The Gaussian conditional independence inference problem

Dissertation

zur Erlangung des akademischen Grades

doctor rerum naturalium
(Dr. rer. nat.)

von M.Sc. Tobias Boege,

geb. am 07. Dezember 1994 in Burg,

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

Gutachter:	Dr. Thomas Kahle
	Dr. Volker Kaibel
	DrSc. Milan Studený

eingereicht am:	23. November 2021
Verteidigung am:	22. April 2022

Zusammenfassung

Die vorliegende Dissertation beschäftigt sich mit Strukturen Gaußscher bedingter Unabhängigkeit und ihrem Inferenzproblem. Bedingte Unabhängigkeit (engl. *conditional independence*, CI) ist ein Begriff aus der Wahrscheinlichkeits- und Informationstheorie und "Gaußsch" bezieht sich auf die bekannte multivariate Normalverteilung. Die CI-Relation einer multivariaten Zufallsvariable ξ, deren Komponenten durch eine endliche Menge N indiziert sind, enthält Informationen darüber, welche Komponenten ξ_I die Verteilung anderer Komponenten ξ_J beeinflussen, wenn der Wert wieder anderer Komponenten ξ_K bekannt ist. Diese Relation wird als $[\xi_\mathsf{I} \perp\!\!\!\perp \xi_\mathsf{J} \mid \xi_\mathsf{K}]$ oder kurz $(\mathsf{I}, \mathsf{J}|\mathsf{K})$ geschrieben. Bedingte Unabhängigkeit ist also eine dreiwertige Relation auf Teilvektoren von ξ, die komplexe Abhängigkeiten zwischen den Variablen in ξ kodiert.

CI-Relationen werden formal in einem Zweig der künstlichen Intelligenz über logische Inferenzregeln studiert. Solche Inferenzregeln nehmen die folgende Form an: "wenn bestimmte bedingte Unabhängigkeiten gelten, welche (Disjunktionen von) anderen Unabhängigkeiten müssen ebenfalls gelten?" Kenntnis dieser Regeln erlaubt die automatische Deduktion von Informationen über die Abhängigkeitsstruktur von beobachteten Zufallsvariablen. Die Regeln, welche für CI-Relationen gelten, hängen von der Art der Wahrscheinlichkeitsverteilung ab. Binäre Verteilungen erfüllen beispielsweise andere Inferenzregeln als die kontinuierlichen Gaußschen Verteilungen.

Eine multivariat Gauß-verteilte Zufallsvariable ξ ist vollständig durch ihre Parameter, den Mittelwert $\mu \in \mathbb{R}^\mathsf{N}$ und die Kovarianzmatrix $\Sigma \in \mathrm{PD}_\mathsf{N}$, bestimmt. Unter dieser speziellen Annahme ist die bedingte Unabhängigkeitsaussage $[\xi_\mathsf{I} \perp\!\!\!\perp \xi_\mathsf{J} \mid \xi_\mathsf{K}]$ äquivalent zu einer Rangbedingung an die Teilmatrix von Σ mit Zeilen $\mathsf{I} \cup \mathsf{K}$ und Spalten $\mathsf{J} \cup \mathsf{K}$, nämlich dass diese Matrix Rang $|\mathsf{K}|$ hat. Dieses Kriterium erlaubt die Behandlung von Gaußscher CI mit Mitteln der kommutativen Algebra, da die Rangbedingung als das Verschwinden einer Reihe von Polynomen in den Einträgen von Σ formuliert werden kann. Das Inferenzproblem wird dann zu einer Frage über die Geometrie spezieller reeller Varietäten innerhalb des Kegels des positiv-definiten Matrizen.

Diese Dissertation behandelt das Gaußsche CI-Inferenzproblem aus kombinatorischer, logischer und geometrischer Sicht. Der Inhalt eines jeden Kapitels wird im Folgenden kurz zusammengefasst.

Kapitel 1 gibt eine Einführung in die Theorie der Strukturen der bedingten Unabhängigkeit im Allgemeinen und von Gaußverteilungen im Besonderen. Elementare Reduktionen der allgemeinen Situation werden hergeleitet und es wird eine Übersicht über frühere Resultate über Gaußsche bedingte Unabhängigkeit gegeben.

Kapitel 2 enthält eine Exposition der Werkzeug aus der Logik, Algebra, Geometrie und Informationstheorie, die wiederholt in der gesamten Arbeit oder in einigen Teilen davon benutzt werden.

