Enumeration and Sparsity in Algebraic Geometry

Dissertation

zur Erlangung des akademischen Grades

doctor rerum naturalium
(Dr. rer. nat.)

von M.Sc. Andreas Kretschmer

geb. am 04.04.1996 in Erlangen

genehmigt durch die Fakultät für Mathematik
der Otto-von-Guericke-Universität Magdeburg

Gutachter: Prof. Dr. Benjamin Nill

Prof. Dr. Thomas Kahle

Prof. Dr. Mateusz Michałek

eingereicht am: 21.02.2024

Verteidigung am: 15.08.2024

Zusammenfassung

Die vorliegende Dissertation behandelt Fragen in der enumerativen algebraischen Geometrie, der kommutativen Algebra, der algebraischen Statistik und der Theorie der Gitterpolytope. Allen gemeinsam sind Verbindungen zur Kombinatorik.

In Kapitel 1 geht es um das Problem der charakteristischen Zahlen für kubische Hyperflächen im projektiven Raum. Wir berechnen unter anderem die Anzahl kubischer Flächen tangential zu 19 Geraden im $\mathbb{P}^3$ und die Anzahl kubischer Dreifaltigkeiten tangential zu 34 Geraden im $\mathbb{P}^4$. Unsere Resultate ermöglichen es prinzipiell, die analogen Fragen in beliebiger Dimension zu beantworten. Dies ist in Teilen gemeinsame Arbeit mit Mara Belotti, Alessandro Danelon und Claudia Fevola.

In Kapitel 2 bestimmen wir die minimale freie Auflösung des Ideals aller $(n-1)$-Minoren einer spärlich besetzten generischen symmetrischen $n \times n$ Matrix. Als Anwendung berechnen wir die erste nicht-triviale charakteristische Zahl für alle Familien spärlich besetzter Quadriken ohne diagonale Nullen. Dies ist gemeinsame Arbeit mit Jiahe Deng.

In Kapitel 3 studieren wir eine gemeinsame Verallgemeinerung von ungerichteten Gaußschen graphischen Modellen und Kovarianzmodellen, bei denen wir Nullen in sowohl der Kovarianzmatrix als auch der Konzentrationsmatrix erlauben. Wir beweisen Strukturresultate für diese Modelle, z.B. Kriterien für Glattheit, Schranken an die Dimension, implizierte Nullen und Blockstrukturen. Dies ist gemeinsame Arbeit mit Tobias Boege, Thomas Kahle und Frank Röttger.

Kapitel 4 handelt von symmetrischen Idealen; dies sind Ideale in einem Polynomring, die invariant unter allen Permutationen der Variablen sind. Wir beweisen, dass Ideale, die von dem Orbit eines generischen homogenen Polynoms erzeugt werden, in einem präzisen Sinn das größtmögliche Radikal und die kleinstmögliche Verschwindungsmenge besitzen.

Kapitel 5 ist ein Beitrag zur lokalen Ehrhart-Theorie. Wir studieren dünne Polytope, also Gitterpolytope, deren lokales h^*-Polynom verschwindet. In Dimension 3 klassifizieren wir dünne Gitterpolytope vollständig und in beliebiger Dimension liefern wir eine Charakterisierung dünner Gorensteinpolytope. Dies ist gemeinsame Arbeit mit Christopher Borger und Benjamin Nill.

Summary

The present thesis deals with questions in enumerative algebraic geometry, commutative algebra, algebraic statistics and the theory of lattice polytopes. All of them share a combinatorial flavor.
Chapter 1 is about characteristic numbers for cubic hypersurfaces in projective space. For instance, we explicitly compute the number of cubic surfaces tangent to 19 lines in $\mathbb{P}^3$ and the number of cubic threefolds tangent to 34 lines in $\mathbb{P}^4$. In principle, our results allow us to answer the analogous question in arbitrary dimensions. This is partly joint work with Mara Belotti, Alessandro Danelon and Claudia Fevola.
Chapter 2 provides the minimal free resolution of the ideal of $(n-1)$-minors of a sparse generic symmetric $n \times n$ matrix. As an application, we compute the first non-trivial characteristic number for all families of sparse quadrics without diagonal zeros. This is joint work with Jiahe Deng.
In Chapter 3 we study a common generalization of undirected Gaussian graphical models and covariance models, allowing for zeros in the covariance and the concentration matrix simultaneously. We prove structural results like smoothness criteria, dimension bounds, implied zeros and block structures. This is joint work with Tobias Boege, Thomas Kahle and Frank Röttger.
Chapter 4 is about symmetric ideals, i.e., ideals in a polynomial ring invariant under all permutations of the variables. We prove that ideals generated by the orbit of a general homogeneous polynomial have, in a precise sense, the largest possible radical and the smallest possible vanishing set.
Chapter 5 is a contribution to local Ehrhart theory. We study thin polytopes, i.e., lattice polytopes whose local h^*-polynomial vanishes. We provide a complete classification in dimension 3 and a characterization of thin Gorenstein polytopes in arbitrary dimension. This is joint work with Christopher Borger and Benjamin Nill.

Authorship

Chapter 1 of this thesis is based on the paper [BDFK23]. While all four authors have contributed significantly to the paper I have taken the leading role, closely followed by Mara Belotti.
In the numbering of the published version [BDFK23], I have contributed the main ideas, results and most of the writing of Section 1. Section 2, except for Subsection 2.6, has been joint work of all four authors but I have taken the leading role, closely followed by Mara Belotti. Subsection 2.6 is entirely my own work and is crucial for the main result Theorem 4.2. Section 3 of the paper has initially been worked out mainly by Mara Belotti with contributions from all other authors. The relevant coding has been done mainly by Mara Belotti and Claudia Fevola. However, I have rewritten and streamlined Section 1.3 which is the thesis version of Section 3 of the paper, and I have written new code that also fits the setting of cubic hypersurfaces of dimensions larger than 2. Section 4 of the paper is joint work of all four authors with essentially equal contributions. I have contributed most of Subsection 4.0, in particular the discussion about cubic surfaces with its consequences for Lemma 4.5. The introduction of the paper has mainly been written by Claudia Fevola with contributions from all other authors.
For Chapter 1, I have rewritten and (sometimes strongly) extended essentially every of the corresponding sections of the paper, to varying degrees. For instance, the introduction of Chapter 1 is a rewritten version of the one in the paper, and Proposition 1.1.1, Lemma 1.1.2 and the proof of Lemma 1.2.1 are new. New are also the entire Subsection 1.2.6, Remark 1.4.2 and Proposition 1.4.3. The main result Theorem 1.4.1 has been significantly extended due to the new Theorem 1.2.24 and the new Proposition 1.3.12.

Chapter 2 is based on the accepted version of the paper [DK23] and sticks to it closely. My coauthor Jiahe Deng did an undergraduate research internship under my supervision for twelve weeks within the RISE Germany 2022 program of the DAAD. The starting point of the paper was my project proposal for this internship. Jiahe Deng was mainly responsible for the coding part and for testing conjectures computationally. I have contributed all the main ideas and have done almost the entire writing. The only exception is Example 2.5.1 which has been contributed by Aldo Conca in personal communication.

Chapter 3 is based on the preprint version of the paper [BKKR23]. In the numbering of Chapter 3, almost all of Section 3.3 and most of Section 3.4 have been contributed by myself. This includes in particular all the main results and their proofs. Exceptions are Remark 3.3.9 which is due to Thomas Kahle and Frank Röttger as well as Remark 3.3.33 which is due to Tobias Boege. Moreover, Tobias Boege and I have contributed equally to Remark 3.3.26, Example 3.3.27, Figure 3.1 and Examples 3.4.1

and 3.4.2. Section 3.2 is almost entirely due to Tobias Boege. Subsection 3.1.1 has been written mostly by Thomas Kahle while Subsection 3.1.2 has been written mainly by myself.
Chapter 3 is, for the most part, very close to the preprint version of the paper (where equal contributions apply). Notable exceptions are Subsection 3.3.1 which I have streamlined and updated as well as the classification in Proposition 3.3.31 which I have corrected. In both the preprint and the published version this classification was incomplete because the cases where $E_G \cap E_H$ forms a star were missing.

Chapter 4 is based on the paper [Kre23] of which I am the only author. Chapter 4 is close to the accepted version up to minor improvements.

Chapter 5 is based on the accepted version of the paper [BKN23] and sticks to it closely. In this case, the main ideas and the biggest part of the writing are due to Benjamin Nill. The notes and insights of Jan Schepers during the collaboration with Benjamin Nill on [NS13] contained in particular Lemma 5.5.5 and were the basis of the proof of Theorem 5.5.3. Lev Borisov contributed the proof of Proposition 5.5.9. My overall contribution is comparatively larger than that of Christopher Borger. The first main result Theorem 5.3.3 is a classification result which is joint work of Christopher Borger and myself with essentially equal contributions; the coding part of the proof, however, has been done almost entirely by Christopher Borger. Benjamin Nill and I have written Section 5.1 with my contribution being comparatively smaller. I have contributed Lemma 5.4.13 with its proof as well as the proof, but not the statement, of Lemma 5.4.18 which is important in the proof of the second main result Theorem 5.5.3. In addition, I have done all the coding apart from that used in the proof of the mentioned classification result Theorem 5.3.3, see also [Kre]. In particular, I contributed Examples 5.2.17, 5.5.10 and 5.5.11 which I have obtained computationally.

Contents

0 Introduction

This thesis comprises several topics in enumerative and combinatorial algebraic geometry. Central themes are *excess intersection* and *syzygies* in Chapters 1 and 2, the structural restrictions imposed by *algebraic group actions* in Chapters 1 and 4, and the presence of *sparsity* in Chapters 2–5.

Chapters 1 and 2 contribute to the characteristic numbers problem in classical enumerative geometry, in the case of cubic hypersurfaces and in the case of sparse quadrics. Contact problems in algebraic geometry have a long history; they date back, if not to antiquity, then at least to the 19^{th} century with early work of Steiner, Chasles, Maillard, Zeuthen and Schubert [Ste48, Cha64, Mai71, Zeu73, Sch79]. Given some family F of subvarieties of $\mathbb{P}^n$, one wishes to count the number of members of F having a contact with each of the given subvarieties $Y_1, \dots, Y_r \subseteq \mathbb{P}^n$. By definition, two subvarieties X and Y of $\mathbb{P}^n$ *make a contact* or *touch* if their conormal varieties in $\mathbb{P}^n \times \check{\mathbb{P}}^n$ intersect [FKM83]. If $\operatorname{codim}(X) + \operatorname{codim}(Y) > n$, then X touches Y if and only if $X \cap Y \neq \emptyset$ which in this case is a non-trivial condition. If on the other hand $\operatorname{codim}(X) + \operatorname{codim}(Y) \leq n$, the notion of a contact formalizes the intuitive meaning of X and Y being *tangent*. It follows from [FKM83] that every contact problem can essentially be reduced to the case where each of the given varieties Y_i is replaced by a linear space of the same or smaller dimension. The resulting numbers in this case are called the *characteristic numbers* of F.

A classical example is Steiner's problem of five conics [Ste48]: How many smooth conics in $\mathbb{P}^2$ are tangent to five given conics in general position? Famously, the correct answer is 3264 [EH16], at least in characteristic $\neq 2$. More generally, we may let F be the family of smooth quadrics in $\mathbb{P}^n$ and ask how many of them are tangent to $r = \dim(F) = \binom{n+2}{2} - 1$ given varieties of various dimensions. An essentially complete solution to these problems is given by the *space of complete quadrics* and its intersection theory [Vai82].

When it comes to higher degree hypersurfaces, much less is known, see Table 1. The only cases of smooth degree $d > 2$ hypersurfaces for which all characteristic numbers are known are plane cubic and plane quartic curves [Alu90, Vak99]. For higher degree plane curves, partial results have been achieved, for instance, by Aluffi [Alu91b, Alu92], but to my knowledge no complete list is known even in degree 5. More generally, in fact, apart from quadrics almost all families for which any non-trivial characteristic numbers are known are families of *curves*.

Our paper [BDFK23] seems to be the first attempt at leaving the realms of curves and quadrics after Coskun [Cos06a, Cos06b] who has computed some of the characteristic numbers for rational normal surface scrolls and del Pezzo surfaces, however allowing only incidence conditions but not tangency conditions. In [BDFK23] we have counted the number of smooth cubic surfaces in $\mathbb{P}^3$ passing through a general points and

(n,d)	**Some references**
$(n,2)$	Chasles (1864), Schubert (1879), Vainsencher (1982), ...
$(2,3)$	Maillard (1871), Zeuthen (1872), Sterz (1986), Kleiman–Speiser (1986, 1988, 1991), Aluffi (1990), Vakil (1999)
$(2,4)$	Zeuthen (1873), van Gastel (1991), Aluffi (1991, 1992), Vakil (1999)
$(2,d)$	Partial results, e.g. Aluffi (1991, 1992)
$(3,3)$	Partial results in our [BDFK23]
$(n,3)$	Partial results in Chapter 1

Table 1: Characteristic numbers for degree d hypersurfaces in $\mathbb{P}^n$

tangent to $b = 19 - a$ general lines, building on the PhD work of Aluffi [Alu90]. In Chapter 1 of the present thesis we will extend this result to arbitrary dimensions, providing methods to compute, in principle, the characteristic numbers with respect to points and lines for the family of smooth cubic hypersurfaces of *any* dimension.

Even in the case of quadrics, there are variations of the characteristic numbers problem which are widely open. For instance, the characteristic numbers for families of *sparse* smooth quadrics are unknown. As an application of the main result of Chapter 2, Theorem 2.1.1, we compute a first non-trivial characteristic number for all sparse families without diagonal zeros, see Section 2.4. This boils down to a degree computation for the ideal $I_{n-1}(A)$ of all $(n-1)$-minors of a sparse generic symmetric $n \times n$ matrix A, i.e., A is a symmetric matrix whose upper triangle is filled with distinct variables or zeros, and the zeros are allowed at off-diagonal places only.
However, Theorem 2.1.1 goes beyond a degree computation. We provide the graded free resolution of $I_{n-1}(A)$, employing the *pruning method* first applied to the ideal of maximal minors of a (rectangular) sparse generic matrix by Boocher in his thesis [Boo12, Boo13]. The word pruning refers to the erasing of certain rows and columns of the matrices defining the resolution in the generic case (without zeros). Contrary to Boocher's situation, the standard $(n-1)$-minors of a generic *symmetric* matrix do not form a universal Gröbner basis. Nonetheless, the pruning procedure still applies here. In particular, the projective dimension and even the individual graded Betti numbers of $I_{n-1}(A)$ can only be less than in the generic case. This statement fails quite drastically already in simple cases if zeros are allowed also on the diagonal, see Example 2.1.2.

My initial motivation to study ideals of certain submaximal minors of A arose in the context of undirected Gaussian graphical models in algebraic statistics. In Chapter 3 we study a new class of statistical models within the realm of Gaussian normal distributions which we termed *double Markovian*, and Definition 3.3.17 introduces the corresponding conditional independence ideals. Double Markovian models are a simultaneous generalization of graphical models and so-called covariance models.
The general aim of algebraic statistics is to study geometric, algebraic and combinatorial properties of statistical models that are given as real (semi-)algebraic sets, see for example [DSS09, Sul18]. For instance, a multivariate Gaussian normal

distribution is determined by its mean vector μ and its *covariance matrix* Σ which is symmetric and positive semi-definite. Therefore, the parameter space of all Gaussian normal distributions is naturally $\mathbb{R}^n \times \mathrm{PSD}_n$, where PSD_n is the cone of symmetric positive semi-definite $n \times n$ matrices. Given an undirected graph G with n vertices, the (undirected) graphical model $\mathcal{M}(G)$ associated to G is the set of all normal distributions such that $(\Sigma^{-1})_{ij} = 0$ whenever $i \neq j$ and ij is not an edge of G [Lau96, Chapter 5]. Given another graph H on the same vertex set, a double Markovian model $\mathcal{M}(G, H)$ also prescribes zeros in Σ as determined by H. From the point of view of real algebraic geometry, double Markovian models are subvarieties of the cone of positive definite matrices defined by coordinate hyperplanes and the vanishing of some submaximal minors of the covariance Σ. In Chapter 3 we prove structural results about these models such as implied zeros and block structures, smoothness criteria and dimension bounds. These are in part algebraic and in part crucially rely on the positivity constraint. An example of the latter is Theorem 3.3.23 which strongly constrains the possible combinations of zero patterns in a positive definite matrix and its inverse. This might be of independent interest. Another example where positivity is essential is Theorem 3.3.8 stating that $\mathcal{M}(G, H)$ is *smooth* if every pair of vertices is joined by an edge in at least one of G and H.

Chapter 4 studies solution sets of *symmetric ideals*, i.e., ideals in a polynomial ring $K[x_1, \ldots, x_n]$ which are invariant under all permutations of the variables. Special classes of symmetric ideals, for instance Specht ideals and Tanisaki ideals, have been studied intensively in the algebraic combinatorics literature and are related to the famous work of Haiman on $n!$, see for example [Hai03, MOY22] and the references therein. General symmetric ideals and chains thereof in increasingly larger polynomial rings are an active topic of research in asymptotic commutative algebra [NR17, LNNR20, LNNR21] and have applications to chemistry and algebraic statistics [AH07, HS12]. One of the most famous results in this area says that an S_∞-invariant ideal of the infinite polynomial ring is finitely generated up to symmetry [Coh67, AH07, HS12, Dra14].
The main result of Chapter 4 concerns the class of symmetric ideals $(S_n \cdot f)$ generated by all permutations of a single polynomial $f \in K[x_1, \ldots, x_n]$ with a prescribed sparsity structure, i.e., only certain monomials are allowed to appear in f. Here, S_n denotes the symmetric group on n elements. Surprisingly, experiments show that the radical $\sqrt{(S_n \cdot f)}$ is very often a *monomial ideal*, necessarily symmetric and square-free. We prove that this phenomenon in fact occurs for general coefficients of f if its support is homogeneous and S_n-stable.

The final Chapter 5 is a contribution to *local Ehrhart theory*. Given a lattice polytope $P \subseteq \mathbb{R}^n$ of dimension d, the number of lattice points in integer dilates kP is a polynomial function of k of degree d by a famous theorem of Ehrhart [Ehr62]. This polynomial is called the *Ehrhart polynomial* of P and agrees with the Hilbert function of the semigroup algebra $\mathbb{C}[P]$ in non-negative degrees. The numerator polynomial of the Hilbert series of $\mathbb{C}[P]$ is called the *h^*-polynomial* of P. Its coefficients are always non-negative integers. This can be deduced from the Cohen–Macaulayness of $\mathbb{C}[P]$ [Hoc72] but today several independent combinatorial proofs are known [BS18]. A subtler and far less studied invariant of a lattice polytope P is its *local h^*-polynomial*, denoted $\ell^*_P(t) \in \mathbb{Z}[t]$. This polynomial has first been introduced by Stanley [Sta92]. It, also, has non-negative coefficients and moreover is always palindromic. From an algebro-geometric perspective, $\ell^*_P(t)$ encodes the numerical information about the mixed Hodge structure on the primitive part of the middle cohomology group of a non-degenerate hypersurface Z of $(\mathbb{C}^*)^d$ defined by a Laurent polynomial with Newton polytope P, see [Bat93, BC94, BB96a] and especially [BM03, Section 5]. In formulas,

$$\ell^*_P(t) = t\sum_{i=0}^{d-1} \dim(PH_c^{i,d-1-i}(Z))t^i.$$

Contrary to the usual h^*-polynomial, for the local h^*-polynomial no combinatorial interpretation of its coefficients is known beyond the simplex case. Moreover, $\ell^*_P(t)$ may vanish identically. In Chapter 5 we study precisely this phenomenon. We say that a lattice polytope P is *thin* if $\ell^*_P(t) = 0$. Our main results include a characterization of thinness in dimension 3 and for Gorenstein polytopes of arbitrary dimension. We have also included an introduction to the local h^*-polynomial with a survey of previous results.

1 Characteristic Numbers for Cubic Hypersurfaces

For questions of authorship, please refer to pages IVf.

This chapter is a rewritten and extended version of the paper [BDFK23]. I have attempted to improve the notation and readability at many places. Several proofs have been streamlined, some shortened, and some (minor) errors have been caught and fixed. Most importantly, the main result has been significantly extended: While the results of [BDFK23] allowed us to provide the characteristic numbers with respect to point and line conditions for smooth cubic *surfaces*, this chapter includes the new Theorem 1.2.24 which makes it possible to compute, in principle, all characteristic numbers with respect to points and lines for smooth cubic hypersurfaces of *any* dimension. The numbers are explicitly derived in the cases of cubic surfaces, threefolds, fourfolds and fivefolds in Theorem 1.4.1.

Contact Problems for Hypersurfaces

A famous moduli space in enumerative geometry is the *space of complete quadrics* [Cha64, Sch79, Vai82]. This is a compactification of the set of smooth quadric hypersurfaces in $\mathbb{P}^n$. One construction starts with the projective space $\mathbb{P}^{\binom{n+2}{2}-1}$ parametrizing all quadrics and iteratively blows up the proper transforms of the loci of quadrics with rank at most $1, 2, \ldots, n-1$. This variety has been used to answer the degree 2 case of the question:

How many smooth degree d hypersurfaces in $\mathbb{P}^n$ are tangent to $\binom{n+d}{d} - 1$ general linear spaces of various dimensions?

Solutions to these kinds of problems are classically called *characteristic numbers* (for the family of, in this case, smooth degree d hypersurfaces). In the case of quadric surfaces, the first complete solution was achieved by Schubert [Sch79] back in 1879 after Chasles had treated the case of plane conics [Cha64] in 1864. Later, these questions have been translated into Chow ring computations on the space of complete quadrics, and beautiful results have been achieved [Sem48, Vai82, DCP85]. For families of quadrics, in effect, an explicit construction of a space is known where the characteristic numbers problem translates, essentially, into a cohomological computation. More recently, the space of complete quadrics has even proved useful in studying some classical problems in algebraic statistics related to maximum likelihood estimation [MMM+23, MMW21].

Much less is understood when it comes to higher degree hypersurfaces. To our knowledge, the only cases where all characteristic numbers are known are plane cubics and plane quartics. The characteristic numbers for cubics were first computed by Maillard [Mai71] in 1871. In 1873, Zeuthen computed the characteristic numbers for plane curves up to degree 4 [Zeu73]. Of course, their methods were relying on assumptions that were not rigorously justified from a modern perspective. It took more than a century to re-prove these numbers using the rigorous theoretical foundations provided by Fulton–MacPherson intersection theory, as for instance in the works of Kleiman and Speiser [KS86, KS88, KS91], and Aluffi [Alu90, Alu91a] in the case of plane cubics. For plane quartics, partial results were achieved in [Alu91b, vG91], and later a full description was given in [Vak99].
A particularly interesting feature of [Alu90] is that the author constructs a *variety of complete plane cubics* whose intersection theory allows to compute the characteristic numbers for, in principle, any family of reduced plane cubics. In a similar fashion as for complete quadrics, the space of complete plane cubics is constructed through a sequence of blow-ups of the projective space parametrizing all cubic forms.
As far as we know, the case of higher-dimensional cubic hypersurfaces has been unexplored. Our aim in this chapter is to generalize the space of complete plane cubics in *loc. cit.* to arbitrary dimensions. The result is a smooth projective variety which we term a *variety of* 1*–complete cubic hypersurfaces*. This is, in essence, a variety whose intersection theory answers the following question:

What is the number of smooth cubic hypersurfaces in $\mathbb{P}^n$ *passing through* n_p *general points and tangent to* $\binom{n+3}{3} - n_p - 1$ *general lines?*

In principle, we are able to answer this question completely, although obtaining the numbers explicitly for high n is computationally demanding. It should be possible to turn these results into closed (but possibly very long) formulas for the characteristic numbers for arbitrary n but I have not pursued this much. An exception in which the formula is nice and the computation remains doable is provided by Proposition 1.4.3.

We give a brief outline of the construction. Let W be a vector space of dimension $n+1$ with $n \geq 2$ over an algebraically closed field K of $\mathrm{char}(K) \neq 2, 3$. The set of all cubic hypersurfaces in $\mathbb{P}(W)$ is naturally parametrized by the projective space $\mathbb{P}(\mathrm{Sym}^3(W^*))$ of dimension $\binom{n+3}{3} - 1$. The classical theory of discriminants of univariate polynomials shows that the subset of $\mathbb{P}(\mathrm{Sym}^3(W^*))$ of cubics tangent to a given line in $\mathbb{P}(W)$ is a degree 4 hypersurface which we call a *line condition*. Similarly, a *point condition* is the hyperplane in $\mathbb{P}(\mathrm{Sym}^3(W^*))$ of all cubics containing a given point in $\mathbb{P}(W)$. We want to count the finite number of *smooth* cubics in the intersection of n_p general point conditions and $n_\ell = \binom{n+3}{3} - 1 - n_p$ general line conditions. However, the intersection of such hypersurfaces in $\mathbb{P}(\mathrm{Sym}^3(W^*))$ is not always zero-dimensional. Instead, the intersection of all line conditions set-theoretically agrees with the locus of non-reduced cubics, i.e., those of the form $\lambda\mu^2$ for linear forms λ and μ. This makes an excess intersection approach necessary. Indeed, the goal of the construction is to obtain a smooth variety birational to $\mathbb{P}(\mathrm{Sym}^3(W^*))$ in such a way that in the new space the proper transforms of the line conditions separate, i.e., the intersection of all these proper transforms is empty. It

turns out that, as in Aluffi's case, it is enough to blow up five times along smooth centers, each of them (apart from the first one) being an irreducible component of the intersection of all proper transforms of the line conditions. The ultimate goal of computing the characteristic numbers is finally achieved by subtracting from the bound provided by Bézout's theorem a certain correction term which can be expressed via Chern classes of the normal bundles arising in the blow-up process.
Let us point out one crucial difference to [Alu90], namely the center of the last blow-up, later denoted B_4. In Aluffi's case $n = 2$, this is a projectivized line bundle on the previous center B_3 while it turns out that, in general, B_4 is the projectivization of a vector bundle $\mathcal{E}$ on B_3 of rank $\binom{n}{2}$. The main difficulty lies in the fact that, a priori, $\mathcal{E}$ is not known explicitly. In particular, there is no obvious method to obtain the Chern classes $c_i(\mathcal{E}) \in \mathrm{CH}^i(B_3)$ which in turn are necessary to compute the non-trivial characteristic numbers. This issue is resolved here in Subsection 1.2.6 by identifying $\mathcal{E}$ as the pullback to B_3 of a normal bundle appearing in the second step of the construction of the space of complete quadrics. This relationship seems interesting in itself. The identification of $\mathcal{E}$ and the computation of its Chern classes have no analog in [Alu90]. In particular, Theorem 1.2.24 and Proposition 1.3.12 provide a new proof of the important Lemma 4.2 of *loc. cit.*
In our presentation we have aimed at providing many details throughout the construction and we have separated the geometric construction from the Chow ring computations as the former is logically independent of the latter.
This chapter is organized as follows. In Section 1.1 we give the definition of a 1-complete variety of cubic hypersurfaces $\tilde{V}$ and Theorem 1.1.6 shows how certain intersection numbers on $\tilde{V}$ yield the characteristic numbers we were aiming for.
Section 1.2 concerns the construction of the 1-complete variety $\tilde{V}$ achieved by performing five blow-ups. In each subsection we spell out the details of each blow-up by giving local coordinates, the support of the intersection of the proper transforms of the line conditions, and local equations for this intersection. An irreducible component of this intersection is then taken to be the center of the next blow-up. The construction ends with Corollary 1.2.23 where we show that the proper transforms of the line conditions do no longer intersect.
Section 1.3 is devoted to the Chow rings of the five centers defined in the previous section and to the computation of the intersection classes needed for the correction term.
In the final Section 1.4, we gather the data computed so far and provide the characteristic numbers for cubic surfaces, threefolds, fourfolds and fivefolds.
The code used in [BDFK23] for the case $n = 3$ of cubic surfaces is available at

`https://mathrepo.mis.mpg.de/CountingCubicHypersurfaces`.

For the more general results presented here, the `Macaulay2` code implementing the necessary Chow ring and Chern class computations for arbitrary n is available from the author upon request.

1.1 First Associated Hypersurfaces and the Hurwitz Map

We fix an integer $d \geq 2$, an algebraically closed field K of characteristic 0 or $> d$, and a K-vector space W of dimension $n+1$ with $n \geq 2$. We refer to [GKZ94, Section 3.2.E] for the notion of higher associated hypersurfaces of a projective variety. Specifically, we are interested in the following case: Let $X \coloneqq \mathcal{V}(f) \subseteq \mathbb{P}(W)$ be an integral hypersurface of degree d, defined by an irreducible homogeneous polynomial f. If X is smooth, its *first associated hypersurface* $\mathcal{Z}_1(X) \subseteq \mathrm{Gr}(2, W)$ consists of all lines $\ell \subseteq \mathbb{P}(W)$ such that ℓ is tangent to X at some point or, more precisely, $\dim(\ell \cap T_x X) = 1$ for some point $x \in \ell \cap X$, where $T_x X$ is the (embedded) tangent space of X at the point x. If instead X is singular, we first consider the lines ℓ for which there exists a smooth point x satisfying the above conditions and then take the Zariski closure of this set in the Grassmannian $\mathrm{Gr}(2, W)$.

Some caution is in order because higher associated hypersurfaces are actually *not* hypersurfaces in their respective ambient Grassmannians in general. The reason for this slightly unfortunate terminology is the erroneous [GKZ94, Proposition 2.11] which would imply that all dual varieties to integral hypersurfaces (over $\mathbb{C}$) are again hypersurfaces which is known to be false. However, the case of lines and hypersurfaces is special, and in this case the mentioned proposition is true, see [Stu17, Theorem 1.1] for a more general result.

We briefly give the argument in the case relevant to us here. Let $U \coloneqq X_{\mathrm{sm}}$ denote the smooth locus of X. Consider the incidence subscheme (with the reduced scheme structure) given by

$$Z = \{(\ell, p) : p \in \ell \subseteq T_p U\} \subseteq \mathrm{Gr}(2, W) \times U,$$

with projections $p : Z \to \mathrm{Gr}(2, W)$ and $q : Z \to U$. Via q, the incidence subscheme Z becomes a $\mathbb{P}^{n-2}$-bundle over U since every local trivialization of the tangent bundle TU provides a local trivialization of the fiber bundle given by q. In particular, Z is irreducible of dimension $\dim(U) + n - 2 = 2n - 3$. This shows that $\mathcal{Z}_1(X) = \overline{p(Z)}$ is irreducible of dimension

$$\dim(\mathcal{Z}_1(X)) = \dim(Z) - \dim(p^{-1}(\ell)) = 2n - 3 - \dim(p^{-1}(\ell))$$

for a general $\ell \in p(Z)$. Hence, $\dim(\mathcal{Z}_1(X)) = \dim(\mathrm{Gr}(2, W)) - 1$ if the general fiber $p^{-1}(\ell)$ is finite. But this must be the case since otherwise the general tangent line to U would have to be contained in X. Since the condition for a line in $\mathrm{Gr}(2, W)$ to be contained in X is closed, *every* tangent line to U would then be contained in X. In particular, for any $p \in U$, every line through p in $T_p U$ is tangent to U and hence contained in X, so $T_p U \subseteq X$, whence $T_p U = X$. This contradicts our assumptions on X. So indeed, $\mathcal{Z}_1(X)$ is an irreducible hypersurface in $\mathrm{Gr}(2, W)$.

If X is smooth, by [Stu17, Theorem 1.1] the hypersurface $\mathcal{Z}_1(X) \subseteq \mathrm{Gr}(2, W)$ is defined by an irreducible element Hu_f of degree $d(d-1)$ in the projective coordinate ring of $\mathrm{Gr}(2, W)$, called the *Hurwitz form*. The latter is a degree $d(d-1)$ homogeneous polynomial in the Plücker coordinates, unique up to the degree $d(d-1)$ piece of the ideal generated by the Plücker relations.

On the open set of $\mathbb{P}(\mathrm{Sym}^d(W^*))$ parametrizing smooth degree d hypersurfaces, we can therefore define a morphism sending X to the degree $d(d-1)$ hypersurface $\mathcal{Z}_1(X)$ of $\mathrm{Gr}(2,W)$. The set of hypersurfaces in $\mathrm{Gr}(2,W)$ of this degree is parametrized by the projective space $|\mathcal{O}_{\mathrm{Gr}(2,W)}(d(d-1))|$. We call this morphism the *Hurwitz map* and view it as a rational map

$$\mathrm{Hu}: \ \mathbb{P}(\mathrm{Sym}^d(W^*)) \dashrightarrow |\mathcal{O}_{\mathrm{Gr}(2,W)}(d(d-1))|, \quad [f] \longmapsto [\mathrm{Hu}_f].$$

For instance, if $n=2$, this map is simply the one taking a degree d plane curve into its dual curve.

Proposition 1.1.1. If $\mathrm{char}(K)=0$ and $d\geq 2$, the Hurwitz map is injective on the locus of smooth degree d hypersurfaces.

Proof. We need to show that we can recover a smooth degree d hypersurface $X\subseteq \mathbb{P}(W)$ from $\mathcal{Z}_1(X)\subseteq \mathrm{Gr}(2,W)$. If $n=2$, i.e., X is a curve, this follows from the biduality theorem, so we may assume $n\geq 3$. As $\mathrm{char}(K)=0$, by the biduality theorem it suffices to show that we can recover the dual variety X^* from $\mathcal{Z}_1(X)$ alone.
For this, given a hyperplane $H\in\mathbb{P}(W^*)$ and a point $p\in H\subseteq \mathbb{P}(W)$, we let

$$G_{p,H} \coloneqq \{\ell\in \mathrm{Gr}(2,W) : p\in \ell\subseteq H\}.$$

Given H, consider the set

$$A_H \coloneqq \{p\in H | G_{p,H}\subseteq \mathcal{Z}_1(X)\}.$$

We claim $A_H\neq\emptyset$ if and only if $H\in X^*$. Since A_H only depends on $\mathcal{Z}_1(X)$ but not on X, this will conclude the proof.
To prove the claim, first observe that $H\in X^*$ is equivalent to H being tangent to X. In fact, the singular points of the scheme-theoretic intersection $X\cap H$ are precisely the points of tangency of H to X. Therefore, if H is tangent to X, every point of tangency will certainly lie in A_H, so $A_H\neq\emptyset$.
Conversely, if $H\notin X^*$, then the scheme-theoretic intersection $X\cap H$ is smooth (and connected, hence irreducible). The claim then follows from Lemma 1.1.2 below with $Y=X\cap H$ and H playing the role of $\mathbb{P}^m$. □

Lemma 1.1.2. Let $\mathrm{char}(K)=0$, $m\geq 2$ and $Y\subseteq\mathbb{P}^m$ a smooth hypersurface. Then, for any point $p\in\mathbb{P}^m$, the general line through p is nowhere tangent to Y.

Proof. Consider the orthogonal projection away from p:

$$\pi: \mathbb{P}^m \dashrightarrow \mathbb{P}^{m-1}.$$

The closure of the fiber over a point q is the line through p and any lift of q. If $p\notin Y$, then π restricts to a morphism $Y\to\mathbb{P}^{m-1}$, and the scheme-theoretic fiber over every closed point is the scheme-theoretic intersection of the corresponding line through p with Y. Since $\dim(Y)=m-1$, the general fiber is finite, and by generic

smoothness in characteristic zero, the general fiber is finite and smooth, i.e., the general line through p is nowhere tangent to Y.
If $p \in Y$, consider instead the morphism $\mathrm{Bl}_p Y \to \mathbb{P}^{m-1}$. Since the exceptional divisor has dimension $m-2$, the general fiber lies in the complement of the exceptional divisor. As before, for dimension reasons and by generic smoothness, the general fiber must be finite and smooth, so the general line through p is not tangent to Y at any intersection point different from p itself. Since the general line through p is clearly not tangent to Y *at p itself*, this concludes the proof. □

Following [Stu17, Example 2.2], Hu_f can be computed as the discriminant of the homogeneous bivariate degree d polynomial in the variables s and t given by $f(sv_0 + tw_0, \dots, sv_n + tw_n)$. As a polynomial in s and t, the coefficients of the latter are bihomogeneous of degree $(1, d)$ in the coefficients of f and the variables v_i, w_i, respectively. It follows that the polynomial Hu_f is bihomogeneous of degree $(2(d-1), 2d(d-1))$ with respect to the aforementioned variables. By SL_2-invariance, Hu_f can even be expressed as a polynomial in the Plücker coordinates $p_{0,1}, p_{0,2}, \dots, p_{n-1,n}$ of the Grassmannian $\mathrm{Gr}(2, W)$ given by the 2×2 minors $p_{ij} = v_i w_j - v_j w_i$. Hence, Hu_f is bihomogeneous of degree $(2(d-1), d(d-1))$ in the coefficients of f and the Plücker coordinates, respectively. Notice that the polynomial Hu_f obtained in this way makes sense for arbitrary degree d hypersurfaces $\mathcal{V}(f)$. It has the property that, for any fixed f, we have $\mathrm{Hu}_f(\ell) = 0$ if and only if $\mathcal{V}(f)$ is tangent to ℓ in the sense that $\ell \subseteq T_p\mathcal{V}(f)$ for some $p \in \ell \cap \mathcal{V}(f)$.
In the same line of [Alu90], we define the *point condition* P^p and the *line condition* L^ℓ as the hypersurfaces in $\mathbb{P}(\mathrm{Sym}^d(W^*))$ consisting of the degree d hypersurfaces, respectively, containing the point p and tangent to the line ℓ.

Lemma 1.1.3. The indeterminacy locus of the Hurwitz map Hu is precisely the scheme-theoretic intersection of all line conditions, which in turn set-theoretically agrees with the subset $S_0 \subseteq \mathbb{P}(\mathrm{Sym}^d(W^*))$ consisting of all non-reduced hypersurfaces, i.e., those defined by a polynomial which is divisible by the square of some non-constant polynomial.

Proof. Fixing a line $\ell \in \mathrm{Gr}(2, W)$, the polynomial $\mathrm{Hu}_f(\ell)$ is a homogeneous degree $2(d-1)$ polynomial in the coefficients of f. Its vanishing set agrees with $L^\ell \subseteq \mathbb{P}(\mathrm{Sym}^d(W^*))$, hence for the first claim it is enough to see that $\mathrm{Hu}_f(\ell)$, for fixed ℓ, is irreducible as a polynomial in the coefficients of f. This is clearly a property invariant under the action of PGL_n, so we can consider the line $\ell = \langle e_0, e_1 \rangle$. Then $\mathrm{Hu}_f(\ell)$ is precisely the discriminant of the generic homogeneous degree d polynomial in two variables x_0, x_1, and this is indeed known to be an irreducible polynomial of degree $2(d-1)$ if $\mathrm{char}(K) \neq 2$.
Set-theoretically, the indeterminacy locus of Hu consists of all $[f]$ such that $\mathrm{Hu}_f(\ell) = 0$ for all lines ℓ. In this case, the singular locus of the closed subscheme $\mathcal{V}(f) \subseteq \mathbb{P}(W)$ must have codimension 0, otherwise the general line would avoid the singular locus of $\mathcal{V}(f)$ and would intersect $\mathcal{V}(f)$ transversally in d distinct smooth points by Bertini's theorem. But the singular locus of $\mathcal{V}(f)$ can only have codimension 0 if $\mathcal{V}(f)$ has an everywhere non-reduced component, so f is divisible by the square of some non-constant polynomial. □

Remark 1.1.4. The rational map induced by the linear system generated by all line conditions in $H^0(\mathbb{P}(\mathrm{Sym}^d(W^*)), \mathcal{O}(2(d-1)))$ is closely related to Hu. Composing the former with a suitable linear embedding into $|\mathcal{O}_{\mathrm{Gr}(2,W)}(d(d-1))|$ gives the latter.

This allows us to present the following definition.

Definition 1.1.5. A 1-*complete variety of degree d hypersurfaces* is a morphism $\pi\colon \tilde{V} \to \mathbb{P}(\mathrm{Sym}^d(W^*))$ from a smooth projective variety $\tilde{V}$ which is an isomorphism outside $\pi^{-1}(S_0)$ resolving Hu, i.e., such that the intersection of the proper transforms of all line conditions L^ℓ in $\tilde{V}$ is empty:

$$\begin{tikzcd} \tilde{V} \arrow[d, "\pi"'] \arrow[drr, "\widetilde{\mathrm{Hu}}"] & & \\ \mathbb{P}(\mathrm{Sym}^d(W^*)) \arrow[rr, dashed, "\mathrm{Hu}"'] & & |\mathcal{O}_{\mathrm{Gr}(2,W)}(d(d-1))|. \end{tikzcd}$$

An analogous construction for tangency with respect to s-dimensional planes instead of lines would lead to the definition of *s-complete varieties of degree d hypersurfaces*. In general, however, the intersection of all s-plane conditions set-theoretically agrees with the subset of $\mathbb{P}(\mathrm{Sym}^d(W^*))$ given by all degree d hypersurfaces with singular locus of dimension $\geq n-s$, by a Bertini type argument. For $s \geq 2$, this set seems hard to describe explicitly. For example, the case $s = 2$ yields the set of *non-normal* degree d hypersurfaces which always contains the reducible and non-reduced hypersurfaces as well as the hypersurfaces singular along a *linear* space of dimension $n-2$. For cubics, this is the full description by [LPS11, Lemma 2.4]. In general, however, there can be more than these classes.

Theorem 1.1.6. We write $V_0 \coloneqq \mathbb{P}(\mathrm{Sym}^d(W^*))$. Let $\tilde{V}$ be a 1-complete variety of degree d hypersurfaces as above and let $F \subseteq V_0 \setminus S_0$ be a locally closed subvariety. Denote by $\tilde{F} \subseteq \tilde{V}$ the proper transform of the closure $\overline{F}$ and by $\tilde{L}^\ell, \tilde{P}^p \subseteq \tilde{V}$ the line and point conditions of $\tilde{V}$, i.e., the proper transforms in $\tilde{V}$ of the irreducible hypersurfaces $L^\ell, P^p \subseteq V_0$ corresponding to line and point conditions of V_0, respectively, for the line $\ell \subseteq \mathbb{P}(W)$ and the point $p \in \mathbb{P}(W)$.

(1) For any finite set of subvarieties $A_1, \ldots, A_r \subseteq \tilde{V}$, there exist a point p and a line ℓ such that $\tilde{P}^p$ and $\tilde{L}^\ell$ both intersect every A_i properly, i.e., in the expected dimension. In fact, this is the case for a general point and a general line.

(2) If $r = \dim(F)$, there exist r lines $\ell_1, \ldots, \ell_r$ such that the corresponding line conditions in $\tilde{V}$ intersect $\tilde{F}$ in finitely many points, mapping to F under π. Again, this is the case for general lines $\ell_1, \ldots, \ell_r$.

(3) The number of elements of F, counted with multiplicity, passing through n_p general points and tangent to n_ℓ general lines such that $n_p + n_\ell = \dim(F)$ equals the degree of the 0-cycle $\tilde{P}^{n_p} \cdot \tilde{L}^{n_\ell} \cdot \tilde{F} \in \mathrm{CH}_0(\tilde{V})$, where $\tilde{P}, \tilde{L}$ denote the cycle classes of any point and line condition $\tilde{P}^p, \tilde{L}^\ell$ in $\tilde{V}$.

(4) Assume $\mathrm{char}(K) = 0$. Then for sufficiently general points and lines, the multiplicity of every element of F from (3) is 1.

Proof. Both (2) and (3) follow from (1), the arguments being the same as in [Alu90, Proposition 1]. For (1), we also mimic the strategy of *loc. cit.* and assume the conclusion is false, i.e., there are subvarieties $A_1, \dots, A_r$ such that for every line ℓ, the line condition $\tilde{L}^\ell$ intersects non-properly at least one of the A_i. As $\tilde{L}^\ell$ is an irreducible hypersurface and all A_i are irreducible, this means that every line condition in $\tilde{V}$ contains at least one of the A_i. Pick a point $p_i \in \widetilde{\mathrm{Hu}}(A_i)$ for each $i = 1, \dots, r$ and denote by $G^\ell \subseteq |\mathcal{O}_{\mathrm{Gr}(2,W)}(d(d-1))|$ the hyperplane given by all degree $d(d-1)$ elements of the projective coordinate ring of $\mathrm{Gr}(2, W)$ vanishing at ℓ. The coefficients of the linear equation defining G^ℓ are the monomials of degree $d(d-1)$ in the Plücker coordinates evaluated at the line ℓ. Now, by construction, $\mathrm{Hu}(L^\ell \setminus S_0) \subseteq G^\ell$ and therefore $\widetilde{\mathrm{Hu}}(\tilde{L}^\ell) \subseteq G^\ell$. In particular, every hyperplane G^ℓ contains at least one of the finitely many points p_i. Dually, in $\check{\mathbb{P}}(H^0(\mathcal{O}_{\mathrm{Gr}(2,W)}(d(d-1))))$ this means that all points corresponding to the hyperplanes G^ℓ are contained in the finite union of hyperplanes corresponding to the points p_i. However, the set of points corresponding to the G^ℓ is the image of the $d(d-1)$-Veronese embedding $\mathrm{Gr}(2, W) \to \check{\mathbb{P}}(H^0(\mathcal{O}_{\mathrm{Gr}(2,W)}(d(d-1))))$ and thus irreducible. Hence, this image would have to be contained in a single hyperplane. In other words, switching back to the primal setting, there exists some p_i, say p_1, that is contained in all hyperplanes G^ℓ. Then p_1 corresponds to a non-zero element in the degree $d(d-1)$ part of the projective coordinate ring of $\mathrm{Gr}(2, W)$ that, as a polynomial in the Plücker coordinates, must vanish at all lines in $\mathbb{P}(W)$, hence on all of $\mathrm{Gr}(2, W)$. This, of course, is impossible.
Point (4) follows from [FKM83, Theorem (d), p. 162]. □

1.2 A 1-Complete Variety of Cubic Hypersurfaces

Let K be an algebraically closed field of characteristic $\neq 2, 3$ and W a K-vector space of dimension $n+1$ with $n \geq 2$. This section is dedicated to the construction of a 1-complete variety of cubic hypersurfaces, starting from the projective space $V_0 := \mathbb{P}(\mathrm{Sym}^3(W^*))$ parametrizing the cubic hypersurfaces in $\mathbb{P}(W)$, by blowing up V_0 five times along smooth centers. At each level except the first, the centers are given by an irreducible component of the intersection of all proper transforms of the line conditions.
We saw in the previous section that S_0 coincides with the set of cubic hypersurfaces divisible by the square of a non-constant polynomial. Hence, S_0 is the image of the morphism

$$\phi_0 : \mathbb{P}(W^*) \times \mathbb{P}(W^*) \longrightarrow \mathbb{P}(\mathrm{Sym}^3(W^*)), \quad ([\lambda], [\mu]) \mapsto [\lambda\mu^2]. \tag{1.1}$$

As ϕ_0 is injective, S_0 is a subvariety of $\mathbb{P}(\mathrm{Sym}^3(W^*))$ of dimension $2n$. Let $\Delta \subseteq \mathbb{P}(W^*) \times \mathbb{P}(W^*)$ be the diagonal. We write $B_0 \subseteq V_0$ for the locus $\phi_0(\Delta)$ of triple hyperplanes. For brevity, we will often write $\mathbb{P}^n = \mathbb{P}(W^*)$. An immediate generalization of [Alu90, Lemma 0.2] implies that the restriction of ϕ_0 to $\mathbb{P}^n \times \mathbb{P}^n \setminus \Delta$ is an isomorphism onto $S_0 \setminus B_0$.
The following result generalizes [Alu90, Lemma 0.1]. We will often use it without explicit reference.

Lemma 1.2.1. The line condition $L^\ell \subseteq \mathbb{P}(\mathrm{Sym}^3(W^*))$ corresponding to the line $\ell \subseteq \mathbb{P}^n$ is an integral degree 4 hypersurface. Let $c \in L^\ell$.

(1) L^ℓ is smooth at c if and only if c does not contain ℓ and intersects ℓ with multiplicity exactly 2 at some (unique) point. In particular, the line conditions are generically smooth along the locus S_0 of non-reduced cubics. More precisely, L^ℓ is smooth at $[\lambda\mu^2] \in S_0$ if and only if $[\lambda] \neq [\mu]$ and ℓ does not intersect $\lambda \cap \mu$.

(2) If c does not contain ℓ and intersects ℓ with multiplicity 3 at some (unique) point, then L^ℓ has multiplicity 2 at c. In particular, the line conditions have multiplicity 2 generically along the locus B_0 of triple hyperplanes.

(3) The tangent hyperplane to L^ℓ at a smooth point c consists of the cubics containing the unique point p of tangency of c to ℓ. The tangent cone in V_0 to L^ℓ at a cubic c not containing ℓ and intersecting ℓ in the unique triple point p is precisely the "doubled" hyperplane in V_0 consisting of all cubics containing p.

Proof. All claims are PGL-equivariant and can hence be checked for the line $\ell = \mathcal{V}(x_2, x_3, \ldots, x_n) \subseteq \mathbb{P}^n$. The equation defining L^ℓ is

$$\begin{aligned} 0 = &- 27a_{(0,0,0)}^2 a_{(1,1,1)}^2 + a_{(0,0,1)}^2 a_{(0,1,1)}^2 - 4a_{(0,0,0)} a_{(0,1,1)}^3 - 4a_{(0,0,1)}^3 a_{(1,1,1)} \\ &+ 18a_{(0,0,0)} a_{(0,0,1)} a_{(0,1,1)} a_{(1,1,1)}. \end{aligned} \tag{1.2}$$

(This is precisely the discriminant of an inhomogeneous univariate cubic polynomial.) A computation in `Macaulay2` shows that the singular locus of L^ℓ is set-theoretically cut out by

$$\begin{aligned} 0 &= 9a_{(0,0,0)} a_{(1,1,1)} - a_{(0,0,1)} a_{(0,1,1)}, \\ 0 &= a_{(0,0,1)}^2 - 3a_{(0,0,0)} a_{(0,1,1)}, \\ 0 &= a_{(0,1,1)}^2 - 3a_{(0,0,1)} a_{(1,1,1)}. \end{aligned}$$

Assume now that $c = a_{(0,0,0)} x_0^3 + \ldots + a_{(n,n,n)} x_n^3$ is a cubic not containing ℓ and tangent to ℓ at p. After acting by an element of PGL preserving ℓ we may assume $p = [0 : 1 : 0 : \cdots : 0]$. The scheme-theoretic intersection of ℓ and c is the subscheme $\mathcal{V}(a_{(0,0,0)} x_0^3 + \ldots + a_{(1,1,1)} x_1^3)$ of $\ell \cong \mathbb{P}^1$. By assumption, the homogeneous bivariate polynomial $a_{(0,0,0)} x_0^3 + \ldots + a_{(1,1,1)} x_1^3$ factors as $x_0^2(\alpha x_0 + \beta x_1)$ for some $[\alpha : \beta] \in \mathbb{P}^1$, hence $a_{(0,1,1)} = a_{(1,1,1)} = 0$. Therefore, c lies in the singular locus of L^ℓ if and only if $a_{(0,0,1)} = 0$. On the other hand, $a_{(0,0,1)} = 0$ if and only if $\beta = 0$, i.e., if and only if p is a triple point. If c is a smooth point of L^ℓ, so $a_{(0,0,1)} \neq 0$, then computing the partial derivatives of the defining equation of L^ℓ above shows that the tangent hyperplane of L^ℓ at c is defined by $a_{(1,1,1)} = 0$, so it consists precisely of all cubics passing through p. If c is a singular point of L^ℓ not containing ℓ, so that p is a triple point of the intersection, then $a_{(0,0,1)} = a_{(0,1,1)} = a_{(1,1,1)} = 0$ and so necessarily $a_{(0,0,0)} \neq 0$ since c does not contain ℓ. In this case, dehomogenizing the defining equation of L^ℓ with respect to $a_{(0,0,0)}$ shows that the tangent cone at any such cubic c is scheme-theoretically defined by $a_{(1,1,1)}^2 = 0$. This proves all claims. $\square$

The next subsections explain in details the construction of a 1-complete variety of cubic hypersurfaces. A schematic overview of this construction and the notation employed is englobed in the following diagram:

$$
\begin{tikzcd}
 & V_5 \arrow[d, "\pi_5"] & & \\
 & V_4 \arrow[d, "\pi_4"] & B_4 = \mathbb{P}(\mathcal{E}) \arrow[l, "j_4"'] \arrow[d] & \\
 & V_3 \arrow[d, "\pi_3"] & B_3 = S_3 \arrow[l, "j_3"'] \arrow[d] & \mathrm{Bl}_\Delta \mathbb{P}^n \times \mathbb{P}^n \arrow[l, "\phi_3"'] \arrow[d, equal] \\
B_2 \arrow[r, "j_2"] \arrow[d] & V_2 \arrow[d, "\pi_2"] & S_2 \arrow[l] \arrow[d] & \mathrm{Bl}_\Delta \mathbb{P}^n \times \mathbb{P}^n \arrow[l, "\phi_2"'] \arrow[d, equal] \\
B_1 \arrow[r, "j_1"] \arrow[d] & V_1 \arrow[d, "\pi_1"] & S_1 \arrow[l] \arrow[d] & \mathrm{Bl}_\Delta \mathbb{P}^n \times \mathbb{P}^n \arrow[l, "\phi_1"'] \arrow[d] \\
B_0 = \nu_3(\mathbb{P}^n) \arrow[r, "j_0"] & V_0 = \mathbb{P}(\mathrm{Sym}^3(W^*)) & S_0 \arrow[l] & \mathbb{P}^n \times \mathbb{P}^n \arrow[l, "\phi_0"']
\end{tikzcd}
$$

The above diagram is analogous to the one in [Alu90, p. 514]. The center of each blow-up is denoted by B_i and we write $V_{i+1} := \mathrm{Bl}_{B_i} V_i$ with blow-up map $\pi_{i+1} : V_{i+1} \to V_i$ for each $i = 0, 1, 2, 3, 4$. Moreover, for $i \leq 3$ the proper transform of S_0 in V_i is denoted S_i.

1.2.0 Cubic Hypersurfaces

By $[a_{(0,0,0)} : \cdots : a_{(n,n,n)}]$ we denote the generic homogeneous coordinate vector of $V_0 = \mathbb{P}(\mathrm{Sym}^3(W^*))$ (after choosing a basis of W). In other words, $a_{(i,j,k)}$ is the coefficient of the monomial $x_i x_j x_k$ in the equation of the generic cubic in $\mathbb{P}^n$, where we assume $i \leq j \leq k$. We write $[n] := \{0, 1, \ldots, n\}$. Let $\mathcal{J}_0$ be the set of multi-indices $(i, j, k) \in [n]^3$ with $i \leq j \leq k$. Let $\mathcal{J}_0^* := \mathcal{J}_0 \setminus \{(0, 0, 0)\}$. Let $\mathcal{J}_1$ be the set of multi-indices $(i, j, k) \in [n]^3$ with $i \leq j \leq k$ and $j \geq 1$. Then, the affine chart $D(a_{(0,0,0)})$ of V_0 is described by the affine coordinates $(a_I)_{I \in \mathcal{J}_0^*}$. Moreover, in this same chart the ideal $\mathcal{I}(B_0)$ of the locus of triple hyperplanes in V_0 is generated by the polynomials $(f_J)_{J \in \mathcal{J}_1}$, where:

$$
\begin{aligned}
f_{(0,i,i)} &:= 3a_{(0,i,i)} - a_{(0,0,i)}^2, && \text{for } i > 0, \\
f_{(0,i,j)} &:= 3a_{(0,i,j)} - 2a_{(0,0,i)} a_{(0,0,j)}, && \text{for } j > i > 0, \\
f_{(i,i,i)} &:= 9a_{(i,i,i)} - a_{(0,0,i)} a_{(0,i,i)}, && \text{for } i > 0, \\
f_{(i,i,j)} &:= 3a_{(i,i,j)} - a_{(0,i,i)} a_{(0,0,j)}, && \text{for } i, j > 0, i \neq j, \\
f_{(i,j,k)} &:= 3a_{(i,j,k)} - a_{(0,i,j)} a_{(0,0,k)}, && \text{for } k > j > i > 0.
\end{aligned}
\tag{1.3}
$$

These polynomials will provide local coordinates for the first blow-up. Note that B_0 is a smooth complete intersection of codimension $\binom{n+3}{3} - 1 - n$ inside the chosen affine chart.

1.2.1 First Blow-up

Let $V_1 = \mathrm{Bl}_{B_0} V_0$ and denote by L_1, P_1 the proper transforms in V_1 of a line condition L and a point condition P, respectively.

Coordinates I. Let $((a_I), [b_J])$ denote the coordinates on $D(a_{(0,0,0)}) \times \mathbb{P}^{r-1}$, where $r = \binom{n+3}{3} - 1 - n$ is the codimension of B_0 as subvariety of V_0 and $I \in \mathcal{J}_0^*,\ J \in \mathcal{J}_1$. Then, by [Eis95, Exercise 17.14(b)], in this open chart the blow-up V_1 is the closed subscheme of the product defined by

$$f_{J_1} b_{J_2} - f_{J_2} b_{J_1} = 0,$$

where $J_1, J_2 \in \mathcal{J}_1$. In the affine open chart $D(a_{(0,0,0)}) \cap D(b_{(0,1,1)})$, introducing a new variable a', the blow-up V_1 can be described by

$$a' - f_{(0,1,1)} = 0, \qquad f_J - b_J a' = 0, \qquad \text{for all } \ J \in \mathcal{J}_1^* \coloneqq \mathcal{J}_1 \setminus \{(0,1,1)\}.$$

Hence, this open chart of V_1 is described by the coordinates $(a_{(0,0,1)}, \ldots, a_{(0,0,n)}, a', b_J)$ with $J \in \mathcal{J}_1^*$, subject to no more conditions. In particular, this open chart of V_1 is just an affine space. The equation for the exceptional divisor E_1 in this chart is $a' = 0$.

Denote by $N_{\mathbb{P}(W^*)}\mathbb{P}(\mathrm{Sym}^d(W^*))$ the normal bundle of the d-th Veronese embedding $\nu_d : \mathbb{P}(W^*) \hookrightarrow \mathbb{P}(\mathrm{Sym}^d(W^*)), [\lambda] \mapsto [\lambda^d]$.

Lemma 1.2.2. Let $e \leq d$ and assume $\mathrm{char}(K) = 0$ or $\mathrm{char}(K) > d$. Then there is an embedding of normal bundles

$$\alpha_{e,d} \ : \ N_{\mathbb{P}(W^*)}\mathbb{P}(\mathrm{Sym}^e(W^*)) \ \hookrightarrow \ N_{\mathbb{P}(W^*)}\mathbb{P}(\mathrm{Sym}^d(W^*)),$$

given by "multiplication by λ^{d-e}" in the fiber over $[\lambda] \in \mathbb{P}(W^*)$.

Proof. We write $R \coloneqq \mathrm{Sym}^\bullet(W) = K[x_0, \ldots, x_n]$ after choosing a basis of W. The pullback of the Euler sequence on $\mathbb{P}(\mathrm{Sym}^d(W^*))$ via ν_d is

$$0 \to \mathcal{O}_{\mathbb{P}(W^*)} \overset{\nu_d^*(\varepsilon)}{\to} \mathrm{Sym}^d(W^*) \otimes \mathcal{O}_{\mathbb{P}(W^*)}(d) \to T\mathbb{P}(\mathrm{Sym}^d(W^*))|_{\mathbb{P}(W^*)} \to 0,$$

where $\nu_d^*(\varepsilon)$ is induced by the graded R-module homomorphism

$$R \to \mathrm{Sym}^d(W^*) \otimes_K R(d),\ f \mapsto \sum_{|I|=d} \binom{d}{I} e_I \otimes (x^I f) = f \cdot (e_0 \otimes x_0 + \ldots + e_n \otimes x_n)^d,$$

where $e_0, \ldots, e_n$ is the basis of W^* dual to that of W. (Note that the "diagonal" $\sum_{i=0}^n e_i \otimes x_i$ is invariant under change of basis since it corresponds to the identity map under the isomorphism $W^* \otimes W \cong \mathrm{Hom}_K(W, W)$.) The fiber of $\nu_d^*(\varepsilon)$ over λ is therefore just multiplication by $\lambda^d = (\lambda_0 e_0 + \ldots + \lambda_n e_n)^d$. More generally, there is a commutative diagram with exact rows

$$\begin{tikzcd}
0 \arrow[r] & \mathcal{O}_{\mathbb{P}(W^*)} \arrow[r] \arrow[d, equal] & W^* \otimes \mathcal{O}_{\mathbb{P}(W^*)}(1) \arrow[r] \arrow[d, "\alpha_{1,e}"] & T\mathbb{P}(W^*) \arrow[r] \arrow[d, "\overline{\alpha_{1,e}}=\mathrm{d}\nu_e"] & 0 \\
0 \arrow[r] & \mathcal{O}_{\mathbb{P}(W^*)} \arrow[r] \arrow[d, equal] & \mathrm{Sym}^e(W^*) \otimes \mathcal{O}_{\mathbb{P}(W^*)}(e) \arrow[r] \arrow[d, "\alpha_{e,d}"] & T\mathbb{P}(\mathrm{Sym}^e(W^*))|_{\mathbb{P}(W^*)} \arrow[r] \arrow[d, "\overline{\alpha_{e,d}}"] & 0 \\
0 \arrow[r] & \mathcal{O}_{\mathbb{P}(W^*)} \arrow[r] & \mathrm{Sym}^d(W^*) \otimes \mathcal{O}_{\mathbb{P}(W^*)}(d) \arrow[r] & T\mathbb{P}(\mathrm{Sym}^d(W^*))|_{\mathbb{P}(W^*)} \arrow[r] & 0.
\end{tikzcd}$$

In here, $\alpha_{e,d}$ is induced by the graded R-module homomorphism which is multiplication by $(e_0 \otimes x_0 + \ldots + e_n \otimes x_n)^{d-e}$. It can be checked that $\overline{\alpha_{1,e}} = \mathrm{d}\nu_e$ is the differential of the e-th Veronese embedding. Then $\alpha_{e,d}$ induces the embedding of normal bundles we are looking for. □

For us, $e = 2$, $d = 3$. The exceptional divisor is $E_1 \cong \mathbb{P}(N_{\mathbb{P}(W^*)}\mathbb{P}(\mathrm{Sym}^3(W^*)))$ and we call B_1 the image of $\mathbb{P}(\alpha_{2,3})$ in E_1. The proper transform of S_0 in V_1 will be denoted by S_1.

Proposition 1.2.3. The intersection of the proper transforms of all line conditions in V_1 set-theoretically agrees with the union $S_1 \cup B_1$.

Proof. It is enough to check that the intersection of the proper transforms of all line conditions and E_1 equals B_1. The intersection of the proper transform L_1 of a line condition L with the fiber over $[\lambda^3] \in B_0$ is the image of the tangent cone of L at the point $[\lambda^3]$ in the projectivized normal bundle $\mathbb{P}(N_{B_0}V_0)$. By definition of $\alpha_{2,3}$ in Lemma 1.2.2, the fiber of B_1 over $[\lambda^3]$ consists of all cubics divisible by λ. Now, Lemma 1.2.1(3) implies that the intersection of all tangent cones at $[\lambda^3]$ of all line conditions is contained in the set of cubics containing the hyperplane λ. This proves that the intersection of E_1 with all line conditions in V_1 is contained in B_1. On the other hand, every individual $\tilde{L}_\ell$ contains a non-empty open of B_1 and hence all of B_1, proving equality. □

Lemma 1.2.4. The natural action of $\mathrm{PGL}_{n+1} = \mathrm{GL}_{n+1}/K^\times$ on the exceptional divisor $e \subseteq \mathrm{Bl}_\Delta \mathbb{P}^n \times \mathbb{P}^n$ is transitive.

Proof. We have $e = \mathbb{P}(N_\Delta \mathbb{P}^n \times \mathbb{P}^n) = \mathbb{P}(T\Delta)$; an isomorphism $N_\Delta \mathbb{P}^n \times \mathbb{P}^n \cong T\Delta$ is provided for example by the difference of the differentials of the projections $d\mathrm{pr}_1 - d\mathrm{pr}_2$. Fix now two points $[\lambda], [\mu] \in \Delta$ and two non-zero normal vectors $(v_1, v_2) \in N_\Delta \mathbb{P}^n \times \mathbb{P}^n|_{[\lambda]}$ and $(w_1, w_2) \in N_\Delta \mathbb{P}^n \times \mathbb{P}^n|_{[\mu]}$. These two normal vectors are represented by two curves $\mathbb{A}^1 \to \mathbb{P}^n \times \mathbb{P}^n, t \mapsto ([\lambda + tv_1], [\lambda + tv_2])$ and $t \mapsto ([\mu + tw_1], [\mu + tw_2])$, respectively. We then only need to find $A \in \mathrm{GL}_{n+1}$ with $A\lambda = \mu$ and $A(v_1 - v_2) = w_1 - w_2$. Such an A exists if $v_1 - v_2$ is not a multiple of λ and $w_1 - w_2$ is not a multiple of μ. Both conditions are satisfied by the requirement that both (v_1, v_2) and (w_1, w_2) are non-zero normal vectors. □

Lemma 1.2.5. We have a commutative diagram

$$\begin{tikzcd} \mathrm{Bl}_\Delta \mathbb{P}^n \times \mathbb{P}^n \arrow[r, "\phi_1", "\cong"'] \arrow[d] & S_1 \arrow[r, hook] \arrow[d] & V_1 \arrow[d] \\ \mathbb{P}^n \times \mathbb{P}^n \arrow[r, "\phi_0"'] & S_0 \arrow[r, hook] & V_0, \end{tikzcd}$$

where ϕ_1 is an isomorphism. In particular, S_1 is smooth.

