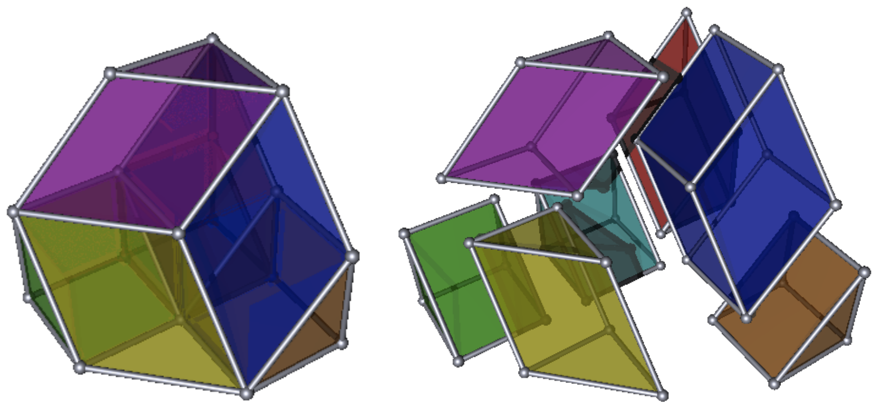

Figure 1: A multiview variety V_A for $n = 3$ cameras degenerates into six copies of $\mathbb{P}^1 \times \mathbb{P}^2$ and one copy of $\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1$.

of dimension $11n - 15$ that compactifies the space of camera positions studied in computer vision. For two cameras, that space is an irreducible cubic hypersurface in $\mathcal{H}_2 \simeq \mathbb{P}^8$.

Section 4 concerns the case when $n \leq 4$ and the focal points f_i are among the coordinate points $(1{:}0{:}0{:}0), \dots, (0{:}0{:}0{:}1)$. Here the multiview variety V_A is a toric threefold, and its degenerations are parametrized by a certain toric Hilbert scheme inside $\mathcal{H}_n$. Each initial monomial ideal of the toric ideal J_A corresponds to a three-dimensional mixed subdivision as seen in Figure 1. A classification of such mixed subdivisions for $n = 4$ is given in Theorem 4.3.

In Section 5 we place our n cameras on a line in $\mathbb{P}^3$. Moving them very close to each other on that line induces a two-step degeneration of the form

$$\text{trinomial ideal} \longrightarrow \text{binomial ideal} \longrightarrow \text{monomial ideal.} \tag{1.2}$$

We present an in-depth combinatorial study of this curve of multiview ideals.

In Section 6 we finally define the Hilbert scheme $\mathcal{H}_n$, and we construct the space of camera positions as a GIT quotient of a Grassmannian. Our main result (Theorem 6.3) states that the latter is an irreducible component of $\mathcal{H}_n$. As a key step in the proof, the tangent space of $\mathcal{H}_n$ at the monomial ideal in (1.2) is computed and shown to have the correct dimension $11n - 15$. Thus, the curve (1.2) consists of smooth points on the distinguished component of $\mathcal{H}_n$. For $n \geq 3$, our Hilbert scheme has multiple components. This is seen from our classification of monomial ideals on $\mathcal{H}_3$, which relates closely to [3, §5].

The triangulation problem in computer vision is the problem of determining the point $\mathbf{x} \in \mathbb{P}^3$ as in (1.1) from a measured point $p = (p_1, p_2, \dots, p_n)$ in the multiview variety V_A. As stated, this reconstruction is a simple exercise in linear algebra, and so the more accurate problem is to consider triangulation when the point p is a noisy measurement and hence does not lie on V_A. Choosing affine coordinates, this can be formulated as a maximum likelihood optimization problem that is constrained over the multiview variety V_A. The equations defining V_A and, in particular, a degree lexicographic Gröbner basis of the multiview ideal J_A, are necessary to initiate certain convex optimization schemes to solve this maximum likelihood problem. This was one of our motivations for embarking on a thorough study of the multiview variety and its ideal. The results obtained here go well beyond this initial goal and expose the rich combinatorial, algebraic, and geometric properties of these ideals and varieties that arise naturally in computer vision.

2 A Universal Gröbner Basis

Let K be any algebraically closed field, $n \geq 2$, and consider the map ϕ_A defined as in (1.1) by a tuple $A = (A_1, A_2, \dots, A_n)$ of 3×4-matrices of rank 3 with entries in K. The subvariety $V_A = \overline{\text{image}(\phi_A)}$ of $(\mathbb{P}^2)^n$ is the *multiview variety*, and its ideal $J_A \subset K[x, y, z]$ is the *multiview ideal*. Note that J_A is prime, because its variety V_A is the image under ϕ_A of an irreducible variety.

We say that the camera configuration A is *generic* if all 4×4-minors of the $(4 \times 3n)$-matrix $\begin{bmatrix} A_1^T & A_2^T & \cdots & A_n^T \end{bmatrix}$ are non-zero. In particular, if A is generic, then the focal points of the n cameras are pairwise distinct in $\mathbb{P}^3$. For any subset $\sigma = \{\sigma_1, \dots, \sigma_s\} \subseteq [n]$ we consider the $3s \times (s+4)$-matrix

$$A_\sigma := \begin{bmatrix} A_{\sigma_1} & p_{\sigma_1} & \mathbf{0} & \cdots & \mathbf{0} \\ A_{\sigma_2} & \mathbf{0} & p_{\sigma_2} & \ddots & \mathbf{0} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ A_{\sigma_s} & \mathbf{0} & \cdots & \mathbf{0} & p_{\sigma_s} \end{bmatrix},$$

where $p_i := \begin{bmatrix} x_i & y_i & z_i \end{bmatrix}^T$ for $i \in [n]$. Assuming $s \geq 2$, each maximal minor of A_σ is a homogeneous polynomial of degree $s = |\sigma|$ that is linear in p_i for $i \in \sigma$. Thus for $s = 2, 3, \dots$, these polynomials are bilinear, trilinear, etc. The matrix A_σ and its maximal minors are considered frequently in multiview geometry [11, 12]. Recall that a *universal Gröbner basis* of an ideal is a subset that is a Gröbner basis of the ideal under all term orders. The following theorem is the main result in this section.

Theorem 2.1 *If A is generic, then the maximal minors of the matrices A_σ for $2 \leq |\sigma| \leq 4$ form a universal Gröbner basis of the multiview ideal J_A.*

The proof rests on a sequence of lemmas. Here is the most basic one.

Lemma 2.2 *The maximal minors of A_σ for $|\sigma| \geq 2$ lie in the prime ideal J_A.*

Proof If $(p_1, \dots, p_n) \in (K^3)^n$ represents a point in image(ϕ_A), then there exists a non-zero vector $q \in K^4$ and non-zero scalars $c_1, \dots, c_n \in K$ such that $A_i q = c_i p_i$ for $i = 1, 2, \dots, n$. This means that the columns of A_σ are linearly dependent. Since A_σ has at least as many rows as columns, the maximal minors of A_σ must vanish at every point $p \in V_A$. ■

Later we shall see that when A is generic, J_A has only one initial monomial ideal up to symmetry. We now identify that ideal. Let M_n denote the ideal in $K[x,y,z]$ generated by the $\binom{n}{2}$ quadrics $x_i x_j$, the $3\binom{n}{3}$ cubics $x_i y_j y_k$, and the $\binom{n}{4}$ quartics $y_i y_j y_k y_l$, where i, j, k, l runs over distinct indices in $[n]$.

We fix the lexicographic term order $\prec$ on $K[x,y,z]$ that is specified by

$$x_1 \succ \cdots \succ x_n \succ y_1 \succ \cdots \succ y_n \succ z_1 \succ \cdots \succ z_n.$$

Our goal is to prove that the initial monomial ideal $\mathrm{in}_\prec(J_A)$ is equal to M_n. We begin with the easier inclusion.

Lemma 2.3 *If A is generic, then $M_n \subseteq \mathrm{in}_\prec(J_A)$.*

Proof The generators of M_n are the quadrics $x_i x_j$, the cubics $x_i y_j y_k$, and the quartics $y_i y_j y_k y_l$. By Lemma 2.2, it suffices to show that these are the initial monomials of maximal minors of $A_{\{ij\}}$, $A_{\{ijk\}}$, and $A_{\{ijkl\}}$ respectively.

For the quadrics this is easy. The matrix $A_{\{ij\}}$ is square, and we have

$$\det(A_{\{ij\}}) = \det \begin{bmatrix} A_i^1 & x_i & 0 \\ A_i^2 & y_i & 0 \\ A_i^3 & z_i & 0 \\ A_j^1 & 0 & x_j \\ A_j^2 & 0 & y_j \\ A_j^3 & 0 & z_j \end{bmatrix} = \det \begin{bmatrix} A_i^2 \\ A_i^3 \\ A_j^2 \\ A_j^3 \end{bmatrix} x_i x_j + \text{ lex. lower terms}, \tag{2.1}$$

where A_t^r is the r-th row of A_t. The coefficient of $x_i x_j$ is non-zero, because A was assumed to be generic. For the cubics, we consider the 9×7-matrix

$$A_{\{ijk\}} = \begin{bmatrix} A_i & p_i & 0 & 0 \\ A_j & 0 & p_j & 0 \\ A_k & 0 & 0 & p_k \end{bmatrix}. \tag{2.2}$$

Now, $x_i y_j y_k$ is the lexicographic initial monomial of the 7×7-determinant formed by removing the fourth and seventh rows of $A_{\{ijk\}}$. Here we are using that, by genericity, the vectors $A_i^2, A_i^3, A_j^3, A_k^3$ are linearly independent.

Finally, for the quartic monomial $y_i y_j y_k y_l$ we consider the 12×8 matrix

$$A_{\{ijkl\}} = \begin{bmatrix} A_i & p_i & 0 & 0 & 0 \\ A_j & 0 & p_j & 0 & 0 \\ A_k & 0 & 0 & p_k & 0 \\ A_l & 0 & 0 & 0 & p_l \end{bmatrix}. \tag{2.3}$$

Removing the first row from each of the four blocks, we obtain an 8×8-matrix whose determinant has $y_i y_j y_k y_l$ as its lex. initial monomial. ∎

The next step towards our proof of Theorem 2.1 is to express the multiview variety V_A as a projection of a diagonal embedding of $\mathbb{P}^3$. This will put us in a position to utilize the results of Cartwright and Sturmfels in [3].

https://doi.org/10.4153/CJM-2012-023-2 Published online by Cambridge University Press

We extend each camera matrix A_i to an invertible 4×4-matrix $B_i = \left[\begin{smallmatrix} b_i \\ A_i \end{smallmatrix}\right]$ by adding a row b_i at the top. Our diagonal embedding of $\mathbb{P}^3$ is the map

$$\psi_B \colon \mathbb{P}^3 \to (\mathbb{P}^3)^n, \quad \mathbf{x} \mapsto (B_1\mathbf{x}, B_2\mathbf{x}, \dots, B_n\mathbf{x}).$$

Let $V^B := \mathrm{image}(\psi_B) \subset (\mathbb{P}^3)^n$ and let $J^B \subset K[w, x, y, z]$ be its prime ideal. Here $(w_i : x_i : y_i : z_i)$ are coordinates on the i-th copy of $\mathbb{P}^3$, and (w, x, y, z) are coordinates on $(\mathbb{P}^3)^n$. The ideal J^B is generated by the 2×2-minors of the $4 \times n$ matrix

$$\left[\; B_1^{-1}\begin{bmatrix} w_1 \\ x_1 \\ y_1 \\ z_1 \end{bmatrix} \;\; B_2^{-1}\begin{bmatrix} w_2 \\ x_2 \\ y_2 \\ z_2 \end{bmatrix} \;\; \cdots \;\; B_n^{-1}\begin{bmatrix} w_n \\ x_n \\ y_n \\ z_n \end{bmatrix} \;\right]. \tag{2.4}$$

These equations can be seen to generate the prime ideal J^B by noting that V^B is nothing more than the image of the diagonal embedding of $\mathbb{P}^3$ in $(\mathbb{P}^3)^n$ after a linear change of coordinates. Now consider the coordinate projection

$$\pi \colon (\mathbb{P}^3)^n \dashrightarrow (\mathbb{P}^2)^n, \quad (w_i : x_i : y_i : z_i) \mapsto (x_i : y_i : z_i) \text{ for } i = 1, \dots, n.$$

The composition $\pi \circ \psi_B$ is a rational map, and it coincides with ϕ_A on its domain of definition $\mathbb{P}^3 \backslash \{f_1, \dots, f_n\}$. Therefore,

$$V_A = \overline{\pi(V^B)} \quad \text{and} \quad J_A = J^B \cap K[x, y, z].$$

The polynomial ring $K[w, x, y, z]$ admits the natural $\mathbb{Z}^n$-grading $\deg(w_i) = \deg(x_i) = \deg(y_i) = \deg(z_i) = e_i$, where e_i is the standard unit vector in $\mathbb{R}^n$. Under this grading, $K[w, x, y, z]/J^B$ has the multigraded Hilbert function

$$\mathbb{N}^n \to \mathbb{N}, (u_1, \dots, u_n) \mapsto \binom{u_1 + \cdots + u_n + 3}{3}.$$

The multigraded Hilbert scheme $H_{4,n}$ that parametrizes $\mathbb{Z}^n$-homogeneous ideals in $K[w, x, y, z]$ with that Hilbert function was studied in [3]. More generally, the multigraded Hilbert scheme $H_{d,n}$ represents degenerations of the diagonal $\mathbb{P}^{d-1}$ in $(\mathbb{P}^{d-1})^n$ for any d and n. For the general definition of multigraded Hilbert schemes, see [10]. It was shown in [3] that $H_{d,n}$ has a unique Borel-fixed ideal $Z_{d,n}$. Here *Borel-fixed* means that $Z_{d,n}$ is stable under the action of $\mathcal{B}^n$ where $\mathcal{B}$ is the group of lower triangular matrices in $\mathrm{PGL}(d, K)$. Here is what we shall need to know about the monomial ideal $Z_{4,n}$.

