Algebraic Geometry of Bayesian Networks

Luis David Garcia–Puente

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Mathematics

Reinhard Laubenbacher, Chair
Ezra Brown
Daniel Farkas
Edward Green
Mark Shimozono

April 01, 2004
Blacksburg, Virginia

Keywords: Algebraic Geometry, Bayesian Networks, Computational Algebra, Statistics

Algebraic Geometry of Bayesian Networks

Luis David Garcia–Puente

(ABSTRACT)

We develop the necessary theory in algebraic geometry to place Bayesian networks into the realm of algebraic statistics. This allows us to create an algebraic geometry–statistics dictionary. In particular, we study the algebraic varieties defined by the conditional independence statements of Bayesian networks. A complete algebraic classification, in terms of primary decomposition of polynomial ideals, is given for Bayesian networks on at most five random variables. Hidden variables are related to the geometry of higher secant varieties. Moreover, a complete algebraic classification, in terms of generating sets of polynomial ideals, is given for Bayesian networks on at most three random variables and one hidden variable. The relevance of these results for model selection is discussed.

Partially supported by the CARGO program of the National Science Foundation (DMS-0138323).

Dedication

To God.

To my wife *Rebecca* for all her love and support.

To my baby daughter *Claudia Isabel*.

To my parents *Isabel* and *Luis David*.

Acknowledgments

First, I want to thank my advisor *Reinhard Laubenbacher* for all his support in these last five years of my life. Reinhard has taught me many things inside and outside of the mathematics world. He has been a good mentor and a good friend.

I want to thank *Bernd Sturmfels* and *Michael Stillman* for all their support. The project undertaken in this thesis was suggested by Bernd and many results were obtained in collaboration with both of them. I also want to thank both for being and endless source of motivation.

I want to thank *Lorenzo Robbiano* for his friendship and all his teachings. Many of the results in this thesis were obtained while I was visiting the Department of Mathematics at the University of Genova.

I want to thank all the professors that form my committee. I also want to thank the *Virginia Polytechnic Institute and State University* and *New Mexico State University* for giving me the opportunity to be part of their wonderful departments of mathematics.

I want to thank all my professors from whom I learned so much mathematics. Also, I want to thank all my friends for their support and their help.

Finally, I want to thank my family: Rebecca and Claudia for being my life. My parents, Christian and Alain for being a sea of love and support.

Contents

List of Tables

Chapter 1

Introduction

The emerging field of *algebraic statistics* [23] advocates polynomial algebra as a tool in the statistical analysis of experiments and discrete data. Statistics textbooks define a *statistical model* as a family of probability distributions, and a closer look reveals that these families are often real algebraic varieties: they are the zeros of some polynomials in the probability simplex [11, 26].

In this work we examine *directed graphical models* for discrete random variables. Such models are also known as *Bayesian networks* and they are widely used in machine learning, bioinformatics and many other applications [20, 22]. Our aim is to place Bayesian networks into the realm of algebraic statistics, by developing the necessary theory in algebraic geometry and by demonstrating the effectiveness of Gröbner bases for this class of models.

Bayesian networks can be described in two possible ways, either by a recursive factorization of probability distributions or by conditional independence statements (local and global Markov properties). This is an instance of the computer algebra principle that varieties can be presented either parametrically or implicitly [4, §3.3]. The equivalence of these two representations for Bayesian networks is a well–known theorem in statistics [20, Theorem 3.27], but, as we shall see, this theorem is surprisingly delicate and no longer holds when probabilities are replaced by negative reals or complex numbers. Hence in the usual setting of algebraic geometry, where the zeros lie in $\mathbb{C}^d$, there are many "distributions" which satisfy the global Markov property but which do not permit a recursive factorization. We explain this phenomenon using primary decomposition of polynomial ideals.

This thesis is organized as follows. In Chapter 2 we review the algebraic theory of conditional independence, and we explicitly determine the Gröbner basis and primary decomposition arising from the *contraction axiom* [22], [31, §2.2.2]. This axiom is shown to fail for negative real numbers. In Chapter 3 we introduce the ideals $I_{\text{local}(G)}$ and $I_{\text{global}(G)}$ which represent a Bayesian network G. When G is a forest then these ideals are the toric ideals derived from undirected graphs as in [12]; see Theorem 7 below.

The recursive factorization of a Bayesian network gives rise to a map between polynomial rings which is studied in Chapter 4. The kernel of this *factorization map* is the *distinguished prime ideal*. We prove that this prime ideal is always a reduced primary component of $I_{\mathrm{local}(G)}$ and $I_{\mathrm{global}(G)}$. Our results in that chapter include the solutions to Problems 8.11 and 8.12 in [33].

In Chapters 5 and 6 we present the results of our computational efforts: the complete algebraic classification of all Bayesian networks on four arbitrary random variables and all Bayesian networks on five binary random variables. The latter involved computing the primary decomposition of 301 ideals generated by a large number of quadrics in 32 unknowns. These large–scale primary decompositions were carried out in `Macaulay2` [14] and Singular [15]. Some of the techniques and software tools we used are described in the Appendix.

The appearance of hidden variables in Bayesian networks leads to challenging problems in algebraic geometry. Statisticians have known for decades that the dimension of the corresponding varieties can unexpectedly drop [13], but the responsible singularities have been studied only quite recently, in [11] and [26]. In Chapter 7 we examine the elimination problem arising from hidden random variables, and we relate it to problems in projective algebraic geometry. We demonstrate that the *naive Bayes model* corresponds to the higher secant varieties of Segre varieties [2, 3], and we present several new results on the dimension and defining ideals of these secant varieties.

In Chapter 8 we expand the results obtained in the previous chapter. We compute the ideal of all polynomial functions which vanish on the space of observable distributions implied by any Bayesian network on three observable variables and one hidden variable. These results extend previous work on the subject, see [10, 11]. In particular, we were able to conclude that all but one of such Bayesian networks correspond to (joins or intersections) of higher secant varieties of Segre varieties. This result is obtained via the computation of the dimension of all ideals involved. It was shown in [10, 24, 25] that the understanding and computation of the dimension and singularities of these varieties is crucial for the relevant problem of model selection. Our results provide new insight in this direction.

Our algebraic theory does not compete with but rather complements other approaches to conditional independence models. An impressive combinatorial theory of such models has been developed by Matúš [21] and Studený [31], culminating in their characterization of all realizable independence models on four random variables. Sharing many of the views expressed by these authors, we believe that exploring the precise relation between their work and ours will be a very fruitful research direction for the near future.

Most of the results contained in the first chapters were obtained in collaboration with Dr. Michael Stillman and Dr. Bernd Sturmfels. A paper containing these results will appear in a special issue of the Journal of Symbolic Computation in the occasion of the MEGA 2003 Conference, see [8]. Chapter 8 consists of original, previously unpublished results. Nevertheless, a paper containing a summary of the results in this last chapter has been submitted to the Conference of Uncertainty in Artificial Intelligence, see [9].

Chapter 2

Ideals, Varieties and Independence Models

2.1 Conditional Independence

We begin by reviewing the general algebraic framework for independence models presented in [33, §8]. Let $X_1, \ldots, X_n$ be discrete random variables where X_i takes values in the finite set $[d_i] = \{1, 2, \ldots, d_i\}$. We write $D = [d_1] \times [d_2] \times \cdots \times [d_n]$ so that $\mathbb{R}^D$ denotes the real vector space of n-dimensional tables of format $d_1 \times \cdots \times d_n$. We introduce an indeterminate $p_{u_1 u_2 \cdots u_n}$ which represents the probability of the event $X_1 = u_1, X_2 = u_2, \ldots, X_n = u_n$. These indeterminates generate the ring $\mathbb{R}[D]$ of polynomial functions on the space of tables $\mathbb{R}^D$. A *conditional independence statement* has the form

$$A \text{ is independent of } B \text{ given } C \qquad (\text{ in symbols: } A \perp\!\!\!\perp B \mid C) \tag{2.1}$$

where A, B and C are pairwise disjoint subsets of $\{X_1, \ldots, X_n\}$. If C is empty then (2.1) means that A is independent of B. By [33, Proposition 8.1], the statement (2.1) translates into a set of homogeneous quadratic polynomials in $\mathbb{R}[D]$, and we write $I_{A \perp\!\!\!\perp B \mid C}$ for the ideal generated by these polynomials.

Many statistical models (see e.g. [20, 31]) can be described by a finite set of independence statements (2.1). An *independence model* is any such set:

$$\mathcal{M} \quad = \quad \{A^{(1)} \perp\!\!\!\perp B^{(1)} \mid C^{(1)}, \ldots, A^{(m)} \perp\!\!\!\perp B^{(m)} \mid C^{(m)}\}.$$

The ideal of the independence model $\mathcal{M}$ is defined as the sum of ideals

$$I_{\mathcal{M}} \quad = \quad I_{A^{(1)} \perp\!\!\!\perp B^{(1)} \mid C^{(1)}} + \cdots + I_{A^{(m)} \perp\!\!\!\perp B^{(m)} \mid C^{(m)}}.$$

We wrote code in `Macaulay2` [14] and `Singular` [15] for generating the ideals $I_{\mathcal{M}}$. The *independence variety* is the set $V(I_{\mathcal{M}})$ of common zeros in $\mathbb{C}^D$ of the polynomials in $I_{\mathcal{M}}$.

Equivalently, $V(I_{\mathcal{M}})$ is the set of all $d_1 \times \cdots \times d_n$-tables with complex number entries which satisfy the conditional independence statements in $\mathcal{M}$. The variety $V(I_{\mathcal{M}})$ has three natural subsets:

- the subset of real tables, denoted $V_{\mathbb{R}}(I_{\mathcal{M}})$,
- the non-negative tables, denoted $V_{\geq}(I_{\mathcal{M}})$,
- the non-negative tables whose entries sum to one, $V_{\geq}(I_{\mathcal{M}} + \langle p - 1\rangle)$,

Here p denotes the sum of all unknowns $p_{u_1 \cdots u_n}$, so that $V_{\geq}(I_{\mathcal{M}} + \langle p - 1\rangle)$ is the subset of the probability simplex specified by the model $\mathcal{M}$.

We illustrate these definitions by analyzing the independence model

$$\mathcal{M} \;=\; \{\; 1 {\perp\!\!\!\perp} 2 \mid 3\,,\;\; 2 {\perp\!\!\!\perp} 3 \,\}$$

for $n = 3$ discrete random variables. Theorem 1 will be cited in Chapter 5 and it serves as a preview to Theorem 12. The ideal $I_{\mathcal{M}}$ lies in the polynomial ring $\mathbb{R}[D]$ in $d_1 d_2 d_3$ unknowns p_{ijk}. Its minimal generators are $\binom{d_1}{2}\binom{d_2}{2} d_3$ quadrics of the form $p_{ijk}\, p_{rsk} - p_{isk}\, p_{rjk}$ and $\binom{d_2}{2}\binom{d_3}{2}$ quadrics of the form $p_{+jk}\, p_{+st} - p_{+jt}\, p_{+sk}$. We change coordinates in $\mathbb{R}[D]$ by replacing each unknown p_{1jk} by $p_{+jk} = \sum_{i=1}^{d_1} p_{ijk}$. This coordinate change transforms $I_{\mathcal{M}}$ into a binomial ideal in $\mathbb{R}[D]$.

Theorem 1. *The ideal $I_{\mathcal{M}}$ has a Gröbner basis consisting of squarefree binomials of degree two, three and four, and it is hence radical. It has $2^{d_3}-1$ minimal primes, each generated by the 2×2-minors of a generic matrix.*

Proof. The minimal primes of $I_{\mathcal{M}}$ will be indexed by proper subsets of $[d_3]$. For each such subset σ we introduce the monomial prime

$$M_\sigma \quad = \quad \langle\, p_{+jk} \;\mid\; j \in [d_2],\, k \in \sigma \,\rangle,$$

and the complementary monomial

$$m_\sigma \quad = \quad \prod_{j=1}^{d_2} \prod_{k \in [d_3] \backslash \sigma} p_{+jk}\,,$$

and we define the ideal

$$P_\sigma \quad := \quad \big((I_{\mathcal{M}} \,+\, M_\sigma) : m_\sigma^\infty \big).$$

It follows from the general theory of binomial ideals [7] that P_σ is a binomial prime ideal. A closer look reveals that P_σ is minimally generated by the $d_2 \cdot |\sigma|$ variables in M_σ together with all the 2×2-minors of the following two-dimensional matrices: the matrix $(\, p_{ijk}\,)$ where the rows are indexed by $j \in [d_2]$ and the columns are indexed by pairs (i,k) with

$i \in \{+, 2, 3, \ldots, d_1\}$ and $k \in [d_3]\backslash\sigma$, and for each $k \in \sigma$, the matrices $(\, p_{ijk}\,)$ where the rows are indexed by $j \in [d_2]$ and the columns are indexed by $i \in \{2, 3, \ldots, d_1\}$.

We partition $V(I_{\mathcal{M}})$ into 2^{d_3} strata, each indexed by a subset σ of $[d_3]$. Namely, given a point (p_{ijk}) in $V(I_{\mathcal{M}})$ we define the subset σ of $[d_3]$ as the set of all indices k such that $(p_{+1k}, p_{+2k}, \ldots, p_{+d_2k})$ is the zero vector. Note that two tables (p_{ijk}) lie in the same stratum if and only if they give the same σ. The stratum indexed by σ is a dense subset in $V(P_\sigma)$. When $\sigma = [d_3]$ the stratum consists of all tables such that the line sums p_{+jk} are all zero, and for each fixed k, the remaining $(d_1 - 1) \times d_2$-matrix (p_{ijk}) with $i \geq 2$ has rank ≤ 1. So this locus is defined by the prime ideal $P_{[d_3]}$. Any point in this stratum satisfies the defining equations of P_σ for any proper subset σ. So the stratum indexed by $[d_3]$ lies in the closure of all other strata. But all remaining $2^{d_3} - 1$ strata have the property that no stratum lies in the closure of any other stratum, since the generic point of P_σ lies in exactly one stratum for any proper subset σ. Hence $V(I_{\mathcal{M}})$ is the irredundant union of the irreducible varieties $V(P_\sigma)$ where σ runs over all proper subsets of $[d_3]$. The second assertion in Theorem 1 now follows from Hilbert's Nullstellensatz.

To prove the first assertion, let us first note that $P_\emptyset$ is the prime ideal of 2×2-minors of the $d_2 \times (d_1 d_3)$-matrix (p_{ijk}) with rows indexed by $j \in [d_2]$ and columns indexed by pairs $(i, k) \in \{+, 2, 3, \ldots, d_1\} \times [d_3]$. Hence

$$P_\emptyset \quad = \quad \big(I_{\mathcal{M}} \, : \, m_\emptyset^\infty \big) \quad = \quad I_{2 \perp\!\!\!\perp \{1,3\}}. \tag{2.2}$$

It is well known (see e.g. [32, Proposition 5.4]) that the quadratic generators

$$p_{ijk}\, p_{rst} \quad - \quad p_{isk}\, p_{rjt} \tag{2.3}$$

form a reduced Gröbner basis for (2.2) with respect to the "diagonal term order". We modify this Gröbner basis to a Gröbner basis for $I_{\mathcal{M}}$ as follows:

- if $k = t$ take (2.3),
- if $i = +$ and $r = +$ take (2.3),
- if $i = +$ and $r \neq +$ and $k \neq t$ take (2.3) times p_{+jt} for any j,
- if $i \neq +$ and $r \neq +$ and $k \neq t$ take (2.3) times $p_{+jt}p_{+sk}$ for any j, s.

All of these binomials lie in $I_{\mathcal{M}}$ (this can be seen by taking S-pairs of the generators) and their S-pairs reduce to zero. By Buchberger's criterion, the given set of quadrics, cubics and quartics is a Gröbner basis, and the corresponding initial monomial ideal is square-free. This implies that $I_{\mathcal{M}}$ is radical (by [33, Proposition 5.3]), and the proof is complete. □

2.2 Contraction Axiom and Semigraphoids

The previous theorem can be regarded as an algebraic refinement of the following well-known rule for conditional independence ([22], [31, §2.2.2]).

Corollary 2. (Contraction Axiom) *If a probability distribution on* $[d_1] \times [d_2] \times [d_3]$ *satisfies* $1 \perp\!\!\!\perp 2 \,|\, 3$ *and* $2 \perp\!\!\!\perp 3$ *then it also satisfies* $2 \perp\!\!\!\perp \{1,3\}$.

Proof. The non-negative points satisfy $V_{\geq}(P_\sigma) \subseteq V_{\geq}(P_\emptyset)$, and this implies

$$V_{\geq}(I_{\mathcal{M}}) \quad = \quad V_{\geq}(I_{2 \perp\!\!\!\perp \{1,3\}}).$$

Intersecting with the probability simplex yields the assertion. □

Theorem 1 shows that the Contraction Axiom fails to hold when probabilities are replaced by negative real numbers. Any general point on $V(P_\sigma)$ for $\sigma \neq \emptyset$ satisfies $1 \perp\!\!\!\perp 2 \,|\, 3$ and $2 \perp\!\!\!\perp 3$ but it does not satisfy $2 \perp\!\!\!\perp \{1,3\}$.

The contraction axiom is one of four axioms that any probability distribution must satisfy. The complete list consists of the axioms of Symmetry (S), Decomposition (D), Weak Union (WU), and Contraction (C).

S: $X_i \perp\!\!\!\perp X_j \mid X_k$ if and only if $X_j \perp\!\!\!\perp X_i \mid X_k$.

D: If $X_i \perp\!\!\!\perp \{X_j, X_k\} \mid X_l$ then $X_i \perp\!\!\!\perp X_j \mid X_l$ and $X_i \perp\!\!\!\perp X_k \mid X_l$.

WU: If $X_i \perp\!\!\!\perp \{X_j, X_k\} \mid X_l$ then $X_i \perp\!\!\!\perp X_j \mid \{X_k, X_l\}$ and $X_i \perp\!\!\!\perp X_k \mid \{X_j, X_l\}$.

C: If $X_i \perp\!\!\!\perp X_j \mid \{X_k, X_l\}$ and $X_i \perp\!\!\!\perp X_k \mid X_l$ then $X_i \perp\!\!\!\perp \{X_j, X_k\} \mid X_l$

Given a model $\mathcal{M}$ defined by a list of conditional independence statements, we can consider the *semigraphoid* $\overline{\mathcal{M}}$ generated by $\mathcal{M}$, that is, the set of independence statements generated by $\mathcal{M}$ with respect to the previous four axioms. Note that $I_{\mathcal{M}} \neq I_{\overline{\mathcal{M}}}$ since the contraction axiom fails in general. Nevertheless, the contraction axiom is the only axiom that fails. This result follows directly from the definition of $I_{\mathcal{M}}$.

Chapter 3

Algebraic Representation of Bayesian Networks

3.1 Bayesian Networks

A *Bayesian network* is an acyclic directed graph G with vertices $X_1, \ldots, X_n$. The following notation and terminology is consistent with Lauritzen's book [20]. The *local Markov property* on G is the set of independence statements

$$\text{local}(G) \quad = \quad \{X_i \perp\!\!\!\perp \text{nd}(X_i) \mid \text{pa}(X_i) \,:\, i = 1, 2, \ldots, n\},$$

where $\text{pa}(X_i)$ denotes the set of parents of X_i in G and $\text{nd}(X_i)$ denotes the set of nondescendents of X_i in G. Here X_j is a *nondescendent* of X_i if there is no directed path from X_i to X_j in G. The *global Markov property*, $\text{global}(G)$, is the set of independence statements $A \perp\!\!\!\perp B \mid C$, for any triple A, B, C of subsets of vertices of G such that A and B are *d-separated* by C. Here two subsets A and B are said to be *d-separated* by C if all chains from A to B are blocked by C. A chain π from X_i to X_j in G is said to be *blocked* by a set C of nodes if it contains a vertex $X_b \in \pi$ such that either

- $X_b \in C$ and arrows of π do not meet head-to-head at X_b, or
- $X_b \notin C$ and X_b has no descendents in C, and arrows of π do meet head-to-head at X_b.

For any Bayesian network G, we have $\text{local}(G) \subseteq \text{global}(G)$, and this implies the following containment relations between ideals and varieties

$$I_{\text{local}(G)} \subseteq I_{\text{global}(G)} \qquad \text{and} \qquad V_{\text{local}(G)} \supseteq V_{\text{global}(G)}. \tag{3.1}$$

The latter inclusion extends to the three real varieties listed above, and we shall discuss when equality holds. First, however, we give an algebraic version of the description of Bayesian networks by recursive factorizations.

Moreover, given a directed, acyclic graph G, then $\overline{\mathrm{local}(G)} = \mathrm{global}(G)$, where $\overline{\mathrm{local}(G)}$ is the semigraphoid generated by $\mathrm{local}(G)$. The previous statement is a different way to state part of the so–called *Factorization Theorem* [20]:

Theorem 3. *Let G be a directed, acyclic graph. For a probability distribution P the following conditions are equivalent:*

DG: P obeys the directed global Markov property, relative to G.

DL: P obeys the directed local Markov property, relative to G.

Hence, we can rewrite the statement (3.1) as follows:

$$I_{\mathrm{local}(G)} \subseteq I_{\overline{\mathrm{local}(G)}} = I_{\mathrm{global}(G)},$$

where the inclusion is proper if and only if the contraction axiom fails for G.

3.2 Recursive Factorization

Consider the set of parents of the j-th node, $\mathrm{pa}(X_j) = \{X_{i_1}, \ldots, X_{i_r}\}$, and consider any event $X_j = u_0$ conditioned on $X_{i_1} = u_1, \ldots, X_{i_r} = u_r$, where $1 \leq u_0 \leq d_j, 1 \leq u_1 \leq d_{i_1}, \ldots, 1 \leq u_r \leq d_{i_r}$. We introduce an unknown $q^{(j)}_{u_0 u_1 \cdots u_r}$ to denote the conditional probability of this event, and we subject these unknowns to the linear relations $\sum_{v=1}^{d_j} q^{(j)}_{v u_1 \cdots u_r} = 1$ for all $1 \leq u_1 \leq d_{i_1}, \ldots, 1 \leq u_r \leq d_{i_r}$. Thus, we have introduced $(d_j - 1)d_{i_1} \cdots d_{i_r}$ unknowns for the vertex j. Let E denote the set of these unknowns $q^{(j)}_{u_0 u_1 \cdots u_r}$ for all $j \in \{1, \ldots, n\}$, and let $\mathbb{R}[E]$ denote the polynomial ring they generate.

If the n random variables are binary ($d_i = 2$ for all i) then the notation for $\mathbb{R}[E]$ can be simplified by dropping the first lower index and writing

$$q^{j}_{u_1 \cdots u_r} \quad := \quad q^{(j)}_{1 u_1 \cdots u_r} \quad = \quad 1 \; - \; q^{(j)}_{2 u_1 \cdots u_r}$$

In the binary case, $\mathbb{R}[E]$ is a polynomial ring in $\sum_{j=1}^{n} 2^{|\mathrm{pa}(X_j)|}$ unknowns.