Proof. We write e for the exceptional divisor of $\mathrm{Bl}_\Delta \mathbb{P}^n \times \mathbb{P}^n$. The map ϕ_0 lifts to a map $\phi_1 : \mathrm{Bl}_\Delta \mathbb{P}^n \times \mathbb{P}^n \to S_1$ via the universal property of blowing up. Indeed, it can be checked that the pullback of the ideal sheaf $\mathcal{I}(B_0)$ via ϕ_0 is precisely the squared ideal sheaf $\mathcal{I}(\Delta)^2$ of the diagonal $\Delta \subseteq \mathbb{P}^n \times \mathbb{P}^n$, in particular the pullback of $\mathcal{I}(B_0)$ to $\mathrm{Bl}_\Delta \mathbb{P}^n \times \mathbb{P}^n$ defines an effective Cartier divisor, as needed. Clearly, ϕ_1 restricts to an isomorphism of $\mathrm{Bl}_\Delta \mathbb{P}^n \times \mathbb{P}^n \setminus e$ onto $S_1 \setminus E_1$. As $\mathrm{Bl}_\Delta \mathbb{P}^n \times \mathbb{P}^n$ and S_1 are projective varieties, ϕ_1 is a closed map, so surjectivity follows. In order to prove the injectivity of ϕ_1 we observe that ϕ_0 is an injective morphism between varieties over an algebraically closed field, hence ϕ_0 is universally injective. Base-changing ϕ_0 along the blow-up map $\pi_1 : V_1 \to V_0$ hence gives an injection $(\mathbb{P}^n \times \mathbb{P}^n) \times_{V_0} V_1 \to V_1$. The blow-up closure lemma ensures that $\mathrm{Bl}_\Delta \mathbb{P}^n \times \mathbb{P}^n$ is naturally a closed subscheme of $(\mathbb{P}^n \times \mathbb{P}^n) \times_{V_0} V_1$, and the composition $\mathrm{Bl}_\Delta \mathbb{P}^n \times \mathbb{P}^n \to V_1$ agrees with ϕ_1, showing that ϕ_1 is injective. By [Har95, Corollary 14.10], it remains to show that $(d\phi_1)_p : T_p(\mathrm{Bl}_\Delta \mathbb{P}^n \times \mathbb{P}^n) \to T_{\phi_1(p)} V_1$ is injective for all p in the exceptional divisor e of $\mathrm{Bl}_\Delta \mathbb{P}^n \times \mathbb{P}^n$. This matter is local and invariant under the PGL-action, so we may choose any point p in the fiber over $([1:0:\cdots:0],[1:0:\cdots:0]) \in \Delta$ by Lemma 1.2.4. Choose local coordinates

$$([1:\lambda_1:\cdots:\lambda_n],[1:\mu_1:\cdots:\mu_n]) \in \mathbb{P}^n \times \mathbb{P}^n.$$

The equations for Δ are $u_i \coloneqq \lambda_i - \mu_i = 0$ for all $i \in \{1,\ldots,n\}$. Thus, $\mathrm{Bl}_\Delta \mathbb{P}^n \times \mathbb{P}^n$ is described by the points $(\mu_1,\ldots,\mu_n,u_1,\ldots,u_n,[s_1,\ldots,s_n])$ such that $u_i s_j - u_j s_i = 0$ for all i,j. In the affine chart $D(s_1)$, the morphism ϕ_1 is given explicitly in the affine coordinates $(\mu_1,\ldots,\mu_n,u_1,s_2,\ldots,s_n)$ by

$$\begin{aligned}
a_{(0,0,1)} &= 3\mu_1 + u_1, & & & b_{(1,1,1)} &= 2\mu_1, & & \\
a_{(0,0,i)} &= 3\mu_i + s_i u_1, & &\text{for } i>1, & b_{(0,1,i)} &= 2s_i, & &\text{for } i>1,\\
a' &= -u_1^2, & & & b_{(0,i,j)} &= 2s_i s_j, & &\text{for } j>i>1,\\
b_{(0,i,i)} &= s_i^2, & &\text{for } i>1, & b_{(i,i,i)} &= 2\mu_i s_i^2, & &\text{for } i>1,\\
b_{(1,i,i)} &= 2\mu_i s_i, & &\text{for } i>1, & b_{(i,i,j)} &= 2\mu_i s_i s_j, & &\text{for } i,j>1, i\neq j,\\
b_{(1,1,i)} &= 2\mu_1 s_i, & &\text{for } i>1, & b_{(i,j,k)} &= 2s_k(\mu_i s_j + \mu_j s_i), & &\text{for } k>j>i>0.
\end{aligned}$$

The exceptional divisor e now has equation $u_1 = 0$. This explicit description of ϕ_1 allows us to conclude the proof by checking the non-degeneracy of the Jacobian at every point. Indeed, the $2n$ row vectors in the Jacobian corresponding to $a_{(0,0,i)}$ for $1 \le i \le n$, to $b_{(0,1,i)}$ for $2 \le i \le n$ and to $b_{(1,1,1)}$ are linearly independent. □

Lemma 1.2.6. The set-theoretic intersection of B_1 and S_1 is $\phi_1(e)$. Moreover, the proper transforms of the line conditions are generically smooth along B_1 and tangent to E_1 along B_1.

Proof. Since ϕ_1 is an isomorphism, we have $\phi_1(e) = S_1 \cap E_1$ and it suffices to show $\phi_1(e) \subseteq B_1$. By PGL-equivariance, it suffices to prove the inclusion for the fiber in E_1 over $[x_0^3] \in B_0$. Using the coordinates described above, the intersection of this fiber with B_1 in V_1 is described by the equations $a_{(0,0,i)} = 0$ for all $1 \le i \le n$, $a' = 0$ and $b_J = 0$ for all those multi-indices $J \in \mathcal{J}_1^*$ with first entry non-zero. The explicit description of ϕ_1 shows that the image of the fiber of $([1:0:\dots:0],[1:0:\dots:0]) \in \mathbb{P}^n \times \mathbb{P}^n$ satisfies all these equations, proving the claim.
Again by PGL-invariance we can verify the second claim for the line $\ell = \mathcal{V}(x_2,\dots,x_n)$. In our affine open chart $D(a_{(0,0,0)}) \cap D(b_{(0,1,1)})$ the equation for the proper transform of the line condition L^ℓ is as follows. We start with (1.2) and plug in $3a_{(0,1,1)} = a' + a_{(0,0,1)}^2$ and $27a_{(1,1,1)} = 3b_{(1,1,1)}a' + a_{(0,0,1)}(a' + a_{(0,0,1)}^2)$, obtaining

$$(a')^2(12b_{(1,1,1)}a_{(0,0,1)} - 4a_{(0,0,1)}^2 - 4a' - 9b_{(1,1,1)}^2) = 0.$$

Outside of E_1 this describes the proper transform L_1^ℓ whose equation is therefore $-4a' - (3b_{(1,1,1)} - 2a_{(0,0,1)})^2 = 0$. Since the equation of E_1 in the local coordinates is $a' = 0$, every point of E_1 belonging to the proper transform is indeed a tangency point. Moreover, the equation shows that the proper transform is smooth in this entire chart.
There is a slightly technical but important point here in that the affine open chart $D(a_{(0,0,0)}) \cap D(b_{(0,1,1)})$ intersects the fiber over $[x_0^3] \in B_0$ in *every rank locus* of $B_1|_{[x_0^3]} = \mathbb{P}(\mathrm{Sym}^2(W^*)/\langle x_0 \cdot W^*\rangle)$, the quadrics on $\{x_0 = 0\}$. Since the orbits of the PGL-action on $B_1 = \mathbb{P}(N_{\mathbb{P}(W^*)}\mathbb{P}(\mathrm{Sym}^2 W^*))$ are precisely the rank loci (across the fibers), we deduce that for every point $p \in B_1$ there is a line ℓ such that $\tilde{L}_\ell$ is smooth at p and we have $T_p\tilde{L}_\ell = T_pE_1$ as subspaces of T_pV_1. ☐

Lemma 1.2.7. The ideal of $B_1 \subseteq V_1$ in the open chart $D(a_{(0,0,0)})$ is generated by the equations

$$\begin{aligned}
&f_J = 0, && \text{for } J \in \mathcal{J}_1\\
&f'_{(i,i,i)} := 3b_{(i,i,i)} - 2a_{(0,0,i)}b_{(0,i,i)} = 0, && \text{for } i > 0,\\
&f'_{(i,i,j)} := 3b_{(i,i,j)} - a_{(0,0,i)}b_{(0,i,j)} = 0, && \text{for } i,j > 0,\ i \neq j,\\
&f'_{(i,j,k)} := 3b_{(i,j,k)} - a_{(0,0,i)}b_{(0,j,k)} - a_{(0,0,j)}b_{(0,i,k)} = 0, && \text{for } k > j > i > 0.
\end{aligned}$$

In the chart $D(a_{(0,0,0)}) \cap D(b_{(0,1,1)})$, the first set of equations can be replaced simply by $a' = 0$. These equations clearly form a regular sequence, in fact they define smooth hypersurfaces intersecting transversally everywhere in the chart.

Proof. From the commutative diagram in the proof of Lemma 1.2.2, the fiber over $[\lambda^3] \in B_0$ of the normal bundle can be naturally identified with the vector space $\mathrm{Sym}^3(W^*)/\langle \lambda^2 x_0, \dots, \lambda^2 x_n\rangle$. Any $k \in \mathrm{Sym}^3(W^*)/\langle \lambda^2 x_0, \dots, \lambda^2 x_n\rangle$ can be uniquely written as a cubic not containing any monomial divisible by x_0^2, so we may uniquely write $k = k_{(0,1,1)}x_0x_1^2 + \ldots + k_{(n,n,n)}x_n^3$, and this normal vector is represented by the germ at 0 of the affine line

$$\mathbb{A}^1 \to D(a_{(0,0,0)}) \subseteq V_0, \quad t \mapsto \lambda^3 + t \cdot k.$$

Note that the $a_{(0,0,i)}$ coordinate of the latter equals $3\lambda_i$ for all $i \geq 1$. The proper transform of the curve therefore has the coordinates

$$\begin{aligned} b_{(i,j,k)} &= k_{(i,j,k)} - \lambda_k k_{(0,i,j)}, && \text{for } k > j > i > 0, \\ b_{(i,i,j)} &= k_{(i,i,j)} - \lambda_j k_{(0,i,i)}, && \text{for } i, j > 0, i \neq j, \\ b_{(i,i,i)} &= 3k_{(i,i,i)} - \lambda_i k_{(0,i,i)}, && \text{for } i > 0, \\ b_{(0,i,j)} &= k_{(0,i,j)}, && \text{for } j \geq i > 0. \end{aligned}$$

On the other hand, it is easy to see that the cubic k is divisible by $\lambda = x_0 + \lambda_1 x_1 + \ldots + \lambda_n x_n$ if and only if k satisfies the equations

$$\begin{aligned} k_{(i,j,k)} &= \lambda_i k_{(0,j,k)} + \lambda_j k_{(0,i,k)} + \lambda_k k_{(0,i,j)}, && \text{for } k > j > i > 0, \\ k_{(i,i,j)} &= \lambda_j k_{(0,i,i)} + \lambda_i k_{(0,i,j)}, && \text{for } i, j > 0, i \neq j, \\ k_{(i,i,i)} &= \lambda_i k_{(0,i,i)}, && \text{for } i > 0. \end{aligned}$$

The claim can be deduced directly from this. □

Lemma 1.2.8. Write $N_2 := N_{\mathbb{P}(W^*)}\mathbb{P}(\mathrm{Sym}^2(W^*))$ and $N_3 := N_{\mathbb{P}(W^*)}\mathbb{P}(\mathrm{Sym}^3(W^*))$ and let $p_1 : B_1 \to B_0$ be the restriction of the canonical map from the projective bundle $E_1 = \mathbb{P}(N_{B_0}V_0)$ to its base $B_0 \cong \mathbb{P}(W^*)$. Then there is a natural isomorphism

$$\begin{aligned} N_{B_1}E_1 &\cong p_1^*(N_3/N_2) \otimes_{\mathcal{O}_{B_1}} \mathcal{O}_{B_1}(1) \\ &\cong p_1^*\left(\frac{\mathrm{Sym}^3(W^*) \otimes \mathcal{O}_{\mathbb{P}(W^*)}(3)}{\mathrm{Sym}^2(W^*) \otimes \mathcal{O}_{\mathbb{P}(W^*)}(2)}\right) \otimes \mathcal{O}_{B_1}(1) \\ &\cong p_1^*(\mathrm{Sym}^3(T\mathbb{P}(W^*))) \otimes \mathcal{O}_{B_1}(1). \end{aligned}$$

Hence, over a point $([\lambda],[q]) \in B_1$, the normal space $N_{B_1}E_1|_{([\lambda],[q])}$ is naturally identified with $\mathrm{Sym}^3(W^*)/(\lambda \cdot \mathrm{Sym}^2(W^*))$. Points in B_2 can be thought of as triples consisting of a hyperplane λ together with a quadric q and a cubic c inside λ.

Proof. The first isomorphism is given by [EH16, Proposition 9.13]. The Euler sequences for $T\mathbb{P}(W^*)$, $T\mathbb{P}(\mathrm{Sym}^2(W^*))$, $T\mathbb{P}(\mathrm{Sym}^3(W^*))$ then give the second and third equality. □

1.2.2 Second Blow-up

Let $V_2 := \mathrm{Bl}_{B_1} V_1$. This is smooth because so is B_1. We denote $\pi_2 : V_2 \to V_1$ the blow-up map, and respectively $\tilde{E}_1, S_2, P_2, L_2$ the proper transforms of E_1, S_1, P_1, L_1. Moreover, we define $B_2 := \tilde{E}_1 \cap E_2 \cong \mathbb{P}(N_{B_1}E_1)$, where E_2 denotes the exceptional divisor in V_2.

Coordinates II. Let $(a_{(0,0,1)}, \ldots, a_{(0,0,n)}, a', b_J, [c_a, c_H])$ denote the coordinates for the product $\big(D(a_{(0,0,0)}) \cap D(b_{(0,1,1)})\big) \times \mathbb{P}^{r-1}$, where $r = \binom{n+3}{3} - \binom{n+2}{2} + 1$ is the codimension of B_1 as subvariety of V_1 and $J \in \mathcal{J}_1^*$. Here, c_a denotes a single variable and c_H is the set of variables indexed by $\mathcal{J}_2 := \{(i,j,k) \in [n]^3 : 1 \leq i \leq j \leq k\}$. Let also $\mathcal{J}_2^* := \mathcal{J}_2 \setminus \{(1,1,1)\}$.

Thanks to Lemma 1.2.7, the blow-up V_2 in the open chart $D(a_{(0,0,0)}) \cap D(b_{(0,1,1)})$ is the closed subvariety of the product given by

$$c_a f'_H - a' c_H = 0, \qquad c_{H_1} f'_{H_2} - c_{H_2} f'_{H_1} = 0,$$

for all $H, H_1, H_2 \in \mathcal{J}_2$. In the affine open chart of V_2 given by $D(c_{(1,1,1)})$, these equations simplify to

$$c_a f'_{(1,1,1)} - a' = 0, \qquad c_H f'_{(1,1,1)} - f'_H = 0,$$

where $H \in \mathcal{J}_2^*$. Introducing the new variable $b' := f'_{(1,1,1)}$, this affine open of V_2 has affine coordinates $(a_{(0,0,i)}, b_{(0,j,k)}, b', c_a, c_H)$ with $H \in \mathcal{J}_2^*$ subject to no relations. In these coordinates, the equation for E_2 in V_2 becomes $b' = 0$ and the equation for the proper transform $\tilde{E}_1$ becomes $c_a = 0$.

Lemma 1.2.9. The intersection of the proper transforms of all line conditions in V_2 set-theoretically agrees with $S_2 \cup B_2$, where $B_2 = \tilde{E}_1 \cap E_2$ is smooth.

Proof. The variety S_2 is clearly a component of the intersection. By Lemma 1.2.6, the line conditions in V_1 are generically smooth along B_1 and generically tangent to E_1 along B_1. Hence, the intersection of the proper transforms of the line conditions with the exceptional divisor E_2 is contained in $\tilde{E}_1$. On the other hand, every line condition in V_2 contains a non-empty open of B_2 and hence all of B_2, proving equality. □

A similar reasoning as in Lemma 1.2.5 shows also the following.

Lemma 1.2.10. The lift $\phi_2 : \mathrm{Bl}_\Delta \mathbb{P}^n \times \mathbb{P}^n \to V_2$ of ϕ_1 is explicitly given by

$$\begin{aligned}
a_{(0,0,1)} &= 3\mu_1 + u_1, & & & c_a &= u_1/2, & &\\
a_{(0,0,i)} &= 3\mu_i + s_i u_1, & &\text{for } i > 1, & c_{(1,1,i)} &= s_i, & &\text{for } i > 1,\\
b_{(0,i,i)} &= s_i^2, & &\text{for } i > 1, & c_{(1,i,i)} &= s_i^2, & &\text{for } i > 1,\\
b_{(0,1,i)} &= 2s_i, & &\text{for } i > 1, & c_{(i,i,j)} &= s_i^2 s_j, & &\text{for } i, j > 1,\\
b_{(0,i,j)} &= 2s_i s_j, & &\text{for } j > i > 1, & c_{(i,i,i)} &= s_i^3, & &\text{for } i \neq 0, 1,\\
b' &= -2u_1, & & & c_{(i,j,k)} &= 2s_i s_j s_k, & &\text{for } k > j > i > 0.
\end{aligned}$$

By Lemma 1.2.6, the set-theoretic intersection of S_1 with B_1 is given by $\phi_1(e)$. Scheme-theoretically, $S_1 \cap B_1 = \phi_1(e)^2$, i.e., the ideal sheaf is the square of the ideal sheaf of the reduced divisor $\phi_1(e)$. It follows that S_2 is isomorphic to S_1, hence to $\mathrm{Bl}_\Delta \mathbb{P}^n \times \mathbb{P}^n$. Abusing notation, we will indicate with e the exceptional divisor of $\mathrm{Bl}_\Delta \mathbb{P}^n \times \mathbb{P}^n$ as well as all its isomorphic images under the maps ϕ_i.

Lemma 1.2.11. The following hold:

(1) Scheme-theoretically, $S_2 \cap B_2 = e$.

(2) The line conditions in V_2 are generically smooth along B_2.

Proof. The first claim follows from Lemma 1.2.10 above using the PGL-action since B_2 is defined by $c_a = b' = 0$ in the chosen chart. For the second claim, observe that the line conditions in V_1 are generically smooth along B_1. The claim then follows from the fact that the blow-up of a smooth variety along a smooth center is again smooth. □

Remark 1.2.12. We denote by $B_2|_{[\lambda,q]}$ the fiber of B_2 over $[\lambda] \in B_0$ and $[q] \in B_1|_{[\lambda]} = \mathbb{P}(\mathrm{Sym}^2(W^*)/(\lambda \cdot W^*))$. Every element $k \in B_2|_{[\lambda,q]} \cong \mathbb{P}(\mathrm{Sym}^3(W^*)/(\lambda \cdot \mathrm{Sym}^2(W^*)))$ is naturally a cubic on λ. In the chart $D(a_{(0,0,0)})$, its defining equation can be uniquely written (up to scaling) in the form $k = k_{(1,1,1)}x_1^3 + k_{(1,1,2)}x_1^2x_2 + \ldots + k_{(n,n,n)}x_n^3$, not containing any monomial divisible by x_0. These coefficients are identified with the projective coordinates $[c_a, c_H]$ in Remark 1.2.2 via $c_a = 0$ and $k_H = 3c_H$ for those $H \in \mathcal{J}_2$ with at least two entries in $H = (i,j,k)$ being distinct, and $k_{(i,i,i)} = c_{(i,i,i)}$ for all $i \geq 1$.
Importantly, $S_2 \cap B_2$ consists of all triples $([\lambda],[q],[k]) = ([\lambda],[g^2],[g^3])$ for some hyperplane $[g] \in \mathbb{P}(W^*/\lambda)$ on λ. Inside our chart this follows in coordinates from the above together with the explicit description of ϕ_2 in Lemma 1.2.10. Using the PGL-action we can deduce the statement everywhere.

Proposition 1.2.13. Let $\overline{\lambda} := ([\lambda],[q],[k])$ be a point of B_2, i.e. a hyperplane λ together with a quadric q on λ and a cubic k on λ. Consider the line condition L_2^ℓ in V_2 corresponding to a line $\ell \subseteq \mathbb{P}(W)$. Then:

(1) The line ℓ intersects λ in the quadric q if and only if L_2^ℓ is tangent to E_2 at $\overline{\lambda}$.

(2) The line ℓ intersects λ in the cubic k if and only if L_2^ℓ is tangent to $\tilde{E}_1$ at $\overline{\lambda}$.

Here, by tangency we mean that the tangent space of E_2 (respectively $\tilde{E}_1$) at $\overline{\lambda}$ is contained in that of L_2^ℓ, hence they agree whenever the latter is smooth at $\overline{\lambda}$.

Proof. We can assume the hyperplane λ to be $\mathcal{V}(x_0)$ and ℓ the line $\mathcal{V}(x_1, x_3, \ldots, x_n)$. By plugging in the equations $c_{(1,2,2)}b' - 3b_{(2,2,2)} + 2a_{(0,0,2)}b_{(0,2,2)} = 0$ and $a' - c_a b' = 0$ in the equation of the proper transform of the line condition in V_1, we get the equation for L_2^ℓ in local coordinates in V_2 as

$$4b_{(0,2,2)}^3 c_a + c_{(2,2,2)}^2 b' = 0.$$

From Lemma 1.2.7 we get that the quadrics on λ intersecting ℓ are given by the equation $b_{(0,2,2)} = 0$. Moreover, from Remark 1.2.12 it follows that the cubics on λ intersecting ℓ are given by the equation $c_{(2,2,2)} = 0$. The statements on the tangency to E_2 and $\tilde{E}_1$ follow from these equations. □

Remark 1.2.14. We can notice that if the line ℓ is not contained in λ and does not intersect q and k in a common point, then the line condition L^ℓ is smooth at $\overline{\lambda}$. This is clear from the proof of the previous lemma when $\lambda = x_0$ and $\ell = \mathcal{V}(x_1, x_3, \ldots, x_n)$. The claim then follows by PGL-invariance.

1.2.3 Third Blow-up

Let $V_3 \coloneqq \mathrm{Bl}_{B_2} V_2$ with exceptional divisor E_3 and blow-up map $\pi_3 : V_3 \to V_2$. We denote by L_3 the proper transform in V_3 of the a line condition $L_2 \subseteq V_2$ and by S_3 the proper transform of S_2.

Coordinates III. In the chosen chart for V_2 the subvariety B_2 is defined by $b' = c_a = 0$. Consider the product $\left(D(a_{(0,0,0)}) \cap D(b_{(0,1,1)}) \cap D(c_{(1,1,1)})\right) \times \mathbb{P}^1$ with coordinates $(a_{(0,0,i)}, b_{(0,j,k)}, b', c_a, c_H, [d_c, d_b])$. The blow-up of B_2 in the chosen chart of V_2 can be described as the subvariety of the product given by

$$b' d_c = d_b c_a.$$

In the affine chart $D(a_{(0,0,0)}) \cap D(b_{(0,1,1)}) \cap D(c_{(1,1,1)}) \cap D(d_c)$ of V_3 we can work with the coordinates $(a_{(0,0,i)}, b_{(0,j,k)}, c_a, c_H, d_b)$ subject to no relations. The exceptional divisor E_3 is cut out by $c_a = 0$ in this chart. (In the chart using $D(d_b)$ instead, E_3 would be cut out by $b' = 0$.)

Remark 1.2.15. The line condition L_3^ℓ corresponding to $\ell \coloneqq \mathcal{V}(x_1, x_3, \ldots, x_n)$ has equation

$$4b_{(0,2,2)}^3 + c_{(2,2,2)}^2 d_b = 0.$$

Therefore, every other line condition obtained from this one by elements of PGL preserving the chart of V_3 will be of the form

$$4f(b_J)^3 + g(c_H)^2 d_b = 0,$$

where f is a linear function in the b_J coordinates and g is a linear function in the c_H coordinates.

We now prove that the intersection of all line conditions coincides with S_3.

Proposition 1.2.16. The intersection of all line conditions in V_3 is supported on the smooth variety S_3 only.

Proof. Clearly, S_3 is an irreducible component of the intersection. To see that it is the only one, recall that $B_2 = E_2 \cap \tilde{E}_1 = \mathbb{P}(N_{B_1} E_1)$ has codimension 2 in V_2. The exceptional divisor $E_3 = \mathbb{P}(N_{B_2} V_2)$ is therefore a $\mathbb{P}^1$-bundle over B_2. By $\overline{\lambda} \coloneqq ([\lambda], [q], [k])$ we denote again a point in B_2, i.e., a hyperplane λ together with a quadric q on λ and a cubic k on λ. Thanks to Remark 1.2.14, a general line condition in V_2 is smooth at $\overline{\lambda} \in B_2$, has codimension 1 in V_2, and contains B_2. Its proper transform intersects the fiber of $\mathbb{P}(N_{B_2} V_2)$ over $\overline{\lambda}$ in at most one point. We need to check that all points of the intersection of E_3 with all line conditions in V_3 is contained in the fibers over $B_2 \cap S_2 = \phi_2(e)$.
As $\tilde{E}_1$ and E_2 are smooth and intersect transversally everywhere, their proper transforms in V_3 cut the fiber of E_3 over any $\overline{\lambda} \in B_2$ in two *different* points r_1 and r_2, respectively. From Proposition 1.2.13 it follows that if a line ℓ intersects q, then the line condition L_3^ℓ contains r_2, while if ℓ intersects k, then the line condition L_3^ℓ contains the point r_1.

We claim that in order for the line conditions to intersect over $\overline{\lambda}$ we must have $q = gh$ and $k = gh^2$ where g, h are linear forms on the hyperplane λ. In fact, suppose there is a point of q which is not in k. Then, we can take a line ℓ in $\mathbb{P}^n$ passing through that point and not contained in λ. Thanks to Remark 1.2.14, the line condition L_2^ℓ is smooth at $\overline{\lambda}$ and L_3^ℓ intersects the fiber over $\overline{\lambda}$ in a unique point, necessarily in r_2. Take now another line condition $L_2^{\ell'}$ in V_2 such that the line ℓ' does not intersect the cubic nor the quadric. The line condition $L_2^{\ell'}$ is a hypersurface which is smooth at $\overline{\lambda}$ and contains B_2. If its proper transform intersects the fiber over $\overline{\lambda}$ in r_2, then $L_2^{\ell'}$ is tangent to E_2, and by Proposition 1.2.13 it would have to intersect the quadric, so this is impossible.
Similarly, we can show that there is no point of q which is not in k. Hence, we proved that in order for the line conditions to intersect over $\overline{\lambda}$ we must have $\mathcal{V}(q) = \mathcal{V}(k)$ set-theoretically. But this is equivalent to $q = gh$ and $k = gh^2$ with g, h linear forms on the hyperplane λ.
By Remark 1.2.12, we just have to show that $g = h$. It is enough to show this for $\lambda = x_0$ because the locus $B_2 \cap S_2$ is invariant under the induced PGL-action on V_2. Consider the point $\overline{x_0} = ([x_0^3], [q], [k])$, where

$$q = (g_1 x_1 + \cdots + g_n x_n)(h_1 x_1 + \cdots + h_n x_n),$$
$$k = (g_1 x_1 + \cdots + g_n x_n)(h_1 x_1 + \cdots + h_n x_n)^2$$

are, respectively, the quadric and the cubic on the hyperplane $x_0 = 0$.
After acting by a suitable element of PGL preserving $x_0 = 0$, we may assume $\overline{x_0}$ belongs to our chart $D(a_{(0,0,0)}) \cap D(b_{(0,1,1)}) \cap D(c_{(1,1,1)}) \subseteq V_2$. In particular, $g_1 \neq 0 \neq h_1$, so after rescaling $g_1 = h_1 = 1$.
We claim that for every index $i \geq 2$ we have $g_i = 0 \Leftrightarrow h_i = 0$. Suppose to the contrary that there is an $i \geq 2$ such that $h_i = 0$ and $g_i \neq 0$ or conversely. Consider the family of line conditions $L_2^{\ell_t}$ in V_2 for the lines $\ell_t = \mathcal{V}(x_1 - t^{-1} x_i, x_2, \ldots, \hat{x}_i, \ldots, x_n)$ with $t \in K \setminus \{0\}$. The equation of $L_2^{\ell_t}$ is

$$4(t^2 b_{(0,i,i)} - t b_{(0,1,i)} + 1)^3 c_a + (t^3 c_{(i,i,i)} - 3t^2 c_{(1,i,i)} + 3t c_{(1,1,i)} - 1)^2 b' = 0$$

in the chosen affine chart. The proper transform $L_3^{\ell_t}$ then intersects the fiber over $\overline{x_0}$ in E_3 in the point

$$[-(1 - 3t c_{(1,1,i)}(\overline{x_0}) + 3t^2 c_{(1,i,i)}(\overline{x_0}))^2 : 4(1 - t b_{(0,1,i)}(\overline{x_0}))^3]$$
$$= [-(1 - t g_i + t^2 h_i^2)^2 : 4(1 - t(g_i + h_i))^3].$$

If $g_i \neq 0$ but $h_i = 0$, it is clear that these would give different points in $\mathbb{P}^1$ for different values of t which means that the intersection of all line conditions above $\overline{x_0}$ is empty. If $g_i = 0$ but $h_i \neq 0$, a short computation shows that the same conclusion holds using $(1 - t^2 h_i^2) = (1 - t h_i)(1 + t h_i)$. So indeed $g_i = 0 \Leftrightarrow h_i = 0$ for all $i \geq 2$.
Finally, assume there is some $i \geq 2$ with $g_i \neq 0 \neq h_i$. We want to show $g_i = h_i$. For this, consider the line condition $L_2^{x_i}$ in V_2 corresponding to the line given by the vanishing of all coordinates except for x_0 and x_i. Its equation is

$$4 b_{(0,i,i)}^3 c_a + c_{(i,i,i)}^2 b' = 0$$

in the chosen chart. By assumption, $b_{(0,i,i)}(\overline{x_0}) = g_i h_i \neq 0 \neq g_i h_i^2 = c_{(i,i,i)}(\overline{x_0})$. Letting t tend to zero, the intersection point in $\mathbb{P}^1$ for the above family of line conditions tends to $[-1:4] \in \mathbb{P}^1$, so if the intersection of all line conditions above $\overline{x_0}$ is non-empty, then every line condition must intersect in the same point, necessarily in $[-1:4]$ by continuity. In particular, we must have

$$[-1:4] = [-(c_{(i,i,i)}(\overline{x_0}))^2 : 4(b_{(0,i,i)}(\overline{x_0}))^3] \Leftrightarrow g_i^2 h_i^4 = g_i^3 h_i^3,$$

and therefore $g_i = h_i$. This proves $g = h$. □

The second part of the last proof where we show $g = h$ seems to be absent from the proof of [Alu90, Proposition 3.2]. I am unsure whether this is a (minor) gap or I have missed an insight which would shorten the argument.

Since S_3 will be the next center for the blow-up, we write $B_3 := S_3$. From Lemma 1.2.11 we deduce $B_3 = S_3 \cong S_2$. The isomorphism $\phi_2 : \mathrm{Bl}_\Delta \mathbb{P}^n \times \mathbb{P}^n \to S_2$ defined in Lemma 1.2.10 lifts to the following map.

Lemma 1.2.17. The lift $\phi_3 : \mathrm{Bl}_\Delta \mathbb{P}^n \times \mathbb{P}^n \to V_3$ of ϕ_2 in the affine open chart is given by:

$$\begin{aligned} a_{(0,0,1)} &= 3\mu_1 + u_1, && \\ a_{(0,0,i)} &= 3\mu_i + s_i u_1, && \text{for } i > 1, \\ b_{(0,i,i)} &= s_i^2, && \text{for } i > 1, \\ b_{(0,1,i)} &= 2s_i, && \text{for } i > 1, \\ b_{(0,i,j)} &= 2s_i s_j, && \text{for } j > i > 1, \\ c_a &= u_1/2, && \end{aligned}$$

$$\begin{aligned} c_{(1,1,i)} &= s_i, && \text{for } i > 1, \\ c_{(1,i,i)} &= s_i^2, && \text{for } i > 1, \\ c_{(i,i,j)} &= s_i^2 s_j, && \text{for } i, j > 1, \\ c_{(i,j,k)} &= 2s_i s_j s_k, && \text{for } k > j > i > 0, \\ d_b &= -4. && \end{aligned}$$

From ϕ_3 we get local equations for $B_3 \subseteq V_3$ as follows.

Remark 1.2.18. The equations

$$\begin{aligned} & d_b + 4 = 0, && \\ g_{(0,1,i)} &:= b_{(0,1,i)} - 2c_{(1,1,i)} = 0, && \text{for } i > 1, \\ g_{(0,i,i)} &:= b_{(0,i,i)} - c_{(1,1,i)}^2 = 0, && \text{for } i > 1, \\ g_{(0,i,j)} &:= b_{(0,i,j)} - 2c_{(1,1,i)}c_{(1,1,j)} = 0, && \text{for } j > i > 1, \\ g_{(1,i,i)} &:= c_{(1,i,i)} - c_{(1,1,i)}^2 = 0, && \text{for } i > 1, \\ g_{(i,i,j)} &:= c_{(i,i,j)} - c_{(1,1,i)}^2 c_{(1,1,j)} = 0, && \text{for } i, j > 1, \\ g_{(i,j,k)} &:= c_{(i,j,k)} - 2c_{(1,1,i)}c_{(1,1,j)}c_{(1,1,k)} = 0, && \text{for } k > j > i > 0, \end{aligned}$$

cut out B_3 in the affine open chart. Clearly, these equations define smooth hypersurfaces intersecting transversally everywhere in the chart, in particular B_3 is a complete intersection of codimension $\binom{n+3}{3} - 1 - 2n$ in the chart.

1.2.4 Fourth Blow-up

Recall that $B_3 = S_3$. Let $V_4 \coloneqq \mathrm{Bl}_{B_3} V_3$. We write E_4 for the exceptional divisor and $\pi_4 : V_4 \to V_3$ for the blow-up map.

Coordinates IV. In the chosen affine chart of V_3 the base locus B_3 is cut out by the equations in Remark 1.2.18. Consider

$$\Big(D(a_{(0,0,0)}) \cap D(b_{(0,1,1)}) \cap D(c_{(1,1,1)}) \cap D(d_c)\Big) \times \mathbb{P}^{\binom{n+3}{3}-2n-2}$$

with coordinates $(a_{(0,0,i)}, b_{(0,j,k)}, c_a, c_H, d_b, [e_d, e_F])$, where e_d is a single variable and F runs through $\mathcal{J}_4 \coloneqq \{(i,j,k) \in [n]^3 : i \leq j \leq k\} \setminus \{(0,0,k), (1,1,k), (0,1,1)\}$. Moreover, let $\mathcal{J}_4^* \coloneqq \mathcal{J}_4 \setminus \{(0,1,2)\}$. The blow-up of V_3 along B_3 in the chosen affine chart can be described as the subvariety determined by

$$\begin{aligned} e_d g_F - (d_b + 4)e_F &= 0, &\quad &\text{for } F \in \mathcal{J}_4, \\ e_{F_1} g_{F_2} - g_{F_1} e_{F_2} &= 0, &\quad &\text{for } F_1, F_2 \in \mathcal{J}_4. \end{aligned}$$

In the affine chart $D(a_{(0,0,0)}) \cap D(y_{(0,1,1)}) \cap D(c_{(1,1,1)}) \cap D(d_c) \cap D(e_{(0,1,2)})$ of V_4, introducing the new variable $e' = g_{(0,1,2)}$, we can work with the coordinates

$$(a_{(0,0,1)}, \ldots, a_{(0,0,n)}, c_a, c_{(1,1,2)}, \ldots, c_{(1,1,n)}, e_d, e_F, e')$$

with $F \in \mathcal{J}_4^*$ subject to no relations. The exceptional divisor E_4 is cut out by $e' = 0$ in this chart.

Proposition 1.2.19. The intersection of all line conditions in V_4 is supported on a smooth subvariety B_4 of codimension $\binom{n+2}{3}$ inside E_4. More precisely, $B_4 = \mathbb{P}(\mathcal{E})$ where $\mathcal{E}$ is a vector subbundle of rank $\binom{n}{2}$ of the normal bundle $N_{B_3}V_3$.

Proof. We generalize the proof of [Alu90, Proposition 4.1]. Let $R_\mu \subseteq V_0$ denote the subvariety of cubics containing the hyperplane μ. Clearly, $R_\mu \cong \mathbb{P}(\mathrm{Sym}^2(W^*))$ is smooth. By Lemma 1.2.1, a line condition L^ℓ is smooth at $[\lambda\mu^2] \in S_0 \setminus B_0$ if the line ℓ intersects μ in a single point outside λ. Clearly, $T_{[\lambda\mu^2]}R_\mu \subseteq T_{[\lambda\mu^2]}L^\ell$ for every line ℓ, and Lemma 1.2.1 shows that

$$\bigcap_{\ell \subseteq \mathbb{P}(W)} T_{[\lambda\mu^2]}L^\ell = T_{[\lambda\mu^2]}R_\mu.$$

Clearly, for any fixed point $[\lambda\mu^2] \in B_3 \setminus e \cong S_0 \setminus B_0$, finitely many line conditions suffice for the intersection of their tangent spaces to agree with $T_{[\lambda\mu^2]}R_\mu$, and for these finitely many line conditions the same will be true in an open neighbourhood of $[\lambda\mu^2]$.

Now, by Proposition 1.2.16, the intersection of the proper transforms L_3^ℓ in V_3 for all lines ℓ agrees set-theoretically with $S_3 = B_3$. The proper transforms L_4^ℓ in V_4 therefore only intersect in the exceptional divisor E_4. We claim that their intersection is precisely the projectivization of a vector subbundle $\mathcal{E} \subseteq N_{B_3}V_3$. We construct $\mathcal{E}$ as the intersection of the images of the tangent sheaves $\mathcal{T}L_3^\ell|_{B_3}$ in $N_{B_3}V_3$ corresponding

to finitely many lines ℓ. The finiteness will ensure that the resulting subsheaf $\mathcal{E}$ of $N_{B_3}V_3$ is coherent.
First, we pick finitely many lines such that the intersection of the tangent spaces over every point $[\lambda\mu^2] \in B_3 \setminus e$ agrees with $T_{[\lambda\mu^2]}R_\mu$. The finiteness can be achieved by the quasi-compactness of $B_3 \setminus e$. The intersection of the images of the tangent sheaves in $N_{B_3}V_3$ of these line conditions then defines a coherent subsheaf $\mathcal{E}'$ which restricts to a vector subbundle over $B_3 \setminus e$. More precisely, by construction

$$\mathcal{E}'|_{[\lambda\mu^2]} \cong T_{[\lambda\mu^2]}R_\mu / T_{[\lambda\mu^2]}S_0 \subseteq N_{B_3}V_3|_{[\lambda\mu^2]}$$

for the geometric fiber of $\mathcal{E}'$ over the point $[\lambda\mu^2] \in B_3 \setminus e$. In particular, the rank of $\mathcal{E}'$ over $B_3 \setminus e$ is $r = \binom{n+2}{2} - 2n - 1 = \binom{n}{2}$. Moreover, by Lemma 1.2.1 and a Zariski closure argument, every other line condition L_4^ℓ contains $\mathbb{P}(\mathcal{E}'|_{B_3 \setminus e})$.
Next, we fix a point $p \in e = B_3 \cap E_3$ lying in our affine open chart. By Remark 1.2.15, in the chosen chart the equation for L_3^ℓ with ℓ any line passing through the point $[1 : 0 : \cdots : 0]$ does not depend on the variable c_a, and the equation determining E_3 in V_3 is exactly $c_a = 0$. The transversality of such line conditions can therefore be checked outside of E_3 and hence in $S_0 \setminus B_0$. Now, by Lemma 1.2.1, for any point $[\lambda\mu^2] \in B_3 \setminus e$ in our chart we even have

$$\bigcap_{[1:0:\cdots:0]\in\ell} T_{[\lambda\mu^2]}L^\ell = T_{[\lambda\mu^2]}R_\mu$$

because the locus of tangency of ℓ and $\lambda\mu^2$ (in case ℓ is not contained in the latter) is the unique point in $\ell \cap \mu$. Hence, we achieve the last equation already by running through all lines ℓ passing through $[1 : 0 : \cdots : 0]$ and any point of $\mu \setminus (\lambda \cap \mu)$ which will be the unique point of tangency. The corresponding line condition L^ℓ will be smooth at $[\lambda\mu^2]$ and its tangent hyperplane is the set of cubics containing the point $\ell \cap \mu$. The intersection of all these tangent hyperplanes is then the set of cubics containing a dense open subset of μ and hence μ itself.
We can therefore deduce that there are $\operatorname{codim}(R_\mu, V_0) = \binom{n+2}{3}$ lines $\ell_i \subseteq \mathbb{P}(W)$ passing through $[1 : 0 : \cdots : 0]$ such that the line conditions $L_3^{\ell_i}$ are all smooth at p and intersect transversally at p. Moreover, employing the PGL-action and using that it acts transitively on e by Lemma 1.2.4, we obtain finitely many more lines such that the intersection of their tangent spaces at every point of e has dimension at most r.
Finally, let $\mathcal{E}$ be the intersection of $\mathcal{E}'$ with the images of the tangent sheaves in $N_{B_3}V_3$ of these new line conditions. Then $\mathcal{E}$ is a coherent subsheaf of $N_{B_3}V_3$ which still restricts over $B_3 \setminus e$ to a vector subbundle of rank r and has rank $\leq r$ over every point of e. By upper semi-continuity of the rank, since $\mathcal{E}$ is coherent, $\mathcal{E}$ is a vector subbundle of $N_{B_3}V_3$ of rank r everywhere. As $\mathbb{P}(\mathcal{E})$ is an irreducible closed subset of V_4, a Zariski closure argument then shows that it is contained in L_4^ℓ for every ℓ, so it is contained in the intersection of all line conditions in V_4. And conversely, by construction $\mathbb{P}(\mathcal{E})$ contains the intersection of some (and hence of all) line conditions in V_4, proving equality. $\square$

Lemma 1.2.20. We have the isomorphism

$$N_{B_4}E_4 \cong (\pi_4|_{B_4})^*(N_{B_3}V_3/\mathcal{E}) \otimes \mathcal{O}_{B_4}(1).$$

Moreover, over $U \coloneqq B_4 \setminus (\pi_4|_{B_4})^{-1}(e)$ the normal bundle $N_{B_4}E_4$ restricts to

$$N_U E_4 \cong (\pi_4|_U)^* \left(\frac{\mathrm{Sym}^3(W^*) \otimes \mathcal{O}(1,2)}{\mathrm{Sym}^2(W^*) \otimes \mathcal{O}(1,1)} \right) \otimes \mathcal{O}_U(1),$$

where the notation $\mathcal{O}(a,b)$ suppresses the restriction to the open $\mathbb{P}^n \times \mathbb{P}^n \setminus \Delta$. In particular, the fiber of $N_{B_4}E_4$ over some point of $B_4 \setminus \pi_4^{-1}(e)$ mapping to $[\lambda\mu^2] \in B_3 \setminus e$ is given by $\mathrm{Sym}^3(W^*)/(\mu \cdot \mathrm{Sym}^2(W^*)) \cong \mathrm{Sym}^3(W^*/\mu)$, the *cubics on* μ.

The proof is similar to that of Lemma 1.2.8.

1.2.5 Fifth Blow-up

Let $V_5 \coloneqq \mathrm{Bl}_{B_4} V_4$ and E_5 the exceptional divisor. Let $\pi_5 : V_5 \to V_4$ be the blow-up map.

Lemma 1.2.21. Fix a line ℓ of $\mathbb{P}^n$ and a cubic $[\lambda\mu^2] \in B_3 \setminus e$ such that ℓ does not intersect $\lambda \cap \mu$. The strict transform L_5^ℓ in V_5 contains a point p in $E_5 \cap \tilde{E}_4 = \mathbb{P}(N_{B_4}E_4)$ with $(\pi_4 \circ \pi_5)(p) = [\lambda\mu^2]$ if and only if the line ℓ intersects the cubic on μ associated with p, i.e., the element of $\mathrm{Sym}^3(W^*)/(\mu \cdot \mathrm{Sym}^2(W^*))$.

Proof. By assumption, L_3^ℓ and its proper transforms are smooth at every point over $[\lambda\mu^2] \in B_3$. We have $(L_5^\ell \cap \tilde{E}_4 \cap E_5)|_{\pi_5(p)} = \mathbb{P}(N_{B_4}(L_4^\ell \cap E_4)|_{\pi_5(p)})$. Since $L_4^\ell \cap E_4|_U = \mathbb{P}(N_{B_3}L_3^\ell|_U)$ on the smooth locus U of L_3^ℓ inside B_3, we have the canonical isomorphisms

$$N_{B_4}(L_4^\ell \cap E_4)|_{\pi_5(p)} \cong ((\pi_4|_{B_4})^*(N_{B_3}L_3^\ell/\mathcal{E}) \otimes \mathcal{O}_{B_4}(1))|_{\pi_5(p)} \cong (N_{B_3}L_3^\ell/\mathcal{E})|_{[\lambda\mu^2]}.$$

Now, the tangent hyperplane $T_{[\lambda\mu^2]}L_3^\ell$ is given by those cubics containing the point $\ell \cap \mu$ by Lemma 1.2.1(3), and the fiber of $\mathcal{E}$ over $[\lambda\mu^2]$ is the quotient of the cubics containing μ by the tangent space of B_3 at $[\lambda\mu^2]$. We conclude that $(N_{B_3}L_3^\ell/\mathcal{E})|_{[\lambda\mu^2]}$ is exactly given by those cubics on μ containing the point $\ell \cap \mu$. □

Lemma 1.2.22. There exists a point $[\lambda\mu^2] \in B_3 \setminus e$ such that for every point $\overline{\lambda\mu^2} \in B_4$ over $[\lambda\mu^2]$, the intersection of all line conditions in V_5 over $\overline{\lambda\mu^2}$, i.e., in the fiber $E_5|_{\overline{\lambda\mu^2}}$, is contained in $\tilde{E}_4$.

Proof. We work in the chart of V_3 given by $D(a_{(0,0,0)}) \cap D(b_{(0,1,1)}) \cap D(c_{(1,1,1)}) \cap D(d_c)$. Inside the intersection of this chart with $B_3 \setminus e$ we choose the point $[\lambda\mu^2]$, where $\lambda = x_0 + x_1 + x_2$ and $\mu = x_0 + x_2$. Let $[q] \in B_4|_{[\lambda\mu^2]} = \mathbb{P}(\mathrm{Sym}^2(W^*)/(\lambda \cdot W^*, \mu \cdot W^*)) = \mathbb{P}(\mathrm{Sym}^2(W^*/\langle \lambda, \mu \rangle))$, i.e., q is a quadric on $\lambda \cap \mu$ determining $\overline{\lambda\mu^2}$. For our choice of λ and μ, q can be uniquely written, up to scaling, as $q = q_{(2,2)}x_2^2 + \ldots + q_{(n,n)}x_n^2$, so that the variables x_0 and x_1 do not occur. After acting by a suitable element of PGL fixing x_0, x_1 and x_2, we may assume $q_{(2,2)} \neq 0$.

We list some of the projective coordinates of the point $\overline{\lambda\mu^2}$ of $B_4|_{[\lambda\mu^2]}$ corresponding to $[q]$ in V_4:

$$\begin{aligned} e_d &= 0, & \\ e_{(0,1,i)} &= 0, & \text{for } i > 1, \\ e_{(0,i,j)} &= -3q_{(i,j)}, & \text{for } i,j > 1, \\ e_{(1,i,j)} &= -\frac{3}{2}q_{(i,j)}, & \text{for } i,j > 1, \\ e_{(i,i,j)} &= 0, & \text{for } i,j > 1. \end{aligned}$$

Moreover, $c_{(1,1,i)} = 0$ for all $i \geq 2$. Note that this point does *not* lie in the chart $D(e_{(0,1,2)})$ of V_4 but this is not a problem.
It suffices now to provide a family of line conditions L_4^ℓ whose tangent spaces at $\overline{\lambda\mu^2}$ converge to the tangent space of E_4 at $\overline{\lambda\mu^2}$. We claim that this is the case for the line conditions L^{ℓ_t} associated to the family of lines $\ell_t := \mathcal{V}(x_1 - tx_2, x_3, \ldots, x_n)$ with $t \in K \setminus \{0\}$. In our affine chart of V_3, the proper transform $L_3^{\ell_t}$ is given by

$$4\left(t^2 - tb_{(0,1,2)} + b_{(0,2,2)}\right)^3 + \left(t^3 - 3t^2c_{(1,1,2)} + 3tc_{(1,2,2)} - c_{(2,2,2)}\right)^2 d_b = 0.$$

We can see that $L_3^{\ell_t}$ is smooth at $[\lambda\mu^2]$ for all $t \neq 0$. Hence, the proper transforms $L_4^{\ell_t}$ in V_4 are smooth at every point $\overline{\lambda\mu^2}$ in $B_4|_{[\lambda\mu^2]}$. It is possible to compute the equation $\mathcal{L} = 0$ of $L_4^{\ell_t}$ in the chart $D(e_{(0,2,2)})$ (where e' corresponds to $g_{(0,2,2)}$). This chart contains our point $\overline{\lambda\mu^2}$ by the assumption that $q_{(2,2)} \neq 0$. The exceptional divisor is still cut out by $e' = 0$ in this chart, and the degree 0 part in the variable e' of $\mathcal{L}$ is given by

$$\begin{gathered} 12(t - c_{(1,1,2)})^4(e_{(0,2,2)} - te_{(0,1,2)}) + e_d(t - c_{(1,1,2)})^6 \\ -8(t - c_{(1,1,2)})^3(3te_{(1,2,2)} - e_{(2,2,2)}). \end{gathered}$$

It can be checked that all occurring terms of the latter which are not divisible by t^3 evaluate to zero at $\overline{\lambda\mu^2}$. Since $\overline{\lambda\mu^2} \in B_4 \subseteq E_4$, evaluating an expression at $\overline{\lambda\mu^2}$ includes setting to zero the variable e'. Hence, for every variable y of our chart which is different from e', we have that $(\partial_y\mathcal{L})(\overline{\lambda\mu^2})$ is divisible by t^3. On the other hand, the linear part of $\mathcal{L}$ in e' is

$$\begin{gathered} e'\Big(12(t - c_{(1,1,2)})^2(e_{(0,2,2)} - te_{(0,1,2)})^2 + 2e_d(t - c_{(1,1,2)})^3(3te_{(1,2,2)} - e_{(2,2,2)}) \\ -4(3te_{(1,2,2)} - e_{(2,2,2)})^2\Big). \end{gathered}$$

Hence,

$$(\partial_{e'}\mathcal{L})(\overline{\lambda\mu^2}) = 12t^2\left(e_{(0,2,2)}(\overline{\lambda\mu^2})^2 - 3e_{(1,2,2)}(\overline{\lambda\mu^2})^2\right) = 27t^2q_{(2,2)}.$$

Since the tangent space of $L_4^{\ell_t}$ at $\overline{\lambda\mu^2}$ is precisely the kernel of the gradient of $\mathcal{L}$ at $\overline{\lambda\mu^2}$, these computations show that $T_{\overline{\lambda\mu^2}}L_4^{\ell_t} \xrightarrow{t\to 0} T_{\overline{\lambda\mu^2}}E_4$ as subspaces of $T_{\overline{\lambda\mu^2}}V_4$. $\square$

Corollary 1.2.23. The intersection of all line conditions in V_5 is empty.

Proof. The line conditions can only intersect in E_5. Thanks to Remark 1.2.15 and the fact that the equations in Remark 1.2.18 do not involve the variable c_a, it suffices to prove the emptiness of the intersection only for the fibers over points $[\lambda\mu^2] \in B_3 \setminus e$. By the PGL-action it suffices to consider only a single such point $[\lambda\mu^2]$, for example the one in Lemma 1.2.22. The claim then follows from Lemmas 1.2.21 and 1.2.22 since the first implies that the intersection of all line conditions is disjoint from $E_5 \cap \tilde{E}_4$ while the second one states that it is contained in the latter. Hence, the intersection must be empty. $\square$

We have thus proved that the line conditions separate in V_5 and hence that V_5 is a 1–complete variety of cubic hypersurfaces.

1.2.6 Identifying the Vector Bundle $\mathcal{E}$

Let $V_0^Q := \mathbb{P}(\operatorname{Sym}^2 W^*)$ and $R_i \subseteq V_0^Q$ for $i \in \{1,2\}$ be the locus of quadrics of rank $\leq i$, i.e, $R_1 \subseteq V_0^Q$ is the image of the second Veronese $R_1 = \nu_2(\mathbb{P}^n)$ and R_2 is the set of reducible quadric forms. The proper transform of R_2 in $V_1^Q := \operatorname{Bl}_{R_1} V_0^Q$ is

$$\tilde{R}_2 = \operatorname{Bl}_{R_1} R_2 \cong (\mathbb{P}^n)^{[2]} \cong (\operatorname{Bl}_\Delta \mathbb{P}^n \times \mathbb{P}^n)/S_2,$$

where S_2 is the symmetric group on two elements swapping the factors of $\mathbb{P}^n \times \mathbb{P}^n$, and $(\mathbb{P}^n)^{[2]}$ is the Hilbert scheme of two points of $\mathbb{P}^n$. Let $q : \operatorname{Bl}_\Delta \mathbb{P}^n \times \mathbb{P}^n \longrightarrow \tilde{R}_2$ be the quotient map. (In fact, q is the universal family of $(\mathbb{P}^n)^{[2]}$.)
We will now describe explicitly an open chart of V_1^Q and its intersection with $\tilde{R}_2$. By $q_{(i,j)}$ for $0 \leq i \leq j \leq n$ we denote the homogeneous coordinates of $\mathbb{P}(\operatorname{Sym}^2(W^*))$ and work in the affine chart $D(q_{(0,0)})$. In this chart, R_1 is cut out by the equations

$$\begin{aligned} Q_{(i,i)} &:= 4q_{(i,i)} - q_{(0,i)}^2 = 0, \\ Q_{(i,j)} &:= 2q_{(i,j)} - q_{(0,i)}q_{(0,j)} = 0, \end{aligned}$$

where $1 \leq i < j \leq n$. These equations define smooth hypersurfaces of $D(q_{(0,0)})$ which intersect transversally everywhere. In particular, they form a regular sequence, so the blow-up can be realized as the closed subscheme of $D(q_{(0,0)}) \times \mathbb{P}^{\binom{n+1}{2}-1}$ given by

$$T_{(i,j)}Q_{(k,l)} - T_{(k,l)}Q_{(i,j)} = 0,$$

where $T_{(i,j)}$ are the projective coordinates on $\mathbb{P}^{\binom{n+1}{2}-1}$, $1 \leq i \leq j \leq n$. The affine chart $D(q_{(0,0)}) \cap D(T_{(1,1)})$ of the blow-up is then isomorphic to the affine space A with coordinates $q_{(0,1)}, \dots q_{(0,n)}, Q_{(1,1)}, T_{(i,j)}$ where $(i,j) \neq (1,1)$ and $1 \leq i \leq j \leq n$. In this chart, $\tilde{R}_2$ is cut out by

$$R_{(i,j)} := T_{(i,j)} - T_{(1,i)}T_{(1,j)} = 0$$

for $2 \leq i \leq j \leq n$. Again, the $R_{(i,j)}$ define smooth hypersurfaces of the chart intersecting transversally everywhere. If $U' \subseteq \operatorname{Bl}_\Delta \mathbb{P}^n \times \mathbb{P}^n$ is the affine space chart from the proof of Lemma 1.2.5 with coordinates $\mu_1, \dots, \mu_n, u_1, s_2, \dots, s_n$, then we have $\tilde{R}_2 \cap A = q(U')$. Both U' and $q(U')$ are affine spaces.

Theorem 1.2.24. We have $\mathcal{E} \cong q^*(N_{\tilde{R}_2} V_1^Q)$.

Proof. Since both $\mathcal{E}$ and $q^*(N_{\tilde{R}_2} V_1^Q)$ are vector bundles on the smooth variety $B_3 = \mathrm{Bl}_\Delta \mathbb{P}^n \times \mathbb{P}^n$, it is enough to give an embedding of geometric vector bundles

$$q^*(N_{\tilde{R}_2} V_1^Q)|_U \hookrightarrow N_{B_3} V_3|_U$$

whose image agrees with $\mathcal{E}|_U$ on an open $U \subseteq B_3$ whose complement has codimension ≥ 2 in B_3.

To achieve this, we write $U = U' \cup U''$ and give two embeddings of $q^*(N_{\tilde{R}_2} V_1^Q)$ into $N_{B_3} V_3$ over U' and U'', respectively, which agree on the overlap $U' \cap U''$ and hence glue to an embedding over U. Here, U' is as above and $U'' := B_3 \setminus e$. Hence, $B_3 \setminus U$ is a proper closed subset of e and thus of codimension ≥ 2 in B_3 because e is irreducible.

The embedding over U'' is the obvious one: Over U'', the differential of $\pi_3 \circ \pi_2 \circ \pi_1$ gives an isomorphism $N_{B_3} V_3|_{U''} \cong N_{S_0 \setminus \nu_3(\mathbb{P}^n)} V_0$. Over $[\lambda\mu^2] \in S_0 \setminus \nu_3(\mathbb{P}^n)$ the fiber of the latter is naturally isomorphic to $\mathrm{Sym}^3(W^*)/\langle \lambda\mu \cdot W^*, \mu^2 \cdot W^* \rangle$. Similarly,

$$q^*(N_{\tilde{R}_2} V_1^Q)|_{[\lambda\mu^2]} \cong \mathrm{Sym}^2(W^*)/\langle \lambda \cdot W^*, \mu \cdot W^* \rangle,$$

and multiplication by μ gives an embedding of the latter into the former. The image agrees with $\mathcal{E}|_{U''}$ by the proof of Proposition 1.2.19.

The open set $U' \subseteq B_3$ is the intersection of B_3 with the affine open chart $\hat{U}' := D(a_{(0,0,0)}) \cap D(b_{(0,1,1)}) \cap D(c_{(1,1,1)}) \cap D(d_c)$ of V_3 from Coordinates 1.2.3. Both U' and $\hat{U}'$ are affine spaces, the latter has coordinates $a_{(0,0,i)}, b_{(0,i,j)}, c_a, c_H, d_b$.

We will now describe an embedding of $q^*(N_{\tilde{R}_2} V_1^Q)|_{U'} \hookrightarrow N_{B_3} V_3|_{U'}$ explicitly in coordinates. The final step will then be to show that both embeddings agree on the overlap $U' \cap U'' = U' \setminus e$.

For this, let J_R be the Jacobian matrix whose rows are the gradients of the $R_{(i,j)}$. Then

$$N_{\tilde{R}_2} V_1^Q|_{q(U')} = TA|_{q(U')}/\ker(J_R) \cong q(U') \times \bigoplus_{2 \leq i \leq j \leq n} K e_{(i,j)},$$

where the last isomorphism is given by multiplication by J_R and the $e_{(i,j)}$ are basis vectors corresponding to the rows of J_R.

Similarly, let J_G be the Jacobian matrix whose rows are the gradients of the equations cutting out B_3 in the open chart $\hat{U}'$ of V_3, see Remark 1.2.18. Analogously we have

$$N_{B_3} V_3|_{U'} = T\hat{U}'|_{U'}/\ker(J_G) \cong U' \times \left(K e_{d_b} \oplus \bigoplus_{(i,j,k) \in \mathcal{F}} K e_{(i,j,k)} \right),$$

where the isomorphism is again given by multiplication by J_G and the $e_{d_b}, e_{(i,j,k)}$ are basis vectors corresponding to the rows of J_G.

The embedding $q^*(N_{\tilde{R}_2} V_1^Q)|_{U'} \hookrightarrow N_{B_3} V_3|_{U'}$ is now defined by sending $(p, r) \mapsto (p, v)$, where $r = (r_{(i,j)})_{2 \leq i \leq j \leq n} \in \bigoplus_{2 \leq i \leq j \leq n} K e_{(i,j)}$ and $v \in K e_{d_b} \oplus \bigoplus_{(i,j,k) \in \mathcal{F}} K e_{(i,j,k)}$ is given by:

$$\begin{aligned}
v_{d_b} &= 0, && \\
v_{(0,1,i)} &= 0, && \text{for } i > 1, \\
v_{(0,i,i)} &= \frac{3}{4} r_{(i,i)}, && \text{for } i > 1, \\
v_{(0,i,j)} &= \frac{3}{2} r_{(i,j)}, && \text{for } j > i > 1, \\
v_{(1,i,i)} &= \frac{3}{8} r_{(i,i)}, && \text{for } i > 1, \\
v_{(1,i,j)} &= \frac{3}{4} r_{(i,j)}, && \text{for } j > i > 1, \\
v_{(i,i,i)} &= \frac{9}{8} s_i r_{(i,i)}, && \text{for } i > 1, \\
v_{(i,i,j)} &= \frac{3}{4} s_i r_{(i,j)} + \frac{3}{8} s_j r_{(i,i)}, && \text{for } i \neq j \text{ and } i,j > 1, \\
v_{(i,j,k)} &= \frac{3}{4} s_k r_{(i,j)} + \frac{3}{4} s_j r_{(i,k)} + \frac{3}{4} s_i r_{(j,k)}, && \text{for } k > j > i > 1.
\end{aligned} \tag{1.4}$$

Here, the $s_i = s_i(p)$ are (some of) the coordinates of the point $p \in U'$. This is clearly an embedding of geometric vector bundles over U'.
Now let $p = [\lambda\mu^2] \in U' \cap U'' = U' \setminus e$. Since the differential of the blow-up map gives an isomorphism $N_{\tilde{R}_2} V_1^Q|_{q(U')\setminus q(e)} \cong N_{R_2 \setminus R_1} D(q_{(0,0)})$, any non-zero normal vector $\mathbf{n} \in N_{\tilde{R}_2} V_1^Q|_{q(p)}$ is represented by the proper transform of an affine line in $D(q_{(0,0)})$, centered at $q(p) = [\lambda\mu]$, which is not contained in R_2. The embedding over U'' then sends $\mathbf{n}$ to the normal vector in $N_{B_3} V_3|_p$ which is represented by the proper transform of the affine line in V_0, centered at $p = [\lambda\mu^2]$, which is just the *multiplication by* μ of the affine line in $D(q_{(0,0)})$. So in order to see that both embeddings agree on $U' \setminus e$, we need to compute the proper transforms of the two curves given by

$$\begin{aligned}
\mathbb{A}^1 &\longrightarrow D(q_{(0,0)}) \subseteq \mathbb{P}(\operatorname{Sym}^2 W^*), && t \mapsto \lambda\mu + t \cdot \tilde{q}, \\
\mathbb{A}^1 &\longrightarrow D(a_{(0,0,0)}) \subseteq \mathbb{P}(\operatorname{Sym}^3 W^*), && t \mapsto \lambda\mu^2 + t \cdot \tilde{q}\mu
\end{aligned}$$

in $A \subseteq V_1^Q$ and $\hat{U}' \subseteq V_3$, respectively. As we shall see, it is enough to consider only those quadrics $\tilde{q}$ not containing x_0 and x_1, i.e., $\tilde{q}_{(0,i)} = \tilde{q}_{(1,j)} = 0$ for all i, j.
Differentiating the proper transforms of both curves with respect to t and then setting $t = 0$ gives a tangent vector in $T_{q(p)}A$ resp. $T_p\hat{U}'$. Finally, multiplying these tangent vectors by J_R resp. J_G, we need to check that the results agree under the embedding (1.4).
Computing these proper transforms and derivatives is not particularly enlightening but also not very hard and can be done, for example, in `Macaulay2`. We only present the result. The $T_{(i,j)}$-coordinates of the proper transform of the first curve are as follows:

$$\begin{aligned}
T_{(i,i)} &= 4t\tilde{q}_{(i,i)} - u_1^2 s_i^2, \\
T_{(i,j)} &= 2t\tilde{q}_{(i,j)} - u_1^2 s_i s_j, \\
T_{(1,i)} &= -u_1^2 s_i.
\end{aligned}$$

Here, $2 \le i < j \le n$. Differentiating and multiplying by J_R gives

$$\begin{aligned} r_{(i,i)} &= -4\tilde{q}_{(i,i)}, \\ r_{(i,j)} &= -2\tilde{q}_{(i,j)}, \end{aligned} \tag{1.5}$$

where $2 \le i < j \le n$. (This shows that it suffices to consider only those $\tilde{q}$ not involving x_0 and x_1.) On the other hand, for the proper transform of the second curve we have the following coordinates:

$$\begin{aligned} d_b &= -4, \\ b_{(0,i,i)} &= -3t\tilde{q}_{i,i} + u_1^2 s_i^2, \\ b_{(0,i,j)} &= -3t\tilde{q}_{i,j} + 2u_1^2 s_i s_j, \\ c_{(1,1,i)} &= u_1^2 s_i, \\ c_{(1,i,i)} &= -\frac{3}{2}t\tilde{q}_{(i,i)} + u_1^2 s_i^2, \\ c_{(i,i,i)} &= -\frac{9}{2}ts_i\tilde{q}_{(i,i)} + u_1^2 s_i^3, \\ c_{(i,i,j)} &= -\frac{3}{2}t\big(s_i\tilde{q}_{(i,j)} + s_j\tilde{q}_{(i,i)}\big) + u_1^2 s_i^2 s_j, \\ c_{(i,j,k)} &= -\frac{3}{2}t\big(s_i\tilde{q}_{(j,k)} + s_j\tilde{q}_{(i,k)} + s_k\tilde{q}_{(i,j)}\big) + 2u_1^2 s_i s_j s_k. \end{aligned}$$

Here, we set $s_1 := 1$. Differentiating and multiplying by J_G gives

$$\begin{aligned} v_{d_b} &= 0, \\ v_{(0,1,i)} &= 0, \\ v_{(0,i,i)} &= -3\tilde{q}_{(i,i)}, \\ v_{(0,i,j)} &= -3\tilde{q}_{(i,j)}, \\ v_{(1,i,i)} &= -\frac{3}{2}\tilde{q}_{(i,i)}, \\ v_{(1,i,j)} &= -\frac{3}{2}\tilde{q}_{(i,j)}, \\ v_{(i,i,i)} &= -\frac{9}{2}s_i\tilde{q}_{(i,i)}, \\ v_{(i,i,j)} &= -\frac{3}{2}s_i\tilde{q}_{(i,j)} - \frac{3}{2}s_j\tilde{q}_{(i,i)}, \\ v_{(i,j,k)} &= -\frac{3}{2}s_i\tilde{q}_{(j,k)} - \frac{3}{2}s_j\tilde{q}_{(i,k)} - \frac{3}{2}s_k\tilde{q}_{(i,j)}. \end{aligned} \tag{1.6}$$

It is now easy to check that (1.5) transforms into (1.6) via (1.4). □

The following can be deduced from Theorem 1.2.24. For a different proof, we refer to [BDFK23, Proposition 2.25]. As we will not make direct use of it, we omit the proof here.