Lemma 2.4 (Cartwright-Sturmfels [3, §2] and Conca [4, §5])

(i) *The unique Borel-fixed monomial ideal $Z_{4,n}$ on $H_{4,n}$ is generated by the following monomials, where i, j, k, l are distinct indices in $[n]$:*

$$w_i w_j, w_i x_j, w_i y_j, x_i x_j, x_i y_j y_k, y_i y_j y_k y_l.$$

(ii) *This ideal $Z_{4,n}$ is the lexicographic initial ideal of J^B when B is sufficiently generic. The lexicographic order here is $w \succ x \succ y \succ z$ with each block ordered lexicographically in increasing order of indices.*

Using these results, it was deduced in [3] that all ideals on $H_{4,n}$ are radical and Cohen–Macaulay, and that $H_{4,n}$ is connected. We now use this distinguished Borel-fixed ideal $Z_{4,n}$ to prove the equality in Lemma 2.3.

Lemma 2.5 *If A is generic, then $M_n = \mathrm{in}_{\prec}(J_A)$.*

Proof We fix the lexicographic term order $\prec$ on $K[w, x, y, z]$ and its restriction to $K[x, y, z]$. Lemma 2.4(i) shows that $M_n = Z_{4,n} \cap K[x, y, z]$. Lemma 2.4(ii) states that $Z_{4,n} = \mathrm{in}_{\prec}(J^B)$ when B is generic. The lexicographic order has the important property that it allows the operations of taking initial ideals and intersections to commute [5, Chapter 3]. Therefore,

$$\mathrm{in}_{\prec}(J_A) = \mathrm{in}_{\prec}\big(J^B \cap K[x, y, z]\big) = \mathrm{in}_{\prec}(J^B) \cap K[x, y, z] = Z_{4,n} \cap K[x, y, z] = M_n.$$

This identity is valid whenever the conclusion of Lemma 2.4(ii) is true. We claim that for this to hold the appropriate genericity notion for B is that all 4×4-minors of the $(4 \times 4n)$-matrix $\begin{bmatrix} B_1^T & B_2^T & \cdots & B_n^T \end{bmatrix}$ are non-zero. Indeed, under this hypothesis, the maximal minors of the $4s \times (s + 4)$-matrix

$$B_\sigma := \begin{bmatrix} B_{\sigma_1} & \widetilde{p}_{\sigma_1} & \mathbf{0} & \cdots & \mathbf{0} \\ B_{\sigma_2} & \mathbf{0} & \widetilde{p}_{\sigma_2} & \ddots & \mathbf{0} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ B_{\sigma_s} & \mathbf{0} & \cdots & \mathbf{0} & \widetilde{p}_{\sigma_s} \end{bmatrix}, \text{ where } \widetilde{p}_i := \begin{bmatrix} w_i & x_i & y_i & z_i \end{bmatrix}^T \text{ for } i \in [n],$$

have non-vanishing leading coefficients. We see that $Z_{4,n} \subseteq \mathrm{in}_{\prec}(J^B)$ by reasoning akin to that in the proof of Lemma 2.3. The equality $Z_{4,n} = \mathrm{in}_{\prec}(J^B)$ is then immediate, since $Z_{4,n}$ is the multigraded generic initial ideal of J^B; see Lemma 2.4(ii). Hence, for any generic camera positions A, we can add a row to A_i and get B_i that are "sufficiently generic" for Lemma 2.4(ii). This completes the proof. ∎

Proof of Theorem 2.1 Lemma 2.5 and the proof of Lemma 2.3 show that the maximal minors of the matrices A_σ for $2 \leq |\sigma| \leq 4$ are a Gröbner basis of J_A for the lexicographic term order. Each polynomial in that Gröbner basis is multilinear, thus the initial monomials remain the same for any term order satisfying $x_i \succ y_i \succ z_i$ for $i = 1, 2, \ldots, n$. So, the minors form a Gröbner basis for that term order. The set of minors is invariant under permuting $\{x_i, y_i, z_i\}$ for each i. Moreover, the genericity of A implies that every monomial that can possibly appear in the support of a minor does so. Hence, these minors form a universal Gröbner basis of J_A. ∎

Remark 2.6 Computer vision experts have known for a long time that multiview varieties V_A are defined set-theoretically by the above multilinear constraints of degree at most 4. We refer to work of Heyden and Åström [12, 13]. What is new here is that these constraints define V_A in the strongest possible sense: they form a universal Gröbner basis for the prime ideal J_A.

The n cameras are in *linearly general position* if no four focal points are coplanar and no three are collinear. While the number of multilinear polynomials in our lex Gröbner basis of J_A is $\binom{n}{2} + 3\binom{n}{3} + \binom{n}{4}$, far fewer suffice to generate the ideal J_A when A is in linearly general position.

Corollary 2.7 *If A is in linearly general position, then the ideal J_A is minimally generated by $\binom{n}{2}$ bilinear and $\binom{n}{3}$ trilinear polynomials.*

Proof This can be shown for $n \leq 4$ by a direct calculation. Alternatively, these small cases are covered by transforming to the toric ideals in Section 4. First map the focal points of the cameras to the torus fixed focal points of the toric case, then multiply each A_i by a suitable $g_i \in \mathrm{PGL}(3, K)$.

Now let $n \geq 5$. For any three cameras i, j, k, the maximal minors of (2.2) are generated by only one such maximal minor modulo the three bilinear polynomials (2.1). Likewise, for any four cameras i, j, k, and l, the maximal minors of (2.3) are generated by the trilinear and bilinear polynomials. This implies that the resulting $\binom{n}{2} + \binom{n}{3}$ polynomials generate J_A, and, by restricting to two or three cameras, we see that they minimally generate. ∎

3 The Generic Initial Ideal

We now focus on combinatorial properties of our special monomial ideal

$$M_n = \left\langle x_i x_j, x_i y_j y_k, y_i y_j y_k y_l : \forall i, j, k, l \in [n] \text{ distinct } \right\rangle .$$

We refer to M_n as the *generic initial ideal* in multiview geometry, because it is the lex initial ideal of any multiview ideal J_A after a generic coordinate change via the group G^n where $G = \mathrm{PGL}(3, K)$. Indeed, consider any rank 3 matrices $A_1, A_2, \ldots, A_n \in K^{3\times 4}$ with pairwise distinct kernels $K\{f_i\}$. If $g = (g_1, g_2, \ldots, g_n)$ is generic in G^n, then $g \circ A$ is generic in the sense that all 4×4-minors of the matrix

$$\begin{bmatrix} (g_1A_1)^T & (g_2A_2)^T & \cdots & (g_nA_n)^T \end{bmatrix}$$

are non-zero. Thus, by the results of Section 2, M_n is the initial ideal of $J_{g\circ A}$, or, using standard commutative algebra lingo, M_n is the generic initial ideal of J_A.

The careful reader will notice that the term "generic initial ideal" is often associated with the action of the full general linear group, whereas here we are using it in reference to our G^n action. For our purposes, the G^n action more naturally captures the structure of the problem at hand. The term *multigraded generic initial ideal* was used for this construction in [3, 4].

Since M_n is a squarefree monomial ideal, it is radical. Hence M_n is the intersection of its minimal primes, which are generated by subsets of the variables x_i and y_j. We begin by computing this prime decomposition.

Proposition 3.1 *The generic initial ideal M_n is the irredundant intersection of $\binom{n}{3} + 2\binom{n}{2}$ monomial primes. These are the monomial primes P_{ijk} and $Q_{ij} \subseteq K[x, y, z]$ defined below for any distinct indices $i, j, k \in [n]$:*

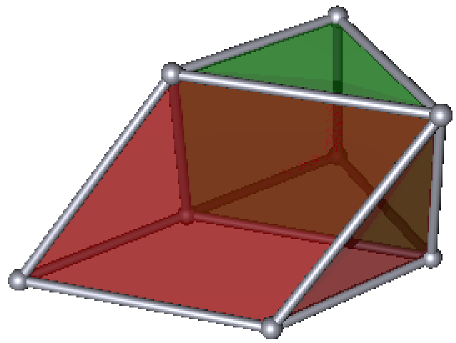

Figure 2: The variety of the generic initial ideal M_2 seen as two adjacent facets of the 4-dimensional polytope $\Delta_2 \times \Delta_2$.

- *P_{ijk} is generated by $x_1, \ldots, x_n$ and all y_l with $l \notin \{i, j, k\}$;*
- *Q_{ij} is generated by all x_l for $l \neq i$ and y_l for $l \notin \{i, j\}$.*

Proof Let L denote the intersection of all P_{ijk} and Q_{ij}. Each monomial generator of M_n lies in P_{ijk} and in Q_{ij}, so $M_n \subseteq L$. For the reverse inclusion, we will show that $V(M_n)$ is contained in $V(L) = (\bigcup V(P_{ijk})) \cup (\bigcup V(Q_{ij}))$.

Let $(\widetilde{x}, \widetilde{y}, \widetilde{z})$ be any point in the variety $V(M_n)$. First suppose $\widetilde{x}_i = 0$ for all $i \in [n]$. Since $\widetilde{y}_i \widetilde{y}_j \widetilde{y}_k \widetilde{y}_l = 0$ for distinct indices, there are at most three indices i, j, k such that $\widetilde{y}_i$, $\widetilde{y}_j$, and $\widetilde{y}_k$ are nonzero. Hence $(\widetilde{x}, \widetilde{y}, \widetilde{z}) \in V(P_{ijk})$.

Next suppose $\widetilde{x}_i \neq 0$. The index i is unique because $x_i x_j \in M_n$ for all $j \neq i$. Since $\widetilde{x}_i \widetilde{y}_j \widetilde{y}_k = 0$ for all $j, k \neq i$, we have $\widetilde{y}_j \neq 0$ for at most one index $j \neq i$. These properties imply $(\widetilde{x}, \widetilde{y}, \widetilde{z}) \in V(Q_{ij})$. ∎

We regard the monomial variety $V(M_n)$ as a threefold inside the product of projective planes $(\mathbb{P}^2)^n$. If the focal points are distinct, V_A has a Gröbner degeneration to the reducible threefold $V(M_n)$. The irreducible components of $V(M_n)$ are

$$V(P_{ijk}) \simeq \mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1 \quad \text{and} \quad V(Q_{ij}) \simeq \mathbb{P}^2 \times \mathbb{P}^1. \tag{3.1}$$

We find it convenient to regard $(\mathbb{P}^2)^n$ as a toric variety so as to identify it with its polytope $(\Delta_2)^n$, a direct product of triangles. The components in (3.1) are 3-dimensional boundary strata of $(\mathbb{P}^2)^n$, and we identify them with faces of $(\Delta_2)^n$. The corresponding 3-dimensional polytopes are the 3-*cube* and the *triangular prism*. The following three examples illustrate this view.