Kapitel 3 führt eine Verallgemeinerung von Gaußschen CI-Relationen auf allgemein Körper ein und gibt elementare Resultate über ihre Struktur, insbesondere werden die Gaussoid-Axiome hergeleitet. Das Inferenzproblem wird in eine geometrische Sprache übersetzt und die Existenz von finalen Polynomen als Korrektheitsbeweise für CI-Inferenzregeln wird bewiesen.

Kapitel 4 setzt den Fokus auf algebraische Konstruktionen Gaußscher CI-Relationen über unendlichen Körpern. Das wichtigste Werkzeug ist ein sogenanntes Transferprinzip, das es erlaubt Konstruktionen von Gaußverteilungen über rationalen Funktionenkörpern auf den Grundkörper zurückzuziehen. Diese Technik wird verwendet um neue Ergebnisse über die Struktur von Gaußschen CI-Relationen zu beweisen. Es folgt, dass die wahren Inferenzregeln für Gaußverteilungen keine endliche, vollständige Axiomatisierung haben, jedoch folgen alle wahren Inferenzregeln mit höchsten zwei Voraussetzungen aus den Gaussoid-Axiomen. Endliche Axiomatisierungen über den zwei kleinsten endlichen Körpern werden hergeleitet, was zeigt dass die Annahme der Unendlichkeit des Körpers signifikant ist. Es wird ein Analogon von Rotas Vermutung aus der Matroidtheorie aufgestellt.

Kapitel 5 widmet sich der Komplexität des Inferenzproblems im für die Statistik gewöhnlichen Rahmen der reellen Zahlen. Aufbauend auf einer Kodierung der von-Staudt-Konstruktionen in der projektiven Geometrie werden drei Universalitätssätze bewiesen, die zeigen dass diese Aufgabe schwer ist, im algorithmischen Sinne (sie ist vollständig für die existentielle Theorie der reellen Zahlen), algebraisch (alle reellen algebraischen Zahlen werden benötigt um Gegenbeispiele für falsche Inferenzen hinzuschreiben), sowie, für eine orientierte Version des Inferenzproblems, topologisch (die Mengen der Gegenbeispiele zu falschen Inferenzregeln können alle Homotopie-Typen von primären semialgebraischen Mengen annehmen). Das algebraische Universalitätsresultat beantwortet eine Frage von Petr Šimeček aus dem Jahr 2006 über rationale Punkte auf Gaußschen CI-Modellen.

Kapitel 6 gibt eine Einführung in eine allgemeine Maschinerie, die aus polynomiellen Relationen auf Unterdeterminanten einer symmetrischen Matrix valide Inferenzregeln erzeugt. Diese Technik wird auf zwei Klassen von polynomiellen Relationen angewendet, was in den Gaussoid- (und den orienteriten Gaussoid-) sowie, respektive, den Semimatroid-Axiomen resultiert. Letztlich wird gezeigt, dass Gaußverteilungen die aus der Informationstheorie stammende Eigenschaft der Selbstadhäsivität haben, welche auf eine Klasse von Inferenzaxiomen angewendet werden kann um potentiell stärkere Axiome zu erzeugen. Trotz der algorithmischen Komplexität des Inferenzproblems im Allgemeinen, sind diese Axiome mithilfe des booleschen Erfüllbarkeitsproblems und linearer Programmierung schnell auffindbar. Über rechnergestützte Resultate basierend auf einer Softwareimplementierung dieser Methoden wird berichtet.

Kapitel 7 fasst die Hauptresultate der Arbeit knapp zusammen und weist auf offene Fragen und künftige Forschungsrichtungen hin.

Summary

The present thesis deals with Gaussian conditional independence structures and their inference problem. Conditional independence (CI) is a notion from statistics and information theory and "Gaussian" refers to the familiar multivariate normal distribution. The conditional independence relation of a multivariate random variable ξ whose components are indexed by a finite set N gives information about which components ξ_I influence the distribution of other components ξ_J given that yet other components ξ_K have been observed. This relation is denoted by $[\xi_\mathsf{I} \perp\!\!\!\perp \xi_\mathsf{J} \mid \xi_\mathsf{K}]$ or just $(\mathsf{I}, \mathsf{J}|\mathsf{K})$ for short. Thus, CI is a ternary relation on subvectors of ξ which encodes complicated dependencies among the variables of ξ.