The factorization of probability distributions according to G defines a polynomial map $\phi : \mathbb{R}^E \to \mathbb{R}^D$. By restricting to non-negative reals we get an induced map $\phi_{\geq 0}$. These maps are specified by the ring homomorphism $\Phi : \mathbb{R}[D] \to \mathbb{R}[E]$ which takes the unknown $p_{u_1 u_2 \cdots u_n}$ to the product of the expressions $q^{(j)}_{u_j u_{i_1} \cdots u_{i_r}}$ as j runs over $\{1, \ldots, n\}$. The image of ϕ lies in the independence variety $V_{\mathrm{global}(G)}$, or, equivalently, the independence ideal $I_{\mathrm{global}(G)}$ is contained in the prime ideal $\ker(\Phi)$. The *Factorization Theorem* for Bayesian networks [20, Theorem 3.27] states:

Theorem 4. *The following four subsets of the probability simplex coincide:*

$$\begin{aligned} V_{\geq}(I_{\mathrm{local}(G)} + \langle p-1\rangle) &= V_{\geq}(I_{\mathrm{global}(G)} + \langle p-1\rangle) \\ = \quad V_{\geq}(\ker(\Phi)) &= \quad \mathrm{image}(\phi_{\geq}). \end{aligned}$$

Example 5. Let G be the network on three binary random variables which has a single directed edge from 3 to 2. The parents and nondescendents are

$$\mathrm{pa}(1)=\emptyset, \mathrm{nd}(1)=\{2,3\},\ \mathrm{pa}(2)=\{3\}, \mathrm{nd}(2)=\{1\},\ \mathrm{pa}(3)=\emptyset, \mathrm{nd}(3)=\{1\}.$$

The resulting conditional independence statements are

$$\mathrm{local}(G) \;=\; \mathrm{global}(G) \;=\; \big\{1 \perp\!\!\!\perp 3,\ 1 \perp\!\!\!\perp 2 \mid 3,\ 1 \perp\!\!\!\perp \{2,3\}\big\}.$$

The ideal expressing the first two statements is contained in the ideal expressing the third statement, and we find that $I_{\mathrm{local}(G)} = I_{1 \perp\!\!\!\perp \{2,3\}}$ is the ideal generated by the six 2×2-subdeterminants of the 2×4-matrix

$$\begin{pmatrix} p_{111} & p_{112} & p_{121} & p_{122} \\ p_{211} & p_{212} & p_{221} & p_{222} \end{pmatrix} \tag{3.2}$$

This ideal is prime and its generators form a Gröbner basis. The Factorization Theorem is understood as follows for this example. We have $E = \{q^1, q_1^2, q_2^2, q^3\}$, and our ring map Φ takes the matrix (3.2) to

$$\begin{pmatrix} q^1 q_1^2 q^3 & q^1 q_2^2 (1-q^3) & q^1(1-q_1^2)q^3 & q^1(1-q_2^2)(1-q^3) \\ (1-q^1)q_1^2 q^3 & (1-q^1)q_2^2(1-q^3) & (1-q^1)(1-q_1^2)q^3 & (1-q^1)(1-q_2^2)(1-q^3) \end{pmatrix}$$

The map ϕ from $\mathbb{R}^4$ to $\mathbb{R}^8$ corresponding to the ring map $\Phi : \mathbb{R}[D] \to \mathbb{R}[E]$ gives a parametrization of all 2×4-matrices of rank 1 whose entries sum to 1. The Factorization Theorem for G is the same statement for non-negative matrices. The kernel of Φ is exactly equal to $I_{\mathrm{local}(G)} + \langle p-1\rangle$. □

Our aim is to decide to what extent the Factorization Theorem is valid over all real and all complex numbers. The corresponding algebraic question is to study the ideal $I_{\mathrm{local}(G)}$ and to determine its primary decomposition. Let us begin by considering all Bayesian networks on three random variables. We shall prove that for such small networks the ideal $I_{\mathrm{local}(G)}$ is always prime and coincides with the kernel of Φ. The following theorem is valid for arbitrary positive integers d_1, d_2, d_3. It is not restricted to the binary case.

Proposition 6. *For any Bayesian network G on three discrete random variables, the ideal $I_{\mathrm{local}(G)}$ is prime, and it has a quadratic Gröbner basis.*

Proof. We completely classify all possible cases. If G is the complete graph, directed acyclically, then local(G) contains no nontrivial independence statements, so $I_{\text{local}(G)}$ is the zero ideal. In what follows we always exclude this case. There are five isomorphism types of (non-complete) directed acyclic graphs on three nodes. They correspond to the rows of the following table:

Graph	Local/Global Markov property	Independence ideal
3 2 1	$1 \perp\!\!\!\perp \{2,3\},\ 2 \perp\!\!\!\perp \{1,3\},\ 3 \perp\!\!\!\perp \{1,2\}$	I_{Segre}
$3 \longrightarrow 2 \quad 1$	$1 \perp\!\!\!\perp 3,\ 1 \perp\!\!\!\perp 2 \mid 3,\ 1 \perp\!\!\!\perp \{2,3\}$	$I_{1 \perp\!\!\!\perp \{2,3\}}$
$3 \longrightarrow 2 \longrightarrow 1$	$1 \perp\!\!\!\perp 3 \mid 2$	$I_{1 \perp\!\!\!\perp 3 \mid 2}$
$1 \longleftarrow 3 \longrightarrow 2$	$1 \perp\!\!\!\perp 2 \mid 3$	$I_{1 \perp\!\!\!\perp 2 \mid 3}$
$3 \longrightarrow 1 \longleftarrow 2$	$2 \perp\!\!\!\perp 3$	$I_{2 \perp\!\!\!\perp 3}$

The third and fourth network represent the same independence model. In all cases except for the first, the ideal $I_{\text{local}(G)}$ is of the form $I_{A \perp\!\!\!\perp B \mid C}$, i.e., it is specified by a single independence statement. It was shown in [33, Lemma 8.2] that such ideals are prime. They are determinantal ideals and well known to possess a quadratic Gröbner basis. The only exceptional graph is the empty graph, which leads to the model of complete independence $1 \perp\!\!\!\perp \{2,3\},\ 2 \perp\!\!\!\perp \{1,3\},\ 3 \perp\!\!\!\perp \{1,2\}$. The corresponding ideal defines the Segre embedding of the product of three projective spaces $\mathbb{P}^{d_1-1} \times \mathbb{P}^{d_2-1} \times \mathbb{P}^{d_3-1}$ into $\mathbb{P}^{d_1 d_2 d_3 - 1}$. This ideal is prime and has a quadratic Gröbner basis. □

A network G is a *directed forest* if every node has at most one parent. The conclusion of Proposition 6 also holds for directed forests on any number of nodes. Proposition 17 will show that the direction of the edges is crucial: it is not sufficient to assume that the underlying undirected graph is a forest.

Theorem 7. *Let G be a directed forest. Then $I_{\text{global}(G)}$ is prime and has a quadratic Gröbner basis. These properties generally fail for $I_{\text{local}(G)}$.*

Proof. For a direct forest, the definition of a *blocked* chain reads as follows. A chain π from X_i to X_j in G is blocked by a set C if it contains a vertex $X_b \in \pi \cap C$. Hence, C d-separates A from B if and only if C separates A from B in the undirected graph underlying G. Thus, [12, Theorem 12] implies that $I_{\text{global}(G)}$ coincides with the distinguished prime ideal $\ker(\Phi)$, this ideal has a quadratic Gröbner basis. The second assertion is proved by the networks 18 and 26 in Table 5.1. See also [33, Example 8.8]. □

We close this chapter with a conjectured characterization of the global Markov property on a Bayesian network G in terms of commutative algebra.

Conjecture 8. $I_{\text{global}(G)}$ *is the ideal generated by all quadrics in* $\ker(\Phi)$.

Chapter 4

The Distinguished Component

4.1 Main Representation Theorem

In what follows we shall assume that every edge (i, j) of the Bayesian network G satisfies $i > j$. In particular, the node 1 is always a sink and the node n is always a source. For any integer $r \in [n]$ and $u_i \in [d_i]$ as before, we abbreviate the *marginalization* over the first r random variables as follows:

$$p_{++\cdots+u_{r+1}\cdots u_n} \quad := \quad \sum_{i_1=1}^{d_1}\sum_{i_2=1}^{d_2}\cdots\sum_{i_r=1}^{d_r} p_{i_1 i_2\cdots i_r u_{r+1}\cdots u_n}.$$

This is a linear form in our polynomial ring $\mathbb{R}[D]$. We denote by $\mathbf{p}$ the product of all of these linear forms. Thus the equation of $\mathbf{p} = 0$ defines a hyperplane arrangement in $\mathbb{R}^D$. We shall prove that the ideal $I_{\text{local}(G)}$ is prime locally outside this hyperplane arrangement, and hence so is $I_{\text{global}(G)}$. The following theorem provides the solution to [33, Problem 8.12].

Theorem 9. *The prime ideal* $\ker(\Phi)$ *is a minimal primary component of both of the ideals* $I_{\text{local}(G)}$ *and* $I_{\text{global}(G)}$*. More precisely,*

$$\left(I_{\text{local}(G)} : \mathbf{p}^\infty\right) \quad = \quad \left(I_{\text{global}(G)} : \mathbf{p}^\infty\right) \quad = \quad \ker(\Phi). \tag{4.1}$$

The prime ideal $\ker(\Phi)$ is called the *distinguished component*. It can be characterized as the set of all homogeneous polynomial functions on $\mathbb{R}^D$ which vanish on all probability distributions that factor according to G.

Proof. We relabel G so that $\text{pa}(1) = \{2, 3, \ldots, r\}$ and $\text{nd}(1) = \{r+1, \ldots, n\}$. Let A denote a set of $(d_1 - 1)d_2 \cdots d_r$ new unknowns $a_{i_1 i_2\cdots i_r}$, for $i_1 > 1$ defining a polynomial ring $\mathbb{R}[A]$. Define $d_2 \cdots d_r$ linear polynomials

$$a_{1\, i_2\cdots i_r} \quad = \quad 1 - \sum_{j=2}^{d_1} a_{j\, i_2\cdots i_r}.$$

Let Q denote a set of $d_2 \cdots d_n$ new unknowns $q_{i_2 \cdots i_r i_{r+1} \cdots i_n} = q_{i_2 \cdots i_n}$, defining a polynomial ring $\mathbb{R}[Q]$. We introduce the partial factorization map

$$\Psi : \mathbb{R}[D] \to \mathbb{R}[A \cup Q]\,,\ p_{i_1 i_2 \cdots i_n} \mapsto a_{i_1 \cdots i_r} \cdot q_{i_2 \cdots i_n}. \tag{4.2}$$

The kernel of Ψ is precisely the ideal $I_1 := I_{1 \coprod \mathrm{nd}(1)|\mathrm{pa}(1)}$. Note that

$$q_{i_2 \cdots i_n} = \Psi(p_{+ i_2 \cdots i_n}).$$

Therefore Ψ becomes an epimorphism if we localize $\mathbb{R}[D]$ at the product $\mathbf{p}_1$ of the $p_{+ i_2 \cdots i_n}$ and we localize R at the product of the $q_{i_2 \cdots i_n}$. This implies that any ideal L in the polynomial ring $\mathbb{R}[D]$ satisfies the identity

$$\Psi^{-1}(\Psi(L)) \quad = \quad \big((L + I_1) : \mathbf{p}_1^\infty\big). \tag{4.3}$$

Let G' denote the graph obtained from G by removing the sink 1 and all edges incident to 1. We regard $I_{\mathrm{local}(G')}$ as an ideal in $\mathbb{R}[Q]$. We modify the set of independence statements local(G) by removing 1 from the sets nd(i) for any $i \geq 2$. Let $J \subset \mathbb{R}[D]$ be the ideal corresponding to these modified independence statements, so that $\Psi(J) = I_{\mathrm{local}(G')}$. Note that

$$J + I_1 \subseteq I_{\mathrm{local}(G)} \subseteq I_{\mathrm{global}(G)} \subseteq \ker(\Phi),$$

so it suffices to show that $(J + I_1) : \mathbf{p}^\infty = \ker(\Phi)$. The map Φ factors as

$$\mathbb{R}[D] \stackrel{\Psi}{\longrightarrow} \mathbb{R}[A \cup Q] \stackrel{\Phi'}{\longrightarrow} \mathbb{R}[A \cup E'] = \mathbb{R}[E], \tag{4.4}$$

where Φ' is the factorization map coming from the graph G', extended to be the identity on the variables A. By induction on the number of vertices, we may assume that Theorem 9 holds for the smaller graph G', i.e.,

$$\ker(\Phi') \quad = \quad (I_{\mathrm{local}(G')} : \mathbf{q}_2^\infty) \quad = \quad \Psi(J : \mathbf{p}_2^\infty), \tag{4.5}$$

where $\mathbf{q}_2 = \Psi(\mathbf{p}_2)$ and $\mathbf{p}_2$ is the product of the linear forms $p_{++\cdots+u_i \cdots u_n}$ with at least two initial $+$'s. Therefore

$$\ker(\Phi) \quad = \quad \Psi^{-1}(\Psi(J : \mathbf{p}_2^\infty)). \tag{4.6}$$

Applying (4.3), we get $\ker(\Phi) = ((J : \mathbf{p}_2^\infty) + I_1) : \mathbf{p_1}^\infty = (J + I_1) : \mathbf{p}^\infty$. □

By following the technique of the proof, we can replace $\mathbf{p}_1$ by the product of a much smaller number of $p_{+u_2 \cdots u_n}$. In fact, we need only take the linear forms $p_{+u_2 \cdots u_r 11 \cdots 1}$. Hence, by induction, $\mathbf{p}$ can be replaced by a much smaller product of linear forms. This observation proved to be crucial for computing some of the tough primary decompositions in Chapter 6.

4.2 Factorization Theorem

As a corollary to the previous theorem we derive an algebraic proof of the Factorization Theorem.

Proof of Theorem 4: We use induction on the number of nodes to show that every point in $V_{\geq}(I_{\mathrm{local}(G)} + \langle p-1\rangle)$ also lies in image$(\phi_{\geq})$. Such a point is a homomorphism $\tau : \mathbb{R}[D] \to \mathbb{R}$ with the property that τ is zero on $I_{\mathrm{local}(G)}$, and its values on the indeterminates $p_{u_1\cdots u_n}$ are non-negative and sum to 1. The map τ can be extended to a homomorphism $\tau' : \mathbb{R}[Q\cup A] \to \mathbb{R}$ as follows. We first set $\tau'(q_{i_2\cdots i_n}) = \tau(p_{+i_2\cdots i_n})$. If that real number is positive then we set $\tau'(a_{i_1\cdots i_r}) = \tau(p_{i_1 i_2\cdots i_n})/\tau(p_{+i_2\cdots i_n})$, and otherwise we set $\tau'(a_{i_1\cdots i_r}) = 0$. Our non-negativity hypothesis implies that τ coincides with the composition of τ' and Ψ, i.e., the point τ is the image of τ' under the induced map $\mathbb{R}^{A\cup Q} \to \mathbb{R}^D$. The conclusion now follows by induction. □

We close this chapter by presenting our solution to [33, Problem 8.11].

Proposition 10. *There exists a Bayesian network G on five binary random variables such that the local Markov ideal $I_{\mathrm{local}}(G)$ is not radical.*

Proof. Let G be the complete bipartite network $K_{2,3}$ with nodes $\{1,5\}$ and $\{2,3,4\}$ and directed edges $(5,2)$, $(5,3)$, $(5,4), (2,1)$, $(3,1)$, $(4,1)$. Then

$$\mathrm{local}(G) \;=\; \big\{\, 1 \perp\!\!\!\perp 5 \,|\, \{2,3,4\},\; 2 \perp\!\!\!\perp \{3,4\} \,|\, 5,\; 3 \perp\!\!\!\perp \{2,4\} \,|\, 5,\; 4 \perp\!\!\!\perp \{2,3\} \,|\, 5 \,\big\}.$$

The polynomial ring $\mathbb{R}[E]$ has 32 indeterminates $p_{11111}, p_{11112}, \ldots, p_{22222}$. The ideal $I_{\mathrm{local}}(G)$ is minimally generated by eight binomial quadrics

$$p_{1u_2u_3u_41} \cdot p_{2u_2u_3u_42} \;-\; p_{1u_2u_3u_42} \cdot p_{2u_2u_3u_41}, \qquad u_2, u_3, u_4 \in \{1,2\},$$

and eighteen non-binomial quadrics

$$\begin{gathered}
p_{+122u_5} \cdot p_{+221u_5} - p_{+121u_5} \cdot p_{+222u_5},\; p_{+212u_5} \cdot p_{+221u_5} - p_{+211u_5} \cdot p_{+222u_5},\\
p_{+112u_5} \cdot p_{+221u_5} - p_{+111u_5} \cdot p_{+222u_5},\; p_{+122u_5} \cdot p_{+212u_5} - p_{+112u_5} \cdot p_{+222u_5},\\
p_{+121u_5} \cdot p_{+212u_5} - p_{+111u_5} \cdot p_{+222u_5},\; p_{+122u_5} \cdot p_{+211u_5} - p_{+111u_5} \cdot p_{+222u_5},\\
p_{+112u_5} \cdot p_{+211u_5} - p_{+111u_5} \cdot p_{+212u_5},\; p_{+121u_5} \cdot p_{+211u_5} - p_{+111u_5} \cdot p_{+221u_5},\\
p_{+112u_5} \cdot p_{+121u_5} - p_{+111u_5} \cdot p_{+122u_5}, \qquad u_5 \in \{1,2\}.
\end{gathered}$$

These nine equations (for fixed value of u_5) define the Segre embedding of $\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1$ in $\mathbb{P}^7$, as in [33, eqn. (8.6), page 103]. Consider the polynomial

$$f \;=\; p_{+1112}p_{+2222}(p_{12221}p_{12212}p_{12122}p_{12111} - p_{12112}p_{12121}p_{12211}p_{12222}).$$

By computing a Gröbner basis, it can be checked that f^2 lies in $I_{\mathrm{local}(G)}$ but f does not lie in $I_{\mathrm{local}(G)}$. Hence $I_{\mathrm{local}(G)}$ is not a radical ideal. The primary decomposition of this ideal will be described in Example 19. □

Chapter 5

Networks on Four Random Variables

5.1 Binary Random Variables

In this chapter we present the algebraic classification of all Bayesian networks on four random variables. In the binary case we have the following result.

Theorem 11. *The local and global Markov ideals of all Bayesian networks on four binary variables are radical. The hypothesis "binary" is essential.*

Thus the solution [33, Problem 8.11] is affirmative for networks on four binary nodes. Proposition 10 shows that the hypothesis "four" is essential. Theorem 11 is proved by exhaustive computations in `Macaulay2`. We summarize the results in Table 5.1. Each row represents one network G on four binary random variables along with some information about its two ideals

$$I_{\mathrm{local}(G)} \;\subseteq\; I_{\mathrm{global}(G)} \;\subseteq\; \mathbb{R}[p_{1111}, p_{1112}, \ldots, p_{2221}, p_{2222}].$$

Here G is represented by the list of sets of children $(\mathrm{ch}(1), \mathrm{ch}(2), \mathrm{ch}(3), \mathrm{ch}(4))$. The information given in the second column corresponds to the *codimension, degree, and number of minimal generators* of the ideal $I_{\mathrm{local}(G)}$. For example, the network in the fourth row has four directed edges $(2,1), (3,1), (4,1)$ and $(4,2)$. Here $I_{\mathrm{local}(G)} = I_{\mathrm{global}(G)} = \ker(\Phi)$. This prime has codimension 3, degree 4 and is generated by the six 2×2-minors of the 2×4-matrix

$$\begin{pmatrix} p_{+111} & p_{+112} & p_{+211} & p_{+212} \\ p_{+121} & p_{+122} & p_{+221} & p_{+222} \end{pmatrix}.$$

Of the 30 local Markov ideals in Table 5.1 all but six are prime. The remaining six ideals are all radical, and the number of their minimal primes is listed. Hence all local Markov ideals are radical. The last column corresponds to the ideal $I_{\mathrm{global}(G)}$. This ideal is equal to the distinguished component for all but two networks, namely `15` and `17`. For these two networks we have $I_{\mathrm{local}(G)} = I_{\mathrm{global}(G)}$. This proves the first assertion of Theorem 11.

Table 5.1: All Bayesian Networks on Four Binary Random Variables

Index	Information	Network	Local	Global
1	1, 2, 1	$\{\}, \{1\}, \{1,2\}, \{1,2\}$	prime	
2	2, 4, 2	$\{\}, \{1\}, \{1\}, \{1,2,3\}$	prime	
3	2, 4, 2	$\{\}, \{1\}, \{1,2\}, \{1,3\}$	prime	
4	3, 4, 6	$\{\}, \{1\}, \{1\}, \{1,2\}$	prime	
5	4, 6, 9	$\{\}, \{1\}, \{1\}, \{1\}$	prime	
6	4, 16, 4	$\{\}, \{\}, \{1,2\}, \{1,2,3\}$	prime	
7	4, 16, 4	$\{\}, \{1\}, \{1,2\}, \{2,3\}$	prime	
8	4, 16, 4	$\{\}, \{1\}, \{2\}, \{1,2,3\}$	prime	
9	5, 32, 5	$\{\}, \{\}, \{1,2\}, \{1,2\}$	prime	
10	5, 32, 5	$\{\}, \{1\}, \{1,2\}, \{2\}$	prime	
11	6, 8, 10	$\{\}, \{1\}, \{1\}, \{2\}$	radical, 5 comp.	prime
12	6, 16, 12	$\{\}, \{\}, \{1\}, \{1,2,3\}$	prime	
13	6, 16, 12	$\{\}, \{\}, \{1,2\}, \{2,3\}$	prime	
14	6, 16, 12	$\{\}, \{1\}, \{2\}, \{2,3\}$	prime	
15	6, 64, 6	$\{\}, \{1\}, \{1\}, \{2,3\}$	radical, 5 comp.	radical
16	6, 64, 6	$\{\}, \{1\}, \{1,2\}, \{3\}$	radical, 9 comp.	prime
17	6, 64, 6	$\{\}, \{1\}, \{2\}, \{1,3\}$	radical, 5 comp.	radical
18	7, 8, 14	$\{\}, \{1\}, \{2\}, \{3\}$	radical, 3 comp.	prime
19	7, 8, 28	$\{\}, \{\}, \{1\}, \{1,3\}$	prime	
20	7, 24, 16	$\{\}, \{\}, \{1\}, \{1,2\}$	prime	
21	7, 32, 13	$\{\}, \{1\}, \{2\}, \{2\}$	prime	
22	8, 14, 31	$\{\}, \{\}, \{1\}, \{1\}$	prime	
23	8, 34, 20	$\{\}, \{\}, \{1\}, \{2,3\}$	prime	
24	8, 36, 18	$\{\}, \{\}, \{\}, \{1,2,3\}$	prime	
25	8, 36, 18	$\{\}, \{\}, \{1,2\}, \{3\}$	prime	
26	9, 20, 27	$\{\}, \{\}, \{1\}, \{2\}$	radical, 5 comp.	prime
27	9, 24, 34	$\{\}, \{\}, \{\}, \{1,2\}$	prime	
28	9, 24, 34	$\{\}, \{\}, \{1\}, \{3\}$	prime	
29	10, 20, 46	$\{\}, \{\}, \{\}, \{1\}$	prime	
30	11, 24, 55	$\{\}, \{\}, \{\}, \{\}$	prime	

5.2 General Random Variables

The main point of this chapter is the second sentence in Theorem 11. Embedded components can appear when the number of levels increases. In the next theorem we let d_1, d_2, d_3 and d_4 be arbitrary positive integers.

Theorem 12. *Of the* 30 *local Markov ideals on four random variables,* 22 *are always prime, five are not prime but always radical (numbers* 10,11,16, 18,26 *in Table 5.1) and three are not radical (numbers* 15,17,21 *in Table 5.1).*

Proof. We prove this theorem by an exhaustive case analysis of all thirty networks. In most cases, the ideal $I_{\text{local}(G)}$ can be made binomial by a suitable coordinate change, just like in the proof of Theorem 1. In fact, let us start with a non-trivial case which is immediately taken care of by Theorem 1.

The network 16*:* Here we have $\text{local}(G) = \big\{1 \perp\!\!\!\perp 4 \,|\, \{2,3\},\, 2 \perp\!\!\!\perp 4 \,|\, 3\big\}$. For fixed value of the third node we get the model $\{1 \perp\!\!\!\perp 4 \,|\, 2,\, 4 \perp\!\!\!\perp 2\}$ whose ideal was shown to be radical in Theorem 1. Hence $I_{\text{local}(G)}$ is the ideal generated by d_3 copies of this radical ideal in disjoint sets of variables. We conclude that $I_{\text{local}(G)}$ is radical and has $(2^{d_2} - 1)^{d_3}$ minimal primes.

The networks 1, 2, 3, 4, 6, 7, 8, 12, 13, 14*:* In each of these ten cases, the ideal $I_{\text{local}(G)}$ is generated by quadratic polynomials corresponding to a single conditional independence statement. This observation implies that $I_{\text{local}(G)}$ is a prime ideal, by [33, Lemma 8.2].

The network 5*:* Here $\text{local}(G)$ specifies the model of complete independence for the random variables X_2, X_3 and X_4. This means that $I_{\text{local}(G)}$ is the ideal of a Segre variety, which is prime and has a quadratic Gröbner basis.

The networks 24 *and* 25*:* Each of these two networks describes the join of d_4 and d_3 Segre varieties. The same reasoning as in case 5 applies.

The network 23*:* Observe that $I_{\text{local}(G)} = I_{\text{global}(G)} = I_{1 \perp\!\!\!\perp \{2,4\}|3} + I_{2 \perp\!\!\!\perp \{1,3\}|4}$. Since G is a directed tree, Theorem 7 implies that $I_{\text{global}(G)}$ coincides with the distinguished prime ideal $\ker(\Phi)$. Therefore, $I_{\text{local}(G)}$ is always prime.