Example 3.2 (Two cameras ($n = 2$)) The variety of $M_2 = \langle x_1 \rangle \cap \langle x_2 \rangle$ is a hypersurface in $\mathbb{P}^2 \times \mathbb{P}^2$. The two components are triangular prisms $\mathbb{P}^2 \times \mathbb{P}^1$, which are glued along a common square $\mathbb{P}^1 \times \mathbb{P}^1$, as shown in Figure 2.

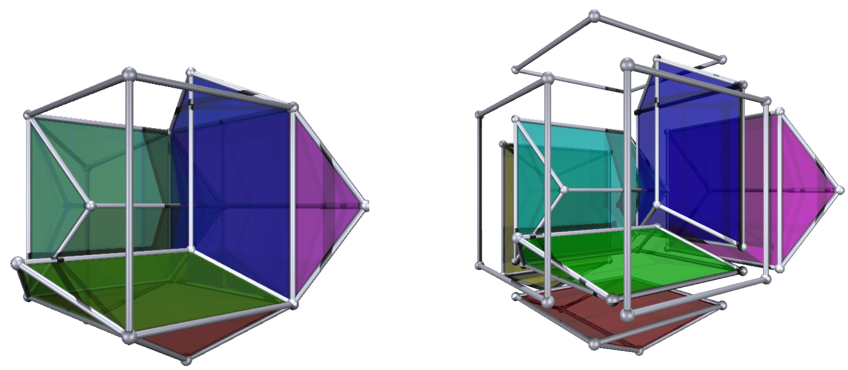

Figure 3: The monomial variety $V(M_3)$ as a subcomplex of $(\Delta_2)^3$.

Example 3.3 (Three cameras ($n = 3$)) The variety of M_3 is a threefold in $\mathbb{P}^2 \times \mathbb{P}^2 \times \mathbb{P}^2$. Its seven components are given by the prime decomposition

$$M_3 = \langle x_1, x_2, y_1\rangle \cap \langle x_1, x_2, y_2\rangle \cap \langle x_1, x_3, y_1\rangle \cap \langle x_1, x_3, y_3\rangle \\ \cap \langle x_2, x_3, y_2\rangle \cap \langle x_2, x_3, y_3\rangle \cap \langle x_1, x_2, x_3\rangle.$$

The last component is a cube $\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1$, and the other six components are triangular prisms $\mathbb{P}^2 \times \mathbb{P}^1$. These are glued in pairs along three of the six faces of the cube. For instance, the two triangular prisms $V(x_1, x_2, y_1)$ and $V(x_1, x_3, y_1)$ intersect the cube $V(x_1, x_2, x_3)$ in the common square face $V(x_1, x_2, x_3, y_1) \simeq \mathbb{P}^1 \times \mathbb{P}^1$. This polyhedral complex lives in the boundary of $(\Delta_2)^3$, and it shown in Figure 3. Compare this picture with Figure 1. ■

Example 3.4 (Four cameras ($n = 4$)) The variety $V(M_4)$ is a threefold in $(\mathbb{P}^2)^4$, regarded as a 3-dimensional subcomplex in the boundary of the 8-dimensional polytope $(\Delta_2)^4$. It consists of four cubes and twelve triangular prisms. The cubes share a common vertex; any two cubes intersect in a square, and each of the six squares is adjacent to two triangular prisms.

From the prime decomposition in Proposition 3.1 we can read off the *multidegree* [17, §8.5] of the ideal M_n. Here and in what follows, we use the natural $\mathbb{Z}^n$-grading on $K[x, y, z]$ given by $\deg(x_i) = \deg(y_i) = \deg(z_i) = e_i$. Each multiview ideal J_A is homogeneous with respect to this $\mathbb{Z}^n$-grading.

Corollary 3.5 *The multidegree of the generic initial ideal M_n is equal to*

$$\mathcal{C}\big(K[x, y, z]/M_n; \mathbf{t}\big) = t_1^2 t_2^2 \cdots t_n^2 \cdot \Big(\sum_{1\le i<j<k\le n} \frac{1}{t_i t_j t_k} + \sum_{1\le i,j\le n} \frac{1}{t_i^2 t_j} \Big) \qquad (3.2)$$

A more refined analysis also yields the Hilbert function in the $\mathbb{Z}^n$-grading.

https://doi.org/10.4153/CJM-2012-023-2 Published online by Cambridge University Press

Theorem 3.6 *The multigraded Hilbert function of $K[x,y,z]/M_n$ equals*

$$\mathbb{N}^n \to \mathbb{N}, (u_1,\dots,u_n) \mapsto \binom{u_1+\cdots+u_n+3}{3} - \sum_{i=1}^n \binom{u_i+2}{3}. \tag{3.3}$$

Proof Fix $u \in \mathbb{N}^n$. A K-basis $\mathfrak{B}_u$ for $(K[x,y,z]/M_n)_u$ is given by all monomials $x^a y^b z^c \notin M_n$ such that $a+b+c=u$. Therefore, either (i) $a=0$ and at most three components of b are non-zero; or (ii) $a \neq 0$, in which case only one a_i can be non-zero and $b_j \neq 0$ for at most one $j \in [n]\backslash\{i\}$.

We shall count the monomials in $\mathfrak{B}_u$. Monomials of type (i) look like $y^b z^c$, with at most three nonzero entries in b. Also, b determines c, since $c_i = u_i - b_i$ for all $i \in [n]$, and so we count the number of possibilities for y^b. There are u_i choices for $b_i \neq 0$, and thus $U := u_1 + \cdots + u_n$ many monomials in the set

$$\mathcal{Y} := \{y_i^{b_i} : 1 \le b_i \le u_i, i = 1,\dots,n\}.$$

The factor y^b in $y^b z^c$ is the product of 0, 1, 2, or 3 monomials from $\mathcal{Y}$ with distinct subscripts.

To resolve over-counting, consider a fixed index i. There are $\binom{u_i}{2}$ ways of choosing two monomials from $\mathcal{Y}$ with subscript i and $\binom{u_i}{3}$ ways of choosing three monomials from $\mathcal{Y}$ with subscript i. Also, there are $\binom{u_i}{2}(U-u_i)$ ways of choosing two monomials from $\mathcal{Y}$ with subscript i and a third monomial with a different subscript. Hence, the number of choices for y^b in $y^b z^c$ is

$$\binom{U}{0} + \binom{U}{1} + \left[\binom{U}{2} - \sum_{i=1}^n \binom{u_i}{2}\right] + \left[\binom{U}{3} - \sum_{i=1}^n \binom{u_i}{3} - U\sum_{i=1}^n \binom{u_i}{2} + \sum_{i=1}^n u_i \binom{u_i}{2}\right].$$

For case (ii) we count all monomials $x^a y^b z^c \in \mathfrak{B}_u$ with $a_i \neq 0$ and all other $a_j = 0$. It suffices to count the choices for the factor $x^a y^b$. For fixed i, there are $\binom{u_i+1}{2}$ monomials of the form $x_i^{a_i} y_i^{b_i}$ with $a_i + b_i \le u_i$ and $a_i \ge 1$. Such a monomial may be multiplied with $y_j^{b_j}$ such that $j \neq i$ and $0 \le b_j \le u_j$. This amounts to choosing zero or one monomial from $\mathcal{Y}\backslash\{y_i, y_i^2, \dots, y_i^{u_i}\}$ for which there are $1 + U - u_i$ choices. Hence, there are

$$[1+U]\sum_{i=1}^n \binom{u_i+1}{2} - \sum_{i=1}^n u_i \binom{u_i+1}{2}$$

monomials in $\mathfrak{B}_u$ of type (ii). Adding the two expressions, we get

$$\begin{aligned} |\mathfrak{B}_u| &= 1 + U + \binom{U}{2} + \binom{U}{3} + (1+U)\sum_{i=1}^n \binom{u_i}{1} - \sum_{i=1}^n u_i \binom{u_i}{1} - \sum_{i=1}^n \binom{u_i}{3} \\ &= 1 + U + \binom{U}{2} + \binom{U}{3} + (1+U)U - \sum_{i=1}^n \binom{u_i+2}{3} \\ &= \binom{U+3}{3} - \sum_{i=1}^n \binom{u_i+2}{3}. \end{aligned}$$

■

Our analysis of M_n has the following implication for the multiview ideals J_A. Note that these are $\mathbb{Z}^n$-homogeneous for any camera configuration A.

Theorem 3.7 *For an n-tuple of camera matrices* $A = (A_1, \dots, A_n)$ *with* $\operatorname{rank}(A_i) = 3$ *for each i, the multiview ideal* J_A *has the Hilbert function* (3.3) *if and only if the focal points of the n cameras are pairwise distinct.*

Proof The if-direction follows from the argument in the first paragraph of this section. If the n camera positions $f_i = \ker(A_i)$ are distinct in $\mathbb{P}^3$, then M_n is the generic initial ideal of J_A, and hence both ideals have the same $\mathbb{Z}^n$-graded Hilbert function. For the only-if-direction we shall use:

(3.4) If $Q \in \mathrm{PGL}(4, K)$ and $AQ := (A_1Q, \dots, A_nQ)$, then $J_A = J_{AQ}$.

This holds because Q defines an isomorphism on $\mathbb{P}^3$, and hence ϕ_A as in (1.1) has the same image in $(\mathbb{P}^2)^n$ as ϕ_{AQ}.

Suppose first that $n = 2$ and A_1 and A_2 have the same focal point and hence the same (three-dimensional) rowspace W. We can map W to the hyperplane $\{x_1 = 0\}$ by some $Q \in \mathrm{PGL}(4, K)$, and (3.4) ensures that $J_A = J_{AQ}$. Thus we may assume that $A_1 = [\,\mathbf{0}\ C_1\,]$ and $A_2 = [\,\mathbf{0}\ C_2\,]$, where C_1 and C_2 are invertible matrices and $\mathbf{0}$ is a column of zeros. Choosing $f_1 = f_2 = (1, 0, 0, 0)$ as the top row of B_1 and B_2 (as in Section 2), we have

$$B_1^{-1} = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & C_1^{-1} \end{bmatrix}, \quad B_2^{-1} = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & C_2^{-1} \end{bmatrix}.$$

The ideal J^B is generated by the 2×2 minors of the matrix (2.4), which is

$$D = \begin{bmatrix} w_1 & w_2 \\ p_1(x_1, y_1, z_1) & q_1(x_2, y_2, z_2) \\ p_2(x_1, y_1, z_1) & q_2(x_2, y_2, z_2) \\ p_3(x_1, y_1, z_1) & q_3(x_2, y_2, z_2) \end{bmatrix},$$

where the p_i's and q_i's are linear polynomials. The ideal I generated by the 2×2 minors of the submatrix of D obtained by deleting the top row lies on the Hilbert scheme $H_{3,2}$ from [3] and hence $K[x, y, z]/I$ has Hilbert function

$$\mathbb{N}^2 \to \mathbb{N}, \quad (u_1, u_2) \mapsto \binom{u_1 + u_2 + 2}{2}.$$

For $(u_1, u_2) = (1, 1)$, this has value 6. Since $I \subseteq J_A = J^B \cap K[x, y, z]$, the Hilbert function of $K[x, y, z]/J_A$ has value ≤ 6, while (3.3) evaluates to 8.

If $n > 2$, we may assume without loss of generality that A_1 and A_2 have the same rowspace. The argument for $n = 2$ shows that $J_A = J^B \cap K[x, y, z] \supseteq I$. The Hilbert function value of $K[x, y, z]/J_A$ in degree $e_1 + e_2$ is again 8, while the Hilbert function value of $K[x, y, z]/I$ in degree $e_1 + e_2$ coincides with the value 6 for $K[x_1, y_1, z_1, x_2, y_2, z_2]/I$. So we again conclude that $K[x, y, z]/J_A$ does not have Hilbert function (3.3). ∎

For $G = \mathrm{PGL}(3, K)$, the product G^n acts on $K[x, y, z]$ by left-multiplication

$$(g_1, \dots, g_n) \cdot \begin{bmatrix} x_i \\ y_i \\ z_i \end{bmatrix} = g_i \begin{bmatrix} x_i \\ y_i \\ z_i \end{bmatrix}.$$

An ideal I in $K[x, y, z]$ is said to be *Borel-fixed* if it is fixed under the induced action of $\mathcal{B}^n$ where $\mathcal{B}$ is the subgroup of lower triangular matrices in G.