Conditional independence relations are formally studied in branches of artificial intelligence by means of logical inference rules. Such inference rules take the form of "given that some conditional independences hold, which other (disjunctions of) conditional independences must also hold?" Knowing these rules allows the reasoning about dependencies among observed random variables to be automated. The rules which are valid for CI relations depends on the kind of probability distribution under consideration. Binary distributions, for example, satisfy different inference rules than the continuous Gaussian distributions.

A multivariate Gaussian random variables ξ is completely given by its two parameters, the mean $\mu \in \mathbb{R}^\mathsf{N}$ and its covariance matrix $\Sigma \in \mathrm{PD}_\mathsf{N}$. In this special setting, the conditional independence statement $[\xi_\mathsf{I} \perp\!\!\!\perp \xi_\mathsf{J} \mid \xi_\mathsf{K}]$ is equivalent to a rank condition on the submatrix of Σ whose rows are $\mathsf{I}\cup\mathsf{K}$ and whose columns are $\mathsf{J}\cup\mathsf{K}$, namely it must have rank $|\mathsf{K}|$. This criterion makes Gaussian conditional independence amenable to methods of commutative algebra because the rank condition can be formulated as the vanishing of a number of polynomials in the entries of Σ. The inference problem then becomes a question of the geometry of certain real varieties inside of the cone of positive-definite matrices.

This thesis studies the Gaussian conditional independence inference problem from a combinatorial, logical and geometric point of view. The content of each chapter is briefly summarized as follows.

Chapter 1 gives an introduction to the theory of conditional independence structures in general and of Gaussians in particular. Elementary simplifications of the general setting are derived and previous results on Gaussian CI are surveyed.

Chapter 2 is an exposition of tools from logic, algebra, geometry and information theory which are used throughout or in various parts of the thesis.

Chapter 3 introduces a generalization of Gaussian CI relations to to arbitrary fields and provides elementary results on their structure, including a derivation of the gaussoid axioms. The inference problem is cast into geometric language and the existence of final polynomials proofs for the validity of CI inference rules is proved.

Chapter 4 shifts the focus to algebraic constructions for Gaussian CI relations which are valid over infinite fields. The main tool is a so-called transfer principle which allows constructions of Gaussian distributions in rational function field extensions to be carried

back into the base field. This technique is used to prove new results on the structure of Gaussian CI relations which are specific to infinite fields. It follows that the valid inference rules for Gaussian have no finite complete axiomatization, but all inference rules with at most two antecedents follow from the gaussoid axioms. Finite axiomatizations are given for the two smallest finite fields, showing that the assumption of infinite cardinality matters. An analogue of Rota's conjecture in matroid theory is posed.

Chapter 5 studies the complexity of the inference problem in the usual statistical setting over the real numbers. Based on an encoding of the von Staudt constructions from projective geometry, three universality theorems are proved which show that this problem is hard algorithmically (it is complete for the existential theory of the reals), algebraically (all real algebraic numbers are necessary to prove the invalidity of proposed inference rules for Gaussians), and an oriented variant of the inference problem is hard topologically (the set of counterexamples to an invalid inference rule assumes the homotopy type of any primary semialgebraic set). The algebraic hardness result answers a question posed in 2006 by Petr Šimeček about rational points on Gaussian conditional independence models.

Chapter 6 introduces a framework for turning polynomial relations on the subdeterminants of a symmetric matrix into inference rules. This is applied to two classes of polynomial relations which result in the gaussoid (and oriented gaussoid) and semimatroid axioms, respectively. Finally, Gaussians are shown to possess an information-theoretic property called selfadhesivity which can be applied to any set of axioms and potentially derives a stronger set. In spite of the hardness of the general inference problem, these axioms can be found much more quickly using solvers for the boolean satisfiability problem and linear programming. Computational results based on a computer implementation of these methods are shown.

Chapter 7 gives a brief summary of the main results and points out some open questions and future directions.

Acknowledgements

I would like to use this page to express my gratitude towards the people and institutions who have supported me in various ways during my PhD years in Magdeburg. There are first of all my advisors Thomas Kahle and Volker Kaibel. I am deeply thankful for all the time and advice you shared with me. Proportionally more thanks are due to Thomas whose group I joined already as a student assistant in 2016. My thanks also go to Bernd Sturmfels for introducing me to the topic of conditional independence in the same year. I also want to thank my family for always supporting me and for leaving me every year with a quiet and idyllic home and garden to do mathematics in during their summer vacation.