Corollary 1.2.25. We have

$$\mathcal{E}|_e \cong \mathrm{Sym}^2(T_{e/\Delta}) \cong \frac{\pi_e^*\left(\mathrm{Sym}^2(T\Delta)\right) \otimes \mathcal{O}_e(2)}{\pi_e^*(T\Delta) \otimes \mathcal{O}_e(1)},$$

where $T_{e/\Delta}$ is the relative tangent bundle of the projective bundle e over Δ.

1.3 Chow Rings and Chern Classes

In this section we determine the intersection-theoretic information needed for the computation of the characteristic numbers. This includes the Chow rings of the centers B_i and the total Chern classes of the normal bundles $N_{B_i}V_i$. From [Alu90, Section 2] we recall the following notion: Let V be a smooth variety and $B \subseteq V$ a smooth subvariety. Moreover, let $X \subseteq V$ be any equidimensional closed subscheme. The *full intersection class $B \circ X$ of X by B in V* is defined as

$$B \circ X := c(N_B V) \cap s(B \cap X, X)$$

in the Chow group $\mathrm{CH}_*(B \cap X)$, where $s(B \cap X, X)$ denotes the Segre class. By abuse of notation we will also write $B \circ X$ for the cycle class $(B \cap X \hookrightarrow B)_*(B \circ X) \in \mathrm{CH}_*(B)$. We will only need the following fact from [Alu90, p. 512].

Remark 1.3.1. If $X = D \subseteq V$ is a divisor and $i : B \hookrightarrow V$ the inclusion, then

$$B \circ D = e_B D[B] + i^*([D]) \in \mathrm{CH}_*(B),$$

where $e_B D$ denotes the multiplicity of D along B. In particular, if B is not contained in D, then $e_B D = 0$, and if D contains B and is generically smooth along B, then $e_B D = 1$.

Along the way, we will compute the full intersection classes $B_i \circ P_i$ and $B_i \circ L_i$ for all $i = 0, 1, 2, 3, 4$.

1.3.0 Chow Ring of B_0

Lemma 1.3.2. The Chow ring of $B_0 \cong \mathbb{P}(W^*)$ is generated by the hyperplane class h and we have $\mathrm{CH}^*(B_0) = \mathbb{Z}[h]/(h^{n+1})$. Moreover,

$$c(N_{B_0} V_0) = (1 + 3h)^{\binom{n+3}{3}} / (1 + h)^{n+1}.$$

Proof. This follows from the Euler sequences for $T\mathbb{P}(W^*)$ and TV_0 pulled back to $\mathbb{P}(W^*)$ via the third Veronese. □

Lemma 1.3.3. The full intersection classes of point and line conditions by B_0 in V_0 are

$$B_0 \circ P_0 = 3h, \qquad B_0 \circ L_0 = 2 + 12h.$$

Proof. This follows from Remark 1.3.1 and the fact that B_0 is contained in L_0 but not in P_0 and that L_0 has multiplicity 2 along B_0 by Lemma 1.2.1. (We are sloppy here in mistaking subschemes for their cycle classes but this simplifies the notation.) Moreover, for the hyperplane class $H \in \mathrm{CH}^1(V_0)$ we have $\nu_3^*(H) = 3h \in \mathrm{CH}^1(B_0)$ because $\nu_3^*(\mathcal{O}_{V_0}(1)) = \mathcal{O}_{B_0}(3)$. □

1.3.1 Chow Ring of B_1

In Subsection 1.2.1 we described the center B_1. Identifying $B_0 \cong \mathbb{P}(W^*)$, by definition $B_1 \cong \mathbb{P}(N_1)$, where $N_1 = N_{\mathbb{P}(W^*)}\mathbb{P}(\mathrm{Sym}^2(W^*))$.

Lemma 1.3.4. We have $\dim(B_1) = \binom{n+2}{2} - 2$. The Chow ring of B_1 is

$$\mathrm{CH}^*(B_1) = \mathrm{CH}^*(B_0)[\epsilon]/(p_1(\epsilon, h)) = \mathbb{Z}[\epsilon, h]/(p_1(\epsilon, h), h^{n+1}),$$
$$p_1(\epsilon, h) := \sum_{i=0}^{\binom{n+1}{2}} (-\epsilon)^{\binom{n+1}{2}-i} \cdot c_i(N_1),$$

where $\epsilon := (B_1 \hookrightarrow V_1)^*(E_1) = c_1(\mathcal{O}_{B_1}(-1))$. Explicitly,

$$c(N_1) = (1 + 2h)^{\binom{n+2}{2}}/(1 + h)^{n+1}.$$

Moreover, $\int_{B_1} h^n(-\epsilon)^{\binom{n+1}{2}-1} = 1$. For the total Chern class of $N_{B_1}V_1$ we have

$$c(N_{B_1}V_1) = (1 + \epsilon)(1 + 3h - \epsilon)^{\binom{n+3}{3}}/(1 + 2h - \epsilon)^{\binom{n+2}{2}}.$$

Proof. For the Chow ring of a projective bundle over a smooth base see [EH16, Theorem 9.6]. Lemma 9.7 in *loc. cit.* shows also $\int_{B_1} h^n(-\epsilon)^{\binom{n+1}{2}-1} = 1$. The total Chern class $c(N_1)$ can be computed from the Euler sequences for $T\mathbb{P}(W^*)$ and $T\mathbb{P}(\mathrm{Sym}^2 W^*)$, pulled back to $\mathbb{P}(W^*)$ along the second Veronese. Moreover, $c(N_{B_1}V_1) = c(N_{E_1}V_1|_{B_1})c(N_{B_1}E_1) = (1 + \epsilon)c(N_{B_1}E_1)$, and the second factor may be computed from Lemma 1.2.8. □

Lemma 1.3.5. We have $\pi_1^*(P_0) = P_1$ and $\pi_1^*(L_0) = L_1 + 2E_1$. The full intersection classes of point and line conditions with respect to B_1 are:

$$B_1 \circ P_1 = 3h, \qquad\qquad B_1 \circ L_1 = 1 + 12h - 2\epsilon.$$

Proof. The equation $\pi_1^*(L_0) = L_1 + 2E_1$ uses that L_0 has multiplicity 2 along B_0. The remaining claims follow again from Remark 1.3.1 together with the fact that B_1 is contained in L_1 but not in P_1 and that L_1 is generically smooth along B_1 by Lemma 1.2.6. □

1.3.2 Chow Ring of B_2

The center $B_2 = E_2 \cap \tilde{E}_1 \cong \mathbb{P}(N_{B_1}E_1)$ was described in Subsection 1.2.2.

Lemma 1.3.6. We have $\dim(B_2) = \binom{n+3}{3} - 3$. For the Chow ring we have

$$\mathrm{CH}^*(B_2) = \mathrm{CH}^*(B_1)[\phi]/(p_2(\phi, \epsilon, h)) = \mathbb{Z}[\phi, \epsilon, h]/(p_2(\phi, \epsilon, h), p_1(\epsilon, h), h^{n+1}),$$
$$p_2(\phi, \epsilon, h) := \sum_{i=0}^{\binom{n+2}{3}} (-\phi)^{\binom{n+2}{3}-i} \cdot c_i(N_{B_1}E_1),$$

where $\phi := (B_2 \hookrightarrow V_2)^*(E_2) = c_1(\mathcal{O}_{B_2}(-1))$ and

$$c(N_{B_1}E_1) = (1+3h-\epsilon)^{\binom{n+3}{3}}/(1+2h-\epsilon)^{\binom{n+2}{2}}.$$

Moreover, $\int_{B_2} h^n(-\epsilon)^{\binom{n+1}{2}-1}(-\phi)^{\binom{n+2}{3}-1} = 1$. For the normal bundle $N_{B_2}V_2$ we have

$$c(N_{B_2}V_2) = (1+\phi)(1+\epsilon-\phi).$$

Proof. The proof is very similar to that of Lemma 1.3.4. For the last claim we observe $c(N_{B_2}V_2) = c(N_{E_2}V_2|_{B_2})c(N_{B_2}E_2) = (1+\phi)c(N_{B_2}E_2)$ since $N_{E_2}V_2|_{B_2} \cong \mathcal{O}_{B_2}(-1)$. Moreover, $N_{B_2}E_2 \cong \pi_2|_{B_2}^*(\mathcal{O}_{B_1}(-1)) \otimes \mathcal{O}_{B_2}(1)$ which follows from the two relative Euler sequences of the projective bundles B_2 and E_2 over B_1. This shows $c(N_{B_2}E_2) = 1+\epsilon-\phi$. □

Lemma 1.3.7. We have $\pi_2^*(P_1) = P_2$ and $\pi_2^*(L_1) = L_2 + E_2$. The full intersection classes of point and line conditions by B_2 in V_2 are:

$$B_2 \circ P_2 = 3h, \qquad\qquad B_2 \circ L_2 = 1 + 12h - 2\epsilon - \phi.$$

Proof. By Lemma 1.2.6, L_1 is generically smooth along B_1, so the same is true for L_2 along B_2. Hence, $(B_2 \hookrightarrow V_2)^*(L_2) = (B_2 \hookrightarrow V_2)^*(\pi_2^*(L_1) - E_2) = 12h - 2\epsilon - \phi$. □

1.3.3 Chow Ring of B_3

Recall from Subsection 1.2.3 that B_3 denotes the fourth center in our sequence of blow-ups, and it is defined as the proper transform in V_3 of S_0. Moreover, in Lemma 1.2.17, we described the isomorphism $\phi_3 : \mathrm{Bl}_\Delta \mathbb{P}^n \times \mathbb{P}^n \xrightarrow{\cong} S_3 = B_3$.

Lemma 1.3.8. We identify $B_3 = \mathrm{Bl}_\Delta \mathbb{P}^n \times \mathbb{P}^n$. Clearly, $\dim(B_3) = 2n$. For the Chow ring of B_3 we have

$$\mathrm{CH}^*(B_3) = \frac{\mathbb{Z}[e,l,m]}{\left(l^{n+1}, m^{n+1}, el - em, \sum_{i=0}^n l^i m^{n-i} + \sum_{i=0}^{n-1} \binom{n+1}{i} l^i(-e)^{n-i}\right)},$$

where l and m are the pullbacks of the hyperplane class of $\mathbb{P}^n$ via the first resp. second projection to $\mathbb{P}^n$, and e is the class of the exceptional divisor. Moreover, $\mathrm{CH}_0(B_3) = \mathbb{Z} \cdot l^n m^n$ and $\int_{B_3} l^n m^n = 1$.

Proof. The identification of the Chow ring follows from [EH16, Theorem 13.14] and we omit the details. Importantly, the class of the diagonal is $[\Delta] = \sum_{i=0}^n l^i m^{n-i}$. Moreover, $el = em$ follows from $el = j_*(j^*(l)) = j_*(h_\Delta) = j_*(j^*(m)) = em$, where h_Δ is the hyperplane class on $\Delta \cong \mathbb{P}^n$ pulled back to e and $j : e \hookrightarrow B_3$ is the inclusion. We have $\int_{B_3} l^n m^n = 1$ since $l^n m^n$ is the pullback of the class represented by any point in $\mathbb{P}^n \times \mathbb{P}^n$. The claim about $\mathrm{CH}_0(B_3)$ follows from the last displayed relation because every monomial in l, m, e in which e has exponent $\geq n$ is equal to a sum of terms in which e only occurs with exponents $< n$. But then, because of $el = em$ and $l^{n+1} = m^{n+1} = 0$, every monomial of degree $\dim(B_3) = 2n$ is equal to an integer multiple of $l^n m^n$. □

Remark 1.3.9. The computation of the total Chern class $c(N_{B_3}V_3)$ can be done using Fulton's formula for blowing up Chern classes [Ful98, Theorem 15.4] several times. The result for $n = 3$ is

$$\begin{aligned} c(N_{B_3}V_3) = {} & 93960l^3m^3 - 215300l^3m^2 - 233400e^2m^3 + 49580l^2m^3 \\ & - 63500l^3m - 33300e^2m^2 - 25900l^2m^2 + 132200em^3 \\ & + 2220lm^3 - 7267l^3 - 4107l^2m + 4796em^2 + 573lm^2 \\ & - 1367m^3 + 705e^2 + 120l^2 - 1880em + 536lm + 610m^2 \\ & - 39e + 16l + 36m + 1. \end{aligned}$$

The computation can be carried out, for instance, in `Macaulay2`, and in principle it is be possible to write down an explicit formula for $c(N_{B_3}V_3)$ for arbitrary n.

Lemma 1.3.10. We have $\pi_3^*(P_2) = P_3$ and $\pi_3^*(L_2) = L_3 + E_3$. The full intersection classes of point and line conditions with respect to B_3 are:

$$B_3 \circ P_3 = l + 2m, \qquad B_3 \circ L_3 = 1 + 4l + 8m - 6e.$$

Proof. In the Chow ring of B_3 we have $j_3^*(P_3) = l + 2m$ and $j_3^*(L_3) = 4(l + 2m) - 2(2e) - e - e = 4l + 8m - 6e$. The assertion then follows from the fact that P_3 does not contain B_3 and that L_3 is generically smooth along B_3. □

1.3.4 Chow Ring of B_4

The center of the last blow up is B_4 which was described in Subsection 1.2.4. Recall that $B_4 = \mathbb{P}(\mathcal{E})$ and that we have we have identified $\mathcal{E}$ in Theorem 1.2.24.

Lemma 1.3.11. We have $\dim(B_4) = \binom{n+2}{2} - 2$. The Chow ring of B_4 is

$$\mathrm{CH}^*(B_4) = \mathrm{CH}^*(B_3)[z]/(p_4(z, e, l, m)),$$
$$p_4(z, e, l, m) = \sum_{i=0}^{\binom{n}{2}} (-z)^{\binom{n}{2}-i} \cdot c_i(\mathcal{E}),$$

where $z := c_1(\mathcal{O}_{B_4}(-1))$. Moreover, we have $\mathrm{CH}_0(B_4) = \mathbb{Z} \cdot z^{\binom{n}{2}-1} l^n m^n$ and $\int_{B_4} (-z)^{\binom{n}{2}-1} l^n m^n = 1$.

Proof. This is similar to the proof of Lemma 1.3.8. □

Proposition 1.3.12. We have $c(\mathcal{E}) = q^*(c(TV_1^Q))/c(q^*(T\tilde{R}_2))$. In here,

$$\begin{aligned} q^*(c(TV_1^Q)) = {} & (1 + l + m)^{\binom{n+2}{2}} \\ & - \sum_{j=0}^{n} \sum_{k=0}^{\binom{n+1}{2}-j} \left(\binom{\binom{n+1}{2} - j}{k} - \binom{\binom{n+1}{2} - j}{k+1} \right) (-2)^{k+1} \alpha_j l^j e^{k+1}, \end{aligned}$$

where

$$\alpha_j := \sum_{i=0}^{j} \binom{n+1}{i} \binom{\binom{n+1}{2}}{j-i} (-1)^i 2^{j-i}.$$

Moreover, $c(q^*(T\tilde{R}_2))$ can be computed from the short exact sequence

$$0 \to TB_3 \to q^*(T\tilde{R}_2) \to i_*(\mathcal{O}_e)(2e) \to 0,$$

where $i : e \hookrightarrow B_3$ is the inclusion. Here,

$$c(i_*(\mathcal{O}_e)(2e)) = \frac{1+2e}{1+e} = 1 + \sum_{i=1}^{2n} (-1)^{i-1} e^i.$$

Proof. This follows from Theorem 1.2.24 and [Ful98, Theorem 15.4] using $q^*(E_1^Q) = 2e$ and $q^*(H^Q) = l + m$, where E_1^Q and H^Q are the class of the exceptional divisor and the pullback of the hyperplane class to V_1^Q, respectively, in $\mathrm{CH}^1(V_1^Q)$. The short exact sequence follows from computing the sheaf of relative Kähler differentials Ω_q of the quotient map $q : B_3 \to \tilde{R}_2$ sitting inside the first fundamental sequence for sheaves of Kähler differentials:

$$0 \to q^*(\Omega_{\tilde{R}_2}) \to \Omega_{B_3} \to \Omega_q \to 0.$$

The exactness of this sequence on the left can be checked on stalks and can hence be verified by a computation in affine open charts. In fact, using the PGL-action, a single chart intersecting e suffices. This computation also shows $\Omega_q = i_* i^* \Omega_q$ and that $i^*\Omega_q$ is a line bundle on e. Moreover, we have the canonical splitting $\Omega_{B_3}|_e = \Omega_e \oplus \mathcal{O}_e(1)$. Since $q|_e$ is an isomorphism onto its image, the image of the map $q^*(\Omega_{\tilde{R}_2})|_e \to \Omega_{B_3}|_e$ necessarily contains Ω_e, so there is a surjection of sheaves $\mathcal{O}_e(1) \to i^*(\Omega_q)$ which is necessarily an isomorphism. Hence, $\Omega_q \cong i_*(\mathcal{O}_e(1)) \cong i_*(\mathcal{O}_e)(-e)$. Another reference is the proof of [SV16, Proposition 11.4].
Now, taking $\mathcal{H}om_{B_3}(-, \mathcal{O}_{B_3})$ of the first fundamental sequence above, the resulting long exact $\mathcal{E}xt$-sequence contains the short exact sequence

$$0 \to TB_3 \to q^*(T\tilde{R}_2) \to \mathcal{E}xt^1(\Omega_q, \mathcal{O}_{B_3}) \to 0.$$

This is because $\mathcal{H}om(\Omega_q, \mathcal{O}_{B_3}) = 0$ as Ω_q is torsion and $\mathcal{E}xt^1(\Omega_{B_3}, \mathcal{O}_{B_3}) = 0$ since Ω_{B_3} is locally free. The term $\mathcal{E}xt^1(\Omega_q, \mathcal{O}_{B_3})$ can be computed to be $i_*(\mathcal{O}_e)(2e)$ using $\Omega_q \cong i_*(\mathcal{O}_e)(-e)$ by tensoring the resolution $0 \to \mathcal{O}_{B_3}(-e) \to \mathcal{O}_{B_3} \to i_*\mathcal{O}_e \to 0$ by $\mathcal{O}_{B_3}(-e)$. ☐

Remark 1.3.13. It is possible, in principle, to write down explicit formulas for $c(\mathcal{E})$ and $c(N_{B_4}V_4)$ for arbitrary n using Lemma 1.2.20. For $n = 3$ we get

$$\begin{aligned} c(\mathcal{E}) = & -22(l^3+m^3) - 2(l^2m + lm^2) + 16em^2 \\ & + 30e^2 + 15(l^2+m^2) - 80em + 26lm \\ & - 9e + 6(l+m) + 1 \end{aligned}$$

and

$$\begin{aligned}
c(N_{B_4}V_4) = {} & 33600z^2l^3m^3 + 22400z^2l^3m^2 - 4200z^2e^2m^3 + 28700z^2l^2m^3 \\
& - 268100zl^3m^3 + 10080z^2l^3m - 840z^2e^2m^2 + 8400z^2l^2m^2 - 97720zl^3m^2 \\
& + 700z^2em^3 - 3360ze^2m^3 - 25900z^2lm^3 + 51520zl^2m^3 + 77560l^3m^3 \\
& + 2730z^2l^3 + 1890z^2l^2m - 26040zl^3m - 420z^2em^2 - 2100ze^2m^2 \\
& - 7910z^2lm^2 + 20440zl^2m^2 + 12600l^3m^2 + 5530z^2m^3 + 15820zem^3 \\
& + 15540e^2m^3 + 19180zlm^3 - 43540l^2m^3 - 270z^2e^2 - 90z^2l^2 - 3060zl^3 \\
& + 720z^2em - 1980z^2lm + 5820zl^2m - 480l^3m + 1550z^2m^2 + 840zem^2 \\
& + 1260e^2m^2 + 2940zlm^2 - 9380l^2m^2 - 8940zm^3 - 7980em^3 + 2300lm^3 \\
& - 135z^2e - 585ze^2 - 180z^2l + 810zl^2 - 1470l^3 + 360z^2m + 1560zem \\
& + 60zlm - 2010l^2m - 1780zm^2 + 720em^2 + 230lm^2 + 1910m^3 + 35z^2 \\
& + 240ze + 405e^2 - 80zl + 45l^2 - 240zm - 1080em + 270lm + 415m^2 \\
& - 9z - 30e + 10l + 30m + 1.
\end{aligned}$$

Lemma 1.3.14. We have $\pi_4^*(P_3) = P_4$ and $\pi_4^*(L_3) = L_4 + E_4$. The full intersection classes of point and line conditions with respect to B_3 are:

$$B_4 \circ P_4 = l + 2m, \qquad\qquad B_4 \circ L_4 = 1 + 4l + 8m - 6e - z.$$

Proof. This follows from the fact that B_4 is not contained in P_4 since B_3 is not contained in P_3 and that L_3 is generically smooth along B_3. The latter also implies that L_4 is generically smooth along B_4. Moreover, $(B_4 \hookrightarrow V_4)^*(E_4) = z$. $\square$

1.4 Some Characteristic Numbers for Cubics

Below we give the characteristic numbers with respect to point and line conditions for smooth cubic hypersurfaces of dimensions 2, 3, 4 and 5 explicitly. Thanks to Theorem 1.1.6 and [Alu90, Theorem II] we obtain the following formula for the number $\mathcal{N}_n(n_p, n_\ell)$ of smooth cubic hypersurfaces in $\mathbb{P}^n$ containing n_p general points and tangent to n_ℓ general lines with $n_p + n_\ell = \binom{n+3}{3} - 1$:

$$\mathcal{N}_n(n_p, n_\ell) = 4^{n_\ell} - \sum_{i=0}^{4} \int_{B_i} \frac{(B_i \circ P_i)^{n_p}(B_i \circ L_i)^{n_\ell}}{c(N_{B_i}V_i)}. \tag{1.7}$$

Theorem 1.4.1. We have the following characteristic numbers.

$$\mathcal{N}_3(n_p, n_\ell) = \begin{cases} 4^{19-n_p}, & n_p \geq 7, \\ 67107584, & n_p = 6, \\ 268391296, & n_p = 5, \\ 1072926016, & n_p = 4, \\ 4266198896, & n_p = 3, \\ 16615227040, & n_p = 2, \\ 61810371328, & n_p = 1, \\ 213642327616, & n_p = 0, \end{cases}$$

$$
\mathcal{N}_4(n_p, n_\ell) = \begin{cases} 4^{34-n_p}, & n_p \geq 9, \\ 4503599627298816, & n_p = 8, \\ 18014398504518144, & n_p = 7, \\ 72057593859735296, & n_p = 6, \\ 288230371735956864, & n_p = 5, \\ 1152916410922381056, & n_p = 4, \\ 4611439266223370624, & n_p = 3, \\ 18440552943629755776, & n_p = 2, \\ 73680611961739695744, & n_p = 1, \\ 293743613260159055616, & n_p = 0, \end{cases}
$$

$$
\mathcal{N}_5(n_p, n_\ell) = \begin{cases} 4^{55-n_p}, & n_p \geq 11, \\ 1237940039285380274890866688, & n_p = 10, \\ 4951760157141521098595270656, & n_p = 9, \\ 19807040628566084336374644736, & n_p = 8, \\ 79228162514264334978909921280, & n_p = 7, \\ 316912650057057265979098451968, & n_p = 6, \\ 1267650600108422245198531502080, & n_p = 5, \\ 5070602390849253903742648557568, & n_p = 4, \\ 20282409173750701918237560930304, & n_p = 3, \\ 81129625967826286127033390178304, & n_p = 2, \\ 324518279007939337388899670437888, & n_p = 1, \\ 1298069289965922934250958638546944, & n_p = 0, \end{cases}
$$

$$
\mathcal{N}_6(n_p, n_\ell) = \begin{cases} 4^{83-n_p}, & n_p \geq 13, \\ 5575186299632655785383929568162088438726656, & n_p = 12, \\ 22300745198530623141535718272647986547785728, & n_p = 11, \\ 89202980794122492566142873090556667222818816, & n_p = 10, \\ 356811923176489970264571492359935951598780416, & n_p = 9, \\ 1427247692705959881058285969326673318416744448, & n_p = 8, \\ 5708990770823839524233143872781898515562758144, & n_p = 7, \\ 22835963083295357547450367470452446432700268544, & n_p = 6, \\ 91343852333181359672918122382354358401157431296, & n_p = 5, \\ 365375409332720862787005988478342433799333150720, & n_p = 4, \\ 1461501637330683302436966294009082115493377474560, & n_p = 3, \\ 5846006549316096042304679826939505378910363975680, & n_p = 2, \\ 23384026197086430915159132013388781299969309016064, & n_p = 1, \\ 93536104784328168306504740799213619812755181666304, & n_p = 0. \end{cases}
$$

Proof. The proof is now merely computational. We only give the individual correction terms for $n = 3$:

$$\int_{B_0} \frac{(3h)^{n_p}(2+12h)^{n_\ell}(1+h)^4}{(1+3h)^{20}} = \begin{cases} 1769472 & n_p = 3 \\ 54263808 & n_p = 2 \\ 877658112 & n_p = 1 \\ 9948889088 & n_p = 0, \end{cases}$$

$$\int_{B_1} \frac{(3h)^{n_p}(1+12h-2\epsilon)^{n_\ell}(1+2h-\epsilon)^{10}}{(1+\epsilon)(1+3h-\epsilon)^{20}} = \begin{cases} 434889 & n_p = 3 \\ 13011156 & n_p = 2 \\ 203305944 & n_p = 1 \\ 2199770536 & n_p = 0, \end{cases}$$

$$\int_{B_2} \frac{(3h)^{n_p}(1+12h-2\epsilon-\phi)^{n_\ell}}{(1+\phi)(1+\epsilon-\phi)} = \begin{cases} 17951031 & n_p = 3 \\ 443328300 & n_p = 2 \\ 5677810728 & n_p = 1 \\ 49885157976 & n_p = 0, \end{cases}$$

$$\int_{B_3} \frac{(l+2m)^{n_p}(1+4l+8m-6e)^{n_\ell}}{c(N_{B_3}V_3)} = \begin{cases} 160 & n_p = 6 \\ 6240 & n_p = 5 \\ 130224 & n_p = 4 \\ 1426504 & n_p = 3 \\ 8284040 & n_p = 2 \\ 7701512 & n_p = 1 \\ -337368096 & n_p = 0, \end{cases}$$

$$\int_{B_4} \frac{(l+2m)^{n_p}(1+4l+8m-6e-z)^{n_\ell}}{c(N_{B_4}V_4)} = \begin{cases} 1120 & n_p = 6 \\ 37920 & n_p = 5 \\ 685584 & n_p = 4 \\ 7186504 & n_p = 3 \\ 45754840 & n_p = 2 \\ 142629112 & n_p = 1 \\ -460870176 & n_p = 0. \end{cases} \qquad \square$$

Remark 1.4.2. In principle, the underlying `Macaulay2` code works for every n but the characteristic numbers grow very fast as can already be seen in Theorem 1.4.1. As n grows, it seems that the non-trivial characteristic numbers, i.e., those in the range $0 \leq n_p \leq 2n$, get very close to the Bézout bound 4^{n_ℓ}. In other words, in (1.7) the correction term

$$f(n_p, n) := \sum_{i=0}^{4} \int_{B_i} \frac{(B_i \circ P_i)^{n_p}(B_i \circ L_i)^{n_\ell}}{c(N_{B_i}V_i)}$$

seems to lose asymptotically against 4^{n_ℓ}. It is an interesting problem to describe $f(n_p, n)$, for every $n_p \in [0, 2n]$, as a function of n. We do not expect $f(n_p, -)$ to be a polynomial but it is possible, at least theoretically, to express $f(n_p, -)$ in terms of "basic" functions involving polynomials in n, exponential functions and binomial coefficients of which both entries possibly depend on n. Proposition 1.4.3 below is slightly different in that n_p depends on n as well. Another general observation is that for $n_p > n$ the correction terms coming from B_0, B_1 and B_2 in (1.7) do not contribute because $(B_i \circ P_i)^{n_p} = (3h)^{n_p} = 0$ for $i = 0, 1, 2$. Moreover, for $n_p > n$, expanding the correction terms for B_3 and B_4, no term involving e will contribute to the degree because $e(l + 2m)^{n_p} = 0$ which follows from $l^{n+1} = m^{n+1} = 0$ and $el = em$. These relations hold in both $\mathrm{CH}^*(B_3)$ and $\mathrm{CH}^*(B_4)$. Therefore, in fact, the characteristic numbers in the range $n + 1 \leq n_p \leq 2n$ can be computed without identifiying the vector bundle $\mathcal{E}$ globally, i.e., without Theorem 1.2.24.

Proposition 1.4.3. Let $n \geq 2$. The function $f(2n, n)$ is precisely the degree of the *scheme-theoretic* intersection of all line conditions in V_0. It takes the simple form

$$f(2n, n) = 2^{\binom{n+1}{2}} \binom{2n}{n}.$$

Proof. We first observe that in both $\mathrm{CH}^*(B_3)$ and $\mathrm{CH}^*(B_4)$ we have $(l + 2m)^{2n} = 2^n \binom{2n}{n} l^n m^n$. Note that multiplying $l^n m^n$ by any homogeneous element of positive degree in $\mathrm{CH}^*(B_3)$ gives zero since $\dim(B_3) = 2n$. The correction term coming from B_3 is therefore precisely $2^n \binom{2n}{n}$. The correction term coming from B_4 is

$$\begin{aligned} &2^n \binom{2n}{n} \int_{B_4} l^n m^n (1 + 4l + 8m - 6e - z)^{\binom{n+3}{3} - 1 - 2n} s(N_{B_4} V_4) \\ &= 2^n \binom{2n}{n} \int_{B_4} l^n m^n (1 - z)^{\binom{n+3}{3} - 1 - 2n} s(N_{B_4} V_4)|_{l=m=e=0}. \end{aligned}$$

We have $s(N_{B_4} V_4) = s(N_{B_4} E_4) \cdot s(N_{E_4} V_4|_{B_4}) = s(N_{B_4} E_4)/(1 + z)$. Moreover, by Lemma 1.2.20 we have $s(N_{B_4} E_4)|_{l=m=e=0} = 1/(1 - z)^{\binom{n+2}{3}}$. We therefore continue with

$$\begin{aligned} &2^n \binom{2n}{n} \int_{B_4} l^n m^n (1 - z)^{\binom{n}{2}} \left(\sum_{j \geq 0} (-z)^j \right) \\ &= 2^n \binom{2n}{n} \int_{B_3} l^n m^n (\pi_4|_{B_4})_* \left(\sum_{k=0}^{\binom{n}{2}-1} \sum_{j \geq 0} \binom{\binom{n}{2}}{k} (-z)^{j+k} \right) \\ &= 2^n \binom{2n}{n} (2^{\binom{n}{2}} - 1) \int_{B_3} l^n m^n = 2^{\binom{n+1}{2}} \binom{2n}{n} - 2^n \binom{2n}{n}, \end{aligned}$$

proving the claim. □

In contrast, the degree of S_0 can be computed to be $2^n \binom{2n}{n}$, as follows. We have the commutative triangle

$$\begin{tikzcd}
\mathrm{CH}_*(\mathbb{P}^n \times \mathbb{P}^n) \arrow[r, "\phi_{0,*}"] \arrow[dr, "\deg"'] & \mathrm{CH}_*(V_0) \arrow[d, "\deg"] \\
 & \mathrm{CH}_*(\mathrm{Spec}(K)) = \mathbb{Z}.
\end{tikzcd}$$

Since $\phi_{0,*}([\mathbb{P}^n \times \mathbb{P}^n]) = [S_0]$ and $\phi_0^*([H]) = l + 2m$, we get

$$\deg(S_0) = \deg([S_0] \cdot [H]^{2n}) = \deg \phi_{0,*}((l+2m)^{2n}) = \deg((l+2m)^{2n}) = 2^n \binom{2n}{n},$$

where in the second equation we have used the projection formula.
Let us denote the scheme-theoretic intersection of all line conditions by $\hat{S}_0$ for the moment. Then both S_0 and $\hat{S}_0$ are PGL-invariant closed subschemes of V_0. Moreover, S_0 consists of precisely two orbits, namely B_0 and $S_0 \setminus B_0$. From this we can deduce that $\hat{S}_0$ is everywhere non-reduced for all $n \geq 2$ because otherwise $\hat{S}_0$ would have to be reduced along all of $S_0 \setminus B_0$ and hence either reduced everywhere (whence $\hat{S}_0 = S_0$) or non-reduced precisely along B_0, making B_0 an embedded component of $\hat{S}_0$. In both cases, however, the degrees of S_0 and $\hat{S}_0$ would have to agree.

Crumbs of hyperplane tangency conditions

We stick to the conventions of Section 1.1. In the case of hyperplane tangency conditions for degree d hypersurfaces, the base locus is in general hard to describe explicitly compared to the base locus of tangency with respect to lines. In fact, hyperplane conditions in $\mathbb{P}(\mathrm{Sym}^d(W^*))$ intersect in the locus B_0^H of hypersurfaces with positive-dimensional singular locus

$$B_0^H(d,n) \;=\; \{[h] \in \mathbb{P}(\mathrm{Sym}^d(W^*)),\ |\dim \mathrm{Sing}(\mathcal{V}(h)) \geq 1\},$$

as a Bertini-type argument shows. This set has been studied in [Sla15, Tse20]. Building on [BS07], it is shown in [LPS11, Lemma 2.4] that an integral cubic hypersurface in $\mathbb{P}^n$ which is not normal, i.e., whose singular locus has dimension $n-2$, is necessarily singular exactly along a *linear* subspace of $\mathbb{P}^n$ of dimension $n-2$. For $n = 3$, this means that the cubic form defining a cubic surface with singular locus of dimension ≥ 1 is either reducible or the singular locus is precisely a line in $\mathbb{P}^3$. This fact is also known classically [Seg42, p. 144]. It is not hard to see that the set of reducible cubic forms has dimension 12. The dimension of the other set is 13 because it is birational to

$$\{(\ell, f) : (\mathrm{grad} f)|_\ell \equiv 0\} \;\subseteq\; \mathrm{Gr}(2,4) \times \mathbb{P}(\mathrm{Sym}^3(W^*)),$$

a $\mathbb{P}^9$-bundle over $\mathrm{Gr}(2,4)$. We deduce that Slavov's theorem [Sla15, Theorem 1.1] is also true for cubic surfaces. The work [Suk20] lists the finitely many PGL-orbits of cubic surfaces with positive-dimensional singular locus and studies containments among their closures.

Remark 1.4.4. The variety of all degree d hypersurfaces in $\mathbb{P}(W)$ which are tangent to H is a hypersurface in $\mathbb{P}(\mathrm{Sym}^d(W^*))$ of degree $n(d-1)^{n-1}$. Indeed, let $H = \mathcal{V}(x_0)$. The set of hypersurfaces in $\mathbb{P}(\mathrm{Sym}^d(W^*))$ which are tangent to H is the vanishing set of the resultant of the polynomials $\partial_{x_i} f(0, x_1, \ldots, x_n)$, $i = 1, \ldots, n$, i.e., the discriminant of $f(0, x_1, \ldots, x_n)$, where f is the generic degree d polynomial. This discriminant has degree $n(d-1)^{n-1}$.

Proposition 1.4.5 ([Tse20] + classical results)**.** Let $\mathrm{char}(K) = 0$. We consider the family of smooth degree d hypersurfaces in $\mathbb{P}^n$ with $n \geq 2$. Assume one of the following holds:

- $d = 2$, $d = 5$ or $d \geq 7$,
- $n = 2$ and $d \geq 2$ arbitrary,
- $d = n = 3$.

Then, if $n_H < n(d-2)+3$, the number $\mathcal{N}^H(n_p, n_H)$ of smooth degree d hypersurfaces in $\mathbb{P}^n$ tangent to n_H general hyperplanes and passing through $n_p = \binom{n+d}{d} - 1 - n_H$ general points equals

$$\mathcal{N}^H(n_p, n_H) = (n(d-1)^{n-1})^{n_H}.$$

Proof. If n_H is strictly less than the codimension of $B_0^H(d, n)$, the claim follows from Bézout's theorem. Now, the codimension of $B_0^H(d, n)$ is known for $d = 5$ and $d \geq 7$ (and arbitrary n) by [Tse20, Theorem 1.6] and equals

$$\mathrm{codim} B_0^H(d, n) = n(d-2) + 3.$$

Moreover, this codimension holds true also for quadric hypersurfaces as well as in the case $d = n = 3$ of cubic surfaces by the above discussion. For $n = 2$, moreover, the base locus $B_0^H(d, 2)$ is the set of non-reduced degree d ternary forms which has codimension $2d - 1 = 2(d-2) + 3$. $\square$

2 Ideals of Submaximal Minors of Sparse Symmetric Matrices

For questions of authorship, please refer to pages IVf.

This chapter is based on the accepted version of [DK23].

2.1 Sparse Determinantal Ideals

The ideal of maximal minors of a generic $m \times n$ matrix is one of the most studied objects in combinatorial commutative algebra due to its geometric significance and its rich combinatorial structure. If one allows the matrix to be *sparse*, i.e. to have some entries replaced by zero, some important tools of study disappear, most notably since the action of the general linear group does not preserve the zero pattern. Nonetheless, many results have been obtained in the sparse case, for example by Giusti and Merle [GM82], Boocher [Boo12] and Conca and Welker [CW19], and in part even for much weaker assumptions on the entries of the matrix, see for example [CDNG15, CDNG17, CDNG18, CDNG20, CDNG22] for the theory of Cartwright–Sturmfels ideals and universal Gröbner bases, [Eis88] for the case of *arbitrary* linear sections of determinantal varieties of low codimension, and [MR07, MR08] for general background on determinantal schemes including homological results. The symmetric case, i.e., the ideal of minors of fixed size of a sparse generic *symmetric* matrix, is much less understood. In [CW19], Conca and Welker study geometric and arithmetic properties like primality, reducedness, codimension and the complete intersection property of these ideals and of other kinds of sparse determinantal ideals. Homological invariants, on the other hand, such as the graded Betti numbers, projective dimension and regularity are in general still unknown in the sparse symmetric case (see, however, the related [MR07, Section 3]). Our main result fills this gap in the case of submaximal minors of a sparse generic symmetric matrix, using ideas of Boocher [Boo12, Boo13] together with a foundational result by Józefiak [Józ78] and combining these with new Gröbner basis results.

More precisely, let K be a field and $R = K[x_{ij} : 1 \leq i \leq j \leq n]$ the polynomial ring, $n \geq 2$. Let $X = (x_{ij})$ be the generic symmetric $n \times n$ matrix, i.e. $x_{ij} := x_{ji}$ for $i > j$. Let G be an undirected simple graph with vertex set $[n] := \{1, 2, \dots, n\}$ and let Z be the set of all off-diagonal variables corresponding to the non-edges of G. We define X_G to be the matrix obtained from X by substituting zeros for all variables in Z, and write $I_{n-1}(X)$ and $I_{n-1}(X_G)$ for the ideals of R generated by all $(n-1)$-minors of X and X_G, respectively.

If $C_1, \dots, C_r$ is the partition of $[n]$ where the C_i are the vertex sets of the connected components of G, we define

$$D_G := \sum_{1 \le s < t \le r} |C_s| \cdot |C_t|.$$

Then $0 \le D_G \le \binom{n}{2}$, and the lower bound is attained if and only if G is connected while the upper bound is attained if and only if G has no edges at all. The following is the main result of this chapter.

Theorem 2.1.1. The minimal graded free resolution of $R/I_{n-1}(X_G)$ is obtained from the one of $R/I_{n-1}(X)$ by substituting zeros for all variables in Z and "pruning" the resulting complex. The graded Betti numbers are

$$\begin{aligned} \beta_{1,n-1}(R/I_{n-1}(X_G)) &= \binom{n+1}{2} - D_G, \\ \beta_{2,n}(R/I_{n-1}(X_G)) &= n^2 - 1 - 2D_G, \\ \beta_{3,n+1}(R/I_{n-1}(X_G)) &= \binom{n}{2} - D_G. \end{aligned}$$

All other graded Betti numbers (apart from $\beta_{0,0} = 1$) are zero. Therefore, $I_{n-1}(X_G)$ has a linear resolution with regularity $\operatorname{reg}(I_{n-1}(X_G)) = n - 1$. For the projective dimension, we deduce $\operatorname{pdim}(R/I_{n-1}(X_G)) = 3$ except if G has no edges at all, in which case $\operatorname{pdim}(R/I_{n-1}(X_G)) = 2$. The quotient ring $R/I_{n-1}(X_G)$ is reduced, and it is Cohen–Macaulay if and only if G is either connected or has no edges at all. Finally,

$$\operatorname{ht}(I_{n-1}(X_G)) = \begin{cases} 3 & \text{if } G \text{ is connected,} \\ 2 & \text{otherwise.} \end{cases}$$

The proof of Theorem 2.1.1 is contained in Sections 2.2 and 2.3. The pruning procedure we referred to is the same as the one used by Boocher in [Boo12, Boo13]: First, we set to zero all variables in Z in all three matrices of the minimal free resolution of $R/I_{n-1}(X)$. Next, all zero columns of the first matrix of the resolution together with the corresponding rows of the second matrix are erased. Then, all emerging zero columns of the cropped second matrix together with the corresponding rows of the third matrix are erased. Finally, we also delete all emerging zero columns of the cropped third matrix. Theorem 2.1.1 implies that the resulting complex is again exact and that, if G is connected, no pruning occurs at all.
If one is only interested in the case where G is connected, then the arguments of Section 2.3 can be avoided. This is because it follows from Section 2.2 that $I_{n-1}(X_G)$ has grade 3. Then Józefiak's result [Józ78, Theorem 3.1] yields that the minimal free resolution of $I_{n-1}(X)$ stays exact after substituting zeros for all variables in Z.

Example 2.1.2. If diagonal variables are set to zero as well, a pruning procedure similar to the above cannot work in general. Explicitly, let $n = 5$ and $Z = \{x_{11}, x_{22}, x_{13}, x_{14}, x_{23}, x_{24}\}$. Let X' be the matrix obtained from X by setting to zero all variables from Z. A computation in `Macaulay2` for $K = \mathbb{Q}$ shows that

$\operatorname{pdim}(R/I_{n-1}(X')) = 4$ while $\operatorname{pdim}(R/I_{n-1}(X)) = 3$. The ideal $I_{n-1}(X')$ is not radical and has minimal primes of heights 2 and 3, so $R/I_{n-1}(X')$ is not Cohen–Macaulay. This contrasts with Theorem 2.1.1.

Let us briefly discuss connections with algebraic statistics since [DK23] was originally inspired by [BKKR23], see Chapter 3. An *undirected Gaussian graphical model* $\mathcal{M}(G)$ associated to G is the set of real symmetric positive definite $n \times n$ matrices Σ (i.e., *covariance matrices*) such that $(\Sigma^{-1})_{ij} = 0$ whenever $i \neq j$ and $ij \notin G$. In other words, $\mathcal{M}(G)$ is the image under the matrix inversion map of a coordinate linear subspace of the cone of positive definite $n \times n$ matrices. Extending the ground field to be $\mathbb{C}$, one can view the matrix inversion map as a rational map

$$\operatorname{inv}\colon \mathbb{P}(\operatorname{Sym}^2(\mathbb{C}^n)) \dashrightarrow \mathbb{P}(\operatorname{Sym}^2(\mathbb{C}^n)),\ [A] \mapsto [A^{-1}].$$

In geometric terms, this map associates to a smooth quadric its dual quadric (which is again smooth). The indeterminacy locus of inv is precisely defined by $I_{n-1}(X)$. This can be seen by observing that the cofactor matrix of A, in case A is invertible, is a scalar multiple of A^{-1}. Restricting the domain of inv to the coordinate linear subspace L_G of G-sparse symmetric matrices, i.e., those A with $A_{ij} = 0$ whenever $i \neq j$ and $ij \notin G$, the indeterminacy locus becomes the vanishing subscheme $V(I_{n-1}(X_G))$. In [DMV21], the degree of the projective variety $\overline{\operatorname{inv}(L_G)}$ is computed in the case where $G = C_n$ is the n-cycle in order to show that the degree of this projective variety can be (much) larger than the conjectural maximum likelihood degree of $\mathcal{M}(C_n)$ [DSS09, Section 7.4] even though for general linear concentration models they agree [SU10, Theorem 1]. The Segre class formula presented in [AGK+21, Theorem 4.2] gives the precise relationship between the degree of $\overline{\operatorname{inv}(L_G)}$ and the maximum likelihood degree of $\mathcal{M}(G)$.
Our Theorem 2.1.1 also sheds some more light on the varieties $\overline{\operatorname{inv}(L_G)}$ for *arbitrary* graphs G. However, the homological information we have obtained is not, to our knowledge, sufficient for computing the degree of $\overline{\operatorname{inv}(L_G)}$ or the maximum likelihood degree of $\mathcal{M}(G)$. The degree of $\overline{\operatorname{inv}(L_G)}$ is equal to the number of smooth G-sparse quadrics in $\mathbb{P}^{n-1}$ tangent to $\dim(L_G)$ general hyperplanes, giving rise to an excess intersection problem. Theorem 2.1.1 provides answers to the simplest non-trivial analogues of this question, namely for 2 hyperplanes and, in case G is connected, also for 3 hyperplanes, see Section 2.4 for details. For more on Gaussian graphical models and generalizations thereof, see Chapter 3.

2.2 Gröbner Bases

We keep the notation of the introduction. For general background on initial ideals and Gröbner bases we refer to [BC03]. For the classical determinantal ideals $I_k(X)$, $1 \leq k \leq n$, of a generic symmetric matrix X (without zeros) Conca proved that the k-minors form a Gröbner basis with respect to diagonal term orders [Con94b], for example the lexicographic term order reading the upper triangle of X row by row. For our purposes, however, it is important to consider certain non-diagonal term orders associated to graphs. Given a graph G on the vertex set $[n]$, let w_G be the weight vector for the polynomial ring $R = K[x_{ij} : 1 \leq i \leq j \leq n]$ which assigns the following weights to the variables:

$$\begin{aligned} w_G(x_{ii}) &= 2 \quad \text{for all } i = 1, \ldots, n, \\ w_G(x_{ij}) &= 2 \quad \text{for all } ij \in G, \\ w_G(x_{ij}) &= 1 \quad \text{for all } ij \notin G. \end{aligned}$$

In other words, the variables in Z have w_G-weight 1 while all others have w_G-weight 2. This is analogous to the weights used by Boocher [Boo12, Boo13]. Moreover, by w_{diag} we denote the weight vector which assigns weight 2 to the diagonal variables and weight 1 to the off-diagonal variables (which is the special case of w_G where G has no edges). Let $T \subseteq G$ be a spanning forest, i.e., the union of one spanning tree for each connected component of G. The main goal of this section is to show that the $\binom{n+1}{2}$ standard $(n-1)$-minors generating $I_{n-1}(X)$ form a Gröbner basis with respect to the weight order given by $<_{T,G} \coloneqq <_{w_{\text{diag}}} \circ <_{w_T} \circ <_{w_G}$ and that $\text{in}_{<_{T,G}}(I_{n-1}(X))$ is a square-free monomial ideal. Here, the composition of the weight orders is to be read from right to left, i.e., two monomials are first compared with respect to their w_G-degrees, then, in case of equal w_G-degrees, with respect to their w_T-degrees, and finally, in case of equal w_T-degrees, with respect to w_{diag}. Of course, this is not a monomial order since it does not yield a total ordering of the set of all monomials in R. Nonetheless, the ideals of leading terms of $I_{n-1}(X)$ and $I_{n-1}(X_G)$ with respect to $<_{T,G}$ are indeed monomial, as we shall see.
We define $I_T \subseteq R$ to be the square-free monomial ideal generated by the following monomials:

$$\begin{aligned} &\frac{x_{11} \cdots x_{nn}}{x_{ii}} && \text{for all } i = 1, \ldots, n, \\ &\frac{x_{11} \cdots x_{nn}}{x_{kk} x_{ll}} x_{kl} && \text{for } 1 \leq k < l \leq n \text{ if in } T \text{ there is no path between } k \text{ and } l, \\ &\left(\prod_{i \in [n] \setminus V(p)} x_{ii} \right) \cdot x_p && \text{for all paths } p \text{ in } T. \end{aligned}$$

We use the word path in its strictest sense, i.e., no vertex appears twice in a path and it contains at least one edge. By $V(p)$ we denote the vertex set of the path p and by $x_p \coloneqq \prod_{ij \in p} x_{ij}$ the product of all off-diagonal variables corresponding to the edges in p. It is easily seen that all generators are square-free monomials of degree $n-1$, none is redundant in I_T and their total number is $\binom{n+1}{2}$. The latter follows from realizing that in a tree every path is uniquely determined by its two end points. We have $I_T \subseteq \text{in}_{<_{T,G}}(I_{n-1}(X))$ since the generators of I_T are precisely the initial terms of the standard $(n-1)$-minors with respect to $<_{T,G}$. This can easily be read off the following formula.

Proposition 2.2.1 ([JW05, BKKR23])**.** For all $1 \leq k < l \leq n$ we have

$$(-1)^{k+l} \det \left((X_G)_{[n] \setminus k, [n] \setminus l} \right) = \sum_{\substack{p \text{ path in } G \\ \text{between } k \text{ and } l}} (-1)^{|V(p)|-1} \cdot \det \left((X_G)_{[n] \setminus V(p), [n] \setminus V(p)} \right) \cdot x_p.$$

Proof. This is the content of [JW05, Theorem 1] and is recalled in [BKKR23, Proposition 3.19], see Proposition 3.3.19. □

Proposition 2.2.2. For any forest T on $[n]$, the ideal I_T has linear quotients, and all graded Betti numbers of I_T and $I_{n-1}(X)$ agree. Moreover, R/I_T is Cohen–Macaulay.

Proof. We first prove that I_T has linear quotients. For this, we introduce shorthand notations: Let $m_i := \frac{x_{11}\cdots x_{nn}}{x_{ii}}$ for all $i = 1, \ldots, n$, and denote by m_p the monomial generator of I_T corresponding to the path p in T. By $\hat{m}_{ij}$ we denote the generator corresponding to the pair of vertices ij if i and j lie in different connected components of T.

We now order the generators of I_T in the following way: First, we take the n monomials m_i only containing diagonal variables, then we take all monomials m_p corresponding to paths p of length 1 (i.e., edges) in T, then all monomials m_p for paths p of length 2 in T, and so on. Finally, we take all monomials $\hat{m}_{ij}$. The order of the generators in each of these individual groups is arbitrary. The first $n-1$ colon ideals are then generated by a single diagonal variable since

$$(m_1, \ldots, m_i) : m_{i+1} = (x_{i+1,i+1})$$

for all $i = 1, \ldots, n-1$. Let $r \geq 1$ and $I := (m_1, \ldots, m_n, m_p |\ \mathrm{length}(p) < r)$. Let J be the ideal of R generated by I and any (possibly empty) collection of generators m_p for paths p of length r. Let q be another path of length r in T with endpoints k and l. Then, we claim

$$J : m_q = (x_{kk}, x_{ll}).$$

The right hand side is contained in the left hand side because the paths $q \setminus k$ and $q \setminus l$ have corresponding generators in J (even for $r = 1$). Conversely, it is enough to show that any square-free monomial multiple m of m_q contained in J is divisible by x_{kk} or x_{ll}. Indeed, if m is a multiple of some m_i, we obviously need to multiply by at least one of x_{kk} and x_{ll} since m_q is not divisible by either. If instead m is a multiple of some m_p, then the vertices corresponding to diagonal variables not dividing m must be contained in p. So if both x_{kk} and x_{ll} do not divide m, then p contains q which is impossible by definition of J.

Finally, let J be the ideal generated by $m_1, \ldots, m_n$, all the m_p and any collection of the $\hat{m}_{ij}$. Let $\hat{m}_{kl}$ be any new generator of this type. This time, it is clear that

$$J : \hat{m}_{kl} = (x_{kk}, x_{ll})$$

because every generator of I_T (and hence of J) other than $\hat{m}_{kl}$ is divisible by at least one of x_{kk} and x_{ll}. This proves that I_T has linear quotients.

From [HH11, Corollary 8.2.2], we can now even deduce the statement about graded Betti numbers. In the notation of *loc. cit.*, we have $r_1 = 0$, $r_2 = \cdots = r_n = 1$ and $r_{n+1} = \cdots = r_{\binom{n+1}{2}} = 2$. Hence, the minimal graded free resolution of I_T has the form

$$0 \to R(-(n+1))^{\binom{n}{2}} \to R(-n)^{n^2-1} \to R(-(n-1))^{\binom{n+1}{2}} \to I_T \to 0,$$

just as for $I_{n-1}(X)$ by [Józ78], see Section 2.3. This proves $\beta_{i,j}(I_T) = \beta_{i,j}(I_{n-1}(X))$ for all i, j. In particular, $\mathrm{pdim}(R/I_T) = 3$. In order to prove that R/I_T is Cohen–Macaulay, by the Auslander–Buchsbaum formula it suffices to show $\mathrm{ht}(I_T) = 3$. This can be seen directly but it also follows from $\mathrm{ht}(I_{n-1}(X)) = 3$ since $\mathrm{in}_{<_{T,G}}(I_{n-1}(X)) = I_T$ by the following corollary. □

Corollary 2.2.3. The $(n-1)$-minors of X form a Gröbner basis of $I_{n-1}(X)$ with respect to $<_{T,G}$. Equivalently, $\mathrm{in}_{<_{T,G}}(I_{n-1}(X)) = I_T$.

Proof. Since $I_T \subseteq \mathrm{in}_{<_{T,G}}(I_{n-1}(X))$, it suffices to show that the two ideals have the same Hilbert functions. But the Hilbert function is determined by the graded Betti numbers, and these coincide for both ideals by Proposition 2.2.2. □

Lemma 2.2.4. Let $J \subseteq K[x_1, \ldots, x_m]$ be an ideal and $g_1, \ldots, g_r$ a Gröbner basis of J with respect to a monomial order $<$. Let Z be a subset of the variables $x_1, \ldots, x_m$ and assume that whenever $g_i|_{Z=0} \neq 0$, we have $\mathrm{in}_<(g_i) = \mathrm{in}_<(g_i|_{Z=0})$. Then, the non-zero $g_i|_{Z=0}$ are a Gröbner basis for $J|_{Z=0}$, both as an ideal of $K[x_1, \ldots, x_m]$ and of $K[x_1, \ldots, x_m]/(Z)$, each time with respect to $<$. In other words, $\mathrm{in}_<(J|_{Z=0}) = \mathrm{in}_<(J)|_{Z=0}$, both as ideals of $K[x_1, \ldots, x_m]$ and of $K[x_1, \ldots, x_m]/(Z)$.

Proof. We first consider $J|_{Z=0}$ as an ideal of the "smaller" polynomial ring, i.e., in $K[x_1, \ldots, x_m]/(Z)$, and we show that the $g_i|_{Z=0}$ form a Gröbner basis with respect to $<$. For this, let $f \in J$. We need to prove that if $f|_{Z=0} \neq 0$, then $\mathrm{in}(f|_{Z=0})$ is divisible by some $\mathrm{in}(g_i|_{Z=0})$. We write $f = p + q$ where p is the sum of all terms of f not divisible by any variable in Z and $q = f - p$, i.e., q is the sum of all terms of f divisible by some variable in Z. Clearly, $p|_{Z=0} = p$ and $q|_{Z=0} = 0$, so $f|_{Z=0} = p$, and we assume $p \neq 0$. If the initial term of f is a summand of p, then we are done since the g_i form a Gröbner basis for J. Otherwise, the leading term m of f is from q. In this case, there is some g_i and a scalar multiple of a monomial, say n, such that $m = n \cdot \mathrm{in}(g_i)$. Since $m|_{Z=0} = 0$, at least one of n and $\mathrm{in}(g_i)$ is divisible by some variable in Z. By the assumption that $\mathrm{in}(g_i)|_{Z=0} = 0 \Rightarrow g_i|_{Z=0} = 0$, this implies that *every* term of $n \cdot \mathrm{in}(g_i)$ is divisible by some variable in Z. Consider now $f' := f - n \cdot g_i \in J$ and write $f' = p' + q'$ as before. Then, $p' = p$ and $\mathrm{in}(q') < \mathrm{in}(q)$. If now $\mathrm{in}(f')$ comes from $p' = p$, we are done. Otherwise the initial term of f' lies in q' again, and we continue in the same way. But we cannot always choose the initial term from the q-part since this would result in an infinite chain of strictly decreasing monomials with respect to the monomial order $<$, proving our claim.
Secondly, the $g_i|_{Z=0}$ even form a Gröbner basis of $J|_{Z=0}$ if the latter is considered as an ideal in $K[x_1, \ldots, x_m]$. Indeed, write J_1 for the ideal in the "smaller" polynomial ring and J_2 for the extended ideal in $K[x_1, \ldots, x_m]$. Then, with respect to the inclusion $K[x_1, \ldots, x_m]/(Z) \hookrightarrow K[x_1, \ldots, x_m]$ we have

$$J_2 \cong J_1 \otimes_{K[x_1,\ldots,x_m]/(Z)} K[x_1, \ldots, x_m].$$

Let now $f \in J_2 \setminus \{0\}$. Then, by the preceeding isomorphism, f can be written uniquely as a sum of monomials in Z with coefficients in J_1. The initial term of f with respect to $<$ is therefore the initial term of some non-zero element of J_1 multiplied by some monomial in the Z-variables. This is a multiple of the initial term of an element of J_1 and therefore divisible by some $\mathrm{in}(g_i|_{Z=0})$ by the first part of the proof. $\square$

Corollary 2.2.5. The $(n-1)$-minors of X_G form a Gröbner basis of $I_{n-1}(X_G)$ with respect to $<_{T,G}$ with square-free initial ideal $I_T|_{Z=0} \subseteq R$. In particular, $I_{n-1}(X_G)$ is always radical, and

$$\mathrm{ht}(I_{n-1}(X_G)) = \begin{cases} 3 & \text{if } G \text{ is connected,} \\ 2 & \text{otherwise.} \end{cases}$$

Proof. It only remains to prove the last claim. Since the polynomial ring R is regular, $\dim(R/J) = \dim(R) - \mathrm{ht}(J)$ for every ideal $J \subseteq R$. This proves $\mathrm{ht}(I_{n-1}(X_G)) = \mathrm{ht}(I_T|_{Z=0})$. If G is connected, T is a spanning tree with vertex set $[n]$, and so $I_T|_{Z=0} = I_T$. Let $\mathfrak{p}$ be a prime ideal above I_T. As $\mathfrak{p}$ contains the first n generators of I_T which are the n possible products of $n-1$ distinct diagonal variables, $\mathfrak{p}$ contains at least two distinct diagonal variables, say x_{ii} and x_{jj}. All generators of I_T which are not divisible by either of x_{ii} and x_{jj} are divisible by x_p, where p is the unique path in T between i and j. Hence, every prime $\mathfrak{p}$ above I_T is generated by at least three distinct variables, and some are indeed generated by *exactly* three distinct variables, so $\mathrm{ht}(I_T) = 3$. If G is not connected, let i and j be two vertices lying in different connected components. Therefore, there is no path of T involving both i and j, hence $\mathfrak{p} = (x_{ii}, x_{jj})$ is a minimal prime above $I_T|_{Z=0}$, so $\mathrm{ht}(I_T|_{Z=0}) = 2$ in this case. $\square$

2.3 Minimal Free Resolution of $I_{n-1}(X_G)$

2.3.1 The Generic Case

We first recall a foundational result by Józefiak. For this result, S denotes any noetherian commutative ring with 1 and W any symmetric $n \times n$ matrix with coefficients in S. As usual, we denote by $S^{n\times n}$ the free S-module of $n \times n$ matrices and by $\mathrm{tr}\colon S^{n\times n} \to S$ the trace map whose kernel is a free S-module of rank $n^2 - 1$. By $\mathrm{Sym}^2(S^n)$ we denote the free S-module of symmetric $n \times n$ matrices and by $A_n(S)$ that of alternating matrices, i.e., the skew-symmetric $n \times n$ matrices with zeros on the diagonal.

Theorem 2.3.1 ([Józ78, Theorem 3.1])**.** Let $W \in \mathrm{Sym}^2(S^n)$ and $Y \in \mathrm{Sym}^2(S^n)$ the cofactor matrix of W. If the grade of $I_{n-1}(W) \subseteq S$ equals 3, then the complex of free S-modules

$$L(W)\colon \quad 0 \to A_n(S) \xrightarrow{d_3} \ker(\mathrm{tr}\colon S^{n\times n} \to S) \xrightarrow{d_2} S^{n\times n}/A_n(S)$$
$$\xrightarrow{d_1} S \to S/I_{n-1}(W) \to 0$$

is exact and provides a free resolution of $S/I_{n-1}(W)$. Here,

$$\begin{aligned} d_1(M \bmod A_n(S)) &:= \mathrm{tr}(YM), \\ d_2(N) &:= WN \bmod A_n(S), \\ d_3(A) &:= AW. \end{aligned}$$

We observe $S^{n\times n}/A_n(S) \cong S^{\binom{n+1}{2}} \cong \mathrm{Sym}^2(S^n)$, where we interpret $S^{\binom{n+1}{2}}$ as the S-module of upper triangular matrices.
Let us return to the case where $S = R$ is the polynomial ring. If all entries of W are homogeneous of the same positive degree and $\mathrm{grade}(I_{n-1}(W)) = 3$, then Józefiak's result shows that $L(W)$ is the minimal graded free resolution of $R/I_{n-1}(W)$. If all entries of W are of degree 1, then all entries of Y are of degree $n-1$, and hence $L(W)$ has the shape

$$0 \to R(-(n+1))^{\binom{n}{2}} \to R(-n)^{n^2-1} \to R(-(n-1))^{\binom{n+1}{2}} \to R \to R/I_{n-1}(X) \to 0.$$

In particular, this is true if W is the sparse generic symmetric matrix X_G where G is *connected on* $[n]$ by Corollary 2.2.5, using that grade and height agree for ideals in a regular ring.