Proposition 3.8 *The generic initial ideal M_n is the unique ideal in $K[x, y, z]$ that is Borel-fixed and has the Hilbert function (3.3) in the $\mathbb{Z}^n$-grading.*

Proof The proof is analagous to that of [3, Theorem 2.1], where $Z_{d,n}$ plays the role of M_n. The ideal M_n is Borel-fixed because it is a generic initial ideal. The same approach as in [6, §15.9.2] can be used to prove this fact. One could also use the generators of M_n to prove Borel-fixedness directly.

The multidegree of any $\mathbb{Z}^n$-graded ideal is determined by its Hilbert series [17, Claim 8.54]. Thus any ideal I with Hilbert function (3.3) has multidegree (3.2). Let I be such a Borel-fixed ideal. This is a monomial ideal.

Each maximum-dimensional associated prime P of I has multidegree either $t_1^2 t_2^2 \cdots t_n^2/(t_i t_j t_k)$ or $t_1^2 t_2^2 \cdots t_n^2/(t_i^2 t_j)$, by [17, Theorem 8.53]. In the first case P is generated by $2n - 3$ indeterminates, one associated with each of the three cameras i, j, k and two each from the other $n - 3$ cameras. Borel-fixedness of I tells us that the generators indexed by each camera must be the most expensive variables with respect to the order $\prec$. Hence $P = P_{ijk}$. Similarly, $P = Q_{ij}$ in the case when P has multidegree $t_1^2 t_2^2 \cdots t_n^2/(t_i^2 t_j)$.

Every prime component of M_n is among the minimal associated primes of I. This yields the containments $I \subseteq \sqrt{I} \subseteq M_n$. Since I and M_n have the same $\mathbb{Z}^n$-graded Hilbert function, the equality $I = M_n$ holds. ■

The *Stanley–Reisner complex* of a squarefree monomial ideal M in a polynomial ring $K[t_1, \dots, t_s]$ is the simplicial complex on $\{1, \dots, s\}$ whose facets are the sets $[s] \backslash \sigma$ where $P_\sigma := \{t_i : i \in \sigma\}$ is a minimal prime of M. A *shelling* of a simplicial complex is an ordering $F_1, F_2, \dots, F_q$ of its facets such that, for each $1 < j \leq q$, there exists a unique minimal face of F_j (with respect to inclusion) among the faces of F_j that are not faces of some earlier facet F_i, $i < j$; see [18, Definition 2.1]. If the Stanley–Reisner complex of M is shellable, then $K[t_1, \dots, t_s]/M$ is Cohen–Macaulay [18, Theorem 2.5].

Proposition 3.9 *The Stanley–Reisner complex of the generic initial ideal M_n is shellable. Hence the quotient ring $K[x, y, z]/M_n$ is Cohen-Macaulay.*

Proof This proof is similar to that for $Z_{d,n}$ given in [3, Corollary 2.6]. Let Δ_n denote the Stanley–Reisner complex of the ideal M_n. By Proposition 3.1, there are two types of minimal primes for M_n, namely P_{ijk} and Q_{ij}, which we describe uniformly as follows. Let $P = (p_{ij})$ be the $3 \times n$ matrix whose i-th column is $[x_i\, y_i\, z_i]^T$. For $u \in \{0, 1, 2\}^n$ define $P_u := \langle p_{ij} : i \leq u_j, 1 \leq j \leq n \rangle$. Then the minimal primes P_{ijk} of M_n are precisely the primes P_u as u varies over all vectors with three coordinates equal to one and the rest equal to two, and the minimal primes Q_{ij} are those P_u where u has one coordinate equal to zero, one coordinate equal to one and the rest equal to two. The facet of Δ_n corresponding to the minimal prime P_u is then $F_u := \{p_{ij} : u_j < i \le 3, 1 \le j \le n\}$. We claim that the ordering of the facets F_u induced by ordering the u's lexicographically starting with $(0,1,2,2,\dots,2)$ and ending with $(2,2,\dots,2,1,0)$ is a shelling of Δ_n.

Consider the face $\eta_u := \{p_{ij} : j > 1, i = u_j + 1 \le 2\}$ of the facet F_u. We will prove that η_u is the unique minimal one among the faces of F_u that have not appeared in a facet $F_{u'}$ for $u' < u$. Suppose G is a face of F_u that does not contain η_u. Pick an element $p_{u_j+1,j} \in \eta_u \backslash G$. Then $j > 1$, $u_j \le 1$, and so if F_u is not the first facet in the ordering, then there exists $i < j$ such that $u_i > 0$, because $u > (0,1,2,2,\dots,2)$ and of the form described above. Pick i such that $i < j$ and $u_i > 0$ and consider $F_{u+e_j-e_i} = F_u \backslash \{p_{u_j+1,j}\} \cup \{p_{u_i,i}\}$. Then $u + e_j - e_i < u$ and G is a face of $F_{u+e_j-e_i}$. Conversely, suppose G is a face of F_u that is also a face of $F_{u'}$ where $u' < u$. Since $\sum u'_j = \sum u_j$, there exists some $j > 1$ such that $u'_j > u_j$. Therefore, G does not contain $p_{u_j+1,j}$ that belongs to η_u. Therefore, η_u is not contained in G. ■

4 A Toric Perspective

In this section we examine multiview ideals J_A that are toric. For an introduction to toric ideals, we refer the reader to [20]. We now assume that, for each camera i, each of the four torus fixed points in $\mathbb{P}^3$ either is the camera position or is mapped to a torus fixed point in $\mathbb{P}^2$. This implies that $n \le 4$. We fix $n = 4$ and $f_i = e_i$ for $i = 1,2,3,4$. Up to permuting and rescaling columns, our assumption implies that the configuration A equals

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

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

For this camera configuration, the multiview ideal J_A is indeed a toric ideal:

Proposition 4.1 *The ideal J_A is obtained by eliminating the diagonal unknowns w_1, w_2, w_3, and w_4 from the ideal of 2×2-minors of the 4×4-matrix*

$$\begin{pmatrix} w_1 & x_2 & x_3 & x_4 \\ x_1 & w_2 & y_3 & y_4 \\ y_1 & y_2 & w_3 & z_4 \\ z_1 & z_2 & z_3 & w_4 \end{pmatrix}. \tag{4.1}$$

This toric ideal is minimally generated by six quadrics and four cubics:

$$J_A = \langle y_1y_4 - x_1z_4,\ y_3x_4 - x_3y_4,\ y_2x_4 - x_2z_4,\ z_1y_3 - x_1z_3,\ z_2x_3 - x_2z_3,\ z_1y_2 - y_1z_2,$$
$$y_2z_3y_4 - z_2y_3z_4,\ y_1z_3x_4 - z_1x_3z_4,\ x_1z_2x_4 - z_1x_2y_4,\ x_1y_2x_3 - y_1x_2y_3\rangle$$

https://doi.org/10.4153/CJM-2012-023-2 Published online by Cambridge University Press

Proof We extend A_i to a 4×4-matrix B_i as in Section 2 by adding the row $b_i = e_i^T$. The B_i's are then all permutation matrices, and the matrix in (2.4) equals the matrix in (4.1). The ideal J^B is generated by the 2×2 minors of that matrix of unknowns. The multiview ideal is $J_A = J^B \cap K[x, y, z]$. We find the listed binomial generators by performing the elimination with a computer algebra package such as `Macaulay2` [8]. Toric ideals are precisely those prime ideals generated by binomials, and hence J_A is a toric ideal. ■

Remark 4.2 The *normalized coordinate system in multiview geometry* proposed by Heyden and Åström [12] is different from ours and does not lead to toric varieties. Indeed, if one uses the camera matrices in [12, §2.3], then J_A is also generated by six quadrics and four cubics, but seven of the ten generators are not binomials. One of the cubic generators has six terms.

In commutative algebra, it is customary to represent toric ideals by integer matrices. Given $\mathcal{A} \in \mathbb{N}^{p \times q}$ with columns $a_1, \dots, a_q$, the *toric ideal* of $\mathcal{A}$ is

$$I_{\mathcal{A}} := \langle t^u - t^v : \mathcal{A}u = \mathcal{A}v, u, v \in \mathbb{N}^q \rangle \subset K[t] := K[t_1, \dots, t_q],$$

where t^u represents the monomial $t_1^{u_1} t_2^{u_2} \cdots t_q^{u_q}$. If $\mathcal{A}'$ is the submatrix of $\mathcal{A}$ obtained by deleting the columns indexed by $j_1, \dots, j_s$ for some $s < q$, then the toric ideal $I_{\mathcal{A}'}$ equals the elimination ideal $I_{\mathcal{A}} \cap K[t_j : j \notin \{j_1, \dots, j_s\}]$; see [20, Prop. 4.13(a)]. The integer matrix $\mathcal{A}$ for our toric multiview ideal J_A in Proposition 4.1 is the following *Cayley matrix* of format 8×12:

$$\mathcal{A} = \begin{bmatrix} A_1^T & A_2^T & A_3^T & A_4^T \\ \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} \end{bmatrix},$$

where $\mathbf{1} = [1\,1\,1]$ and $\mathbf{0} = [0\,0\,0]$. This matrix $\mathcal{A}$ is obtained from the following 8×16 matrix by deleting columns $1, 6, 11$, and 16:

$$\begin{bmatrix} I_4 & I_4 & I_4 & I_4 \\ \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} \end{bmatrix}. \tag{4.2}$$

The vectors $\mathbf{1}$ and $\mathbf{0}$ now have length four, I_4 is the 4×4 identity matrix, and we assume that the columns of (4.2) are indexed by

$$w_1, x_1, y_1, z_1, x_2, w_2, y_2, z_2, x_3, y_3, w_3, z_3, x_4, y_4, z_4, w_4.$$

The matrix (4.2) represents the direct product of two tetrahedra, and its toric ideal is known (by [20, Prop. 5.4]) to be generated by the 2×2 minors of (4.1). Its elimination ideal in the ring $K[x, y, z]$ is $I_{\mathcal{A}}$, and hence $J_A = I_{\mathcal{A}}$.

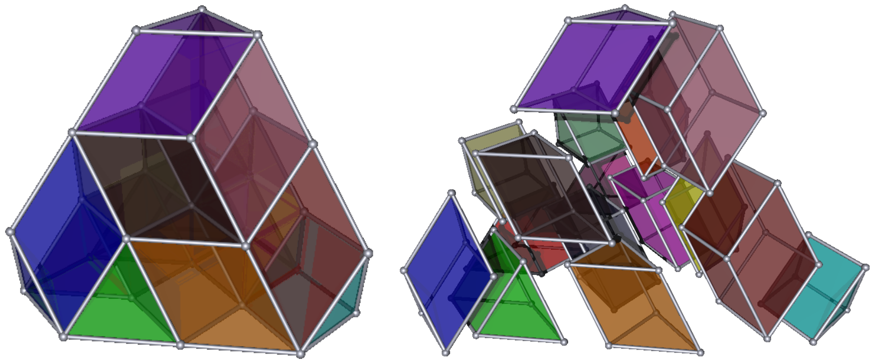

Figure 4: Initial monomial ideals of the toric multiview variety correspond to mixed subdivisions of the truncated tetrahedron P. These have 4 cubes and 12 triangular prisms.