I gratefully acknowledge funding for my position in the MathCoRe research training group from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under project № 314838170, GRK 2297 MathCoRe. And I thank the other fellows and PIs, and in particular my mentor Eliana Duarte, for creating a stimulating research environment out of this grant. It is hard to imagine how this thesis would look like had I never heard about mathematical complexity reduction.

Among my colleagues, I want to thank especially Andreas Kretschmer for delivering countless lessons in commutative algebra and algebraic geometry over lunch, Xiangying Chen for discussions about matroids and gaussoids, and Yuri Santos Rego for always having an open ear towards my topology problems. I also thank my seniors Frank Röttger and Christopher Borger for many interesting mathematical discussions.

Likewise, I am indebted to Milan Studený and Gennadiy Averkov for many helpful correspondences. I am grateful for the hospitality of the Institute of Information Theory and Automation which enabled a long-anticipated stay in Prague in September 2021.

Since my work involved many costly computer experiments, I want to thank the Otto-von-Guericke-Universität Magdeburg and the IT personnel of the Mathematics department for providing me with access to a range of computing servers.

Lastly, it is my special desire to express my high esteem for the late František Matúš whose work on the interface of conditional independence and matroid theory had a great influence on me — thematically and in the form of his admirably uncompromising writing style and accuracy. On many occasions during the last three years, I would have liked to ask him a question. Unfortunately we never met.

Tobias Boege

Contents

Index of symbols

The following notational conventions are employed throughout:

- N a finite ground set
- $\mathsf{I}, \mathsf{J}, \mathsf{K}, \mathsf{L}$ subsets of N
- $\mathsf{i}, \mathsf{j}, \mathsf{k}, \mathsf{l}$ elements of N not distinguished from singleton subsets
- $\mathsf{KL} = \mathsf{K} \cup \mathsf{L}$
- $\langle \mathsf{KL} \rangle = \mathsf{K} \cap \mathsf{L}$
- $\mathsf{K} \oplus \mathsf{L} = (\mathsf{K} \setminus \mathsf{L}) \cup (\mathsf{L} \setminus \mathsf{K})$
- $\mathsf{K}^{\mathsf{ij}} = \mathsf{K} \setminus \mathsf{ij}$
- $\mathsf{K}^{\mathsf{c}} = \mathsf{N} \setminus \mathsf{K}$
- $\overset{!}{=}$ to emphasize the part of an equation which is to be characterized

- $\mathbb{K}^{\times}$ the multiplicative group of the field $\mathbb{K}$
- $\operatorname{char} \mathbb{K}$ characteristic of a field
- $\overline{\mathbb{K}}$ algebraic closure
- $\widetilde{\mathbb{K}}$ real closure
- $\mathbb{K}(x_1, \dots, x_p)$ field of rational functions
- $\mathcal{R}$, $\mathcal{I}$, $\mathcal{U}$ objects of commutative algebra
- $\mathcal{J}_n$ the ideal of relations among principal and almost-principal minors

- $\mathbb{1}_{\mathsf{N}}$ the $\mathsf{N} \times \mathsf{N}$ identity matrix
- $\Sigma_{\mathsf{I},\mathsf{J}}$ the submatrix of Σ with rows I and columns J
- σ_{ij} the (i, j)-entry of Σ
- $\Sigma[\mathsf{K}] = \det \Sigma_{\mathsf{K},\mathsf{K}}$
- $\Sigma[\mathsf{ij}|\mathsf{K}] = \det \Sigma_{\mathsf{iK},\mathsf{jK}}$ (observe the Sign Convention)
- $[\![\Sigma]\!]$ the conditional independence structure
- Sym symmetric, PR principally regular, and PD positive-definite matrices

- $\mathcal{F}_k^{\mathsf{N}}$ the set of k-dimensional faces of the cube C_{N}
- $(\mathsf{I}|\mathsf{K})$ an $|\mathsf{I}|$-dimensional face of C_{N} (Section 1.2.2)
- $\mathcal{L} \setminus \mathsf{k}$, $\mathcal{L} \,/\, \mathsf{k}$, $\mathcal{L}{\downarrow}(\mathsf{I}|\mathsf{K})$ minors (Section 1.2.3)
- $\mathcal{L}{\uparrow}(\mathsf{I}|\mathsf{K})$, $\iota_{(\mathsf{I}|\mathsf{K})}$ embedding map (Section 4.3)
- $\triangle(\mathsf{ij}|\mathsf{K})$ difference functional on rank functions (Section 2.5)