2.3.2 Matrices Representing the d_i

We choose bases of the free R-modules in Józefiak's minimal graded free resolution of $R/I_{n-1}(X)$. For $R^{n\times n}/A_n(R) = R(-(n-1))^{\binom{n+1}{2}}$ we choose the graded basis E_{ij} for $i \leq j$, where E_{ij} is the matrix with a 1 in position (i,j) and zeros everywhere else, with the following order:

$$E_{11}, E_{22}, \ldots, E_{nn}, E_{12}, \ldots, E_{1,n}, \ldots, E_{n-1,n}.$$

For $\ker(\mathrm{tr} : R^{n\times n} \to R) \cong R(-n)^{n^2-1}$ we take the graded basis

$$E_{22} - E_{11}, E_{33} - E_{11}, \ldots, E_{nn} - E_{11}, E_{12}, E_{21}, \ldots, E_{1,n}, E_{n,1}, \ldots, E_{n-1,n}, E_{n,n-1}.$$

Finally, for $A_n(R) \cong R(-(n+1))^{\binom{n}{2}}$ we take the graded basis

$$E_{12} - E_{21}, \ldots, E_{1,n} - E_{n,1}, \ldots, E_{n-1,n} - E_{n,n-1}.$$

The matrix of d_1, with respect to these bases, is simply the row vector whose entries are the cofactors of X, ordered in the same way as the basis of $R^{n\times n}/A_n(R)$, i.e., the principal minors come first and then all the others with the appropriate signs in the usual lexicographic order.
For d_2 we have

$$d_2(E_{ii} - E_{11}) = \sum_{\substack{k=1\\k\neq i}}^{n} (-x_{1k})E_{1k} + \sum_{k=2}^{i} x_{ki}E_{ki} + \sum_{k=i+1}^{n} x_{ik}E_{ik}$$

for all $i > 1$, and

$$d_2(E_{ij}) = \sum_{k=1}^{j} x_{ki}E_{kj} + \sum_{k=j+1}^{n} x_{ki}E_{jk}.$$

for all $i \neq j$. The matrix of d_2 can hence be written as a block matrix

$$[d_2] = \left(\begin{array}{cccccc|c} -x_{11} & -x_{11} & -x_{11} & \dots & -x_{11} & & \\ x_{22} & 0 & 0 & \dots & 0 & & \\ 0 & x_{33} & 0 & \dots & 0 & & \\ 0 & 0 & x_{44} & \dots & 0 & & \Gamma_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & & \\ 0 & 0 & 0 & \dots & x_{nn} & & \\ \hline & & * & & & & \Delta_2 \end{array}\right),$$

where the only non-zero entries of Γ_2 in the row corresponding to E_{ii} are in the columns corresponding to E_{ji} for some j, and the entry there is precisely x_{ij} (the order of indices matters for the basis elements but not for the variables!). The only non-zero entries of Δ_2 in the row corresponding to E_{ij}, $i < j$, are in the columns corresponding to E_{ij}, E_{ji}, E_{ki} for $k \neq i, j$, and E_{kj} for $k \neq i, j$, with entries x_{ii}, x_{jj}, x_{jk}, and x_{ik}, respectively. In particular, all non-zero entries of $[d_2]$ are simply variables up to sign. Moreover, all non-zero entries of Γ_2 are *off-diagonal*, and no variable ever appears twice in the same row of Γ_2 or Δ_2. This implies that, by setting to zero any set of entries in Γ_2, Δ_2 and the lower left block $*$, those columns of $[d_2]$ that do not become identically zero will remain linearly independent over the ground field K. Note also that by setting to zero any set of off-diagonal variables, no column will become identically zero.
For d_3 and $i < j$ we compute

$$\begin{aligned} d_3(E_{ij} - E_{ji}) = x_{ij}(E_{ii} - E_{11}) - x_{ij}(E_{jj} - E_{11}) + x_{jj}E_{ij} - x_{ii}E_{ji} \\ + \sum_{\substack{k=1 \\ k \neq i,j}}^{n} x_{jk}E_{ik} + \sum_{\substack{k=1 \\ k \neq i,j}}^{n} (-x_{ik})E_{jk}. \end{aligned}$$

So again, all non-zero entries of $[d_3]$ are variables up to sign. If we write $[d_3]$ as a block matrix

$$[d_3] = \begin{pmatrix} \Gamma_3 \\ \hline \Delta_3 \end{pmatrix},$$

where the rows of Γ_3 correspond to the basis elements $E_{ii} - E_{11}$, $i = 2, \dots, n$, and the rows of Δ_3 correspond to the remaining basis elements of ker(tr), then no variable ever occurs twice in the same row of $[d_3]$. Moreover, all variables appearing in Γ_3 are off-diagonal and precisely one diagonal variable appears in every row of Δ_3.

2.3.3 Homogenizing $L(X)$

Let G be any undirected, simple graph on $[n]$. We will now homogenize $L(X)$ with respect to the weight w_G from Section 2.2 by introducing a new variable t with weight 1. We start by homgenizing all $(n-1)$-cofactors which are the entries of the row vector $[d_1]$. We fix one cofactor. All terms of the latter with w_G-degree strictly less than the w_G-degree of the cofactor itself are now multiplied by the appropriate power of t, as usual. The resulting row vector is denoted $[d_1]^h$. Next, we fix a column of $[d_2]$. Let d be the maximum w_G-degree of the product of an entry of this column with the corresponding $(n-1)$-cofactor of $[d_1]^h$. Then we multiply every entry of this column of $[d_2]$ by the appropriate power of t such that the product of the resulting entry and the corresponding minor has combined degree equal to d. This we do for every column of $[d_2]$ and call the result $[d_2]^h$. Note that no column of $[d_2]^h$ is divisible by t. Similarly, we fix a column of $[d_3]$ and we choose any row of $[d_2]^h$. Then we homogenize this column of $[d_3]$ in the same way as we did for $[d_2]$. The result does not depend on the chosen row of $[d_2]^h$. Again, no column of $[d_3]^h$ will be divisible by t. By construction, the resulting sequence of matrices still gives a *complex* $L(X)^h$, i.e. $[d_2]^h[d_3]^h = 0$ and $[d_1]^h[d_2]^h = 0$.

Lemma 2.3.2. Let R be a positively graded finitely generated algebra over the field $R_0 = K$ and let $\mathfrak{m} = R_{>0}$ be its homogeneous maximal ideal. View $K = R/\mathfrak{m}$ as a graded R-module concentrated in degree 0. Let $\varphi : N \hookrightarrow M$ be a graded injection of finitely generated graded R-modules. If all graded pieces of $N \otimes_R K$ and $M \otimes_R K$ have the same dimension as K-vector spaces, then $\overline{\varphi} : N \otimes_R K \to M \otimes_R K$ (and hence φ) is an isomorphism.

Proof. Since N and M are finitely generated, both N_d and M_d are zero in sufficiently negative degrees. Inductively, we can therefore assume that we are given some d such that $\overline{\varphi}$ is an isomorphism in all degrees $< d$ and we have to prove the same is true in degree d. First we observe that this implies that φ is an isomorphism in all degrees $< d$ by the graded Nakayama lemma. Moreover, by assumption it suffices to show that $\overline{\varphi}_d : (N \otimes_R K)_d \to (M \otimes_R K)_d$ is an injection. For this, in turn, let $x \in N_d$ such that $\varphi(x) \in M_d \cap (\mathfrak{m}M)$. This means $\varphi(x) = \sum_i r_i m_i$ for some $r_i \in \mathfrak{m}$ and $m_i \in M_{<d}$ such that $\deg(r_i) + \deg(m_i) = d$. Hence, $0 = \varphi(x - \sum_i r_i \varphi^{-1}(m_i))$, so the injectivity of φ_d implies $x = \sum_i r_i \varphi^{-1}(m_i) \in N_d \cap (\mathfrak{m}N)$, i.e., $\overline{x} = 0$ in $N \otimes_R K$. ☐

Proposition 2.3.3. The complex $L(X)^h|_{t=0}$ is exact and therefore the graded minimal free resolution of $R/\operatorname{in}_{w_G}(I_{n-1}(X))$. Moreover, $R/\operatorname{in}_{w_G}(I_{n-1}(X))$ is Cohen–Macaulay.

A consequence of this and the graded Nakayama lemma is that $L(X)^h$ is exact as well.

Proof. The reasoning is similar to the proof of [Boo12, Proposition 3.8]. Clearly, $L(X)^h|_{t=0}$ is a complex of free R-modules. The image of $[d_1]^h|_{t=0}$ agrees with $\operatorname{in}_{w_G}(I_{n-1}(X)) \subseteq R$ by Corollary 2.2.3 since $<_{T,G}$ refines the weight order $<_{w_G}$. But we can say more. Recall that $I_T = \operatorname{in}_{<_{T,G}}(I_{n-1}(X)) = \operatorname{in}_{<_{T,G}} \operatorname{in}_{<_{w_G}}(I_{n-1}(X))$. From Proposition 2.2.2 and the monotonicity of graded Betti numbers under weight degenerations [HH11, Theorem 3.3.1] we obtain that all graded Betti numbers of $I_{n-1}(X)$, $\operatorname{in}_{w_G}(I_{n-1}(X))$ and I_T coincide. Moreover, as a consequence of the Auslander–Buchsbaum formula, all three ideals define Cohen–Macaulay rings, proving the second claim.
From this the first claim now follows via Lemma 2.3.2 with $N = \operatorname{im}(d_i^h|_{t=0})$ and $M = \ker(d_{i-1}^h|_{t=0})$ because the dimension of the degree j-part of both $\operatorname{im}(d_i^h|_{t=0}) \otimes_R K$ and $\ker(d_{i-1}^h|_{t=0}) \otimes_R K$ is precisely $\beta_{i,j}(R/I_{n-1}(X))$ for $i = 2, 3$ and all j. For the former this is a consequence of the fact that (1) every non-zero entry of $[d_i]^h|_{t=0}$ is a

variable up to sign, (2) no column of $[d_i]^h|_{t=0}$ is identically zero and (3) no row of Γ_2, Δ_2 or $[d_3]$ contains any variable more than once. $\square$

Theorem 2.3.4. Let G be any undirected, simple graph on $[n]$. The minimal graded free resolution of $R/I_{n-1}(X_G)$ is obtained from $L(X)$ via Boocher's pruning procedure.

The proof below is exactly the same as Boocher's proof of [Boo12, Theorem 4.1] and follows from a careful study of the w_G-grading of $L(X)^h$ together with Proposition 2.3.3. We repeat the argument for the sake of completeness.

Proof. We first observe that all principal minors have w_G-weight $2n-2$ and the same is true for $\det(X_{[n]\setminus k,[n]\setminus l})$, $k<l$, whenever there is a path in G between k and l. All the other $(n-1)$-minors of X have w_G-weight $2n-3$. The minors of the smaller w_G-weight $2n-3$ are precisely those which vanish identically after substituting zeros for all variables in Z. All non-zero entries of the matrices $[d_2]$ and $[d_3]$ are simply variables up to sign and hence have w_G-weight at most 2. Therefore, the minimal occurring w_G-degree in the w_G-graded free $R[t]$-module of $L(X)^h$ at the i-th place is $-2(n-1)-2(i-1)$ for all $i\geq 1$. Therefore, with respect to the w_G-grading, $L(X)^h$ has the following shape:

$$\begin{matrix} \bigoplus_{c_j<2n+2} R[t](-c_j) \\ \bigoplus \\ \bigoplus R[t](-2n-2) \end{matrix} \xrightarrow{[d_3]^h} \begin{matrix} \bigoplus_{b_j<2n} R[t](-b_j) \\ \bigoplus \\ \bigoplus R[t](-2n) \end{matrix} \xrightarrow{[d_2]^h} \begin{matrix} \bigoplus_{a_j<2(n-1)} R[t](-a_j) \\ \bigoplus \\ \bigoplus R[t](-2(n-1)) \end{matrix} \xrightarrow{[d_1]^h} R[t].$$

We write

$$[d_i]^h = \left(\begin{array}{c|c} A_i & B_i \\ \hline C_i & D_i \end{array}\right)$$

for $i=2,3$ and $[d_1]^h = \left(C_1 \mid D_1\right)$, each time according to the direct sum decomposition. Just as Boocher, we can deduce three things from the grading alone:

- For $i=2,3$, every non-zero entry of B_i has w_G-weight at least 3 and is therefore divisible by t. In particular, $B_i|_{t=0}=0$.
- For $i=2,3$, every non-zero entry of D_i has w_G-weight 2 and is therefore either of the form $\pm x_{ij}$ for $ij\in G$ or $i=j$ or of the form $\pm t x_{ij}$ for $i<j$ and $ij\notin G$. In particular, $D_i|_{t=0}=D_i|_{Z=0}$. The latter is true also for $i=1$.
- For $i=2,3$, all non-zero entries of C_i necessarily have w_G-weight 1 and are hence of the form $\pm x_{ij}$ for $i<j$ and $ij\notin G$. Therefore, $C_i|_{Z=0}=0$ for all $i=1,2,3$.

After setting $t=0$ in $L(X)^h$, we obtain the maps

$$\begin{aligned} [d_1]^h|_{t=0} &= \left(C_1 \mid D_1|_{t=0}\right), \\ [d_i]^h|_{t=0} &= \left(\begin{array}{c|c} A_i|_{t=0} & 0 \\ \hline C_i & D_i|_{t=0} \end{array}\right) \quad \text{for } i=2,3, \end{aligned}$$

giving an exact sequence by Proposition 2.3.3. Now, tensoring with $R/(Z)$ does not change $D_i|_{t=0}$ for $i = 2, 3$ while C_i becomes the zero matrix for all $i = 1, 2, 3$. The pruning procedure will hence first erase all columns of C_1, so that $[d_1]^h|_{t=0}$ becomes simply $D_1|_{t=0}$, whose image is $I_{n-1}(X_G)$. The corresponding rows of $[d_2]^h|_{t=0}$, which are precisely those of A_2, will be erased as well. Proceeding in the same way, the resulting pruned complex P_G eventually only consists of the three maps D_1, D_2, D_3 with all variables from Z replaced by zeros. The complex P_G is then still exact: Denote by $\pi_i\colon (L(X)^h|_{t=0})^i \to P_G^i$ the projection onto the second big direct summand. Then $\pi_i([d_{i+1}]^h|_{t=0}(v)) - D_{i+1}|_{Z=0}(\pi_i(v))$ lies in the image of C_{i+1} whose non-zero entries are variables in Z up to sign. If, over $R/(Z)$, the element w is in the kernel of D_i, we consider the obvious lift of w to $(0, w)$ in the kernel of $[d_i]^h|_{t=0}$, now over R. By exactness of $L(X)^h|_{t=0}$, there exists v such that $(0, w) = [d_{i+1}]^h|_{t=0}(v)$ and hence

$$w = \pi_i((0, w)) = \pi_i([d_{i+1}]^h|_{t=0}(v)) = D_{i+1}|_{Z=0}(\pi_i(v)) \bmod Z,$$

so that over $R/(Z)$ the element w lies in the image of $D_{i+1}|_{Z=0}$. This shows that P_G is exact over $R/(Z)$ and hence also over R because R is a free module over $R/(Z)$. $\square$

Corollary 2.3.5. The graded Betti numbers of $R/I_{n-1}(X_G)$ are those stated in Theorem 2.1.1.

Proof. By Theorem 2.3.4, it is enough to understand which columns of the matrices $[d_i]|_{Z=0}$ are identically zero. First, the entries of the row vector $[d_1]|_{Z=0}$ are $\pm\det((X_G)_{[n]\setminus k,[n]\setminus l})$ for $k \leq l$ and the latter vanishes if and only if $k \neq l$ and there is no path between k and l in G. Next, we consider our explicit description of the matrix $[d_2]$. We claim that, after pruning, the column of $[d_2]|_{Z=0}$ corresponding to the basis element E_{ij}, $i \neq j$, is zero precisely if there is no path in G between i and j. The entry in the row corresponding to E_{ij}, if $i < j$, resp. to E_{ji}, if $i > j$, is the diagonal variable x_{ii}. Thus, for the column to vanish this row must have been erased in the pruning process, which is the case if and only if the minor corresponding to ij vanishes identically after substituting zero for all variables in Z. This, as we saw, is equivalent to i and j lying in different connected components of G. Conversely, if this is the case, then the column of $[d_2]|_{Z=0}$ indeed vanishes after pruning because the remaining variables appearing in this column are x_{ki} for $k \neq i$ which appears only in the row corresponding to E_{kj} if $k \leq j$ resp. E_{jk} if $k > j$. But either $x_{ki} \in Z$ or $ki \in G$, and in the last case k and j necessarily lie in different connected components of G, so the row corresponding to E_{kj} resp. E_{jk} must have been erased in the pruning process. The argument for $[d_3]$ is similar. $\square$

2.4 A Characteristic Number for Sparse Quadrics

Let $K = \overline{K}$ and $\mathrm{char}(K) \neq 2$ in this section. The set of quadric hypersurfaces in $\mathbb{P}_K^{n-1}$ is identified with the set of non-zero symmetric $n \times n$ matrices over K up to scaling, i.e., with $\mathbb{P}(\mathrm{Sym}^2(K^n))$. A G-sparse quadric is one where all off-diagonal entries of its associated symmetric matrix corresponding to the non-edges of G are

zero. Geometrically, this is a coordinate linear subspace of $\mathbb{P}(\mathrm{Sym}^2(K^n))$, canonically identified with $\mathrm{Proj}(R_G) = \mathbb{P}^{N_G-1}$, where $R_G := R/(Z)$ and $N_G := \dim(R_G) = |E_G| + n$, E_G being the edge set of G.
An immediate consequence of Theorem 2.1.1 is that we can compute the degree of the vanishing subscheme $V(I_{n-1}(X_G)) \subseteq \mathbb{P}^{N_G-1}$.

Proposition 2.4.1. If G is a disconnected graph on $[n]$, then the codimension of $V(I_{n-1}(X_G)) \subseteq \mathbb{P}^{N_G-1}$ is 2 and its degree is D_G. If G is connected on $[n]$, then the codimension of $V(I_{n-1}(X_G)) \subseteq \mathbb{P}^{N_G-1}$ is 3 and its degree is $\binom{n+1}{3}$.

Proof. By Theorem 2.1.1, the minimal graded free resolution of $R_G/I_{n-1}(X_G)$ over R_G (which looks the same as over R) has the form

$$0 \to R_G(-(n+1))^{\binom{n}{2}-D_G} \to R_G(-n)^{n^2-1-2D_G} \to R_G(-(n-1))^{\binom{n+1}{2}-D_G} \to R_G.$$

For the Hilbert series of $R_G/I_{n-1}(X_G)$ we deduce

$$\begin{aligned}\mathrm{HS}(R_G/I_{n-1}(X_G)) &= \frac{1-\left(\binom{n+1}{2}-D_G\right)t^{n-1}+(n^2-1-2D_G)t^n-\left(\binom{n}{2}-D_G\right)t^{n+1}}{(1-t)^{N_G}}\\ &= \frac{t^{n-1}(1-t)^2D_G+(1-t)^3\sum_{k=0}^{n-2}\binom{k+2}{2}t^k}{(1-t)^{N_G}}.\end{aligned}$$

Observing that $\sum_{k=0}^{n-2}\binom{k+2}{2} = \binom{n+1}{3}$, the last computation gives all claims after canceling $(1-t)^2$ for $D_G \neq 0$ or $(1-t)^3$ for $D_G = 0$. □

An application of Proposition 2.4.1 is the following geometric result.

Corollary 2.4.2. Let $n \geq 3$ and $\mathrm{char}(K) = 0$. For any graph G on $[n]$, the number of smooth G-sparse quadrics in $\mathbb{P}^{n-1}$ tangent to 2 general hyperplanes and passing through $N_G - 3$ general points is

$$(n-1)^2 - D_G.$$

If G is a connected graph on $[n]$, then moreover the number of smooth G-sparse quadrics in $\mathbb{P}^{n-1}$ tangent to 3 general hyperplanes and passing through $N_G - 4$ general points is

$$(n-1)^3 - \binom{n+1}{3} = \frac{(n-1)(n-2)(5n-3)}{6}.$$

Proof. This follows from Proposition 2.4.1. The condition of passing through a given point is a hyperplane in $\mathbb{P}^{N_G-1}$ while the condition of being tangent to a given hyperplane is a hypersurface of degree $n-1$ in $\mathbb{P}^{N_G-1}$ (defined by some linear combination of the $(n-1)$-minors of X_G). By Theorem 2.1.1, the codimension of $V(I_{n-1}(X_G))$ in $\mathbb{P}^{N_G-1}$ is 3 or 2 according to whether G is connected or not. Hence, for codimension 2, cutting $V(I_{n-1}(X_G))$ by the $N_G - 3$ hyperplanes corresponding to the same number of general points gives us a set of $\deg(V(I_{n-1}(X_G)))$ many points, counted with multiplicities. The two tangency conditions will contain all these points in $V(I_{n-1}(X_G))$ but will cut down the complement (which is an open of a 2-dimensional linear space) to a zero-dimensional set of $(n-1)^2 - \deg(V(I_{n-1}(X_G)))$ many points, counted according to their multiplicities. This is an instance of residual intersection, see [Ful98, Proposition 9.1.2]. If the points and hyperplanes are sufficiently general, none of the points outside $V(I_{n-1}(X_G))$ corresponds to a *singular* G-sparse quadric because the intersection of all point conditions is empty and the intersection of all hyperplane tangency conditions is precisely $V(I_{n-1}(X_G))$, so any fixed proper subset of $\mathbb{P}^{N_G-1} \setminus V(I_{n-1}(X_G))$ can be avoided, such as the set of rank $n-1$ quadrics. Finally, by [FKM83, Theorem (d), p. 162], for sufficiently general points and hyperplanes, the multiplicity of each of the points outside $V(I_{n-1}(X_G))$ is 1; note that this uses $\mathrm{char}(K) = 0$. If $V(I_{n-1}(X_G))$ has codimension 3, the argument is analogous. □

Let us remark that the statement about multiplicities at the end of the proof is also true in every positive characteristic $p > 2$ which does not divide the respective characteristic number. This is again a consequence of [FKM83, Theorem (c)+(d)]. In particular, for the first of the two given characteristic numbers this is the case for all $p > (n-1)^2$ and for the second one $p > 5n - 3$ will suffice.

2.5 Open Questions

2.5.1 Primality

It would be desirable to have a combinatorial characterization for when $I_{n-1}(X_G)$ is prime, and this same question can of course be asked for the ideals of minors of arbitrary size $I_k(X_G)$, $1 \le k \le n$. It is clear that $I_1(X_G)$ is always a prime ideal but even for $k = 2$ and $k = n$ the answer is not entirely trivial. A necessary condition for $I_k(X_G)$ to be prime where $2 \le k \le n$ is that G is $(n-k+1)$-connected, i.e., for any subset $M \subseteq [n]$ of cardinality $|M| = k$, the induced subgraph of G on M is connected. Indeed, if $G|_M$ is disconnected, then the principal minor $\det((X_G)_{M,M})$ factors as a product of two lower-order principal minors because after some permutation of M the matrix $(X_G)_{M,M}$ is block-diagonal. But none of the two factors can be contained in $I_k(X_G)$ for degree reasons. This observation is also present in [CW19, Lemma 7.10]. The following example was provided by Aldo Conca in personal communication. It shows that this necessary combinatorial condition is *not* sufficient in general.

Example 2.5.1. Let $n = 8$, $k = 5$ and $\mathrm{char}(K) = 0$. Let H be the complement graph of the complete bipartite graph $K_{4,4}$ with partition $[8] = \{1, 2, 3, 4\} \sqcup \{5, 6, 7, 8\}$. Adding to H the four additional edges 15, 26, 37 and 48, we obtain a graph G. The induced subgraph of G on any set of 5 vertices is connected, so G is 4-connected.

The sparse generic symmetric matrices for G and H are:

$$X_G = \begin{pmatrix} x_{11} & x_{12} & x_{13} & x_{14} & x_{15} & 0 & 0 & 0 \\ x_{12} & x_{22} & x_{23} & x_{24} & 0 & x_{26} & 0 & 0 \\ x_{13} & x_{23} & x_{33} & x_{34} & 0 & 0 & x_{37} & 0 \\ x_{14} & x_{24} & x_{34} & x_{44} & 0 & 0 & 0 & x_{48} \\ x_{15} & 0 & 0 & 0 & x_{55} & x_{56} & x_{57} & x_{58} \\ 0 & x_{26} & 0 & 0 & x_{56} & x_{66} & x_{67} & x_{68} \\ 0 & 0 & x_{37} & 0 & x_{57} & x_{67} & x_{77} & x_{78} \\ 0 & 0 & 0 & x_{48} & x_{58} & x_{68} & x_{78} & x_{88} \end{pmatrix},$$

$$X_H = \begin{pmatrix} x_{11} & x_{12} & x_{13} & x_{14} & 0 & 0 & 0 & 0 \\ x_{12} & x_{22} & x_{23} & x_{24} & 0 & 0 & 0 & 0 \\ x_{13} & x_{23} & x_{33} & x_{34} & 0 & 0 & 0 & 0 \\ x_{14} & x_{24} & x_{34} & x_{44} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & x_{55} & x_{56} & x_{57} & x_{58} \\ 0 & 0 & 0 & 0 & x_{56} & x_{66} & x_{67} & x_{68} \\ 0 & 0 & 0 & 0 & x_{57} & x_{67} & x_{77} & x_{78} \\ 0 & 0 & 0 & 0 & x_{58} & x_{68} & x_{78} & x_{88} \end{pmatrix}.$$

We claim that $I_5(X_G)$ is not prime. For this, let $J = (x_{15}, x_{26}, x_{37}, x_{48})$. Then, clearly $I_5(X_G) + J = I_5(X_H) + J$. Since X_H is block-diagonal, every 5-minor of X_H factors as a product of either a variable in one block and the determinant of the other block or as the product of a 2-minor of one block and a 3-minor of the other block. Laplace expansion shows that the determinant of a square matrix is contained in the ideal of all submaximal minors. In particular, $I_5(X_H)$ is contained in $I_1 + I_2$, where I_1, I_2 are the ideals of all 3-minors of the two blocks, respectively. It is well-known that I_1 and I_2 are of height 3. In particular, $I_1 + I_2 + J$ is a prime ideal of height $3 + 3 + 4 = 10$ since the three ideals are geometrically prime and involve disjoint sets of variables. We obtain $I_5(X_G) \subseteq I_5(X_G) + J = I_5(X_H) + J \subseteq I_1 + I_2 + J$. On the other hand, a computation in `Macaulay2` for $K = \mathbb{Q}$ gives $\operatorname{ht} I_5(X_G) = 10$ over $\mathbb{Q}$ and hence over any field K of $\operatorname{char}(K) = 0$. So if $I_5(X_G)$ was prime, necessarily $I_5(X_G) = I_1 + I_2 + J$, which is impossible for degree reasons.

Nonetheless, for some values of k the necessary combinatorial condition is actually sufficient as the next result states.

Proposition 2.5.2. For $k = n$, the principal ideal $I_n(X_G) = (\det(X_G))$ is prime if and only if G is connected. In case $\operatorname{char}(K) = 0$ and $k = 2, 3$, again $I_k(X_G)$ is prime if and only if G is $(n - k + 1)$-connected.

Proof. For $k = n$, it is enough to prove that $\det(X_G)$ is an irreducible polynomial in R. We adapt a combinatorial proof of the case $G = K_n$.[1] First, we recall the following elementary fact: Let S be any integral domain and $f \in S[x]$ a polynomial over S in a single variable x. If $f = ax + b$ for $a, b \in S$, $a \neq 0$, and f factors as $f = gh$ in $S[x]$, then precisely one of g and h is linear in x, i.e. of the form $cx + d$, $c \neq 0$, and the other one does not involve x at all.
Let now $f := \det(X_G)$ where G is connected. We have the following two facts:

[1]see the accepted answer of the following stackexchange post:
math.stackexchange.com/questions/1893344/determinant-of-symmetric-matrix-is-an-irreducible-polynomial

- f is linear in x_{ii} for all $i = 1, \dots, n$.
- For all $i < j$ such that $ij \in G$, we have that f contains terms which are divisible by x_{ij} but no term of f is divisible by $x_{ii}x_{ij}$ or $x_{jj}x_{ij}$.

Now we assume that $f = gh$ in R. Without loss of generality, we assume that g is linear in x_{11}, hence h is independent of x_{11} by the above. Therefore, h is also independent of x_{1i} for all $i > 1$ such that $1i \in G$ since otherwise f would contain some term divisible by $x_{11}x_{1i}$. Hence, g contains terms divisible by x_{1i} for all $i > 1$ such that $1i \in G$ since f contains such terms and h does not. We conclude that all non-zero entries of X_G in the first row appear only in g but not in h. Next, for every i such that 1 is incident to i in G, h must be independent of x_{ii} as well, otherwise f would again contain terms of the form $x_{ii}x_{1i}$. Hence, g must be linear in x_{ii} for all i incident to 1 in G. With the same argument as before, h is then also independent of all x_{ij} such that $ij \in G$ while g contains terms divisible by each of these variables. We conclude that all non-zero entries of X_G in the i-th row only appear in g but not in h for every i such that $1i$ is an edge of G. Continuing in this way, since G is connected, we will eventually reach every row of X_G, implying that h is constant, which concludes the proof.
For $k = 2, 3$ and $\mathrm{char}(K) = 0$, the second claim is just a reformulation of [CW19, Theorem 7.8] since their graph is precisely the complement graph G^c of G. The important observation for $k = 2, 3$ is that the complement of any $(n-k+1)$-connected graph G is a forest of maximal degree at most $k-2$. Indeed, for $k = 2$ the graph G is $(n-1)$-connected if and only if G induces a connected graph on any pair of vertices. This means that $G = K_n$ is the complete graph. Hence, G^c has no edges at all, so G^c is clearly a forest of maximal degree 0. For $k = 3$ the graph G is $(n-2)$-connected if and only if G induces a connected graph on any triple of vertices, so either the triangle or the path. For G^c this means that there is at most one edge between any three vertices. In particular, every vertex has degree at most 1 in G^c, so G^c is a forest of maximal degree 1. $\square$

Motivated by these results, we ask the following question.

Question 2.5.3. Is the ideal $I_{n-1}(X_G)$ prime whenever G is 2-connected?

For $K = \mathbb{Q}$, we checked with `Macaulay2` that the answer is affirmative if $n \leq 6$.

2.5.2 More Sparsity

Graphical models in algebraic statistics motivate the study of ideals generated by only *some* submaximal minors of a generic symmetric matrix. If in addition the matrix is allowed to be sparse, one arrives at the notion of a so-called *double Markovian model* [BKKR23], see Chapter 3. The following result answers a combinatorial question raised in [BKKR23, Remark 9], showing that the positivity assumption of Corollary 3.3.24 can be relaxed.

Proposition 2.5.4. Let $A \in \mathrm{Sym}^2(K^n)$ be a symmetric $n \times n$ matrix which is principally regular, i.e., all principal minors of A are non-zero. If for all $1 \leq i < j \leq n$ we have $A_{ij} \cdot (A^{-1})_{ij} = 0$, then A is a diagonal matrix.

Proof. We write $A = (a_{ij})_{i,j\in[n]}$. First note that any principal submatrix of a principally regular symmetric matrix is also principally regular by definition. Moreover, an invertible symmetric matrix is principally regular if and only if so is its inverse. The latter follows from the formula

$$\det((A^{-1})_{I,I}) = \frac{\det(A_{[n]\setminus I,[n]\setminus I})}{\det(A)}$$

for any $I \subseteq [n]$. If A is not diagonal, then after a permutation of $[n]$ we may assume that there exists a non-zero off-diagonal entry in the first row. We can even assume that there is $k \geq 2$ such that $a_{1i} = 0$ for all $i \geq k+1$ and $a_{1i} \neq 0$ for all $2 \leq i \leq k$. Then

$$A_{[n]\setminus 1,[n]\setminus i} = \left(\begin{array}{c|c} a_{12} & \\ a_{13} & \\ \vdots & \\ a_{1k} & A_{[n]\setminus 1,[n]\setminus 1i} \\ 0 & \\ \vdots & \\ 0 & \end{array}\right).$$

By hypothesis it now follows that $(A^{-1})_{1i} = 0$ for all $2 \leq i \leq k$ which translates into

$$0 = (-1)^i \det(A_{[n]\setminus 1,[n]\setminus i}) = \sum_{j=2}^{k} (-1)^{i+j} a_{1j} \det(A_{[n]\setminus 1j,[n]\setminus 1i})$$

for all $2 \leq i \leq k$. Equivalently, in matrix form,

$$\Big((-1)^{i+j}\det(A_{[n]\setminus 1j,[n]\setminus 1i})\Big)_{i,j=2,3,\dots,k} \cdot \begin{pmatrix} a_{12} \\ a_{13} \\ \vdots \\ a_{1k} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{2.1}$$

We observe that

$$(-1)^{i+j}\det(A_{[n]\setminus 1j,[n]\setminus 1i}) = \det(A_{[n]\setminus 1,[n]\setminus 1}) \cdot ((A_{[n]\setminus 1,[n]\setminus 1})^{-1})_{i-1,j-1}.$$

But $(A_{[n]\setminus 1,[n]\setminus 1})^{-1}$ is a principally regular symmetric matrix, hence the $(k-1)\times(k-1)$ matrix in (2.1) is invertible, implying $a_{12} = a_{13} = \cdots = a_{1k} = 0$. □

3 The Geometry of Gaussian Double Markovian Distributions

For questions of authorship, please refer to pages IVf.

This chapter is based on the preprint version of [BKKR23]. The main differences are the inclusion of the missing cases in the classification of all double Markovian models $\mathcal{M}(G,H)$ with $|E_G \cap E_H| = 3$ in Proposition 3.3.31, a new version of Section 3.3.1 and a reference to Proposition 2.5.4 in Remark 3.3.25. Proposition 2.5.4 answers the question raised in [BKKR23, Remark 9].

3.1 Mixing Graphical and Covariance Models

Let G and H be two undirected, simple graphs on the vertex set $[n] = \{1, \ldots, n\}$. Denote by PD_n the set of real symmetric positive definite $n \times n$ matrices. In this chapter, we study the following statistical models.

Definition 3.1.1. The *Gaussian double Markovian model of G and H* is

$$\mathcal{M}(G,H) \coloneqq \left\{\Sigma \in \mathrm{PD}_n : \left(\Sigma^{-1}\right)_{ij} = 0 \text{ for all } ij \notin G,\ \Sigma_{kl} = 0 \text{ for all } kl \notin H\right\}.$$

When writing $ij \in G$ or $ij \notin G$, we always assume $i \neq j$ (and similarly for k and l), so that diagonal entries are unconstrained since otherwise the set would be empty. Denoting the complete graph on $[n]$ by K_n, ordinary undirected Gaussian graphical models occur in Definition 3.1.1 as the special case $\mathcal{M}(G) \coloneqq \mathcal{M}(G, K_n)$. Covariance models are precisely the models $\mathcal{M}(K_n, H)$, so both model classes are unified and generalized here.

3.1.1 Conditional Independence and Graphical Modeling

Conditional independence (CI) constraints are a central tool in mathematical modeling of random events. For random variables $X_1, \ldots, X_n$, a conditional independence statement prescribes that one random variable X_i be independent of another variable X_j given a collection of remaining variables $(X_k)_{k\in K}$ where $i, j \in [n]$ and $K \subseteq [n] \setminus \{i,j\}$. CI constraints for discrete or normally distributed random variables translate into polynomial conditions on elementary probabilities in the discrete case and on covariance matrices in the Gaussian case. One aim of algebraic statistics is to understand the algebraic and geometric properties of conditional independence models and to relate them to properties of statistical inference procedures.

We exclusively treat Gaussian random variables in this chapter, i.e., we assume that $(X_1, \ldots, X_n)$ has a multivariate normal distribution with a positive definite covariance matrix $\Sigma \in \mathrm{PD}_n$. Since the theory of CI is insensitive to the mean, we may restrict to centered distributions, i.e., distributions with mean zero.

In graphical modeling, edges and paths represent correlation or interaction and, conversely, notions of disconnectedness represent independence. Double Markovian models are conditional independence models whose CI statements are of the following two forms: either $X_k \perp\!\!\!\perp X_l$, for each non-edge kl of H, or $X_i \perp\!\!\!\perp X_j \,\middle|\, X_{[n]\setminus\{i,j\}}$, for each non-edge ij of G. Pairwise conditional independence as in the second type is common in graphical modeling [Lau96]. The resulting *Gaussian graphical model* of a simple undirected graph G on $[n]$ with edge set E_G is

$$\mathcal{M}(G) = \mathcal{M}(G, K_n) = \left\{\Sigma \in \mathrm{PD}_n : (\Sigma^{-1})_{ij} = 0 \text{ for all } ij \notin G\right\}.$$

In words, the non-edges of G specify zeros of the inverse covariance matrix Σ^{-1}, also called the *concentration matrix*. Gaussian graphical models first appeared in [SK86], and [Uhl19] is a modern survey containing many connections to e.g. optimization and matrix completion. Marginal independence constraints (as in the first case) also appear, for example with bidirected [DR02] or dashed graphs [CW93]. These models $\mathcal{M}(K_n, H)$ encode marginal independence constraints and are also known as covariance graph models [Kau96, LZ22].

A model similar to double Markovian models appeared recently in [LZ22, Example 3.4], where the authors consider graphical models with some entries of Σ^{-1} zero and complementary entries of Σ nonnegative. They investigate efficient estimation procedures. These models go back to work of Kauermann [BR03, Kau96]. One way to describe them is via mixed parametrizations: the regular exponential family of all multivariate mean-zero Gaussians can be parametrized by the mean parameter $\Sigma = (\sigma_{ij})$ or the natural parameter $\Sigma^{-1} = (k_{st})$. One can also employ a mixed parametrization, using σ_{ij} and k_{st} for $ij \in A$ and $st \in B$, where $A \dot\cup B$ is a partition of the entries of an $n \times n$-symmetric matrix. Double Markovian models with $G \cup H = K_n$ arise from imposing zeros in the mixed parametrization. In general they do not form regular exponential families, though. In the terminology of [LZ22], linear constraints on mixed parameters define *mixed linear exponential families*. Such models also appear in causality theory [PW94].

We study geometric properties of statistical models since they can imply favorable statistical properties. The asymptotic behavior of M-estimators like the MLE depends on properties of tangent cones that go under the name *Chernoff regularity* in [Gey94]. Drton has shown that the nature of singularities determines the large sample asymptotics of likelihood ratio tests [Drt09]. Smoothness of log-linear models for discrete random variables has been studied by examining the parametrization by marginals [Eva15, For12]. Generally, smoothness is favorable because estimation procedures using analytical techniques like gradient descent rely on it.

In several occasions, for example in Corollary 3.3.21 geometric niceness results follow because either $\mathcal{M}(G, H)$ or its inverse $\mathcal{M}(H, G)$ is an ordinary graphical model and thus irreducible, connected, and smooth. In these cases one has found an effective new parametrization of $\mathcal{M}(G, H)$. This theme has occurred in the literature. For

example [DR08] asks for a Markov equivalent directed and undirected graph to the bidirected (i.e. covariance) graph. Our geometric niceness results, however, go beyond recognizing disguised graphical models.
A systematic analysis of smoothness of Gaussian CI models has been initiated in [DX10]. That paper treats the $n = 4$ case in detail. It relies on similar algebraic techniques as we do here, but also on the characterization of realizable 4-gaussoids from [LM07]. We deal with a smaller class of models here, but achieve results independent of the number of random variables, aiming to understand how the geometry of $\mathcal{M}(G, H)$ depends on G and H. In particular, we are interested in dimension, smoothness, irreducible decompositions and other basic geometric facts that seem useful and interesting for inference methodology.
Algebraically, a double Markovian model $\mathcal{M}(G, H)$ is the vanishing set inside PD_n of an ideal generated by some entries of $\Sigma \in \mathrm{PD}_n$ and some more of its inverse. The latter is an algebraic condition as it can be encoded as the vanishing of submaximal minors of Σ. This puts us broadly in the framework of sparse determinantal ideals, see Definition 3.3.17 for the concrete class of ideals we are concerned with. The sparsity is twofold in our setting: our ideals are generated by only *some* minors of a *sparse* generic symmetric matrix, i.e., a symmetric matrix whose entries in the upper triangle are either distinct variables or zero. To our knowledge, no systematic study of these ideals has been carried out, even in the case of submaximal minors only. Minors of symmetric matrices are a classical topic in commutative algebra, see for example [Con94b, CW19]. Our results, in particular Theorem 3.3.20, can be viewed as a further step towards the study of this class of sparse determinantal ideals.

We illustrate our results on a simple preliminary example. Section 3.4 contains further examples.

Example 3.1.2. Let $G =$  be a star with edge set $\{12, 13, 14\}$ and $H =$  a path with edge set $\{12, 23, 34\}$. To study the model $\mathcal{M}(G, H)$, consider two indeterminate symmetric 4×4-matrices $\Sigma = (\sigma_{ij}), K = (k_{ij})$ representing covariance and concentration matrices. The non-edges of G dictate the zeros of K and the non-edges of H those of Σ. Algebraically (which means ignoring the positive definiteness for a moment), the model is specified by the equations

$$\Sigma K = \mathbb{1}_4, \quad k_{23} = k_{24} = k_{34} = \sigma_{13} = \sigma_{14} = \sigma_{24} = 0. \tag{3.1}$$

These equations can be solved in `Macaulay2` [GS] using primary decomposition algorithms. This computation shows that the complex algebraic variety defined by (3.1) consists of two irreducible components. In one of the components, $k_{22} = s_{11} = 0$ holds. So, if this component contains covariance matrices at all, then they are of non-regular Gaussians. We do not consider this further in this example, although boundary components can be important with respect to marginalization; see Example 3.2.16. The other component consists of the matrices Σ of the form

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & 0 & 0 \\ \sigma_{12} & \sigma_{22} & 0 & 0 \\ 0 & 0 & \sigma_{33} & 0 \\ 0 & 0 & 0 & \sigma_{44} \end{pmatrix}, \quad K = \begin{pmatrix} k_{11} & k_{12} & 0 & 0 \\ k_{12} & k_{22} & 0 & 0 \\ 0 & 0 & k_{33} & 0 \\ 0 & 0 & 0 & k_{44} \end{pmatrix},$$

subject to the constraints $\Sigma K = \mathbb{1}_4$. This component is 5-dimensional and contains positive definite matrices. If one additionally normalizes the variances as $\sigma_{ii} = 1$ for $i = 1, \dots, 4$, then Σ is positive definite exactly if $\sigma_{12} \in (-1, 1)$. The (complex) singular locus of the interesting component is empty. In fact, smoothness is clear from the simple parametrization.

Several features of this example follow from general results in this chapter. For example, the vanishing ideal of the model is a monomial ideal by Theorem 3.3.20. That $\mathcal{M}_1(G, H)$ is a curve segment is explained by Proposition 3.3.28. The block-diagonal structure follows from Theorem 3.3.23.

Remark 3.1.3. For computation it is sometimes useful to work in (Σ, K)-space as in Example 3.1.2. Let $\Sigma = (\sigma_{ij})$ and $K = (k_{st})$ be generic symmetric matrices (possibly with ones on the diagonal). To computationally answer algebraic questions about constrained covariance matrices, one considers the rings $\mathbb{R}[\sigma_{ij} : i \leq j]$ and $\mathbb{R}[\sigma_{ij}, k_{st} : i \leq j, s \leq t]$ of polynomials with real coefficients and whose indeterminates stand for entries of the symmetric matrices $\Sigma = (\sigma_{ij})$ and $K = (k_{st})$. To impose equational constraints, one forms quotients by ideals generated by the equations. For example, the relation that $\Sigma K = \mathbb{1}_n$ is implemented by construction of the quotient ring $\mathbb{R}[\sigma_{ij}, k_{st} : i \leq j, s \leq t]/(\Sigma K - \mathbb{1}_n)$. Another useful trick is to impose non-vanishing or invertibility of certain polynomials, for example $\det(\Sigma)$. This is achieved by localization. The ring $\mathbb{R}[\sigma_{ij} : i \leq j]_{\det(\Sigma)}$ is an enlarged version of $\mathbb{R}[\sigma_{ij} : i \leq j]$ where now $\det(\Sigma)$ is invertible. In fact, the natural map

$$\mathbb{R}[\sigma_{ij} : i \leq j]_{\det(\Sigma)} \xrightarrow{\cong} \mathbb{R}[\sigma_{ij}, k_{st} : i \leq j, s \leq t]/(\Sigma K - \mathbb{1}_n),$$

defined by mapping σ_{ij} to σ_{ij} is an isomorphism, as it should be because the constraint that $K = \Sigma^{-1}$ makes all variables k_{st} functions of the σ_{ij}. If $I \subseteq \mathbb{R}[\sigma_{ij} : i \leq j]$ is an ideal, then the restriction of its extension in the localization at $\det(\Sigma)$ agrees with its saturation at $\det(\Sigma)$. This provides a way to study conditional independence ideals in the ring $\mathbb{R}[\sigma_{ij}, k_{st} : i \leq j, s \leq t]/(\Sigma K - \mathbb{1}_n)$, where almost-principal minors of high degree are directly available as the k_{st} variables. Saturation at $\det(\Sigma)$, however, is of course not equivalent to the saturation at all principal minors. We recommend [Sul18, Chapter 3] for a general introduction to computational methods of commutative algebra with a view towards statistics.

Remark 3.1.4. The term *double Markov property* appears in [KR13, Lemma 1] based on [CK11, Exercise 16.25, p. 392] where it is called *double Markovity*. It describes constraints on three random variables which are in a special *pair of Markov chains*. This notion is unrelated to the Markovness with respect to two undirected graphs studied here. We judge the potential for confusion low enough to reuse this term.

3.1.2 Overview of Results

The core of our work are several geometric and algebraic insights having implications for statistical procedures dealing with $\mathcal{M}(G, H)$. In these results it is sometimes useful to restrict dimension and work with correlation matrices, which are covariance matrices with ones on the diagonal. We write $\mathrm{PD}_{n,1}$ for the set of positive definite matrices with ones on the diagonal. It is bounded and known as the *elliptope*. Then $\mathcal{M}_1(G,H) := \mathcal{M}(G,H) \cap \mathrm{PD}_{n,1}$ is a *correlation model*. Let E_G denote the edge set of G and $G \cap H$ the graph on $[n]$ with edge set $E_G \cap E_H$, and similarly $G \cup H$ the graph with edge set $E_G \cup E_H$. An important insight is that geometric and algebraic properties of double Markovian models often depend on features or the simplicity of $G \cap H$. The first result is a decomposition theorem relying on a notion of direct sum defined via block matrices in Section 3.2.3.

Theorem A (Theorem 3.3.23 and Corollary 3.3.24)**.** Let $V_1, \dots, V_r$ be a partition of $[n]$ such that each V_i is the vertex set of a connected component of $G \cap H$. Then

$$\mathcal{M}(G,H) = \bigoplus_{i=1}^{r} \mathcal{M}(G|_{V_i}, H|_{V_i}),$$

i.e., every $\Sigma \in \mathcal{M}(G,H)$ has a block-diagonal structure with r diagonal blocks having rows and columns indexed by the V_i. In particular, the correlation model satisfies $\mathcal{M}_1(G,H) = \{\mathbb{1}_n\}$ if and only if $E_G \cap E_H = \emptyset$.

The next result exhibits that also the union of G and H contributes. If it is complete, then the imposed constraints are simple enough to show, for example, smoothness.

Theorem B (Theorems 3.3.8, 3.3.11 and 3.3.15)**.** For all graphs G and H we have $\dim(\mathcal{M}(G,H)) \leq |E_G \cap E_H| + n$. If $G \cup H = K_n$, then $\mathcal{M}(G,H)$ is smooth and all connected components have dimension $|E_G \cap E_H| + n$. Conversely, if $\mathcal{M}(G,H)$ attains this maximal dimension, its top-dimensional connected components are smooth with irreducible Zariski closure.

In Section 3.3.3 we initiate the study of connectedness of $\mathcal{M}(G,H)$ in the euclidean topology. We conjecture that all double Markovian correlations models are connected (Conjecture 3.4.5). This contrasts with the fact that, allowing semi-definite matrices, similarly defined variants of $\mathcal{M}_1(G,H)$ can consist of isolated points as in Example 3.4.4. We have the following results.

Theorem C (Corollary 3.3.21, Theorem 3.3.22, and Propositions 3.3.28, 3.3.29 and 3.3.31)**.** The double Markovian model $\mathcal{M}(G,H)$ is connected in the following cases:

(1) For every non-edge kl of G there is at most one path p in H connecting k and l (or if this holds for G and H exchanged).

(2) There is a vertex $i \in [n]$ such that for all non-edges kl of G, every path in H connecting k and l contains i (or if this holds for G and H exchanged).

(3) $|E_G \cap E_H| \leq 3$.

The next theorem is our main algebraic result, see Theorem 3.3.20 for a more general version. Here a forest is a (not necessarily connected) graph with no cycles.

Theorem D (Theorem 3.3.20)**.** Let G be any graph and H a forest. Then the vanishing ideal of $\mathcal{M}_1(G, H)$ is the square-free monomial ideal

$$\mathcal{I}(\mathcal{M}_1(G, H)) = (\sigma_{ij}, \sigma_p : ij \notin H,\ p \text{ path in } H \text{ with } e(p) \notin G).$$

Here, σ_p is the product of variables corresponding to edges in p, and $e(p)$ denotes its endpoints.

Finally, in Propositions 3.3.28 to 3.3.31 we give a classification up to symmetry and matrix inversion of all double Markovian models with $|E_G \cap E_H| \leq 3$.

3.2 Preliminaries on CI Structures

3.2.1 Gaussian Conditional Independence

Gaussian graphical models as well as the double Markovian models are conditional independence models: they are sets of Gaussian distributions specified by conditional independence assumptions derived from a graph or pair of graphs. The conditional independence relations of random variables can be studied combinatorially, using abstract properties of conditional independence instead of concrete numerical data like a density function or covariance matrix. To this end, we introduce formal symbols $(ij|K)$ where $i \neq j \in [n]$ and $K \subseteq [n] \setminus ij$. These formal symbols are subject to the efficient *Matúš set notation* where union is written as concatenation and singletons are written without curly braces. For example, ijK is shorthand for $\{i\} \cup \{j\} \cup K$. The symbol $(ij|K)$ shall represent the conditional independence $X_i \perp\!\!\!\perp X_j \,|X_K$ where $X_K = (X_k)_{k \in K}$. For Gaussian random variables $X_1, \ldots, X_n$, the CI statement $X_I \perp\!\!\!\perp X_J \,|X_K$ is equivalent to $\operatorname{rk} \Sigma_{IK,JK} = |K|$ by [Sul18, Proposition 4.1.9]. Using the adjoint formula for the inverse of a matrix, it can be seen that a statement $(ij|[n] \setminus ij)$ (as it appears in the definition of a graphical model) is equivalent to $(\Sigma^{-1})_{ij} = 0$.
If $I = i$ and $J = j$ are singletons, the rank condition is equivalent to the vanishing of the determinant of the square submatrix $\Sigma_{iK,jK}$. These determinants are *almost-principal minors*. It is well-known that the statements $(ij|K)$ completely describe the entire CI relation of a random vector [Mat97, Section 2]. The set of all conditional independence statements among n random variables is $\mathcal{A}_{[n]} = \{(ij|K) : i \neq j \in [n], K \subseteq [n] \setminus ij\}$. An abstract conditional independence relation is a subset of $\mathcal{A}_{[n]}$. Fundamental problems in the intersection of probability, computer science and information theory concern the set of *realizable* subsets $\mathfrak{R} \subseteq 2^{\mathcal{A}_{[n]}}$, meaning that for $\mathcal{R} \in \mathfrak{R}$ there is a random vector X satisfying all statements in $\mathcal{R}$ and none of those in $\mathcal{A}_{[n]} \setminus \mathcal{R}$.
To each positive definite $n \times n$ matrix Σ we associate a corresponding CI relation

$$\langle\!\langle \Sigma \rangle\!\rangle := \{(ij|K) : \operatorname{rk} \Sigma_{iK,jK} = |K|\} \subseteq \mathcal{A}_{[n]},$$

consisting exactly of the CI statements satisfied by Σ. Conversely, the covariance matrices satisfying all statements of a CI relation $\mathcal{R}$ form its *Gaussian conditional* *independence model*:

$$\begin{aligned}\mathcal{M}(\mathcal{R}) &\coloneqq \{\Sigma \in \mathrm{PD}_n : \det(\Sigma_{iK,jK}) = 0 \text{ for all } (ij|K) \in \mathcal{R}\} \\ &\subseteq \mathrm{PD}_n \subseteq \mathrm{Sym}^2(\mathbb{R}^n) \cong \mathbb{R}^{\binom{n+1}{2}}.\end{aligned}$$

We consider this set together with the subset topology with respect to the euclidean topology on the set of symmetric matrices $\mathrm{Sym}^2(\mathbb{R}^n)$. The cone PD_n is open in $\mathrm{Sym}^2(\mathbb{R}^n)$ as it is the preimage of $\mathbb{R}^n_{>0} \subseteq \mathbb{R}^n$ under the continuous map which sends $\Sigma \in \mathrm{Sym}^2(\mathbb{R}^n)$ to the vector in $\mathbb{R}^n$ of all leading principal minors, using Sylvester's criterion. Writing $\Sigma = (\sigma_{ij})$, the associated *correlation matrix* of Σ has $\frac{\sigma_{ij}}{\sqrt{\sigma_{ii}}\sqrt{\sigma_{jj}}}$ as its ij-entry. Its diagonal consists of ones and all non-diagonal entries lie in $(-1, 1)$ as $\det(\Sigma_{ij,ij}) > 0$. Therefore, the set of correlation matrices $\mathrm{PD}_{n,1}$ is an intersection of PD_n with an affine linear subspace of $\mathrm{Sym}^2(\mathbb{R}^n)$. This yields a subspace topology and also a canonical smooth structure on $\mathrm{PD}_{n,1}$, making it into a smooth submanifold of PD_n of codimension n. It is often convenient to work with the bounded set of correlation matrices in the model:

$$\mathcal{M}_1(\mathcal{R}) \coloneqq \{\Sigma \in \mathrm{PD}_{n,1} : \det(\Sigma_{iK,jK}) = 0 \text{ for all } (ij|K) \in \mathcal{R}\} = \mathrm{PD}_{n,1} \cap \mathcal{M}(\mathcal{R}).$$

Many favorable properties transfer between $\mathcal{M}_1(\mathcal{R})$ and $\mathcal{M}(\mathcal{R})$, especially if they are of differentiable nature, see Lemma 3.3.1. Some care is necessary when considering algebraic properties such as the number of Zariski irreducible components. $\mathcal{M}_1(\mathcal{R})$ is a linear section of $\mathcal{M}(\mathcal{R})$, so its algebraic properties may differ; see e.g. Example 3.4.1. There is no finite axiomatic characterization of the set $\mathfrak{R}$ of realizable CI relations that is valid for all n. Neither in general [Stu92] nor for Gaussians specifically [Sul09]. Closure properties of CI relations often have cryptic names going back to the search for a finite axiomatization. In this terminology, for $\Sigma \in \mathrm{PD}_n$, the relation $\langle\!\langle \Sigma \rangle\!\rangle$ is a weakly transitive, compositional graphoid. The compound of these properties is also the definition of *gaussoid* [LM07]. Gaussoids approximate Gaussian conditional independence in a similar way to matroids approximating linear independence [BDKS19].

3.2.2 Undirected Graphical Models

If G is a graph on $[n]$, then

$$\langle\!\langle G \rangle\!\rangle \coloneqq \{(ij|K) : K \text{ separates } i \text{ and } j \text{ in } G\}$$

denotes the CI *separation statements*, those that follow from separation in the graph. We refer to CI relations of the form $\langle\!\langle G \rangle\!\rangle$ as *Markov relations*. See [Lau96] for all details on modeling CI by graphs. Any relation $\langle\!\langle G \rangle\!\rangle$ is realizable, meaning that it equals $\langle\!\langle \Sigma \rangle\!\rangle$ for some $\Sigma \in \mathrm{PD}_n$ and one can even pick Σ with all positive correlations [BDKS19, Theorem 4]. The models realizing $\langle\!\langle G \rangle\!\rangle$ are smooth as they are inverse linear spaces and thus parametrized by a diffeomorphism (see Proposition 3.2.2). Since the CI relation $\langle\!\langle G \rangle\!\rangle$ of any graph G is realizable, it follows that $\langle\!\langle G \rangle\!\rangle$ is a gaussoid. In addition, graph separation is upward-stable, meaning that $(ij|L)$ implies $(ij|kL)$

for all $k \in [n] \setminus ijL$, and being an upward-stable gaussoid is even a characterization of being of the form $\langle\!\langle G \rangle\!\rangle$ for some undirected graph G by [Mat97, Proposition 2]. A *pseudographoid* is an abstract CI structure which satisfies the intersection property, which together with the semigraphoid axiom forms the definition of *graphoid*; see [LM07, Remark 1] for the terminology. The following lemma states that it is sufficient to verify that Σ satisfies the maximal separation statements (which correspond to the non-edges in G), for it to satisfy all separation statements for G. In this case Σ is *Markovian for G*.

Lemma 3.2.1. Let G be an undirected graph and Σ a complex symmetric matrix with non-vanishing principal minors and let $m = \{(ij|[n] \setminus ij) : i \neq j\}$ denote the set of maximal CI statements. Then $\langle\!\langle G \rangle\!\rangle \cap m \subseteq \langle\!\langle \Sigma \rangle\!\rangle$ implies $\langle\!\langle G \rangle\!\rangle \subseteq \langle\!\langle \Sigma \rangle\!\rangle$.

Proof. By [Mat05, Corollary 1] the CI structure $\langle\!\langle \Sigma \rangle\!\rangle$ is a weakly transitive, compositional graphoid already when Σ is a complex symmetric matrix with non-vanishing principal minors (which includes the real positive definite case). The lemma follows from [LM07, Lemma 3], by choosing $\mathcal{M} = \langle\!\langle G \rangle\!\rangle \cap m$. Then G is a graph with i and j adjacent if and only if $(ij|[n] \setminus ij) \notin \mathcal{M}$. Since $\mathcal{M} \subseteq \langle\!\langle \Sigma \rangle\!\rangle$ by assumption, it follows that $\langle\!\langle G \rangle\!\rangle \subseteq \langle\!\langle \Sigma \rangle\!\rangle$. □

By Lemma 3.2.1, Σ is Markovian for G if and only if $\langle\!\langle G \rangle\!\rangle \cap m \subseteq \langle\!\langle \Sigma \rangle\!\rangle$. The maximal CI statements $\langle\!\langle G \rangle\!\rangle \cap m$ point out precisely the non-edges in G and therefore Σ being Markovian for G is equivalent to $(\Sigma^{-1})_{ij} = 0$ for all $ij \notin G$. This shows that $\mathcal{M}(G) = \mathcal{M}(\langle\!\langle G \rangle\!\rangle)$.

Proposition 3.2.2. Every Markov relation $\langle\!\langle G \rangle\!\rangle$ is realizable by a regular Gaussian distribution. For each graph G, the model $\mathcal{M}(G)$ is irreducible and smooth.

Proof. Realizations for $\langle\!\langle G \rangle\!\rangle$ were constructed from (inverses of) generalized adjacency matrices in [LM07, Theorem 1]. By Lemma 3.2.1 the set $\mathcal{M}(G)^{-1} = \{\Sigma^{-1} : \Sigma \in \mathcal{M}(G)\}$ is a linear subspace intersected with the cone PD_n. It is the interior of a *spectrahedron*. As such it is an irreducible semi-algebraic set and smooth. These properties are transferred to the inverse $\mathcal{M}(G)$ by Lemmas 3.3.1 and 3.3.2 below. □

Matúš [Mat12, Theorem 2] proved a geometric characterization of the sets $\mathcal{M}(G)^{-1}$: among all Gaussian conditional independence models, they are precisely those which are convex subsets of PD_n.

3.2.3 Minors, Duality and Direct Sums

Marginalization and conditioning are natural operations on random vectors and can also be carried out on conditional independence structures. These abstract operations mimic the effect of statistical operations on a purely formal level.

Definition 3.2.3. Let $\mathcal{R} \subseteq \mathcal{A}_{[n]}$ and $k \in [n]$. The *marginal* and the *conditional* of $\mathcal{R}$ on $[n] \setminus k$ are, respectively,

$$\mathcal{R} \setminus k := \left\{(ij|K) \in \mathcal{A}_{[n]\setminus k} : (ij|K) \in \mathcal{R}\right\},$$
$$\mathcal{R} \,/\, k := \left\{(ij|K) \in \mathcal{A}_{[n]\setminus k} : (ij|k \cup K) \in \mathcal{R}\right\}.$$

Any set $\mathcal{R}' \subseteq \mathcal{A}_S$, $S \subseteq [n]$, obtained from $\mathcal{R}$ by a sequence of marginalization and conditioning operations is a *minor* of $\mathcal{R}$.

On covariance matrices, marginalizing away a variable $k \in [n]$ is achieved by taking the principal submatrix $\Sigma \setminus k := \Sigma_{[n]\setminus k}$. The conditional distribution on k is the Schur complement of the $k \times k$ entry $\Sigma \,/\, k := \Sigma_{[n]\setminus k} - \sigma_{kk}^{-1}\Sigma_{[n]\setminus k,k} \cdot \Sigma_{k,[n]\setminus k}$. This is proven in [Sul18, Theorem 2.4.2].

Lemma 3.2.4. For $\Sigma \in \mathrm{PD}_n$ we have $\langle\!\langle \Sigma \rangle\!\rangle \setminus k = \langle\!\langle \Sigma \setminus k \rangle\!\rangle$ and $\langle\!\langle \Sigma \rangle\!\rangle \,/\, k = \langle\!\langle \Sigma \,/\, k \rangle\!\rangle$. In particular, minors of realizable CI relations are realizable.

For any Gaussian distribution, matrix inversion exchanges the covariance and concentration matrices. The combinatorial version of this operation furnishes an involution on CI relations.

Definition 3.2.5. The *dual* of $\mathcal{R} \subseteq \mathcal{A}_{[n]}$ is $\mathcal{R}^\rceil := \{(ij|[n] \setminus ijK) : (ij|K) \in \mathcal{R}\} \subseteq \mathcal{A}_{[n]}$.

This involution turns a covariance matrix Σ which is Markovian for a graph G into a concentration matrix $K = \Sigma^{-1}$ such that $K_{ij} = 0$ for all $ij \notin G$. It also exchanges marginal and conditional [LM07, Lemma 1]:

Lemma 3.2.6. For any $\mathcal{R} \subseteq \mathcal{A}_{[n]}$ and $k \in [n]$ we have $\mathcal{R}^\rceil \setminus k = (\mathcal{R} \,/\, k)^\rceil$. If Σ is positive definite, then $\langle\!\langle \Sigma \rangle\!\rangle^\rceil = \langle\!\langle \Sigma^{-1} \rangle\!\rangle$. In particular, duals of realizable CI relations are realizable.

The final operation of interest is concatenating two independent Gaussian random vectors $(X_i)_{i\in S}$ and $(Y_i)_{i\in T}$ which are indexed by disjoint ground sets S and T. This is called *direct sum* in the structure theory of CI relations [Mat94]. The corresponding CI relation is as follows.

Definition 3.2.7. Let $\mathcal{R}$ and $\mathcal{R}'$ be two CI structures on disjoint ground sets S and T, respectively. Their *direct sum* is the CI structure

$$\begin{aligned}
\mathcal{R} \oplus \mathcal{R}' := \; & \{(ij|K) \in \mathcal{A}_{ST} : i \in S, j \in T\} \\
\cup \; & \{(ij|KL) \in \mathcal{A}_{ST} : (ij|K) \in \mathcal{R}, L \subseteq T\} \\
\cup \; & \{(ij|KL) \in \mathcal{A}_{ST} : (ij|K) \in \mathcal{R}', L \subseteq S\} \subseteq \mathcal{A}_{ST}.
\end{aligned}$$

On the level of covariance matrices, the direct sum imposes a block-diagonal structure with the summands on the diagonal. For $\Sigma \in \mathrm{PD}_S$ and $\Sigma' \in \mathrm{PD}_T$ let $\Sigma \oplus \Sigma' = \left(\begin{smallmatrix} \Sigma & 0 \\ 0 & \Sigma' \end{smallmatrix}\right)$. The following lemma is immediate.

Lemma 3.2.8. For $\Sigma \in \mathrm{PD}_S$ and $\Sigma' \in \mathrm{PD}_T$ we have $\langle\!\langle \Sigma \oplus \Sigma' \rangle\!\rangle = \langle\!\langle \Sigma \rangle\!\rangle \oplus \langle\!\langle \Sigma' \rangle\!\rangle$. In particular, the direct sum $\mathcal{R} \oplus \mathcal{R}'$ is realizable if and only if $\mathcal{R}$ and $\mathcal{R}'$ are both realizable. Moreover, the direct sum commutes with duality and minors.