The networks 19, 22, 27, 28, 29, 30*:* Each of these six networks has an isolated vertex. This means that $I_{\text{local}(G)}$ is the ideal of the Segre embedding of the product of two smaller varieties namely, the projective space $\mathbb{P}^{d_i-1}$ corresponding to the isolated vertex i and the scheme specified by the local ideal of the remaining network on three nodes. The latter ideal is prime and has a quadratic Gröbner basis, by Proposition 6, and hence so is $I_{\text{local}(G)}$.

The network 20*:* The ideal $I_{\text{local}(G)}$ is binomial in the coordinates p_{ijkl} with $i \in \{+, 2, \ldots, d_1\}$. Generators are $p_{i_1j_2kl}p_{i_2j_1kl} - p_{i_1j_1kl}p_{i_2j_2kl}$, $p_{i_1j_2k_1l}p_{i_2j_1k_2l} - p_{i_1j_1k_1l}p_{i_2j_2k_2l}$, and $p_{+j_1k_2l_1}p_{+j_2k_1l_2} - p_{+j_1k_1l_1}p_{+j_2k_2l_2}$. The S-pairs within each group reduce to zero by the Gröbner basis property of the 2×2-minors of a generic matrix. It can be checked easily that the crosswise reverse lexicographic S-pairs also reduce to zero. We conclude that the given set of irreducible quadrics is a reverse lexicographic Gröbner basis. In view of [32, Lemma 12.1], the lowest variable is not a zero-divisor, and hence by symmetry none of the variables p_{ijkl} is zero-divisor. It now follows from equation (4.1) in Theorem 9 that $I_{\text{local}(G)}$ coincides with the prime ideal $\ker(\Phi)$.

The network `9`*:* The ideal $I_{\text{local}(G)}$ is generated by the quadratic polynomials $p_{i_1j_2kl}p_{i_2j_1kl} - p_{i_1j_1kl}p_{i_2j_2kl}$, $p_{++k_1l_2}p_{++k_2l_1} - p_{++k_1l_1}p_{++k_2l_2}$. These generators form a Gröbner basis in the reverse lexicographic order. Indeed, assuming that $i_1 < i_2$, $j_1 < j_2$, $k_1 < k_2$, $l_1 < l_2$, the leading terms are $p_{i_1j_2kl}p_{i_2j_1kl}$ and $p_{11k_1l_2}p_{11k_2l_1}$. Hence no leading term from the first group of quadrics shares a variable with a leading term from the second group. Hence the crosswise S-pairs reduce to zero by [4, Prop. 4, §2.9]. The S-pairs within each group also reduce to zero by the Gröbner basis property of the 2×2-minors of a generic matrix. Hence the generators are a Gröbner basis. Since the leading terms are square-free, we see that the ideal is radical. An argument similar to the previous case shows that $I_{\text{local}(G)}$ is prime.

The network `18`*:* Here G is a directed chain of length four. We claim that $I_{\text{local}(G)}$ is the irredundant intersection of $2^{d_2} - 1$ primes, and it has a Gröbner basis consisting of square-free binomials of degree two, three and four. We give an outline of the proof. We first turn $I_{\text{local}(G)}$ into a binomial ideal by taking the coordinates to be p_{ijkl} with $i \in \{+, 2, 3, \ldots, d_1\}$. The minimal primes are indexed by proper subsets of $[d_2]$. For each such subset σ we introduce the monomial prime $M_\sigma = \langle p_{+jkl} : j \in \sigma, k \in [d_3], l \in [d_4]\rangle$ and the complementary monomial $m_\sigma = \prod_{j\in[d_2]\setminus\sigma} \prod_{k\in[d_3]} \prod_{l\in[d_4]} p_{+jkl}$, and we define the ideal $P_\sigma = \big((I_{\text{local}(G)} + M_\sigma) : m_\sigma^\infty\big)$. These ideals are prime, and the union of their varieties is irredundant and equals the variety of $I_{\text{local}(G)}$. Using Buchberger's S-pair criterion, we check that the following four types of square-free binomials are a Gröbner basis:

- the generators $p_{i_1jk_1l_1}p_{i_2jk_2l_2} - p_{i_1jk_2l_2}p_{i_1jk_2l_2}$ encoding $1 \perp\!\!\!\perp \{3, 4\} \mid 2$,
- the generators $p_{+j_1kl_1}p_{+j_2kl_2} - p_{+j_1kl_2}p_{+j_2kl_1}$ encoding $2 \perp\!\!\!\perp 4 \mid 3$,
- the cubics $\big(p_{+j_1kl_1}p_{ij_2kl_2} - p_{+j_1kl_2}p_{ij_2kl_1}\big) \cdot p_{+j_2k_3l_3}$,
- the quartics $\big(p_{i_1j_1kl_1}p_{i_2j_2kl_2} - p_{i_1j_1kl_2}p_{i_2j_2kl_1}\big) \cdot p_{+j_1l_3k_3} \cdot p_{+j_2l_4k_4}$.

The network `10`*:* The ideal $I_{\text{local}(G)}$ is generated by

$$p_{i_1jkl_2}p_{i_2jkl_1} - p_{i_1jkl_1}p_{i_2jkl_2} \qquad \text{and} \qquad p_{++k_1l_2}p_{++k_2l_1} - p_{++k_1l_1}p_{++k_2l_2}.$$

In general, this ideal is not prime, but it is always radical. If $d_4 = 2$ then the ideal is always prime, If $d_4 > 2$, $I_{\text{local}(G)}$ is the intersection of the distingushed component and 2^{d_3-1} prime ideals indexed by all proper subsets $\sigma \subset [d_3]$ as in the previous network.

The network `11`*:* Here, $\text{local}(G) = \big\{1 \perp\!\!\!\perp 4 \mid \{2, 3\},\ 2 \perp\!\!\!\perp 3 \mid 4,\ 3 \perp\!\!\!\perp \{2, 4\}\big\}$. The ideal $I_{\text{local}(G)}$ is binomial in the coordinates p_{ijkl} with $i \in \{+, 2, \ldots, d_1\}$. It is generated by the binomials $p_{i_1jkl_1}p_{i_2jkl_2} - p_{i_1jkl_2}p_{i_2jkl_1}$, $p_{+j_1k_1l_1}p_{+j_2k_2l_2} - p_{+j_1k_2l_1}p_{+j_2k_1l_2}$ encoding the first and third independent statements. The minimal primes are indexed by pairs of proper subsets of $[d_2]$ and $[d_3]$. For each such pair of subsets (σ, τ) we introduce the monomial prime $M_{(\sigma,\tau)} = \langle p_{+jkl} : j \in \sigma, k \in \tau, l \in [d_4]\rangle$ and the complementary monomial $m_{(\sigma,\tau)} = \prod_{j\in[d_2]\setminus\sigma} \prod_{k\in[d_3]\setminus\tau} \prod_{l\in[d_4]} p_{+jkl}$, and we define the ideal $P_{(\sigma,\tau)} = \big((I_{\mathrm{local}(G)} + M_{(\sigma,\tau)}) : m^{\infty}_{(\sigma,\tau)}\big)$. These ideals are prime, and the union of their varieties equals the variety of $I_{\mathrm{local}(G)}$. Moreover, the ideal $I_{\mathrm{local}(G)}$ is equal to the intersection of the minimal primes which are indexed by the following pairs: For each proper $\tau \subset [d_3]$ the pair $(\emptyset, \tau)$, and for each nonempty proper $\sigma \subset [d_2]$ the pairs (σ, τ) where $\tau \subset [d_3]$ is any subset of cardinality at most $d_3 - 2$. In particular, for $d_2 = d_3 = 3$, and arbitrary d_1, d_4, the ideal $I_{\mathrm{local}(G)}$ has 31 prime components. For $d_2 = 2, d_3 = 4$, $I_{\mathrm{local}(G)}$ has 37 prime components, and for $d_2 = 4, d_3 = 2$, $I_{\mathrm{local}(G)}$ has 17 prime components.

The network `26`*:* The ideal $I_{\mathrm{local}(G)}$ is a radical ideal. The minimal primes are indexed by all pairs of proper subsets of $[d_3]$ and $[d_4]$. For each such pair (σ, τ) we introduce the monomial primes $M_\sigma = \langle p_{+jkl} : k \in \sigma, j \in [d_2], l \in [d_4]\rangle$, $M_\tau = \langle p_{i+kl} : l \in \tau, i \in [d_1], k \in [d_3]\rangle$, and $M_{(\sigma,\tau)} = M_\sigma + M_\tau$. Just as before, we introduce the complementary monomial $m_{(\sigma,\tau)}$, and the ideal $P_{(\sigma,\tau)} = \big((I_{\mathrm{local}(G)} + M_{(\sigma,\tau)}) : m^{\infty}_{(\sigma,\tau)}\big)$. The ideal $I_{\mathrm{local}(G)}$ is equal to the intersection of all these prime ideals.

The network `21`*:* Here, $\mathrm{local}(G) = \big\{1 \perp\!\!\!\perp \{3,4\} \,|\, 2,\ 3 \perp\!\!\!\perp 4\big\}$. The ideal $I_{\mathrm{local}(G)}$ is generated by the binomials $p_{i_1jk_2l_2}p_{i_2jk_1l_1} - p_{i_1jk_1l_1}p_{i_2jk_2l_2}$, and the polynomials $p_{++k_1l_2}p_{++k_2l_1} - p_{++k_1l_1}p_{++k_2l_2}$. This ideal is not radical, in general. The first counterexample occurs for the case $d_1 = d_2 = d_3 = 2$ and $d_4 = 3$. Here $I_{\mathrm{local}(G)}$ is generated by 33 quadratic polynomials in 24 unknowns. The degree reverse lexicographic Gröbner basis of this ideal consists of 123 polynomials of degree up to 8. In this case, $I_{\mathrm{local}(G)}$ is the intersection of the distinguished component and the P-primary ideal $Q = I_{1 \perp\!\!\!\perp \{3,4\} \,|\, 2} + P^2$, where P is the prime ideal generated by the 12 linear forms p_{+jkl}.

The networks `15` *and* `17`*:* Here, after relabeling network `17`,

$$\mathrm{local}(G) = \big\{1 \perp\!\!\!\perp 4 \,|\, \{2,3\},\ 2 \perp\!\!\!\perp 3 \,|\, 4\big\}.$$

The ideal $I_{\mathrm{local}(G)}$ is binomial in the coordinates p_{ijkl} with $i \in \{+, 2, \ldots, d_1\}$. It is generated by the binomials $p_{i_1jkl_1}p_{i_2jkl_2} - p_{i_1jkl_2}p_{i_2jkl_1}$, $p_{+j_1k_1l}p_{+j_2k_2l} - p_{+j_1k_2l}p_{+j_2k_1l}$. This ideal is not radical, in general. The first counterexample occurs for the case $d_1 = 2$ and $d_2 = d_3 = d_4 = 3$. Here $I_{\mathrm{local}(G)}$ is generated by 54 quadratic binomials in 54 unknowns. The reverse lexicographic Gröbner basis consists of 13,038 binomials of degree up to 14. One of the elements in the Gröbner basis is

$$p_{+111}p_{+223}(p_{+331})^2 \cdot \big(p_{2122}p_{2133}p_{2323}p_{2332} - p_{2333}p_{2322}p_{2132}p_{2123}\big).$$

Removing the square from the third factor, we obtain a polynomial f of degree 7 such that that $f \notin I$ but $f^2 \in I$. This proves that I is not radical. The number of minimal primes of $I_{\mathrm{local}(G)}$ is equal to $2^{d_2} + 2^{d_3} - 3$. □

In the 22 cases where I_{local} is prime, it follows from Theorem 9 that the global Markov ideal I_{global} is prime as well. Among the remaining cases, we have $I_{\mathrm{local}(G)} = I_{\mathrm{global}(G)}$ for networks `10`, `15`, `17`, `21`, and we have $I_{\mathrm{local}} \neq I_{\mathrm{global}} = \ker(\Phi)$ for networks `11`, `16`, `18`, `26`. This discussion implies:

Corollary 13. *Of the* 30 *global Markov ideals on four random variables,* 26 *are always prime, one is not prime but always radical (number* `10` *in Table 5.1) and three are not radical (numbers* `15`,`17`,`21` *in Table 5.1).*

It is instructive to examine the distinguished prime ideal $P = \ker(\Phi)$ in the last case `15`, `17`. Assume for simplicity that $d_1 = 2$ but d_2, d_3 and d_4 are arbitrary positive integers. We rename the unknowns $x_{jkl} = p_{2jkl}$ and $y_{jkl} = p_{+jkl}$. Then we can take Φ to be the following monomial map:

$$\mathbb{R}[x_{jkl}, y_{jkl}] \to \mathbb{R}[u_{jk}, v_{jl}, w_{kl}],\ x_{jkl} \mapsto u_{jk}v_{jl}w_{kl},\ y_{jkl} \mapsto v_{jl}w_{kl}, \tag{5.1}$$

For example, for $d_2 = d_3 = 3$ and $d_4 = 2$, the ideal $P = \ker(\Phi)$ has 361 minimal generators, of degrees ranging from two to seven. One generator is

$$x_{111}x_{132}x_{222}x_{312}x_{321}y_{221}y_{331} - x_{112}x_{131}x_{221}x_{311}x_{322}y_{232}y_{321}.$$

Among the 361 minimal generators, there are precisely 15 which do not contain any variable y_{ijk}, namely, there are nine quartics and six sextics like

$$x_{112}x_{121}x_{211}x_{232}x_{322}x_{331} - x_{111}x_{122}x_{212}x_{231}x_{321}x_{332}.$$

These 15 generators form the Markov basis for the $3 \times 3 \times 2$-tables in the no-three-way interaction model. See [32, Corollary 14.12] for a discussion.

The ideal for the no-three-way interaction model of $d_2 \times d_3 \times d_4$-tables always coincides with the elimination ideal $P \cap \mathbb{R}[x_{ijk}]$ and, moreover, every generating set of P contains a generating set for $P \cap \mathbb{R}[x_{ijk}]$. In view of [32, Proposition 14.14], this shows that the maximal degree among minimal generators of P exceeds any bound as d_2, d_3, d_4 increases. In practical terms, it is hard to compute these generators even for $d_2 = d_3 = d_4 = 4$. We refer to the web page `http://math.berkeley.edu/~seths/ccachallenge.html`.

Chapter 6

Networks on Five Binary Random Variables

6.1 Global Markov Ideals

In this chapter we discuss the global Markov ideals of all Bayesian networks on five binary random variables. In each case we computed the primary decomposition. In general, the built-in primary decomposition algorithms in current computer algebra systems cannot compute the primary decompositions of most of these ideals. In the Appendix, we outline some techniques that allowed us to compute these decompositions. The primary decompositions of the local Markov ideals of these networks could also be computed, but they have less regular structure and are in general more complicated.

There are 301 distinct non-complete networks on five random variables, up to isomorphism of directed graphs. We have placed descriptions of these networks and of the primary decompositions of their global Markov ideals on the website

`http://math.cornell.edu/~mike/bayes/global5.html`.

In this chapter, we refer to the graphs as $G_0, G_1, \ldots, G_{300}$, the indices matching the information on the website. We summarize our results in a theorem.

Theorem 14. *Of the* 301 *global Markov ideals on five binary random variables,* 220 *are prime,* 68 *are radical but not prime, and* 13 *are not radical.*

Proof. The proof is via direct computation with each of these ideals in `Macaulay2`. Some of these require little or no computation: if G is a directed forest, or if there is only one independence statement, then the ideal is prime. Others require substantial computation and some ingenuity to find the primary decomposition. Results are posted at the website cited above.

To prove primality, it suffices to compute the ideal quotient of $I = I_{\mathrm{global}(G)}$ with respect to a small subset of the $p_{+++u_r\cdots u_n}$. Alternatively, one may birationally project I by eliminating variables, as in Proposition 25. In either case, if a zero divisor x is found, the ideal is not prime. If some ideal quotient satisfies $(I : x^2) \neq (I : x)$, then I is not radical. □

The numbers of prime components of the 288 radical global Markov ideals range from 1 to 39. The distribution is given in the following table:

# of components	1	3	5	7	17	25	29	33	39
# of ideals	220	8	41	3	9	1	2	3	1

Theorem 15. *Conjecture 8 is true for Bayesian networks G on five binary random variables. In each of the* 301 *cases, the distinguished prime ideal* $\ker(\Phi)$ *is generated by homogeneous polynomials of degree at most eight.*

Proof. We compute the distinguished component from $I_{\mathrm{global}(G)}$ by saturation, and we check the result by using the techniques in the Appendix. The computation of the distinguished component of the 81 non-prime examples yields that 64 of these ideals are generated in degrees ≤ 4, twelve are generated in degrees ≤ 6, and five are generated in degrees ≤ 8. □

Theorem 9 says that we can decide primality or find the distinguished component of $I_{\mathrm{global}(G)}$ by inverting each of the $p_{+++u_i\cdots u_n}$. With some care, it is possible to reduce this to a smaller set. Still, the following is unexpected.

Proposition 16. *For all but two networks on five binary random variables,* p_{+1111} *is a non-zero divisor on* $I = I_{\mathrm{global}(G)}$ *if and only if* I *is prime. In all but these two examples,* I *is radical if and only if* $(I : p_{+1111}^2) = (I : p_{+1111})$.

Proof. The networks which do not satisfy the given property are

$$G_{201} = \big(\{\}, \{1\}, \{1,2\}, \{1,2\}, \{3,4\}\big) \qquad \text{and} \qquad G_{214} = \big(\{\}, \{1\}, \{1,2\}, \{3\}, \{1,2,4\}\big).$$

After permuting the nodes 4, 5, both the local and global independence statements of G_{214} are the same as those for G_{201}. The global independence statements for G_{201} are $\big\{\{1,2\} \perp\!\!\!\perp 5 \,|\, \{3,4\},\ 3 \perp\!\!\!\perp 4 \,|\, 5\big\}$. The primary decomposition for the radical ideal $I = I_{\mathrm{global}(G_{201})}$ is

$$I \;=\; \ker(\Phi) \cap (I + P_{++1\bullet\bullet}) \cap (I + P_{++2\bullet\bullet}) \cap (I + P_{++\bullet 1\bullet}) \cap (I + P_{++\bullet 2\bullet}),$$

where $\ker(\Phi)$ is the distinguished prime component,

$$P_{++1\bullet\bullet} \quad = \quad \langle p_{++111}, p_{++112}, p_{++121}, p_{++122} \rangle,$$

and the other three components are defined in an analogous manner. Therefore, p_{+1111} is a non-zero divisor modulo I. By examining all 81 non-prime ideals, we see that all except these two have a minimal prime containing p_{+1111}. The final statement also follows from direct computation. □

We have searched for conditions on the network which would characterize under what conditions the global Markov ideal is prime, or fails to be prime. Theorem 7 states that if the network is a directed forest, then the global Markov ideal is prime. Two possible conditions, the first for primality, and the second for non-primality, are close, but not quite right. We present them, with their counterexamples, in the following two propositions.

Proposition 17. *There is a unique network G on 5 binary nodes whose underlying undirected graph is a tree, but $I_{\text{global}(G)}$ is not radical. Every other network whose underlying graph is a tree has prime global Markov ideal.*

Proof. The unique network is $G_{23} = \big(\{\}, \{1\}, \{2\}, \{2\}, \{2\}\big)$. Its local and global Markov independent statements coincide and are equal to

$$\big\{1 \perp\!\!\!\perp \{3,4,5\} \,|\, 2,\ 3 \perp\!\!\!\perp \{4,5\},\ 4 \perp\!\!\!\perp \{3,5\},\ 5 \perp\!\!\!\perp \{3,4\}\big\}.$$

Computation using `Macaulay2` reveals

$$I_{\text{global}(G_{23})} \quad = \quad \ker(\Phi) \cap (I_{\text{global}(G_{23})} + (P_{+\bullet\bullet\bullet\bullet})^2),$$

where $P_{+\bullet\bullet\bullet\bullet}$ is the ideal generated by the 16 linear forms $p_{+u_2u_3u_4u_5}$. Inspecting the 81 non-prime ideals shows that G_{23} is the only example. □

We say that the network G *has an induced r-cycle* if there is an induced subgraph H of G with r vertices which consists of two disjoint directed paths which share the same start point and end point.

Proposition 18. *Of the* 301 *networks on five nodes,* 70 *have an induced* 4*-cycle or* 5*-cycle. For exactly two of these, the ideal $I_{\text{global}(G)}$ is prime.*

Proof. Once again, this follows by examination of the 301 cases. The graphs which have an induced 4-cycle but whose global Markov ideal is prime are

$$G_{265} = \{\{\}, \{1\}, \{1,2\}, \{1,2\}, \{2,3,4\}\}$$
$$\text{and} \quad G_{269} = \{\{\}, \{1\}, \{1,2\}, \{2,3\}, \{1,2,4\}\}.$$

Removing node 2 results in a 4-cycle. The local and global Markov statements are all the same up to relabeling: $\big\{1 \perp\!\!\!\perp 5 \,|\, \{2,3,4\},\ 3 \perp\!\!\!\perp 4 \,|\, 5\big\}$. □

There are four graphs with three induced 4-cycles, namely, G_{138}, G_{139}, G_{150}, G_{157}. The first two graphs give rise to the same (global or local) independence statements, and similarly for the last two. The ideal $I_{\text{global}(G_{138})}$ has the most components of any of the 301 ideals considered in this chapter.

Example 19. The network $G_{138} = \big(\{\}, \{1\}, \{1\}, \{1\}, \{2,3,4\}\big)$ is isomorphic to the one in Proposition 10. Its ideal $I_{\text{global}(G_{138})}$ has 207 minimal primes, and 37 embedded primes. Each of the 207 minimal primary components are prime. We will describe the structure of these components.

Let $F_{i_1i_2i_3} = \det \begin{pmatrix} p_{+i_1i_2i_31} & p_{+i_1i_2i_32} \\ p_{2i_1i_2i_31} & p_{2i_1i_2i_32} \end{pmatrix}$. Let J_i be the ideal generated by the 2×2 minors located in the first two rows or columns of the matrix

$$\begin{pmatrix} p_{+111i} & p_{+112i} & p_{+211i} & p_{+212i} \\ p_{+121i} & p_{+122i} & p_{+221i} & p_{+222i} \\ p_{+211i} & p_{+212i} & * & * \\ p_{+221i} & p_{+222i} & * & * \end{pmatrix}.$$

We have

$$I \quad := \quad I_{\text{global}(G_{138})} \quad = \quad J_1 \; + \; J_2 \; + \; \langle F_{111}, F_{112}, \ldots, F_{222} \rangle.$$

Each J_i is minimally generated by 9 quadrics, so that I is minimally generated by 26 quadrics. Each J_i is prime of codimension 4, and so $J_1 + J_2$ is prime of codimension 8. Since there are only 8 more quadrics, Krull's principal ideal theorem tells us that all minimal primes have codimension at most 16, which is also the codimension of the distinguished component. Note that I is a binomial ideal in the unknowns $p_{+u_2u_3u_4u_5}$ and $p_{2u_2u_3u_4u_5}$.