- $\mathfrak{g}_{\mathbb{K}}^{+}$ positive Gaussians, $\mathfrak{g}_{\mathbb{K}}^{*}$ algebraic Gaussians over a field, (Section 3.1)
- $\mathfrak{g}^{\bullet}$ either of the above
- $\mathcal{R}^{+}$, $\mathcal{R}^{*}$, $\mathcal{R}^{\bullet}$ CI model of a constraint system (Section 3.4)
- $\mathfrak{g}$ gaussoids, $\mathfrak{o}$ orientable gaussoids (Section 1.3.2, Section 6.2.2)
- $\mathfrak{sg}$ semigraphoids, $\mathfrak{sm}$ semimatroids (Section 2.5)
- $[\varphi]$, $[\mathcal{L}]$ context of an inference formula or CI structure (Section 4.4)

1

Introduction

This chapter functions as an introduction to the objects, language and the concerns of the theory of conditional independence structures. Only after the landscape of the general theory is sketched, the subject of this thesis, Gaussian CI structures, are defined and located on the map in Section 1.3. The goal of this section is also to give an account of the directions, methods and results about Gaussian CI structures prior to this thesis.

1.1 Synthetic probability theory

Conditional independence (CI) is a basic notion in statistics. Given random variables $\xi_1, \dots, \xi_n$ indexed by $N = [n]$, conditional independence is a **ternary** relation on subcollections ξ_I, ξ_J, ξ_K indexed by subsets $I, J, K \subseteq N$ and denoted by $[\xi_I \perp\!\!\!\perp \xi_J \mid \xi_K]$. This symbol is read as "ξ_I is (conditionally) independent of ξ_J given ξ_K". It means, informally, that whenever the outcome of the *conditioning variables* ξ_K is known, then the outcomes of ξ_I and ξ_J are stochastically *independent*. In other words, if K is known or controlled, then learning I does not affect our uncertainty about J and vice versa.

This relation makes an information-theoretic assertion about random events which are coupled through a joint probability distribution. It restricts the **relevance** of some factors of the distribution to other factors, given control over the conditioning set. This notion can be used in a number of ways in the sciences. In the probabilistic approach to artificial intelligence, more specifically reasoning under uncertainty, conditional independence relations may be learned from samples. This is done, for example, in graphical modeling [Lau96], where one seeks to discover the relevance or even causal structure of the observed factors by itself. Another application are probabilistic expert systems [Pea88] which are based on a statistical model of a, say, natural phenomenon derived from assertions about the interplay of its factors from domain experts, such as biologists. The system is then fed with observational data and produces conclusions based on the data and model assumptions. Here, the conditional independence structure is part of the model. Pearl emphasizes its use in designing more performant storage and processing schemes for incoming data, by exploiting the knowledge about irrelevance relations among certain factors of the distribution. Similar ideas are used in normalization of relational databases, where *embedded multivalued dependency* constraints replace conditional independence [Fag77, ITK83]. See also the historical overview in [Stu19].

Next to its applications in reasoning and modeling, the notion of conditional independence has grown into a mathematical discipline in its own right. This development was anticipated in Dawid's seminal paper [Daw79] on the role of conditional independence in statistical theory, in which he explains a number of seemingly unrelated statistical concepts through conditional independence. A line of research into formal, or logical, properties of CI relations was initiated by Pearl, Paz and Geiger [PP85, GP93]. They were motivated by

Dawid's observation that the set of all CI statements which are valid for a probability distribution satisfies certain closure axioms. In his book [Pea88, Chapter 3], Pearl further makes the point that the notion of CI provides a qualitative, as opposed to numeric, measure of independence in artificial intelligence. It allows the above-mentioned relevance reasoning tasks to be performed by applying universally valid and **human-verifiable**, logical deduction rules, in discrete steps and without the need for computing numerically with probabilities. The distribution of discrete random variables $(\xi_i)_{i \in N}$, each taking q_i different values, requires a $q_1 \times \cdots \times q_n$ tensor to store. Computations of marginal and conditional probabilities on this tensor quickly become infeasible and prone to numerical problems, in particular since the determination of a conditional independence requires an **equation** to hold. This is a difficult condition to establish under inexact arithmetic and the logical structure of CI relations provides an exact alternative. Similar issues have been recognized and treated in computational geometry; see, e.g., the failure of a Delaunay triangulation algorithm in [Sch00] due to numerical issues resulting in **inconsistent** geometric predicate computations; see also the excellent exposition in [KMP+08]. Knuth [Knu92] resorts, similarly to Pearl, to an axiomatic framework for reasoning about geometric predicates.