Lemma 3.2.9. For $V = \mathcal{M}(\mathcal{R})$ and $V' = \mathcal{M}(\mathcal{R}')$ on disjoint ground sets S and T, respectively, the direct sum $U = V \oplus V'$ on ST is smooth if and only if V and V' are both smooth.

Proof. From the block-diagonal shape of matrices in U and Lemma 3.2.8 it follows that:

(1) V and V' are irreducible if and only if U is irreducible,

(2) $T_{\Sigma \oplus \Sigma'} U = T_\Sigma V \oplus T_{\Sigma'} V'$, and

(3) $\dim U = \dim V + \dim V'$.

Given irreducibility, smoothness means equality of the tangent space dimension to the model dimension. Therefore the smoothness conditions are equivalent. □

Remark 3.2.10. Any direct summand of a CI relation is a marginalization. Marginalizations in general need not preserve smoothness, as Example 3.2.16 below shows. But, one direction of Lemma 3.2.9 yields that a direct summand of a smooth model is smooth. Consequently, the non-zero entries in off-diagonal blocks are obstructions to smoothness of marginalizations.

The corresponding operations on graphs have been explained in [Mat97]. For a graph G and a vertex k, write $G \setminus k$ for the graph G where vertex k and all incident edges are deleted and $G \,/\, k$ for the graph G where vertex k is deleted and all vertices previously adjacent to k are connected to form a clique. The direct sum $G \oplus G'$ of graphs G, G' on disjoint ground sets S, T, respectively, consists of the disjoint unions of the vertex and edge sets of G and G' forming an undirected graph on ST. The operations on CI relations, positive definite matrices and graphs are all aligned:

Lemma 3.2.11. Let G be a graph on a vertex set $[n]$ and $k \in [n]$. We have $\langle\!\langle G \rangle\!\rangle \setminus k = \langle\!\langle G \,/\, k \rangle\!\rangle$, $\langle\!\langle G \rangle\!\rangle \,/\, k = \langle\!\langle G \setminus k \rangle\!\rangle$ and $\langle\!\langle G \oplus G' \rangle\!\rangle = \langle\!\langle G \rangle\!\rangle \oplus \langle\!\langle G' \rangle\!\rangle$. In particular, Markov relations are closed under forming minors and direct sums.

Duality has no counterpart in undirected graphical models. If undirected graphical models are referred to as "concentration models", their duals are "covariance models". Sometimes they are written with bidirected edges instead of undirected ones.

3.2.4 Double Markovian Models

In a double Markovian model a pair of graphs (G, H) specifies vanishing conditions via G on the concentration matrix and via H on the covariance matrix. This can be expressed using duality on ordinary Markov relations:

Definition 3.2.12. Let G and H be undirected graphs on vertex set $[n]$. Their *double Markov relation* is $\langle\!\langle G, H \rangle\!\rangle = \langle\!\langle G \rangle\!\rangle \cup \langle\!\langle H \rangle\!\rangle^{\rceil}$.

By Lemma 3.2.1, $\mathcal{M}(G, H) = \mathcal{M}(\langle\!\langle G, H \rangle\!\rangle)$ and the following is similar to Lemma 3.2.11.

Lemma 3.2.13. Let G and H be graphs on the vertex set $[n]$. Then

$$\langle\!\langle G, H \rangle\!\rangle^{\rceil} = \langle\!\langle H, G \rangle\!\rangle, \quad \langle\!\langle G, H \rangle\!\rangle \setminus k = \langle\!\langle G \,/\, k, H \setminus k \rangle\!\rangle, \quad \langle\!\langle G, H \rangle\!\rangle \,/\, k = \langle\!\langle G \setminus k, H \,/\, k \rangle\!\rangle.$$

Hence the class of double Markov relations is minor-closed.

All Markov relations are realizable and their models are smooth. These desirable properties fail to hold for double Markov relations. The CI relation $\langle\!\langle G, H \rangle\!\rangle$ gives only partial information about the geometry of the model. It is incomplete in the sense that $\langle\!\langle G, H \rangle\!\rangle$ does not necessarily contain all CI statements which are true on the model $\mathcal{M}(G, H)$. The remainder of this section contains examples of pathological behavior in the double Markovian setting.
Our first example of this is a double Markov relation which is not even a semigraphoid because it violates the following basic rule, which is satisfied by any probability distribution (Gaussian or not) and therefore must hold on the model $\mathcal{M}(G, H)$:

$$(12|) \wedge (13|2) \;\Rightarrow\; (13|) \wedge (12|3).$$

Example 3.2.14. Consider the two graphs $G =$  and $H =$ , where the vertices are labeled $1, 2, 3, 4$, clockwise starting at the top left. The vertex 4 can be replaced by any graph as long as it is not connected to 123. Because $(13|) \notin \langle\!\langle G \rangle\!\rangle$ and $(13|24) \notin \langle\!\langle H \rangle\!\rangle$, $(13|) \notin \langle\!\langle G, H \rangle\!\rangle$. However, $(13|2)$ holds in G and $(12|L)$ holds in H for any $L \subseteq 34$. But $(12|)$ and $(13|2)$ without $(13|)$ in $\langle\!\langle G, H \rangle\!\rangle$ contradict the semigraphoid property, because by contraction $(12|) \wedge (13|2) \Rightarrow (1, 23|)$ and then by decomposition $(1, 23|) \Rightarrow (13|)$ and by weak union $(1, 23|) \Rightarrow (12|3)$.

The semigraphoid closure of the CI structure in Example 3.2.14 (with 4 being an isolated vertex) is $\mathcal{A}_4$ and its model consists only of the identity matrix. As explained by Corollary 3.3.24, this happens because the edge sets of G and H are disjoint.
The next example uses the weak transitivity axiom $(12|) \wedge (12|3) \Rightarrow (13|) \vee (23|)$ to construct a family of graph pairs (G, H) whose model is non-smooth. The basic idea is that the logical OR in the conclusion of weak transitivity produces two components of the model. We verify that the CI model of $\{(12|), (12|3)\}$ has two irreducible components which intersect in the positive definite cone. In particular, this model is not smooth.

Example 3.2.15. Pick $G = H =$  which connects 1 with 2 via 3. Making only small changes to the following arguments, the fourth node can be replaced by any (possibly empty) graph that is not connected to 123. The CI structure is

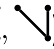
$$\langle\!\langle G, H \rangle\!\rangle = (\mathcal{A}_4 \setminus \{(13|L) : L \subseteq 24\}) \setminus \{(23|L) : L \subseteq 14\}\,.$$

In particular the formula $(12|) \wedge (12|3) \wedge \neg(13|) \wedge \neg(23|)$ holds for $\mathcal{R} = \langle\!\langle G, H \rangle\!\rangle$, which violates weak transitivity. Because it violates weak transitivity, $\mathcal{R}$ is not a gaussoid and not realizable by a regular Gaussian distribution. There are two gaussoid extensions of $\mathcal{R}$ to consider: $\mathcal{R}_1$ which adds $(13|L)$ and $\mathcal{R}_2$ which adds $(23|L)$, for all L, respectively, to $\mathcal{R}$. These extensions are isomorphic by exchanging the roles of 1 and 2. They are Markov relations corresponding to the complete graph with one edge removed. Hence, they are realizable and their models are irreducible and smooth by Proposition 3.2.2. However, the model of $\mathcal{R}$ consists of two copies of this smooth model, intersecting at the identity matrix, which makes it a singular point.

Unlike being double Markovian, the condition on a CI relation to be double Markovian *and* having a smooth model is not minor-closed. This is because for instance marginalizations of irreducible models can be reducible.

Example 3.2.16. The two graphs $G =$  and $H =$  impose the following relations on a positive definite 4×4 matrix $\Sigma = (\sigma_{ij})$ in their double Markovian model:

$$\sigma_{12} = \sigma_{14} = \sigma_{24} = \sigma_{34} = 0, \text{ from } \langle\!\langle H \rangle\!\rangle^{\rceil},$$
$$\sigma_{13}\sigma_{23}\sigma_{44} = 0, \ \sigma_{13}\sigma_{22}\sigma_{44} = 0, \text{ from } \langle\!\langle G \rangle\!\rangle.$$

The bounded model $\mathcal{M}_1(G, H) = \mathcal{M}(G, H) \cap \mathrm{PD}_{4,1}$ is a curve segment parametrized by $\sigma_{23} \in (-1, 1)$ since σ_{13} is forced to zero on $\mathrm{PD}_{4,1}$. The marginal CI structure on 123 is $\langle\!\langle G, H \rangle\!\rangle \setminus 4 = \langle\!\langle G \,/\, 4, H \setminus 4 \rangle\!\rangle = \langle\!\langle$ ,  $\rangle\!\rangle$, the one from Example 3.2.15. Its model has *two* components which intersect in the identity matrix and is therefore not smooth.
To understand this phenomenon one has to distinguish the model of the marginal CI structure, $\mathcal{M}(G/4, H\setminus 4)$, from the pointwise marginalization of $\mathcal{M}(G, H)$. What is discussed above is the former. It is reducible and properly contains the latter model as one of its two components.
The "unexpected" component of $\mathcal{M}(G \,/\, 4, H \setminus 4)$ arises from semi-definite matrices on the boundary of $\mathcal{M}(G, H)$ which become regular after marginalization. Namely, the two equations

$$\sigma_{13}\sigma_{23}\sigma_{44} = 0, \ \sigma_{13}\sigma_{22}\sigma_{44} = 0$$

imply $\sigma_{13} = 0$ on positive definite matrices, but there are semi-definite solutions to them where (1) $\sigma_{22} = 0$ and $\sigma_{23} = 0$, or (2) $\sigma_{44} = 0$, and σ_{13} and σ_{23} are arbitrary. Thus there are three types of solutions:

$$\begin{pmatrix} \sigma_{11} & 0 & 0 & 0 \\ 0 & \sigma_{22} & \sigma_{23} & 0 \\ 0 & \sigma_{23} & \sigma_{33} & 0 \\ 0 & 0 & 0 & \sigma_{44} \end{pmatrix}, \quad \begin{pmatrix} \sigma_{11} & 0 & \sigma_{13} & 0 \\ 0 & 0 & 0 & 0 \\ \sigma_{13} & 0 & \sigma_{33} & 0 \\ 0 & 0 & 0 & \sigma_{44} \end{pmatrix}, \quad \begin{pmatrix} \sigma_{11} & 0 & \sigma_{13} & 0 \\ 0 & \sigma_{22} & \sigma_{23} & 0 \\ \sigma_{13} & \sigma_{23} & \sigma_{33} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The first type is visible in PD_4 and in $\mathcal{M}(G, H)$ and its marginalization forms one of the components of $\mathcal{M}(G \,/\, 4, H \setminus 4)$. The second type of solutions is not visible in the marginalization because it contains no PD_3 matrices. Marginalizing 4 from the CI structure removes the last row and column of these matrices and imposes additional constraints, in this case $\sigma_{13}\sigma_{23} = 0$. This turns the third type of solution positive definite and reduces the dimension by one. In this case, both the first and second component arise from this 2-dimensional boundary component.

Next, we present a family of double Markov relations which are realizable but whose model is singular at the identity matrix. Moreover, this gives an infinite family of realizable, non-smooth models all of whose proper minors are realizable and smooth.

Example 3.2.17. By [DX10, Proposition 4.2], the model of $\mathcal{R} = \{(12|), (12|[n] \setminus 12)\}$ is singular at the identity matrix. It can be represented by the double Markov relation $\langle\!\langle G, G \rangle\!\rangle$ where G is the graph on $[n]$ whose only non-edge is 12. If $k = 1, 2$, the marginalization and conditioning of G by k are both the complete graph, so their double Markovian model is the entire cone PD_n and smooth. If $k \in [n] \setminus 12$, then G / k is a complete graph and $G \setminus k$ is another graph on $[n] \setminus k$ whose only non-edge is 12. Since a complete graph imposes no relations, we find that $\langle\!\langle G, G \rangle\!\rangle \setminus k = \langle\!\langle G \setminus k \rangle\!\rangle$ is a Markov relation and thus its model is smooth by Proposition 3.2.2. It is clear that all further minors of these two cases yield smooth models as well. This gives infinitely many examples of minor-minimal, non-smooth double Markovian CI structures. Explicit computation in `Macaulay2` for $n = 4, 5, 6$ verifies that (the Zariski closure of) the model is an irreducible variety and this shows that singularities are not always caused by an intersection of irreducible components. In Example 3.4.2 we examine the singular locus of $\mathcal{M}(G, G)$ further. It turns out to be another conditional independence model.

We have no general proof for the irreducibility of the models in the preceding example when $n \geq 7$. It was shown in [Boe22, Lemma 6.4] that the CI structures $\{(12|), (12|[n] \setminus 12)\}$ are realizable by Gaussian distributions for all $n \geq 4$. Thus the models are irreducible in the finite "lattice of CI relations" studied by Drton and Xiao in [DX10, Section 2.1]. This is a coarsening of the lattice of the usual Zariski topology induced on PD_n, so CI-irreducibility provides a necessary condition for Zariski-irreducibility.
There is a Galois connection between CI structures $\mathcal{R} \subseteq \mathcal{A}_{[n]}$ and Gaussian CI models $\mathcal{M}(\mathcal{R}) \subseteq \mathrm{PD}_n$. The closed CI structures under this connection are termed *complete relations* by Drton and Xiao, and their [DX10, Theorem 2.2] characterizes them as the intersections of realizable gaussoids. The completion of a double Markovian relation $\langle\!\langle G, H \rangle\!\rangle$ adds all CI statements which hold on every matrix in its model $\mathcal{M}(G, H)$. For single Markovian relations $\langle\!\langle G \rangle\!\rangle$ is always complete because it is realizable by Proposition 3.2.2 and its elements can be read off from G.
On the other hand, Example 3.2.14 exhibits a pair of graphs G and H whose CI structure does not satisfy the semigraphoid property. Since the semigraphoid axiom holds for every Gaussian distribution, it follows that $\langle\!\langle G, H \rangle\!\rangle$ need not be complete. Moreover, even if the set $\langle\!\langle G, H \rangle\!\rangle$ is closed under the compositional graphoid axioms, which hold for all Gaussians (called *semigaussoid axioms* by Drton and Xiao), it may still be incomplete:

Example 3.2.18. Let $G =$  and $H =$  both be 4-cycles. We have $\langle\!\langle G \rangle\!\rangle = \{(13|24), (24|13)\}$ and $\langle\!\langle H \rangle\!\rangle^{\rceil} = \{(12|), (34|)\}$. Their union $\mathcal{R}$ is the set of antecedents to an instance of inference rule [LM07, Lemma 10 (17)]:

$$(12|) \wedge (34|) \wedge (13|24) \wedge (24|13) \Rightarrow (13|)$$

This formula is valid for all regular Gaussians. Since $\mathcal{R}$ is not closed under this rule, it is not realizable by a positive definite matrix. It is not complete either because the inference rule (17) is a *Horn clause*, i.e., it has a unique consequence $(13|)$ which every realizable superset of $\mathcal{R}$ and hence $\mathcal{R}$ as their intersection would have to contain if it were complete.

However, $\mathcal{R}$ is a gaussoid (therefore closed under the compositional graphoid axioms) and it is realizable by a complex matrix with non-vanishing principal minors. Consequently, working with equations in a computer algebra system like `Macaulay2`, one cannot deduce any further CI statements from $\langle\!\langle G, H \rangle\!\rangle$. Positive definiteness has to be taken into account.
The completion of $\mathcal{R}$ in the positive definite setting can be computed as its closure under the semigaussoid axioms [LM07, Definition 1 (7)–(9)] and the higher inference rules [LM07, Lemma 10 (17)–(21)] of Lněnička and Matúš. It equals $\overline{\mathcal{R}} = \mathcal{A}_4 \setminus (\{(14|L) : L \subseteq 23\} \cup \{(23|L) : L \subseteq 14\})$. This structure is a self-dual Markov relation and hence can be written as the relation of two identical graphs (J, J) such that $\langle\!\langle J \rangle\!\rangle = \langle\!\langle J \rangle\!\rangle^{\rceil}$. Indeed $J = G \cap H$ gives $\langle\!\langle J, J \rangle\!\rangle = \overline{\mathcal{R}}$. This shows that nevertheless $\mathcal{M}(G, H) = \mathcal{M}(\overline{\mathcal{R}}) = \mathcal{M}(J, J) = \mathcal{M}(G \cap H)$ is smooth.

Question 3.2.19. Is there a combinatorial criterion similar to separation in undirected graphs to derive a complete set of valid CI statements for $\mathcal{M}(G, H)$?

A first step is Corollary 3.3.24 where the triviality of the model is characterized by disjointness of the edge sets. In this case, every CI statement is a consequence of the statements in $\langle\!\langle G, H \rangle\!\rangle$.

3.3 Geometry of the Models $\mathcal{M}(G, H)$

Our study of the smoothness of double Markovian models starts with the known observation that one may as well work with the bounded set of correlation matrices.

Lemma 3.3.1 ([DX10, Lemma 3.2])**.** The set $\mathcal{M}(\mathcal{R})$ is a smooth submanifold of PD_n if and only if $\mathcal{M}_1(\mathcal{R})$ is a smooth submanifold.

Proof. The map $\mathrm{PD}_n \to \mathbb{R}^n_{>0} \times \mathrm{PD}_{n,1}$, sending Σ to the pair consisting of its diagonal and its associated correlation matrix, is a diffeomorphism. Consider the following commutative diagram:

$$
\begin{tikzcd}
\mathrm{PD}_n \arrow[r, "\cong"] & \mathbb{R}^n_{>0} \times \mathrm{PD}_{n,1} & \{(1,1,\dots,1)\} \times \mathrm{PD}_{n,1} \arrow[l, hook'] \arrow[r, equal] & \mathrm{PD}_{n,1} \\
\mathcal{M}(\mathcal{R}) \arrow[u, hook] \arrow[r, "\cong"'] & \mathbb{R}^n_{>0} \times \mathcal{M}_1(\mathcal{R}) \arrow[u, hook] & \{(1,1,\dots,1)\} \times \mathcal{M}_1(\mathcal{R}) \arrow[u, hook] \arrow[l, hook'] \arrow[r, equal] & \mathcal{M}_1(\mathcal{R}). \arrow[u, hook]
\end{tikzcd}
$$

All maps are topological embeddings or homeomorphisms, and the upper row consists of diffeomorphisms and embeddings of smooth submanifolds. If $\mathcal{M}(\mathcal{R})$ is a smooth submanifold, $\mathbb{R}^n_{>0} \times \mathcal{M}_1(\mathcal{R})$ inherits a smooth manifold structure making the second vertical inclusion an embedding of a smooth submanifold. Now, the product $\{(1, 1, \dots, 1)\} \times \mathcal{M}_1(\mathcal{R})$ is the preimage of a regular value under the projection map $\mathbb{R}^n_{>0} \times \mathcal{M}_1(\mathcal{R}) \to \mathbb{R}^n_{>0}$. This again can be seen from the composition

$$\mathcal{M}(\mathcal{R}) \hookrightarrow \mathrm{PD}_n \xrightarrow{\cong} \mathbb{R}^n_{>0} \times \mathrm{PD}_{n,1} \to \mathbb{R}^n_{>0},$$

sending $\Sigma \in \mathcal{M}(\mathcal{R})$ to its diagonal. The claim follows because for each $\Sigma \in \mathcal{M}(\mathcal{R})$ having only ones on the diagonal and an arbitrary positive diagonal matrix D, we

have $D\Sigma D \in \mathcal{M}(\mathcal{R})$ (by the proof of [DX10, Lemma 3.1]). Choosing an appropriate smooth path of diagonal matrices passing through the identity matrix, we obtain every possible tangent vector in $T_{(1,1,\ldots,1)}\mathbb{R}^n_{>0}$.
If $\mathcal{M}_1(\mathcal{R})$ is a smooth submanifold of PD_n, then also of $\mathrm{PD}_{n,1}$, and the product $\mathbb{R}^n_{>0} \times \mathcal{M}_1(\mathcal{R})$ inherits a canonical smooth structure making the second vertical inclusion an embedding of a smooth submanifold. Then clearly the induced smooth structure on $\mathcal{M}(\mathcal{R})$ makes the leftmost vertical inclusion an embedding of a smooth submanifold as well. $\square$

The proof shows that if $\mathcal{M}_1(\mathcal{R})$ is a smooth submanifold of PD_n, then it is in fact a smooth submanifold of $\mathcal{M}(\mathcal{R})$ via the inclusion. The following lemma is easily verified.

Lemma 3.3.2 ([DX10, Lemma 3.3])**.** There is a self-inverse diffeomorphism $\mathrm{PD}_{n,1} \to \mathrm{PD}_{n,1}$, given by matrix inversion followed by forming the correlation matrix, mapping $\mathcal{M}_1(\mathcal{R})$ onto $\mathcal{M}_1(\mathcal{R}^\rceil)$.

In particular, $\mathcal{M}(\mathcal{R})$ and $\mathcal{M}_1(\mathcal{R})$ are smooth if and only if $\mathcal{M}(\mathcal{R}^\rceil)$ and $\mathcal{M}_1(\mathcal{R}^\rceil)$ are. As an image of a linear space under the inversion map, any graphical model $\mathcal{M}(G)$ is smooth. Moreover, the bijective morphism of semi-algebraic sets $\mathbb{R}^n_{>0} \times \mathcal{M}_1(\mathcal{R}) \to \mathcal{M}(\mathcal{R}), (D, \Sigma) \mapsto D\Sigma D$ can be used to show that the dimensions of $\mathcal{M}_1(\mathcal{R})$ and $\mathcal{M}(\mathcal{R})$ always differ by n.

3.3.1 Basics from Real Algebraic Geometry

We collect several foundational definitions and results from [BCR98] and refer to this textbook for an extensive treatment.
A *real algebraic set* $Z \subseteq \mathbb{R}^n$ is the vanishing set $\mathcal{V}(S)$ of a collection $S \subseteq \mathbb{R}[x_1, \ldots, x_n]$ of polynomials, and S may be replaced by the ideal (S) it generates. Real algebraic sets are the closed sets of the *Zariski topology* on $\mathbb{R}^n$. If $\Theta \subseteq \mathbb{R}^n$ is any subset, its ideal is $\mathcal{I}(\Theta) = \{f \in \mathbb{R}[x_1, \ldots, x_n] : f(x) = 0, \text{ for all } x \in \Theta\}$. The real algebraic set of $\mathcal{I}(\Theta)$ is the *Zariski closure* $\overline{\Theta}$ *of* Θ. Every irreducible component of the Zariski closure of Θ in $\mathbb{R}^n$ intersects Θ. If Z is irreducible in this topology, Z is a *real algebraic variety.* A set of the form

$$\Theta = \{x \in \mathbb{R}^n : f_1(x) = \cdots = f_r(x) = 0, g_1(x) > 0, \ldots, g_s(x) > 0\},$$

where $f_i, g_j \in \mathbb{R}[x_1, \ldots, x_n]$ are real polynomials is a *basic semi-algebraic set.* A finite union of basic semi-algebraic sets in a fixed $\mathbb{R}^n$ is a *semi-algebraic set.* The *dimension* of a semi-algebraic set $\Theta \subseteq \mathbb{R}^n$ is the dimension of its coordinate ring $\mathbb{R}[x_1, \ldots, x_n]/\mathcal{I}(\Theta)$. The dimension of Θ then equals the dimension of its Zariski closure as $\mathcal{I}(\Theta) = \mathcal{I}(\overline{\Theta})$. A semi-algebraic set $\Theta \subseteq \mathbb{R}^n$ is *semi-algebraically connected* if for every two *disjoint* semi-algebraic subsets $A, B \subseteq \Theta$ which are closed in Θ and satisfy $A \cup B = \Theta$, one has $A = \Theta$ or $B = \Theta$. According to [BCR98, Theorem 2.4.5], a semi-algebraic set $\Theta \subseteq \mathbb{R}^n$ is semi-algebraically connected if and only if it is connected with respect to the euclidean topology on $\mathbb{R}^n$.

Definition 3.3.3. For a real algebraic set $V \subseteq \mathbb{R}^n$ with vanishing ideal $\mathcal{I}(V) = (f_1, \dots, f_r)$, the *Zariski tangent space T_pV at $p \in V$* is the kernel of the Jacobian matrix
$$J_p = \left(\frac{\partial f_i}{\partial x_j}(p) \right)_{\substack{i=1,\dots,r, \\ j=1,\dots,n}}.$$
An algebraic set V is *smooth at $p \in V$* if p is contained in some irreducible component Z of V with $\dim(Z) = \dim(T_pV)$. Finally, V is *smooth* if it is smooth at every point $p \in V$.

The principal ideal theorem implies that $\dim(T_pV) \geq \dim(Z)$ for all irreducible components Z of V containing p. If $p \in V$ is a smooth point and Z an irreducible component of V containing p with $\dim(Z) = \dim(T_pV)$, then $\mathrm{rk}(J_p) = n - \dim(Z)$ and Z is the *only* irreducible component of V containing p. The latter is a consequence of the fact that regular local rings are integral domains [Kem11, Corollary 13.6]. Another general fact is that the rank of the Jacobian at p does not depend on the set of generators of $\mathcal{I}(V)$. This follows from the fact that adding an element of the ideal $\mathcal{I}(V)$ to a list of generators does not change the rank because the gradient of the additional polynomial is linearly dependent on the gradients of the generators at every point $p \in V$. This reasoning also proves the following.

Lemma 3.3.4. Let $V \subseteq \mathbb{R}^n$ be a real algebraic set with generators $f_1, \dots, f_r$ of $\mathcal{I}(V)$ and let $g_1, \dots g_s \in \mathcal{I}(V)$ be arbitrary. Then for all $p \in V$ we have
$$\mathrm{rk} \left(\frac{\partial f_i}{\partial x_j}(p) \right)_{\substack{i=1,\dots,r, \\ j=1,\dots,n}} \geq \mathrm{rk} \left(\frac{\partial g_i}{\partial x_j}(p) \right)_{\substack{i=1,\dots,s, \\ j=1,\dots,n}}.$$

The natural application of Lemma 3.3.4 is to bound $\dim(V)$ when no generating set of $\mathcal{I}(V)$ is known.

Similarly to Definition 3.3.3, a semi-algebraic set $\Theta \subseteq \mathbb{R}^n$ is *smooth at $p \in \Theta$* if p is a smooth point of the Zariski closure $\overline{\Theta}$. If Θ is smooth at every point, then Θ is *smooth*. The set of smooth points of Θ is its *smooth locus*, denoted Θ_{sm}. The smooth locus V_{sm} of a non-empty real algebraic set V is a non-empty Zariski-open subset of V by [BCR98, Proposition 3.3.14] and is always Zariski-dense in V because V_{sm} intersects every irreducible component of V. We stress that our definition of the smooth locus of a real algebraic set V differs from that of [BCR98, Notation 3.3.13] where all irreducible components of dimension less than $\dim(V)$ are counted as singular. By our definition, instead, a point $p \in V$ is smooth if and only if the local ring $(\mathbb{R}[x_1, \dots, x_n]/\mathcal{I}(V))_{\mathfrak{m}_p}$ is a regular local ring (of any dimension), which is natural from a scheme-theoretic point of view. Importantly, the sheaf of regular functions of V in [BCR98, Definition 3.2.1] is different from its scheme-theoretic structure sheaf. However, the stalks at all points $p \in V$ coincide for both sheaves [BCR98, p. 65]. Therefore, the two definitions of the smooth locus of V coincide if and only if V is equidimensional, so in particular if V is irreducible.
The smooth locus V_{sm} is a smooth submanifold of $\mathbb{R}^n$ (possibly disconnected with connected components of different dimensions) by [BCR98, Proposition 3.3.11]. From this it follows that, if $\Theta \subseteq \mathbb{R}^n$ is a basic semi-algebraic set, then Θ_{sm} is a smooth

submanifold of $\mathbb{R}^n$ by viewing $\Theta_{\mathrm{sm}} = \Theta \cap (\overline{\Theta})_{\mathrm{sm}}$ as an open subset (in the euclidean topology) of $(\overline{\Theta})_{\mathrm{sm}}$.

3.3.2 Smoothness of the Models $\mathcal{M}(G, H)$

The geometry of double Markovian models is best understood in terms of semi-algebraic sets. We can identify

$$\mathcal{M}(G, H) = \mathcal{M}(G) \cap \mathcal{M}(H)^{-1}$$

with

$$\widetilde{\mathcal{M}(G, H)} := \left\{ (\Sigma, \Sigma^{-1}) : \Sigma_{ij} = 0 \text{ for all } ij \notin H \text{ and } (\Sigma^{-1})_{kl} = 0 \text{ for all } kl \notin G \right\},$$

and the latter is a smooth submanifold of $\mathrm{PD}_n \times \mathrm{PD}_n$ if and only if $\mathcal{M}(G, H)$ is a smooth submanifold of PD_n, as follows from the diagram:

$$\begin{tikzcd} \mathrm{PD}_n \arrow[r, hook, "\mathrm{id} \times \mathrm{inv}"] & \mathrm{PD}_n \times \mathrm{PD}_n \\ \mathcal{M}(G, H) \arrow[u, hook] \arrow[r, "\cong"] & \widetilde{\mathcal{M}(G, H)}. \arrow[u, hook] \end{tikzcd}$$

Because of this, it suffices to study the smoothness of $\mathcal{M}(G, H)$.
Both $\mathcal{M}(G)$ and $\mathcal{M}(H)^{-1}$ are smooth submanifolds of PD_n. It is sufficient for the intersection to be smooth that the intersection be transverse at every common point [GP10, Chapter 1 §5], meaning that the sum of the tangent spaces equals the tangent space of the ambient manifold. This criterion yields Theorem 3.3.8 below.

Proposition 3.3.5. A basis of the tangent space $T_P\mathcal{M}(G)$ is given by the matrices

$$M^{ij} := P^i \cdot P_j + P^j \cdot P_i, \qquad \text{for } i = j \text{ or } ij \in E_G,$$

where P^i is the i-th column of P and P_j is the j-th row of P. A basis of the tangent space $T_P\mathcal{M}(H)^{-1}$ consists of $E^{ij} := E_{ij} + E_{ji}$ for $i = j$ or $ij \in E_H$, where E_{ij} is the $n \times n$ matrix having a 1 at the (i, j)-th position and zeros everywhere else.

Proof. We view both tangent spaces as subspaces of $T_P \mathrm{PD}_n = T_P \operatorname{Sym}^2(\mathbb{R}^n) \cong \operatorname{Sym}^2(\mathbb{R}^n)$. As $\mathcal{M}(H)^{-1}$ is just the intersection of PD_n with the linear subspace of $\operatorname{Sym}^2(\mathbb{R}^n)$ given by the vanishing of the ij-entries for all non-edges ij of H, the second claim follows. For $T_P\mathcal{M}(G)$, we use the differential of the matrix inversion $\mathrm{inv}\colon \mathrm{PD}_n \to \mathrm{PD}_n$. In coordinates $X = (x_{ij})$ it is

$$\frac{\partial\, \mathrm{inv}}{\partial x_{ij}}(X) = \frac{\partial (X^{-1})}{\partial x_{ij}} = -X^{-1} E^{ij} X^{-1} = -(X^{-1})^i \cdot (X^{-1})_j.$$

Now $T_{P^{-1}}\mathcal{M}(H)^{-1}$ is generated by E^{ij} for $i = j$ or $ij \in E_H$, and the differential at P^{-1} maps $T_{P^{-1}}\mathcal{M}(H)^{-1}$ isomorphically onto $T_P\mathcal{M}(G)$, and the image of E^{ij} is $-M^{ij}$. $\square$

Remark 3.3.6. The proof also shows $\dim(\mathcal{M}(G)) = |E_G| + n$. Moreover, $G \cup H = K_n$ if and only if $T_{\mathbb{1}_n}\mathcal{M}(G) + T_{\mathbb{1}_n}\mathcal{M}(H)^{-1} = T_{\mathbb{1}_n}PD_n$, that is, if and only if the intersection is transverse at $P = \mathbb{1}_n$. In particular, under this assumption $\mathcal{M}(G,H) = \mathcal{M}(G) \cap \mathcal{M}(H)^{-1}$ is smooth at $\mathbb{1}_n$ and thus at all positive diagonal matrices by means of the map $P \mapsto DPD$ for a fixed positive diagonal matrix D. This is a diffeomorphism of PD_n, restricting to a homeomorphism $\mathcal{M}(G,H) \to \mathcal{M}(G,H)$, mapping smooth points to smooth points and $\mathbb{1}_n$ to D^2.

Remark 3.3.7. For $H = K_n$, the proof of Proposition 3.3.5 implies that the matrix

$$M := (M_{kl}^{ij})_{kl,ij \in \left[\binom{n+1}{2}\right]}$$

is symmetric and invertible whenever P is symmetric and invertible – a fact that might be difficult to prove directly. If P is positive definite, M is the inverse information matrix appearing, for example, in the Cramér–Rao inequality and in [Drt09, Example 2.2].

The following theorem was the original motivation for our paper [BKKR23].

Theorem 3.3.8. If $G \cup H = K_n$ is the complete graph, then the model $\mathcal{M}(G,H)$ is smooth. In fact, the intersection of $\mathcal{M}(G)$ and $\mathcal{M}(H)^{-1}$ is transverse at every intersection point.

Proof. The inverse information matrix M at a positive definite matrix $P \in \mathrm{PD}_n$ is positive definite. The matrix $\left(M_{kl}^{ij}\right)_{kl,ij \in E_G \setminus E_H}$ is its principal $(E_G \setminus E_H \times E_G \setminus E_H)$-submatrix and therefore also positive definite and in particular invertible. This implies $T_P\mathcal{M}(G) + T_P\mathcal{M}(H)^{-1} = T_P \,\mathrm{PD}_n$, so the intersection of the smooth submanifolds $\mathcal{M}(G)$ and $\mathcal{M}(H)^{-1}$ of PD_n is transverse at every common point P. Therefore, the intersection $\mathcal{M}(G,H) = \mathcal{M}(G) \cap \mathcal{M}(H)^{-1}$ is a smooth submanifold of PD_n (of constant dimension) by [GP10, Chapter 1 §5]. $\square$

Remark 3.3.9. Multivariate centered Gaussian random vectors form a regular exponential family with mean parameter Σ and natural parameter $K = \Sigma^{-1}$ [Sun19, Chapter 3]. According to [Sun19, Corollary 3.17], a mixed parametrization

$$(\sigma_{ij}, k_{st})_{ij \in A, st \in B}$$

with $A \cup B = E_{K_n}$ and $A \cap B = \emptyset$ is a valid parametrization for the exponential family. In the situation of Theorem 3.3.8, when $G \cup H = K_n$, the non-edges of G and H are disjoint. Therefore one could pick $A, B \subseteq E_{K_n}$ such that the non-edges of G are contained in B and the non-edges of H in A. However, when zero constraints on both mean and natural parameters are imposed, the result is in general not a regular exponential family. Therefore smoothness results like Theorem 3.3.8 do not simply follow from general theory.

In Theorem 3.3.8 we have $\operatorname{codim}(\mathcal{M}(G,H)) = \operatorname{codim}(\mathcal{M}(G)) + \operatorname{codim}(\mathcal{M}(H))$ and so, regarding dimensions, by Remark 3.3.6,

$$\begin{aligned}\dim(\mathcal{M}(G,H)) &= \frac{n^2+n}{2} - \left(\frac{n^2-n}{2} - |E_G| + \frac{n^2-n}{2} - |E_H|\right)\\ &= |E_G| + |E_H| - \frac{n^2-3n}{2}\\ &= |E_G \cap E_H| + n,\end{aligned}$$

where we have used $|E_G| + |E_H| = \binom{n}{2} + |E_G \cap E_H|$. The intersection with $\mathrm{PD}_{n,1}$ satisfies

$$\dim(\mathcal{M}_1(G,H)) = \dim(\mathcal{M}(G,H) \cap \mathrm{PD}_{n,1}) = \dim(\mathcal{M}(G,H)) - n = |E_G \cap E_H|.$$

If one is not in the favorable situation $G \cup H = K_n$, the dimension computation becomes more involved. Let G and H now be arbitrary graphs on $[n]$, and denote by E_G^{c} the edge complement $\left[\binom{n}{2}\right] \setminus E_G$ and similarly for E_H. The following lemma is a technical core for dimension bounds.

Lemma 3.3.10. With $\Sigma = (\sigma_{st})_{st \in \left[\binom{n+1}{2}\right]}$, let $f_{ij} = \sigma_{ij}$ for $ij \in E_H^{\mathsf{c}}$ and $g_{kl} = \det(\Sigma_{[n]\setminus k,[n]\setminus l})$ for $kl \in E_G^{\mathsf{c}}$. Consider the Jacobian matrices

$$J_H = \left(\frac{\partial f_{ij}}{\partial \sigma_{st}}\right)_{ij \in E_H^{\mathsf{c}}, st \in \left[\binom{n+1}{2}\right]} \quad \text{and} \quad J_G = \left(\frac{\partial g_{kl}}{\partial \sigma_{st}}\right)_{kl \in E_G^{\mathsf{c}}, st \in \left[\binom{n+1}{2}\right]}.$$

For every $\Sigma \in \mathcal{M}(G,H)$, define the $(|E_G^{\mathsf{c}}| + |E_H^{\mathsf{c}}|) \times \binom{n+1}{2}$ matrix

$$\widetilde{J}_\Sigma := \begin{pmatrix} J_G \\ J_H \end{pmatrix}.$$

Then $\operatorname{rk} \widetilde{J}_\Sigma \geq \binom{n}{2} - |E_G \cap E_H|$.

Proof. The kernel of $\widetilde{J}_\Sigma$ is the intersection of the kernels of J_G and J_H. The Zariski closures $\overline{\mathcal{M}(G)}$ and $\overline{\mathcal{M}(H)^{-1}}$ in $\mathrm{Sym}^2(\mathbb{R}^n) \cap \mathrm{GL}(\mathbb{R}^n)$ of $\mathcal{M}(G)$ and $\mathcal{M}(H)^{-1}$ are both irreducible varieties. The Zariski tangent spaces have been computed in Proposition 3.3.5. Using $\dim(U \cap V) = \dim(U) + \dim(V) - \dim(U+V)$ for finite-dimensional vector spaces U and V inside a common vector space, we compute

$$\begin{aligned}\dim(\ker(\widetilde{J}_\Sigma)) &= \dim(\mathrm{span}(E^{ij} : ij \in E_H \text{ or } i=j) \cap \mathrm{span}(M^{kl} : kl \in E_G \text{ or } k=l))\\ &= (|E_G| + n) + (|E_H| + n)\\ &\quad - \dim(\mathrm{span}(E^{ij} : ij \in E_H \text{ or } i=j) + \mathrm{span}(M^{kl} : kl \in E_G \text{ or } k=l))\\ &\leq |E_G \cap E_H| + n.\end{aligned}$$

In the last step we used that the dimension of the sum of the two vector spaces is at least $|E_H \cup (E_G \setminus E_H)| + n = |E_G \cup E_H| + n$ because the matrix $(M^{kl}_{st})_{kl,st \in E_G \setminus E_H}$ is a principal submatrix of the inverse of the information matrix and therefore invertible. ◻

Theorem 3.3.11. We have $\dim(\mathcal{M}(G,H)) \leq |E_G \cap E_H| + n$.

Proof. The polynomials f_{ij} and g_{kl} from Lemma 3.3.10 lie in the vanishing ideal $\mathcal{I}(\mathcal{M}(G,H)) \subseteq \mathbb{R}[\sigma_{st} : st \in \left[\binom{n+1}{2}\right]]$, and hence in the prime ideal of every irreducible component Z of the Zariski closure $\mathcal{V}(\mathcal{I}(\mathcal{M}(G,H)))$ inside the affine space $\mathrm{Sym}^2(\mathbb{R}^n)$. Then the Jacobian matrix J_Σ at $\Sigma \in Z$ of a generating set of $\mathcal{I}(Z)$ satisfies $\mathrm{rk}\, J_\Sigma \geq \mathrm{rk}\, \widetilde{J}_\Sigma$ by Lemma 3.3.4. By Lemma 3.3.10, for $\Sigma \in \mathcal{M}(G,H)$, $\mathrm{rk}\, \widetilde{J}_\Sigma \geq \binom{n}{2} - |E_G \cap E_H|$. Then Krull's principal ideal theorem implies that the Zariski tangent space satisfies $\dim(T_\Sigma Z) \geq \dim(Z)$, and so

$$\dim(Z) \leq \dim(T_\Sigma Z) = \binom{n+1}{2} - \mathrm{rk}(J_\Sigma) \leq \binom{n+1}{2} - \mathrm{rk}(\widetilde{J}_\Sigma) \leq |E_G \cap E_H| + n. \qquad \square$$

Corollary 3.3.12. We have $\dim(\mathcal{M}_1(G,H)) \leq |E_G \cap E_H|$.

Remark 3.3.13. The inequalities in Corollary 3.3.12 and Theorem 3.3.11 can be strict. Example 3.2.15 contains a model $\mathcal{M}_1(G,H)$ of dimension 1 with $|E_G \cap E_H| = 2$. It is reducible with two irreducible components which intersect only in the identity matrix $\mathbb{1}_4$.

It is not hard to see that $E_G \cap E_H = \emptyset$ if $\mathcal{M}_1(G,H)$ is zero-dimensional, i.e., a union of finitely many points (including $\mathbb{1}_n$). Indeed, if $E_G \cap E_H \neq \emptyset$, then $\dim(\mathcal{M}_1(G,H)) \geq 1$ because after some permutation of $[n]$ one can assume that $12 \in E_G \cap E_H$. Then

$$\Sigma = \begin{pmatrix} 1 & \sigma_{12} & 0 & \dots & 0 \\ \sigma_{12} & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

lies in $\mathcal{M}_1(G,H)$ for any $\sigma_{12} \in (-1,1)$. Corollary 3.3.24 below strengthens this further. In fact, $E_G \cap E_H = \emptyset$ if and only if the model consists only of $\mathbb{1}_n$.

Lemma 3.3.14. Let $V := \overline{\mathcal{M}(G,H)}$ and $\Sigma \in \mathcal{M}(G,H)$. Then,

$$\dim(T_\Sigma V) \leq |E_G \cap E_H| + n.$$

Proof. The Zariski tangent space $T_\Sigma V$ is the kernel of the Jacobian matrix J_Σ at $\Sigma \in \mathcal{M}(G,H)$ of a generating set of $\mathcal{I}(V)$. By Lemma 3.3.4 and Lemma 3.3.10, we have $\mathrm{rk}\, J_\Sigma \geq \mathrm{rk}\, \widetilde{J}_\Sigma \geq \binom{n}{2} - |E_G \cap E_H|$ which yields the claimed inequality. $\square$

Theorem 3.3.15. Every connected component of $\mathcal{M}(G,H)$ of dimension $|E_G \cap E_H| + n$ is smooth and has irreducible Zariski closure.

Proof. Let $V := \overline{\mathcal{M}(G,H)} \subseteq \mathrm{Sym}^2(\mathbb{R}^n)$. Let M be a connected component of $\mathcal{M}(G,H)$ and Z an irreducible component of its Zariski closure $\overline{M}$ with $\dim(Z) = |E_G \cap E_H| + n$. Then at every point $\Sigma \in M \cap Z$, we have

$$|E_G \cap E_H| + n = \dim(Z) \leq \dim(T_\Sigma Z) \leq \dim(T_\Sigma V) \leq |E_G \cap E_H| + n,$$

hence all inequalities are equalities, proving that the local rings $\mathcal{O}_{V,\Sigma}$ with $\Sigma \in M \cap Z$ are regular. Regular local rings are integral domains by [Kem11, Corollary 13.6]. Hence, for every $\Sigma \in M \cap Z$, there is only one irreducible component of $\overline{M}$ containing Σ, so Z does not intersect any other irreducible component of $\overline{M}$ inside M. Therefore, as M is connected, $\overline{M} = Z$ is irreducible, and $M = M \cap Z$ is smooth. □

Remark 3.3.16. The proof of Theorem 3.3.15 also shows that a connected component of $\mathcal{M}_1(G,H)$ of dimension $|E_G \cap E_H|$ is smooth and hence has irreducible Zariski closure. Proposition 3.3.31 (8) contains a smooth model $\mathcal{M}_1(G,H)$ on 4 vertices of dimension $3 = |E_G \cap E_H|$ with $G \cup H \neq K_4$, so the converse of Theorem 3.3.8 is false, see however Remark 3.3.32. Even when $G \cup H \neq K_n$, Theorem 3.3.15 still provides a sufficient criterion for smoothness. In fact, we know of no example of a smooth model $\mathcal{M}(G,H)$ (resp. $\mathcal{M}_1(G,H)$) having dimension less than $|E_G \cap E_H| + n$ (resp. $|E_G \cap E_H|$).

We now move on to approximations of the vanishing ideal $\mathcal{I}(\mathcal{M}_1(G,H))$ of double Markovian models. Ordinary Gaussian graphical models have rational parametrizations and their vanishing ideals are prime. Vanishing ideals of double Markovian models need not be prime. They arise from conditional independence ideals by removing components whose varieties do not intersect PD_n and taking the radical. We do not expect double Markovian CI ideals to be radical in general.
For the rest of the section we restrict to the normalized variance case, i.e., $\mathcal{M}_1(G,H)$ as opposed to $\mathcal{M}(G,H)$. This removes duplications from the statements. In most cases only small changes are necessary to change a result for $\mathcal{M}_1(G,H)$ into one for $\mathcal{M}(G,H)$.

Definition 3.3.17. Let G and H be two graphs on $[n]$ and $\Sigma = (\sigma_{ij})$ a generic symmetric matrix with ones on the diagonal. The *saturated conditional independence ideal* $\mathrm{CI}_{G,H} \subseteq \mathbb{R}[\sigma_{ij} : i < j]$ is the saturation of the ideal $(\sigma_{ij}, \det(\Sigma_{kC,lC}) : ij \in E_H^{\mathsf{c}}, kl \in E_G^{\mathsf{c}}$ separated by $C)$ at the product of all principal minors of Σ. Similarly, the *simplified saturated conditional independence ideal* $\mathrm{SCI}_{G,H} \subseteq \mathbb{R}[\sigma_{ij} : i < j]$ is the saturation of $(\sigma_{ij}, \det(\Sigma_{[n]\setminus l,[n]\setminus k}) : ij \in E_H^{\mathsf{c}}, kl \in E_G^{\mathsf{c}})$ at the product of all principal minors of Σ.

We have $\mathrm{SCI}_{G,H} \subseteq \mathrm{CI}_{G,H}$, and both ideals have the same vanishing sets in the affine space of symmetric matrices with ones on the diagonal over both $\mathbb{R}$ and $\mathbb{C}$ by Lemma 3.2.1. The ideals $\mathrm{SCI}_{G,H}$ and $\mathrm{CI}_{G,H}$ are saturations of determinantal ideals of symmetric matrices. Certain classes of determinantal ideals of generic symmetric matrices have been featured already in the early work of Conca [Con94a, Con94b, Con94c]. Double Markovian models might provide an incentive to further study ideals generated by only *some* minors of sparse generic symmetric matrices. This is related to Chapter 2 but very likely much more difficult.
In the following lemma we consider the affine scheme defined by $\mathrm{SCI}_{G,H}$ or $\mathrm{CI}_{G,H}$. Picking the latter, the dual of the Zariski tangent space at the identity $\mathbb{1}_n$ is $\mathfrak{m}/(\mathrm{CI}_{G,H} + \mathfrak{m}^2)$ where $\mathfrak{m} := (\sigma_{st} : s < t)$ is the maximal ideal in $A := \mathbb{R}[\sigma_{st} : s < t]$ corresponding to $\mathbb{1}_n$. Its dimension is also known as the *embedding dimension* of $(A/\mathrm{CI}_{G,H})_{\mathfrak{m}}$.

Lemma 3.3.18. Let V be the closed subscheme of the affine space of symmetric $n \times n$ matrices with ones on the diagonal defined by $\mathrm{SCI}_{G,H}$ or $\mathrm{CI}_{G,H}$. Then the embedding dimension of V at $\mathbb{1}_n$ equals $|E_G \cap E_H|$.

Proof. The short exact sequence

$$0 \to (\mathrm{CI}_{G,H} + \mathfrak{m}^2)/\mathfrak{m}^2 \to \mathfrak{m}/\mathfrak{m}^2 \to \mathfrak{m}/(\mathrm{CI}_{G,H} + \mathfrak{m}^2) \to 0$$

shows that

$$\operatorname{edim}((A/\mathrm{CI}_{G,H})_{\mathfrak{m}}) = \binom{n}{2} - \dim_{\mathbb{R}}((\mathrm{CI}_{G,H} + \mathfrak{m}^2)/\mathfrak{m}^2),$$

so it suffices to compute $d := \dim_{\mathbb{R}}((\mathrm{CI}_{G,H} + \mathfrak{m}^2)/\mathfrak{m}^2)$. But $\mathfrak{m}^2$ contains all products of two or more σ_{st} and therefore $\det(\Sigma_{kC,lC}) \equiv_{\mathfrak{m}^2} \pm\sigma_{kl}$. As every principal minor contains 1 as a monomial, saturation does not change d. Hence, $d = |E_G^{\mathsf{c}} \cup E_H^{\mathsf{c}}| = \binom{n}{2} - |E_G \cap E_H|$. The proof for $\mathrm{SCI}_{G,H}$ is the same. □

We recall next Proposition 2.2.1, expressing almost-principal minors of symmetric matrices via paths in a graph. We keep the convention that a path p traverses no vertex more than once and denote by $V(p) \subseteq [n]$ the set of vertices in p. Moreover, let $e(p) = kl$ denote the endpoints of p and $\sigma_p = \prod_{ij \in p} \sigma_{ij}$ the product over the variables corresponding to the edges in p. The sign of p is $\operatorname{sgn}(p) := (-1)^{|V(p)|-1}$.

Proposition 3.3.19. Let H be a graph on the vertex set $[n]$ and let $\Sigma = (\sigma_{ij})$ be a generic $n \times n$ symmetric matrix with $\sigma_{ij} = 0$ for all $ij \notin E_H$. Then

$$(-1)^{k+l} \det(\Sigma_{[n]\setminus k,[n]\setminus l}) = \sum_{\substack{p \text{ path in } H,\\ e(p)=kl}} \operatorname{sgn}(p) \cdot \det(\Sigma_{[n]\setminus V(p),[n]\setminus V(p)}) \cdot \sigma_p.$$

This formula appears first in [JW05]. We include a quick proof disregarding the signs since we do not use them in the sequel.

Proof. By the Leibniz formula, $\det(\Sigma_{[n]\setminus l,[n]\setminus k}) = \sum_\tau \operatorname{sgn}(\tau) \prod_{i \in [n]\setminus l} \sigma_{i,\tau(i)}$, where the sum is over all bijective $\tau \colon [n] \setminus l \to [n] \setminus k$. The summand corresponding to τ is non-zero if and only if each $\{i, \tau(i)\}$ with $\tau(i) \neq i$ is an edge of H. Starting at the vertex k, the sequence $k, \tau(k), \tau^2(k), \dots$ is a path from k to l in H, showing that

$$\begin{aligned} \det(\Sigma_{[n]\setminus l,[n]\setminus k}) &= \sum_{\substack{p \text{ path in } H,\\ e(p)=kl}} \pm\sigma_p \cdot \sum_{\tau' : [n]\setminus V(p) \xrightarrow{\cong} [n]\setminus V(p)} \operatorname{sgn}(\tau') \prod_{i \in [n]\setminus V(p)} \sigma_{i,\tau'(i)} \\ &= \sum_{\substack{p \text{ path in } H,\\ e(p)=kl}} \pm\sigma_p \cdot \det(\Sigma_{[n]\setminus V(p),[n]\setminus V(p)}). \end{aligned}$$ □

If all but one term in every generator $\det(\Sigma_{[n]\setminus k,[n]\setminus l})$ vanish, the saturated CI ideal is a monomial ideal and in fact agrees with the vanishing ideal.

Theorem 3.3.20. Let G and H be graphs on $[n]$ such that for every non-edge kl of G, there is at most one path p in H connecting k and l. Then

$$\mathrm{SCI}_{G,H} = \mathrm{CI}_{G,H} = \mathcal{I}(\mathcal{M}_1(G,H)) = (\sigma_{ij}, \sigma_p : ij \notin E_H,\ p \text{ path in } H \text{ with } e(p) \notin E_G).$$

Proof. Let Σ be the generic symmetric matrix with ones on the diagonal and zeros corresponding to non-edges of H. If $ij \notin E_H$, all terms of almost-principal minors of Σ which contain σ_{ij} can be neglected since $\sigma_{ij} \in \mathrm{SCI}_{G,H} \subseteq \mathrm{CI}_{G,H}$. By Proposition 3.3.19, if p is the unique path connecting k and l inside H, then $\det(\Sigma_{[n]\setminus k,[n]\setminus l}) = \pm \det(\Sigma_{[n]\setminus V(p),[n]\setminus V(p)}) \cdot \sigma_p$, so σ_p lies in the saturated simplified conditional independence ideal. If there exists no such path, this almost-principal minor vanishes. Since the square-free monomial ideal

$$(\sigma_{ij}, \sigma_p : ij \notin E_H, p \text{ path in } H \text{ with } e(p) \notin E_G)$$

agrees with its saturation at the product of all principal minors of Σ, it equals $\mathrm{SCI}_{G,H}$. To show that this (radical) ideal equals the vanishing ideal, it suffices to see that each of its components intersects $\mathrm{PD}_{n,1}$ in a smooth real point. This is clear since all irreducible components of $\mathcal{V}_{\mathbb{C}}(\mathrm{SCI}_{G,H})$ are coordinate subspaces. $\square$

The hypothesis of Theorem 3.3.20 may seem restrictive but, for example, it includes the case that H is a forest and G is arbitrary. On the other hand, it is easy to find an example on four vertices where H is a cycle and $\mathrm{SCI}_{G,H}$ is not a monomial ideal. Even determining vanishing ideals of ordinary Gaussian graphical models can be complicated, see for example [MS21].

3.3.3 Connectedness

In this subsection we study the connectedness of $\mathcal{M}(G,H)$ in the euclidean topology. If the vanishing ideal is known and simple enough, the results are easy as in the next corollary. Theorem 3.3.22 contains a sufficient condition based on connectedness in G and H.

Corollary 3.3.21. Under the hypotheses of Theorem 3.3.20, the model $\mathcal{M}(G,H)$ is connected. Moreover the following are equivalent.

(1) $\mathcal{M}(G,H)$ is smooth.

(2) $\mathcal{M}(G,H)$ is irreducible.

(3) $\mathcal{M}(G,H)$ has the maximal dimension $|E_G \cap E_H| + n$.

In this case, $\mathcal{M}(G,H) = \mathcal{M}(G \cap H)^{-1}$ is an inverse graphical model, hence a spectrahedron.

Proof. As a union of coordinate subspaces intersected with PD_n, $\mathcal{M}(G,H)$ is star-shaped with respect to the identity matrix and thus connected. Therefore, having the maximal dimension implies smoothness by Theorem 3.3.15, and smoothness together with connectedness implies irreducibility because regular local rings are integral domains [Kem11, Corollary 13.6]. By Theorem 3.3.20, irreducibility implies that the square-free monomial ideal $\mathrm{SCI}_{G,H}$ is prime. This is equivalent to the condition that for a path p inside H with $e(p) = kl \in E_G^{\mathsf{c}}$ there exists an edge $ij \in p \cap E_G^{\mathsf{c}}$. Thus, if G' is the graph on $[n]$ which is obtained from G by adding all edges in $E_G^{\mathsf{c}} \cap E_H^{\mathsf{c}}$, then $\mathcal{M}(G,H) = \mathcal{M}(G',H) = \mathcal{M}(G \cap H)^{-1}$, in particular $\mathcal{M}(G,H)$ has dimension $|E_G \cap E_H| + n$. $\square$

Theorem 3.3.22. Let G and H be graphs on $[n]$ with the property that there exists $i \in [n]$ such that for all non-edges $kl \in E_G^{\mathsf{c}}$, every path in H connecting k and l contains i. Then the model $\mathcal{M}(G,H)$ is connected.

Proof. In this proof we denote by $H \setminus i$ the graph on $[n]$ obtained from H by deleting all edges incident with i but keeping i as a vertex. The model $\mathcal{M}(H \setminus i)^{-1}$ is connected as it is the intersection of a linear space with the convex set PD_n, and the intersection of convex sets is convex, hence connected. Clearly, $\Sigma \in \mathcal{M}(H \setminus i)^{-1}$ if and only if $\Sigma \in \mathcal{M}(H)^{-1}$ and $\sigma_{ij} = 0$ for all $j \neq i$. The determinantal identity of Proposition 3.3.19, and the assumptions on G and H imply $\mathcal{M}(H \setminus i)^{-1} \subseteq \mathcal{M}(G,H)$. Now, let $\Sigma \in \mathcal{M}(G,H)$ be arbitrary. It suffices to find a path from Σ to some matrix in $\mathcal{M}(H \setminus i)^{-1}$.
For $\varepsilon \in [0,1]$ consider

$$\Sigma^\varepsilon := \Sigma \odot \begin{pmatrix} 1 & \dots & 1 & \varepsilon & 1 & \dots & 1 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 1 & \dots & 1 & \varepsilon & 1 & \dots & 1 \\ \varepsilon & \dots & \varepsilon & 1 & \varepsilon & \dots & \varepsilon \\ 1 & \dots & 1 & \varepsilon & 1 & \dots & 1 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 1 & \dots & 1 & \varepsilon & 1 & \dots & 1 \end{pmatrix},$$

where the second factor has entries ε in the i-th row and column, but 1 in entry ii. The symbol $\odot$ denotes the Hadamard product which multiplies matrices entry-wise. Then Σ^ε is symmetric and positive definite for all $\varepsilon \in [0,1]$ as it is the Hadamard product of a positive definite matrix and a positive semi-definite matrix with strictly positive diagonal entries. Moreover, $\Sigma^1 = \Sigma$ and $\Sigma^0 \in \mathcal{M}(H \setminus i)^{-1}$, so it suffices to show that $\Sigma^\varepsilon \in \mathcal{M}(G,H)$ for all $\varepsilon \in [0,1]$. This follows from the assumptions on G and H as for all $kl \in E_G^{\mathsf{c}}$ with $k \neq i \neq l$ we have

$$\begin{aligned} \det((\Sigma^\varepsilon)_{[n]\setminus k,[n]\setminus l}) &= \sum_{\substack{p \text{ path in } G \\ e(p)=kl}} \operatorname{sgn}(p) \cdot \sigma_p^\varepsilon \cdot \det((\Sigma^\varepsilon)_{[n]\setminus V(p),[n]\setminus V(p)}) \\ &= \varepsilon^2 \cdot \sum_{\substack{p \text{ path in } G \\ e(p)=kl}} \operatorname{sgn}(p) \cdot \sigma_p \cdot \det(\Sigma_{[n]\setminus V(p),[n]\setminus V(p)}) \\ &= \varepsilon^2 \cdot \det(\Sigma_{[n]\setminus k,[n]\setminus l}) = 0, \end{aligned}$$

using in the second step that, by assumption, $i \in V(p)$ for all occurring paths p, so that the occurring principal minors of Σ^ε agree with the corresponding principal minors of Σ. Moreover, each monomial σ_p^ε contains exactly two variables that are scaled by ε. If one of k and l agrees with i, the same calculation works with ε instead of ε^2. □

3.3.4 A Decomposition Theorem

Section 3.3.5 contains a classification of models with small $|E_G \cap E_H|$. This is based on the following decomposition theorem, whose proof also works for complex hermitian positive definite matrices.

Theorem 3.3.23. Let G, H be two graphs on the vertex set $[n]$. Let $V_1, \dots, V_r$ be a partition of $[n]$ such that each V_i is the vertex set of a connected component of $G \cap H$, considered as the graph on $[n]$ with edge set $E_G \cap E_H$. Then

$$\mathcal{M}(G, H) = \bigoplus_{i=1}^{r} \mathcal{M}(G|_{V_i}, H|_{V_i}).$$

In words, every $\Sigma \in \mathcal{M}(G, H)$ is block-diagonal with r blocks whose rows and columns are indexed by the V_i.

Proof. Inductively, it suffices to consider the case $r = 2$. We set $V := V_1$ and $W := V_2$. It is then enough to show that every matrix

$$\Sigma = \left(\begin{array}{c|c} \Sigma_{VV} & \Sigma_{VW} \\ \hline \Sigma_{VW}^t & \Sigma_{WW} \end{array}\right) \in \mathcal{M}(G, H)$$

is block-diagonal, i.e., $\Sigma_{VW} = 0$. We partition Σ^{-1} in the same way as Σ. By assumption, $G \cap H$ contains no edges between V and W. This implies that the matrix $\Sigma_{VW}(\Sigma^{-1})^t_{VW}$ has only zeros on the diagonal as for all $v \in V$,

$$(\Sigma_{VW}(\Sigma^{-1})^t_{VW})_{vv} = \sum_{w \in W} \Sigma_{vw}(\Sigma^{-1})_{vw} = \sum_{w \in W} 0 = 0.$$

In particular, $\operatorname{tr}(\Sigma_{VW}(\Sigma^{-1})^t_{VW}) = 0$. As Σ_{VV} and Σ_{WW} are positive definite, there exist symmetric square roots A_V and A_W such that $A_V^2 = \Sigma_{VV}$ and $A_W^2 = \Sigma_{WW}$. We now define

$$\Sigma' := \left(\begin{array}{c|c} A_V^{-1} & 0 \\ \hline 0 & A_W^{-1} \end{array}\right) \cdot \Sigma \cdot \left(\begin{array}{c|c} A_V^{-1} & 0 \\ \hline 0 & A_W^{-1} \end{array}\right) = \left(\begin{array}{c|c} \mathbb{1}_V & A_V^{-1}\Sigma_{VW}A_W^{-1} \\ \hline A_W^{-1}\Sigma_{VW}^t A_V^{-1} & \mathbb{1}_W \end{array}\right).$$

Clearly, $\operatorname{tr}(\Sigma') = |V| + |W| = n$. For the inverse matrix we have

$$\Sigma'^{-1} = \left(\begin{array}{c|c} A_V & 0 \\ \hline 0 & A_W \end{array}\right) \cdot \Sigma^{-1} \cdot \left(\begin{array}{c|c} A_V & 0 \\ \hline 0 & A_W \end{array}\right) = \left(\begin{array}{c|c} (\Sigma'^{-1})_{VV} & A_V(\Sigma^{-1})_{VW}A_W \\ \hline A_W(\Sigma^{-1})^t_{VW}A_V & (\Sigma'^{-1})_{WW} \end{array}\right).$$

Now observe that $\Sigma'_{VW}(\Sigma'^{-1})^t_{VW} = A_V^{-1}\Sigma_{VW}(\Sigma^{-1})^t_{VW}A_V$ as the product $A_W^{-1}A_W = \mathbb{1}_W$ in the middle cancels out. Since the trace is cyclic, we have

$$\operatorname{tr}(\Sigma'_{VW}(\Sigma'^{-1})^t_{VW}) = \operatorname{tr}(A_V^{-1}\Sigma_{VW}(\Sigma^{-1})^t_{VW}A_V) = \operatorname{tr}(\Sigma_{VW}(\Sigma^{-1})^t_{VW}A_VA_V^{-1}) = 0.$$

Moreover,

$$\mathbb{1}_V = (\Sigma'\Sigma'^{-1})_{VV} = (\Sigma'^{-1})_{VV} + \Sigma'_{VW}(\Sigma'^{-1})^t_{VW},$$

implying $\operatorname{tr}((\Sigma'^{-1})_{VV}) = \operatorname{tr}(\mathbb{1}_V) = |V|$. Similarly, $\operatorname{tr}((\Sigma'^{-1})_{WW}) = |W|$, so $\operatorname{tr}(\Sigma'^{-1}) = |V| + |W| = n$. As Σ' is real symmetric positive definite there exists a symmetric square root T with $T^2 = \Sigma'$ and thus also $T^{-2} = \Sigma'^{-1}$. Using the inner product $\langle X, Y\rangle = \operatorname{tr}(XY) = \sum_{i,j=1}^n x_{ij}y_{ij}$ on the space of real symmetric matrices,

$$\langle T, T\rangle = \operatorname{tr}(T^2) = \operatorname{tr}(\Sigma') = n,$$
$$\langle T^{-1}, T^{-1}\rangle = \operatorname{tr}(T^{-2}) = \operatorname{tr}(\Sigma'^{-1}) = n,$$
$$\langle T, T^{-1}\rangle = \operatorname{tr}(\mathbb{1}_n) = n.$$

Therefore, we have equality in the Cauchy–Schwarz inequality

$$n^2 = \langle T, T^{-1} \rangle^2 \leq \langle T, T \rangle \cdot \langle T^{-1}, T^{-1} \rangle = n^2,$$

implying that T and T^{-1} are linearly dependent as matrices, i.e., $T = \lambda T^{-1}$ for some $\lambda \in \mathbb{R}$. This implies $\Sigma' = T^2 = \lambda \mathbb{1}_n$. In particular, $0 = \Sigma'_{VW} = A_V^{-1} \Sigma_{VW} A_W^{-1}$ as matrices. But this is equivalent to $\Sigma_{VW} = 0$, as desired. □

If all V_i are single vertices we get the following.