The matrix $\mathcal{A}$ has rank 7 and its columns determine a 6-dimensional polytope $\operatorname{conv}(\mathcal{A})$ with 12 vertices. The normalized volume of $\operatorname{conv}(\mathcal{A})$ equals 16, and this is the degree of the 6-dimensional projective toric variety in $\mathbb{P}^{11}$ defined by J_A. In our context, we do not care about the 6-dimensional variety in $\mathbb{P}^{11}$, but are interested in the threefold in $\mathbb{P}^2\times\mathbb{P}^2\times\mathbb{P}^2\times\mathbb{P}^2$ cut out by J_A. To study this combinatorially, we apply the *Cayley trick*. This means we replace the 6-dimensional polytope $\operatorname{conv}(\mathcal{A})$ by the 3-dimensional polytope

$$P = \operatorname{conv}(A_1^T) + \operatorname{conv}(A_2^T) + \operatorname{conv}(A_3^T) + \operatorname{conv}(A_4^T).$$

This is the Minkowski sum of the four triangles that form the facets of the standard tetrahedron. Equivalently, P is the scaled tetrahedron $4\Delta_3$ with its vertices sliced off. Triangulations of $\mathcal{A}$ correspond to mixed subdivisions of P. Each 6-simplex in $\mathcal{A}$ becomes a cube or a triangular prism in P. Each mixed subdivision has four cubes $\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1$ and twelve triangular prisms $\mathbb{P}^2 \times \mathbb{P}^1$. Such a mixed subdivision of P is shown in Figure 4. Note the similarities and differences relative to the complex $V(M_4)$ in Example 3.4.

We worked out a complete classification of all mixed subdivisions of P:

Theorem 4.3 *The truncated tetrahedron P has* 1068 *mixed subdivisions, one for each triangulation of the Cayley polytope* $\operatorname{conv}(\mathcal{A})$. *Precisely* 1002 *of the* 1068 *triangulations are regular. The regular triangulations form* 48 *symmetry classes, and the non-regular triangulations form* 7 *symmetry classes.*

We offer a brief discussion of this result and how it was obtained. Using the software `Gfan` [15], we found that $I_{\mathcal{A}}$ has 1002 distinct monomial initial ideals. These ideals fall into 48 symmetry classes under the natural action of $(S_3)^4 \rtimes S_4$ on $K[x, y, z]$, where the i-th copy of S_3 permutes the variables x_i, y_i, z_i, and S_4 permutes the labels of the cameras. The matrix $\mathcal{A}$ being unimodular, each initial ideal of $I_{\mathcal{A}}$ is squarefree and each triangulation of $\mathcal{A}$ is unimodular. To calculate all non-regular triangulations, we used the bijection between triangulations and $\mathcal{A}$-graded monomial ideals in

https://doi.org/10.4153/CJM-2012-023-2 Published online by Cambridge University Press

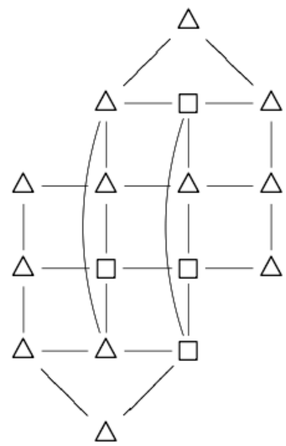

Figure 5: The dual graph of the mixed subdivision given by Y_1.

[20, Lemma 10.14]. Namely, we ran a second computation using the software package `CaTS` [14] that lists all $\mathcal{A}$-graded monomials ideals, and we found their number to be 1068, and hence $\mathcal{A}$ has 66 non-regular triangulations.

The 48 distinct initial monomial ideals of the toric multiview ideal J_A can be distinguished by various invariants. First, their numbers of generators range from 12 to 15. There is precisely one initial ideal with 12 generators:

$$\begin{aligned} Y_1 = \langle & y_1z_2, z_1y_3, x_1z_4, z_2x_3, y_2x_4, x_3y_4, \\ & x_1y_2x_3, z_1y_2x_3, x_1z_2x_4, z_1x_3z_4, z_2y_3x_4, z_2y_3z_4 \rangle. \end{aligned}$$

At the other extreme, there are two classes of initial ideals with 15 generators. These are the only classes having quartic generators, as all ideals with $\leq$ 14 generators require only quadrics and cubics. A representative is

$$\begin{aligned} Y_2 = \langle & z_1y_2, x_1z_3, x_1z_4, x_2z_3, y_2x_4, y_3x_4, \ y_1z_2x_3y_4, \\ & x_1y_2x_3, x_1z_2x_3, x_1z_2x_4, x_4z_2y_1, y_1z_3x_4, y_1z_3y_4, y_2x_3y_4, y_2z_3y_4 \rangle. \end{aligned}$$

All non-regular $\mathcal{A}$-graded monomial ideal have 14 generators. One of them is

$$\begin{aligned} Y_3 = \langle & z_1y_2, z_1y_3, x_1z_4, x_2z_3, x_2z_4, y_3x_4, \ x_1y_2z_3, y_1x_2y_3, \\ & x_1y_2x_4, x_1z_2x_4, x_1z_3x_4, y_1z_3x_4, y_2z_3x_4, y_2z_3y_4 \rangle. \end{aligned}$$

A more refined combinatorial invariant of the 55 types is the dual graph of the mixed subdivision of P. The 16 vertices of this graph are labeled with squares and triangles to denote cubes and triangular prisms respectively, and edges represent common facets. The graph for Y_1 is shown in Figure 5.

For complete information on the classification in Theorem 4.3, see the website www.math.washington.edu/~aholtc/HilbertScheme.

That website also contains the same information for the toric multiview variey in the easier case of $n = 3$ cameras. Taking A_1, A_2, and A_3 as camera matrices, the corresponding Cayley matrix has format 7×9 and rank 6:

$$\mathcal{A} = \begin{bmatrix} A_1^T & A_2^T & A_3^T \\ \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} \end{bmatrix} = \begin{bmatrix} 0&0&0&1&0&0&1&0&0 \\ 1&0&0&0&0&0&0&1&0 \\ 0&1&0&0&1&0&0&0&0 \\ 0&0&1&0&0&1&0&0&1 \\ 1&1&1&0&0&0&0&0&0 \\ 0&0&0&1&1&1&0&0&0 \\ 0&0&0&0&0&0&1&1&1 \end{bmatrix}.$$

This is the transpose of the matrix $A_{\{123\}}$ in (2.2) when evaluated at $x_1 = y_1 = \cdots = z_3 = 1$. The corresponding 6-dimensional Cayley polytope $\mathrm{conv}(\mathcal{A})$ has 9 vertices and normalized volume 7, and the toric multiview ideal equals

$$J_A = \langle z_1 y_3 - x_1 z_3, z_2 x_3 - x_2 z_3, z_1 y_2 - y_1 z_2, x_1 y_2 x_3 - y_1 x_2 y_3 \rangle.$$

We note that the quadrics cut out V_A plus an extra component $\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1$:

$$\langle z_1 y_3 - x_1 z_3, z_2 x_3 - x_2 z_3, z_1 y_2 - y_1 z_2 \rangle = J_A \cap \langle z_1, z_2, z_3 \rangle.$$

This equation is precisely [12, Theorem 5.6] but written in toric coordinates.

The toric ideal J_A has precisely 20 initial monomial ideals, in three symmetry classes, one for each mixed subdivision of the 3-dimensional polytope

$$P = \mathrm{conv}(A_1^T) + \mathrm{conv}(A_2^T) + \mathrm{conv}(A_3^T).$$

Thus P is the Minkowski sum of three of the four triangular facets of the regular tetrahedron. Each mixed subdivision of P uses one cube $\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1$ and six triangular prisms $\mathbb{P}^2 \times \mathbb{P}^1$. A picture of one of them is seen in Figure 1.

Remark 4.4 Our toric study in this section is universal in the sense that **every** multiview variety V_A for $n \leq 4$ cameras in linearly general position in $\mathbb{P}^3$ is isomorphic to the toric multiview variety under a change of coordinates in $(\mathbb{P}^2)^n$. This fact can be proved using the coordinate systems for the Grassmannian $\mathrm{Gr}(4, 3n)$ furnished by the construction in [21, §4]. Here is how it works for $n = 4$. The coordinate change via $\mathrm{PGL}(3, K)^4$ gives

$$\begin{bmatrix} A_1^T & A_2^T & A_3^T & A_4^T \end{bmatrix} = \begin{bmatrix} 0&0&0&*&*&*&*&*&*&*&*&* \\ *&*&*&0&0&0&*&*&*&*&*&* \\ *&*&*&*&*&*&0&0&0&*&*&* \\ *&*&*&*&*&*&*&*&*&0&0&0 \end{bmatrix}, \tag{4.3}$$

where the 3×3-matrices indicated by the stars in the four blocks are invertible. Now, the 4×12-matrix (4.3) gives a *support set* Σ that satisfies the conditions in [21, Proposition 3.1]. The corresponding Zariski open set $\mathcal{U}_\Sigma$ of the Grassmannian $\mathrm{Gr}(4, 12)$

https://doi.org/10.4153/CJM-2012-023-2 Published online by Cambridge University Press

is non-empty. In fact, by [21, Remark 4.9(a)], the set $\mathcal{U}_\Sigma$ represents configurations whose cameras f_1, f_2, f_3, f_4 are not coplanar. Now, [21, Theorem 4.6] completes our proof because (the universal Gröbner basis of) the ideal J_A depends only on the point in $\mathcal{U}_\Sigma \subset \mathrm{Gr}(4, 12)$ represented by (4.3) and not on the specific camera matrices $A_1, \ldots, A_4$. ■

5 Degeneration of Collinear Cameras

In this section we consider a family of collinear camera positions. The degeneration of the associated multiview variety will play a key role in proving our main results in Section 6, but they may also be of independent interest. Collinear cameras have been studied in computer vision, for example, in [11].

Let ε be a parameter and fix the configuration $A(\varepsilon) := (A_1, \ldots, A_n)$ where

$$A_i := \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ \varepsilon^{n-i} & 0 & 0 & 1 \end{bmatrix}$$

The focal point of camera i is $f_i = (-1{:}1{:}1{:}\varepsilon^{n-i})$, and hence the n cameras given by $A(\varepsilon)$ are collinear in $\mathbb{P}^3$. Note that these camera matrices stand in sharp contrast to those for which A is generic, which was the focus of Sections 2 and 3. They also differ from the toric situation in Section 4.

We consider the multiview ideal $J_{A(\varepsilon)}$ in the polynomial ring $K(\varepsilon)[x, y, z]$, where $K(\varepsilon)$ is the field of rational functions in ε with coefficients in K. Then $J_{A(\varepsilon)}$ has the Hilbert function (3.3) by Theorem 3.7. Let $\mathcal{G}_n$ be the set of polynomials in $K(\varepsilon)[x, y, z]$ consisting of the $\binom{n}{2}$ quadratic polynomials

$$x_i y_j - x_j y_i \qquad \text{for } 1 \le i < j \le n \tag{5.1}$$

and the $3\binom{n}{3}$ cubic polynomials below for all choices of $1 \le i < j < k \le n$:

$$\begin{gathered} (\varepsilon^{n-k} - \varepsilon^{n-i})x_i z_j x_k + (\varepsilon^{n-j} - \varepsilon^{n-k})z_i x_j x_k + (\varepsilon^{n-i} - \varepsilon^{n-j})x_i x_j z_k \\ (\varepsilon^{n-k} - \varepsilon^{n-i})y_i z_j y_k + (\varepsilon^{n-j} - \varepsilon^{n-k})z_i y_j y_k + (\varepsilon^{n-i} - \varepsilon^{n-j})y_i y_j z_k \\ (\varepsilon^{n-k} - \varepsilon^{n-i})y_i z_j x_k + (\varepsilon^{n-j} - \varepsilon^{n-k})z_i y_j x_k + (\varepsilon^{n-i} - \varepsilon^{n-j})y_i x_j z_k \end{gathered} \tag{5.2}$$

Let L_n be the ideal generated by (5.1) and the following binomials from the first two terms in (5.2):

$$L_n := \left\langle x_i y_j - x_j y_i : 1{\le}i{<}j{\le}n \right\rangle + \left\langle \begin{matrix} x_i z_j x_k - z_i x_j x_k, \\ y_i z_j y_k - z_i y_j y_k, \\ y_i z_j x_k - z_i y_j x_k \end{matrix} \;:\; 1{\le}i{<}j{<}k{\le}n \right\rangle.$$

Let N_n be the ideal generated by the leading monomials in (5.1) and (5.2):

$$N_n := \left\langle x_i y_j : 1{\le}i{<}j{\le}n \right\rangle + \left\langle x_i z_j x_k,\, y_i z_j y_k,\, y_i z_j x_k : 1{\le}i{<}j{<}k{\le}n \right\rangle.$$

The main result in this section is the following construction of a two-step flat degeneration $J_{A(\varepsilon)} \to L_n \to N_n$. This gives an explicit realization of (1.2). We note that $V_{A(\varepsilon)}$ can be seen as a variant of the *Mustafin varieties* in [2].