Table 6.1: All 207 minimal primes of the ideal $I_{\text{global}(G_{138})}$

# primes	codim	degree	faces
6	14	48	(f, f), f a facet
12	14	4	(e, e), e an edge
24	16	15	(f_1, f_2), $f_1 \cap f_2$ is an edge
48	16	4	(f, e), $f \cap e$ is a point
12	16	1	(e_1, e_2), 2 antipodal edges
48	16	1	(e_1, e_2), 2 non-parallel disjoint edges
48	16	1	(e, p), point p on the edge antipodal to e
8	16	1	(p_1, p_2), antipodal points
1	16	2316	distinguished component

Let Δ be the unit cube, with vertices $(1,1,1), (1,1,2), \ldots, (2,2,2)$. If $\sigma \subset \Delta$ is a face, define $P_{\sigma,i}$ to be the monomial prime generated by $\{p_{+vi} \mid v \notin \sigma\}$, for $i \in \{1,2\}$. If P is a minimal prime of I, which is not the distinguished component, then P must contain some $p_{+v_1v_2v_31}$, and also contain some $p_{+u_1u_2u_32}$. Therefore, there are faces σ_1 and σ_2 of Δ such that P contains $P_{\sigma_1,1} + P_{\sigma_2,2}$, and does not contain any other elements p_{+vi}. Let $m_{\sigma_1\sigma_2}$ be the product of all of the p_{+vi} such that $v \in \sigma_i$ for $i = 1, 2$. It turns out that every minimal prime ideal of I has the form

$$P_{\sigma_1,\sigma_2} \quad := \quad \big((I + P_{\sigma_1,1} + P_{\sigma_2,2}) : m_{\sigma_1\sigma_2}^{\infty}\big)$$

for some pair σ_1, σ_2 of proper faces of the cube Δ. However, not all pairs of faces correspond to minimal primes. There are 27 proper faces of the cube, and so there are $27^2 = 729$ possible minimal primes. Only 206 of these occur. The list of minimal primes is given in Table 6.1. □

6.2 Double Join of a Toric Ideal

Bayesian networks give rise to very interesting (new and old) constructions in algebraic geometry. In the next chapter, we shall encounter secant varieties. Here, we offer a generalization of Example 19 to arbitrary toric varieties. Let $I_A \subset \mathbb{R}[z_1, \ldots, z_n]$ be any *toric ideal*, specified as in [32] by a point configuration $A = \{a_1, \ldots, a_n\} \subset \mathbb{Z}^d$. Let Δ be the convex hull of A in $\mathbb{R}^d$. We define the *double join* of the toric ideal I_A to be the new ideal

$$I_A(x) + I_A(y) + \langle F_1, \ldots, F_n \rangle \;\subset\; \mathbb{R}[x_1, \ldots, x_n, y_1, \ldots, y_n, a_1, \ldots a_n, b_1, \ldots, b_n]$$

where $F_i = \det \begin{pmatrix} x_i & a_i \\ y_i & b_i \end{pmatrix}$, and $I_A(x)$ and $I_A(y)$ are generated by copies of I_A in $\mathbb{R}[x_1, \ldots, x_n]$ and $\mathbb{R}[y_1, \ldots, y_n]$ respectively. The ideal I in Example 19 is the double join of the Segre variety $\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1 \subset \mathbb{P}^7$, which is the toric variety whose polytope Δ is the 3-cube. In general, the minimal primes of the double join of I_A are indexed by pairs of faces of the polytope Δ. We believe that this construction deserves the attention of algebraic geometers.

Chapter 7

Hidden Variables and Higher Secant Varieties

7.1 Ring of Observable Probabilities

Let G be a Bayesian network on n discrete random variables and let $P_G = \ker(\Phi)$ be its homogeneous prime ideal in the polynomial ring $\mathbb{R}[D]$, whose indeterminates $p_{i_1 i_2 \cdots i_n}$ represent probabilities of events $(i_1, i_2, \ldots, i_n) \in D$. We now consider the situation when some of the random variables are hidden. After relabeling we may assume that the variables corresponding to the nodes $r+1, \ldots, n$ are hidden, while the random variables corresponding to the nodes $1, \ldots, r$ are observed. Thus the *observable probabilities* are

$$p_{i_1 i_2 \cdots i_r ++\cdots+} \quad = \quad \sum_{j_{r+1} \in [d_{r+1}]} \sum_{j_{r+2} \in [d_{r+2}]} \cdots \sum_{j_n \in [d_n]} p_{i_1 i_2 \cdots i_r j_{r+1} j_{r+2} \cdots j_n}.$$

We write $D' = [d_1] \times \cdots \times [d_r]$ and $\mathbb{R}[D']$ for the polynomial subring of $\mathbb{R}[D]$ generated by the observable probabilities $p_{i_1 i_2 \cdots i_r ++\cdots+}$. Let $\pi : \mathbb{R}^D \to \mathbb{R}^{D'}$ denote the canonical linear epimorphism induced by the inclusion of $\mathbb{R}[D']$ in $\mathbb{R}[D]$. We are interested in the following inclusions of semi-algebraic sets:

$$\pi(V_{\geq 0}(P_G)) \;\subset\; \pi(V(P_G))_{\geq 0} \;\subset\; \pi(V(P_G)) \;\subset\; \overline{\pi(V(P_G))} \;\subset\; \mathbb{R}^{D'}. \tag{7.1}$$

These inclusions are generally all strict. In particular, the space $\pi(V_{\geq 0}(P_G))$ which consists of all *observable probability distributions* is often much smaller than the space $\pi(V(P_G))_{\geq 0}$ which consists of probability distributions on D' which would be observable if non-negative or complex numbers were allowed for the hidden parameters. However, they have the same Zariski closure:

Proposition 20. *The set of all polynomial functions which vanish on the space $\pi(V_{\geq 0}(P_G))$ of observable probability distributions is the prime ideal*

$$Q_G \quad = \quad P_G \cap \mathbb{R}[D']. \tag{7.2}$$

Proof. The elimination ideal $Q_G \subset \mathbb{R}[D']$ is prime because $P_G \subset \mathbb{R}[D]$ was a prime ideal. By the Closure Theorem of Elimination Theory [4, Theorem 3, §3.2], the ideal Q_G is the vanishing ideal of the image $\pi(V(P_G))$. Since $V_{\geq 0}(P_G)$ is Zariski dense in $V(P_G)$, by the Factorization Theorem 4, and π is a linear map, it follows that $\pi(V_{\geq 0}(P_G))$ is Zariski dense in $\pi(V(P_G))$. □

7.2 Hidden Naive Bayesian Networks

We wish to demonstrate how computational algebraic geometry can be used to study hidden random variables in Bayesian networks. To this end we apply the concepts introduced above to a standard example from the statistics literature [11], [26], [27]. We fix the network G which has $n+1$ random variables $F_1, \ldots, F_n, H$ and n directed edges (H, F_i), $i = 1, 2, \ldots, n$. This is the *naive Bayes model.* The variable H is the hidden variable, and its levels $1, 2, \ldots, d_{n+1} =: r$ are called the *classes.* The observed random variables $F_1, \ldots, F_n$ are the *features* of the model. In this example, the prime ideal P_G coincides with the local ideal $I_{\mathrm{local}(G)}$ which is specified by requiring that, for each fixed class, the features are completely independent:

$$F_1 \perp\!\!\!\perp F_2 \perp\!\!\!\perp \cdots \perp\!\!\!\perp F_n \,|\, H.$$

This ideal is obtained as the kernel of the map $p_{i_1 i_2 \cdots i_n k} \mapsto x_{i_1} y_{i_2} \cdots z_{i_n}$, one copy for each fixed class k, and then adding up these r prime ideals. Equivalently, P_G is the ideal of the join of r copies of the *Segre variety*

$$X_{d_1, d_2, \ldots, d_n} \quad := \quad \mathbb{P}^{d_1 - 1} \times \mathbb{P}^{d_2 - 1} \times \cdots \times \mathbb{P}^{d_n - 1} \quad \subset \quad \mathbb{P}^{d_1 d_2 \cdots d_n - 1}. \tag{7.3}$$

The points on $X_{d_1, d_2, \ldots, d_n}$ represent tensors of rank ≤ 1. Our linear map π takes an r-tuple of tensors of rank ≤ 1 and it computes their sum, which is a tensor of rank $\leq r$. The closure of the image of π is what is called a *higher secant variety* in the language of algebraic geometry [16, Example 11.30].

Corollary 21. *The naive Bayes model with r classes and n features corresponds to the r-th secant variety of a Segre product of n projective spaces:*

$$\overline{\pi(V(P_G))} \quad = \quad \mathrm{Sec}^r(X_{d_1, d_2, \ldots, d_n}) \tag{7.4}$$

The case $n = 2$ of two features is a staple of classical projective geometry. In that special case, the image of π is closed, and $\pi(V(P_G)) = \mathrm{Sec}^r(X_{d_1, d_2})$ consists of all real $d_1 \times d_2$-matrices of rank at most r. This variety has codimension $(d_1 - r)(d_2 - r)$, provided $r \leq \min(d_1, d_2)$. Its ideal Q_G is generated by the $(r+1) \times (r+1)$-minors of the $d_1 \times d_2$ matrix (p_{ij+}). The dimension formula of Settimi and Smith [27, Theorem 1] follows immediately. For instance, in the case of two ternary features ($d_1 = d_2 = 3, r = 2$), discussed in different guises in [27, §4.2] and [16, Example 11.26], the *observable space* is the cubic hypersurface defined by the 3×3-determinant $\det(p_{ij+})$.

Table 7.1: The prime ideal defining the secant lines to the Segre variety (7.3)

$dim(X)$	$dim(\text{Sec}^2(X))$	$\prod_{i=1} d_i$	$(d_1, \ldots, d_n)$	*degree*	*cubics*
4	9	12	$(2, 2, 3)$	6	4
4	9	16	$(2, 2, 2, 2)$	64	32
5	11	16	$(2, 2, 4)$	20	16
5	11	18	$(2, 3, 3)$	57	36
5	11	24	$(2, 2, 2, 3)$	526	184
5	11	32	$(2, 2, 2, 2, 2)$	3256	768
6	13	20	$(2, 2, 5)$	50	40
6	13	24	$(2, 3, 4)$	276	120
6	13	27	$(3, 3, 3)$	783	222
6	13	32	$(2, 2, 2, 4)$	2388	544
6	13	36	$(2, 2, 3, 3)$	6144	932

The leftmost inclusion in (7.1) leads to difficult open problems even for $n = 2$ features. Here, $\pi(V(P_G))_{\geq 0}$ is the set of all non-negative $d_1 \times d_2$-matrices of rank at most r, while $\pi(V_{\geq 0}(P_G))$ is the subset consisting of all matrices of *non-negative rank* at most r. Their difference consists of non-negative matrices of rank $\leq r$ which cannot be written as the sum of r non-negative matrices of rank 1. In spite of recent progress by Barradas and Solis [1], there is still no practical algorithm for computing the non-negative rank of a $d_1 \times d_2$-matrix. Things get even harder for $n \geq 3$, when testing membership in $\pi(V_{\geq 0}(P_G))$ means computing *non-negative tensor rank*.

We next discuss what is known about the case of $n \geq 3$ features. The *expected dimension* of the secant variety (7.4) is

$$r \cdot (d_1 + d_2 + \cdots + d_n - n + 1) \; - \; 1. \tag{7.5}$$

This number is always an upper bound, and it is an interesting problem, studied in the statistics literature in [11], to characterize those cases $(d_1, \ldots, d_n; r)$ when the dimension is less than the expected dimension. We note that the results on dimension in [11] are all special cases of results by Catalisano, Geramita and Gimigliano [3], and the results on singularities in [11] follow from the geometric fact that the r-th secant variety of any projective variety is always singular along the $(r-1)$-st secant variety. The statistical problem of *identifiability*, addressed in [26], is related to the beautiful work of Strassen [30] on tensor rank, notably his Theorem 2.7 on *optimal computations*.

In Table 7.1 we display the range of straightforward `Macaulay2` computations when

$$dim(X) = d_1 + \cdots + d_n - 1$$

is small. First consider the case of two classes ($r = 2$), which corresponds to secant lines on $X = \mathbb{P}^{d_1-1} \times \cdots \times \mathbb{P}^{d_n-1}$. In each of these cases, the ideal Q_G is generated by cubic polynomials, and each of these cubic generators is the determinant of a two-dimensional matrix obtained by flattening the tensor $(p_{i_1 i_2 \cdots i_n})$. The column labeled "cubics" lists the number of minimal generators. For example, in the case ($d_1 = d_2 = d_3 = 3$), we can flatten (p_{ijk}) in three possible ways to a 3×9-matrix, and these have $3 \cdot \binom{9}{3} = 252$ maximal subdeterminants. The vector space spanned by these subdeterminants has dimension 222, the listed number of minimal generators. The column "degree" lists the degree of the projective variety $\mathrm{Sec}^2(X)$, which is 783 in the previous example. These computational results in Table 7.1 lead us to make the following conjecture:

Conjecture 22. *The prime ideal Q_G of any naive Bayes model G with $r = 2$ classes is generated by the 3×3-subdeterminants of any two-dimensional table obtained by flattening the n-dimensional table $(p_{i_1 i_2 \cdots i_n})$.*

This conjecture was proved set–theoretically in all cases and ideal–theoretically for $n = 3$ by Landsberg and Manivel in [19].

It was proved by Catalisano, Geramita and Gimigliano that the variety $\mathrm{Sec}^2(X)$ always has the expected dimension (7.5) when $r = 2$. A well-known example (see [13, page 221]) when the dimension is less than expected occurs for four classes and three binary features ($r = 3, n = 4, d_1 = d_2 = d_3 = d_4 = 2$). Here (7.5) evaluates to 14, but $\dim(\mathrm{Sec}^3(X)) = 13$ for $X = \mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1$. The corresponding ideal Q_G is a complete intersection generated by any two of the three 4×4-determinants obtained by flattening the $2 \times 2 \times 2 \times 2$-table (p_{ijkl}). The third is a signed sum of the other two.

The problem of identifying explicit generators of Q_G is much more difficult when $r \geq 3$, i.e., when the hidden variable has three or more levels. We present the complete solution for the case of three ternary features. Here (p_{ijk}) is an indeterminate $3 \times 3 \times 3$-tensor which we wish to write as a sum of r rank one tensors. The following solution is derived from a result of Strassen [30, Theorem 4.6]. Let $A = (p_{ij1})$, $B = (p_{ij2})$ and $C = (p_{ij3})$ be three 3×3-matrices obtained by taking slices of the $3 \times 3 \times 3$-table (p_{ijk}).

Proposition 23. *Let Q_G be the ideal of $\mathrm{Sec}^r(\mathbb{P}^2 \times \mathbb{P}^2 \times \mathbb{P}^2)$, the naive Bayes model with $n = 3$ ternary features with r classes. If $r = 2$ then Q_G is generated by the cubics described in Conjecture 22. If $r = 3$ then Q_G is generated by the quartic entries of the various 3×3-matrices of the form $A \cdot \mathrm{adj}(B) \cdot C - C \cdot \mathrm{adj}(B) \cdot A$. If $r = 4$ then Q_G is the principal ideal generated by the following homogeneous polynomial of degree 9 with 9,216 terms:*

$$\det(B)^2 \cdot \det\big(A \cdot B^{-1} \cdot C - C \cdot B^{-1} \cdot A\big).$$

If $r \geq 5$ then Q_G is the zero ideal.

Chapter 8

Networks on Three Observable Variables and One Hidden Variable

8.1 Model Selection

Statisticians and other applied researchers are often faced with the problem of choosing the appropriate model that best fits a given set of observations. One possible approach to model selection is the Bayesian approach by which a model $\mathcal{M}$ is chosen according to the maximum posteriori probability, given the observed data D:

$$p(\mathcal{M}|D) \propto p(\mathcal{M}, D) = p(\mathcal{M})p(D|\mathcal{M}).$$

The probability $p(D|\mathcal{M})$ that the data D is generated by the model $\mathcal{M}$ is obtained via integration over all possible parameter values with which the model is compatible. A closed–form formula for the *marginal likelihood* $p(D|\mathcal{M})$ is not known in general for Bayesian networks with hidden variables. But Schwarz [28] derived an asymptotic formula for the marginal likelihood when the model is an affine subspace of the natural parameter space of an exponential family. This formula is known as the *Bayesian Information Criteria score*. Later on, Haughton [17] established, under some regularity conditions that the BIC score is a valid asymptotic rule for selecting models from *curved exponential families*. In particular, this result holds for any Bayesian network without hidden variables, see [10, Theorem 4]. We note that although researchers have been using the BIC score for selecting models among Bayesian networks with hidden variables, it was shown in [24] that the BIC score is generally not valid for statistical models that belong to a *stratified exponential family (SEF)*. Moreover, it was shown in [10] that Bayesian networks with hidden variables are stratified exponential families.

There are two primary reasons for which the BIC score is generally not valid for Bayesian networks with hidden variables:

1. The number of network parameters is no longer the dimension of the associated variety. In other words, there is a drop in the dimension of the variety as explained in the previous chapter.

2. The associated variety is singular along the set of maximum likelihood points.

Hence, the understanding and computation of the dimension and sigularities of these varieties is crucial for the problem of model selection. This chapter provides a solution to the first problem for any Bayesian network on four random variables when one variable is assumed to be hidden.

8.2 Polynomial Constraints and Dimension

Bayesian networks with hidden variables are usually defined parametrically because the independence and non-independence constraints on the distributions over the observable variables are not easily established. Since these constraints vary from one model to another they can be used to distinguish between models. Moreover, since these constraints are over the observable variables, their fit to data can be measured directly with some specially-designed statistical tests.

In this chapter, we will study all Bayesian Networks G on three observable random variables and one hidden variable. We give a generating set for the ideal of all polynomial functions which vanish on the space of observable probability distributions. This prime ideal is given by the equation (7.2) in Proposition 20

$$Q_G = P_G \cap \mathbb{R}[D'],$$

where P_G is the distinguished ideal

$$P_G = I_{\mathrm{local}(G)} : \mathbf{p}^\infty = I_{\mathrm{global}(G)} : \mathbf{p}^\infty,$$

as introduced in Theorem 9. Throughout this chapter we will make use of the projection

$$\pi : \mathbb{C}^D \longrightarrow \mathbb{C}^{D'}.$$

We also give formulas for the *dimension* of all the ideals involved in these computations. An interesting ideal appears at the end of our list, namely $P_{G_{17}} \cap \mathbb{C}[D']$ which is generated by irreducible **sextic** polynomials.

Some networks require the hidden variable to be binary for some of our results to hold. This restriction is due to a fundamental distinction between bilinear and multilinear objects. The next paragraph appears in [16, p. 100]:

> While a bilinear object $\phi \in V \otimes W$ is completely determined (up to the action of $\mathrm{GL}(V) \times \mathrm{GL}(W)$) by its one invariant, i.e., the *rank*. The situation with trilinear objects $\omega \in U \otimes V \otimes W$ is more subtle. The action of $\mathrm{GL}(V) \times \mathrm{GL}(V) \times \mathrm{GL}(W)$ on $U \otimes V \otimes W$ will not in general have a finite number of orbits, but rather a continuously varying family of them, whose invariants are far from being completely understood. One exception to this is the special case where the dimension of one of the factors is 2.

Table 8.1 gives all the non-isomorphic directed acyclic graphs on 4 vertices, except the graphs arising from the complete graph. This table is part of table 5.1, but it is given here in terms of diagrams for the convenience of the reader.

Conjecture 22 states that the prime ideal Q_G of any *naive Bayes model* G with $r = 2$ classes is generated by the 3×3-subdeterminants of any two-dimensional table obtained by flattening the n-dimensional table $(p_{i_1 i_2 \cdots i_n})$. Here, we will discuss this conjecture for the particular case of $n = 3$, see treatment for G_{24} below. For this we recall [16, Proposition 12.2, Exercise 11.29].

Proposition 12.2. *The variety $M_k \subset M$ of $m \times n$ matrices of rank at most k is irreducible of codimension $(m-k)(n-k)$ in M.*

Exercise 11.29. *Let M be the projective space of $m \times n$ matrices, and let $M_k \subset M$ be the subvariety of matrices of rank at most k. Take k so that $2k < \min(m, n)$. Show that the secant variety $S(M_k)$ is equal to the subvariety $M_{2k} \subset M$ of matrices of rank at most $2k$. Indeed, show that for a map $A : K^m \to K^n$ of rank $2k$ and any pair of complementary k-dimensional subspaces $\Lambda, \Xi \subset \mathrm{Im}(A)$ the composition of A with projections of Λ and Ξ give an expression of A as a sum of two matrices of rank k. Deduce that in the case $\mathbb{P}^n = M, X = M_k$, if $\mathcal{S}(X)$ denotes the variety of secant lines to X, the fiber of the incidence correspondence*

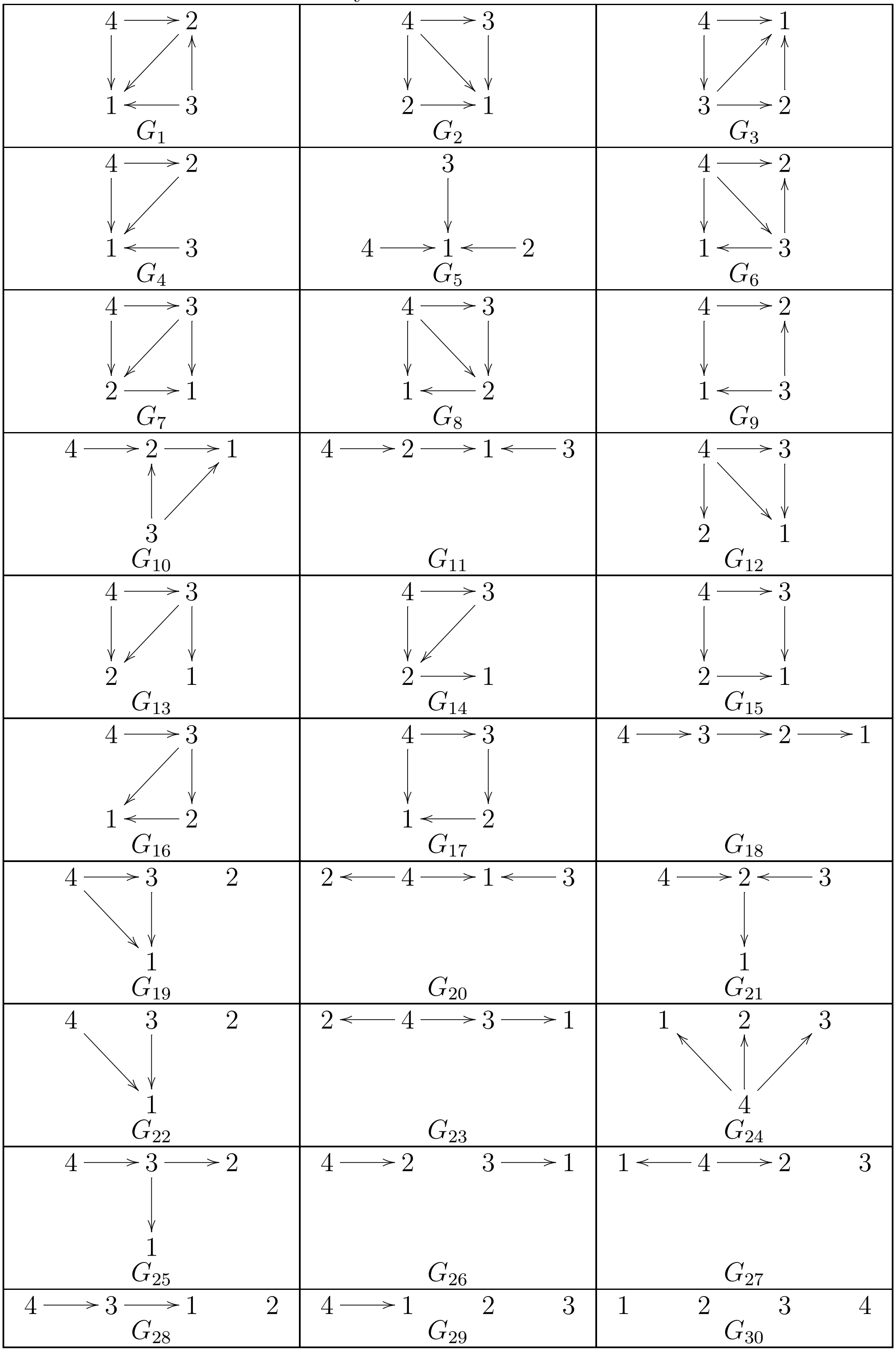
$$\Sigma = \{(l,p) : \ p \in l\} \subset \mathcal{S}(X) \times M \subset \mathbb{G}(1,n) \times M$$

over a general point $S(X)$ has dimension $2k^2$, and hence that

$$\dim(S(M_k)) \leq 2\dim(M_k) + 1 - 2k^2.$$

Proof. The secant variety is contained in the set of matrices which are sum of two matrices of rank at most k, so any matrix in $S(M_k)$ has rank at most $2k$. On the other hand, since $2k < \min(m,n)$ and A has rank $2k$, A can be decomposed as the sum of two rank k matrices by using the first k independent rows and setting the remaining rows to 0, or the last k independent rows and the remaining rows 0. Thus the secant variety $S(M_k) = M_{2k}$. Now consider Σ, $\pi_2(\Sigma) \subset M$ is the secant variety $S(X) = M_{2k}$. The fiber dimension over a general point is equal to the dimension of the space of decompositions (Λ, Ξ) of the image of A as the direct sum of subspaces. The Grassmannian $G(k, 2k)$ of k planes in $2k$ space has dimension k^2, see [16, p.138], so the fiber dimension over a general point in $S(X)$ is $2k^2$. Thus by [16, Theorem 11.12], since $S(X)$ is irreducible, $\dim(\Sigma) = \dim(S(X)) + 2k^2$. So, $\dim(S(X)) = 2\dim(M_k) + 1 - 2k^2$., see [16, Prop. 11.24]. □

Table 8.1: All Bayesian Networks on four variables



The main result of this chapter is the following theorem, which states that if G is a Bayesian network on four random variables ($G \neq G_{17}$) where one of the variables is a hidden binary variable, then the variety associated to G is the join (or the intersection) of several higher–secant varieties of Segre varieties.