Corollary 3.3.24. We have $\mathcal{M}_1(G, H) = \{\mathbb{1}_n\}$ if and only if $E_G \cap E_H = \emptyset$.

In other words, if Σ is a symmetric positive definite $(n \times n)$-matrix with the property that every off-diagonal entry vanishes either in Σ or in Σ^{-1} (or both), then Σ is a diagonal matrix. We have not found this result in the literature.

Remark 3.3.25. A natural question is whether the assumption of positive definiteness in Theorem 3.3.23 is necessary. Example 3.4.4 shows that Corollary 3.3.24 does not hold for positive *semi*-definite matrices, and Example 3.3.27 below shows that Theorem 3.3.23 does not hold for principally regular matrices, that is, matrices whose principal minors do not vanish but that might not be positive definite. Perhaps surprisingly, however, Corollary 3.3.24 *does* hold for principally regular matrices over *any* field by Proposition 2.5.4.

Remark 3.3.26. A simpler variant of Theorem 3.3.23 can be proven by recursive direct sum decomposition and duality: To every pair of graphs (G, H) on $[n]$ there exists a partition $V_1, \ldots, V_r$ of $[n]$ such that

(1) $G_i = G|_{V_i}$ and $H_i = H|_{V_i}$ are connected.

(2) $\mathcal{M}(G, H)$ is smooth if and only if all $\mathcal{M}(G_i, H_i)$ are smooth.

(3) $\mathcal{M}(G, H)$ is connected if and only if all $\mathcal{M}(G_i, H_i)$ are connected.

The merit of this simpler assertion is that it does not require positive definiteness. It also holds for principally regular models of $\langle\!\langle G, H \rangle\!\rangle$ over $\mathbb{C}$ because the proof uses only elementary operations on CI relations introduced in Section 3.2.3.

Example 3.3.27. Consider the graph in Fig. 3.1. We study the variety of $\langle\!\langle G, H \rangle\!\rangle$ in `Macaulay2`:

```
R = QQ[x11,x12,x13,x14,x15,x16,  x22,x23,x24,x25,x26,
x33,x34,x35,x36,            x44,x45,x46,
x55,x56,                    x66]
X = genericSymmetricMatrix(R,x11,6)

-- Impose the relations from $H$ directly on the matrix
X = sub(X, { x16=>0, x24=>0, x25=>0, x26=>0, x35=>0 })
```

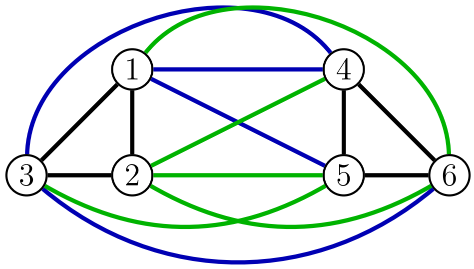


Figure 3.1: The graphs for Example 3.3.27. The edges of $G \cap H$ are drawn black, $G \setminus H$ in green and $H \setminus G$ in blue. Then $G \cap H = K_{123} \oplus K_{456}$ and $G \cup H = K_{1\cdots 6}$.

```
-- Pick an affine slice of the model which is likely to contain
-- positive definite matrices by diagonal dominance
X = sub(X, {
x11=>10, x22=>10, x33=>10, x44=>10, x55=>10, x66=>10,
x12=>1,  x13=>1,  x23=>1,  x45=>1,  x46=>1,  x56=>1
})
```

$$\begin{pmatrix} 10 & 1 & 1 & x_{14} & x_{15} & 0 \\ 1 & 10 & 1 & 0 & 0 & 0 \\ 1 & 1 & 10 & x_{34} & 0 & x_{36} \\ x_{14} & 0 & x_{34} & 10 & 1 & 1 \\ x_{15} & 0 & 0 & 1 & 10 & 1 \\ 0 & 0 & x_{36} & 1 & 1 & 10 \end{pmatrix}$$

Some of the variables are specified to ensure quick termination of the following computations. If Theorem 3.3.23 held for principally regular matrices, x_{14}, x_{15}, x_{34} and x_{36} would vanish on every principally regular matrix satisfying the equations of $\langle\!\langle G \rangle\!\rangle$.

```
-- The relations imposed by $G$
I = radical ideal(
det submatrix'(X, {0}, {3}),  -- $14$
det submatrix'(X, {0}, {4}),  -- $15$
det submatrix'(X, {2}, {3}),  -- $34$
det submatrix'(X, {2}, {5})   -- $36$
)
-- Saturation at each of the principal minors
J = fold((I,f) -> I : f, I, subsets(numRows(X)) / (K -> det X_K^K))
decompose J
```

$$\begin{aligned} &(x_{14}, x_{15}, x_{34}, x_{36}) \\ \cap\ &(1210x_{14}^2 - 999, -11x_{14} + x_{15}, -x_{14} + x_{34}, -11x_{14} + x_{36}) \\ \cap\ &(1210x_{14}^2 - 981, -11x_{14} + x_{15}, x_{14} + x_{34}, 11x_{14} + x_{36}) \end{aligned}$$

The first component has the desired block structure of $K_{123} \oplus K_{456}$, but the other components contain real points as well. Consider the last component. It consists of two real points:

$$x_{14} = \pm\sqrt{\frac{981}{1210}}, \quad x_{15} = 11x_{14}, \quad x_{34} = -x_{14}, \quad x_{36} = -11x_{14}.$$

This yields a real matrix satisfying the equations of $\langle\!\langle G, H \rangle\!\rangle$ and whose principal minors are non-zero. However, the determinant of the entire matrix equals $-\frac{4374}{55}$, which is not positive.
This shows that $\langle\!\langle G, H \rangle\!\rangle$ has real, principally regular solutions without block-diagonal structure. The positive definite matrices in the affine slice J of the model all fall into the first component and do have the block structure. A purely algebraic computation, without taking positive definiteness into account, would not be able to prove Theorem 3.3.23.

3.3.5 Classification of $\mathcal{M}(G, H)$ for $|E_G \cap E_H| \leq 3$

Theorem 3.3.11 bounds the model dimension in terms of $|E_G \cap E_H|$. We finish our analysis of the geometry of $\mathcal{M}(G, H)$ with a classification of models with small intersections of the edge sets. In view of Theorem 3.3.23 we can restrict to the cases where $E_G \cap E_H$ defines a connected graph on the subset of vertices of $[n]$ incident to some edge in $E_G \cap E_H$. For any $N' \subseteq [n]$, we also write $\mathrm{PD}_{N'}$ for the set of positive definite matrices with rows and columns indexed by N'. For disjoint subsets N' and N'', a direct sum $\mathrm{PD}_{N'} \oplus \mathrm{PD}_{N''}$ indicates the set of block-diagonal positive definite matrices inside $\mathrm{PD}_{N' \cup N''}$ with the rows and columns of the two blocks indexed, respectively, by N' and N'', and similarly for $\mathrm{PD}_{N',1} \oplus \mathrm{PD}_{N'',1}$ if we restrict to ones on the diagonal.

Proposition 3.3.28. Let $E_G \cap E_H = \{ij\}$ consist of a single edge. Then

$$\mathcal{M}(G, H) = \mathrm{PD}_{ij}\,.$$

In particular, $\mathcal{M}_1(G, H)$ is connected and smooth and has the maximal dimension $|E_G \cap E_H| = 1$.

Proof. Immediate from Theorem 3.3.23 and the definitions. □

Proposition 3.3.29. Let $|E_G \cap E_H| = 2$ and so $E_G \cap E_H = \{ij, jk\}$ with distinct i, j, k.

(1) If $ik \in E_G \setminus E_H$, then $\mathcal{M}(G, H) = \mathcal{M}(G \cap H)^{-1}$ is an inverse graphical model.

(2) In case $ik \in E_H \setminus E_G$, symmetrically $\mathcal{M}(G, H) = \mathcal{M}(G \cap H)$ is a graphical model.

(3) If $ik \notin E_G \cup E_H$, then $\mathcal{M}_1(G, H)$ decomposes as

$$\mathcal{M}_1(G, H) = \{\mathbb{1}_n + tE^{ij} : t \in (-1, 1)\} \cup \{\mathbb{1}_n + tE^{jk} : t \in (-1, 1)\}.$$

The Zariski closure of $\mathcal{M}_1(G, H)$ is a pair of lines intersecting in $\mathbb{1}_n$. Thus $\mathcal{M}_1(G, H)$ is connected of dimension one, with reducible Zariski closure.

Proof. After a suitable permutation of $[n]$, we assume $ij = 12$ and $jk = 23$. Any $\Sigma \in \mathcal{M}_1(G,H)$ has the block-diagonal form consisting of an upper left (3×3)-block and an identity matrix. Therefore we can assume $n = 3$. Then, in the first case, we have

$$\Sigma = \begin{pmatrix} 1 & \sigma_{12} & 0 \\ \sigma_{12} & 1 & \sigma_{23} \\ 0 & \sigma_{23} & 1 \end{pmatrix}, \quad \operatorname{adj}(\Sigma) = \begin{pmatrix} 1-\sigma_{23}^2 & -\sigma_{12} & \sigma_{12}\sigma_{23} \\ -\sigma_{12} & 1 & -\sigma_{23} \\ \sigma_{12}\sigma_{23} & -\sigma_{23} & 1-\sigma_{12}^2 \end{pmatrix}.$$

As $13 \in E_G$, we have that $G = K_3$ is complete, so there are no further restrictions and we obtain $\mathcal{M}(G,H) = \mathcal{M}(G\cap H)^{-1}$. In the second case, the same is true if we replace Σ by Σ^{-1} everywhere, so $\mathcal{M}(G,H) = \mathcal{M}(G\cap H)$. Finally, in the third case, $13 \notin E_G \cup E_H$, so we additionally get $\sigma_{12} = 0$ or $\sigma_{23} = 0$, obtaining the union of two line segments, as desired. □

Remark 3.3.30. Proposition 3.3.29 shows that in some non-obvious cases double Markovian models are graphical or inverse graphical models. This theme has occurred in the literature. For example, in [DR08, Proposition 12] it is shown that the only way that a covariance graph model $\mathcal{M}(K_n, H)$ is a graphical model $\mathcal{M}(G, K_n)$ is if covariance and concentration matrices have aligned block structures and the model is a product of PD cones (in particular, $G = H$ is a disjoint union of cliques).

The same ideas also prove the following via direct computations.

Proposition 3.3.31. Let $|E_G \cap E_H| = 3$. If $E_G \cap E_H = \{ij, ik, jk\}$ forms a 3-clique, then $\mathcal{M}(G,H) = \mathrm{PD}_{ijk}$. Otherwise $E_G \cap E_H = \{ij, jk, kl\}$ with distinct i,j,k,l forms a path or $E_G \cap E_H = \{ij, ik, il\}$ is a star. Up to swapping G and H, we can restrict to the case where $H|_{ijkl}$ has equally many or more non-edges than $G|_{ijkl}$ (i.e. at least as many prescribed zeros in the covariance matrix as in the concentration matrix). We can restrict moreover to $n = 4$ and assume $(i,j,k,l) = (1,2,3,4)$. If $E_G \cap E_H = \{12, 23, 34\}$ is a path, we have the following cases up to symmetry and inversion:

(1) $E_H = E_G \cap E_H$ and $E_G = K_{1234}$. Here, $\mathcal{M}(G,H) = \mathcal{M}(G\cap H)^{-1}$.

(2) $E_H = E_G \cap E_H$ and $E_G = K_{1234} \setminus \{13\}$. Here, $\mathcal{M}(G,H) = (\mathrm{PD}_{12} \oplus \mathrm{PD}_{34}) \cup \mathcal{M}(\{23,34\})^{-1}$.

(3) $E_H = E_G \cap E_H$ and $E_G = K_{1234} \setminus \{14\}$. Here, $\mathcal{M}(G,H) = (\mathrm{PD}_{12} \oplus \mathrm{PD}_{34}) \cup \mathcal{M}(\{12,23\})^{-1} \cup \mathcal{M}(\{23,34\})^{-1}$.

(4) $E_H = E_G \cap E_H$ and $E_G = K_{1234} \setminus \{13,14\}$. Here, $\mathcal{M}(G,H)$ is as in the previous case.

(5) $E_H = E_G \cap E_H$ and $E_G = K_{1234} \setminus \{13,24\}$. Here, $\mathcal{M}(G,H) = (\mathrm{PD}_{12} \oplus \mathrm{PD}_{34}) \cup \mathrm{PD}_{23}$.

(6) $E_G = E_H = E_G \cap E_H$. Here, $\mathcal{M}(G,H)$ is as in the previous case.

(7) $E_H = (E_G \cap E_H) \cup \{13\}$ and $E_G = K_{1234} \setminus \{13\}$. Here,

$$\mathcal{M}_1(G,H) = \left\{ \begin{pmatrix} 1 & \sigma_{12} & \sigma_{12}\sigma_{23} & 0 \\ \sigma_{12} & 1 & \sigma_{23} & 0 \\ \sigma_{12}\sigma_{23} & \sigma_{23} & 1 & \sigma_{34} \\ 0 & 0 & \sigma_{34} & 1 \end{pmatrix} : \sigma_{12} \in (-1,1), \sigma_{23}^2 + \sigma_{34}^2 < 1 \right\}.$$

(8) $E_H = (E_G \cap E_H) \cup \{13\}$ and $E_G = K_{1234} \setminus \{13, 14\}$. Here, $\mathcal{M}_1(G,H)$ is as in the previous case.

(9) $E_H = (E_G \cap E_H) \cup \{13\}$ and $E_G = K_{1234} \setminus \{13, 24\}$. Here,

$$\begin{aligned}\mathcal{M}_1(G,H) = {} & (\mathrm{PD}_{\{12\},1} \oplus \mathrm{PD}_{\{34\},1}) \\ & \cup \left\{ \begin{pmatrix} 1 & \sigma_{12} & \sigma_{12}\sigma_{23} & 0 \\ \sigma_{12} & 1 & \sigma_{23} & 0 \\ \sigma_{12}\sigma_{23} & \sigma_{23} & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} : \sigma_{12}, \sigma_{23} \in (-1,1) \right\}.\end{aligned}$$

(10) $E_H = (E_G \cap E_H) \cup \{14\}$ and $E_G = K_{1234} \setminus \{14\}$. Here,

$$\mathcal{M}_1(G,H) = \left\{ \begin{pmatrix} 1 & \sigma_{12} & 0 & \frac{-\sigma_{12}\sigma_{23}\sigma_{34}}{1-\sigma_{23}^2} \\ \sigma_{12} & 1 & \sigma_{23} & 0 \\ 0 & \sigma_{23} & 1 & \sigma_{34} \\ \frac{-\sigma_{12}\sigma_{23}\sigma_{34}}{1-\sigma_{23}^2} & 0 & \sigma_{34} & 1 \end{pmatrix} : \sigma_{12}^2 + \sigma_{23}^2 < 1, \sigma_{23}^2 + \sigma_{34}^2 < 1 \right\}.$$

(11) $E_H = (E_G \cap E_H) \cup \{14\}$ and $E_G = K_{1234} \setminus \{13, 14\}$. Here,

$$\mathcal{M}_1(G,H) = (\mathrm{PD}_{\{12\},1} \oplus \mathrm{PD}_{\{34\},1}) \cup \left\{ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \sigma_{23} & 0 \\ 0 & \sigma_{23} & 1 & \sigma_{34} \\ 0 & 0 & \sigma_{34} & 1 \end{pmatrix} : \sigma_{23}^2 + \sigma_{34}^2 < 1 \right\}.$$

If $E_G \cap E_H = \{12, 13, 14\}$ is a star, we have the following cases up to symmetry and inversion:

(1) $E_H = E_G \cap E_H$ and $E_G = K_{1234}$. Here, $\mathcal{M}(G,H) = \mathcal{M}(G \cap H)^{-1}$.

(2) $E_H = E_G \cap E_H$ and $E_G = K_{1234} \setminus \{23\}$. Then

$$\begin{aligned}\mathcal{M}_1(G,H) = {} & \left\{ \begin{pmatrix} 1 & 0 & \sigma_{13} & \sigma_{14} \\ 0 & 1 & 0 & 0 \\ \sigma_{13} & 0 & 1 & 0 \\ \sigma_{14} & 0 & 0 & 1 \end{pmatrix} : \sigma_{13}^2 + \sigma_{14}^2 < 1 \right\} \\ & \cup \left\{ \begin{pmatrix} 1 & \sigma_{12} & 0 & \sigma_{14} \\ \sigma_{12} & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sigma_{14} & 0 & 0 & 1 \end{pmatrix} : \sigma_{12}^2 + \sigma_{14}^2 < 1 \right\},\end{aligned}$$

a union of two discs intersecting in a line segment.

(3) $E_H = (E_G \cap E_H) \cup \{23\}$ and $E_G = K_{1234} \setminus \{23\}$. Then

$$\mathcal{M}_1(G,H) = \left\{ \begin{pmatrix} 1 & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{12} & 1 & \frac{\sigma_{12}\sigma_{13}}{1-\sigma_{14}^2} & 0 \\ \sigma_{13} & \frac{\sigma_{12}\sigma_{13}}{1-\sigma_{14}^2} & 1 & 0 \\ \sigma_{14} & 0 & 0 & 1 \end{pmatrix} \in \mathrm{PD}_4 \right\}.$$

(4) $E_H = E_G \cap E_H$ and $E_G = K_{1234} \setminus \{23, 24\}$. Then

$$\mathcal{M}_1(G,H) = \left\{ \begin{pmatrix} 1 & \sigma_{12} & 0 & 0 \\ \sigma_{12} & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} : \sigma_{12} \in (-1,1) \right\}$$
$$\cup \left\{ \begin{pmatrix} 1 & 0 & \sigma_{13} & \sigma_{14} \\ 0 & 1 & 0 & 0 \\ \sigma_{13} & 0 & 1 & 0 \\ \sigma_{14} & 0 & 0 & 1 \end{pmatrix} : \sigma_{13}^2 + \sigma_{14}^2 < 1 \right\}.$$

(5) $E_H = (E_G \cap E_H) \cup \{23\}$ and $E_G = K_{1234} \setminus \{23, 24\}$. Then

$$\mathcal{M}_1(G,H) = \left\{ \begin{pmatrix} 1 & \sigma_{12} & \sigma_{13} & 0 \\ \sigma_{12} & 1 & \sigma_{12}\sigma_{13} & 0 \\ \sigma_{13} & \sigma_{12}\sigma_{13} & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} : \sigma_{12}, \sigma_{13} \in (-1,1) \right\}$$
$$\cup \left\{ \begin{pmatrix} 1 & 0 & \sigma_{13} & \sigma_{14} \\ 0 & 1 & 0 & 0 \\ \sigma_{13} & 0 & 1 & 0 \\ \sigma_{14} & 0 & 0 & 1 \end{pmatrix} : \sigma_{13}^2 + \sigma_{14}^2 < 1 \right\}.$$

(6) $E_G = E_H = E_G \cap E_H$. Then $\mathcal{M}_1(G,H)$ is the union of the three coordinate line segments parametrized by σ_{12}, σ_{13} and σ_{14}, all contained in $(-1,1)$.

In particular, for $|E_G \cap E_H| \leq 3$ the models $\mathcal{M}_1(G,H)$ are all connected in the euclidean topology. Moreover, in this range $\mathcal{M}_1(G,H)$ is smooth if and only if its Zariski closure is irreducible if and only if $\dim(\mathcal{M}_1(G,H)) = |E_G \cap E_H|$ is maximal. However, Example 3.4.2 below shows that there exist irreducible but singular double Markovian models on four vertices. Proposition 3.3.31 shows that double Markovian models need not be equidimensional.

Remark 3.3.32. The preceding classification for $|E_G \cap E_H| \leq 3$ and computer-assisted computations on $n \leq 4$ vertices show that in this range for every smooth model $\mathcal{M}(G,H)$ with $G \cap H$ connected, there exist graphs G', H' such that

$$\mathcal{M}(G,H) = \mathcal{M}(G',H')$$

and $G' \cup H' = K_n$. Hence, in hindsight, the smoothness of the geometric model follows from Theorem 3.3.8. It is unknown whether there are smooth models without this property.

Remark 3.3.33. Any classification project would profit from a complete list of models of double Markovian CI structures or equivalently the set of completions of relations $\langle\!\langle G, H \rangle\!\rangle$. There is no efficient, combinatorial algorithm to compute the completion of $\langle\!\langle G, H \rangle\!\rangle$, although computing the completion can be reduced to invocations of the Positivstellensatz and thus quantifier eliminiation [BCR98, Chapter 4]. One such project would be to classify smoothness of double Markovian models on small vertex sets. Exploiting Theorem 3.3.23, this reduces to multiple smoothness queries for smaller models. It seems like a worthwhile computational challenge to compile a table of the pairs of small connected graphs which have a smooth model. We determined that for $n = 3, 4, 5$ there are $4+55+2644$ pairs of connected graphs which induce pairwise inequivalent CI structures modulo isomorphy and duality. For more information on computations see `https://gaussoids.de/doublemarkov.html`.

3.4 Examples, Counterexamples, and Conjectures

Example 3.4.1. Consider the double Markovian CI structure arising from $G = H =$ , namely $\{(14|), (14|23), (23|), (23|14)\}$. A computation in `Macaulay2` shows that (the Zariski closure of) the model $\mathcal{M}(G, H)$ has three irreducible components while it was determined in [DX10, Example 4.1] that the correlation model $\mathcal{M}_1(G, H) = \mathcal{M}(G, H) \cap \mathrm{PD}_{4,1}$ has four. Thus there are algebraic differences between $\mathcal{M}_1(G, H)$ and $\mathcal{M}(G, H)$.

Example 3.4.2. Continuing Example 3.2.17, suppose that $G = H$ is the complete graph on $[n]$ minus the edge 12. These double Markovian models are singular at the identity matrix but the models of all proper minors of $\langle\!\langle G, G \rangle\!\rangle$ are smooth. According to [DX10, Proposition 4.2] with $C_1 = \emptyset$ and $C_2 = [n] \setminus 12$, the singular locus of this model is again a Gaussian CI model and it is described as a submodel of $\mathcal{M}(G, G)$ by the CI statements

$$\begin{aligned} &(12|) \text{ and } (12|C_2) \text{ from } \langle\!\langle G, G \rangle\!\rangle, \\ &(1j|) \text{ and } (2j|) \text{ for all } j \in C_2. \end{aligned} \tag{$*$}$$

All but one of them are simple zero constraints on the covariance matrix. By Schur complement, the remaining almost-principal minor equals

$$\det \Sigma_{12|C_2} = (\sigma_{12} - \Sigma_{1,C_2} \cdot \Sigma_{C_2}^{-1} \cdot \Sigma_{C_2,2}) \det \Sigma_{C_2}.$$

By all the marginal independence statements in ($*$), the vectors Σ_{1,C_2} and $\Sigma_{C_2,2}$ as well as the entry σ_{12} are zero, so the right-hand side of this equality vanishes, and $(12|C_2)$ is implied by the marginal statements. Thus, the singular locus is in fact a *linear subspace* of $\mathrm{Sym}^2(\mathbb{R}^n)$ of codimension $2n-3$, intersected with PD_n. This shows that double Markovian models can have singular loci of arbitrarily large dimension. It is instructive to compute the concrete case of $n = 4$. The maximal possible dimension of the correlation model in this case is $|E_G| = 5$. However, $\mathcal{M}_1(G, G)$ is of dimension $4 < 5$. Indeed, the only conditions on a positive definite matrix $\Sigma \in \mathcal{M}_1(G, G)$ are $\sigma_{12} = 0$ and $(\Sigma^{\mathrm{adj}})_{12} = 0$, which, using $\sigma_{12} = 0$, writes as

$$f \coloneqq \sigma_{13}(\sigma_{24}\sigma_{34} - \sigma_{23}) + \sigma_{14}(\sigma_{23}\sigma_{34} - \sigma_{24}) = 0.$$

This is an irreducible polynomial, so the Zariski closure of $\mathcal{M}_1(G,G)$, which can be viewed as the vanishing set of the above polynomial inside the affine space where we forget the variable σ_{12}, is irreducible. In this case, the ideal $\mathrm{SCI}_{G,G}$ is prime even without saturation, and it coincides with $\mathcal{I}(\mathcal{M}_1(G,G)) = (\sigma_{12}, f)$. Moreover, we can see again that $\mathcal{M}_1(G,G)$ is connected: for every positive definite matrix Σ satisfying $\sigma_{12} = 0$ and $f = 0$, scale all variables *except* for σ_{34} by some ε tending to 0. This preserves the two equations and establishes a path inside $\mathcal{M}_1(G,G)$ connecting Σ to a matrix with only the entry $\sigma_{34} \in (-1,1)$ possibly non-zero. The set of these matrices is clearly connected. As computed above, in the singular locus all variables but σ_{34} are forced to zero, showing that it is a line inside PD_4. In particular, failure of smoothness is not always due to reducibility for double Markovian models.

In Example 3.4.2, we have $G = H =$ an almost complete graph, where only one edge is missing. This model is singular at the identity matrix and therefore shows that the sufficient condition for smoothness in Theorem 3.3.8, namely that $G \cup H = K_n$, cannot be weakened. The singular locus in this example is a submodel described by the occurrence of additional CI statements. However, this is not always the case, as [DX10, Example 4.3] discovered.

Example 3.4.3. Let $G =$  and $H =$ . The model $\mathcal{M}_1(G,H)$ agrees with its inverse model up to permutation of $[n]$. Moreover, $G \cup H = K_4$ is the complete graph, so this model is smooth of the expected dimension $|E_G \cap E_H| = 4$. However, neither $\mathcal{M}_1(G,H)$ nor its inverse lie in the graphical model $\mathcal{M}(G \cap H)$. Indeed, for every $\Sigma \in \mathcal{M}_1(G,H)$ the condition $(\Sigma^{-1})_{12} = 0$ translates into $\sigma_{12} = \sigma_{13}\sigma_{23} + \sigma_{14}\sigma_{24}$, in particular $\sigma_{12}, \sigma_{13}, \sigma_{14}, \sigma_{23}, \sigma_{24}$ can all be non-zero at the same time. The same is true for the inverse model, as it arises from the permutation exchanging 1 with 4 and 2 with 3.

Example 3.4.4. Consider the disjoint graphs $G =$  and $H =$ . The semi-definite model with ones on the diagonal consists of the five points

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & \pm 1 & 0 & 0 \\ \pm 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & \mp 1 \\ 0 & 0 & \mp 1 & 1 \end{pmatrix},$$

where the signs in the first two rows and in last two rows agree, respectively, but are independent of each other. Thus semi-definite models can be disconnected even in dimension zero.

We have no analogue of Example 3.4.4 with positive definite matrices, and semi-definite models with no restriction on the diagonal are connected as they are star-shaped with respect to the zero matrix. The question if double Markovian models are connected in general is open.

Conjecture 3.4.5. For any G, H, the models $\mathcal{M}(G,H)$ and $\mathcal{M}_1(G,H)$ are connected in the euclidean topology.

The connectedness of $\mathcal{M}(G,H)$ is equivalent to that of $\mathcal{M}_1(G,H)$. Moreover, $\mathcal{M}_1(G,H) \setminus \{\mathbb{1}_n\}$ can be disconnected as already evidenced by the 1-dimensional smooth case. Connectedness, even for the Zariski topology, implies:

Conjecture 3.4.6. The model $\mathcal{M}(G,H)$ (resp. $\mathcal{M}_1(G,H)$) is smooth if it has the maximal dimension $|E_G \cap E_H| + n$ (resp. $|E_G \cap E_H|$).

4 Symmetric Ideals

For questions of authorship, please refer to pages IVf.

This chapter is based on the accepted version of [Kre23].

4.1 Monomial Radicals of Symmetric Ideals

A symmetric ideal is an ideal in a polynomial ring $K[x_1, \ldots, x_n]$ which is stable under all permutations of the variables. Special classes of symmetric ideals, for instance *Specht ideals* [MOY22] and *Tanisaki ideals* [Tan82, BG92, GP92], have been studied intensively in the algebraic combinatorics literature and are related to Kostka–Macdonald polynomials and the famous work of Haiman on $n!$, see [Hai03] and the references therein. In commutative algebra, symmetric ideals are mainly studied for their asymptotic properties. We refer to [NR17, NR19, LNNR20, LNNR21, NS21] for plenty of examples as well as to [CF13, CEF15] for foundational results on representation stability and FI-modules. One of the most well-known results on symmetric ideals is that an infinite chain of symmetric ideals $I_n \subseteq K[x_1, \ldots, x_n]$ in increasingly many variables with $I_{n-1} \subseteq I_n$ for all n eventually stabilizes in the sense that $I_n = (S_n \cdot I_{n-1})$ for all n large enough [Coh67, AH07, HS12, Dra14]. Here, S_n is the symmetric group on n elements.
As opposed to such stabilization phenomena, in this chapter we investigate a certain rigidity statement for symmetric ideals in a *fixed* polynomial ring. We indicate that solution sets to symmetric systems of *general* polynomial equations should be very simple, in the precise sense of Theorem 4.1.1 and Conjecture 4.1.2. Proposition 4.1.4 is an example for how the infinite case can yield simpler and more explicit results. For a focus on Specht polynomials instead of monomials, see for example [MRV21].

Let us first introduce some relevant notation. The set of occurring monomials in a polynomial $f \in K[x_1, \ldots, x_n]$ over a field K is called its *support* and written $\operatorname{supp}(f)$. We may identify a monomial in $K[x_1, \ldots, x_n]$ with its exponent vector in $(\mathbb{Z}_{\geq 0})^n$ and call any non-empty, finite subset $\mathcal{A} \subseteq (\mathbb{Z}_{\geq 0})^n$ a *support set*. A support set $\mathcal{A}$ is called *homogeneous of degree d* if $\sum_{i=1}^n a_i = d$ for all $a \in \mathcal{A}$. It is called *symmetric* if all permutations of its elements belong to $\mathcal{A}$ as well. We identify $K^{\mathcal{A}}$ with the set of all polynomials $f \in K[x_1, \ldots, x_n]$ with $\operatorname{supp}(f) \subseteq \mathcal{A}$. We write S_∞ for the (small) infinite symmetric group. Moreover, we denote by A_n the alternating group, i.e., the index 2 subgroup of S_n consisting of all even permutations. The action of $\sigma \in S_n$ on a polynomial f is written $\sigma \cdot f$, and it is induced by letting $\sigma \cdot x_i := x_{\sigma(i)}$. For the orbit of f under the action of a subgroup $G \subseteq S_n$ we write $G \cdot f$. Listing the exponents of a monomial in decreasing order gives a partition of the degree of the monomial, and we call this partition the *type* of the monomial. Two monomials have the same type if and only if they are permutations of one another. For example, $x_1^2 x_2$ and $x_1 x_3^2$ are of the same type while $x_1 x_2 x_3$ has different type. Whenever we use the term *general*, the field K is assumed to be infinite. Fixing a support set $\mathcal{A}$, an assertion about polynomials $f \in K[x_1, \ldots, x_n]$ with $\operatorname{supp}(f) \subseteq \mathcal{A}$ holds for a *general* $f \in K^{\mathcal{A}}$ if the subset of $K^{\mathcal{A}}$ for which the assertion holds contains a non-empty Zariski-open subset. Given a set of polynomials S, by $\mathcal{V}(S) \subseteq K^n$ we denote its vanishing set.

Theorem 4.1.1. Let $\operatorname{char}(K) = 0$ and $n \geq 5$. Let $\mathcal{A} \subseteq (\mathbb{Z}_{\geq 0})^n$ be a homogeneous, symmetric support set, and denote by k the minimal number of distinct variables dividing some monomial in $\mathcal{A}$. Then, for general $f \in K^{\mathcal{A}}$,

$$\sqrt{(S_n \cdot f)} = (S_n \cdot x_1 x_2 \cdots x_k).$$

In particular, $\mathcal{V}(S_n \cdot f)$ is the union of all $(k-1)$-dimensional coordinate subspaces of K^n.

Theorem 4.1.1 indicates that the surprising behavior of [JKLR20, Example 2.6] might not be so rare. We even expect the following to hold true.

Conjecture 4.1.2. Let $\mathcal{A} \subseteq (\mathbb{Z}_{\geq 0})^n$ be any support set and k as in Theorem 4.1.1. Then, for a general $f \in K^{\mathcal{A}}$, one of the following holds:

- If $\mathcal{A}$ is homogeneous, then

$$\sqrt{(S_n \cdot f)} = (S_n \cdot x_1 x_2 \cdots x_k).$$

- If $\mathcal{A}$ is inhomogeneous, then

$$\mathcal{V}(S_n \cdot f) \subseteq \mathcal{V}(S_n \cdot x_1 x_2 \cdots x_k) \cup \mathcal{V}(x_i^e - x_j^e : i, j = 1, \ldots, n)$$

  for some $1 \leq e \leq \deg(f)$.

Remark 4.1.3. We collect several easy facts.

(1) Assume $\mathcal{A}$ is homogeneous but not necessarily symmetric and contains a power of some variable x_i^d. Let $G \subseteq S_n$ act transitively on the variables. Then, for the general $f \in K^{\mathcal{A}}$, $\sqrt{(G \cdot f)} = (x_1, \ldots, x_n)$ is the homogeneous maximal ideal, in particular $\mathcal{V}(G \cdot f) = \{0\}$. Indeed, for every $r \geq d$ the dimension of the degree r piece of the ideal $(G \cdot f)$ is precisely the rank of the matrix whose rows are the coefficients of the polynomials $m \cdot (g \cdot f)$, written in the monomial basis, where $g \in G$ and m is a monomial of degree $r - d$. Since the rank is lower semicontinuous, for any fixed r the general $f \in K^{\mathcal{A}}$ satisfies

$$\dim_K K[x_1, \ldots, x_n]_r / (G \cdot f)_r \leq \dim_K K[x_1, \ldots, x_n]_r / (G \cdot x_i^d)_r.$$

For $r \geq nd$, the right hand side is zero. Therefore, for any fixed r large enough, $(G \cdot f)_r$ is spanned by all monomials of degree r, including the powers of the variables. This implies immediately $\sqrt{(G \cdot f)} = (x_1, \ldots, x_n)$.

(2) Similarly, if all monomials in $\mathcal{A}$ are of the same type and $G \subseteq S_n$ acts transitively on the set of monomials of this type, then for general $f \in K^{\mathcal{A}}$ the ideal $(G \cdot f)$ is monomial, generated by the orbit of any term of f.

(3) It is a natural question whether the ideal generated by the orbit of a homogeneous polynomial with sufficiently general coefficients does not only have a monomial radical but is itself monomial. The answer is negative as Example 4.3.2 shows.

(4) If f is inhomogeneous, then $(S_n \cdot f)$ usually does not contain *any* monomial. Indeed, if f has at least two homogeneous parts f_i and f_j which do not vanish at $(1, 1, \dots, 1)$, then $f(t, t, \dots, t)$ is an inhomogeneous univariate polynomial which therefore has a non-zero solution in $\overline{K}$, in particular $\mathcal{V}_{\overline{K}}(S_n \cdot f)$ intersects the torus $(\overline{K} \setminus \{0\})^n$.

(5) Theorem 4.1.1 suggests that ideals generated by the orbit of a single polynomial should rarely be expected to be radical, see however Proposition 4.1.4 for a notable exception.

Proposition 4.1.4. Let $\operatorname{char}(K) = 0$ or $\operatorname{char}(K) > n$. Let $f \in K[x_1, \dots, x_n]$ be a homogeneous polynomial of degree d having only square-free terms. Assume $f(1, 1, \dots, 1) \neq 0$. Then, for all $N \geq n + d$, we have

$$(S_N \cdot f) = (S_N \cdot x_1 x_2 \cdots x_d) \subseteq K[x_1, \dots, x_N].$$

Proposition 4.1.4 can also be viewed as a representation-theoretic statement about the S_N-representation with a basis given by all d-element subsets of $\{1, \dots, N\}$. The character of this representation is known explicitly,[1] nonetheless the result does not seem to follow in a straightforward way. The proof given below is purely combinatorial.

Remark 4.1.5. If $\operatorname{char}(K) = 0$, Proposition 4.1.4 implies that the K-linear S_∞-representation V_d given by all d-element subsets of the natural numbers $\mathbb{N}$ has a unique maximal proper subrepresentation, namely the subvector space of V_d defined by all coefficients summing to zero. Indeed, the theorem implies that any element which does not lie in this subvector space generates all of V_d. We expect this to be known to experts on the representation theory of the infinite symmetric group.

Remark 4.1.6. Proposition 4.1.4 provides an example that the statements of Theorem 4.1.1 and Conjecture 4.1.2 might become simpler, with the possibility of obtaining the genericity conditions explicitly, in *large enough* polynomial rings, i.e., if the ideals are extended to symmetric ideals in polynomial rings with sufficiently many variables, or even to the infinite polynomial ring. This feature also appears in [MRV21].

Remark 4.1.7. Let $\mathcal{A} \subseteq \mathbb{Z}^n$ be a homogeneous support set and $f \in K^{\mathcal{A}}$. Experimentally, even the *saturation* of $(S_n \cdot f)$ at the homogeneous maximal ideal $(x_1, \dots, x_n)$ is very often a monomial ideal. It is an open problem whether this statement holds for the general $f \in K^{\mathcal{A}}$.

[1] see for example the following mathoverflow post:
mathoverflow.net/questions/123721/permutation-character-of-the-symmetric-group-on-subsets-of-certain-size

4.2 Proofs

For $n \geq 2$, there are precisely two S_n-representations of dimension 1, the trivial representation and the sign representation. For $n \geq 3$, $n \neq 4$, every irreducible S_n-representation of dimension > 1 has dimension at least $n-1$ (assuming characteristic zero). This follows from the classical representation theory of the symmetric groups, see for example [FH91]. If m_0 is a monomial, the *permutation module* M^{m_0} is the K-vector space generated by all permutations of m_0. The multiplicities of the irreducible S_n-representations in M^{m_0} are classically known as the *Kostka numbers*. We will use that the trivial representation has multiplicity 1 in M^{m_0} and is spanned by the monomial symmetric polynomial inside M^{m_0}. The sign representation has multiplicity 0 if m_0 has at least two equal exponents (also counting zeros). If on the other hand all exponents of m_0 are distinct, then the sign representation has multiplicity 1 in M^{m_0} and is spanned by the polynomial

$$\sum_{m \in A_n \cdot m_0} m - \sum_{m \in A_n \cdot m_0} (1,2).m,$$

having coefficient 1 in front of all even permutations of m_0 and coefficient -1 in front of all odd permutations of m_0. In particular, the sum of the isotypic components of the trivial and the sign representation in M^{m_0} has the property that for each element all coefficients of the even permutations of m_0 agree (and the same for the odd permutations).

Proof of Theorem 4.1.1. We assume $k \leq n-1$ since otherwise we may divide f by the appropriate power of $x_1 x_2 \cdots x_n$ and proceed with the resulting polynomial. We can also assume that $\mathcal{A}$ contains at least two monomials of different types since otherwise the conclusion is clear, see also Remark 4.1.3(ii). Moreover, we first let $K = \overline{K}$ and give the standard reduction to the case of an arbitrary field of characteristic zero at the end of the proof.
Elements of $K^{\mathcal{A}}$ are interpreted as polynomial functions with support contained in $\mathcal{A}$ and are written as coefficient vectors $(c_m)_{m \in \mathcal{A}}$. As $\mathcal{A}$ is symmetric, $K^{\mathcal{A}}$ is naturally an S_n-representation. We write elements of its dual representation $K^{\mathcal{A},*}$ in the dual basis $(y_m)_{m \in \mathcal{A}}$. We consider the partial Veronese

$$\varphi \colon \mathbb{P}^{n-1} \dashrightarrow \mathbb{P}K^{\mathcal{A},*},$$

sending $[x_1 : \cdots : x_n]$ to the homogeneous coordinate vector $[y_m]$ of all monomials in $\mathcal{A}$, so $y_m = m(x_1, \ldots, x_n)$. Set-theoretically, the indeterminacy locus of φ is then precisely $\mathcal{V}(S_n \cdot x_1 x_2 \cdots x_k) \subseteq \mathbb{P}^{n-1}$. We denote its complement by U and observe that $\varphi \colon U \to \mathbb{P}K^{\mathcal{A},*}$ is S_n-equivariant. With the idea of an incidence subset, we define the constructible subset $X \subseteq \mathbb{P}K^{\mathcal{A}} \times \mathbb{P}K^{\mathcal{A},*}$ as

$$X = \{([c_m], [y_m]) : [y_m] \in \varphi(U), \sum_{m \in \mathcal{A}} c_m y_{\sigma \cdot m} = 0 \text{ for all } \sigma \in S_n\}.$$

Denoting by pr_1, pr_2 the projections onto the factors, the constructible set $\mathrm{pr}_1(X) \subseteq \mathbb{P}K^{\mathcal{A}}$ is precisely the set of polynomials g for which $\sqrt{(S_n \cdot g)} \neq (S_n \cdot x_1 x_2 \cdots x_k)$.

Hence, it is enough to see that the dimension of X, and thus of $\mathrm{pr}_1(X)$, is at most $\dim(\mathbb{P}K^{\mathcal{A}}) - 1$. To prove the latter, we write X as a union of three constructible subsets which all satisfy this dimension bound. More precisely, we distinguish the following three sets of points $[y_m] \in \varphi(U) \subseteq \mathbb{P}K^{\mathcal{A},*}$:

(1) First, assume there is some $m_0 \in \mathcal{A}$ which is not a power of $x_1x_2\cdots x_n$ and such that $0 \neq y_{m_0} = y_{\sigma\cdot m_0}$ for all $\sigma \in A_n$. There are in fact only finitely many such points $[y_m] \in \varphi(U)$. In order to see this, for fixed m_0 let $Z \subseteq \mathbb{P}^{n-1}$ be the preimage under φ of the set of all these points $[y_m] \in \varphi(U)$. Then Z is contained in the intersection of $\mathcal{V}(m_0 - \sigma\cdot m_0 : \sigma \in A_n) \subseteq \mathbb{P}^{n-1}$ with the algebraic torus. Write $m_0 = x_1^{e_1}\cdots x_n^{e_n}$. Without loss of generality assume $e_1 \geq e_2 \geq e_3$ and $e_1 > e_3$. Then $m_0 - (1,2,3)\cdot m_0$ is a monomial multiple of $m' := x_1^{e_1-e_3} - x_2^{e_1-e_2}x_3^{e_2-e_3}$. Hence, Z is contained in the subset of the algebraic torus defined by all A_n-permutations of m'. Let $a := e_1 - e_2$ and $b := e_2 - e_3$. From the two equations $x_1^{a+b} = x_2^a x_3^b$ and $x_3^{a+b} = x_1^a x_2^b$ we deduce

$$x_1^{(a+b)^2} = x_2^{a(a+b)}(x_3^{a+b})^b = x_2^{a(a+b)}(x_1^a x_2^b)^b = x_2^{a^2+b^2+ab}x_1^{ab},$$

and hence $x_1^{a^2+b^2+ab} = x_2^{a^2+b^2+ab}$. Therefore, Z is contained in the closed subset of the algebraic torus defined by $x_1^e = x_2^e = \cdots = x_n^e$ with $e = (e_1-e_3)^2 - (e_1 - e_2)(e_2-e_3)$. Clearly, there are only finitely many solutions to these equations in $\mathbb{P}^{n-1}$, so the image of Z under φ is also a finite set of points. The fiber of $X \xrightarrow{\mathrm{pr}_2} \mathbb{P}K^{\mathcal{A},*}$ over each of these points is contained in some hyperplane inside $\mathbb{P}K^{\mathcal{A}}$.

(2) Second, assume there is some $m_0 \in \mathcal{A}$ which is not a power of $x_1x_2\cdots x_n$ and such that

- $\sum_{m\in S_n\cdot m_0} y_m \neq 0$ and
- not all y_m with m an even permutation of m_0 agree (or, without loss of generality, the same statement with respect to all odd permutations of m_0).

Then the fiber of $X \xrightarrow{\mathrm{pr}_2} \mathbb{P}K^{\mathcal{A},*}$ over such a point is a linear space of codimension at least n in $\mathbb{P}K^{\mathcal{A}}$. Indeed, this fiber is the projectivization of the kernel of the matrix Y whose rows are all S_n-permutations of $(y_m)_{m\in\mathcal{A}}$, and the codimension of this kernel in $K^{\mathcal{A}}$ is precisely the rank of Y. We claim that even the submatrix Y' of Y whose rows are all the S_n-permutations of $(y_m)_{m\in S_n\cdot m_0}$ has rank at least n. This is because the S_n-representation given by the span of all rows of Y' is a subrepresentation of the permutation module M^{m_0} which contains the trivial representation (just sum all the rows of Y') and also some other irreducible representation which is neither the trivial nor the sign representation. The latter follows from the assumption since every element of M^{m_0} lying in the sum of the isotypic components of the trivial and the sign representation would have equal coefficients for all monomials of the even permutations of m_0. Hence, using $\dim(\varphi(U)) \leq n-1$, the dimension of the preimage under $X \xrightarrow{\mathrm{pr}_2} \mathbb{P}K^{\mathcal{A},*}$ of the set of all such points $[y_m] \in \varphi(U)$ is at most

$$\dim(\varphi(U)) + (\dim(\mathbb{P}K^{\mathcal{A}}) - n) \leq \dim(\mathbb{P}K^{\mathcal{A}}) - 1.$$

(3) If the two cases above do not apply, then for all $m_0 \in \mathcal{A}$ which is not a power of $x_1 x_2 \cdots x_n$ we have $\sum_{m \in S_n \cdot m_0} y_m = 0$ but not all y_m with m an even permutation of m_0 agree. This translates into the codimension 1 condition $\sum_{m \in S_n \cdot m_0} m(x_1, \ldots, x_n) = 0$ on $\mathbb{P}^{n-1}$. With a similar argument as in the previous case, the fiber of any such point $[y_m] \in \varphi(U)$ has codimension at least $n-1$, making the dimensions add up to at most $\dim(\mathbb{P}K^{\mathcal{A}}) - 1$ again.

Now, we deduce the claim for arbitrary fields K of characteristic zero. Let $N := |\mathcal{A}|$. By the above, there is a non-empty principal open subset $D_{\overline{K}}(\alpha) \subseteq \mathbb{A}^N_{\overline{K}}$ for which the assertion of the theorem holds. But $\alpha \in \overline{K}[c_1, \ldots c_N]$ only has finitely many coefficients, so there is a finite field extension L of K such that $\alpha \in L[c_1, \ldots c_N]$. The integral ring extension $L[c_1, \ldots, c_N] \hookrightarrow \overline{K}[c_1, \ldots, c_N]$ induces the surjective morphism of affine schemes $b : \mathbb{A}^N_{\overline{K}} \to \mathbb{A}^N_L$, and as $b^{-1}(D_L(\alpha)) = D_{\overline{K}}(\alpha)$, we obtain $b(D_{\overline{K}}(\alpha)) = D_L(\alpha)$. The finite ring extension $K[c_1, \ldots, c_N] \hookrightarrow L[c_1, \ldots c_N]$ induces the finite surjective morphism $b' : \mathbb{A}^N_L \to \mathbb{A}^N_K$ of finite type K-schemes, to which Chevalley's theorem on constructible subsets applies. The dimensions of $\mathbb{A}^N_L$ and $\mathbb{A}^N_K$ agree, and by finiteness of b' the constructible image of the open dense subset $D_L(\alpha)$ under b' is necessarily dense in $\mathbb{A}^N_K$, hence contains a non-empty principal open $D_K(\beta) \subseteq \mathbb{A}^N_K$. Restricting to the set of K-rational points, $D_K(\beta) \cap K^N$ is still a non-empty open of the irreducible space K^N since K is infinite.
Finally, let $f \in D_K(\beta) \cap K^N$. Then f is identified with its corresponding polynomial $f \in K[x_1, \ldots, x_n]$ having $\operatorname{supp}(f) \subseteq \mathcal{A}$. As $f \in D_K(\beta)$, we have $\sqrt{(S_n \cdot f)} = (S_n \cdot x_1 x_2 \cdots x_k)$ in the polynomial ring over $\overline{K}$. Completing $\{1\}$ to a K-basis of $\overline{K}$, this equality also follows in the polynomial ring over K. $\square$

For smaller n or symmetric *inhomogeneous* support sets $\mathcal{A}$, the above proof can be adapted in many cases but the details are tedious and hence omitted.
Proposition 4.1.4 is a consequence of the following more special result. In order to state it, denote by $e^d_n(x_1, \ldots, x_n)$ the elementary symmetric polynomial of degree d in n variables, $1 \leq d \leq n$.

Proposition 4.2.1. Let $I = (S_N \cdot e^d_n(x_1, \ldots, x_n)) \subseteq K[x_1, \ldots, x_N]$ and $N \geq n + d$. If $\operatorname{char}(K) = 0$ or $\operatorname{char}(K) > n$, then

$$I = (S_N \cdot x_1 \cdots x_d).$$

Lemma 4.2.2. We have the identity

$$\binom{n-1}{d} \sum_{j=0}^{d} (-1)^j \frac{\binom{d-a}{j}\binom{n-d+a}{d-j}}{\binom{n-1}{d-j}} = \begin{cases} \binom{n}{d} & \text{for } a = d \\ 0 & \text{for all } 0 \leq a \leq d-1 \end{cases}$$

for all $1 \leq d \leq n-1$.

Proof. For $a = d$ this can be checked easily and the other case is equivalent to the well-known formula

$$\sum_{j=0}^{r} (-1)^j \binom{r}{j} (s+j)^{\underline{r-1}} = 0,$$

where $(s+j)^{\underline{r-1}}$ denotes the falling factorial and $r := d-a \geq 1$, $s := n-1-d \geq 0$. To see this, observe that this sum is just the r-th discrete derivative of the polynomial $s^{\underline{r-1}}$ which is of degree $r-1$ in s. Here, the discrete derivative of a polynomial $f(s)$ is defined as $(\Delta f)(s) = f(s+1) - f(s)$. Clearly, $\deg(\Delta f) \leq \deg(f) - 1$. It follows that $\Delta^{(r)} s^{\underline{r-1}} = 0$. □

Proof of Proposition 4.2.1. For a given field K, our goal is to find a K-linear combination of the polynomials $\sigma \cdot e_n^d$ with $\sigma \in S_{n+d}$ which equals a non-zero K-multiple of the monomial $x_1 \cdots x_d$. Let first $K = \mathbb{Q}$. We want to find coefficients $c_j \in \mathbb{Q}$ such that

$$\binom{n}{d} x_1 \cdots x_d = \sum_{j=0}^{d} (-1)^j c_j \sum_{|J_1|=d-j} e_n^d(x_J), \tag{4.1}$$

where the second sum ranges over all subsets $J \subseteq \{1, \ldots, n+d\}$ of cardinality n such that $J_1 := J \cap \{1, \ldots, d\}$ is of the given cardinality. Moreover, by $e_n^d(x_J)$ we denote the elementary symmetric polynomial of degree d in the n variables indexed by J. It is easy to see that only the summand with $j = 0$ contributes elementary symmetric polynomials that contain the monomial $x_1 \cdots x_d$, and there are precisely $\binom{n}{d}$ of those, which forces $c_0 = 1$ for equation (4.1) to hold. Note moreover that in the j-th summand of (4.1), all occuring $e_n^d(x_J)$ only contain monomials containing at most $d-j$ of the variables $x_1, \ldots, x_d$. More precisely, given a square-free monomial $x_A x_B$ with $A \subseteq \{1, \ldots d\}$, $B \subseteq \{d+1, \ldots, d+n\}$ of degree d we write $a = |A|$, so $|B| = d - a$. Then in the j-th summand of the sum in (4.1), the monomial $x_A x_B$ occurs exactly $\binom{d-a}{j}\binom{n-d+a}{d-j}$ times as a counting argument shows. Obviously, this does not depend on the sets A and B but only on the cardinality a of A. Now define c_1 in a way such that the monomials $x_A x_B$ with $a = d-1$ in the $j = 1$ summand cancel with the corresponding terms in the $j = 0$ summand. Clearly, there exists a unique such $c_1 \in \mathbb{Q}$, and the $j = 1$ summand does not contribute monomials with $a = d$. Similarily, define then $c_2 \in \mathbb{Q}$ to be the unique rational number such that the monomials $x_A x_B$ with $a = d-2$ in the $j = 2$ summand cancel out all the corresponding terms in the $j = 0$ and $j = 1$ summands. Again, the $j = 2$ summand cannot contribute any monomials with $a \geq d-1$. Continuing in this way, we define unique numbers $c_0, \ldots, c_d \in \mathbb{Q}$ depending only on n and d with $c_0 = 1$ and such that (4.1) must hold by construction. Now fix a monomial $x_A x_B$. Then the fact that this monomial has coefficient $\binom{n}{d}$ if $a = d$ and coefficient 0 otherwise on the right hand side of the equation (4.1) precisely translates into the binomial identity

$$\sum_{j=0}^{d} (-1)^j c_j \binom{d-a}{j} \binom{n-d+a}{d-j} = \begin{cases} \binom{n}{d} & \text{for } a = d \\ 0 & \text{for all } 0 \leq a \leq d-1 \end{cases}$$

for all $1 \leq d \leq n-1$. From the uniqueness of the c_j and Lemma 4.2.2 we then obtain $c_j = \frac{\binom{n-1}{d}}{\binom{n-1}{d-j}}$. Next, as we now know that the binomial identity of Lemma 4.2.2 holds over every field K in which all expressions are defined, if additionally $\operatorname{char}(K) \nmid \binom{n}{d}$, then the monomial $x_1 \cdots x_d$ lies in the ideal I by (4.1). Equivalently, a sufficient assumption on $\operatorname{char}(K)$ for the elimination in (4.1) to work is that $\operatorname{char}(K)$ does not divide $\binom{n}{d}$ nor any denominator of the reduced fractions c_j for all $0 \le j \le d$. In particular, $\mathrm{char}(K) = 0$ or $\mathrm{char}(K) > n$ will suffice. $\square$

Proof of Proposition 4.1.4. Let $\mathcal{J}$ be the set of all d-element subsets of $\{1, \ldots, n\}$ corresponding to terms of f, so that there are $c_J \in K$ for all $J \in \mathcal{J}$ such that $f = \sum_{J \in \mathcal{J}} c_J x_J$. We write $c := f(1, 1, \ldots, 1) = \sum_{J \in \mathcal{J}} c_J$. The following polynomial clearly lies in the ideal $(S_N \cdot f)$:

$$\begin{aligned}\sum_{\sigma \in S_n} \sigma \cdot f &= \sum_{J \in \mathcal{J}} c_J \sum_{\sigma \in S_n} \sigma \cdot x_J \\ &= \left(\sum_{J \in \mathcal{J}} c_J \right) d!(n-d)! e_n^d(x_1, \ldots, x_n) = c \cdot d!(n-d)! e_n^d(x_1, \ldots, x_n).\end{aligned}$$

If $c \neq 0$ in K, we deduce $e_n^d(x_1, \ldots, x_n) \in (S_N \cdot f)$ and so $(S_N \cdot f) = (S_N \cdot x_1 x_2 \cdots x_d)$ by Proposition 4.2.1. $\square$

4.3 Examples

We start with a small addition to Proposition 4.2.1 with slightly weaker assumptions on the characteristic of the field and on the number of variables N.

Lemma 4.3.1. Let $I = (S_N \cdot e_n^d(x_1, \ldots, x_n)) \subseteq K[x_1, \ldots, x_N]$ where either $d > 1$ and $N \ge n + d - 1$ or $d = 1$ and $N \ge n + 1$. If $\mathrm{char}(K) \nmid \binom{n}{d}$, then

$$\sqrt{I} = (S_N \cdot x_1 \cdots x_d).$$

Otherwise, if $\mathrm{char}(K) \mid \binom{n}{d}$, then for all $N \ge n$ the ideal $\sqrt{I}$ does not contain any monomial.

Proof. We proceed in several steps.

- First, we show that the polynomial $f := (x_1 - x_{n+1}) \cdots (x_d - x_{n+d})$ lies in the ideal I and hence so do all its permutations. Indeed, write

  $$\begin{aligned}f_1 &:= e_n^d(x_1, \ldots, x_n) - (1, n+1) \cdot e_n^d(x_1, \ldots, x_n) \\ &= (x_1 - x_{n+1}) e_{n-1}^{d-1}(x_2, \ldots, x_n) \in I.\end{aligned}$$

  Then, similarly,

  $$\begin{aligned}f_2 &:= f_1 - (2, n+2) \cdot f_1 \\ &= (x_1 - x_{n+1})(x_2 - x_{n+2}) e_{n-2}^{d-2}(x_3, \ldots, x_n) \in I.\end{aligned}$$

  Inductively, we obtain $f \in I$, as desired.

- Next, we show that every non-zero entry of $y \in \mathcal{V}(I) \subseteq K^N$ occurs at most $n-1$ times. Indeed, if an entry $t \in K$ occurs at n places $i_1 < \cdots < i_n$, then

  $$0 = e_n^d(y_{i_1}, \ldots, y_{i_n}) = \binom{n}{d} t^d,$$

  hence $t = 0$ by our assumption on $\mathrm{char}(K)$.

- In fact, we claim that every non-zero entry of $y \in \mathcal{V}(I)$ can occur at most $d-1$ times. Otherwise, if the entry $t \neq 0$ occurred r times, where $d \leq r \leq n-1$, then there would be at least $N-r \geq d$ entries of y different from t since $N \geq n+d-1$. This, however, contradicts $f(\sigma \cdot y) = 0$ for all σ because we can find a permutation σ after which $y_1 = \cdots = y_d = t$ and $y_{n+1}, \ldots, y_{n+d}$ are all different from t and so we would have $f(y) \neq 0$ which is impossible. Therefore, $\mathcal{V}(I) = \mathcal{V}(S_N \cdot x_1 \cdots x_d)$ and hence $\sqrt{I} = (S_N \cdot x_1 \cdots x_d)$.
- If $\mathrm{char}(K) \mid \binom{n}{d}$, we observe $e_n^d(1,1,\ldots,1) = \binom{n}{d} = 0$ in K, so $(1,1,\ldots,1)$ lies in $\mathcal{V}(I)$. $\square$

Example 4.3.2. Let $f \coloneqq x_1^2 + tx_1x_2$. Then for all $t \neq 0$ the ideal $(S_n \cdot f)$ does not contain x_1^2 for any $n \geq 2$. Suppose to the contrary that x_1^2 can be written as

$$x_1^2 = \sum_{\sigma \in S_n} c_\sigma (\sigma \cdot f) = \sum_{\sigma \in S_n} c_\sigma x_{\sigma(1)}^2 + t \sum_{\sigma \in S_n} c_\sigma x_{\sigma(1)} x_{\sigma(2)}$$

for some $c_\sigma \in K$. Then $\sum_{\sigma(1)=1} c_\sigma = 1$ and $\sum_{\sigma(1)\neq 1} c_\sigma = 0$ by looking at the sum only involving squared variables. The sum only involving square-free variables becomes

$$0 = tx_1 \sum_{\sigma(1)=1} c_\sigma x_{\sigma(2)} + t \sum_{\sigma(1)\neq 1} c_\sigma x_{\sigma(1)} x_{\sigma(2)}.$$

Now, we set all variables $x_n = \cdots = x_3 \coloneqq x_2$ (but x_1 stays unchanged). Then the first sum becomes simply tx_1x_2, and the second sum, after splitting it up as

$$t \sum_{\sigma(1)\neq 1 \neq \sigma(2)} c_\sigma x_{\sigma(1)} x_{\sigma(2)} + t \sum_{\sigma(2)=1} c_\sigma x_{\sigma(1)} x_{\sigma(2)},$$

becomes

$$tx_2^2 \left(\sum_{\sigma(1)\neq 1 \neq \sigma(2)} c_\sigma \right) + tx_1x_2 \sum_{\sigma(2)=1} c_\sigma.$$

From $t \neq 0$ it follows that $\sum_{\sigma(1)\neq 1\neq \sigma(2)} c_\sigma = 0$ and $\sum_{\sigma(2)=1} c_\sigma = -1$. But then,

$$-1 = \sum_{\sigma(1)\neq 1 \neq \sigma(2)} c_\sigma + \sum_{\sigma(2)=1} c_\sigma = \sum_{\sigma(1)\neq 1} c_\sigma = 0,$$

a contradiction.

Example 4.3.3. The genericity assumptions of Theorem 4.1.1 and Conjecture 4.1.2 are necessary. The obvious examples here are symmetric polynomials and polynomials vanishing at $(1,1,\ldots,1)$. For a less obvious example, we can take $f \coloneqq x_1^2x_2 + x_1x_2^2$ and $I = (S_N \cdot f) \subseteq \mathbb{Q}[x_1, \ldots, x_N]$, $N \geq 3$. For $N = 3$, a computation in `Macaulay2` shows

$$\sqrt{I} = (S_3 \cdot f, x_1x_2x_3),$$

so for all $N \geq 3$ we have $x_1x_2x_3 \in \sqrt{I}$. However, no permutation of x_1x_2 lies in $\sqrt{I}$ because all permutations of $(1,-1,0,0,\ldots,0)$ lie in $\mathcal{V}(S_N \cdot f)$. Hence, all monomials in I and $\sqrt{I}$ are divisible by at least 3 distinct variables.

Example 4.3.4. It is possible even for the ideal generated by the orbit of an inhomogeneous polynomial to be monomial although it is a rare phenomenon as explained by Remark 4.1.3. An example is given by $I = (S_3 \cdot (x_1 + x_2 + x_1^2 - x_2^2)) \in K[x_1, x_2, x_3]$ for any field K of $\mathrm{char}(K) \neq 2$. Indeed, one has $I = (x_1, x_2, x_3)$ as follows from

$$2x_1 = (x_1 + x_2 + x_1^2 - x_2^2) + (x_3 + x_1 + x_3^2 - x_1^2) - (x_3 + x_2 + x_3^2 - x_2^2) \in I.$$

Example 4.3.5. Proposition 4.1.4 and Proposition 4.2.1 are false in general for $N < n + d$. Consider the case $n = 3$, $d = 2$, $N = 4 < n + d$ for the ideal $I = (S_N \cdot e_3^2) \subseteq \mathbb{Q}[x_1, x_2, x_3, x_4]$. Then a computation in `Macaulay2` shows that all permutations of $x_1^2 x_2$ lie in I but no monomial of degree 2. In particular, I is not radical. Note that Lemma 4.3.1 still applies.

Example 4.3.6. The statement of Proposition 4.2.1 is false in general under the weaker assumption on the characteristic $\mathrm{char}(K) \nmid \binom{n}{d}$. As an example, consider the case $K = \mathbb{Z}/2\mathbb{Z}$ with $n = 3$, $d = 2$ and $N = n + d = 5$ for the polynomial $e_3^2 = x_1x_2 + x_1x_3 + x_2x_3$. Then $(x_1x_2)^2 \in I = (S_N \cdot e_3^2)$ but $x_1x_2 \notin I$, as can be checked by computing a Gröbner basis of I. The same is still true for $N = 6, 7$.

5 Thin Lattice Polytopes

For questions of authorship, please refer to pages IVf.

This chapter is based on the accepted version of [BKN23].

The Local h^*-Polynomial

We propose to investigate *thin polytopes*: lattice polytopes with vanishing local h^*-polynomials. Local h^*-polynomials are also called ℓ^*-polynomials or $\tilde{S}$-polynomials. In the case of lattice simplices, they equal the so-called box polynomial, see Example 5.1.15. Thin simplices were first defined in the context of regular A-determinants and A-discriminants by Gelfand, Kapranov and Zelevinsky [GKZ94, Section 11.4.B] as those lattice simplices whose Newton numbers are zero, see Remark 5.2.2. As has been noted in [GKZ94], "a classification of thin lattice simplices seems to be an interesting problem in the geometry of numbers." In this chapter, we extend this endeavor to thin lattice polytopes, which we throughout refer to for simplicity as thin polytopes. Our main results are a complete classification of thin polytopes up to dimension 3 (Theorem 5.3.3) and a characterization of thin Gorenstein polytopes in any dimension (Theorem 5.5.3). The latter relies crucially on a recent non-negativity result by Katz and Stapledon [KS16, Theorem 6.1]. As a consequence, we solve the original problem of [GKZ94] in these two cases and answer questions posed by Borisov, Nill and Schepers that came up in the investigation of stringy E-polynomials of Gorenstein polytopes.

We hope for a renewed interest in the study of the local h^*-polynomial as a fundamental invariant of a lattice polytope with many fruitful connections as pioneered in the work of Stanley [Sta92], Karu [Kar08], Batyrev, Borisov, Mavlyutov [BN08, BM03], Schepers [Sch12, NS13], and Katz, Stapledon [KS16].