Theorem 5.1 *The three ideals $J_{A(\varepsilon)}$, L_n and N_n satisfy the following:*

(i) *The multiview ideal $J_{A(\varepsilon)}$ is generated by the set $\mathcal{G}_n$.*

(ii) *The binomial ideal L_n equals the special fiber of $J_{A(\varepsilon)}$ for $\varepsilon = 0$.*

(iii) *The monomial ideal N_n is the initial ideal of L_n, in the Gröbner basis sense, with respect to the lexicographic term order with $x \succ y \succ z$.*

The rest of this section is devoted to explaining and proving these results. Let us begin by showing that $\mathcal{G}_n$ is a subset of $J_{A(\varepsilon)}$. The determinant of

$$A(\varepsilon)_{\{ij\}} = \begin{bmatrix} A_i & p_i & \mathbf{0} \\ A_j & \mathbf{0} & p_j \end{bmatrix}$$

equals $(\varepsilon^{n-j} - \varepsilon^{n-i})(x_i y_j - x_j y_i)$. Hence $J_{A(\varepsilon)}$ contains (5.1), by the argument in Lemma 2.2. Similarly, for any $1 \leq i < j < k \leq n$, consider the 9×7 matrix

$$A(\varepsilon)_{\{ijk\}} = \begin{bmatrix} 1 & 1 & 0 & 0 & x_i & 0 & 0 \\ 1 & 0 & 1 & 0 & y_i & 0 & 0 \\ \varepsilon^{n-i} & 0 & 0 & 1 & z_i & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & x_j & 0 \\ 1 & 0 & 1 & 0 & 0 & y_j & 0 \\ \varepsilon^{n-j} & 0 & 0 & 1 & 0 & z_j & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & x_k \\ 1 & 0 & 1 & 0 & 0 & 0 & y_k \\ \varepsilon^{n-k} & 0 & 0 & 1 & 0 & 0 & z_k \end{bmatrix}.$$

The three cubics (5.2), in this order and up to sign, are the determinants of the 7×7 submatrices of $A(\varepsilon)_{\{ijk\}}$ obtained by deleting the rows corresponding to y_j and y_k, the rows corresponding to x_j and x_k, and the rows corresponding to x_i and y_k respectively. We conclude that $\mathcal{G}_n$ lies in $J_{A(\varepsilon)}$.

We next discuss Theorem 5.1(ii). Every rational function $c(\varepsilon) \in K(\varepsilon)$ has a unique expansion as a Laurent series $c_1\varepsilon^{a_1} + c_2\varepsilon^{a_2} + \cdots$ where $c_i \in K$ and $a_1 < a_2 < \cdots$ are integers. The function $\mathrm{val}\colon K(\varepsilon) \to \mathbb{Z}$ given by $c(\varepsilon) \mapsto a_1$ is then a valuation on $K(\varepsilon)$, and $K[[\varepsilon]] = \{c \in K(\varepsilon) : \mathrm{val}(c) \geq 0\}$ is its valuation ring. The unique maximal ideal in $K[[\varepsilon]]$ is $m = \langle c \in K(\varepsilon) : \mathrm{val}(c) > 0 \rangle$. The residue field $K[[\varepsilon]]/m$ is isomorphic to K, so there is a natural map $K[[\varepsilon]] \to K$ that represents the evaluation at $\varepsilon = 0$. The *special fiber* of an ideal $I \subset K(\varepsilon)[x, y, z]$ is the image of $I \cap K[[\varepsilon]][x, y, z]$ under the induced map $K[[\varepsilon]][[x, y, z] \to K[x, y, z]$. The special fiber is denoted $\mathrm{in}(I)$. It can be computed from I by a variant of Gröbner bases (cf. [16, §2.4]).

What we are claiming in Theorem 5.1(ii) is the following identify

$$\mathrm{in}(J_{A(\varepsilon)}) = L_n \qquad \text{in } K[x, y, z].$$

It is easy to see that the left-hand side contains the right-hand side; indeed, by multiplying the trinomials in (5.2) by ε^{k-n} and then evaluating at $\varepsilon = 0$, we obtain the binomial cubics among the generators of L_n.

Finally, what is claimed in Theorem 5.1(iii) is the following identity:

$$\mathrm{in}_{\prec}(L_n) = N_n \qquad \text{in } K[x, y, z].$$

Here, $\mathrm{in}_{\prec}(L_n)$ is the lexicographic initial ideal of L_n, in the usual Gröbner basis sense. Again, the left-hand side contains the right-hand side because the initial monomials of the binomial generators of L_n generate N_n.

Note that N_n is distinct from the generic initial ideal M_n. Even though M_n played a prominent role in Sections 2 and 3, the ideal N_n will be more useful in Section 6. The reason is that M_n is the most singular point on the Hilbert scheme $\mathcal{H}_n$ while, as we shall see, N_n is a smooth point on $\mathcal{H}_n$.

In summary, what we have shown thus far is the following inclusion:

$$N_n \subseteq \mathrm{in}_{\prec}\big(\mathrm{in}(J_{A(\varepsilon)})\big)\,. \tag{5.3}$$

We seek to show that equality holds. Our proof rests on the following lemma.

Lemma 5.2 *The monomial ideal N_n has the $\mathbb{Z}^n$-graded Hilbert function* (3.3).

Proof Let $u = (u_1, \dots, u_n) \in \mathbb{N}^n$, and let $\mathfrak{B}_u$ be the set of all monomials of multidegree u in $K[x, y, z]$ that are not in N_n. We need to show that

$$|\mathfrak{B}_u| = \binom{u_1 + \cdots + u_n + 3}{3} - \sum_{i=1}^{n} \binom{u_i + 2}{3}.$$

It can be seen from the generators of N_n that the monomials in $\mathfrak{B}_u$ are of the form $z^a y^b x^c z^d$ for $a, b, c, d \in \mathbb{N}^n$ such that $u = a + b + c + d$ and

$$\begin{aligned} a &= (a_1, \dots, a_i, 0, \dots, 0)\\ b &= (0, \dots, 0, b_i, \dots, b_j, 0, \dots, 0)\\ c &= (0, \dots, 0, c_j, \dots, c_k, 0, \dots, 0)\\ d &= (0, \dots, 0, d_k, \dots, d_n)\end{aligned}$$

for some triple i, j, k with $1 \le i \le j \le k \le n$.

We count the monomials in $\mathfrak{B}_u$ using a combinatorial "stars and bars" argument. Each monomial can be formed in the following way. Suppose there are $u_1 + \cdots + u_n + 3$ blank spaces laid left to right. Fill exactly three spaces with bars. This leaves $u_1 + \cdots + u_n$ open blanks to fill in, which is the total degree of a monomial in $\mathfrak{B}_u$. The three bars separate the blanks into four compartments, some possibly empty. From these compartments we greedily form a, b, c, and d to make $z^a y^b x^c z^d$ as described below.

In what follows, $\star$ is used as a placeholder symbol. Fill the first u_1 blanks with the symbol $\star_1$, the next u_2 blanks with $\star_2$, and continue to fill up until the last u_n blanks are filled with $\star_n$. Now we pass once more through these symbols and replace each $\star_i$ with either x_i, y_i, or z_i such that all variables in the first compartment are z's, those in the second are y's, then x's and in the fourth compartment z's. Removing the bars gives $z^a y^b x^c z^d$ in $\mathfrak{B}_u$.

There are $\binom{u_1+\cdots+u_n+3}{3}$ ways of choosing the three bars. The monomials in $\mathfrak{B}_u$ are overcounted only when $i = j = k$ if z_i appears in both the first and fourth compartments. Indeed, in such cases if we require $a_i = 0$, the monomial is uniquely represented, so we are overcounting by the $\binom{u_i+2}{3}$ choices when $a_i \neq 0$. ∎

https://doi.org/10.4153/CJM-2012-023-2 Published online by Cambridge University Press

We are now prepared to derive the main result of this section.

Proof of Theorem 5.1 Lemma 5.2 and Theorem 3.7 tell us that N_n and $J_{A(\varepsilon)}$ have the same $\mathbb{Z}^n$-graded Hilbert function (3.3). We also know from [16, §2.4] that $\mathrm{in}(J_{A(\varepsilon)})$ has the same Hilbert function, just as passing to an initial monomial ideal for a term order preserves Hilbert function. Hence the equality $N_n \subseteq \mathrm{in}_\prec\big(\mathrm{in}(J_{A(\varepsilon)})\big)$ holds in (5.3). This proves parts (ii) and (iii). We have shown that $\mathcal{G}_n$ is a Gröbner basis for the homogeneous ideal $J_{A(\varepsilon)}$ in the valuative sense of [16, §2.4]. This implies that $\mathcal{G}_n$ generates $J_{A(\varepsilon)}$. ■

Remark 5.3 The polyhedral subcomplexes of $(\Delta_2)^n$ defined by the binomial ideal L_n and the monomial ideal N_n are combinatorially interesting. For instance, L_n has prime decomposition $I_3 \cap I_4 \cap \cdots \cap I_n \cap I_{n+1}$, where

$$\begin{aligned} I_t := \langle x_i, y_i : i = t, t+1, \dots, n\rangle + \langle x_i y_j - x_j y_i : 1 \le i < j < t\rangle \\ + \langle x_i z_j - x_j z_i, y_i z_j - y_j z_i : 1 \le i < j < t-1\rangle. \end{aligned}$$

The monomial ideal N_n is the intersection of $\mathrm{in}_\prec(I_t)$ for $t = 3, \dots, n+1$.

6 The Hilbert Scheme

We define $\mathcal{H}_n$ to be the multigraded Hilbert scheme that parametrizes all $\mathbb{Z}^n$-homogeneous ideals in $K[x, y, z]$ with the Hilbert function in (3.3). According to the general construction given in [10], $\mathcal{H}_n$ is a projective scheme. The ideals J_A and $\mathrm{in}_\prec(J_A)$ for n distinct camera positions, as well as the combinatorial ideals M_n, L_n and N_n all correspond to closed points on $\mathcal{H}_n$.

Our Hilbert scheme $\mathcal{H}_n$ is closely related to the Hilbert scheme $H_{4,n}$ which was studied in [3]. We already utilized results from that paper in our proof of Theorem 2.1. Note that $H_{4,n}$ parametrizes degenerations of the diagonal $\mathbb{P}^3$ in $(\mathbb{P}^3)^n$ while $\mathcal{H}_n$ parametrizes blown-up images of that $\mathbb{P}^3$ in $(\mathbb{P}^2)^n$.

Let $G = \mathrm{PGL}(3, K)$ and $\mathcal{B} \subset G$ the Borel subgroup of lower-triangular 3×3 matrices modulo scaling. The group G^n acts on $K[x, y, z]$ and this induces an action on the Hilbert scheme $\mathcal{H}_n$. Our results concerning the ideal M_n in Section 3 imply the following corollary, which summarizes the statements analogous to [3, Theorem 2.1 and Corollaries 2.4 and 2.6].

Corollary 6.1 *The multigraded Hilbert scheme $\mathcal{H}_n$ is connected. The point representing the generic initial ideal M_n lies on each irreducible component of $\mathcal{H}_n$. All ideals that lie on $\mathcal{H}_n$ are radical and Cohen–Macaulay.*

In particular, every monomial ideal in $\mathcal{H}_n$ is squarefree and can hence be identified with its variety in $(\mathbb{P}^2)^n$, or, equivalently, with a subcomplex in the product of triangles $(\Delta_2)^n$. One of the first questions one asks about any multigraded Hilbert scheme, including $\mathcal{H}_n$, is to list its monomial ideals.