For the networks 15, 17, 20, 21, 23, and 27, we have conjectures about the generating set based on extensive computations for particular cases and the computation of the dimension of the corresponding ideals in general. We should also point out that a proof for any of the last four networks would also yield a proof for the remaining three.

Theorem 24. *Of the 30 Bayesian networks on three random variables and one hidden variable*

(I) five networks G always give zero ideals $Q_G = P_G \cap \mathbb{C}[D']$, regardless of the number of levels of the random variables. (numbers 1, 3, 7, 10, 16 *in Table 8.1).*

(II) Fourteen networks G always give ideals Q_G generated by quadratic polynomials arising from the 2×2 minors of certain matrices of indeterminates (numbers 4, 5, 11, 13, 14, 18, 19, 21, 22, 25, 26, 28, 29, 30 *in Table 8.1).*

(III) If the hidden variable is binary, ten networks G give ideals Q_G generated by quadratic and cubic polynomials arising from the 2×2 or 3×3 minors of certain matrices of indeterminates (numbers 2, 6, 8, 9, 12, 15, 20, 23, 24, 27 *in Table 8.1).*

(IV) The network G_{17} gives an ideal $Q_{G_{17}}$ generated by irreducible sextic polynomials and cubic polynomials, if the hidden variable is binary.

Proof. We prove this theorem by an exhaustive case analysis of all thirty networks.

Network 1*:* The ideal $I_{\mathrm{local}(G_1)}$ is a prime ideal equal to $I_{3 \perp\!\!\!\perp 4}$. It is generated by the 2×2-minors of the matrix (p_{++kl}), where the rows are indexed by $k \in [d_3]$ and the columns are indexed by $l \in [d_4]$. We claim that $I_{3 \perp\!\!\!\perp 4} \cap \mathbb{C}[D'] = 0$. It is enough to show that every $d_1 \times d_2 \times d_3$-table $a = (a_{ijk})$ is the projection of a table in $V(I_{3 \perp\!\!\!\perp 4})$. Let A be the $d_1 \times d_2 \times d_3 \times d_4$-table defined by $A_{ijkl} = a_{ijk}/d_4$, so $A_{++kl} = a_{++k}/d_4$. Then, the d_4 rows of $A' = (A_{++kl})$ are equal to each other, that is, A' has rank 1. So $A \in V(I_{3 \perp\!\!\!\perp 4})$ and $\pi(A) = a$.

Networks 3, 7, 16*:* The ideal $I_{\mathrm{local}(G_3)}$ is a prime ideal equal to $I_{2 \perp\!\!\!\perp 4|3}$. It is generated by the 2×2-minors of the d_3 matrices (p_{+jk_0l}), where the rows are indexed by $j \in [d_2]$, the columns are indexed by $l \in [d_4]$, and $k_0 \in [d_3]$. We claim that $I_{2 \perp\!\!\!\perp 4|3} \cap \mathbb{C}[D'] = 0$. It is enough to show that every $d_1 \times d_2 \times d_3$-table $a = (a_{ijk})$ is the projection of a table in $V(I_{2 \perp\!\!\!\perp 4|3})$. Let A be the $d_1 \times d_2 \times d_3 \times d_4$-table defined by $A_{ijkl} = a_{ijk}/d_4$, so $A_{+jkl} = a_{+jk}/d_4$. For each $k_0 \in [d_3]$, the d_4 rows of $A' = (A_{+jk_0l})$ are equal to each other, that is, A' has rank 1. So $A \in V(I_{2 \perp\!\!\!\perp 4|3})$ and $\pi(A) = a$. The ideal $I_{\mathrm{local}(G_7)}$ is a prime ideal equal to $I_{1 \perp\!\!\!\perp 4|\{2,3\}}$. A similar argument shows that $I_{1 \perp\!\!\!\perp 4|\{2,3\}} \cap \mathbb{C}[D'] = 0$. Finally, observe that the prime ideal $I_{\mathrm{global}(G_{16})} = I_{\{1,2\} \perp\!\!\!\perp 4|3}$. Thus we also have $Q_{16} = 0$.

Network 10*:* The ideal $I_{\mathrm{local}(G_{10})}$ is a radical ideal equal to $I_{1\perp\!\!\!\perp 4|\{2,3\}} + I_{3\perp\!\!\!\perp 4}$. We claim that $I_{\mathrm{local}(G_{10})} \cap \mathbb{C}[D'] = 0$. It is enough to show that every $d_1 \times d_2 \times d_3$-table $a = (a_{ijk})$ is the projection of a table in $V(I_{\mathrm{local}(G_{10})})$. Let A be the $d_1 \times d_2 \times d_3 \times d_4$-table defined by $A_{ijkl} = a_{ijk}/d_4$. We saw above that $A \in V(I_{1\perp\!\!\!\perp 4|\{2,3\}})$ and $A \in V(I_{3\perp\!\!\!\perp 4})$, so

$$A \in V(I_{\mathrm{local}(G_{10})}) = V(I_{1\perp\!\!\!\perp 4|\{2,3\}}) \cap V(I_{3\perp\!\!\!\perp 4}).$$

Theorem 12 shows that if $d_4 = 2$, then $I_{\mathrm{local}(G_{10})}$ is a prime ideal. In this case, $Q_{10} = I_{\mathrm{local}(G_{10})} \cap \mathbb{C}[D'] = 0$. If $d_4 > 2$, $I_{\mathrm{local}(G_{10})}$ is the intersection of the distinguished component $P_{G_{10}}$ and 2^{d_3-1} prime ideals P_σ indexed by all proper subsets $\sigma \subset [d_3]$. By construction, the ideal $M_\sigma = \langle p_{+jkl} : j \in [d_2], k \in \sigma, l \in [d_4]\rangle$ is contained in P_σ, so $M_\sigma \cap \mathbb{C}[D'] = \langle p_{+jk+} : j \in [d_2], k \in \sigma\rangle$ is a subset of $P_\sigma \cap \mathbb{C}[D']$. Then

$$\dim(P_\sigma \cap \mathbb{C}[D']) \leq \dim(M_\sigma \cap \mathbb{C}[D']) < d_1 d_2 d_3$$

Moreover, since $V(Q_{10}) \cup \cup_{\sigma\in[d_3]} V(P_\sigma \cap \mathbb{C}[D']) = \mathbb{C}[D']$, then $V(Q_{10}) = \mathbb{C}[D']$, that is, $Q_{10} = 0$.

Networks 2, 12*:* The ideal $I_{\mathrm{local}(G_2)}$ is a prime ideal equal to $I_{2\perp\!\!\!\perp 3|4}$. It is generated by the 2×2-minors of the d_4 matrices (p_{+jkl_0}), where the rows are indexed by $j \in [d_2]$ and the columns are indexed by $k \in [d_3]$. If $d_4 = 2$, then by Exercise 11.29 above

$$V(Q_2) = \overline{\pi(V(P_{G_2}))} = S(M_1) = M_2,$$

where M_k is the variety of $d_2 \times d_3$ matrices of rank at most k. Thus the ideal Q_2 is given by the 3×3-minors of the $d_2 \times d_3$ matrix $(p + jk+)$. The prime ideal $I_{\mathrm{local}(G_{12})} = I_{2\perp\!\!\!\perp \{1,3\}|4}$, so we can proceed in exactly as for G_2 to find a set of generators for Q_{12}.

Networks 6, 8*:* The ideal $I_{\mathrm{local}(G_6)}$ is a prime ideal equal to $I_{1\perp\!\!\!\perp 2|\{3,4\}}$. It is generated by the 2×2-minors of the $d_3 d_4$ matrices $(p_{ijk_0l_0})$, where the rows are indexed by $i \in [d_1]$ and the columns are indexed by $j \in [d_2]$. Assume $d_4 = 2$. For each $k_0 \in [d_3]$, let I_{k_0} be the ideal generated by the 2×2-minors of the 2 matrices (p_{ijk_0l}). Then just as for G_2

$$V(I_{k_0} \cap \mathbb{C}[D']) = \overline{\pi(V(I_{k_0}))} = S(M_1) = M_2, \tag{8.1}$$

where M_k is the variety of $d_1 \times d_2$ matrices of rank at most k. Note that $I_{\mathrm{local}(G_6)} = \sum_{k\in[d_3]} I_k$, and the ideals I_k are defined in pairwise disjoint set of indeterminates. For each $k_0 \in [d_3]$, the fiber dimension over a general point in $V(I_{k_0} \cap \mathbb{C}[D'])$ is equal to 2. Thus the fiber dimension over a general point in $V(I_{\mathrm{local}(G_6)} \cap \mathbb{C}[D'])$ is equal to $2d_3$. Moreover,

$$\begin{aligned}\operatorname{codim}(I_{\mathrm{local}(G_6)}) &= \operatorname{codim}(\sum_{k\in[d_3]} I_k) = \sum_{k\in[d_3]} \operatorname{codim}(I_k)\\ &= \sum_{k\in[d_3]} 2(d_1-1)(d_2-1) = 2d_1d_2d_3 - 2d_1d_3 - 2d_2d_3 + 2d_3\end{aligned}$$

So $\dim(I_{\text{local}(G_6)}) = 2d_1d_3 + 2d_2d_3 - 2d_3$, and

$$\operatorname{codim}\Big(\sum_{k\in[d_3]}(I_k \cap \mathbb{C}[D'])\Big) = \sum_{k\in[d_3]}\operatorname{codim}(I_k\cap\mathbb{C}[D']) = d_3(d_1-2)(d_2-2)$$

So $\dim\big(\sum_{k\in[d_3]}(I_k\cap\mathbb{C}[D'])\big) = 2d_1d_3 + 2d_2d_3 - 4d_3$. Also,

$$\dim(Q_6) = \dim(I_{\text{local}(G_6)}) - 2d_3 = 2d_1d_3 + 2d_2d_3 - 4d_3.$$

Hence $\sum_{k\in[d_3]}(I_k\cap\mathbb{C}[D']) \subseteq Q_6$. Moreover, both prime ideals have the same dimension, thus

$$Q_6 = \sum_{k\in[d_3]} I_k\cap\mathbb{C}[D'],$$

and equation (8.1) gives a set of generators for this ideal. The prime ideal $I_{\text{local}(G_8)} = I_{1 \perp\!\!\!\perp 3|\{2,4\}}$. So we can proceed in a similar way as for G_6 to find a set of generators for Q_8.

Network 9*:* The ideal $I_{\text{local}(G_9)}$ is a prime ideal equal to $I_{1\perp\!\!\!\perp 2|\{3,4\}} + I_{3\perp\!\!\!\perp 4}$. We know that $I_{3\perp\!\!\!\perp 4}\cap\mathbb{C}[D'] = 0$. We claim that if $d_4 = 2$, then

$$Q_9 = I_{1\perp\!\!\!\perp 2|\{3,4\}}\cap\mathbb{C}[D'] = Q_6.$$

Let $a\in V(Q_6)$, that is, for each $k_0\in[d_3]$ the $d_1\times d_2$-matrix (a_{ijk_0}) has rank 1. Let the $d_1\times d_2\times d_3\times d_4$-table A be defined as $A_{ijkl} = a_{ijk}/2$. Then $A\in V(I_{1\perp\!\!\!\perp 2|\{3,4\}})$, because for each $(k_0,l_0)\in[d_3]\times[d_4]$ the $d_1\times d_2$-matrix $(A_{ijk_0l_0}) = (a_{ijk_0}/2)$ has rank 1. Moreover, $A_{++k1} = A_{++k2} = a_{++k}/2$, so the $d_3\times 2$-matrix (A_{++kl}) has also rank 1. Thus, $A\in V(I_{3\perp\!\!\!\perp 4})\cap V(I_{1\perp\!\!\!\perp 2|\{3,4\}}) = V(I_{\text{local}(G_9)})$ and $a = \pi(A)\in V(Q_9)$.

Network 24*:* It is worth mentioning that this result was proved set–theoretically in all cases and ideal–theoretically for $n = 3$ by Landsberg and Manivel in [19]. Here, we compute the dimension of all the ideals involved. We also give some insight that will be helpful for other Bayesian networks.

The graph G_{24} corresponds to the *naive Bayes model* with d_4 classes and 3 features, we will examine the case $d_4 = 2$. The prime ideal $I_{\text{local}(G_{24})}$ is equal to the sum

$$I_{\text{local}(G_{24})} = I_{1\perp\!\!\!\perp\{2,3\}|4} + I_{2\perp\!\!\!\perp\{1,3\}|4} + I_{3\perp\!\!\!\perp\{1,2\}|4}$$

Let $I_1 = I_{1\perp\!\!\!\perp\{2,3\}|4}$, $I_2 = I_{2\perp\!\!\!\perp\{1,3\}|4}$, and $I_3 = I_{3\perp\!\!\!\perp\{1,2\}|4}$, note that $I_{\text{local}(G_{24})} = I_1 + I_2 = I_1 + I_3 = I_2 + I_3$. Assume $d_4 = 2$, then by [16, Proposition 12.2] we have that $\operatorname{codim}(I_1) = 2(d_1-1)(d_2d_3-1)$, so $\dim(I_1) = 2d_1 + 2d_2d_3 - 2$. The ideal I_1 is generated by the 2×2-minors of the $d_1\times d_2d_3$-matrices $M_{l_0} = (p_{ijkl_0})$, where $l_0\in\{1,2\}$. Similarly, The ideal I_2 is generated by the 2×2-minors of the $d_2\times d_1d_3$-matrices $N_{l_0} = (p_{ijkl_0})$, where $l_0\in\{1,2\}$. Note that for each $(k_0,l_0)\in[d_3]\times[d_4]$, the $d_1\times d_2$-matrix $M_{k_0l_0} = (p_{ijk_0l_0})$ is the transpose of the $d_2 \times d_1$-matrix $N_{k_0l_0} = (p_{ijk_0l_0})$. Hence, for each $k \geq 2$, the 2×2-minors of N_{kl_0} lowers the dimension of I_1 by $d_2 - 1$. Thus

$$\begin{aligned}\dim(I_{\text{local}(G_{24})}) &= \dim(I_1 + I_2) = 2d_1 + 2d_2d_3 - 2 - 2(d_2 - 1)(d_3 - 1) \\ &= 2d_1 + 2d_2 + 2d_3 - 4.\end{aligned}$$

Let $\widetilde{I_r} = I_r \cap \mathbb{C}[D']$ for $r = 1, 2, 3$. Recall that Exercise 11.29 gives a generating set for $\widetilde{I_r}$ and its dimension. Thus $\widetilde{I_1}$ is generated by the 3×3-minors of the $d_1 \times d_2d_3$-matrix $\tilde{M} = (p_{ijk+})$, and $\dim(\widetilde{I_1}) = \dim(I_1) - 2 = 2d_1 + 2d_2d_3 - 4$. This result can be obtained directly by applying [16, Proposition 12.2], but we chose to apply [16, Theorem 11.12] which implies $\dim(I_1) = \dim(\widetilde{I_1}) + \mu$, where μ is the minimum of the fiber dimensions of π through $A \in V(I_1)$, that is, $\mu = \min\{\dim(\pi^{-1}\pi(A)) : A \in V(I_1)\}$. Exercise 11.29 shows that $\mu = 2$ which also implies our assertion about $\dim(\widetilde{I_1})$. Proceeding in a similar way as for $\dim(I_1 + I_2)$, we conclude that $\widetilde{I_2}$ lowers the dimension of $\widetilde{I_1}$ by $2(d_2 - 2)(d_3 - 1)$, so $\dim(\widetilde{I_1} + \widetilde{I_2}) = 2d_1 + 2d_2 + 4d_3 - 8$. The ideal $\widetilde{I_3}$ is generated by the 3×3-minors of the $d_3 \times d_1d_2$-matrix $L = (p_{ijk+})$. Note that the k_0 row of L can be obtained by flattening the $d_1 \times d_2$-matrix $\tilde{M}_{k_0} = (p_{ijk_0+})$. Hence, the ideal $\widetilde{I_3}$ lowers the dimension of $\widetilde{I_1} + \widetilde{I_2}$ by $2(d_3 - 2)$. Thus,

$$\dim(\widetilde{I_1} + \widetilde{I_2} + \widetilde{I_3}) = 2d_1 + 2d_2 + 2d_3 - 4.$$

Clearly, $\sum_{s=1}^{3} \widetilde{I_s} \subseteq Q_{24}$, and $\dim(\sum_{s=1}^{3} \widetilde{I_s}) = \dim(I_{\text{local}\, G_{24}}) \geq \dim(Q_{24})$. But the ideal–theoretic result in [19] states that

$$\sum_{s=1}^{3} \widetilde{I_s} = Q_{24}.$$

Thus, $\dim(Q_{24}) = 2d_1 + 2d_2 + 2d_3 - 4$.

Network 4*:* The ideal $I_{\text{local}(G_4)}$ is a prime ideal equal to $I_{3 \perp\!\!\!\perp \{2,4\}}$. It is generated by the 2×2-minors of the matrix $M = (p_{+jkl})$, where the rows are labeled by $k \in [d_3]$ and the columns are labeled by the pairs $(j, l) \in [d_2] \times [d_4]$. First, change coordinates in $\mathbb{R}[D]$ by replacing each unknown p_{1jkl} by $p_{jkl} = \sum_{i=1}^{d_1} p_{ijkl}$. This coordinate change transforms $I_{local(G_4)}$ into a binomial ideal in $\mathbb{R}[D]$. In Theorem 12, we proved that $I_{local(G_4)}$ is a prime ideal and hence equal to P_{G_4}.

We want to show that $Q_4 = P_{G_4} \cap \mathbb{C}[D']$ is equal to the ideal I generated by the 2×2-minors of the matrix $N = (p_{jk+})$ where the rows are indexed by $k \in [d_3]$ and the columns are indexed by $j \in [d_2]$. Each column of this matrix is obtained by taking the sum of the corresponding d_4 columns of M. Let $f = p_{j_1k_1+}p_{j_2k_2+} - p_{j_1k_2+}p_{j_2k_1+}$ be an arbitrary generator of I. Note that f can be written as the sum of binomials of the form $p_{j_1k_1l_1}p_{j_2k_2l_2} - p_{j_1k_2l_1}p_{j_2k_1l_2}$. These binomials are generators of the ideal P_{G_4}, thus $f \in P_{G_4}$. But $f \in \mathbb{R}[D']$, therefore $f \in Q_4$. We have shown that $I \subset Q_4$.

The ideal Q_4 is prime as shown in Proposition 20. The ideal I is prime since it is generated by the 2×2-minors of a generic matrix [6]. Hence, to show equality between ideals, it would suffice to show that $V(I) \subset V(Q_4)$. Let $a \in V(I)$, then a is a $d_3 \times d_2$-matrix of rank 1 since all the 2×2-minors of a vanish. Let A be the $d_3 \times d_2d_4$-matrix defined by $A_{jkl} = a_{jk}/d_4$. Clearly, A has rank 1, so $A \in V(P_{G_4})$. Thus $a = \pi(A) \in V(Q_4)$, since $V(Q_4) = \overline{\pi(V(P_{G_4}))}$ by the Closure Theorem of Elimination theory [4, Theorem 3, §3.2].

Network 5*:* First, change coordinates in $\mathbb{R}[D]$ by replacing each indeterminate p_{1jkl} by $p_{jkl} = \sum_{i=1}^{d_1} p_{ijkl}$. The prime ideal $I = I_{\mathrm{local}(G_5)}$ is equal to the sum of ideals $I_{3 \perp\!\!\!\perp \{2,4\}} + I_{2 \perp\!\!\!\perp \{3,4\}}$. The first ideal equals $I_{\mathrm{local}(G_4)}$, and the corresponding ideal in the subring $\mathbb{C}[D']$ is given by the 2×2-minors of the matrix $N = (p_{jk+})$ where the rows are indexed by $k \in [d_3]$ and the columns are indexed by $j \in [d_2]$. Similarly, $I_{2 \perp\!\!\!\perp \{3,4\}} \cap \mathbb{C}[D']$ is given by the 2×2-minors of the matrix (p_{jk+}) where the rows are indexed by $j \in [d_2]$ and the columns are indexed by $k \in [d_3]$. This matrix is the transpose of N, hence the ideal of the 2×2-minors of this matrix equals Q_4. Hence $Q_5 = Q_4$.

We can take a different approach to obtain this result. Observe that $I = I_{3 \perp\!\!\!\perp \{2,4\}} + I_{4 \perp\!\!\!\perp \{2,3\}}$. We claim that $I_{4 \perp\!\!\!\perp \{2,3\}} \cap \mathbb{C}[D'] = 0$. It is enough to show that every $d_2 \times d_3$-matrix $a = (a_{jk})$ is the projection of a table in $V(I_{4 \perp\!\!\!\perp \{2,3\}})$. Let A be the $d_4 \times d_2d_3$-matrix defined by $A_{jkl} = a_{jk}/d_4$. The d_4 rows of A are equal to each other, so A has rank 1. So $A \in V(I_{4 \perp\!\!\!\perp \{2,3\}})$ and $\pi(A) = a$. Observe that $I_{3 \perp\!\!\!\perp \{2,4\}} \cap \mathbb{C}[D'] \subseteq I \cap \mathbb{C}[D']$, that is, $V(Q_5) \subseteq V(Q_4)$. On the other hand, if $a \in V(Q_4)$ then for A as defined above, $A \in V(P_{G_4})$, so $A \in V(P_{G_4}) \cap V(I_{4 \perp\!\!\!\perp \{2,3\}})$, that is, $A \in V(I)$, and $a = \pi(A) \in V(Q_5)$. Hence $Q_5 = Q_4$.

Network 11*:* First, change coordinates in $\mathbb{R}[D]$ by replacing each unknown p_{1jkl} by $p_{jkl} = \sum_{i=1}^{d_1} p_{ijkl}$. The prime ideal $I_{\mathrm{global}(G_{11})}$ is equal to the sum of ideals $I_{3 \perp\!\!\!\perp \{2,4\}} + I_{\{1,3\} \perp\!\!\!\perp 4|2}$. The first ideal equals $I_{\mathrm{local}(G_4)}$. We will show that $I_{\{1,3\} \perp\!\!\!\perp 4|2} \cap \mathbb{C}[D'] = 0$. It is enough to show that every $d_1d_3 \times d_2$-matrix $a = (a_{ijk})$ is the projection of a matrix in $V(I_{\{1,3\} \perp\!\!\!\perp 4|2})$. Let A be the $d_1d_3 \times d_2d_4$-matrix defined by $A_{ijkl} = a_{ijk}/d_4$. Hence for each $j \in [d_2]$ the $d_1d_3 \times d_4$-matrix (A_{ijkl}) has rank 1. Actually, the d_4 columns of this matrix are equal. Hence $A \in V(I_{\{1,3\} \perp\!\!\!\perp 4|2})$. Moreover, if $a \in V(Q_4)$ then $A \in V(P_{G_4})$, so $Q_{11} = Q_4$.