Let us give an overview of this chapter. In Section 5.1 we give a comprehensive survey on the local h^*-polynomial of a lattice polytope. In Section 5.2 we define thin polytopes, present the main examples and discuss several open questions (e.g., Question 5.2.16). Section 5.3 contains the complete classification of three-dimensional thin polytopes. In particular, we prove that three-dimensional lattice simplices are thin if and only if they are lattice pyramids (Corollary 5.3.10). Section 5.5.1 presents the characterization of thin Gorenstein polytopes (Theorem 5.5.3). In particular, we deduce that thin Gorenstein polytopes have lattice width 1 (Corollary 5.5.7), being thin is invariant under the duality of Gorenstein polytopes (Corollary 5.5.12) and show that Gorenstein simplices are thin if and only if they are lattice pyramids (Corollary 5.5.15). To obtain these results, we study the behavior of local h^*-polynomials under joins, particularly for Gorenstein polytopes in Section 5.4.

5.1 A Primer on Combinatorial Invariants of Lattice Polytopes

Since the local h^*-polynomial is still not as well known in Ehrhart theory as the usual h^*-polynomial and also has been studied with different names and notations, we will give a slightly more thorough account on previous research than strictly necessary for the mere purpose of our results.

5.1.1 Toric g- and h-Polynomials of Lower Eulerian Posets

In [Sta87], Stanley generalized the notion of h-vectors of simplicial complexes and simplicial polytopes significantly. For this, let us recall some basic terminology.

Definition 5.1.1. The dual of a finite poset $\mathcal{P}$ is denoted $\mathcal{P}^*$. A finite poset $\mathcal{P}$ is *locally graded* if every inclusion-maximal chain in every interval $[x, y]$ has the same length $r(x, y)$. The *rank* $\operatorname{rk}(\mathcal{P})$ is the length of the longest chain in $\mathcal{P}$. If in addition there exists a rank function $\rho : \mathcal{P} \to \mathbb{Z}$, i.e., $r(x, y) = \rho(y) - \rho(x)$ for every interval $[x, y]$, then $\mathcal{P}$ is called *ranked*. If $\mathcal{P}$ is ranked and every interval $[x, y]$ with $x \neq y$ has the same number of even rank and odd rank elements, then $\mathcal{P}$ is *locally Eulerian*. If $\mathcal{P}$ is locally Eulerian and contains a minimal element $\hat{0}$, then it is called *lower Eulerian*. If it also contains a maximum $\hat{1}$, then $\mathcal{P}$ is called *Eulerian*. In presence of a minimum $\hat{0}$ in a ranked poset $\mathcal{P}$, we will always assume that the rank function satisfies $\rho(\hat{0}) = 0$.

Here is the definition of the g-polynomial and the h-polynomial for lower Eulerian posets according to Stanley [Sta87].

Definition 5.1.2. Let $\mathcal{P}$ be a lower Eulerian poset with rank function ρ and rank d. We define the *g-polynomial* $g_{\mathcal{P}}(t)$ and the *h-polynomial* $h_{\mathcal{P}}(t)$ recursively by introducing a third polynomial $f_{\mathcal{P}}(t)$ as an intermediate step. Let

$$f_{\emptyset}(t) = g_{\emptyset}(t) = h_{\emptyset}(t) = 1$$

and if $\mathcal{P} \neq \emptyset$, we set

$$f_{\mathcal{P}}(t) = \sum_{x \in \mathcal{P}} (t-1)^{d-\rho(x)} g_{[\hat{0},x)}(t)$$

and define for $f_{\mathcal{P}}(t) = \sum_{i=0}^{d} f_i t^i$,

$$g_{\mathcal{P}}(t) = \sum_{i=0}^{\lfloor d/2 \rfloor} (f_i - f_{i-1}) t^i, \text{ and}$$

$$h_{\mathcal{P}}(t) = \sum_{i=0}^{d} f_{d-i} t^i.$$

Hence, $h_{\mathcal{P}}(t)$ is a polynomial with constant term 1 of degree $\leq d$, and $g_{\mathcal{P}}(t)$ is a polynomial with degree $\leq d/2$.

Remark 5.1.3. If for $x \in \mathcal{P}$, the interval $[\hat{0}, x]$ is boolean, then $g_{[\hat{0},x)}(t) = 1$, see [Sta87, Proposition 2.1].

Remark 5.1.4. Let us recall the situation of simplicial complexes Δ (see [Sta87]), where the previous definition of the h-polynomial agrees with the usual one. For this, we identify Δ with its face poset which is a lower Eulerian poset with minimum $\emptyset \in \Delta$. Throughout, we use the convention that $\dim(\emptyset) = -1$. It follows from Remark 5.1.3 that $g_{[\emptyset,\sigma)}(t) = 1$ for all faces σ of Δ. If Δ has dimension $d-1$, we get $f_\Delta(t) = \sum_{i=0}^{d} f_{i-1}(t-1)^{d-i}$, where f_j denotes the number of faces of Δ of dimension j. Hence, this implies

$$h_\Delta(t) = \sum_{\sigma \in \Delta} t^{\dim(\sigma)+1}(1-t)^{d-1-\dim(\sigma)},$$

which indeed equals the usual h-polynomial of Δ, and where its coefficients form the usual h-vector of Δ [Sta87, p. 199]. For instance, if Δ is the boundary complex of a d-dimensional simplex, then $h_\Delta(t) = 1 + t + \cdots + t^d$. Let us also give one example to illustrate that the previous formula for $h_\Delta(t)$ fails in the non-simplicial situation. Let P be the pyramid over the square. In this case, the h-polynomial of the boundary complex of P equals $1 + 2t + 2t^2 + t^3$, while the previous formula would give $1 + 2t + t^2 + t^3$. Note that the h-polynomial is palindromic while the latter expression is not.

Stanley proved in [Sta87, Theorem 2.4] the following combinatorial palindromicity result generalizing the Dehn–Sommerville equations for face numbers of simplicial polytopes.

Theorem 5.1.5. Let $\hat{\mathcal{P}}$ be an Eulerian poset and $\mathcal{P} := \hat{\mathcal{P}} \setminus \hat{1}$ with $\mathrm{rk}(\mathcal{P}) = d$. Then the h-polynomial $h_{\mathcal{P}}(t) = \sum_{i=0}^{d} h_i t^i$ is palindromic of degree d, i.e. $h_i = h_{d-i}$ for all $i = 0, \ldots, d$.

In particular, we have $f_{\mathcal{P}}(t) = h_{\mathcal{P}}(t)$ in this case.

Remark 5.1.6. We emphasize that in the situation of Theorem 5.1.5 it is important to distinguish between the g- and h-polynomials of $\mathcal{P}$ and $\hat{\mathcal{P}}$. Indeed, $g_{\hat{\mathcal{P}}}(t) = 0$ and $h_{\hat{\mathcal{P}}}(t) = g_{\mathcal{P}}(t)$. Unfortunately, in this regard the different notations employed in the literature can be confusing. Our notation follows that of Stanley while Katz and Stapledon in [KS16] write $g(\hat{\mathcal{P}}; t)$ for our $g_{\mathcal{P}}(t)$ but also use $h(\mathcal{P}; t)$ for our $h_{\mathcal{P}}(t)$. In Borisov and Mavylutov [BM03], as well as in [BN08, NS13], our $g_{\mathcal{P}}(t)$ and $h_{\mathcal{P}}(t)$ would be $g_{\hat{\mathcal{P}}^*}(t)$ and $h_{\hat{\mathcal{P}}^*}(t)$.

Let us give the definition of h- and g-polynomials of polytopes.

Definition 5.1.7. For P a polytope we define its *(toric) h-polynomial* $h_P(t)$, and its *(toric) g-polynomial* $g_P(t)$ as the h-, resp., g-polynomial of the face lattice $[\emptyset, P)$ of proper faces of P. Note that $g_P(t) = h_{[\emptyset,P]}(t)$, see Remark 5.1.6.

Note that by Remark 5.1.3, we have $g_P(t) = 1$ if P is a simplex.

Theorem 5.1.8. Let P be a polytope of dimension d. Then $h_P(t) = \sum_{i=0}^{d} h_i t^i$ is a palindromic polynomial with positive integer coefficients that form a unimodal sequence, i.e.,

$$1 = h_0 \leq h_1 \leq \cdots \leq h_{\lfloor \frac{d}{2} \rfloor}.$$

Equivalently, $g_P(t)$ has non-negative coefficients.

Proof. Palindromicity follows from Theorem 5.1.5. For rational polytopes P, nonnegativity follows from the interpretation of the coefficients of $h_P(t)$ as the dimensions of the even intersection cohomology groups of the toric variety associated with P and the unimodality property follows from the hard Lefschetz theorem [Sta87, Theorem 3.1, Corollary 3.2]. The non-rational case has been treated by Karu in [Kar04]. □

Let us mention the following less well-known duality property of g-polynomials that will be of importance in Section 5.5.1. This is a result by Kalai, published in [Bra06, Theorem 4.5] as a consequence of the main result in that paper by Braden. Here, $\mathcal{P}^*$ denotes the *dual* poset of a poset $\mathcal{P}$.

Theorem 5.1.9. Let P be a polytope. Then

$$\deg(g_{[\emptyset,P)}) = \deg(g_{(\emptyset,P]^*}).$$

In other words, if Q is any polytope which is combinatorially dual to P, then $\deg(g_P) = \deg(g_Q)$.

5.1.2 (Relative) Local h-Polynomials of Polyhedral Subdivisions

We give the definition of the local h-polynomial (and its generalized relative version) of a polyhedral subdivision Δ of a polytope P, following [Sta92] and [KS16] (the relative version was introduced independently in [Ath12] and [NS12]). Here, we define the *link* of a face $\sigma \in \Delta$ as $\operatorname{link}(\Delta, \sigma) := \{\rho \in \Delta : \sigma \subseteq \rho\}$. We view $\operatorname{link}(\Delta, \sigma)$ as a lower Eulerian poset with minimum σ.

Definition 5.1.10. Let P be a polytope, Δ a polyhedral subdivision of P, and $\sigma \in \Delta$. The *relative local h-polynomial* of Δ with respect to σ is defined as

$$\ell_{\Delta,\sigma}(t) := \sum_{\sigma \subseteq F \leq P} (-1)^{\dim(P)-\dim(F)} h_{\operatorname{link}(\Delta_F,\sigma)}(t)\, g_{(F,P]^*}(t),$$

where $F \leq P$ means that F is a face of P (including $\emptyset$ and P) and $\Delta_F := \{\rho \in \Delta \colon \rho \subseteq F\}$.
We call $\ell_\Delta(t) := \ell_{\Delta,\emptyset}$ the *local h-polynomial* of Δ.

We suppress P in this notation as it equals $|\Delta| = \bigcup_{\sigma \in \Delta} \sigma$, the support of Δ. We remark that the same definition of the local h-polynomial can be extended to so called strong formal subdivisions of Eulerian posets, see [KS16, Definition 4.1].

Theorem 5.1.11. Let P be a polytope of dimension d, Δ a polyhedral subdivision of P, and $\sigma \in \Delta$. Then we can write $\ell_{\Delta,\sigma}(t) = \sum_{i=0}^{d-\dim(\sigma)} \ell_i t^i$. Moreover, the following holds:

(1) $\ell_{\Delta,\sigma}(t)$ has nonnegative integer coefficients.

(2) $\ell_{\Delta,\sigma}(t)$ is palindromic, i.e. $\ell_i = \ell_{d-\dim(\sigma)-i}$ for $i = 0, \ldots, d - \dim(\sigma)$.

(3) If Δ is regular, then the coefficients of $\ell_{\Delta,\sigma}(t)$ form a unimodal sequence.

Proof. Let us give the references. (2): For the local h-polynomial this is a special case of [Sta92, Corollary 7.7], for the relative local h-polynomial see [KS16, Corollary 4.5]. (1) and (3): For Δ a rational polyhedral subdivision this has been proven in [Sta92, Theorem 7.9], respectively, [KS16, Theorem 6.1] using the decomposition theorem (cf. [BBD82, dCM09, dCMM18]). As pointed out in [KS16, Remark 6.6], the only missing ingredient to drop the rationality hypothesis was the relative hard Lefschetz theorem for the intersection cohomology of fans which was subsequently proven in [Kar19]. □

The following decomposition theorem was one of the main motivations of Stanley for the notion of *local* h-vectors. This is proven in [Sta92, Theorem 7.8], and the general version in [KS16] (see, e.g., second equation in proof of Lemma 6.4). To stress the analogy to Theorem 5.1.17, we state the equality also using the h-polynomial.

Proposition 5.1.12. Let P be a polytope of dimension d, Δ a polyhedral subdivision of P, and $\sigma \in \Delta$. Then

$$h_{\mathrm{link}(\Delta,\sigma)}(t) = \sum_{\sigma \subseteq F \leq P} \ell_{\Delta_F,\sigma}(t) g_{[F,P)}(t) = \sum_{\sigma \subseteq F \leq P} \ell_{\Delta_F,\sigma}(t) h_{[F,P]}(t).$$

In particular for $\sigma = \emptyset$, we get $\ell_\Delta(t) \leq h_\Delta(t)$ and $g_P(t) = h_{[\emptyset,P]} \leq h_\Delta(t)$ coefficientwise.

As a consequence of the above nonnegativity results, Stanley and later Katz and Stapledon show that h-polynomials as well as relative local h-polynomials of polyhedral subdivisions are nonnegative and coefficientwise monotone under subdivision refinement [KS16, Corollary 6.10].

5.1.3 The h^*-Polynomial of a Lattice Polytope

We quickly recall the basic notions of Ehrhart theory. Let $P \subseteq \mathbb{R}^d$ be a lattice polytope with respect to some lattice $M \subseteq \mathbb{R}^d$ of maximal rank d, usually $M = \mathbb{Z}^d$. The *Ehrhart series of* P (with respect to M) is the formal power series

$$\mathrm{Ehr}_P(t) := 1 + \sum_{n \geq 1} |(nP) \cap M| t^n \in \mathbb{Z}[[t]].$$

By a theorem of Ehrhart [Ehr62], the map $\mathrm{ehr}_P \colon \mathbb{Z}_{\geq 1} \to \mathbb{Z},\ n \mapsto |(nP) \cap M|$ is a polynomial in n, called the *Ehrhart polynomial of* P. It has degree $\dim(P)$, constant term 1 and leading coefficient equal to the volume of P normalized with respect to M. It follows that

$$\mathrm{Ehr}_P(t) = \frac{h^*_P(t)}{(1-t)^{\dim(P)+1}},$$

for a unique polynomial $h^*_P(t) \in \mathbb{Z}[t]$ of degree at most $\dim(P)$, called the *h^*-polynomial of P*. Here, $h^*_P(t)$ has non-negative integer coefficients by [Sta80]. Moreover, $h^*_P(0) = 1$ and $h^*_P(1)$ equals the *lattice volume* $\mathrm{vol}_{\mathbb{Z}}(P) \in \mathbb{Z}_{\geq 1}$, which is defined as $\dim(P)!$ times the volume of P normalized with respect to M. Note that the lattice volume of a lattice point equals 1. The *degree of P*, denoted $\deg(P)$, is the degree of its h^*-polynomial $h^*_P(t)$. The *codegree of P*, denoted $\mathrm{codeg}(P)$, is the smallest integer $k \geq 1$ such that the k-th dilate kP of P contains a lattice point of M in its relative interior. By convention, a point has codegree 1. It follows from Ehrhart–MacDonald reciprocity [Mac71] that

$$\deg(P) + \mathrm{codeg}(P) = \dim(P) + 1.$$

Let us recall that two lattice polytopes P and Q (with respect to the lattice M) are called *isomorphic* (or *unimodularly equivalent*) if there is an affine lattice automorphism of M that maps the vertices of P to the vertices of Q. We say P is a *unimodular simplex* if P is isomorphic to the convex hull of an affine lattice basis of M. Now, P is a unimodular simplex if and only if $h^*_P(t) = 1$, or equivalently, $\deg(P) = 0$. For $M = \mathbb{Z}^d$ let us also define the *standard unimodular simplex* $\Delta_d := \mathrm{conv}(0, e_1, \ldots, e_d)$ for the standard lattice basis $e_1, \ldots, e_d$.

5.1.4 The Local h^*-Polynomial of a Lattice Polytope

Let us introduce our main player, see [Sta92, Example 7.13] and [KS16, Definition 7.2].

Definition 5.1.13. Let P be a lattice polytope. The *local h^*-polynomial* or *ℓ^*-polynomial* of P is defined as

$$\ell^*_P(t) := \sum_{\emptyset \leq F \leq P} (-1)^{\dim(P)-\dim(F)} h^*_F(t) g_{(F,P]^*}(t).$$

Let us note that the local h^*-polynomial of the empty face equals 1, while for a point it equals 0. We also emphasize the analogy of Definition 5.1.13 with Definition 5.1.10 above. See also Subsection 5.1.5 for precise relationships between the h, ℓ, h^*, ℓ^*-polynomials.

Remark 5.1.14. The local h^*-polynomial has been studied by Batyrev, Borisov, Mavlyutov, Nill and Schepers under the name *$\tilde{S}$-polynomial*, see [BM03, Definition 5.3]. It was used by Borisov and Mavlyutov to simplify the formulas for the stringy E-polynomial of Calabi–Yau complete intersections in Gorenstein toric Fano varieties originally described via so-called B-polynomials [BB96a]. We remark that the reader should be aware that in these papers in the definition of h- and g-polynomials the dual poset was used compared to the one given here.

Example 5.1.15. For lattice simplices P (of dimension $d > 0$) the h^*- and ℓ^*-polynomial can be easily understood, as in this case the face posets are all Boolean. Let Π denote the half-open parallelepiped spanned by the vertices of $P \times \{1\}$. Then $h^*_P(t)$ (resp., $\ell^*_P(t)$) enumerates the number of lattice points in Π (resp., in the interior of Π). More precisely, we have $h^*_P(t) = \sum_{i=0}^{d+1} h^*_i t^i$ and $\ell^*_P(t) = \sum_{i=0}^{d+1} \ell^*_i t^i$, where for $i = 0, \dots, d+1$ the coefficient h^*_i (resp. ℓ^*_i) equals the number of lattice points in Π (resp., in the interior of Π) with last coordinate i. We refer to [BN08, Prop. 4.6]. This polynomial $\ell^*_P(t)$ of a lattice simplex P is also often called *box polynomial*, cf. [Bra16, Sol19, GS20]. For instance, we have $h^*_P(t) = 1$ if and only if P is a unimodular simplex; in this case, $\ell^*_P(t) = 0$. Let us note that for h^*-polynomials this combinatorial interpretation of its coefficients can also be extended to arbitrary lattice polytopes, e.g., by half-open decompositions [KV08]. On the other hand, there is not yet a combinatorial counting interpretation for the coefficients of the local h^*-polynomial of lattice polytopes known.

Let us summarize some of the basic properties of the local h^*-polynomial. Throughout, we use the convention that the degree of the zero-polynomial is zero.

Theorem 5.1.16. Let P be a lattice polytope of dimension $d > 0$. Then we can write the local h^*-polynomial $\ell^*_P(t) = \sum_{i=1}^{d} \ell^*_i t^i$. Moreover, the following holds:

(1) $\ell^*_P(t)$ has nonnegative integer coefficients.

(2) $\ell^*_P(t)$ is palindromic: we have $\ell^*_i = \ell^*_{d+1-i}$ for $i = 1, \dots, d$.

(3) If $\ell^*_P(t)$ does not vanish, then the degree of $\ell^*_P(t)$ equals at most the degree of $h^*_P(t)$, and its subdegree (i.e., the smallest i such that the i-th coefficient of $\ell^*(t)$ is non-zero) is at least the codegree of P.

(4) $\ell^*_1 = \ell^*_d$ equals the number of lattice points in the interior of P.

Proof. Let us give the corresponding references: (1) This was conjectured by Stanley [Sta92, Conjecture 7.14] and proven by Karu [Kar08]. Using the $\tilde{S}$-notation for ℓ^* it also follows from its interpretation as the Hilbert function of a graded vector space by Borisov, Mavlyutov [BM03, Proposition 5.5]. (2) This was observed in [BM03, Remark 5.4]. (3) This follows directly, see also [NS13, Corollary 2.16(2)]. (4) For this observation, see [BN08, Example 4.7]. □

In particular, the number $\mathrm{int}_{\mathbb{Z}}(P)$ of interior lattice points completely determines the local h^*-polynomial up to dimension 2. If $d = 0$, then $\ell^*_P(t) = 0$; if $d = 1$, then $\ell^*_P(t) = \mathrm{int}_{\mathbb{Z}}(P)t$; and if $d = 2$, then $\ell^*_P(t) = \mathrm{int}_{\mathbb{Z}}(P)t + \mathrm{int}_{\mathbb{Z}}(P)t^2$.

5.1.5 Decomposing and Relating the h, ℓ, h^*, ℓ^*-Polynomials

The following classical result by Betke and McMullen, generalized by Katz and Stapledon, explains the relation of h^*-polynomials to h-polynomials of a *lattice* subdivision (i.e., a polyhedral subdivision whose vertices are lattice points). Recall that a lattice triangulation is called *unimodular* if all its simplices are unimodular simplices.

Theorem 5.1.17. Let P be a lattice polytope with a lattice subdivision Δ. Then the following holds:

$$h^*_P(t) = \sum_{\sigma \in \Delta} \ell^*_\sigma(t)\, h_{\mathrm{link}(\Delta,\sigma)}(t).$$

In particular, we have $h_\Delta(t) \leq h^*_P(t)$ coefficientwise, where we have equality if and only if the local h^*-polynomial of every non-empty face of Δ vanishes. If Δ is a lattice triangulation, then this is equivalent to Δ being a unimodular triangulation.

Proof. This is Lemma 7.12(3) of [KS16], generalizing [BM85]. We recall that the consequence follows from the nonnegativity of the occuring polynomials and the fact that the h-polynomials have constant term 1. Second, the combinatorial description of the h^*- and ℓ^*-polynomial of a lattice simplex, Example 5.1.15, implies that a lattice simplex S is a unimodular simplex if and only if $\ell^*_\sigma(t) = 0$ for all non-empty faces σ of S. □

Let us note that this result was one motivation for Stanley to define local h-polynomials, as these allowed him to prove an analogous result in the combinatorial setting, namely, Proposition 5.1.12 above. And similar to that formula positively expressing the h-polynomial of a subdivision into local h-polynomials and toric h-polynomials, one can also decompose the h^*-polynomial of a lattice polytope positively into local h^*-polynomials and toric h-polynomials of its face poset.

Corollary 5.1.18. Let P be a lattice polytope. Then

$$h^*_P(t) = \sum_{\emptyset \leq F \leq P} \ell^*_F(t) g_{[F,P)}(t) = \sum_{\emptyset \leq F \leq P} \ell^*_F(t) h_{[F,P]}(t).$$

In particular, $\ell^*_P(t) + g_P(t) = \ell^*_P(t) + h_{[\emptyset,P]}(t) \leq h^*_P(t)$ coefficientwise.

A proof in greater generality is given in [Sch12, Proposition 2.9], see also [Kar08, Corollary 1.1] and [NS13, Proposition 2.5]. We remark that Corollary 5.1.18 gives significance to thinking of the local h^*-polynomial as the "Ehrhart core" of the h^*-polynomial. This is most prominently clear in the case of lattice simplices, see Example 5.1.15.
Now, just as the (generalized) Betke–McMullen formula transparently separates the lattice data (the ℓ^*-polynomials of the cells) and combinatorial data (the h-polynomials of the links of the cells) of the h^*-polynomial of the support of a lattice subdivision, the same can be done for the local h^*-polynomial. This was observed in [NS12], see also [KS16, Lemma 7.12(4)].

Proposition 5.1.19. Let P be a lattice polytope with a lattice subdivision Δ. Then the following holds:

$$\ell^*_P(t) = \sum_{\sigma \in \Delta} \ell^*_\sigma(t)\; \ell_{\Delta,\sigma}(t).$$

In particular, $\ell_\Delta(t) \leq \ell^*_P(t)$ coefficientwise, with equality if Δ is a unimodular triangulation.

Here, we critically used the nonnegativity of the relative local h-polynomial, Theorem 5.1.11(1), for the consequence. In particular, we get another proof of the nonnegativity of the local h^*-polynomial. Moreover, as already observed in [NS12],this implies that the unimodality of the ℓ^*-vector is an *intrinsic* obstruction for a lattice polytope to have a unimodular triangulation (apply Theorem 5.1.11(3) with $\sigma = \emptyset$). In fact, it is enough to have unimodality of the "local box polynomials" (this is Remark 7.23 in [KS16]). Such triangulations where called *box unimodal* in [SVL13].

Corollary 5.1.20. If P admits a regular triangulation such that the local h^*-polynomials of each cell have unimodal coefficients (e.g., the triangulation is unimodular), then its local h^*-polynomial has unimodal coefficients.

This was used in [GS20] to prove the unimodality of the local h^*-polynomial of s-lecture hall order polytopes.

For our purposes, the following innocent looking consequence of the nonnegativity of relative local h-polynomials is crucial.

Corollary 5.1.21. Let P and P' be lattice polytopes such that P' is obtained from P by refining the lattice. Then $h^*_P(t) \leq h^*_{P'}(t)$ and $\ell^*_P(t) \leq \ell^*_{P'}(t)$ coefficientwise.

Proof. This follows from Theorem 5.1.17, respectively, Proposition 5.1.19, the explicit combinatorial description of the ℓ^*-polynomial of a lattice simplex, see Example 5.1.15, and the nonnegativity of the h-polynomial, respectively, of the relative local h-polynomial. $\square$

We remark that for h^*-polynomials this lattice-monotonicity can also easily be seen combinatorially, e.g., using half-open decompositions (see [BS18]). However, for local h^*-polynomials there seems to be no such combinatorial argument known. This is also true for the next result. Note that by Stanley's famous monotonicity result, the h^*-polynomial is coefficientwise monotone with respect to inclusion. However, this is not true for the local h^*-polynomial. Still, it holds when one considers subpolytopes that do not lie on the boundary.

Corollary 5.1.22. Let P and Q be lattice polytopes such that $\operatorname{relint}(Q) \subseteq \operatorname{int}(P)$ (for instance, $\dim(Q) = \dim(P)$). Then $\ell^*_Q(t) \leq \ell^*_P(t)$ coefficientwise.

Proof. Choose a lattice subdivision Δ of P that contains Q as a cell. Then by the nonnegativity of the appearing polynomials, we see from Proposition 5.1.19 that $\ell^*_Q(t)\ell_{\Delta,Q}(t) \leq \ell^*_P(t)$ coefficientwise. It remains to observe that since Q is in the relative interior of P, it follows directly from Definition 5.1.10 that $\ell_{\Delta,Q}(t) = h_{\operatorname{link}(\Delta,Q)}(t)$ and hence has constant coefficient 1. $\square$

Finally, let us just shortly mention that in the Katz–Stapledon paper [KS16], motivated by algebraic and tropical geometry, the h^*- and ℓ^*-polynomials are further refined to bivariate (and even trivariate) polynomials leading to the notion of h^*- and ℓ^*-diamonds. As we will use the following notation later for the computation of the local h^*-polynomial in dimension three, let us introduce it here.

Definition 5.1.23. Let P be a lattice polytope of dimension d. Then we define

$$h^*_P(u,v) := \sum_{F \le P} v^{\dim(F)+1} \ell^*_F(uv^{-1}) g_{[F,P)}(uv).$$

Note that $h^*_P(t,1) = h^*_P(t)$. In [KS16, Remark 7.7] it is shown that

$$h^*_P(u,v) = 1 + uv \sum_{0 \le p,q \le d-1} h^*_{p,q} u^p v^q,$$

where $h^*_{i,d-1-i} = \ell^*_i$ for $i = 1, \ldots, d$. These refined invariants satisfy many beautiful properties. Let us present here at least one such consequence, namely, the following lower bound theorem on the coefficients of the ℓ^*-polynomial [KS16, p.184].

Theorem 5.1.24. Let P be a lattice polytope of dimension d with $\ell^*_P(t) = \sum_{i=1}^d \ell^*_i t^*$. Then

$$\ell^*_1 \le \ell^*_i \quad \text{for } i = 2, \ldots, d.$$

5.2 Definition, Basic Properties and Examples of Thin Polytopes

5.2.1 Main Definition and Known Results

The following notion is the main focus of this chapter.

Definition 5.2.1. A lattice polytope P is called *thin* if its local h^*-polynomial ℓ^*_P vanishes. By the nonnegativity of the coefficients, Theorem 5.1.16(1), this is equivalent to $\ell^*_P(1) = 0$.

Let us note that lattice polytopes of dimension 0 as well as unimodular simplices are thin, see Example 5.1.15. We remark that thin polytopes naturally appear in Theorem 5.1.17.

Remark 5.2.2. Thin simplices were first investigated in [GKZ94, Section 11.4.B] in the context of regular A-determinants and A-discriminants, more precisely, in the characterization of so-called D-equivalence classes of regular triangulations of A. There a lattice simplex S was defined to be *thin* if its *Newton number* $\nu(S)$ equals zero. Here, the Newton number is defined as follows:

$$\nu(S) := \sum_{\emptyset \le F \le S} (-1)^{\dim(S)-\dim(F)} \mathrm{vol}_{\mathbb{Z}}(F) = 0, \tag{5.1}$$

where, $\mathrm{vol}_{\mathbb{Z}}(\emptyset) := 1$ (also $\mathrm{vol}_{\mathbb{Z}}(F) = 1$ if $\dim(F) = 0$). Recall from Example 5.1.15, that $\mathrm{vol}_{\mathbb{Z}}(F) = h^*_F(1)$ counts the number of lattice points in the half-open parallelepiped over F. Hence, by inclusion-exclusion, it is straightforward to deduce $\nu(S) = \ell^*_S(1)$, the number of interior lattice points in the half-open parallelepiped over S. Thus, for lattice simplices the definitions agree. Let us note that in [GKZ94] the nonnegativity of $\nu(S)$ follows from quite deep algebro-geometric arguments, while it is combinatorially obvious from the interpretation of ℓ^*_S as the box polynomial of the lattice simplex S. The reader should also be warned that the expression in equation (5.1) may be negative for lattice *polytopes*. For instance, it equals -1 for the 0/1-cube $[0,1]^3$.

Thin simplices were classified in [GKZ94] up to dimension 2. Here, we can deduce the following statement directly from Theorem 5.1.16(4). Let us define a lattice polytope to be *hollow* if it has no lattice points in its interior. Here, a 0-dimensional lattice polytope is not hollow (but thin).

Proposition 5.2.3. Thin polytopes of dimension > 0 are hollow. The converse also holds in dimensions 1 and 2.

In particular, Δ_1 is the only thin polytope of dimension 1. Hollow polytopes in dimension 2 are well-known. They are either isomorphic to $2\Delta_2$ or have lattice width 1 (i.e., all vertices lie on two parallel hyperplanes of lattice distance one). Note that hollow three-dimensional lattice polytopes do not have to be thin., e.g., $2\Delta_3$ and $[0,1]^3$ are not thin.
One important construction for thin polytopes is to take lattice pyramids.

Definition 5.2.4. Let $P \subset \mathbb{R}^d$ be a lattice polytope. Then

$$\operatorname{conv}(P \times \{0\}, \{0\} \times \{1\}) \subset \mathbb{R}^d \times \mathbb{R}$$

is called the *lattice pyramid* over P. By convention, a lattice point is also considered a lattice pyramid.

It is well-known that the h^*-polynomial, and particularly the degree, does not change under taking lattice pyramids. The following result has already been observed in [GKZ94] for lattice simplices and in general in [BN08] for lattice polytopes.

Proposition 5.2.5. Lattice pyramids over arbitrary lattice polytopes are thin.

Using this notation we can state the classification of thin simplices up to dimension two as follows.

Corollary 5.2.6. A lattice simplex of dimension at most ≤ 2 is thin if and only if it is isomorphic to $2\Delta_2$ or it is a lattice pyramid.

Remark 5.2.7. In [GKZ94] *thin triangulations* were intensively studied. Recently, this notion has also been investigated by [dMGP+20] where it was completely characterized up to dimension 3. As Stanley observed at the end of Section 7 in [Sta92], a thin triangulation may be defined by the vanishing of its local h-polynomial. Now, it follows from Proposition 5.1.19 that *all* lattice triangulations of thin polytopes are thin. This seems to be a quite strong combinatorial obstruction worth of further study.

Remark 5.2.8. By Corollary 5.1.21, a thin polytope stays thin if the lattice is coarsened. We do not know of a purely combinatorial proof of this fact.

Remark 5.2.9. If a lattice polytope is contained in a thin polytope but not in its boundary, then it is also thin. This non-trivial fact follows from Corollary 5.1.22.

Remark 5.2.10. Let us note that thin simplices turn up in [SVL13] when studying conditions for unimodality of (local) h^*-polynomials in the context of box unimodal triangulations mentioned before Corollary 5.1.20. Here, let us recall the following observation: if P admits a regular triangulation Δ such that every non-empty face of Δ is thin, then its local h^*-polynomial equals the local h-polynomial of Δ and its h^*-polynomial equals the h-polynomial of Δ, see Proposition 5.1.19 and Theorem 5.1.17. Now, in [SVL13] it is asked whether every IDP lattice polytope has a regular triangulation into lattice simplices that have vanishing or monomial ℓ^*-polynomial. The motivation was that the existence of a box unimodal triangulation of an IDP reflexive polytope implies unimodality of its h^*-polynomial. While the previous question is still open, a proof of the latter result using completely different methods was recently announced in [APPS21].

5.2.2 Two Classes of Examples of Thin Polytopes

Let us describe two ways to get thin polytopes in higher dimensions.
The first observation is that lattice polytopes of small degree (in other words, "very hollow" lattice polytopes) are always thin.

Definition 5.2.11. We say, P is *trivially thin* if $\dim(P) \geq 2\deg(P)$.

Proposition 5.2.12. Trivially thin polytopes are thin.

Proof. A lattice polytope P is trivially thin if and only if $\deg(P) < \operatorname{codeg}(P)$. Now, the statement follows from Theorem 5.1.16(3). □

Typical examples of trivially thin polytopes are Cayley polytopes with many factors. We will talk about Cayley polytopes with two factors in much more detail later (see Definition 5.4.7 and Remark 5.4.8), however, let us already now give the definition of a Cayley polytope to make the previous statement precise. For this, we denote by a *lattice projection* $\mathbb{R}^d \to \mathbb{R}^m$ an affine-linear map mapping $\mathbb{Z}^d$ surjectively onto $\mathbb{Z}^m$. If there is a lattice projection mapping a d-dimensional lattice polytope P onto a unimodular simplex Δ_k with $k \geq 1$, then P is called a *Cayley polytope* with $k+1$ factors (namely, the fibers of the vertices of Δ_k). One can easily deduce from [BN08, Proposition 1.12] that P is trivially thin if $k \geq d/2$. An alternative way to view this is also the following. Take r lattice polytopes $P_0, \ldots, P_{r-1}$ in $\mathbb{R}^m$. Then the *Cayley sum* of $P_0, \ldots, P_{r-1}$ is defined as the convex hull of $P_0 \times \{0\}$ and $P_i \times \{e_i\}$ for $i = 1, \ldots, r-1$ in $\mathbb{R}^{m+r-1}$. It is trivially thin if $r \geq m+1$. Note that a Cayley sum is a Cayley polytope, and every Cayley polytope is isomorphic to a Cayley sum.

A second way to get high-dimensional thin polytopes is to use free joins.

Definition 5.2.13. Let $P \subset \mathbb{R}^n$ and $Q \subset \mathbb{R}^m$ be lattice polytopes. We call

$$P \circ_{\mathbb{Z}} Q := \operatorname{conv}(P \times \{0\} \times \{0\}, \{0\} \times P \times \{1\}) \subset \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R},$$

the *free join* of P and Q.

For instance, the free join of $[0,1]$ with itself is a unimodular 3-simplex. Note that isomorphic factors lead to isomorphic free joins. From the Ehrhart-theoretic viewpoint the free join construction is important because of the following multiplicativity property, see [HT09, Lemma 1.3] and [NS13, Remark 4.6(5)].

Proposition 5.2.14. Let $P \subset \mathbb{R}^n$ and $Q \subset \mathbb{R}^m$ be lattice polytopes. Then

$$h^*_{P \circ_{\mathbb{Z}} Q}(t) = h^*_P(t)\, h^*_Q(t), \quad \text{and} \quad \ell^*_{P \circ_{\mathbb{Z}} Q}(t) = \ell^*_P(t)\, \ell^*_Q(t).$$

Corollary 5.2.15. The free join of two lattice polytopes is thin if and only if at least one of the two factors is thin.

As a lattice pyramid is the free join of a point (which is thin) and a lattice polytope, this generalizes Proposition 5.2.5.

5.2.3 Are There Other Examples of Thin Polytopes?

It is not trivial to give examples of thin polytopes (such as Example 5.2.17 below) that do not fall in above described two classes. In order to formulate a natural question in this respect, let us recall two notions. First, a lattice polytope P is called *spanning* if every lattice point in its affine hull is an integer affine combination of the lattice points in P. Note that every lattice polytope becomes spanning after a possible coarsening of the ambient lattice (we refer to [HKN18] for more background and results on spanning lattice polytopes). Second, let us call a lattice polytope P a *join* if there are two non-empty faces F and G of P such that the free join of F and G is affinely-isomorphic to P. Let us remark that if P is spanning and a join of F and G where every lattice point in P is contained in F or G, then P is the free join of F and G.

Question 5.2.16.

(1) Is every thin polytope trivially thin or a join?

(2) Is every spanning thin polytope trivially thin or a free join?

Both questions are closely related but not directly. The reason is that the degree of the polytope can drop under coarsenings of the lattice, so a non-spanning thin but not trivially thin polytope could be trivially thin with respect to its spanning lattice. We also note that trivially thin polytopes are often not joins. For example the unit square $[0,1]^2$ has degree 1 and is hence trivially thin, while triangles are the only polygons which are joins.
As the following example shows, the spanning hypothesis in the second part of Question 5.2.16 is indeed important. It is one of the apparently rare thin polytopes that are not trivially thin and not a free join.

Example 5.2.17. Consider the 4-simplex $P = \operatorname{conv}(0, e_1, e_2, (1,2,4,0), (2,1,0,4)) \subseteq \mathbb{R}^4$. The sublattice N of $\mathbb{Z}^4$ spanned by all lattice points of P has index 2 and the quotient $\mathbb{Z}^4/N \cong \mathbb{Z}/2\mathbb{Z}$ is generated by $\overline{e_3} = \overline{e_4}$. A computation in `SageMath` with backend `Normaliz` shows that P is thin and $h^*_P(t) = t^3 + 11t^2 + 3t + 1$, in particular $\deg(P) = 3$, so P is not trivially thin. A computation in `Polymake` shows that the lattice width of P is 2, so that P is not a Cayley polytope, in particular not a free join. It can be checked that with respect to N, P is the lattice pyramid over a reflexive 3-simplex of lattice volume 8.

Question 5.2.16 should be understood as a guiding question for finding interesting high-dimensional thin polytopes. Let us discuss this problem in more detail below. As being hollow is equivalent to $\deg(P) < \dim(P)$, it is evident that every hollow lattice polytope in dimension ≤ 2 is trivially thin. Hence, by Proposition 5.2.3 every thin polytope in dimension ≤ 2 is trivially thin. It will be proven in our first main result Theorem 5.3.3 that in dimension 3 all non-trivially thin polytopes are lattice pyramids. In particular, Question 5.2.16 has an affirmative answer in dimensions ≤ 3. Note that $\operatorname{conv}(e_1, e_2, -e_1 - e_2) \circ_{\mathbb{Z}} 2\Delta_2$ is an example of a thin simplex in dimension 5 that is not trivially thin (it has degree 3), but is not a lattice pyramid, while being a free join with a (trivially) thin factor.

In higher dimensions our second main result shows that non-trivially thin Gorenstein polytopes are so-called Gorenstein joins (see Definition 5.4.11) with a trivially thin factor, so that Question 5.2.16 has an affirmative answer also in the Gorenstein case (see Corollary 5.5.4).

Computationally, we have verified that Question 5.2.16 has an affirmative answer for all 4-dimensional lattice polytopes of lattice volume ≤ 21, for all 5-dimensional lattice simplices of lattice volume ≤ 20 and for all 6-dimensional lattice simplices of lattice volume ≤ 16. We provide some of the relevant data at [Kre].

5.2.4 Interesting Thin Empty Simplices?

A lattice simplex is called *empty* if its vertices are its only lattice points. Among the hollow polytopes this is the class of lattice simplices that has been studied most intensively, see e.g. [IVnS21] and the references therein. However, it turns out that there are no interesting thin empty simplices in dimension at most 4. Let us give the easy reasoning. For this, we recall that the *quotient group* of a d-dimensional lattice simplex $P \subset \mathbb{R}^d$ is defined as the quotient of $\mathbb{Z}^{d+1}$ by the subgroup generated by the vertices of $P \times \{1\}$.

Proposition 5.2.18. Let P be a lattice simplex with cyclic quotient group. Then P is thin if and only if P is a lattice pyramid.

Proof. Let $P \subset \mathbb{R}^d$ be d-dimensional. We denote by Π the half-open parallelepiped from Example 5.1.15. Clearly, every element in the quotient group of P has a unique representative in $\Pi \cap \mathbb{Z}^{d+1}$. Let $g \in \Pi \cap \mathbb{Z}^{d+1}$ be the representative of a generator of the quotient group of P. We assume that P is thin. Hence, there is a proper, non-empty subset V' of the vertex set of $S \times \{1\}$ such that g is a linear combination of vertices of V'. In particular, this also holds for the representatives of all the elements in the quotient group of P. Now, it follows from [Nil08, Lemma 12] that P is a lattice pyramid. $\square$

It is well-known that all empty lattice simplices in dimension at most 4 have cyclic quotient group [BBBK11]. As also in higher dimensions most empty simplices constructed (but not all of them) have this property (see e.g. [DKNS21]), it seems to be a challenge to find examples of empty simplices that are thin but not simply lattice pyramids.

5.2.5 Are Thin Polytopes 'Flat'?

We observed above that all thin polytopes in dimension at most two have lattice width 1 except for $2\Delta_2$. We leave it as an exercise to the reader to show that $2\Delta_d$ for even d is the only thin simplex among all lattice simplices of the form $\mathrm{conv}(0, k_1e_1, \ldots, k_de_d) \subset \mathbb{R}^d$ with $k_1, \ldots, k_d \in \mathbb{Z}_{\geq 1}$ that are not lattice pyramids (i.e., $k_i > 1$ for all i). It will follow from our main results that thin polytopes in dimension three (Corollary 5.3.4) as well as thin Gorenstein polytopes in arbitrary dimension (Corollary 5.5.7) have lattice width 1. In dimension four Example 5.2.17 has lattice width 2. As thin polytopes (of dimension > 0) are hollow, in fixed dimension their lattice width is bounded. Now, our lack of 'non-flat' examples motivates the following question.

Question 5.2.19. Can one find (spanning) thin polytopes of arbitrarily large lattice width?

We expect that such examples with increasing lattice width should exist with increasing dimension. Note that if one assumes that Question 5.2.16(2) has an affirmative answer, then for Question 5.2.19 it would be important to find the maximum width of trivially thin spanning polytopes P. However, it is a folklore open question, often called 'the' Cayley conjecture (see [DN10, HNP09, Hig19]), that any lattice polytope with $\dim(P) > 2\deg(P)$ has lattice width 1. Thus, assuming also that the Cayley conjecture holds essentially reduces the previous question to the study of spanning lattice polytopes with $\dim(P) = 2\deg(P)$.

5.3 Classification of Thin Polytopes in Dim. 3

As observed above, 3-dimensional lattice polytopes P that are lattice pyramids over polygons or have degree at most one are thin. Our first main result, Theorem 5.3.3, shows that in dimension three indeed all the thin polytopes are of this type.
Lattice polytopes of degree at most one are completely known in any dimension. For this, let us recall the following definition.

Definition 5.3.1. A *Lawrence prism* is a d-dimensional lattice polytope in $\mathbb{R}^d$ isomorphic to the convex hull $\mathrm{conv}(0, e_1, \ldots, e_{d-1}, k_0e_d, e_1 + k_1e_d, \ldots, e_{d-1} + k_{d-1}e_d)$ for some $k_0, k_1, \ldots, k_{d-1} \in \mathbb{Z}_{\geq 1}$.

The following result was proven in [BN07].

Theorem 5.3.2. Any lattice polytope of degree 1 is either a lattice pyramid, a Lawrence prism or isomorphic to $2\Delta_2$.

Here is the main result of this section.

Theorem 5.3.3. Let P be a 3-dimensional lattice polytope. Then P is thin if and only if P is a lattice pyramid over a lattice polygon (i.e., over a lattice polytope of dimension 2) or $\deg(P) \leq 1$. Equivalently, P is thin if and only if

- P is a lattice pyramid over a lattice polygon, or
- P is a Lawrence prism.

Corollary 5.3.4. Every 3-dimensional thin polytope has lattice width 1.

The proof of Theorem 5.3.3 relies on two instances that seem to be exceptional to small dimensions. First, in dimension three all the coefficients of the local h^*-polynomial can be explicitly determined.

Proposition 5.3.5. Let $P \subseteq \mathbb{R}^3$ be a 3-dimensional lattice polytope. Then

$$\ell^*_P(t) = |\operatorname{int}_{\mathbb{Z}}(P)|(t+t^3) + \left(|\operatorname{int}_{\mathbb{Z}}(2P)| - 4|\operatorname{int}_{\mathbb{Z}}(P)| - \sum_{F \leq P \text{ facet}} |\operatorname{int}_{\mathbb{Z}}(F)| \right) t^2.$$

Proof. Recall from Theorem 5.1.11 that $\ell^*_1 = \ell^*_3 = h^*_3 = |\operatorname{int}_{\mathbb{Z}}(P)|$. Hence, we need only determine ℓ^*_2. From Stanley reciprocity, we deduce $h^*_2 = |\operatorname{int}_{\mathbb{Z}}(2P)| - 4|\operatorname{int}_{\mathbb{Z}}(P)|$. Now, in the notation of the h^*-diamond introduced in [KS16] (see Definition 5.1.23) we have $h^*_2 = h^*_{1,0} + h^*_{1,1}$, where $h^*_{1,1} = \ell^*_2$ and $h^*_{1,0} = \sum_{F \leq P \text{ facet}} |\operatorname{int}_{\mathbb{Z}}(F)|$ by [KS16, Example 8.9]. This implies the statement. $\square$

Let us note that we get from the lower bound theorem of Katz–Stapledon, Theorem 5.1.24, $\ell^*_1 \leq \ell^*_2$. This leads to the following non-obvious corollary. It would be very interesting to find a purely combinatorial proof.

Corollary 5.3.6. Let $P \subseteq \mathbb{R}^3$ be a 3-dimensional lattice polytope. Then

$$|\operatorname{int}_{\mathbb{Z}}(2P)| \geq 5\,|\operatorname{int}_{\mathbb{Z}}(P)| + \sum_{F \leq P \text{ facet}} |\operatorname{int}_{\mathbb{Z}}(F)|.$$

For our purposes, let us note the following numerical characterization of thinness in dimension three.

Corollary 5.3.7. Let $P \subseteq \mathbb{R}^3$ be a 3-dimensional lattice polytope. Then P is thin if and only if P is hollow and

$$|\operatorname{int}_{\mathbb{Z}}(2P)| = \sum_{F \leq P \text{ facet}} |\operatorname{int}_{\mathbb{Z}}(F)|.$$

The second result that is not yet available in higher dimensions is a complete classification of hollow 3-dimensional lattice polytopes.

Theorem 5.3.8 ([AWW11])**.** Let $P \subseteq \mathbb{R}^3$ be a 3-dimensional hollow lattice polytope. Then one of the following holds:

(1) P is contained in one of the 12 maximal hollow lattice polytopes classified in [AWW11].

(2) There is a lattice projection $\mathbb{R}^3 \to \mathbb{R}^1$ mapping P onto Δ_1.

(3) There is a lattice projection $\mathbb{R}^3 \to \mathbb{R}^2$ mapping P onto $2\Delta_2$.

Before giving the proof of Theorem 5.3.3 let us also recall the following well-known formula for the mixed volume (e.g. [Nil20, Corollary 3.2]):

Lemma 5.3.9. Let $P_1, P_2 \subseteq \mathbb{R}^2$ be lattice polytopes. Then

$$\begin{aligned}\mathrm{MV}(P_1, P_2) = 1 &+ (-1)^{\dim(P_1+P_2)} |\operatorname{int}_{\mathbb{Z}}(P_1 + P_2)| \\ &+ (-1)^{\dim(P_1)-1} |\operatorname{int}_{\mathbb{Z}}(P_1)| + (-1)^{\dim(P_2)-1} |\operatorname{int}_{\mathbb{Z}}(P_2)|.\end{aligned}$$

Proof of Theorem 5.3.3. By Corollary 5.3.7, P is hollow. We treat the three cases of Theorem 5.3.8 separately. A direct computation in `Magma` deals with case 1, see [Kre].

For case 2, denote by $P_1, P_2 \subseteq \mathbb{R}^2$ the preimages in P of the vertices of Δ_1. Note that P_1 and P_2 are faces of P such that every lattice point of P is either contained in P_1 or P_2. (We remark that P is a Cayley polytope of P_1 and P_2 in the notation of Definition 5.4.7.) We denote by $\operatorname{int}^2_{\mathbb{Z}}(Q)$ the set of lattice points in the *absolute* interior of a lattice polytope $Q \subseteq \mathbb{R}^2$. Then

$$\begin{aligned}\sum_{F \leq P \text{ facets}} |\operatorname{int}_{\mathbb{Z}}(F)| &= |\operatorname{int}^2_{\mathbb{Z}}(P_1)| + |\operatorname{int}^2_{\mathbb{Z}}(P_2)|, \\ |\operatorname{int}_{\mathbb{Z}}(2P)| &= |\operatorname{int}_{\mathbb{Z}}(P_1 + P_2)|,\end{aligned}$$

where the second equation follows from the so-called Cayley trick. Therefore, Corollary 5.3.7 translates into

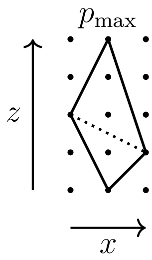
$$|\operatorname{int}^2_{\mathbb{Z}}(P_1)| + |\operatorname{int}^2_{\mathbb{Z}}(P_2)| = |\operatorname{int}_{\mathbb{Z}}(P_1 + P_2)|.$$

In case $\dim(P_1) = \dim(P_2) = 2$, plugging this into Lemma 5.3.9 yields $\mathrm{MV}(P_1, P_2) = 1$, thus $(P_1, P_2) \cong (\Delta_2, \Delta_2)$ by [CCD+13, Proposition 2.7], hence $\deg(P) = 1$.

If $\dim(P_1) = 2$ and $\dim(P_2) = 1$, then Lemma 5.3.9 yields $\mathrm{MV}(P_1, P_2) = 1 + |\operatorname{int}_{\mathbb{Z}}(P_2)|$. On the other hand, $\mathrm{MV}(P_1, P_2) = V(\pi_{P_2}(P_1))(|\operatorname{int}_{\mathbb{Z}}(P_2)| + 1)$ by [Sch13, Theorem 5.3.1], where π_{P_2} is a lattice projection along the line segment P_2 and $V(\pi_{P_2}(P_1))$ denotes the lattice volume. Hence, $V(\pi_{P_2}(P_1)) = 1$ and therefore $\pi_{P_2}(P_1) \cong \Delta_1$. The lattice projection of P along P_2 is then a lattice projection onto Δ_2. Thus, $\operatorname{codeg}(P) \geq 3$ and hence $\deg(P) \leq 1$, so either $\deg(P) = 1$ or $P \cong \Delta_3$ is a lattice pyramid.

The case $\dim(P_1) = \dim(P_2) = 1$ is similar. Lemma 5.3.9 yields $\mathrm{MV}(P_1, P_2) = 1 + |\operatorname{int}_{\mathbb{Z}}(P_1)| + |\operatorname{int}_{\mathbb{Z}}(P_2)|$. On the other hand, we again have $\mathrm{MV}(P_1, P_2) = V(\pi_{P_2}(P_1))(|\operatorname{int}_{\mathbb{Z}}(P_2)|+1)$. We may assume $|\operatorname{int}_{\mathbb{Z}}(P_1)| \leq |\operatorname{int}_{\mathbb{Z}}(P_2)|$. If $V(\pi_{P_2}(P_1)) \geq 2$ or $V(\pi_{P_2}(P_1)) = 0$, we obtain a contradiction, so $V(\pi_{P_2}(P_1)) = 1$. The same argument as above shows $\deg(P) \leq 1$.
Finally, if one of the P_i is zero-dimensional, then P is a lattice pyramid.

It is left to study case 3, and we may assume P to be of lattice width at least 2 because width 1 is equivalent to P being a Cayley polytope which is precisely case 2. We distinguish several cases and always start by showing how, in each case, we can associate to each lattice point in the interior of a facet of P, in an injective way, a lattice point in the interior of $2P$. We then prove that there always exists an additional lattice point in the interior of $2P$, therefore showing that case 3 does not yield any new thin polytopes by Corollary 5.3.7.
We may assume that P projects onto $2\Delta_2$ along the z-axis. As lattice projections map interior lattice points to interior lattice points, all interior lattice points of a facet of P are of the form $x_1^a = (1, 0, a)$, $x_2^a = (0, 1, a)$, or $x_3^a = (1, 1, a)$ for suitable $a \in \mathbb{Z}$. By fixing vertices $v_1 = (0, 2, \alpha)$, $v_2 = (2, 0, \beta)$, $v_3 = (0, 0, \gamma)$ of P we hence obtain points $\frac{1}{2}(x_i^a + v_i) \in \operatorname{int}(P)$, and therefore $(x_i^a + v_i) \in \operatorname{int}_{\mathbb{Z}}(2P)$ for all $a \in \mathbb{Z}$ such that x_i^a is an interior point of a facet of P. Then $(x_i^a + v_i) \neq (x_j^b + v_j)$ if $i \neq j$ or $a \neq b$.
Now we show the existence of an additional interior lattice point. Indeed, P can have at most three facets containing interior lattice points, namely at most those facets, if there are such, that project to one of the three edges of $2\Delta_2$.
We proceed by distinguishing these different cases. If there is no such facet at all, then Corollary 5.3.7 implies that $2P$ is hollow, so $\deg(P) \leq 1$, contradicting the fact that there is no lattice polytope of degree ≤ 1 with width > 1 by Theorem 5.3.2.
Next, assume that P has exactly two facets containing interior lattice points, and say these are the facets opposite to v_1 and v_2. We pick two such points $(0, 1, q)$ and $(1, 0, r)$. Then we obtain as many interior lattice points in $2P$ of the form $x_1^a + v_1$ or $x_2^a + v_2$ as there are points in the interiors of facets of P, and $(1, 1, q + r)$ is an additional interior lattice point of $2P$.
Next, assume P has three facets containing interior lattice points. If there exists $i \in \{1, 2, 3\}$ such that the fiber of $2P$ containing v_i contains more than one lattice point, then we can similarly construct an additional interior lattice point of $2P$ by considering the two points of minimal and maximal height in this fiber. Therefore, we may assume v_1, v_2, v_3 to be the unique lattice points of P over $(0, 2)$, $(2, 0)$ and $(0, 0)$, respectively. This implies that all three facets F_1, F_2, F_3 of P lying over the three edges of $2\Delta_2$ have a special form. E.g., the facet projecting to the edge $[(0, 0), (2, 0)]$ is a quadrangle or triangle which, after a unimodular equivalence, looks similar to the following:

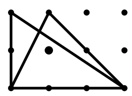


Now, we fix one interior lattice point $u_i := x_i^{a_i}$ in each of the facets for some suitable $a_i \in \mathbb{Z}$. The three maps $x_i^a \mapsto x_i^a + u_j$ for $(i,j) \in \{(1,2),(2,3),(3,1)\}$ map each interior lattice point of the three facets F_i injectively to an interior lattice point of $2P$. Moreover, the images of these three maps are disjoint by construction. Now, we may assume that in the facet projecting to the edge $[(0,0),(2,0)]$, the lattice point $p_{\max}$ lying over $(1,0)$ with maximal third coordinate is a vertex as in the picture. But then we obtain the additional interior lattice point $p_{\max} + u_2$ of $2P$ not covered by the images of the three maps above.

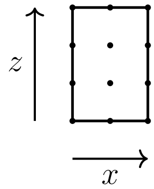
The only remaining case is the one where only one facet F of P contains interior lattice points. We may assume F is the facet which projects onto the edge $[(0,0),(2,0)]$. Then all interior lattice points of F project to $(1,0)$ and are of the form x_1^a for $a \in \mathbb{Z}$ ranging in a suitable interval. From this we obtain $|\operatorname{int}_{\mathbb{Z}}(F)|$ interior lattice points $v_1 + x_1^a$ of $2P$. Observe that all of them have second coordinate 2. Therefore, it is enough to show that $2P$ contains an interior lattice point with second coordinate 1. Again we proceed by distinguishing cases. If F contains at least three lattice points projecting to $(1,0)$, then it contains the convex hull of the points $(1,0,a)$, $(1,0,a+1)$, $(1,0,a+2)$, $(2,0,b)$ for some $a,b \in \mathbb{Z}$. Recall that $v_1 = (0,2,\alpha)$. Then the point

$$2 \cdot \left(\frac{1}{4}v_1 + \frac{1}{4}(1,0,a) + \frac{1}{4}(2,0,b) + \frac{1}{4}(1,0,a+n)\right) = (2,1,a+\frac{\alpha+b+n}{2})$$

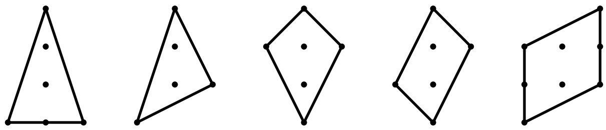
is an interior lattice point of $2P$ with second coordinate 1 for precisely one choice of $n \in \{1,2\}$. Hence, by Corollary 5.3.7, P is not thin if F contains at least three lattice points over $(1,0)$, which in particular includes the case $|\operatorname{int}_{\mathbb{Z}}(F)| \geq 3$.
Before we proceed further let us observe that the fibers of P over the points $(0,0)$ and $(2,0)$ both consist of at most three lattice points because otherwise another facet than F would contain interior lattice points. This is because, up to unimodular equivalence, there are exactly two lattice pyramids of height 2 over a lattice segment of length ≥ 3, and both contain an interior lattice point:



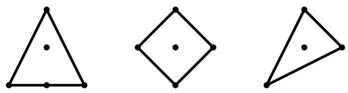
Now, we can deal with the remaining case $|\operatorname{int}_{\mathbb{Z}}(F)| \in \{1,2\}$. Let $F' \subseteq F$ run through the inclusion-minimal subpolygons of F which contain the same interior lattice points as F. We will show that there are only two possibilities for F'.
Let us first consider the case $|\operatorname{int}_{\mathbb{Z}}(F)| = 2$. By the previous argument, after a suitable shear fixing the $x = 1$ line (within the x-z-plane, i.e. $y = 0$), F' fits inside the following box:



As F' is inclusion-minimal, it is isomorphic to one of the following polygons:



But only the last one does not contain three lattice points over $(1, 0)$.
Let us also consider the case of $|\operatorname{int}_{\mathbb{Z}}(F)| = 1$. Here, we can fit F' inside the standard square $[0, 2]^2$ (inside the x-z-plane) after a suitable shear fixing the $x = 1$ line, so F' can be taken to be one of the following polygons:



Again, only the last one does not contain three lattice points over $(1, 0)$.
Lastly, for each of these two remaining polygons F' we may choose a lattice subpolytope P' of P that is a pyramid of height 2 over F'. We observe that for given F' there are at most four non-isomorphic possibilities for P' to consider as the first two coordinates of an apex in $\mathbb{R}^2 \times \{2\}$ over a base polytope in $\mathbb{R}^2 \times \{0\}$ may be chosen by a unimodular shearing to be in $\{(0, 0), (1, 0), (0, 1), (1, 1)\}$. Now, an explicit computation in `SageMath` shows that for all these at most eight cases we have $|\operatorname{int}_{\mathbb{Z}}(2P')| > |\operatorname{int}_{\mathbb{Z}}(F')| = |\operatorname{int}_{\mathbb{Z}}(F)|$, concluding the proof. $\square$

We can now answer the original question in [GKZ94] in dimension 3.

Corollary 5.3.10. A three-dimensional lattice simplex is thin if and only if it is a lattice pyramid.

This follows directly from Theorem 5.3.3. The reader is cautioned not to jump to the conclusion that the same result may be true in higher dimensions. In dimension 4, $[-1, 1] \circ_{\mathbb{Z}} 2\Delta_2$ is an example of a (trivially) thin simplex that is not a lattice pyramid.

5.4 Thin Gorenstein Polytopes and Gorenstein Joins

5.4.1 Gorenstein Polytopes and Their Duals

Definition 5.4.1. A lattice polytope P is called *Gorenstein* if h^*_P is palindromic.

Let us recall that *reflexive polytopes* are precisely the Gorenstein polytopes of codegree one. For more background on reflexive and Gorenstein polytopes, its relevance in toric geometry and mirror symmetry, as well as alternative characterizations we refer to [BB96b, BN08, NS13]. Here, let us summarize definitions and properties of the dual Gorenstein polytope. We remark that the codegree of a Gorenstein polytope is often called its index.

Definition 5.4.2. Let $P \subset \mathbb{R}^d$ be a d-dimensional Gorenstein polytope. In this case, the dilate $\operatorname{codeg}(P) \cdot P$ is a reflexive polytope (up to lattice translation), and we denote its unique interior lattice point by w. Then

$$P^\times := \{y \in (\mathbb{R}^{d+1})^* \ : \ \langle y, w\rangle = 1 \text{ and } \langle y, x\rangle \geq 0 \, \forall \, x \in P \times \{1\}\}$$

is called the *dual Gorenstein polytope* of P.

Proposition 5.4.3. Let $P \subset \mathbb{R}^d$ be a d-dimensional Gorenstein polytope. Then $P^\times$ is a Gorenstein polytope of the same dimension and degree as P, and it is combinatorially dual to P.

Note that, if a Gorenstein polytope is lower-dimensional, we consider, as usual, its ambient lattice in order to get its dual Gorenstein polytope.

Definition 5.4.4. If F is a face of P, we denote by F^* the *dual face*, i.e., the corresponding face of $P^\times$.

Attention: it is important to distinguish the dual face F^* from $F^\times$, the latter being defined only if F is itself a Gorenstein polytope which is not true in general. Even if this is the case, the two polytopes might have completely different dimensions (since the one definition is relative to P while the other one is intrinsic).
Local h^*-polynomials of Gorenstein polytopes (often called $\tilde{S}$-polynomials) allow to give an elegant formula for computing stringy E-polynomials of Calabi-Yau complete intersections in toric Gorenstein Fano varieties (we refer to [BM03, BN08]). In this context, several questions about stringy E-polynomials are still open, see [BN08, NS13]. Here, we make some progress in this direction by addressing the question when the local h^*-polynomial of a Gorenstein polytope vanishes. As one consequence of our main result, Theorem 5.5.3, we will see that not only the degree of the h^*-polynomials of Gorenstein polytopes and their duals are the same but also of their ℓ^*-polynomials (Corollary 5.5.13).

5.4.2 Joins, Cayley Polytopes, and Cayley Joins

In the sequel let us discuss some important notions of decomposing lattice polytopes that turn up naturally when studying Gorenstein polytopes (we refer to [BN08, NS13]). Let us first introduce a formal notation for a lattice polytope being a join (as already defined in Subsection 5.2.3).

Definition 5.4.5. Let $P \subseteq \mathbb{R}^d$ be a polytope and F, G non-empty subsets of P. Then P is the *join of F and G*, written $P = F \circ G$, if $P = \operatorname{conv}(F, G)$ and $\dim(P) = \dim(F) + \dim(G) + 1$.

Equivalently, P is affinely-isomorphic to the free join $F \circ_{\mathbb{Z}} G$. In particular, F and G are automatically faces of P.

Remark 5.4.6. Note that the join property is associative. Namely, given faces F, G, H of P, then $P = F \circ (G \circ H)$, respectively, $P = (F \circ G) \circ H$, are both equivalent to $P = \operatorname{conv}(F, G, H)$ and $\dim(P) = \dim(F) + \dim(G) + \dim(H) + 2$.

Let us also give the formal notation for a lattice polytope being a Cayley polytope. We recall that the notion of Cayley polytopes and Cayley sums was already shortly mentioned and defined in Subsection 5.2.2. Here, we will solely focus on the case of two factors. Note that if a Cayley polytope has more than two factors, it is still a Cayley polytope with two factors.