This task is easy for the first case, $n = 2$. The Hilbert scheme $\mathcal{H}_2$ parametrizes $\mathbb{Z}^2$-homogeneous ideals in $K[x, y, z]$ having Hilbert function

$$h_2 \colon \mathbb{N}^2 \to \mathbb{N}, (u_1, u_2) \mapsto \binom{u_1 + u_2 + 3}{3} - \binom{u_1 + 2}{3} - \binom{u_2 + 2}{3}.$$

https://doi.org/10.4153/CJM-2012-023-2 Published online by Cambridge University Press

There are exactly nine monomial ideals on $\mathcal{H}_2$, namely

$$\langle x_1x_2\rangle, \langle x_1y_2\rangle, \langle x_1z_2\rangle, \langle y_1x_2\rangle, \langle y_1y_2\rangle, \langle y_1z_2\rangle, \langle z_1x_2\rangle, \langle z_1y_2\rangle, \langle z_1z_2\rangle.$$

In fact, the ideals on $\mathcal{H}_2$ are precisely the principal ideals generated by bilinear forms, and $\mathcal{H}_2$ is isomorphic to an 8-dimensional projective space

$$\mathcal{H}_2 = \left\{ \langle c_0x_1x_2 + c_1x_1y_2 + \cdots + c_8z_1z_2\rangle : (c_0 : c_1 : \cdots : c_8) \in \mathbb{P}^8 \right\}.$$

The principal ideals J_A that actually arise from two cameras form a cubic hypersurface in this $\mathcal{H}_2 \simeq \mathbb{P}^8$. To see this, we write A_i^j for the j-th row of the i-th camera matrix and $[A_{i_1}^{j_1}A_{i_2}^{j_2}A_{i_3}^{j_3}A_{i_4}^{j_4}]$ for the 4×4-determinant formed by four such row vectors. The bilinear form can be written as

$$\mathbf{x}_2^T F \mathbf{x}_1 = \begin{bmatrix} x_2 & y_2 & z_2 \end{bmatrix} \begin{bmatrix} c_0 & c_3 & c_6 \\ c_1 & c_4 & c_7 \\ c_2 & c_5 & c_8 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix},$$

where F is the *fundamental matrix* [11]. In terms of the camera matrices,

$$F = \begin{bmatrix} [A_1^2A_1^3A_2^2A_2^3] & -[A_1^1A_1^3A_2^2A_2^3] & [A_1^1A_1^2A_2^2A_2^3] \\ -[A_1^2A_1^3A_2^1A_2^3] & [A_1^1A_1^3A_2^1A_2^3] & -[A_1^1A_1^2A_2^1A_2^3] \\ [A_1^2A_1^3A_2^1A_2^2] & -[A_1^1A_1^3A_2^1A_2^2] & [A_1^1A_1^2A_2^1A_2^2] \end{bmatrix}. \tag{6.1}$$

This matrix has rank ≤ 2, and every 3×3-matrix of rank ≤ 2 can be written in this form for suitable camera matrices A_1 and A_2 of size 3×4.

The formula in (6.1) defines a map $(A_1, A_2) \mapsto F$ from pairs of camera matrices with distinct focal points into the Hilbert scheme $\mathcal{H}_2$. The closure of its image is a compactification of the space of camera positions. We now precisely define the corresponding map for arbitrary $n \geq 2$. The construction is inspired by the construction due to Thaddeus discussed in [3, Example 7].

Let $\mathrm{Gr}(4, 3n)$ denote the Grassmannian of 4-dimensional linear subspaces of K^{3n}. The n-dimensional algebraic torus $(K^*)^n$ acts on this Grassmannian by scaling the coordinates on K^{3n}, where the i-th factor K^* scales the coordinates indexed by $3i-2$, $3i-1$, and $3i$. Thus, if we represent each point in $\mathrm{Gr}(4, 3n)$ as the row space of a $(4 \times 3n)$-matrix $\begin{bmatrix} A_1^T & A_2^T & \cdots & A_n^T \end{bmatrix}$, then $\lambda = (\lambda_1, \ldots, \lambda_n) \in (K^*)^n$ sends this matrix to $\begin{bmatrix} \lambda_1 A_1^T & \lambda_2 A_2^T & \cdots & \lambda_n A_n^T \end{bmatrix}$. The multiview ideal J_A is invariant under this action by $(K^*)^n$. In symbols, $J_{\lambda \circ A} = J_A$. In the next lemma, GIT stands for *geometric invariant theory*.

Lemma 6.2 *The assignment $A \mapsto J_A$ defines an injective rational map γ from a GIT quotient* $\mathrm{Gr}(4, 3n)//(K^*)^n$ *to the multigraded Hilbert scheme $\mathcal{H}_n$.*

Proof For the proof it suffices to check that $J_A \neq J_{A'}$ whenever A and A' are generic camera configurations that are not in the same $(K^*)^n$-orbit. ∎

https://doi.org/10.4153/CJM-2012-023-2 Published online by Cambridge University Press

We call γ the camera map. Since we need γ only as a rational map, the choice of linearization does not matter when we form the GIT quotient. The closure of its image in $\mathcal{H}_n$ is well defined and independent of that choice of linearization. We define the *compactified camera space*, for n cameras, to be

$$\Gamma_n := \overline{\gamma(\mathrm{Gr}(4,3n)//(K^*)^n)} \subseteq \mathcal{H}_n.$$

The projective variety Γ_n is a natural compactification of the parameter space studied by Heyden in [13]. Since the torus $(K^*)^n$ acts on $\mathrm{Gr}(4,3n)$ with a one-dimensional stabilizer, Lemma 6.2 implies that the compactified space of n cameras has the dimension we expect from [13], namely,

$$\dim(\Gamma_n) = \dim(\mathrm{Gr}(4,3n)) - (n-1) = 4(3n-4) - (n-1) = 11n - 15.$$

We regard the following theorem as the main result in this paper.

Theorem 6.3 *For $n \geq 3$, the compactified camera space Γ_n appears as a distinguished irreducible component in the multigraded Hilbert scheme $\mathcal{H}_n$.*

Note that the same statement if false for $n = 2$: Γ_2 is not a component of $\mathcal{H}_3 \simeq \mathbb{P}^8$. It is the hypersurface consisting of the fundamental matrices (6.1).

Proof By definition, the compactified camera space Γ_n is a closed subscheme of $\mathcal{H}_n$. The discussion above shows that the dimension of any irreducible component of $\mathcal{H}_n$ that contains Γ_n is no smaller than $11n - 15$. We shall now prove the same $11n - 15$ as an upper bound for the dimension. This is done by exhibiting a point in Γ_n whose tangent space in the Hilbert scheme $\mathcal{H}_n$ has dimension $11n - 15$. This will imply the assertion.

For any ideal $I \in \mathcal{H}_n$, the tangent space to the Hilbert scheme $\mathcal{H}_n$ at I is the space of $K[x, y, z]$-module homomorphisms $I \to K[x, y, z]/I$ of degree $\mathbf{0}$. In symbols, this space is $\mathrm{Hom}(I, K[x, y, z]/I)_{\mathbf{0}}$. The K-dimension of the tangent space provides an upper bound for the dimension of any component on which I lies. It remains to specifically identify a point on Γ_n that is smooth on $\mathcal{H}_n$, an ideal that has tangent space dimension exactly $11n - 15$.

It turns out that the monomial ideal N_n described in the previous section has this desired property. Lemmas 6.4 and 6.5 give the details. ∎

Lemma 6.4 *The ideals L_n and N_n from the previous section lie in Γ_n.*

Proof The image of γ in $\mathcal{H}_n$ consists of all multiview ideals J_A, where A runs over configurations of n distinct cameras, by Theorem 3.7. Let $A(\varepsilon)$ denote the collinear configuration in Section 5 and consider any specialization of ε to a non-zero scalar in K. The resulting ideal $J_{A(\varepsilon)}$ is a K-valued point of Γ_n, for any $\varepsilon \in K \backslash \{0\}$. The special fiber $J_{A(0)} = L_n$ is in the Zariski closure of these points, because, locally, any regular function vanishing on the coordinates of $J_{A(\varepsilon)}$ for all $\varepsilon \neq 0$ will vanish for $\varepsilon = 0$. We conclude that L_n is a K-valued point in the projective variety Γ_n. Likewise, since $N_n = \mathrm{in}_{\prec}(L_n)$ is an initial monomial ideal of L_n, it also lies on Γ_n. ∎

Lemma 6.5 *The tangent space of the multigraded Hilbert scheme $\mathcal{H}_n$ at the point represented by the monomial ideal N_n has dimension $11n - 15$.*

Proof The tangent space at N_n equals $\operatorname{Hom}(N_n, K[x, y, z]/N_n)_{\mathbf{0}}$. We shall present a basis for this space that is broken into three distinct classes: those homomorphisms that act nontrivially only on the quadratic generators, those that act nontrivially only on the cubics, and those with a mix of both.

Each $K[x, y, z]$-module homomorphism $\varphi\colon N_n \to K[x, y, z]/N_n$ below is described by its action on the minimal generators of N_n. Any generator not explicitly mentioned is mapped to 0 under φ. One checks that each is in fact a well-defined $K[x, y, z]$-module homomorphism from N_n to $K[x, y, z]/N_n$.

Class I: For each $1 \le i < n$, we define the following maps:

- $\alpha_i\colon x_i y_k \mapsto y_i y_k$ for all $i < k \le n$,
- $\beta_i\colon x_i y_{i+1} \mapsto x_{i+1} y_i$.

For each $1 < k \le n$, we define the following map:

- $\gamma_k\colon x_i y_k \mapsto x_i x_k$ for all $1 \le i < k$.

We define two specific homomorphisms:

- $\delta_1\colon x_1 y_2 \mapsto y_1 z_2$,
- $\delta_2\colon x_{n-1} y_n \mapsto z_{n-1} x_n$.

Class II: For each $1 < j < n$, we define the following maps. Each homomorphism is defined on every pair (i, k) such that $1 \le i < j < k \le n$.

- $\rho_j\colon x_i z_j x_k \mapsto x_i x_j x_k$ and $y_i z_j x_k \mapsto y_i x_j x_k$,
- $\sigma_j\colon x_i z_j x_k \mapsto x_i x_j z_k$ and $y_i z_j x_k \mapsto y_i x_j z_k$,
- $\tau_j\colon x_i z_j x_k \mapsto x_i z_j z_k$ and $y_i z_j x_k \mapsto y_i z_j z_k$,
- $\nu_j\colon y_i z_j x_k \mapsto y_i y_j x_k$ and $y_i z_j y_k \mapsto y_i y_j y_k$,
- $\mu_j\colon y_i z_j x_k \mapsto z_i y_j x_k$ and $y_i z_j y_k \mapsto z_i y_j y_k$,
- $\pi_j\colon y_i z_j x_k \mapsto z_i z_j x_k$ and $y_i z_j y_k \mapsto z_i z_j y_k$.

Class III: For each $1 \le i < n$, we define the map

- $\epsilon_i\colon x_i y_k \mapsto z_i y_k$ and $x_i z_j x_k \mapsto z_i z_j x_k$ for $i < k \le n$ and $i < j < k$.

For each $1 < k \le n$, we define the map

- $\zeta_k\colon x_i y_k \mapsto x_i z_k$ and $y_i z_j y_k \mapsto y_i z_j z_k$ for $1 \le i < k$ and $i < j < k$.

All of these maps are linearly independent over the field K. There are $n - 1$ maps each of type α_i, β_i, γ_k, ϵ_i, and ζ_k, for a total of $5(n - 1)$ different homomorphisms. Each subclass of maps in class II has $n - 2$ members, adding $6(n - 2)$ more homomorphisms. Finally adding δ_1 and δ_2, we arrive at the total count of $5(n - 1) + 6(n - 2) + 2 = 11n - 15$ homomorphisms.

We claim that any $K[x, y, z]$-module homomorphism $N_n \to K[x, y, z]/N_n$ can be recognized as a K-linear combination of those from the three classes described above. To prove this, suppose that $\varphi\colon N_n \to K[x, y, z]/N_n$ is a module homomorphism. For $1 \le i < k \le n$, we can write $\varphi(x_i y_k)$ as a linear combination of monomials of multidegree $e_i + e_k$ that are not in N_n. By subtracting appropriate multiples of α_i, ϵ_i, γ_k, and ζ_k, we can assume that

$$\varphi(x_i y_k) = a y_i x_k + b y_i z_k + c z_i x_k + d z_i z_k$$

for some scalars $a, b, c, d \in K$. We show that this can be written as a linear combination of the maps described above by considering a few cases.