Network 20*:* First, change coordinates in $\mathbb{R}[D]$ by replacing each unknown p_{1jkl} by $p_{jkl} = \sum_{i=1}^{d_1} p_{ijkl}$. The binomial prime ideal $I_{\mathrm{local}(G_{20})}$ is equal to the sum of ideals $I = I_{2 \perp\!\!\!\perp \{1,3\}|4} = I_{\mathrm{local}(G_{12})}$ and $J = I_{3 \perp\!\!\!\perp \{2,4\}} = I_{\mathrm{local}(G_4)}$. Denote by $\tilde{I} = I \cap \mathbb{C}[D']$ and $\tilde{J} = J \cap \mathbb{C}[D']$. Recall that the ideal I is generated by the 2×2-minors of the d_4 matrices $M_l = (p_{ijkl})$ where the rows are indexed by $(i,k) \in [d_1] \times [d_3]$, the columns by $j \in [d_2]$, and $l \in [d_4]$ is fixed for each matrix. The ideal J is generated by the 2×2-minors of the matrix $N = (p_{1jkl})$ where the rows are indexed by $k \in [d_3]$ and the columns by $(j,l) \in [d_2] \times [d_4]$. For each l, the ideal generated by the 2×2-minors of M_l has codimension $(d_1d_3 - 1)(d_2 - 1)$. Moreover, since the entries of each matrix are pairwise disjoint, the codimension of I equals $d_4(d_1d_3 - 1)(d_2 - 1)$. Hence

$$\dim(I) = d_1d_3d_4 + d_2d_4 - d_4 \tag{8.2}$$

Similarly, the codimension of J equals $(d_3 - 1)(d_2d_4 - 1)$, so

$$\dim(J) = d_1d_2d_3d_4 - d_2d_3d_4 + d_2d_4 + d_3 - 1 \tag{8.3}$$

Let $M_{i_0l_0}$ be the $d_3 \times d_2$-matrix $(p_{i_0jkl_0})$, then

$$M_l = \begin{pmatrix} M_{1l} \\ M_{2l} \\ \vdots \\ M_{d_1l} \end{pmatrix} \quad \text{and} \quad N = (M_{11}M_{12}\cdots M_{1d_4}).$$

Hence, just as for G_{24}, the ideal J removes $d_3 - 1$ parameters of all but one of the matrices M_l. Thus,

$$\dim(I + J) = \dim(I) - (d_3 - 1)(d_4 - 1) = d_1d_3d_4 + d_2d_4 - d_3d_4 + d_3 - 1. \tag{8.4}$$

Let $d_4 = 2$. Then the prime ideal $\tilde{I}$ is generated by the 3×3-minors of the two dimensional table $M_+ = (p_{ijk+})$ where the rows are indexed by $j \in [d_2]$ and the columns are indexed by pairs $(i, k) \in [d_1] \times [d_3]$, that is, M_+ is obtained by flattening the 3-dimensional table (p_{ijk+}) according to the relation $2 \perp\!\!\!\perp \{1, 3\}$. Hence $\operatorname{codim}(\tilde{I}) = (d_1d_3 - 2)(d_2 - 2)$, so

$$\dim(\tilde{I}) = 2d_1d_3 + 2d_2 - 4.$$

Similarly, since $\tilde{J}$ is generated by the 2×2-minors of the $d_2 \times d_3$-matrix $N_+ = (p_{1jk+})$, then $\operatorname{codim}(\tilde{J}) = (d_2 - 1)(d_3 - 1)$, so

$$\dim(\tilde{J}) = d_1d_2d_3 - d_2d_3 + d_2 + d_3 - 1.$$

Recall that I, J, and $I + J$ are prime ideals, then [16, Thm. 11.12] implies the following

$$\begin{aligned} \dim(V(I)) &= \dim(V(\tilde{I})) + \mu_1, \\ \dim(V(J)) &= \dim(V(\tilde{J})) + \mu_2, \\ \dim(V(I + J)) &= \dim(V(Q_{20})) + \mu \end{aligned}$$

We have seen that $\mu_1 = 2$. Also, since $d_4 = 2$ then $\mu_2 = \dim(J) - \dim(\tilde{J}) = d_1d_2d_3 - d_2d_3 + d_2$, so $\mu_2 > \mu_1$. We conjecture that $\mu = 2$ which implies

$$\dim(Q_{20}) = \dim(I + J) - 2 = 2d_1d_3 + 2d_2 - d_3 - 3.$$

Let M_{i_0+} be the $d_2 \times d_3$-matrix (p_{i_0jk+}), then $M_+ = (M_{1+}M_{2+}\cdots M_{d_1+})$, and $M_{1+} = N_+$. A similar argument as for the ideal $I + J$ shows that the ideal $\tilde{J}$ lowers the dimension of $\tilde{I}$ by $d_3 - 1$. Hence

$$\dim(\tilde{I} + \tilde{J}) = 2d_1d_3 + 2d_2 - d_3 - 3 = \dim(Q_{20}).$$

Note that $\tilde{I} + \tilde{J} \subseteq I_{\mathrm{local}(G_{20})} \cap \mathbb{C}[D']$ and both ideals have the same dimension. One can check that $\tilde{I} + \tilde{J}$ is a radical ideal by Gröbner basis methods. In fact, if $<$ denotes the degree reverse lexicographic ordering, then the (quadratic) generators of $\tilde{J}$ and the (cubic) generators of $\tilde{I}$ form a Gröbner basis of $\tilde{I} + \tilde{J}$. Therefore, the initial ideal $\mathrm{in}_<(\tilde{I} + \tilde{J})$ is square–free, which implies that $\tilde{I} + \tilde{J}$ is a radical ideal, see [33, Proposition 5.3]. Moreover, a set–theoretic result as in [19] would imply that the ideal $Q_{20} = I_{\mathrm{local}(G_{20})} \cap \mathbb{C}[D']$ equals the sum

$$Q_{20} = I_{\mathrm{local}(G_4)} \cap \mathbb{C}[D'] + I_{2 \perp\!\!\!\perp \{1,3\}|4} \cap \mathbb{C}[D'],$$

because a subvariety of an irreducible variety X having the same dimension as X is equal to X.

Network 13*:* The prime ideal $I_{\mathrm{local}(G_{13})} = I_{1 \perp\!\!\!\perp \{2,4\}|3}$. We claim that the ideal $Q_{13} = I_{\mathrm{local}(G_{13})} \cap \mathbb{C}[D']$ is equal to the ideal I generated by the 2×2-minors of the d_3 matrices of the form (p_{ijk+}) where the rows are indexed by $i \in [d_1]$, the columns are indexed by $j \in [d_2]$, and k is fixed. We argue as for G_4 that $I \subset Q_{13}$. The ideal Q_{13} is prime as shown in Proposition 20. The ideal I is prime since it describes the join of d_3 varieties, each defined by the 2×2-minors of a matrix of indeterminates and hence irreducible, see [6]. Thus, to show equality between ideals, it suffices to show that $V(I) \subset V(Q_{13})$. Let $a \in V(I)$. Then a is a $d_1 \times d_2 \times d_3$-table, such that for each $k_0 \in [d_3]$ the slice $a_k = (p_{ijk_0})$ is a $d_1 \times d_2$-matrix of rank 1. Let A be the $d_1 \times d_2 d_4 \times d_3$-table defined by $A_{ijkl} = a_{ijk}/d_4$. Clearly, for each $k_0 \in [d_3]$ the slice $A_{k_0} = (p_{ijk_0 l})$ has rank 1, so $A \in V(P_{G_{13}})$. Thus $a = \pi(A) \in V(Q_{13})$.

Network 23*:* The binomial prime ideal $I_{local(G_{23})}$ is the sum of two prime ideals $I = I_{2 \perp\!\!\!\perp \{1,3\}|4}$ and $J = I_{local(G_{13})}$. Let $\tilde{I} = I \cap \mathbb{C}[D']$. We showed in G_{20} that if $d_4 = 2$, the ideal $\tilde{I}$ is generated by the 3×3-minors of the $d_1 d_3 \times d_2$-matrix $M_+ = (p_{ijk+})$ obtained by flattening the 3-dimensional table (p_{ijk+}) according to the relation $\{1,3\} \perp\!\!\!\perp 2$. Note that

$$\tilde{I} + \tilde{J} \subseteq I_{\mathrm{local}(G_{23})} \cap \mathbb{C}[D'].$$

Moreover, a similar argument as for G_{20} shows that the ideal $Q_{23} = I_{\mathrm{local}(G_{23})} \cap \mathbb{C}[D']$ equals the sum of ideals $I_{\mathrm{local}(G_{13})} \cap \mathbb{C}[D'] + I_{2 \perp\!\!\!\perp \{1,3\}|4} \cap \mathbb{C}[D']$. We sketch the proof below, giving formulas for the dimension of each of the previous ideals.

The ideal I is generated by the 2×2-minors of the d_4 matrices $M_l = (p_{ijkl})$ where the rows are indexed by $(i,k) \in [d_1] \times [d_3]$, the columns by $j \in [d_2]$ and l is fixed. Recall that $\dim(I) = d_1 d_3 d_4 + d_2 d_4 - d_4$. The ideal J is generated by the 2×2-minors of the d_3 matrices of the form $N_k = (p_{ijkl})$ where the rows are indexed by $i \in [d_1]$, the columns are indexed by $(j,l) \in [d_2] \times [d_4]$, and k is fixed. The codimension of J equals $d_3(d_1 - 1)(d_2 d_4 - 1)$, so

$$\dim(J) = d_2 d_3 d_4 + d_1 d_3 - d_3$$

For each k_0, l_0, let $M_{k_0 l_0}$ be the $d_1 \times d_2$-matrix $(p_{ijk_0l_0})$, then

$$M_l = \begin{pmatrix} M_{1l} \\ M_{2l} \\ \vdots \\ M_{d_1 l} \end{pmatrix} \quad \text{and} \quad N_k = (M_{11} M_{12} \cdots M_{1d_4}).$$

Thus the following $d_1 d_3 \times d_2 d_4$-matrices are equal

$$\begin{pmatrix} M_1 & \cdots & M_{d_4} \end{pmatrix} = \begin{pmatrix} N_1 \\ \vdots \\ N_{d_3} \end{pmatrix}$$

Hence, just as for G_{20}, the ideal I lowers the dimension of J by $d_4(d_2-1)(d_3-1)$. Thus,

$$\dim(I+J) = \dim(J) - d_4(d_2-1)(d_3-1) = d_1 d_3 + d_2 d_4 + d_3 d_4 - d_3 - d_4.$$

When $d_4 = 2$, then $\dim(I+J) = d_1 d_3 + 2d_2 + d_3 - 2$. Recall that $\dim(\tilde{I}) = 2d_1 d_3 + 2d_2 - 4$. Moreover, since $\tilde{J}$ is generated by the 2×2-minors of the d_3 matrices $N_{k+} = (p_{ijk+})$, then $\operatorname{codim}(\tilde{J}) = d_3(d_1-1)(d_2-1)$, so $\dim(\tilde{J}) = d_1 d_3 + d_2 d_3 - d_3$.

Recall that I, J, and $I+J$ are prime ideals, then [16, Thm. 11.12] implies the following

$$\begin{aligned} \dim(V(I)) &= \dim(V(\tilde{I})) + \mu_1, \\ \dim(V(J)) &= \dim(V(\tilde{J})) + \mu_2, \\ \dim(V(I+J)) &= \dim(V(Q_{23})) + \mu \end{aligned}$$

We have seen that $\mu_1 = 2$. Also, as $d_4 = 2$, $\mu_2 = \dim(J) - \dim(\tilde{J}) = d_2 d_3$, so $\mu_2 > \mu_1$. Just as for G_{20}, we conjecture that $\mu = \mu_1 = 2$ which implies

$$\dim(Q_{23}) = \dim(I+J) - 2 = d_1 d_3 + 2d_2 + d_3 - 4.$$

Observe that $M_+ = (N_{1+} N_{2+} \cdots N_{d_3+})$. So $\tilde{I}$ lowers the dimension of $\tilde{J}$ by $(d_2-2)(d_3-2)$. Hence

$$\dim(\tilde{I} + \tilde{J}) = d_1 d_3 + 2d_2 + d_3 - 4 = \dim(Q_{23}).$$

Hence, the ideal $Q_{23} = \tilde{I} + \tilde{J}$.

Network 25*:* We have the following equalities of prime ideals

$$I_{\text{local}(G_{25})} = I_{1 \perp\!\!\!\perp \{2,4\}|3} + I_{2 \perp\!\!\!\perp \{1,4\}|3} = I_{\text{local}(G_{13})} + I_{2 \perp\!\!\!\perp \{1,4\}|3}.$$

The ideal $I = I_{\text{local}(G_{13})} \cap \mathbb{C}[D']$ is equal to the ideal generated by the 2×2-minors of the d_3 matrices of the form (p_{ijk+}) where the rows are indexed by $i \in [d_1]$ and the columns are indexed by $j \in [d_2]$ and k is fixed. Similarly, the ideal $J = I_{2 \perp\!\!\!\perp \{1,4\}|3} \cap \mathbb{C}[D']$ equals the ideal generated by the 2×2-minors of the d_3 matrices of the form (p_{ijk+}) where the rows are indexed by $j \in [d_2]$ and the columns are indexed by $i \in [d_1]$ and k is fixed. Note that the matrix which defines J is the transpose of the matrix that defines I, so $I = J$. The ideal $Q_{25} = I_{\mathrm{local}(G_{25})} \cap \mathbb{C}[D']$ is prime as shown in Proposition 20. Recall also that the ideal $I \subseteq Q_{25}$ is prime. Thus, to show equality between ideals, it suffices to show that $V(I) \subset V(Q_{25})$. Let $a \in V(I)$, then a is a $d_1 \times d_2 \times d_3$-table, such that for each $k_0 \in [d_3]$ the slice $a_k = (p_{ijk_0})$ is a $d_1 \times d_2$-matrix of rank 1. Let A be the 4-dimensional $d_1 \times d_2 \times d_3 \times d_4$-table defined by $A_{ijkl} = a_{ijk}/d_4$. For each $k_0 \in [d_3]$ consider the slice $A_{k_0} = (p_{ijk_0l})$. The $d_1 \times d_2d_4$-matrix (p_{ijk_0l}) obtained by flattening the 3-dimensional table A_{k_0} according to $1 \perp\!\!\!\perp \{2,4\}$ has rank 1. Also, the $d_2 \times d_1d_4$-matrix (p_{ijk_0l}) obtained by flattening the 3-dimensional table A_{k_0} according to $2 \perp\!\!\!\perp \{1,4\}$ has rank 1. Thus $A \in V(P_{G_{25}})$ which implies $a = \pi(A) \in V(Q_{25})$.

Network 14*:* The ideal $I_{\mathrm{local}(G_{14})}$ is a prime ideal equal to $I_{1 \perp\!\!\!\perp \{3,4\}|2}$. It is generated by the 2×2-minors of the d_2 slices $M_j = (p_{ijkl})$, where the rows are labeled by $i \in [d_1]$ and the columns are labeled by the pairs $(k,l) \in [d_3] \times [d_4]$. In Theorem 12 we proved that $I_{local(G_{14})}$ is a prime ideal and hence equal to $P_{G_{14}}$.

The ideal $Q_{14} = I_{local(G_{14})} \cap \mathbb{C}[D']$ is equal to the ideal I generated by the 2×2-minors of the d_2 slices $N_j = (p_{ijk+})$ where the rows are indexed by $i \in [d_1]$ and the columns are indexed by $k \in [d_3]$. Each column of this matrix is obtained by taking the sum of the corresponding d_4 columns of M_j. Using the same argument as for G_4 we have that $I \subset Q_{14}$.

The ideal Q_{14} is prime as shown in Proposition 20. The ideal I is prime, since the variety $V(I)$ is irreducible. To see this observe that $V(I)$ is the join of d_2 disjoint irreducible varieties, each described by the 2×2-minors of a generic matrix, the result follows from [16, Theorem 11.14]. Hence, to show equality between ideals, it suffices to show that $V(I) \subset V(Q_{14})$. Let $a \in V(I)$, then a is a $d_1 \times d_2 \times d_3$-table such that each slice $a_{j_0} = (a_{ij_0k})$ has rank 1. Let A be the $d_1 \times d_2 \times d_3 \times d_4$-table defined by $A_{ijkl} = a_{ijk}/d_4$. For each slice A_{j_0} consider the flattening $A'_{j_0} = (A_{ij_0kl})$ where the rows are indexed by $i \in [d_1]$ and the columns are indexed by pairs $(k,l) \in [d_3] \times [d_4]$. Clearly, A'_{j_0} has rank 1, so $A \in V(P_{G_4})$. Thus $a = \pi(A) \in V(Q_{14})$.

Network 18*:* The prime ideal $I_{\mathrm{global}(G_{18})}$ is equal to the sum $I_{1 \perp\!\!\!\perp \{3,4\}|2} + I_{\{1,2\} \perp\!\!\!\perp 4|3}$. The first ideal equals $I_{\mathrm{local}(G_{14})}$. We will show that $I_{\{1,2\} \perp\!\!\!\perp 4|3} \cap \mathbb{C}[D'] = 0$. It is enough to show that every $d_1d_2 \times d_3$-matrix $a = (a_{ijk})$ is the projection of a table in $V(I_{\{1,2\} \perp\!\!\!\perp 4|3})$. Let A be the $d_1d_2 \times d_3d_4$-matrix defined by $A_{ijkl} = a_{ijk}/d_4$. Hence for each $k_0 \in [d_3]$ the $d_1d_2 \times d_4$-matrix (A_{ijk_0l}) has rank 1. In fact, the d_4 columns of this matrix are equal. Thus $A \in V(I_{\{1,2\} \perp\!\!\!\perp 4|3})$, so $Q_{G_{18}} = Q_{G_{14}}$.

Network 21*:* The ideal $I = I_{\mathrm{local}(G_{21})} = I_{\mathrm{global}(G_{21})}$ is the sum of two ideals, namely $I_1 = I_{1 \perp\!\!\!\perp \{3,4\}|2}$ and $J = I_{3 \perp\!\!\!\perp 4}$. This ideal is not radical in general as seen in Theorem 12. A closer look at the proof of Theorem 9 reveals that the distinguished component $P_{G_{21}}$ equals

$$P_{G_{21}} = (J + I_1) : p_1^{\infty} = I_{\mathrm{local}(G_{21})} : p_1^{\infty}.$$

Where p_1 is the product of p_{+jkl} for all $j \in [d_2], k \in [d_3]$ and $l \in [d_4]$. Hence $I_{local(G_{21})} = P_{G_{21}} \cap (I, p_1^{e_1})$ for some e_1. So we have the following equality of varieties

$$V(I) = V(P_{G_{21}}) \cup V(I, p_1) \tag{8.5}$$

$$\overline{\pi(V(I))} = V(Q_{21}) \cup \overline{\pi(V(I, p_1))} \tag{8.6}$$

We know that $I_{3 \perp\!\!\!\perp 4} \cap \mathbb{C}[D'] = 0$. Hence following a similar argument as in G_9 we see that $I \cap \mathbb{C}[D'] = I_1 \cap \mathbb{C}[D'] = Q_{14}$. So the equation (8.6) can be rewritten as

$$V(Q_{14}) = V(Q_{21}) \cup \overline{\pi(V(I, p_1))}.$$

Therefore, $V(Q_{21})$ is a subvariety of the irreducible variety $V(Q_{14})$. Moreover, we conjecture that $V(Q_{21}) = V(Q_{14})$. For this we need to show that $\dim(V(Q_{14})) = \dim(V(Q_{21}))$. Since both ideals are prime, this would imply $Q_{21} = Q_{14}$.

Network 19*:* The ideal $I_{\text{global}(G_{19})}$ is a prime ideal equal to $I_{2 \perp\!\!\!\perp \{1,3,4\}}$. It is generated by the 2×2-minors of the $d_2 \times d_1 d_3 d_4$-matrix $M = (p_{ijkl})$, where the rows are labeled by $j \in [d_2]$ and the columns are labeled by the tuples $(i, k, l) \in [d_1] \times [d_3] \times [d_4]$. Following a similar argument as for the graph G_4 we see that the ideal $Q_{19} = I_{\text{global}(G_{19})} \cap \mathbb{C}[D']$ is generated by the 2×2-minors of the $d_2 \times d_1 d_3$-matrix $M_+ = (p_{ijk+})$ where the rows are indexed by $j \in [d_2]$ and the columns are indexed by pairs $(i, k) \in [d_1] \times [d_3]$. Each column of this matrix is obtained by taking the sum of the corresponding d_4 columns of M.

Network 22*:* The ideal $I_{\text{global}(G_{22})}$ is a prime ideal equal to $I = I_{2 \perp\!\!\!\perp \{1,3,4\}} + I_{4 \perp\!\!\!\perp \{2,3\}}$. From our analysis of graph G_5, we know that $I_{4 \perp\!\!\!\perp \{2,3\}} \cap \mathbb{C}[D'] = 0$. Moreover, following a similar argument we see that $Q_{22} = I_{\text{global}(G_{22})} \cap \mathbb{C}[D'] = Q_{19}$.

Network 26*:* The ideal $I_{\text{global}(G_{26})}$ is a prime ideal equal to $I = I_{\{1,3\} \perp\!\!\!\perp \{2,4\}}$. It is generated by the 2×2-minors of the $d_2 d_4 \times d_1 d_3$-matrix $M = (p_{ijkl})$, where the rows are labeled by $(j, l) \in [d_2] \times [d_4]$, and the columns are labeled by $(i, k) \in [d_1] \times [d_3]$. Following a similar argument as for the graph G_4 we see that the ideal $Q_{26} = I_{\text{global}(G_{26})} \cap \mathbb{C}[D']$ is generated by the 2×2-minors of the $d_2 \times d_1 d_3$-matrix $N = (p_{ijk+})$ where the rows are indexed by $j \in [d_2]$ and the columns are indexed by pairs $(i, k) \in [d_1] \times [d_3]$. Each row of this matrix is obtained by taking the sum of the corresponding d_4 rows of M. Hence $Q_{26} = Q_{19}$.

Network 28*:* The ideal $I_{\text{global}(G_{28})}$ is a prime ideal equal to $I_{2 \perp\!\!\!\perp \{1,3,4\}} + I_{\{1,2\} \perp\!\!\!\perp 4|3}$. From our analysis of graph G_{18}, we know that $I_{\{1,2\} \perp\!\!\!\perp 4|3} \cap \mathbb{C}[D'] = 0$. Moreover, following a similar argument we see that $Q_{28} = I_{\text{global}(G_{28})} \cap \mathbb{C}[D'] = Q_{19}$.

Network 27*:* The prime ideal $I_{\text{local}(G_{27})}$ is equal to the sum $I_{3 \perp\!\!\!\perp \{1,2,4\}} + I_{2 \perp\!\!\!\perp \{1,3\}|4}$. Proceeding in a similar way as for graph G_4, we see that $I = I_{3 \perp\!\!\!\perp \{1,2,4\}} \cap \mathbb{C}[D']$ is generated by the 2×2-minors of the matrix (p_{ijk+}) where the rows are indexed by $k \in [d_3]$ and the columns are indexed by $(i, j) \in [d_1] \times [d_2]$. We have seen that if $d_4 = 2$ then the prime ideal $J = I_{2 \perp\!\!\!\perp \{1,3\}|4} \cap \mathbb{C}[D']$ is generated by the 3×3-minors of the two dimensional table (p_{ijk+}) where the rows are indexed by $j \in [d_2]$ and the columns are indexed by the pairs $(i,k) \in [d_1] \times [d_3]$. Observe that $I + J = I_{3 \perp\!\!\!\perp \{1,2,4\}}) \cap \mathbb{C}[D'] + I_{2 \perp\!\!\!\perp \{1,3\}|4} \cap \mathbb{C}[D'] \subseteq I_{\mathrm{local}(G_{27})} \cap \mathbb{C}[D']$. Moreover, following a similar procedure as for graph G_{20}, we conjecture that the ideal $I_{\mathrm{local}(G_{27})} \cap \mathbb{C}[D']$ equals the sum of ideals $I_{3 \perp\!\!\!\perp \{1,2,4\}} \cap \mathbb{C}[D'] + I_{2 \perp\!\!\!\perp \{1,3\}|4} \cap \mathbb{C}[D']$.

Network 29*:* The prime ideal $I_{\mathrm{local}(G_{29})}$ is equal to the sum of ideals $I_{2 \perp\!\!\!\perp \{1,3,4\}} + I_{3 \perp\!\!\!\perp \{1,2,4\}}$. Proceeding in a similar way as for graph G_4, we see that $I = I_{2 \perp\!\!\!\perp \{1,3,4\}} \cap \mathbb{C}[D']$ is generated by the 2×2-minors of the matrix (p_{ijk+}) where the rows are indexed by $j \in [d_2]$ and the columns are indexed by $(i,k) \in [d_1] \times [d_3]$. Moreover, following a similar argument as in [19], the ideal $I_{\mathrm{local}(G_{29})} \cap \mathbb{C}[D']$ equals the sum of ideals $I_{3 \perp\!\!\!\perp \{1,2,4\}} \cap \mathbb{C}[D'] + I_{2 \perp\!\!\!\perp \{1,3,4\}} \cap \mathbb{C}[D']$.