Definition 5.4.7. Let $P \subseteq \mathbb{R}^d$ be a lattice polytope and F, G non-empty subsets of P. Then P is the *Cayley polytope* of (factors) F and G, written $P = F * G$, if $P = \operatorname{conv}(F, G)$ and there exists an affine-linear map $\mathbb{R}^d \to \mathbb{R}$ mapping $\mathbb{Z}^d \to \mathbb{Z}$, such that $F \mapsto 0$ and $G \mapsto 1$. In other words, P is a Cayley polytope if and only if there is a lattice projection mapping P onto Δ_1.

If $P = F * G$, then F and G are necessarily faces of P. Cayley polytopes can also be characterized as lattice polytopes with lattice width one. Cayley sums are explicit descriptions of Cayley polytopes.

Remark 5.4.8. Given lattice polytopes F and G in $\mathbb{R}^d$, the convex hull of $F \times \{0\}$ and $G \times \{1\}$ is called the *Cayley sum* of F and G. Its dimension is one larger than the dimension of the Minkowski sum of F and G. If $P = F * G$, then P is isomorphic to the Cayley sum of F and G.

Cayley sums are important in the construction of high-dimensional Gorenstein polytopes, see e.g. [BN08, Theorem 2.6]. Note that the degree of a Cayley polytope is at most the dimension of the Minkowski sum of its factors, see Proposition [BN07, Proposition 1.12].

Definition 5.4.9. Let $P \subseteq \mathbb{R}^d$ be a full-dimensional lattice polytope and $F, G \subseteq P$ faces. Then P is the *Cayley join* of F and G, written $P = F \circ_{\mathrm{Cay}} G$, if $P = F \circ G$ and $P = F * G$.

Clearly, the notion of a Cayley join is more restrictive than that of a Cayley polytope (e.g., $[0,1]^2$ is a Cayley polytope of two edges but not a Cayley join). The reader should be aware that Cayley polytopes and Cayley joins are in general not associative in the sense of Remark 5.4.6, see Example 5.4.17 below.

Let us recall some properties of a Gorenstein polytope that is a join or Cayley join, see [NS13, Lemma 4.8, Proposition 4.9].

Proposition 5.4.10. Let P be a Gorenstein polytope.

- If $P = F \circ G$, then $P^\times = F^* \circ G^*$ with F and G^* (respectively, G and F^*) being combinatorially dual to each other.

- If $P = F \circ_{\mathrm{Cay}} G$, then F^* is a Gorenstein polytope with dual Gorenstein polytope $(F^*)^\times$, which can be identified with the lattice polytope G with respect to a refined lattice.

In the last statement, the lattice does not have to be refined if the Cayley join is even a free join.

5.4.3 Gorenstein Joins

The following notion is defined in [NS13].

Definition 5.4.11. Let F and G be faces of a Gorenstein polytope P. We say P is a *Gorenstein join* of F and G, denoted by $P = F \circ_{\mathrm{Gor}} G$, if $P = F \circ_{\mathrm{Cay}} G$ and $P^\times = F^* \circ_{\mathrm{Cay}} G^*$. We call F and G the *factors* of the Gorenstein join.

We remark that Gorenstein joins do not have to be free joins, see [NS13, Example 4.14]. The following result, a strengthening of Stanleys monotonicity theorem in the case of faces of Gorenstein polytopes, motivated the previous definition of a Gorenstein join and gives a direct enumerative criterion for its existence ([NS13, Theorems 3.6 and 4.12]).

Theorem 5.4.12. Let P be a Gorenstein polytope and F a non-empty proper face of P. Then $\mathrm{codeg}(P) \leq \mathrm{codeg}(F) + \mathrm{codeg}(F^*)$ (equivalently, $\deg(F) + \deg(F^*) \leq \deg(P)$), with equality if and only if P is a Gorenstein join with factor F. In this case, F is a Gorenstein polytope.

Gorenstein polytopes that are not Gorenstein joins have been previously also called *irreducible* in [NS13]. As we see from the following result it is not necessary to compute the dual Gorenstein polytope to check whether a Cayley join is a Gorenstein join.

Lemma 5.4.13. Let $P = F \circ_{\mathrm{Cay}} G$ be a Gorenstein polytope which is the Cayley join of two faces $F, G \leq P$. Then $P = F \circ_{\mathrm{Gor}} G$ if and only if $\mathrm{codeg}(P) = \mathrm{codeg}(F) + \mathrm{codeg}(G)$ (or equivalently, $\deg(P) = \deg(F) + \deg(G)$).

Proof. By Theorem 5.4.12, $P = F *_{\mathrm{Gor}} G$ if and only if $\mathrm{codeg}(F) + \mathrm{codeg}(F^*) = \mathrm{codeg}(P) =: r$, and in this case by [NS13, Theorem 4.12] $\mathrm{codeg}(G) = \mathrm{codeg}(P) - \mathrm{codeg}(F) = \mathrm{codeg}(F^*)$. Conversely, assume $\mathrm{codeg}(G) = \mathrm{codeg}(P) - \mathrm{codeg}(F)$. By Theorem 5.4.12, the inequality $\mathrm{codeg}(F) + \mathrm{codeg}(F^*) \geq r$ always holds, so that by Theorem 5.4.12 again we only need to prove $\mathrm{codeg}(F^*) \leq \mathrm{codeg}(G)$. As P is the Cayley join of F and G, Proposition 5.4.10 yields that $\mathrm{codeg}(F^*) = \mathrm{codeg}((F^*)^\times) \leq \mathrm{codeg}(G)$ as the codegree can only decrease under refinements of the lattice. □

Remark 5.4.14. As we will need it for the upcoming proofs, let us recall how to characterize Gorenstein polytope via cones. For more details, we refer to [BN08]. The *cone over* P is denoted by $C_P \subseteq \mathbb{R}^{d+1}$ spanned by $P \times \{1\} \subset \mathbb{R}^{d+1}$. Any polyhedral cone in $\mathbb{R}^{d+1}$ that is unimodularly equivalent to some C_P is called a *Gorenstein cone*. Now, P is a Gorenstein polytope if and only if the dual cone $C_P^\vee = \{y \in (\mathbb{R}^{d+1})^* : \langle y, x\rangle \geq 0 \;\forall\, x \in C_P\}$ is a Gorenstein cone. In this case, $C_P^\vee$ is unimodularly equivalent to the cone over $P^\times$.

The following proposition contains a positive result regarding associativity of Gorenstein joins. In general, however, we do not expect associativity to hold.

Proposition 5.4.15. Let $P \subseteq \mathbb{R}^d$ be a d-dimensional Gorenstein polytope with faces $F, G, H \leq P$ such that $P = (F *_{\mathrm{Gor}} G) *_{\mathrm{Gor}} H$. If F is a vertex or H is a vertex, then $P = F *_{\mathrm{Gor}} (G *_{\mathrm{Gor}} H)$. In particular, Gorenstein joins are associative for $\dim(P) \leq 3$.

Proof. By Lemma 5.4.13, the faces $\mathrm{conv}(F, G)$ and H of P are themselves Gorenstein polytopes with $r := \mathrm{codeg}(P) = \mathrm{codeg}(\mathrm{conv}(F, G)) + \mathrm{codeg}(H)$. Applying the same result to the Gorenstein polytope $\mathrm{conv}(F, G) = F *_{\mathrm{Gor}} G$, we obtain that F and G are Gorenstein polytopes with $\mathrm{codeg}(\mathrm{conv}(F, G)) = \mathrm{codeg}(F) + \mathrm{codeg}(G)$. Hence, $r = \mathrm{codeg}(F) + \mathrm{codeg}(G) + \mathrm{codeg}(H)$. By Lemma 5.4.13, it hence suffices to show $P = F *_{\mathrm{Cay}} (G *_{\mathrm{Cay}} H)$. That the join of G and H is a Cayley join is immediate from the assumption, so it is enough to show that the join of F and $\mathrm{conv}(G, H)$ is a Cayley join.

Let first F be a vertex. Hence, P is a pyramid with vertex F and base $\mathrm{conv}(G, H)$. Consider the cone $C_P \subset \mathbb{R}^{d+1}$. As P is Gorenstein of codegree r, there is a unique interior lattice point $p \in C_P$ on height $p_{d+1} = r$. Similarly, let f, g and h be the unique interior lattice points of $C_F, C_G, C_H \subseteq C_P$ on heights $\mathrm{codeg}(F), \mathrm{codeg}(G), \mathrm{codeg}(H)$, respectively. Then necessarily $f + g \in C_{\mathrm{conv}(F,G)} \subseteq C_P$ is the unique interior lattice point on height $\mathrm{codeg}(\mathrm{conv}(F, G))$ of the Gorenstein polytope $\mathrm{conv}(F, G)$. Hence, $p = f + g + h$. Let $u \in (\mathbb{Z}^{d+1})^*$ be the primitive inner facet normal of the hyperplane containing $\mathrm{conv}(G, H)$. As u is a vertex of $P^\times$, we have $\langle u, p\rangle = 1$ (cf. Definition 5.4.2). Therefore,

$$1 = \langle u, p\rangle = \langle u, f\rangle + \langle u, g + h\rangle = \langle u, f\rangle.$$

But this means that F and $\mathrm{conv}(G, H)$ have lattice distance equal to 1, i.e., the combinatorial join $F \circ \mathrm{conv}(G, H)$ is a Cayley join.

Let now H be a vertex, so P is a lattice pyramid with vertex H and base $\mathrm{conv}(F, G)$. As $F \circ_{\mathrm{Cay}} G$, we may assume that $\mathrm{lin}(F, G) = \mathbb{R}^{d-1} \times \{0\}$, $H = \{e_d\}$, and there exists some $u \in (\mathbb{Z}^{d-1} \times \{0\})^*$ such that $\langle u, F\rangle = 0$ and $\langle u, G\rangle = 1$. Now, $\langle u + e_d^*, F\rangle = 0$ and $\langle u + e_d^*, \mathrm{conv}(G, H)\rangle = 1$. In particular, the join of F and $\mathrm{conv}(G, H)$ is a Cayley join.

Finally, if $d = \dim(P) \leq 3$ and P is the join of F, G and H, then necessarily at least one of F and H is a vertex for dimension reasons, concluding the proof. $\square$

We observe that a Gorenstein polytope P is a lattice pyramid over a face F with apex a vertex v of P if and only if P is a Gorenstein join of F and v. Hence, the previous result has the following consequence.

Corollary 5.4.16. If P is a Gorenstein join of two faces with one face a lattice pyramid, then P is also a lattice pyramid.

This result is already contained in the master's thesis [Mic21]. Let us give an example that shows that it fails for Cayley joins.

Example 5.4.17. Consider $F := \operatorname{conv}(e_1, e_2) \times \{0\}$ and $G := \operatorname{conv}(0, -e_1 - e_2) \times \{1\}$ in $\mathbb{R}^3$. Then its convex hull P is a tetrahedron that is a Gorenstein polytope of lattice volume 2 (with $h^*_P(t) = 1 + t^2$ and $\ell^*_P(t) = t^2$). It is a Cayley join $P = F \circ_{\mathrm{Cay}} G$ but not a Gorenstein join as $\deg(P) = 2 \neq 0 = \deg(F) + \deg(G)$. Note that F and G are lattice pyramids, but P is not. In particular, this example shows that the Cayley join property is not associative, and moreover, a Cayley join does not have to be thin if a factor of the Cayley join is thin.

5.4.4 Local h^*-Polynomials of Joins

Recall that h^*- and ℓ^*-polynomials are multiplicative with respect to free joins (Proposition 5.2.14). For general joins, one still gets inequalities.

Lemma 5.4.18. Let P be a lattice polytope which is the join of two faces F and G. Then

$$\ell^*_F(t) \cdot \ell^*_G(t) \leq \ell^*_P(t) \quad \text{and} \quad h^*_F(t) \cdot h^*_G(t) \leq h^*_P(t).$$

If moreover P is a Gorenstein polytope which is the Gorenstein join of F and G, then also

$$\ell^*_{P^\times}(t) \leq \ell^*_{F^\times}(t) \cdot \ell^*_{G^\times}(t) \quad \text{and} \quad h^*_{P^\times}(t) \leq h^*_{F^\times}(t) \cdot h^*_{G^\times}(t).$$

Proof. We will use the notation of [NS13]. Let $\overline{M} = \mathbb{Z}^{d+1}$, and $M(F)$ denote the sublattice of $\overline{M}$ spanned by the lattice points in the linear hull of $F \times \{1\}$. Relative to the sublattice $M(F) \oplus_{\mathbb{Z}} M(G)$ of $\overline{M}$, P becomes the free join of F and G. Recall that by Corollary 5.1.21 both the h^*-polynomial and the ℓ^*-polynomial are (weakly) monotonically increasing under refinements of the lattice. It hence follows from Proposition 5.2.14 that $\ell^*_F(t) \cdot \ell^*_G(t) \leq \ell^*_P(t)$ and $h^*_F(t) \cdot h^*_G(t) \leq h^*_P(t)$.

For the second claim, assume that $P \subseteq \mathbb{R}^d$ is a full-dimensional Gorenstein polytope of codegree r with respect to the lattice $M = \mathbb{Z}^d \subseteq \mathbb{R}^d$. By assumption, P is the Gorenstein join of two faces F and G. By Theorem 5.4.12 and Lemma 5.4.13, F and G are Gorenstein polytopes and $\operatorname{codeg}(F) + \operatorname{codeg}(G) = r$. For $\overline{M} = \mathbb{Z}^{d+1}$ we define the dual lattice $\overline{N} := \operatorname{Hom}_{\mathbb{Z}}(\overline{M}, \mathbb{Z}) \subseteq (\mathbb{R}^{d+1})^*$. By definition, as P is a Gorenstein polytope, $C_P^\vee$ is a Gorenstein cone with respect to $\overline{N}$. Let $n = e^*_{d+1} \in \overline{N}$ be the unique interior lattice point of $C_P^\vee$ with $P \times \{1\} = C_P \cap \{x \in \mathbb{R}^{d+1} : \langle n, x\rangle = 1\}$. In the same way, we denote by $m \in \overline{M}$ the unique interior lattice point of C_P such that $P^\times = C_P^\vee \cap \{y \in (\mathbb{R}^{d+1})^* : \langle y, m\rangle = 1\}$. Recall that $\langle n, m\rangle = r$ and hence $m = (p, r) \in M \oplus \mathbb{Z} = \overline{M}$ with $p \in M$ the unique interior lattice point of the r-th dilate rP of P.
Now, let us consider the sublattice $M(F) \oplus M(G) \subseteq \overline{M}$. With respect to this coarser lattice, the polytope P is the *free* join of F and G, and this is clearly a Gorenstein polytope of codegree $\operatorname{codeg}(F) + \operatorname{codeg}(G) = r$. Hence, the r-th dilate rP contains a unique interior lattice point in the original as well as in the coarser lattice. These two points must therefore agree, so the unique interior lattice point $m = (p, r) \in (rP) \times \{r\} \subseteq C_P$ with respect to the original lattice actually lies in $M(F) \oplus M(G)$. Let now $\tilde{N} \subset (\mathbb{R}^{d+1})^*$ be the dual lattice of $M(F) \oplus M(G)$. Hence, $P^\times$ is with respect to the finer lattice $\tilde{N}$ the dual Gorenstein polytope of the Gorenstein polytope P considered with respect to the coarser lattice $M(F) \oplus M(G)$. By Proposition 5.4.10,

the Gorenstein dual of the free join of the Gorenstein polytopes F and G, is the free join of the Gorenstein duals $F^\times$ and $G^\times$. Again, monotonicity and multiplicativity proves the second claim: $\ell^*_{P^\times}(t) \le \ell^*_{F^\times}(t) \cdot \ell^*_{G^\times}(t)$ and $h^*_{P^\times}(t) \le h^*_{F^\times}(t) \cdot h^*_{G^\times}(t)$, where $P^\times$ is considered with respect to the original lattice $\overline{N}$ again. □

Remark 5.4.19. It follows in the situation of the second part of Lemma 5.4.18 from Proposition 5.4.10 and Remark 5.2.8 that $\ell^*_F(t) \le \ell^*_{(G^*)^\times}(t)$ and $\ell^*_G(t) \le \ell^*_{(F^*)^\times}(t)$, because $(G^*)^\times$ is just the polytope F with a possibly finer lattice, and analogously for $(F^*)^\times$ and G. The same holds for the h^*-polynomial.

5.5 Characterization of Thin Gorenstein Polytopes

5.5.1 The Main Result

The following notion will occur naturally in the proof of Theorem 5.5.3.

Definition 5.5.1. A lattice polytope P is called *g-thin* if $\deg(g_P) = \deg(P)$.

For instance, any unimodular simplex is g-thin. By Corollary 5.1.18, we always have $\deg(g_P) \le \deg(P)$. Since by Definition 5.1.2 we have $\deg(g_P) \le \dim(P)/2$, we deduce:

$$g\text{-thin} \quad \Longrightarrow \quad \text{trivially thin}$$

Example 5.5.2. Let P denote the lattice pyramid over $[-1, 1]$. Then P has dimension 2, degree 1 and $\deg(g_P) = 0$ as it is a simplex. This is an example of a trivially thin Gorenstein polytope that is not g-thin. For another example, consider $2\Delta_2$ which is a trivially thin (non-Gorenstein) simplex that is not g-thin. This example shows that a spanning thin polytope that is not a free join does not have to be g-thin (i.e., in Question 5.2.16 'trivially thin' cannot be strengthened by 'g-thin').

Here is our main result.

Theorem 5.5.3. Let P be a Gorenstein polytope. Then the following are equivalent:

(i) P is thin,

(ii) P is trivially thin or $P = F *_{\mathrm{Gor}} G$ with at least one factor trivially thin,

(iii) P is g-thin or $P = F *_{\mathrm{Gor}} G$ with $\deg(\ell^*_F) = \deg(F)$ and G g-thin.

Moreover, if P is not thin, then $\deg(\ell^*_P) = \deg(P)$.

Let us remark that if P is not thin, the last statement implies that $\ell^*_P(t)$ and $h^*_P(t)$ have the same leading coefficient 1, as $h^*_P(t)$ is palindromic with constant coefficient 1.

We have to leave it as an open question whether it is possible to strengthen in the previous result 'Gorenstein join' to 'free join'. Let us point out the following situation in which a Gorenstein join (or even just a spanning join of faces) is already a free join.

Corollary 5.5.4. Let P be a spanning Gorenstein polytope. Then P is thin if and only if it is trivially thin or a free join with a trivially thin factor (necessarily also a spanning Gorenstein polytope).

The proof of Theorem 5.5.3 relies critically on the decomposition of the h^*-polynomial into ℓ^*-polynomials and g-polynomials (Corollary 5.1.18), valid also for general lattice polytopes.

Lemma 5.5.5. Let P be a lattice polytope with $\deg(\ell^*_P) < \deg(P)$. Then P is g-thin or there exists a non-empty, proper face F of P with $\deg(P) = \deg(\ell^*_F) + \deg(g_{[F,P)})$.

Here, we recall $\deg(\ell^*_\emptyset) = \deg(1) = 0$ and $\deg(g^*_\emptyset) = \deg(1) = 0$.

Proof. By the nonnegativity of ℓ^*- and g-polynomials, Corollary 5.1.18 implies that there exists a face F of P with $\deg(\ell^*_F) + \deg(g_{[F,P)}) = \deg(h^*_P)$. By our assumption, $F \neq P$. If $F = \emptyset$, then $\deg(h^*_P) = \deg(g_{[\emptyset,P)})$, so P is g-thin. □

Lemma 5.5.6. Let $P = F *_{\mathrm{Gor}} G$. Then the Gorenstein polytopes F, G^*, $F^\times$, $(G^*)^\times$ have the same degree, dimension, and degree of their g-polynomials. In particular, if any of these Gorenstein polytopes are trivially thin (respectively, g-thin), then all of them are.

Proof. The Gorenstein property follows from Proposition 5.4.10. It is well-known, cf. [BN08], that duality of Gorenstein polytopes keeps dimension and degree invariant. It follows from Theorem 5.1.9 that this is also true for the degree of the g-polynomial. By Proposition 5.4.10, F and G^* are combinatorially dual to each other, hence, have the same dimension and by Theorem 5.1.9 the same degree of the g-polynomial. Finally, by Lemma 5.4.13 and Theorem 5.4.12,

$$\deg(P) - \deg(F) = \deg(G) = \deg(P) - \deg(G^*),$$

hence, $\deg(F) = \deg(G^*)$. □

Proof of Theorem 5.5.3. The implication (iii) ⇒ (ii) is immediate.
(ii) ⇒ (i): Let $P = F *_{\mathrm{Gor}} G$ with F trivially thin. By Lemma 5.5.6, it follows that $(G^*)^\times$ is trivially thin as well. Now, applying Lemma 5.4.18 to the factorization $P^\times = F^* *_{\mathrm{Gor}} G^*$ yields $\ell^*_P(t) \leq \ell^*_{(F^*)^\times}(t) \cdot \ell^*_{(G^*)^\times}(t) = 0$, so P is thin.
(i) ⇒ (iii): Let P be a Gorenstein polytope. We assume only that $\deg(\ell^*_P) < \deg(P)$ and will deduce (iii), so that P is in particular thin by the implications we already proved (and thus, if P is not thin, then $\deg(\ell^*_P) = \deg(P)$). Let us assume that P is not g-thin. Now, by Lemma 5.5.5 there exists a non-empty, proper face F of P with $\deg(P) = \deg(\ell^*_F) + \deg(g_{[F,P)})$. Theorem 5.1.9 shows that $\deg(g_{[F,P)}) = \deg(g_{(F,P]^*}) = \deg(g_{F^*})$. Thus, Corollary 5.1.18 and Theorem 5.4.12 imply that

$$\deg(F^*) \geq \deg(g_{F^*}) = \deg(P) - \deg(\ell^*_F) \geq \deg(P) - \deg(F) \geq \deg(F^*).$$

Therefore, F^* is g-thin, $\deg(\ell^*_F) = \deg(F)$, and $\deg(F) + \deg(F^*) = \deg(P)$, which implies (iii) by Theorem 5.4.12 and Proposition 5.4.10 (with the roles of F and G exchanged). □

Corollary 5.5.7. Every thin Gorenstein polytope (of dimension > 0) has lattice width 1.

Proof. As Gorenstein joins are Cayley joins, by Theorem 5.5.3 it remains to show that a trivially thin Gorenstein polytope of dimension > 0 is a Cayley polytope. This is precisely the statement of Theorem 3.1 in [HNP09]. $\square$

Example 5.5.8. Let us illustrate Theorem 5.5.3 by showing that all thin Gorenstein polytopes P of dimension $d = 3$ are lattice pyramids over Gorenstein polygons (without using Theorem 5.3.3 directly). Let us assume otherwise. If P is trivially thin, then $\deg(P) \leq 1$, so by Theorem 5.3.2 P is a Lawrence prism. Palindromicity implies $h^*_P(t) = 1 + t$, so lattice volume 2, which is a contradiction because any three-dimensional Lawrence prism has at least lattice volume 3. Hence, by Theorem 5.5.3 P must be a lattice pyramid or a Gorenstein join of two Gorenstein intervals one of them being thin. As a thin interval is a unimodular simplex, thus a lattice pyramid, also P is a lattice pyramid by Corollary 5.4.16.

5.5.2 Borisov's Proof of $\deg(\ell^*)$ for Gorenstein Polytopes

Theorem 5.5.3 answers affirmatively Question 6.3(b) in [NS13] asking whether for Gorenstein polytopes having a non-vanishing ℓ^*-polynomial forces its degree to be maximal (i.e., equal to the degree of the h^*-polynomial). Lev Borisov has provided us with an alternative algebraic proof of this fact that we reproduce here. It uses the description of the local h^*-polynomial of a lattice polytope as a Hilbert series of a graded ideal given in [BM03].

Proposition 5.5.9. Let $P \subseteq \mathbb{R}^d$ be a Gorenstein polytope of codegree r. Then either P is thin or $\ell^*_P(t)$ starts with t^r. In this case, $\ell^*_P(t)$ has degree $\deg(P)$ and leading coefficient 1.

Proof. Let $K \subseteq \mathbb{Z}^{d+1}$ be the lattice points in the Gorenstein cone over $P \times \{1\}$. Denote by $\mathbb{C}[K]$ the associated affine semi-group algebra with $\mathbb{N}_0$-grading given by the exponent of x_{d+1}, viewing $\mathbb{C}[K] \subseteq \mathbb{C}[x_1^{\pm 1}, \ldots, x_d^{\pm 1}, x_{d+1}]$. As in [BM03, Section 4], we let $f \in \mathbb{C}[K]_1$ be non-degenerate and $I \subseteq \mathbb{C}[K]$ the homogeneous ideal generated by the so called logarithmic derivatives of f. Let moreover $J \subseteq \mathbb{C}[K]$ be the homogeneous ideal generated by all lattice points in the relative interior K° of K. Then Borisov and Mavlyutov define $R_1(f, K)$ to be the image of J in the quotient ring $\mathbb{C}[K]/I$, i.e., $R_1(f, K)$ is the homogeneous ideal $(I + J)/I$ of $\mathbb{C}[K]/I$.
Now, by [BM03, Proposition 5.5], $\ell^*_P(t)$ is the Hilbert series of $R_1(f, K)$. Moreover, as P is Gorenstein of codegree r, we have $K^\circ = (p, r) + K$, where $p \in (rP) \cap \mathbb{Z}^d$ is the unique interior lattice point of rP. Therefore, $R_1(f, K)$ is just the image of the principal ideal $(x^p x^r_{d+1})$ in the quotient $\mathbb{C}[K]/I$. Hence, $R_1(f, K)$ is 0 if and only if $x^p x^r_{d+1} \in I$, and otherwise the lowest degree of its non-zero homogeneous components is r with $R_1(f, K)_r = \langle x^p x^r_{d+1} \rangle$ of vector space dimension 1. This proves the first claim, and the second follows from reciprocity, Theorem 5.1.16(2). $\square$

In dimensions ≤ 4 it is a consequence of the reciprocity of $\ell^*_P(t)$ that for any lattice polytope P either P is thin or $\deg(\ell^*_P) = \deg(P)$. In higher dimensions, this property fails for non-Gorenstein lattice polytopes.

Example 5.5.10. Consider the full-dimensional lattice simplex $P \subseteq \mathbb{R}^5$ given as the convex hull $P = \operatorname{conv}(0, e_1, e_2, e_3, (0,1,1,2,0), (5,3,3,2,6))$. Then $\ell^*_P(t) = 4t^3$ while $h^*_P(t) = t^4 + 5t^3 + 4t^2 + t + 1$. In particular, P is not thin but $\deg(\ell^*_P) < \deg(h^*_P) = \deg(P)$. This is the only such example among lattice simplices of dimension 5 with lattice volume ≤ 15. It was found using the database [Bal]. The computations were performed in `SageMath` with backend `Normaliz`.

Example 5.5.11. Consider the full-dimensional lattice simplex $P \subseteq \mathbb{R}^5$ given as the convex hull $P = \operatorname{conv}(0, e_1, e_2, (1,1,2,0,0), (3,5,6,8,0), (1,1,0,0,2))$. Then $\ell^*_P(t) = t^3$ while $h^*_P(t) = 7t^3 + 19t^2 + 5t + 1$. So $\deg(\ell^*_P) = \deg(h^*_P)$ but the leading coefficient of ℓ^*_P is strictly smaller.

5.5.3 Thinness Is Invariant under Duality

It was noted in Lemma 5.5.6 that being trivially thin as well as being g-thin is invariant under duality of Gorenstein polytopes. Let us explain how this allows us to deduce that also thinness has this beautiful duality property:

Corollary 5.5.12. Let P be a Gorenstein polytope. Then P is thin if and only if $P^\times$ is thin.

Proof. By Theorem 5.5.3(ii) we may assume that P is a thin Gorenstein polytope such that P is a Gorenstein join of faces F and G with F being trivally thin. Hence, by Lemma 5.5.6 we also have $P^\times = F^* *_{\mathrm{Gor}} G^*$ with G^* being trivially thin. Again, by Theorem 5.5.3 this implies that $P^\times$ is thin. □

Having such a direct proof answers a question of Lev Borisov, who communicated to us that this statement might also be proven using vertex algebra techniques.
In particular, as Theorem 5.5.3 implies that there are only two choices for the degree of the ℓ^*-polynomial of a Gorenstein polytope we see that its degree is also invariant under duality (as it holds for the degrees of the h^*-polynomial and the g-polynomial).

Corollary 5.5.13. Let P be a Gorenstein polytope. Then $\deg(\ell^*_P) = \deg(\ell^*_{P^\times})$.

Example 5.5.14. The reader should be aware that the local h^*-polynomials of a Gorenstein polytope P and its dual $P^\times$ may differ. For instance, for $P = [-1,1]^3$ we have $\ell^*_P(t) = t + 17t^2 + t^3$ and $\ell^*_{P^\times}(t) = t + 3t^2 + t^3$.

5.5.4 Thin Gorenstein Simplices

For the special case of Gorenstein simplices, we can answer the original question in [GKZ94] about classifying thin simplices.

Corollary 5.5.15. Let P be a Gorenstein simplex. Then P is thin if and only if P is a lattice pyramid.

Proof. Let P be thin. If P is g-thin, then $\deg(P) = \deg(g_P) = 0$ as P is a simplex. Hence, P is a unimodular simplex, in particular, a lattice pyramid.

Otherwise, Theorem 5.5.3(iii) implies that there are faces F and G of P such that $P = F *_{\mathrm{Gor}} G$ with G g-thin. As G is also a simplex, the previous consideration shows that G is a unimodular simplex, thus, a lattice pyramid. Hence, Corollary 5.4.16 implies that P is also a lattice pyramid. $\square$

In particular, if a Gorenstein simplex P satisfies $\dim(P) \geq 2\deg(P)$, then P is a lattice pyramid. This statement can also be deduced from [DRHNP13, Corollary 3.10(2)].

Bibliography

[AAZ19] Arash Amini, Bryon Aragam, and Qing Zhou, *On perfectness in Gaussian graphical models*, arXiv:1909.01978 (2019).

[AGK+21] Carlos Améndola, Lukas Gustafsson, Kathlén Kohn, Orlando Marigliano, and Anna Seigal, *The maximum likelihood degree of linear spaces of symmetric matrices*, Le Matematiche **76** (2021), no. 2, 535–557.

[AH07] Matthias Aschenbrenner and Christopher Hillar, *Finite generation of symmetric ideals*, Trans. Amer. Math. Soc. **359** (2007), no. 11, 5171–5192.

[Alu90] Paolo Aluffi, *The enumerative geometry of plane cubics. I. Smooth cubics*, Transactions of the American Mathematical Society **317** (1990), no. 2, 501–539.

[Alu91a] ———, *The enumerative geometry of plane cubics. II. Nodal and cuspidal cubics*, Math. Ann. **289** (1991), no. 4, 543–572.

[Alu91b] ———, *Some characteristic numbers for nodal and cuspidal plane curves of any degree*, Manuscripta Math. **72** (1991), no. 4, 425–444.

[Alu92] ———, *Two characteristic numbers for smooth plane curves of any degree*, Trans. Amer. Math. Soc. **329** (1992), no. 1, 73–96.

[APPS21] Karim Adiprasito, Stavros Papadakis, Vasiliki Petrotou, and Johanna Steinmeyer, *Beyond positivity in Ehrhart Theory*, arXiv:2103.14925 (2021).

[Ath12] Christos Athanasiadis, *Flag subdivisions and γ-vectors*, Pac. J. Math. **259** (2012), no. 2, 257–278.

[AWW11] Gennadiy Averkov, Christian Wagner, and Robert Weismantel, *Maximal lattice-free polyhedra: finiteness and an explicit description in dimension three*, Mathematics of Operations Research **36** (2011), no. 4, 721–742.

[Bal] Gabriele Balletti, *A database of "small" lattice polytopes*, `https://github.com/gabrieleballetti/small-lattice-polytopes`.

[Bat93] Victor V. Batyrev, *Variations of the mixed Hodge structure of affine hypersurfaces in algebraic tori*, Duke Math. J. **69** (1993), no. 2, 349–409.

[Bat06] Victor Batyrev, *Lattice polytopes with a given h^*-polynomial*, Algebraic and geometric combinatorics, Contemp. Math., vol. 423, Amer. Math. Soc., Providence, RI, 2006, pp. 1–10.

[BB96a] Victor Batyrev and Lev Borisov, *Mirror duality and string-theoretic Hodge numbers*, Invent. Math. **126** (1996), no. 1, 183–203.

[BB96b] ______, *On Calabi-Yau complete intersections in toric varieties*, Higher-dimensional complex varieties (Trento, 1994), de Gruyter, Berlin, 1996, pp. 39–65.

[BBBK11] Margherita Barile, Dominique Bernardi, Alexander Borisov, and Jean-Michel Kantor, *On empty lattice simplices in dimension 4*, Proc. Amer. Math. Soc. **139** (2011), no. 12, 4247–4253.

[BBD82] Alexander Beĭlinson, Joseph Bernstein, and Pierre Deligne, *Faisceaux pervers*, Analysis and topology on singular spaces, I (Luminy, 1981), Astérisque, vol. 100, Soc. Math. France, Paris, 1982, pp. 5–171.

[BC94] Victor V. Batyrev and David A. Cox, *On the Hodge structure of projective hypersurfaces in toric varieties*, Duke Math. J. **75** (1994), no. 2, 293–338.

[BC03] Winfried Bruns and Aldo Conca, *Gröbner bases and determinantal ideals*, Commutative algebra, singularities and computer algebra (Sinaia, 2002), NATO Sci. Ser. II Math. Phys. Chem., vol. 115, Kluwer Acad. Publ., Dordrecht, 2003, pp. 9–66.

[BCR98] Jacek Bochnak, Michel Coste, and Marie-Françoise Roy, *Real algebraic geometry*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3), vol. 36, Springer-Verlag, Berlin, 1998.

[BDFK23] Mara Belotti, Alessandro Danelon, Claudia Fevola, and Andreas Kretschmer, *The enumerative geometry of cubic hypersurfaces: point and line conditions*, Collect. Math. (2023), doi.org/10.1007/s13348-023-00401-z.

[BDKS19] Tobias Boege, Alessio D'Alì, Thomas Kahle, and Bernd Sturmfels, *The Geometry of Gaussoids*, Found. Comput. Math. **19** (2019), no. 4, 775–812.

[BG92] Nantel Bergeron and Adriano M. Garsia, *On certain spaces of harmonic polynomials*, Hypergeometric functions on domains of positivity, Jack polynomials, and applications (Tampa, FL, 1991), Contemp. Math., vol. 138, Amer. Math. Soc., Providence, RI, 1992, pp. 51–86.

[BH93] Winfried Bruns and Jürgen Herzog, *Cohen–Macaulay rings*, Cambridge Studies in Advanced Mathematics, vol. 39, Cambridge University Press, Cambridge, 1993.

[BK17] Samir Bouchiba and Salah-Eddine Kabbaj, *Embedding dimension and codimension of tensor products of algebras over a field*, Rings, polynomials, and modules, Springer, Cham, 2017, pp. 53–77.

[BK20] Tobias Boege and Thomas Kahle, *Construction Methods for Gaussoids*, Kybernetika **56** (2020), no. 6, 1045–1062.

[BKKR23] Tobias Boege, Thomas Kahle, Andreas Kretschmer, and Frank Röttger, *The geometry of Gaussian double Markovian distributions*, Scand. J. Stat. **50** (2023), no. 2, 665–696.

[BKN23] Christopher Borger, Andreas Kretschmer, and Benjamin Nill, *Thin polytopes: Lattice polytopes with vanishing local h^*-polynomial*, to appear in Int. Math. Res. Not. IMRN (2023), arXiv:2207.09323.

[BM85] Ulrich Betke and Peter McMullen, *Lattice points in lattice polytopes*, Monatsh. Math. **99** (1985), 253–265.

[BM03] Lev Borisov and Anvar Mavlyutov, *String cohomology of Calabi-Yau hypersurfaces via mirror symmetry*, Adv. Math. **180** (2003), no. 1, 355–390.

[BN07] Victor Batyrev and Benjamin Nill, *Multiples of lattice polytopes without interior lattice points*, Moscow Mathematical Journal **7** (2007), no. 2, 195–207.

[BN08] ———, *Combinatorial aspects of mirror symmetry*, Contemporary Mathematics **452** (2008), 35–66.

[Boe22] Tobias Boege, *Gaussoids are two-antecedental approximations of Gaussian conditional independence structures*, Ann. Math. Artif. Intell. **90** (2022), no. 6, 645–673.

[Boo12] Adam Boocher, *Free resolutions and sparse determinantal ideals*, Math. Res. Lett. **19** (2012), no. 4, 805–821.

[Boo13] ———, *Superflatness*, ProQuest LLC, Ann Arbor, MI, 2013, Ph.D. Thesis – University of California, Berkeley.

[BR03] Moulinath Banerjee and Thomas Richardson, *On a dualization of graphical Gaussian models: A correction note*, Scandinavian Journal of Statistics **30** (2003), no. 4, 817–820.

[Bra06] Tom Braden, *Remarks on the combinatorial intersection cohomology of fans*, Pure Appl. Math. Q. **2** (2006), no. 4, 1149–1186.

[Bra16] Benjamin Braun, *Unimodality problems in Ehrhart theory*, Recent trends in combinatorics, IMA Vol. Math. Appl., vol. 159, Springer, Cham, 2016, pp. 687–711.

[BS07] Markus Brodmann and Peter Schenzel, *Arithmetic properties of projective varieties of almost minimal degree*, J. Algebraic Geom. **16** (2007), no. 2, 347–400.

[BS18] Matthias Beck and Raman Sanyal, *Combinatorial reciprocity theorems*, Graduate Studies in Mathematics, vol. 195, American Mathematical Society, Providence, RI, 2018, An invitation to enumerative geometric combinatorics.

[BT18] Paul Breiding and Sascha Timme, *HomotopyContinuation.jl: A package for homotopy continuation in Julia*, International Congress on Mathematical Software, Springer, 2018, pp. 458–465.

[CCD+13] Eduardo Cattani, María Angélica Cueto, Alicia Dickenstein, Sandra Di Rocco, and Bernd Sturmfels, *Mixed discriminants*, Mathematische Zeitschrift **274** (2013), no. 3-4, 761–778.

[CDNG15] Aldo Conca, Emanuela De Negri, and Elisa Gorla, *Universal Gröbner bases for maximal minors*, Int. Math. Res. Not. IMRN (2015), no. 11, 3245–3262.

[CDNG17] ———, *Multigraded generic initial ideals of determinantal ideals*, Homological and computational methods in commutative algebra, Springer INdAM Ser., vol. 20, Springer, Cham, 2017, pp. 81–96. MR 3751880

[CDNG18] ———, *Cartwright-Sturmfels ideals associated to graphs and linear spaces*, J. Comb. Algebra **2** (2018), no. 3, 231–257.

[CDNG20] ———, *Universal Gröbner bases and Cartwright-Sturmfels ideals*, Int. Math. Res. Not. IMRN (2020), no. 7, 1979–1991.

[CDNG22] ———, *Radical generic initial ideals*, Vietnam J. Math. **50** (2022), no. 3, 807–827.

[CEF15] Thomas Church, Jordan S. Ellenberg, and Benson Farb, *FI-modules and stability for representations of symmetric groups*, Duke Math. J. **164** (2015), no. 9, 1833–1910.

[CF13] Thomas Church and Benson Farb, *Representation theory and homological stability*, Adv. Math. **245** (2013), 250–314.

[Cha64] Michel Chasles, *Dètermination du nombre de sections coniques qui doivent toucher cinq courbes donneés d'ordre quelconque, ou satisfaire à diverses autres conditions*, C.R. Acad. Sci. Paris **58** (1864), 222–226.

[CHT06] Aldo Conca, Serkan Hoşten, and Rekha Thomas, *Nice initial complexes of some classical ideals*, Algebraic and geometric combinatorics, Contemp. Math., vol. 423, Amer. Math. Soc., Providence, RI, 2006, pp. 11–42.

[CK11] Imre Csiszár and János Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, 2 ed., Cambridge University Press, 2011.

[CLO96] David Cox, John Little, and Donal O'Shea, *Ideals, varieties, and algorithms*, second ed., UTM, Springer, New York, 1996.

[Coh67] Daniel Cohen, *On the laws of a metabelian variety*, J. Algebra **5** (1967), 267–273.

[Con94a] Aldo Conca, *Divisor class group and canonical class of determinantal rings defined by ideals of minors of a symmetric matrix*, Archiv der Mathematik **63** (1994), no. 3, 216–224.

[Con94b] ______, *Gröbner bases of ideals of minors of a symmetric matrix*, J. Algebra **166** (1994), no. 2, 406–421.

[Con94c] ______, *Symmetric ladders*, Nagoya Mathematical Journal **136** (1994), 35–56.

[Cos06a] Izzet Coskun, *Degenerations of surface scrolls and the Gromov-Witten invariants of Grassmannians*, J. Algebraic Geom. **15** (2006), no. 2, 223–284.

[Cos06b] ______, *The enumerative geometry of Del Pezzo surfaces via degenerations*, Amer. J. Math. **128** (2006), no. 3, 751–786.

[CRS19] Rainelly Cunha, Zaqueu Ramos, and Aron Simis, *Symmetry preserving degenerations of the generic symmetric matrix*, J. Algebra **523** (2019), 154–191.

[CV20] Aldo Conca and Matteo Varbaro, *Square-free Gröbner degenerations*, Invent. Math. **221** (2020), no. 3, 713–730.

[CW93] David Cox and Nanny Wermuth, *Linear Dependencies Represented by Chain Graphs*, Statistical Science **8** (1993), no. 3, 204 – 218.

[CW19] Aldo Conca and Volkmar Welker, *Lovász-Saks-Schrijver ideals and coordinate sections of determinantal varieties*, Algebra Number Theory **13** (2019), no. 2, 455–484.

[dCM09] Mark Andrea de Cataldo and Luca Migliorini, *The decomposition theorem, perverse sheaves and the topology of algebraic maps*, Bull. Am. Math. Soc., New Ser. **46** (2009), no. 4, 535–633.

[dCMM18] Mark Andrea de Cataldo, Luca Migliorini, and Mircea Mustaţă, *Combinatorics and topology of proper toric maps*, J. Reine Angew. Math. **744** (2018), 133–163.

[DCP85] Claudio De Concini and Corrado Procesi, *Complete symmetric varieties. II. Intersection theory*, Algebraic groups and related topics, Adv. Stud. Pure Math., vol. 6, Amsterdam, 1985, pp. 481–513.

[DK23] Jiahe Deng and Andreas Kretschmer, *Ideals of submaximal minors of sparse symmetric matrices*, to appear in J. Pure Appl. Algebra (2023), arXiv:2211.07020.

[DKNS21] Joseph Doolittle, Lukas Katthän, Benjamin Nill, and Francisco Santos, *Empty simplices of large width*, arXiv:2103.14925 (2021).

[dMGP+20] André de Moura, Elijah Gunther, Sam Payne, Jason Schuchardt, and Alan Stapledon, *Triangulations of simplices with vanishing local h-polynomial*, Algebr. Comb. **3** (2020), no. 6, 1417–1430.

[DMV21] Rodica Dinu, Mateusz Michałek, and Martin Vodička, *Geometry of the Gaussian graphical model of the cycle*, arXiv:2111.02937 (2021).

[DN10] Alicia Dickenstein and Benjamin Nill, *A simple combinatorial criterion for projective toric manifolds with dual defect*, Math. Res. Lett. **17** (2010), no. 3, 435–448.

[DR02] Mathias Drton and Thomas Richardson, *A new algorithm for maximum likelihood estimation in Gaussian graphical models for marginal independence*, Proceedings of the Nineteenth Conference on Uncertainty in Artificial Intelligence (San Francisco, CA, USA), UAI'03, Morgan Kaufmann Publishers Inc., 2002, p. 184–191.

[DR08] ______, *Graphical methods for efficient likelihood inference in Gaussian covariance models*, Journal of Machine Learning Research **9** (2008), no. 29, 893–914.

[Dra14] Jan Draisma, *Noetherianity up to symmetry*, Combinatorial algebraic geometry, Lecture Notes in Math., vol. 2108, Springer, Cham, 2014, pp. 33–61.

[DRHNP13] Sandra Di Rocco, Christian Haase, Benjamin Nill, and Andreas Paffenholz, *Polyhedral adjunction theory*, Algebra Number Theory **7** (2013), no. 10, 2417–2446.

[Drt09] Mathias Drton, *Likelihood ratio tests and singularities*, The Annals of Statistics **37** (2009), no. 2, 979–1012.

[DSS09] Mathias Drton, Bernd Sturmfels, and Seth Sullivant, *Lectures on algebraic statistics*, Oberwolfach Seminars, vol. 39, Birkhäuser Verlag, Basel, 2009.

[DX10] Mathias Drton and Han Xiao, *Smoothness of Gaussian conditional independence models*, Contemp. Math **516** (2010), 155–177.

[Efr78] Bradley Efron, *The Geometry of Exponential Families*, The Annals of Statistics **6** (1978), no. 2, 362–376.

[EH16] David Eisenbud and Joe Harris, *3264 and all that—a second course in algebraic geometry*, Cambridge University Press, Cambridge, 2016.

[Ehr62] Eugène Ehrhart, *Sur les polyèdres rationnels homothétiques à n dimensions*, C. R. Acad. Sci. Paris **254** (1962), 616–618.

[Eis88] David Eisenbud, *Linear sections of determinantal varieties*, Amer. J. Math. **110** (1988), no. 3, 541–575.

[Eis95] ———, *Commutative Algebra, with a View Toward Algebraic Geometry*, Graduate Texts in Mathematics, vol. 150, Springer-Verlag, New York, 1995.

[EKS14] Alexander Engström, Thomas Kahle, and Seth Sullivant, *Multigraded commutative algebra of graph decompositions*, Journal of Algebraic Combinatorics **39** (2014), no. 2, 335–372.

[Eva15] Robin Evans, *Smoothness of marginal log-linear parameterizations*, Electron. J. Stat. **9** (2015), no. 1, 475–491.

[FH91] William Fulton and Joe Harris, *Representation Theory, A First Course*, Graduate Texts in Mathematics, vol. 129, Springer-Verlag, New York, 1991.

[FKM83] William Fulton, Steven Kleiman, and Robert MacPherson, *About the enumeration of contacts*, Algebraic geometry—open problems (Ravello, 1982), Lecture Notes in Math., vol. 997, Springer, Berlin, 1983, pp. 156–196.

[For12] Antonio Forcina, *Smoothness of conditional independence models for discrete data*, J. Multivariate Anal. **106** (2012), 49–56.

[Ful98] William Fulton, *Intersection theory*, second ed., Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics, vol. 2, Springer-Verlag, Berlin, 1998.

[Gey94] Charles Geyer, *On the asymptotics of constrained M-estimation*, The Annals of Statistics **22** (1994), no. 4, 1993–2010.

[GKZ94] Israel Gelfand, Mikhail Kapranov, and Andrei Zelevinsky, *Discriminants, resultants, and multidimensional determinants*, Mathematics: Theory & Applications, Birkhäuser Boston, Inc., Boston, MA, 1994.

[GM82] Marc Giusti and Michel Merle, *Singularités isolées et sections planes de variétés déterminantielles. II. Sections de variétés déterminantielles par les plans de coordonnées*, Algebraic geometry (La Rábida, 1981), Lecture Notes in Math., vol. 961, Springer, Berlin, 1982, pp. 103–118.

[GP92] Adriano M. Garsia and Claudio Procesi, *On certain graded S_n-modules and the q-Kostka polynomials*, Adv. Math. **94** (1992), no. 1, 82–138.

[GP10] Victor Guillemin and Alan Pollack, *Differential topology*, AMS Chelsea Publishing, Providence, RI, 2010, Reprint of the 1974 original.

[GS] Daniel R. Grayson and Michael E. Stillman, *Macaulay2, a software system for research in algebraic geometry*, Available at http://www.math.uiuc.edu/Macaulay2/.

[GS20] Nils Gustafsson and Liam Solus, *Derangements, Ehrhart theory, and local h-polynomials*, Adv. Math. **369** (2020), 107169, 35.

[Hai03] Mark Haiman, *Combinatorics, symmetric functions, and Hilbert schemes*, Current developments in mathematics, 2002, Int. Press, Somerville, MA, 2003, pp. 39–111.

[Har77] Robin Hartshorne, *Algebraic geometry*, Graduate Texts in Mathematics, No. 52, Springer-Verlag, New York-Heidelberg, 1977.

[Har95] Joe Harris, *Algebraic geometry. a first course*, Graduate Texts in Mathematics, vol. 133, Springer-Verlag, New York, 1995.

[HH11] Jürgen Herzog and Takayuki Hibi, *Monomial ideals*, Graduate Texts in Mathematics, vol. 260, Springer-Verlag London, Ltd., London, 2011.

[HHS23] Katherine Harris, Jonathan Hauenstein, and Agnes Szanto, *Smooth points on semi-algebraic sets*, J. Symbolic Comput. **116** (2023), 183–212.

[Hig19] Akihiro Higashitani, *Lattice simplices of maximal dimension with a given degree*, Michigan Math. J. **68** (2019), no. 1, 193–210.

[HKN18] Johannes Hofscheier, Lukas Katthän, and Benjamin Nill, *Ehrhart theory of spanning lattice polytopes*, Int. Math. Res. Not. **2018** (2018), no. 19, 5947–5973.

[HNP09] Christian Haase, Benjamin Nill, and Sam Payne, *Cayley decompositions of lattice polytopes and upper bounds for h^*-polynomials*, J. Reine Angew. Math. **637** (2009), 207–216.

[Hoc72] Melvin Hochster, *Rings of invariants of tori, Cohen-Macaulay rings generated by monomials, and polytopes*, Ann. of Math. (2) **96** (1972), 318–337.

[HS12] Christopher Hillar and Seth Sullivant, *Finite Gröbner bases in infinite dimensional polynomial rings and applications*, Adv. Math. **229** (2012), no. 1, 1–25.

[HT09] Martin Henk and Makoto Tagami, *Lower bounds on the coefficients of Ehrhart polynomials*, Eur. J. Comb. **30** (2009), no. 1, 70–83.

[IVnS21] Óscar Iglesias-Valiño and Francisco Santos, *The complete classification of empty lattice 4-simplices*, Rev. Mat. Iberoam. **37** (2021), no. 6, 2399–2432.

[JKLR20] Martina Juhnke-Kubitzke, Dinh Van Le, and Tim Römer, *Asymptotic behavior of symmetric ideals: a brief survey*, Combinatorial structures in algebra and geometry, Springer Proc. Math. Stat., vol. 331, Springer, Cham, 2020, pp. 73–94.

[Józ78] Tadeusz Józefiak, *Ideals generated by minors of a symmetric matrix*, Comment. Math. Helv. **53** (1978), 595–607.

[JW05] Beatrix Jones and Mike West, *Covariance decomposition in undirected Gaussian graphical models*, Biometrika **92** (2005), no. 4, 779–786.

[Kar04] Kalle Karu, *Hard Lefschetz theorem for nonrational polytopes*, Invent. Math. **157** (2004), no. 2, 419–447.

[Kar08] ______, *Ehrhart analogue of the h-vector*, Integer points in polyhedra—geometry, number theory, representation theory, algebra, optimization, statistics, Contemp. Math., vol. 452, Amer. Math. Soc., Providence, RI, 2008, pp. 139–146.

[Kar19] ______, *Relative hard Lefschetz theorem for fans*, Adv. Math. **347** (2019), 859–903.

[Kau96] Göran Kauermann, *On a dualization of graphical Gaussian models*, Scandinavian Journal of Statistics **23** (1996), no. 1, 105–116.

[Kem11] Gregor Kemper, *A course in commutative algebra*, Graduate Texts in Mathematics, vol. 256, Springer, Heidelberg, 2011.

[KR13] Tarik Kaced and Andrei Romashchenko, *Conditional information inequalities for entropic and almost entropic points*, IEEE Transactions on Information Theory **59** (2013), no. 11, 7149–7167.

[Kre] Andreas Kretschmer, *Database of thin lattice polytopes*, https://github.com/A-Kretschmer/thin-lattice-polytopes.

[Kre23] ______, *When are symmetric ideals monomial?*, Journal of Commutative Algebra **15** (2023), no. 3, 367 – 376.

[KRS14] Thomas Kahle, Johannes Rauh, and Seth Sullivant, *Positive margins and primary decomposition*, Journal of Commutative Algebra **6** (2014), no. 2, 173–208.

[KS86] Steven Kleiman and Robert Speiser, *Enumerative geometry of cuspidal plane cubics*, Proceedings of the 1984 Vancouver conference in algebraic geometry, CMS Conf. Proc., vol. 6, Amer. Math. Soc., Providence, RI, 1986, pp. 227–268.

[KS88] ______, *Enumerative geometry of nodal plane cubics*, Algebraic geometry (Sundance, UT, 1986), Lecture Notes in Math., vol. 1311, Springer, Berlin, 1988, pp. 156–196.

[KS91] ______, *Enumerative geometry of nonsingular plane cubics*, Algebraic geometry: Sundance 1988, Contemp. Math., vol. 116, Amer. Math. Soc., Providence, RI, 1991, pp. 85–113.

[KS16] Eric Katz and Alan Stapledon, *Local h-polynomials, invariants of subdivisions, and mixed Ehrhart theory*, Adv. Math. **286** (2016), 181–239.

[KV08] Matthias Köppe and Sven Verdoolaege, *Computing parametric rational generating functions with a primal Barvinok algorithm*, Electron. J. Comb. **15** (2008), no. 1, research paper r16, 19.

[Lau96] Steffen Lauritzen, *Graphical models*, vol. 17, Oxford University Press, 1996.

[LM07] Radim Lněnička and František Matúš, *On Gaussian conditional independence structures*, Kybernetika **43** (2007), no. 3, 327–342.

[LNNR20] Dinh Van Le, Uwe Nagel, Hop D. Nguyen, and Tim Römer, *Codimension and projective dimension up to symmetry*, Math. Nachr. **293** (2020), no. 2, 346–362.

[LNNR21] ______, *Castelnuovo-Mumford regularity up to symmetry*, Int. Math. Res. Not. IMRN (2021), no. 14, 11010–11049.

[LPS11] Wanseok Lee, Euisung Park, and Peter Schenzel, *On the classification of non-normal cubic hypersurfaces*, J. Pure Appl. Algebra **215** (2011), no. 8, 2034–2042.

[LS18] Steffen Lauritzen and Kayvan Sadeghi, *Unifying Markov properties for graphical models*, Ann. Statist. **46** (2018), no. 5, 2251–2278.

[LZ22] Steffen Lauritzen and Piotr Zwiernik, *Locally associated graphical models and mixed convex exponential families*, Ann. Statist. **50** (2022), no. 5, 3009–3038.

[Mac71] Ian Macdonald, *Polynomials associated with finite cell-complexes*, J. London Math. Soc. (2) **4** (1971), 181–192.

[Mai71] Synère-Nicolas Maillard, *Recherches des caractéristiques des systèmes élémentaires de courbes planes du troisième ordre*, Ph.D. thesis, Imprimerie Cusset, 1871.

[Mat94] František Matúš, *Probabilistic conditional independence structures and matroid theory: Background*, Int. J. Gen. Syst. **22** (1994), no. 2, 185–196.

[Mat97] ______, *Conditional independence structures examined via minors*, Ann. Math. Artif. Intell. **21** (1997), no. 1, 99–30.

[Mat05] ———, *Conditional independences in Gaussian vectors and rings of polynomials*, Conditionals, Information, and Inference (Gabriele Kern-Isberner, Wilhelm Rödder, and Friedhelm Kulmann, eds.), Springer, 2005, pp. 152–161.

[Mat12] ———, *On conditional independence and log-convexity*, Ann. Inst. Henri Poincaré, Probab. Stat. **48** (2012), no. 4, 1137–1147.

[Mic21] Anna Franziska Michael, *Polytope von kleinem kombinatorischem Grad und die Batyrev-Juny-Vermutung*, Master's thesis, OVGU Magdeburg, 2021.

[MMM+23] Laurent Manivel, Mateusz Michałek, Leonid Monin, Tim Seynnaeve, and Martin Vodička, *Complete quadrics: Schubert calculus for Gaussian models and semidefinite programming*, J. Eur. Math. Soc. (2023), doi.org/10.4171/jems/1330.

[MMW21] Mateusz Michałek, Leonid Monin, and Jarosław A Wiśniewski, *Maximum likelihood degree, complete quadrics, and $\mathbb{C}^*$-action*, SIAM Journal on Applied Algebra and Geometry **5** (2021), no. 1, 60–85.

[MOY22] Satoshi Murai, Hidefumi Ohsugi, and Kohji Yanagawa, *A note on the reducedness and Gröbner bases of Specht ideals*, Comm. Algebra **50** (2022), no. 12, 5430–5434.

[MR07] Rosa M. Miró-Roig, *Betti numbers of determinantal ideals*, J. Algebra **318** (2007), no. 2, 653–668.

[MR08] ———, *Determinantal ideals*, Progress in Mathematics, vol. 264, Birkhäuser Verlag, Basel, 2008.

[MR22] Satoshi Murai and Claudiu Raicu, *An equivariant Hochster's formula for $\mathfrak{S}_n$-invariant monomial ideals*, J. London Math. Soc. **105** (2022), no. 3, 1974–2010.

[MRV21] Philippe Moustrou, Cordian Riener, and Hugues Verdure, *Symmetric ideals, Specht polynomials and solutions to symmetric systems of equations*, J. Symbolic Comput. **107** (2021), 106–121.

[MS21] Pratik Misra and Seth Sullivant, *Gaussian graphical models with toric vanishing ideals*, Annals of the Institute of Statistical Mathematics **73** (2021), 757–785.

[Nil08] Benjamin Nill, *Lattice polytopes having h^*-polynomials with given degree and linear coefficient*, European J. Combin. **29** (2008), no. 7, 1596–1602.

[Nil20] ———, *The mixed degree of families of lattice polytopes*, Annals of Combinatorics **24** (2020), no. 1, 203–216.

[NR17] Uwe Nagel and Tim Römer, *Equivariant Hilbert series in non-noetherian polynomial rings*, J. Algebra **486** (2017), 204–245.

[NR19] ———, *FI- and OI-modules with varying coefficients*, J. Algebra **535** (2019), 286–322.

[NS12] Benjamin Nill and Jan Schepers, *Combinatorial questions related to stringy E-polynomials of Gorenstein polytopes*, Oberwolfach Reports **21** (2012), 62 – 64.

[NS13] ———, *Gorenstein polytopes and their stringy E-functions*, Math. Ann. **355** (2013), no. 2, 457–480.

[NS21] Rohit Nagpal and Andrew Snowden, *Symmetric ideals of the infinite polynomial ring*, arXiv:2107.13027 (2021).

[NZ11] Benjamin Nill and Günter Ziegler, *Projecting lattice polytopes without interior lattice points*, Mathematics of Operations Research **36** (2011), no. 3, 462–467.

[PW94] Judea Pearl and Nanny Wermuth, *When can association graphs admit a causal interpretation?*, Selecting Models from Data, Springer, 1994, pp. 205–214.

[Sad17] Kayvan Sadeghi, *Faithfulness of probability distributions and graphs*, The Journal of Machine Learning Research **18** (2017), no. 1, 5429–5457.

[Sch79] Hermann Schubert, *Kalkül der abzählenden Geometrie*, BG Teubner, 1879.

[Sch12] Jan Schepers, *Stringy Hodge numbers of strictly canonical nondegenerate singularities*, J. Algebraic Geom. **21** (2012), no. 2, 273–297.

[Sch13] Rolf Schneider, *Convex Bodies: The Brunn–Minkowski Theory*, 2nd ed., Encyclopedia of Mathematics and its Applications, Cambridge University Press, 2013.

[Seg42] Beniamino Segre, *The Non-singular Cubic Surfaces*, Oxford University Press, Oxford, 1942.

[Sem48] John Semple, *On complete quadrics (i)*, Journal of the London Mathematical Society **1** (1948), no. 4, 258–267.

[SK86] Terence Speed and Harri Kiiveri, *Gaussian Markov distributions over finite graphs*, The Annals of Statistics **14** (1986), no. 1, 138 – 150.

[Sla15] Kaloyan Slavov, *The moduli space of hypersurfaces whose singular locus has high dimension*, Mathematische Zeitschrift **279** (2015), no. 1, 139–162.

[Sol19] Liam Solus, *Local h^*-polynomials of some weighted projective spaces*, Algebraic and geometric combinatorics on lattice polytopes. Proceedings of the summer workshop on lattice polytopes, Osaka, Japan, July 23 – August 10, 2018, World Scientific, 2019, pp. 382–399.

[Sta75] Richard Stanley, *The upper bound conjecture and Cohen-Macaulay rings*, Stud. Appl. Math. **54** (1975), 135–142.

[Sta80] ———, *Decompositions of rational convex polytopes*, Ann. Discrete Math. **6** (1980), 333–342.

[Sta87] ———, *Generalized H-vectors, intersection cohomology of toric varieties, and related results*, Commutative algebra and combinatorics (Kyoto, 1985), Adv. Stud. Pure Math., vol. 11, North-Holland, Amsterdam, 1987, pp. 187–213.

[Sta92] ———, *Subdivisions and local h-vectors*, J. Amer. Math. Soc. **5** (1992), no. 4, 805–851.

[Ste48] Jakob Steiner, *9. Elementare Lösung einer geometrischen Aufgabe, und über einige damit in Beziehung stehende Eigenschaften der Kegelschnitte*, J. Reine Angew. Math. **37** (1848), 161–192b.

[Ste86] Ulrich Sterz, *Berührungsvervollständigung für ebene Kurven dritter Ordnung. IV. Berührungsbedingungen*, Beiträge zur Algebra und Geometrie **21** (1986), 91–108.

[Stu92] Milan Studený, *Conditional independence relations have no finite complete characterization*, Information Theory, Statistical Decision Functions and Random Processes. Transactions of the 11th Prague Conference (S. Kubik and J.A. Visek, eds.), vol. B, Kluwer, Dordrecht, 1992, pp. 377–396.

[Stu17] Bernd Sturmfels, *The Hurwitz form of a projective variety*, J. Symbolic Comput. **79** (2017), no. 1, 186–196.

[SU10] Bernd Sturmfels and Caroline Uhler, *Multivariate Gaussian, semidefinite matrix completion, and convex algebraic geometry*, Ann. Inst. Statist. Math. **62** (2010), no. 4, 603–638.

[Suk20] Eunice Sukarto, *Orbit and Orbit Closure Containments for Cubic Surfaces*, arXiv:2006.11688 (2020).

[Sul07] Seth Sullivant, *Toric fiber products*, J. Algebra **316** (2007), no. 2, 560–577.

[Sul09] ———, *Gaussian conditional independence relations have no finite complete characterization*, Journal of Pure and Applied Algebra **213** (2009), no. 8, 1502–1506.

[Sul18] ———, *Algebraic statistics*, Graduate Studies in Mathematics, vol. 194, American Mathematical Society, 2018.

[Sun19] Rolf Sundberg, *Statistical modelling by exponential families*, Institute of Mathematical Statistics Textbooks, Cambridge University Press, 2019.

[SV16] Mingmin Shen and Charles Vial, *The Fourier transform for certain hyperkähler fourfolds*, Mem. Amer. Math. Soc. **240** (2016), no. 1139, vii+163.

[SVL13] Jan Schepers and Leen Van Langenhoven, *Unimodality questions for integrally closed lattice polytopes*, Ann. Comb. **17** (2013), no. 3, 571–589.

[Tan82] Toshiyuki Tanisaki, *Defining ideals of the closures of the conjugacy classes and representations of the Weyl groups*, Tohoku Math. J. (2) **34** (1982), no. 4, 575–585.

[Tse20] Dennis Tseng, *Collections of hypersurfaces containing a curve*, Int. Math. Res. Not. (2020), no. 13, 3927–3977.

[Uhl19] Caroline Uhler, *Gaussian graphical models: An algebraic and geometric perspective*, Handbook of Graphical Models (Mathias Drton, Steffen Lauritzen, Marloes Maathuis, and Martin Wainwright, eds.), Chapman Hall, 2019.

[Vai82] Israel Vainsencher, *Schubert calculus for complete quadrics*, Enumerative Geometry and Classical Algebraic Geometry, Springer, 1982, pp. 199–235.

[Vak99] Ravi Vakil, *The characteristic numbers of quartic plane curves*, Canad. J. Math. **51** (1999), no. 5, 1089–1120.

[vdV98] Aad van der Vaart, *Asymptotic Statistics*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, 1998.

[vG91] Leendert van Gastel, *Characteristic numbers of plane curves*, Enumerative Algebraic Geometry: Proceedings of the 1989 Zeuthen Symposium, vol. 123, American Mathematical Soc., 1991, p. 259.

[Zeu72] Hieronymus Zeuthen, *Déterminations des caractéristique des systémes élémentaires de cubiques.*, C. R. Acad. Sci., Paris **74** (1872), 521–524.

[Zeu73] ———, *Almindelige Egenskaber ved Systemer af plane Kurver*, Kongelige Danske Videnskabernes Selskabs Skrifter – Naturvidenskabelig og Mathematisk **10** (1873), 285–393.