In the first case we assume $i + 1 < k$. We use $K[x, y, z]$-linearity to infer

$$\varphi(x_i y_{i+1} y_k) = a y_i y_{i+1} x_k + b y_i y_{i+1} z_k + c z_i y_{i+1} x_k + d z_i y_{i+1} z_k = y_k \varphi(x_i y_{i+1}).$$

Specifically, y_k divides the middle polynomial. But none of the four monomials are zero in the quotient $K[x, y, z]/N_n$. Hence, $0 = a = b = c = d$.

For the subsequent cases we assume $k = i+1$. This allows us to further assume that $a = 0$, since we can subtract $a\beta_i(x_i y_{i+1})$. Now suppose that we have strict inequality $k < n$. As before, the $K[x, y, z]$-linearity of φ gives

$$\varphi(x_i y_k y_n) = d z_i z_k y_n = y_k \varphi(x_i y_n).$$

Specifically, y_k divides the middle term. Hence, $d = 0$. Similarly, $c = 0$:

$$\varphi(x_i y_k z_k x_n) = c z_i x_k z_k x_n = y_k \varphi(x_i z_k x_n).$$

Suppose we further have the strict inequality $1 < i$. Then necessarily $b = 0$:

$$\varphi(y_1 z_i x_i y_k) = b y_1 z_i y_i z_k = x_i \varphi(y_1 z_i y_k).$$

However, if $i = 1$ and $k = 2$, we have that $\varphi(x_1 y_2) = b\delta_1(x_1 y_2)$.

The only case that remains is $k = n$ and $i = n - 1$. Here, we can also assume that $c = 0$ by subtracting $c\delta_2(x_{n-1} y_n)$. We will show that $d = 0 = b$ by once more appealing to the fact that φ is a module homomorphism

$$\varphi(x_1 x_{n-1} y_n) = d x_1 z_{n-1} z_n = x_{n-1} \varphi(x_1 y_n),$$

which gives $d = 0$. This subsequently implies the desired $b = 0$, because

$$\varphi(y_1 x_i z_i y_n) = b y_1 y_i z_i z_n = x_i \varphi(y_1 z_i y_n).$$

This has finally put us in a position where we can assume that $\varphi(x_i y_k) = 0$ for all $1 \le i < k \le n$. To finish the proof that φ is a linear combination of the $11n - 15$ classes described above, we need to examine what happens with the cubics. Suppose $1 \le i < j < k \le n$ and consider $\varphi(y_i z_j x_k)$. This can be written as a linear sum of the 17 standard monomials of multidegree $e_i + e_j + e_k$ that are not in N_n. Explicitly, these standard monomials are:

$$\begin{array}{lllll} x_i x_j x_k, & x_i x_j z_k, & x_i z_j z_k, & y_i x_j x_k, & y_i x_j z_k \\ y_i y_j x_k, & y_i y_j y_k, & y_i y_j z_k, & y_i z_j z_k, & \\ z_i x_j x_k, & z_i x_j z_k, & z_i y_j x_k, & z_i y_j y_k, & \\ z_i y_j z_k, & z_i z_j x_k, & z_i z_j y_k, & z_i z_j z_k. & \end{array}$$

https://doi.org/10.4153/CJM-2012-023-2 Published online by Cambridge University Press

By subtracting multiples of the maps ρ_j, σ_j, τ_j, ν_j, μ_j, and π_j, we can assume that this is a sum of the 11 monomials remaining after removing $y_i x_j x_k$, $y_i x_j z_k$, $y_i z_j z_k$, $y_i y_j x_k$, $z_i y_j x_k$, and $z_i z_j x_k$. However, we now note that

$$\varphi(x_i y_i z_j x_k) = x_i \varphi(y_i z_j x_k) = y_i \varphi(x_i z_j x_k).$$

This means that for every one of the 11 monomials m appearing in the sum, either $x_i m = 0$ or y_i divides m. Similarly,

$$\varphi(y_i z_j x_k y_k) = y_k \varphi(y_i z_j x_k) = x_k \varphi(y_i z_j y_k),$$

and so either $y_k m = 0$ or x_k divides m. Taking these both into consideration actually kills every one of the 11 possible standard monomials (we spare the reader the explicit check), and hence we can assume that $\varphi(y_i z_j x_k) = 0$.

Now consider what happens with $\varphi(x_i z_j x_k)$. Indeed,

$$0 = x_i \varphi(y_i z_j x_k) = \varphi(x_i y_i z_j x_k) = y_i \varphi(x_i z_j x_k).$$

So for every one of the 17 standard monomials m that possibly appears in the support of $\varphi(x_i z_j x_k)$ we must have that $y_i m = 0$ in $K[x, y, z]/N_n$. This actually leaves us with only two possible such standard monomials, namely $z_i z_j x_k$ and $z_i z_j y_k$. We write $\varphi(x_i z_j x_k) = a z_i z_j x_k + b z_i z_j y_k$.

The fact that we assume $\varphi(x_i y_k) = 0$ implies $a = 0 = b$. This is because

$$0 = z_j x_k \varphi(x_i y_k) = \varphi(x_i z_j x_k y_k) = y_k \varphi(x_i z_j x_k).$$

To sum up, we have shown that, under our assumptions, if $\varphi(y_i z_j x_k) = 0$ holds, then it also must be the case that $\varphi(x_i z_j x_k) = 0$. We can prove in a similar manner that $\varphi(y_i z_j y_k) = 0$, and this finishes the proof that φ can be written as a K-linear sum of the $11n - 15$ classes of maps described. ■

We reiterate that Theorem 6.3 fails for $n = 2$, since $\mathcal{H}_2 \simeq \mathbb{P}^8$, and Γ_2 is a cubic hypersurface cutting through $\mathcal{H}_2$. We offer a short report for $n = 3$.

Remark 6.6 The Hilbert scheme $\mathcal{H}_3$ contains 13,824 monomial ideals. These come in 16 symmetry classes under the action of $(S_3)^3 \rtimes S_3$. A detailed analysis of these symmetry classes and how we found the 13,824 ideals appears on the website www.math.washington.edu/$\sim$aholtc/HilbertScheme. For seven of the symmetry classes, the tangent space dimension is less than $\dim(\Gamma_3) = 18$. From this we infer that $\mathcal{H}_3$ has components other than Γ_3.

We note that the number 13,824 is exactly the number of monomial ideals on $H_{3,3}$ as described in [3]. Moreover, the monomial ideals on $H_{3,3}$ also fall into 16 distinct symmetry classes. We do not yet fully understand the relationship between $\mathcal{H}_n$ and $H_{3,n}$ suggested by this observation.

Moreover, it would be desirable to coordinatize the inclusion $\Gamma_3 \subset \mathcal{H}_3$ and to relate it to the equations defining *trifocal tensors*, as seen in [1,13]. It is our intention to investigate this topic in a subsequent publication.

Our study was restricted to cameras that take 2-dimensional pictures of 3-dimensional scenes. Yet, residents of *flatland* might be more interested in taking 1-dimensional pictures of 2-dimensional scenes. From a mathematical perspective, generalizing to arbitrary dimensions makes sense: given n matrices of format $r \times s$ we get a map from $\mathbb{P}^{s-1}$ into $(\mathbb{P}^{r-1})^n$, and one could study the Hilbert scheme parametrizing the resulting varieties. Our focus on $r = 3$ and $s = 4$ was motivated by the context of computer vision.

Acknowledgments Aholt and Thomas thank Fredrik Kahl for hosting them at Lund in February 2011 and pointing them to the work of Heyden and Åström. They also thank Sameer Agarwal for introducing them to problems in computer vision and continuing to advise them in this field. Sturmfels thanks the Mittag-Leffler Institute, where this project started, and MATHEON Berlin for their hospitality. We are indebted to the makers of the software packages `CaTS`, `Gfan`, `Macaulay2`, and `Sage` [19] which allowed explicit computations that were crucial in discovering our results.

References

[1] A. Alzati and A. Tortora, *A geometric approach to the trifocal tensor.* J. Math. Imaging Vision **38**(2010), no. 3, 159–170. http://dx.doi.org/10.1007/s10851-010-0216-4

[2] D. Cartwright, M. Häbich, B. Sturmfels, and A. Werner, *Mustafin varieties.* Selecta Math. **17**(2011), no. 4, 757–793. http://dx.doi.org/10.1007/s00029-011-0075-x

[3] D. Cartwright and B. Sturmfels, *The Hilbert scheme of the diagonal in a product of projective spaces.* Int. Math. Res. Not. **2010**, no. 9, 1741–1771.

[4] A. Conca, *Linear spaces, transversal polymatroids and ASL domains.* J. Algebraic Combin. **25**(2007), no. 1, 25–41. http://dx.doi.org/10.1007/s10801-006-0026-3

[5] D. Cox, J. Little, and D. O'Shea, *Ideals, varieties, and algorithms. An introduction to computational algebraic geometry and commutative algebra.* Third ed. Undergraduate Texts in Mathematics, Springer, New York, 2007.

[6] D. Eisenbud, *Commutative algebra with a view toward algebraic geometry.* Graduate Texts in Mathematics, 150, Springer-Verlag, New York, 1995.

[7] O. Faugeras and Q-T. Luong, *The geometry of multiple images. The laws that govern the formation of multiple images of a scene and some of their applications.* MIT Press, Cambridge, MA, 2001.

[8] D. R. Grayson and M. E. Stillman, *Macaulay2, a software system for research in algebraic geometry.* http://www.math.uiuc.edu/Macaulay2/.

[9] F. D. Grosshans, *On the equations relating a three-dimensional object and its two-dimensional images.* Adv.in Appl. Math. **34**(2005), no. 2, 366–392. http://dx.doi.org/10.1016/j.aam.2004.07.005

[10] M. Haiman and B. Sturmfels, *Multigraded Hilbert schemes.* J. Algebraic Geom. **13**(2004), no. 4, 725–769. http://dx.doi.org/10.1090/S1056-3911-04-00373-X

[11] R. Hartley and A Zisserman, *Multiple view geometry in computer vision.* Second ed. Cambridge University Press, Cambridge, 2003.

[12] A. Heyden and K. Åström, *Algebraic properties of multilinear constraints.* Math. Methods Appl. Sci. **20**(1997), no. 13, 1135–1162. http://dx.doi.org/10.1002/(SICI)1099-1476(19970910)20:13⟨1135::AID-MMA908⟩3.0.CO;2-9

[13] A. Heyden, *Tensorial properties of multiple view constraints.* Math. Methods Appl. Sci. **23**(2000), no. 2, 169–202. http://dx.doi.org/10.1002/(SICI)1099-1476(20000125)23:2⟨169::AID-MMA110⟩3.0.CO;2-Y

[14] A. Jensen, *CaTS, a software system for toric state polytopes.* http://www.soopadoopa.dk/anders/cats/cats.html.

[15] ———, *Gfan, a software system for Gröbner fans and tropical varieties.* http://www.math.tu-berlin.de/~jensen/software/gfan/gfan.html.

[16] D. Maclagan and B. Sturmfels, *Introduction to tropical geometry.* http://www.warwick.ac.uk/staff/D.Maclagan/papers/papers.html.

[17] E. Miller and B. Sturmfels, *Combinatorial commutative algebra.* Graduate Texts in Mathematics, 227, Springer, New York, 2005.
[18] R. Stanley, *Combinatorics and commutative algebra.* Second ed., Progress in Mathematics, 41, Birkhäuser Boston Inc., Boston, 1996.
[19] W. Stein *et. al*, *Sage Mathematics Software (Version 4.7).* The Sage Development Team, 2011, http://www.sagemath.org.
[20] B. Sturmfels, *Gröbner bases and convex polytopes.* University Lecture Series, 8, American Mathematical Society, Providence, RI, 1996.
[21] B. Sturmfels and A. Zelevinsky, *Maximal minors and their leading terms.* Adv. Math. **98**(1993), no. 1, 65–112. http://dx.doi.org/10.1006/aima.1993.1013

Mathematics, University of Washington, Seattle, WA 98195, USA
e-mail: aholtc@uw.edu rrthomas@uw.edu

Mathematics, University of California, Berkeley, CA 94720, USA
e-mail: bernd@math.berkeley.edu