Network 30*:* The prime ideal $I_{\mathrm{local}(G_{30})}$ is equal to the sum of ideals $I_{2 \perp\!\!\!\perp \{1,3,4\}} + I_{3 \perp\!\!\!\perp \{1,2,4\}} + I_{4 \perp\!\!\!\perp \{1,2,3\}}$. Proceeding in a similar way as for graph G_5, we see that $I = I_{4 \perp\!\!\!\perp \{1,2,3\}} \cap \mathbb{C}[D'] = 0$. This implies that

$$I_{\mathrm{local}(G_{30})} \cap \mathbb{C}[D'] = \big(I_{2 \perp\!\!\!\perp \{1,3,4\}} + I_{3 \perp\!\!\!\perp \{1,2,4\}}\big) \cap \mathbb{C}[D'].$$

Moreover, following a similar argument as in [19], we conclude that

$$I_{\mathrm{local}(G_{30})} \cap \mathbb{C}[D'] = I_{3 \perp\!\!\!\perp \{1,2,4\}} \cap \mathbb{C}[D'] + I_{2 \perp\!\!\!\perp \{1,3,4\}} \cap \mathbb{C}[D'].$$

Thus, $Q_{30} = Q_{29}$.

Network 15*:* The ideal $I_{\mathrm{local}(G_{15})} = I_{\mathrm{global}(G_{15})}$ is the sum of two ideals, namely $I = I_{1 \perp\!\!\!\perp 4|\{2,3\}}$ and $J = I_{2 \perp\!\!\!\perp 3|4}$. This ideal is not radical in general as seen in Theorem 12. Hence $I_{\mathrm{local}(G_{15})} = P_{G_{15}} \cap L$ for some ideal L. So we have the following equality of varieties

$$V(I_{\mathrm{local}(G_{15})}) = V(P_{G_{15}}) \cup V(L) \tag{8.7}$$

$$\overline{\pi(V(I_{local(G_{15})}))} = V(Q_{15}) \cup \overline{\pi(V(L))} \tag{8.8}$$

Note that $I = I_{\mathrm{local}(G_7)}$, so $I \cap \mathbb{C}[D'] = 0$. Moreover, if we assume $d_4 = 2$, then a similar argument as for G_{11} shows that $I_{\mathrm{local}(G_{15})} \cap \mathbb{C}[D'] = J \cap \mathbb{C}[D']$, and $J \cap \mathbb{C}[D'] = Q_2$ is generated by the 3×3-minors of the $d_2 \times d_3$-matrix (p_{+jk+}). Thus $V(Q_2) = V(Q_{15}) \cup \overline{\pi(V(L))}$, that is, $V(Q_{15})$ is a subvariety of the irreducible variety $V(Q_2)$. Similar to G_{21}, we conjecture that in fact $Q_{15} = Q_2$.

Network 17*:* The ideal $I_{\mathrm{local}(G_{17})} = I_{\mathrm{global}(G_{17})} = I + J$, where $I = I_{1 \perp\!\!\!\perp 3|\{2,4\}}$ and $J = I_{2 \perp\!\!\!\perp 4|3}$. This ideal is not radical in general as seen in Theorem 12. Hence $I_{\mathrm{local}(G_{17})} = P_{G_{17}} \cap L$ for some ideal L. So we have the following equality of varieties

$$V(I_{\mathrm{local}(G_{17})}) = V(P_{G_{17}}) \cup V(L) \tag{8.9}$$

$$\overline{\pi(V(I_{local(G_{17})}))} = V(Q_{17}) \cup \overline{\pi(V(L))} \tag{8.10}$$

Note that $J = I_{\mathrm{local}(G_3)}$, so $J \cap \mathbb{C}[D'] = 0$. Also $I = I_{\mathrm{local}(G_8)}$ and we have given a generating set for Q_8 for the case $d_4 = 2$. We claim that $I_{\mathrm{local}(G_{17})} \cap \mathbb{C}[D'] = Q_8$. Let $a \in V(Q_8)$, for each

$j_0 \in [d_2]$ the slice a_{j_0} of a, that is, the $d_1 \times d_3$-matrix a_{ij_0k} has rank 1. Define a $d_1 \times d_2 \times d_3 \times d_4$ table A as $A_{ijkl} = a_{ijk}/2$. Then for each $(j_0, l_0) \in [d_2] \times [d_4]$ the $d_1 \times d_3$ slice $A_{j_0l_0} = (A_{ij_0kl_0})$ has rank 1, so $A \in V(I)$. Note that $A \in V(J)$, thus $a = \pi(A) \in \pi(V(I_{\mathrm{local}(G_{17})}))$. Hence $V(Q_{17})$ is an irreducible subvariety of the irreducible variety Q_8. But, opposed to all the previous varieties, in general $V(Q_{17})$ will be a proper subvariety of $V(Q_8)$. We have a conjecture for the case $d_1 = d_4 = 2$, note that for this case $Q_8 = 0$, that is, $V(Q_8) = \mathbb{C}[D']$. To simplify notation, we will set $p_{ijk} = p_{ijk+}$.

The ideal Q_{17} is generated by $\binom{d_2}{2}\binom{d_3}{3}$ *sextic* polynomials constructed as follows. For each $j_0 \in [d_2]$, let M_{j_0} be the $d_1 \times d_3$-matrix $M_{j_0} = (p_{ij_0k})$. Each $j_1, j_2 \in [d_2]$, $j_1 \neq j_2$ specify two matrices M_{j_1} and M_{j_2}. Also, each triplet k_1, k_2, k_3 of distinct elements in $[d_3]$ specify three columns on each $2 \times d_3$-matrix M_{j_1} and M_{j_2}. So we get the following two 2×3 submatrices N_{j_1} and N_{j_2}

$$\begin{pmatrix} p_{1j_1k_1} & p_{1j_1k_2} & p_{1j_1k_3} \\ p_{2j_1k_1} & p_{2j_1k_2} & p_{2j_1k_3} \end{pmatrix} \text{ and } \begin{pmatrix} p_{1j_2k_1} & p_{1j_2k_2} & p_{1j_2k_3} \\ p_{2j_2k_1} & p_{2j_2k_2} & p_{2j_2k_3} \end{pmatrix}$$

The irreducible sextic polynomial arising from these two submatrices is given by the following alternating sum

$$p_{+j_1k_1}U_1V_1 - p_{+j_1k_2}U_2V_2 + p_{+j_1k_3}U_3V_3.$$

Where the polynomial U_s is the determinant of the 2×2-submatrix of N_{j_1} obtained by eliminating the s-th column. The polynomial V_s is the determinant of the 2×2-matrix N'_{j_2} where the first column of N'_{j_2} equals the s-th column of N_{j_2} and the second column of N'_{j_2} is the product of the remaining two columns of N_{j_2}. □

Bibliography

[1] I. Barradas and F. Solis: Nonnegative rank, *International Mathematical Journal* **1** (2002) 601–610.

[2] M. Catalano-Johnson: The homogeneous ideals of higher secant varieties, *Journal of Pure and Applied Algebra* **158** (2001) 123–129.

[3] M.V. Catalisano, A.V. Geramita and A. Gimigliano: Ranks of tensors, secant varieties and fat points, *Linear Algebra Appl.*, to appear.

[4] D. Cox, J. Little and D. O'Shea: *Ideals, Varieties and Algorithms*, Springer Undergraduate Texts in Mathematics, Second Edition, 1997.

[5] W. Decker, G.-M. Greuel and G. Pfister: Primary decomposition: algorithms and comparisons. *Algorithmic Algebra and Number Theory (Heidelberg, 1997)*, 187–220, Springer, Berlin, 1999.

[6] C. DeConcini, D. Eisenbud and C. Procesi: Young Diagrams and Determinantal Varieties. *Inventiones Mathematicae* **56** (1980) 129-16

[7] D. Eisenbud and B. Sturmfels: Binomial ideals, *Duke Math. Journal* **84** (1996) 1–45.

[8] L. D. Garcia, M. Stillman and B. Sturmfels: Algebraic Geometry of Bayesian Networks, *Journal of Symbolic Computation* (2004) pages to appear.

[9] L. D. Garcia: Algebraic Statistics in Model Selection, submitted.

[10] D. Geiger and C. Meek: Graphical models and exponential families, *Proceedings of the Fourteenth Annual Conference on Uncertainty in Artificial Intelligence* (1998) 156–165, San Francisco, CA. Morgan Kaufmann Publishers.

[11] D. Geiger, D. Heckerman, H. King and C. Meek: Stratified exponential families: graphical models and model selection, *Annals of Statistics* **29** (2001) 505–529.

[12] D. Geiger, C. Meek and B. Sturmfels: On the toric algebra of graphical models, Manuscript, 2002.

[13] L. Goodman: Explanatory latent structure analysis using both identifiable and unidentifiable models, *Biometrika* **61** (1974) 215–231.

[14] D. Grayson and M. Stillman: Macaulay2: a system for computation in algebraic geometry and commutative algebra, 1996, available at `http://www.math.uiuc.edu/Macaulay2/`.

[15] G.-M. Greuel, G. Pfister and H. Schönemann: Singular 2.0: A computer algebra system for polynomial computations, University of Kaiserslautern, 2001, `http://www.singular.uni-kl.de`.

[16] J. Harris: *Algebraic Geometry: A First Course*, Springer Graduate Texts in Mathematics, 1992.

[17] D. Haughton: On the choice of a model to fit data from an exponential family, *Annals of Statistics* **16** (1988) 342–555.

[18] T. Kočka and N. L. Zhang: Dimension correction for hierarchical latent class models, *Proceedings of the Eighteenth Conference on Uncertainty in Artificial Intelligence (UAI-02)* (2002).

[19] J. M. Landsberg and L. Manivel: On the ideals of secant varieties of Segre varieties, sumbitted (2003).

[20] S. L. Lauritzen: *Graphical Models*, Oxford University Press, 1996.

[21] F. Matúš: Conditional independences among four random variables. III. Final conclusion. *Combin. Probab. Comput.* **8** (1999) 269–276.

[22] J. Pearl: *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann, San Mateo, CA, 1988

[23] G. Pistone, E. Riccomagno and H. Wynn: *Algebraic Statistics: Computational Commutative Algebra in Statistics*, Chapman and Hall, Boca Raton, 2001.

[24] D. Rusakov and D. Geiger: Asymptotic model selection for naive Bayesian networks, *Proceedings of the Eighteenth Annual Conference on Uncertainty in Artificial Intelligence (UAI-02)* (2002).

[25] D. Rusakov and D. Geiger: Automated analytic asymptotic evaluation of the marginal likelihood for latent models, *Proceedings of the Nineteenth Annual Conference on Uncertainty in Artificial Intelligence (UAI-03)* (2003).

[26] R. Settimi and J.Q. Smith: Geometry, moments and conditional independence trees with hidden variables, *Annals of Statistics* **28** (2000) 1179–1205.

[27] R. Settimi and J.Q. Smith: On the geometry of Bayesian graphical models with hidden variables, In *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence*, Morgan Kaufmann Publishers, San Francisco, CA (1998) 472–479.

[28] G. Schwarz: Estimating the dimension of a model, *Annals of Statistics* **6(2)** (1978) 461–464.

[29] T. Shimoyama and K. Yokoyama: Localization and primary decomposition of polynomial ideals. *J. Symbolic Comput.* **22** (1996) 247–277.

[30] V. Strassen: Rank and optimal computation of generic tensors, *Linear Algebra and its Applications* **52/53** (1983) 645–685.

[31] M. Studený: *On Mathematical Description of Probabilistic Conditional Independence Structures*, Dr.Sc. thesis, Prague, May 2001, posted at `http://www.utia.cas.cz/user_data/studeny/studeny_home.html`.

[32] B. Sturmfels: *Gröbner Bases and Convex Polytopes*, American Mathematical Society, University Lectures Series, No. 8, Providence, Rhode Island, 1996.

[33] B. Sturmfels, *Solving Systems of Polynomial Equations*, CBMS Lectures Series, American Mathematical Society, 2002.

Appendix A

Techniques for Primary Decomposition

The ideals in this paper present a challenge for present day computer algebra systems. Their large number of variables (e.g. 32 in Chapter 6), combined with the sometimes long polynomials which arise are difficult to handle with built-in primary decomposition algorithms. Even the standard implementations of factorization of multivariate polynomials have difficulty with some of the long polynomials. This is only a problem with current implementations, which are generally not optimized for large numbers of variables.

For the computations performed in Chapters 5 and 6, it was necessary to write special code (in `Macaulay2`) in order to compute the components and primary decompositions of these ideals. We also have some code in `Macaulay2` or `Singular` for generating the ideals $I_{\text{local}(G)}$ or $I_{\text{global}(G)}$ from the graph G and the integers $d_1, d_2, \ldots, d_n$. In this appendix we indicate some techniques and tricks that were used to compute with these ideals.

The first modification which simplifies the problems dramatically is to change coordinates so that the indeterminates are $p_{2u_2\cdots u_n}$ and $p_{+u_2\cdots u_n}$, instead of $p_{u_1\cdots u_n}$. This change of variables sometimes takes a Markov ideal into a binomial ideal, which is generally much simpler to compute with. Computing any one Gröbner basis, ideal quotient, or intersection of our ideals is not too difficult. Therefore, our algorithms make use of these operations. All ideals examined in this project have the property that every component is rational. The distinguished component $\ker(\Phi)$ is more complicated than any of the other components, in terms of the number of generators and their degrees, and it cannot be computed by implicitization.

The first problem is to decide whether an ideal is prime (i.e. whether it equals the unknown ideal $\ker(\Phi)$). There are several known methods for deciding primality (see [5] for a nice exposition). The standard method is to reduce to a zero-dimensional problem. This entails either a generic change of coordinates, or factorization over extension fields. We found that the current implementations of these methods fail for the majority of the 301 examples in Chapter 6. The technique that did work for us is to search for birational projections. This either produces a zero divisor, or a proof that the ideal is prime. It can sometimes be used to count the components (both minimal and embedded), without actually producing the components.

The following result is proved by localizing with respect to powers of g. This defines a *birational projection* $(x_1, x_2, \ldots, x_n) \mapsto (x_2, \ldots, x_n)$ for J.

Proposition 25. *Let $J \subset \mathbb{R}[x_1, \ldots, x_n]$ be an ideal, containing a polynomial $f = gx_1 + h$, with g, h not involving x_1, and g a non-zero divisor modulo J. Let $J_1 = J \cap \mathbb{R}[x_2, \ldots, x_n]$ be the elimination ideal. Then*

(a) $J = \big(\langle J_1, gx_1 + h\rangle : g^\infty\big)$,

(b) J is prime if and only if J_1 is prime.

(c) J is primary if and only if J_1 is primary.

(d) Any irredundant primary decomposition of J_1 lifts to an irredundant primary decomposition of J.

Our algorithm to check primality starts by searching for variables which occur linearly, checking that its lead coefficient is not a zero divisor and then eliminating that variable as in Proposition 25. In almost all of the Markov ideals that we have studied, iterative use of this technique proves or disproves primality. A priori, one might not be able to find a birational projection at all, but this never happened for any of our examples.

The second problem is to compute the minimal primes or the primary decomposition. Finding the minimal primes is the first step in computing a primary decomposition, using the technique of [29], which is implemented in several computer algebra systems, including `Macaulay2`. Here, we have not found a single method that always works best. One method that worked in most cases is based on splitting the ideal into two parts. Given an ideal I, if there is an element f of its Gröbner basis which factors as $f = f_1 f_2$, then

$$\sqrt{I} \quad = \sqrt{\langle I, f_1\rangle} \;\cap\; \sqrt{\langle I, f_2\rangle : f_1^\infty}.$$

We keep a list of ideals whose intersection has the same radical as I. We process this list of ideals by ascending order on its codimension. For each ideal, we keep a list of the elements that we have inverted by so far (e.g. f_1 in the ideal $\big(\langle I, f_2\rangle : f_1^\infty\big)$) and saturate at each step with these elements.

If there is no element which factors, then we search for a variable to birationally project away from, as in Proposition 25. If its lead coefficient g is a zero divisor, use this element to split the ideal via

$$\sqrt{I} \quad = \quad \sqrt{I : g} \;\cap\; \sqrt{\langle I, g\rangle}.$$

As we go, we only process ideals which do not contain the intersection of all known components computed so far.

If we cannot find any birational projection or reducible polynomial, then we have no choice but to decompose the ideal using the built-in routines, which are based on characteristic sets. However, in none of the examples of this paper was this final step reached. This method works in a reasonable amount of time for all but about 10 to 15 of the 301 ideals in Chapter 6.

Appendix B

A Singular Package for Algebraic Statistics

In this chapter we include the code written in the computer algebra system *Singular* needed to perform most of the computations in this thesis. This package is in process of being included in the main distribution of Singular. Due to space limitations, we do not included all the functions that comprise the package, but only the main functions. Moreover, together with Michael Stillman we have also implemented this package in the computer algebra system *Macaulay2*.

```
// Luis David Garcia, last modified: 03.12.04
///////////////////////////////////////////////////////////////////////
version="$Id: Markov-relations.lib,v 1.0.0.0 2004/03/12 13:04:15
          levandov Exp $";
category="Algebraic Statistics";
info="
LIBRARY:  Markov.lib   Markov Relations for Bayesian Networks

PROCEDURES:
 info(I)               ideal, displays codim, degree, #mingens.

 bnet(n,u)             int, intvec, creates a nxn-matrix whose lower
                       triangle is given by the entries of u. Bayesian
                       net in topological order.

 pairMarkov(m)         intmat, computes the pairwise Markov relations
                       of a Bayesian network.

 localMarkov(m)        intmat, computes the local Markov relations
```

```
                          of a Bayesian network.


 globalMarkov(m)          intmat, computes the global Markov relations
                          of a Bayesian network.

 probring(d,#)            intvec, list, creates the ring
                          QQ[p11..1, ...,pd1d2..dn] where
                          d = {d1,d2,..,dn} and # is a list of options
                          {name-of-ring <pdR>, name-of-variable <p>,
                          ordering <dp>}.

 Prob(l,d)                list, intvec, computes the marginal p(l) where
                          l is an instance of the random variables of the
                          form l = IND,IND,1,1, where IND is a string
                          denoting +.

 MarkovIdeal(M,d)         list, intvec, computes the Markov ideal I_{M}
                          for a list of independent statements M. To get
                          an example of the usage of this package type
                          example MarkovIdeal;

 torideal(I,d)            ideal, intvec, computes the distinguished
                          component of the ideal I.
 ";

LIB "general.lib";
LIB "elim.lib";
LIB "primdec.lib";
LIB "standard.lib";
LIB "presolve.lib";
/////////////////////////////////////////////////////////////////////

// GENERAL FUNCTIONS

// General function to return information about the ideal

proc info (ideal I)
{
  int c = nvars(basering) - dim(I);
```

```
  int d = degree(I);
  int s = size(minbase(I));
  list l = c,d,s;
  return(l);
}

/* Definition of the matrix corresponding to the Bayesian Network.
   This procedure assumes that the DAG is in topological order
   n > n-1 > ... > 1 */

proc bnet (int n, intvec u)
{
  int i,j,k;
  intmat m[n][n];
  for (i=2; i<=n; i++)
    {
      for (j=1; j<i; j++)
        {
          k++;
          m[i,j] = u[k];
        }
    }
  return(m);
}

/////////////////////////////////////////////////////////////////
/* Definition of the Pairwise Markov Relations */
/////////////////////////////////////////////////////////////////

proc nondec (int v, intmat m)
{
  int n = ncols(m);
  list l,s, visited;
  int i,j,k;
  s = v;
  for (i=1; i<=n; i++)
    {
      visited[i] = 0;
    }
  while (size(s) != 0)
    {
      k = s[1];
```

```
      s = delete(s,1);
      visited[k] = 1;
      for (j=1; j<=n; j++)
        {
          if (m[k,j] == 1)
            {
              if (visited[j] == 0)
                {
                  s = insert(s,j);
                }
            }
        }
    }
  for (i=1; i<=n; i++)
    {
      if (visited[i] == 0)
        {
          l = insert(l,i,size(l));
        }
    }
  return(l);
}

proc pairMarkov (intmat m)
{
  int n = ncols(m);
  list nd, l, s, e, e2;
  int i, j, k, c, check;
  for (i=1; i<=n; i++)
    {
      nd = nondec(i,m);
      if (size(nd) != 0)
        {
          for (j=1; j<=size(nd); j++)
            {
              if (m[nd[j],i] == 0)
                {
                  s = list();
                  s = insert(s,list(i));
                  s = insert(s, list(nd[j]), size(s));
                  s = insert(s, delete(nd,j), size(s));
                  if (nd[j] > i)
```

```
          {
            l = insert(l,s,size(l));
          }
        else
          {
            c = nd[j];
            e = nondec(c,m);
            if (size(e) == 0)
              {
                l = insert(l,s,size(l));
              }
            else
              {
                for (k=1; k<=size(e); k++)
                  {
                    if (e[k] == i)
                      {
                        check = 1;
                        break;
                      }
                  }
                if (check == 0)
                  {
                    l = insert(l,s,size(l));
                  }
                else
                  {
                    e = delete(e,k);
                    e2 = delete(nd,j);
                    if (size(e) != size(e2))
                      {
                        l = insert(l,s,size(l));
                      }
                    else
                      {
                        if (size(e) != 0)
                          {
                            check = 0;
                            for (k=1; k<=size(e); k++)
                              {
                                if (e[k] != e2[k])
                                  {
```

```
                                      check = 1;
                                      break;
                                    }
                                }
                              if (check == 1)
                                {
                                  l = insert(l,s,size(l));
                                }
                            }
                        }
                    }
                }
            }
          }
        }
      }
    }
  return(l);
}

//////////////////////////////////////////////////////////////////
/* Definition of the Local Markov Relations */
//////////////////////////////////////////////////////////////////

proc parent (int v, intmat m)
{
  int n = ncols(m);
  list l;
  int i;
  for (i=1; i<=n; i++)
    {
      if (m[i,v] == 1)
        {
          l = insert(l,i,size(l));
        }
    }
  return(l);
}

proc nondecminusparents (int v, intmat m)
{
  int n = ncols(m);
```

```
  list l,s, visited;
  int i,j,k;
  s = v;
  for (i=1; i<=n; i++)
    {
      visited[i] = 0;
    }
  while (size(s) != 0)
    {
      k = s[1];
      s = delete(s,1);
      visited[k] = 1;
      for (j=1; j<=n; j++)
        {
          if (m[k,j] == 1)
            {
              if (visited[j] == 0)
                {
                  s = insert(s,j);
                }
            }
        }
    }
  for (i=1; i<=n; i++)
    {
      if (visited[i] == 0 and m[i,v] == 0)
        {
          l = insert(l,i,size(l));
        }
    }
  return(l);
}

proc localMarkov (intmat m)
{
  int n = ncols(m);
  list pa, s, l, nd, e;
  int i,c,check;
  for (i=1; i<=n; i++)
    {
      s = list();
      s = insert(s,list(i));
```

```
nd = nondecminusparents(i,m);
if (size(nd) != 0)
  {
    pa = parent(i,m);
    s = insert(s,nd,size(s));
    s = insert(s,pa,size(s));
    if (size(nd) > 1)
      {
        l = insert(l,s,size(l));
      }
    else
      {
        c = nd[1];
        if (c > i)
          {
            l = insert(l,s,size(l));
          }
        else
          {
            e = parent(c,m);
            if (size(e) != size(pa))
              {
                l = insert(l,s,size(l));
              }
            else
              {
                if (size(e) != 0)
                  {
                    check = 1;
                    while (check < size(e) and e[check] == pa[check])
                      {
                        check++;
                      }
                    if (check != size(e))
                      {
                        l = insert(l,s,size(l));
                      }
                  }
                e = nondecminusparents(c,m);
                if (size(e) != 1)
                  {
                    l = insert(l,s,size(l));
```

```
                        }
                      else
                        {
                          if (e[1] != i)
                            {
                              l = insert(l,s,size(l));
                            }
                        }
                    }
                }
            }
        }
    }
  return(l);
}

///////////////////////////////////////////////////////////////////
/* Construction of the Global Markov Relations */
///////////////////////////////////////////////////////////////////

proc subset (int k, list X)
{
  if (size(X) == 0)
    {
      return(list());
    }
  int n = size(X);
  intvec bin;
  int q,i;
  for (i=1; i<=n; i++)
    {
      bin[i] = 0;
    }
  q = k;
  i = 1;
  while (q != 0)
    {
      bin[i] = q%2;
      q = q / 2;
      i++;
    }
  list res;
```

```
  for (i=1; i<=n; i++)
    {
      if (bin[i] == 1)
        {
          res = insert(res, X[i] ,size(res));
        }
    }
  return(res);
}

proc children (int v, intmat m)
{
  int n = ncols(m);
  list l;
  int i;
  for (i=1; i<=n; i++)
    {
      if (m[v,i] == 1)
        {
          l = insert(l,i,size(l));
        }
    }
  return(l);
}

proc Bayes_ball (list A, list C, intmat m)
{
  int n = ncols(m);
  int i,v;
  list B, pa,ch,vqueue;
  intvec visited, blocked, up, down, top, bottom;
  for (i=1; i<=n; i++)
    {
      visited[i] = 0;
      blocked[i] = 0;
      up[i] = 0;
      down[i] = 0;
      top[i] = 0;
      bottom[i] = 0;
    }
  for (i=1; i<=size(C); i++)
    {
```

```
      blocked[C[i]] = 1;
    }
  for (i=1; i<=size(A); i++)
    {
      vqueue = insert(vqueue, A[i]);
      up[A[i]] = 1;
    }
  while (size(vqueue) != 0)
    {
      v = vqueue[size(vqueue)];
      vqueue = delete(vqueue,size(vqueue));
      visited[v] = 1;
      if (!blocked[v] and up[v])
        {
          if (!top[v])
            {
              top[v] = 1;
              pa = parent(v,m);
              for (i=1; i<=size(pa); i++)
                {
                  vqueue = insert(vqueue, pa[i]);
                  up[pa[i]] = 1;
                }
            }
          if (!bottom[v])
            {
              bottom[v] = 1;
              ch = children(v,m);
              for (i=1; i<=size(ch); i++)
                {
                  vqueue = insert(vqueue, ch[i]);
                  down[ch[i]] = 1;
                }
            }
        }
      if (down[v])
        {
          if (blocked[v] and !top[v])
            {
              top[v] = 1;
              pa = parent(v,m);
              for (i=1; i<=size(pa); i++)
```

```
              {
                vqueue = insert(vqueue, pa[i]);
                up[pa[i]] = 1;
              }
          }
        if (!blocked[v] and !bottom[v])
          {
            bottom[v] = 1;
             ch = children(v,m);
            for (i=1; i<=size(ch); i++)
              {
                vqueue = insert(vqueue, ch[i]);
                down[ch[i]] = 1;
              }
          }
      }
    }
  for (i=1; i<=n; i++)
    {
      if(!bottom[i] and !blocked[i])
        {
          B = insert(B,i,size(B));
        }
    }
  return(B);
}

proc globalMarkov (intmat m)
{
  int n = ncols(m);
  int i,j,k,d,flag;
  list X,Y,A,B,C,l,s;
  for (i=1; i<=n; i++)
    {
      X[i] = i;
    }
  for (i=1; i<2^n-1; i++)
    {
      A = subset(i,X);
      Y = subset(2^n-i-1, X);
      d = size(Y);
      for (j=0; j<2^d-1; j++)
```

```
        {
          C = subset(j,Y);
          B = Bayes_ball(A,C,m);
          if (size(B)!=0)
            {
              flag = 0;
              s = list();
              s = insert(s,A);
              s = insert(s,B,size(s));
              s = insert(s,C,size(s));
              for (k=1; k<=size(l); k++)
                {
                  if (equivStatements(s,l[k]))
                    {
                      flag = 1;
                      break;
                    }
                }
              if (!flag)
                {
                  l = insert(l,s,size(l));
                }
            }
        }
    }
  return(l);
}

proc equivStatements(list s, list t)
{
  int i;
  if (size(s[1])!=size(t[2]) or size(s[2])!=size(t[1])
      or size(s[3])!=size(t[3]))
    {
      return(0);
    }
  for (i=1; i<=size(s[1]); i++)
    {
      if (s[1][i] != t[2][i])
        {
          return(0);
        }
```

```
    }
  for (i=1; i<=size(s[2]); i++)
    {
      if (s[2][i] != t[1][i])
        {
          return(0);
        }
    }
  for (i=1; i<=size(s[3]); i++)
    {
      if (s[3][i] != t[3][i])
        {
          return(0);
        }
    }
  return(1);
}

//////////////////////////////////////////////////////////////////
/* Construction of the probability distribution ring */
//////////////////////////////////////////////////////////////////

/* This procedure computes the index of the next variable in the ring */

proc next (intvec u, int j, intvec d)
{
  intvec v = u;
  if (j > 1)
    {
      v[j] = (v[j]+1)%(d[j]+1);
      if (v[j] == 0)
        {
          v[j] = 1;
          v = next (v, j-1, d);
        }
    }
  else
    {
      int check = (v[j]+1)%(d[j]+1);
      if (check != 0)
        {
          v[j] = check;
```

```
        }
    }
  return(v);
}

proc sdec (intvec id)
{
  int n = size(id);
  int dec;
  for (int i=n; i>=1; i--)
    {
      dec = dec + id[i]*(10)^(n-i);
    }
  return(string(dec));
}

/* This procedure has as its inpute the list of d_i's */

proc probring (intvec d, list#)
{
  if (size(#)==0 or size(#)>3) { #[1] = "pdR"; #[2] = "p"; #[3] = "dp";}
  if (size(#)==1) { #[2] = "p"; #[3] = "dp";}
  if (size(#)==2) { #[3] = "dp";}
  int i;
  intvec idx;
  for (i=1; i<=size(d); i++)
    {
      idx[i] = 1;
    }
  string ringconstructor = "ring" + " " + #[1] + " = 0,(" +
                           #[2] + sdec(idx);
  for (i=2; i<=product(d); i++)
    {
      ringconstructor = ringconstructor + ",";
      idx = next(idx,size(d),d);
      ringconstructor = ringconstructor + #[2] + sdec(idx);
    }
  ringconstructor = ringconstructor + ")," + #[3] + ";";
  execute(ringconstructor);
  keepring basering;
}
```

```
/* This procedure returns the index of the corresponding indet. */

proc index (list linput, intvec d)
{
  int i;
  int base = 1;
  list shift = linput;
  for (i=1; i<=size(shift); i++)
    {
      shift[i] = shift[i]-1;
    }
  int n = size(shift);
  int idx = shift[n];
  for (i=n-1; i>=1; i--)
    {
      base = base*d[i+1];
      idx = idx + shift[i]*base;
    }
  idx = idx+1;
  return(idx);
}


//////////////////////////////////////////////////////////////////
/* Computation of the ideal I_M for all the models defined above */
//////////////////////////////////////////////////////////////////

/* This procedure computes the cartesian product of a list of lists */

proc cartesian (list linput)
{
  int i,j,k;
  if (size(linput) == 1)
    {
      list l = linput[1];
      for (i=1; i<=size(l); i++)
        {l[i] = list(l[i]);}
      return(l);
    }
  list head = linput[1];
  list tail = cartesian(delete(linput,1));
  list final, each;
```

```
  for (i=1; i<=size(head); i++)
    {
    for (j=1; j<=size(tail); j++)
      {
        each = insert(tail[j], head[i]);
        final = insert(final, each, size(final));
      }
    }
  return(final);
}

/* This procedure computes the set of all pairs of a given list */

proc Pairs (list L)
{
  int i,j;
  list result;
  for (i=1; i<=size(L)-1; i++)
    {
      for (j=i+1; j<=size(L); j++)
        {
          result = insert(result, list(L[i],L[j]), size(result));
        }
    }
  return(result);
}

/* This little procedure computes the levels of the random variable Xi */

proc levels (int di)
{
  list l;
  for (int i=1; i<=di; i++)
    {
      l[i] = i;
    }
  return(l);
}

/* This procedure computes the linear form Prob(A)
   for an instantiation of a subset A of the random variables */
```

```
proc Prob (list linput, intvec d)
{
  int i;
  list l=linput;
  for (i=1; i<=size(linput); i++)
    {
      if (typeof(l[i]) == "string")
        {
          l[i] = levels(d[i]);
        }
      else
        {
          l[i] = list(l[i]);
        }
    }
  l = cartesian(l);
  poly pr;
  for (i=1; i<=size(l); i++)
    {
      pr = pr + var(index(l[i],d));
    }
  return(pr);
}

/* This procedure computes the quadric associated to the fixed event
   P(A=a,B=b,C=c) */

proc Quad (list A, list a, list B, list b, list C, list c, intvec d)
{
  int i,j;
  list Q1, Q2, Q3, Q4, l;
  poly P1,P2,P3,P4;
  for (i=1; i<=size(d); i++)
    {
      Q1[i] = "IND";
      Q2[i] = "IND";
      Q3[i] = "IND";
      Q4[i] = "IND";
    }
  if (size(C) != 0)
    {
      for (i=1; i<=size(C); i++)
```

```
        {
          j = C[i];
          Q1[j]=c[i]; Q2[j]=c[i];
          Q3[j]=c[i]; Q4[j]=c[i];
        }
    }
  for (i=1; i<=size(B); i++)
    {
      j = B[i];
      l = b[1];
      Q1[j] = l[i]; Q3[j] = l[i];
      l = b[2];
      Q2[j] = l[i]; Q4[j] = l[i];
    }
  for (i=1; i<=size(A); i++)
    {
      j = A[i];
      l = a[1];
      Q1[j] = l[i]; Q4[j] = l[i];
      l = a[2];
      Q2[j] = l[i]; Q3[j] = l[i];
    }
  P1 = Prob(Q1,d);
  P2 = Prob(Q2,d);
  P3 = Prob(Q3,d);
  P4 = Prob(Q4,d);
  return(P1*P2 - P3*P4);
}

/* This procedure computes the list of all quadrics associated to the
   probability of the event P(A,B,C) */

proc StatementQuadrics (list A, list B, list C, intvec d)
{
  int i,j,k;
  list a,b,c,result;
  for (i=1; i<=size(A); i++)
    {
      a[i] = levels(d[A[i]]);
    }
  a = Pairs(cartesian(a));
  for (i=1; i<=size(B); i++)
```

```
    {
      b[i] = levels(d[B[i]]);
    }
  b = Pairs(cartesian(b));
  if (size(C) == 0)
    {
      for (i=1; i<=size(a); i++)
        {
          for (j=1; j<=size(b); j++)
            {
              result = insert(result, Quad(A, a[i], B, b[j], C, list(), d),
                                size(result));
            }
        }
    }
  else
    {
      for (i=1; i<=size(C); i++)
        {
          c[i] = levels(d[C[i]]);
        }
      c = cartesian(c);
      for (k=1; k<=size(c); k++)
        {
          for (i=1; i<=size(a); i++)
            {
              for (j=1; j<=size(b); j++)
                {
                  result = insert(result, Quad(A, a[i], B, b[j], C, c[k], d),
                                    size(result));
                }
            }
        }
    }
  return(result);
}

// This function computes the Markov ideal I_{L}
// for a list of independent statements L.

proc MarkovIdeal (list L, intvec d)
{
```

```
  option(redSB);
  int i,j;
  ideal result;
  list l,genlist;
  for (i=1; i<=size(L); i++)
    {
      l = L[i];
      genlist = StatementQuadrics(l[1],l[2],l[3],d);
      for (j=1; j<=size(genlist); j++)
        {
          result = result,genlist[j];
        }
      result = interred(result);
    }
  //  result = groebner(result);
  return(result);
}
example
{
  "EXAMPLE:";
  echo = 2;
  intvec d = 2,2,2,2; int n = size(d);
  probring(d);
  intvec v15 = 1,1,0,0,1,1;
  intmat m15 = bnet(n,v15);
  list l15 = localMarkov(m15);
  list pw15 = pairMarkov(m15);
  list g15 = globalMarkov(m15);
  ideal I15 = MarkovIdeal(l15,d);
  info(I15);
  ideal G15 = MarkovIdeal(g15,d);
  info(G15);
  quotient(I15,G15);
  ideal T15 = torideal(I15,d);
  quotient(I15,T15);
  ideal Q15 = sat(I15,T15)[1];
  list pd15 = primdecGTZ(Q15);
  info(T15)[1];
  for (int i=1; i<=size(pd15); i++)
    {
      info(std(pd15[i][1]))[1];
    }
```

```
}

proc torideal (ideal I, intvec d)
{
  list l,f;
  int i,j;
  ideal t = I;
  for (i=1; i<=size(d); i++)
    {
      for (j=1; j<i; j++)
        {
          l[j] = "IND";
        }
      for (j=i; j<=size(d); j++)
        {
          l[j] = levels(d[j]);
        }
      f = f + cartesian(l);
    }
  for (i=1; i<=size(f); i++)
    {
      t = sat(t,Prob(f[i],d))[1];
    }
  return(t);
}
```

Vita

Home Address:	125 Fisher Street No. 4, Christiansburg, VA 24073 (540) 239–3256
Office Address:	Department of Mathematics, Virginia Tech 407A McBryde Hall (0123), Blacksburg, VA 24061 (540) 231–5456, lgarcia@math.vt.edu, http://www.math.vt.edu/people/lgarcia
Date of Birth:	25 July 1976
Place of Birth:	Mexico City, Mexico
Citizenship:	Mexican (applying for U.S. permanent residency)
Visa Status:	F1

Education:

2004 Ph.D. **Virginia Polytechnic Institute and State University**, Mathematics (expected)
Advisor: Reinhard Laubenbacher
Dissertation: Algebraic Geometry of Bayesian Networks.

1999 B.Sc. **UNAM**, Mexico City, Mexico, Mathematics
with an honors thesis in Gröbner basis.
Double major and course work in Computer Science.

Scientific/Academic Honors:

Fall 2002 **Research Fellowship** (Department of Mathematics, University of Genova, Italy)

2002 – 2004 **Graduate Teaching Assistantship**
(Department of Mathematics, Virginia Polytechnic Institute and State University)

1999 – 2001 **Graduate Teaching Assistantship**
(Department of Mathematical Sciences, New Mexico State University)

2000 **2000 Outstanding Scholars of the 21^{ST} Century**
(International Biographical Centre, Cambridge, England)

1999 **Sotero Prieto Award for the Best Undergraduate Thesis of the Year**
(Nationwide honor awarded by the Mexican Mathematical Society)

1997 – 1999 **Undergraduate Teaching Assistantship**
(Science School, National Autonomous University of Mexico)

Research Interests:

Algebraic Geometry, Algebraic Statistics, Combinatorics,
Commutative Algebra, Discrete Dynamical Systems, Symbolic Computation

Publications/Preprints:

2003 1. Toric Ideals of Spanning Trees of a Graph. In preparation.

2. Algebraic Geometry of Bayesian Networks (with M. Stillman, B. Sturmfels). Submitted, 30 pages.

2002 3. Cohen-Macaulay Deformations and Resolutions of Matroid Ideals. In preparation.

2001 4. Classification of Finite Dynamical Systems (with A. Jarrah, R. Laubenbacher). Submitted, 13 pages, arXiv:math.DS/0112216.

5. Topics in Discrete Dynamical Systems (with A. Jarrah, R. Laubenbacher). Manuscript, 63 pages.

6. The Additive Structure of Indecomposable Modules over a Dedekind–like Ring (with M. A. Aviño Diaz). Submitted, 8 pages.

2000 7. Bases de Gröbner Asociadas a Módulos Finitos, in *Miscelánea Matemática (MMS)* **30** (2000), pp. 65–70.

Selected Lectures:

2003 Dec. **Computational Algebraic Statistics (American Institute of Mathematics, Palo Alto, CA)**: Independence varieties of Bayesian networks on three observable variables and one hidden variable.

Aug. **Informal Geometry Seminar (Georgia Tech., GA)**: Algebraic geometry of Bayesian networks.

June **Challenges in Stochastic Computation Closing Workshop (SAMSI, NC)**: Algebraic geometry of Bayesian networks with hidden variables.

—— **MEGA 2003 (Kaiserslautern, Germany)**: Algebraic geometry of Bayesian networks.

May **Seminario de Algebra (Instituto de Matemáticas, Morelia, Mexico)**: Algebraic geometry of Bayesian networks.

Feb. **Grostat VI (Nice, France)**: Algebraic classification of Bayesian networks.

Jan. **Mini-workshop on Algebraic Statistics (U. C. Berkeley, CA)**: Algebraic geometry of Bayesian networks.

2002 Dec. **U. Cantabria (Santander, Spain)**: Algebraic geometry of Bayesian networks.

—— **U. Cantabria (Santander, Spain)**: Resolutions of Cohen-Macaulay deformations of matroid ideals.

Nov. **Algebraic Statistics Seminar (Politecnico di Torino, Italy)**: Algebraic geometry of Bayesian networks.

Aug. **MSRI seminar on Combinatorial Commutative Algebra (Mathematical Science Research Institute, Berkeley, California)**: Resolutions of matroid ideals.

—— **SIAM (San Diego, CA)**: Classification of finite dynamical systems.

Mar. **SIAM Graduate Student Seminar (Virginia Tech)**: Resolutions of matroid ideals.

—— **SIAM Graduate Student Seminar (Virginia Tech)**:Combinatorics in the primary decomposition of Cohen-Macaulay monomial ideals.

—— **SIAM Graduate Student Seminar (Virginia Tech)**: Monomial ideals and free resolutions.

2001 Oct. **U. Bordeaux I, (Bordeaux, France)**: Mathematical foundations for computer simulations.

Sept. **SACNAS (Phoenix, AZ)**: Mathematical foundations for computer simulations.

2000 Oct. **SACNAS (Atlanta, GA)**: Combinatorial tools for the analysis of decision systems.

1999 June **U. Wyoming**: Computing Gröbner bases associated to finite modules.

—— **U. Wyoming**: Computing syzygies *à* la Gauß-Jordan.

Feb. **CIMAT (Guanajuato, México)**: Gröbner bases associated to finite modules.

Teaching/Research Activities:

2003 Fall **Graduate Teaching Assistant, Virginia Tech**. Taught a Calculus course, Math 1205.

2003 Fall **Organized (with R. Laubenbacher) and presented** a weekly seminar on *Algebraic Statistics*, Virginia Tech.

2003 Fall **Conducted research** in the area of *Non-commutative Gröbner Bases* under the supervision of Ed Green, Virginia Tech.

2003 Fall Attended Virginia Tech Mathematics Department training course for teaching assistants.

2003 Spr. **Research Fellowship, Virginia Tech**. Conducted research on the project *Algebraic Algorithms for Cell Complexes* under the supervision of R. Laubenbacher and B. Sturmfels. This project received funds from the NSF, through the Computational and Algebraic Representation of Geometric Objects program.

2002 Fall **Research Fellowship, University of Genova, Italy**. Collaborator in the CoCoA *group* under the supervision of L. Robbiano. Worked on the implementation of several procedures related to primary decomposition of ideals in the computer algebra system CoCoA.

2002 Fall **Organized (with L. Robbiano) and presented** a weekly seminar on *Algebraic Geometry of Graphical Models*, DIMA, University of Genova, Italy.

2002 Sum. **Teaching Assistantship, Virginia Tech**. Worked under the supervision of Eileen Shugart. Revised and updated the *Virginia Tech faculty teaching manual.*

2002 Spr. **Organized the Graduate Student Seminar** sponsored by the Virginia Tech SIAM University Chapter. In this weekly seminar, senior graduate students presented their own research.

2002 Spr. **Graduate Teaching Assistant, Virginia Tech**. Teaching Assistant at the Math Emporium, a facility with 500 high-performance computers where students of calculus and linear algebra work on their course assignments. *Matlab* assistant for the undergraduate calculus course Math 1205.

2001 Fall **Research Fellowship, New Mexico State University.** Collaborator in the *biocomplexity research program*, under the supervision of R. Laubenbacher. Developed several modules of a software package designed to study the behavior of biological interaction networks.

2001 Oct. **Collaborator in the Tulip group**, as part of the biocomplexity research project. Tulip is a graph visualization software written and maintained at the University of Bordeaux I.

2001 Sum. **Seminar Associate at the Summer Institute in Mathematics for Undergraduates, NSF/REU program, University of Puerto Rico–Humacao.** Helped develop an intensive 6-week course on computational algebra for undergraduates, under the supervision of R. Laubenbacher. **Designed and conducted (with R. Garcia)** computer laboratories on using and programming in the computer algebra system *Singular*.

2001 Spr. **Organized (with R. Laubenbacher) and presented** a weekly seminar on *Gröbner bases and convex polytopes*, New Mexico State University.

2000 – 2001 **Research Fellowship**. Conducted research in the area of *Sequential Dynamical Systems*, which provide a mathematical foundation for the analysis of computer simulations. Supported by a contract from Los Alamos National Laboratory.

2000 Sum. **Research Fellowship, NMSU Physical Science Laboratory.** Developed a software package to compute combinatorial homotopy of simplicial complexes.

1999 – 2000 **Graduate Teaching Assistant, New Mexico State University.** Lectured several sections of the following undergraduate courses: Intermediate algebra, Precalculus, Trigonometry.

1999 Sum. **Teaching Fellowship, New Mexico State University.** Translated a series of mathematical articles to Spanish as part of a long-term binational research and education project conducted by D. Finston.

1999 Sum. Attended New Mexico State University training course for teaching assistants.

1997 – 1998 **Undergraduate Teaching Assistant, Universidad Nacional Autónoma de México, México.** Led recitation sections for the following undergraduate courses: C++ Programming language, Data Structures, College Algebra, Linear Algebra.

Conference/Workshops Attended:

2004 Jan. *Introductory Workshop in Topological Aspects of Real Algebraic Geometry*, MSRI, University of California at Berkeley (supported).

—— *AMS/MAA Joint Mathematics Meeting*, Phoenix, AZ.

2003 Dec. *Computational Algebraic Statistics*, American Institute of Mathematics, Palo Alto, CA (supported).

June *Challenges in Stochastic Computation Closing Workshop, June 26-28th 2003*, SAMSI, NC (supported).

—— *MEGA 2003 Conference*, Kaiserslautern, Germany (supported).

—— *COCOA VIII International School on Computer Algebra and Conference on Computational Commutative Algebra*, Cádiz, Spain (supported).

Mar. *MSRI's workshop on Computational Commutative Algebra*, University of California at Berkeley (supported).

Feb. *Grostat VI*, Department STID, University of Nice-Sophia Antipolis, Nice, France (supported).

2002 Aug. *SIAM Conference on Discrete Mathematics*, San Diego, CA.

July *Symbolic Computational Algebra 2002*, University of Western Ontario, London, Ontario, Canada (supported).

June *Learning Stacks and Computational Methods through problem solving*, University of Illinois at Urbana-Champaign (supported).

May *CBMS Lectures at Texas A&M on Solving Systems of Polynomial Equations*, Texas A&M University, College Station (supported).

2001 Sept. *2001 SACNAS National Conference*, Phoenix, Arizona (supported).

—— *Grostat V*, Tulane University, New Orleans, Louisiana (supported).

Jan. *AMS/MAA Joint Mathematics Meeting*, New Orleans, Louisiana (supported).

2000 Oct. *2000 SACNAS National Conference*, Atlanta, Georgia (supported).

Jan. *AMS/MAA Joint Mathematics Meeting*, Washington, D.C. (supported).

1999 June *Rocky Mountain Mathematics Consortium Summer Conference on Computational Algebra with Applications*, University of Wyoming, Laramie (supported).

Feb. *Conference on Gröbner Bases*, CIMAT, Guanajuato, Gto., México (supported).

1998 Nov. *Combinatorial and Computational Aspects of Optimization, Topology, and Algebra*, Oaxaca, Oax., México (supported).

1996 Oct. XXIX *National Conference of the Mexican Mathematical Society*, Universidad Aut. de San Luis Potosí, SLP, México (supported).

Computing:

Computer Algebra Systems: CoCoA GAP Macaulay2 Maple Risa/Asir Singular
Numerical Systems: Matlab
Programming Languages: C C++
Operating Systems: Linux Windows

References:

Lorenzo Robbiano
Department of Mathematics
University of Genova
Via Dodecaneso 35
16146 Genova, Italy
robbiano@dima.unige.it

Bernd Sturmfels
Department of Mathematics
925 Evans Hall
University of California
Berkeley, CA 94720-3840
bernd@math.berkeley.edu

Michael Stillman
Department of Mathematics
Cornell University
Ithaca, NY 14853-4201
mike@math.cornell.edu

Reinhard Laubenbacher
Virginia Bioinformatics Institute
1880 Pratt Drive
Blacksburg, VA 24061
reinhard@vbi.vt.edu

Edward Green
Department of Mathematics
Virginia Tech
Blacksburg, VA 24061-0123
green@math.vt.edu

Eileen Shugart (teaching)
Department of Mathematics
Virginia Tech
Blacksburg, VA 24061-0123
shugart@math.vt.edu