The general theory of conditional independence, outside of applications and graphical models, has subsequently been cultivated by Studený, Matúš, their students in Prague and collaborators. This mathematical area may be called, in the spirit of Dawid, **synthetic probability theory** and it is the area this thesis is set in. In synthetic probability, random variables and special relations among them, such as conditional or functional dependence, are studied **descriptively** rather than analytically as properties of their distributions. This point of view is entirely analogous to synthetic geometry which avoids the usage of coordinates in the statement of its results. CI statements $[\xi_I \perp\!\!\!\perp \xi_J \mid \xi_K]$ are treated as (and abbreviated to) combinatorial objects $(I, J|K)$ and fundamental laws of probability are formulated in terms of *exchange laws* such as $(I, JK|L) \Rightarrow (I, J|L) \wedge (I, K|JL)$, which expresses that the two CI statements on the right-hand side are consequences of the one on the left-hand side for every probability distribution. The term "exchange law" refers to the fact that the subcollection-indexing sets I, J, K, L which are present in the premise of this law are recombined to form the conclusions. Tools from algebra, information theory, measure theory, combinatorics and geometry are established in the study of these discrete structures and their probabilistic representations. In synthetic geometry [BS89], *linear independence* is the fundamental relation among points in space. The corresponding combinatorial study of their laws was initiated by Whitney [Whi35] and is now known as **matroid theory**. It has inspired in particular the work of Matúš [Mat94, Mat97, Mat99b] and continues to be a shaping force in this thesis. The analogy between synthetic probability and geometry does not exist in ideas only but also in methodology and mathematical content. This is because for certain types of distributions, such as discrete and Gaussian ones, a CI statement $[\xi_I \perp\!\!\!\perp \xi_J \mid \xi_K]$ corresponds to certain polynomial equations on their parameters. This observation is a cornerstone of the treatment of conditional independence in *algebraic statistics* and one reason why the corresponding CI theories are relatively well developed; see [Sul18].

Matúš set notation. The objects of interest, over which CI relations are studied, are indexed by a finite *ground set*, usually denoted N or M. Upper-case sans-serif letters $I, J, K, \ldots$ stand for subsets of the ground set and lower-case letters $i, j, k, \ldots$ for elements. Elements and singleton subsets are not distinguished in notation. The following abbreviations for common operations apply:

$$KL := K \cup L, \quad K^c := N \setminus K, \quad N^{ij} := N \setminus ij.$$

In particular, $ij = ji$ stands for the two-element subset $\{i, j\} \subseteq N$. The set union notation KL suggests that K and L are disjoint, unless otherwise stated.

1.2 Conditional independence structures

1.2.1 The semigraphoid axioms. We consider finitely many, jointly distributed random variables indexed by the set N. As in the previous section, conditional independence statements are abbreviated as $(\mathsf{I}, \mathsf{J}|\mathsf{K})$ with subsets $\mathsf{I}, \mathsf{J}, \mathsf{K} \subseteq \mathsf{N}$. We will additionally assume that these sets are disjoint because otherwise degenerate CI statements like $(\mathsf{I}, \mathsf{I}|\mathsf{K})$ may arise (and will arise given the Decomposition axiom below). These statements are *functional dependence (FD)* statements which express the condition that there exists a deterministic function f such that $\xi_{\mathsf{I}} = f(\xi_{\mathsf{K}})$. While these statements and their interplay with conditional independence are interesting, they are trivial in the regular Gaussian case which this thesis focuses on. It is therefore justified and it simplifies this introductory treatment to focus on "pure" conditional independence statements with disjoint $\mathsf{I}, \mathsf{J}, \mathsf{K}$.

The following fundamental axioms of probabilistic CI were found by Dawid [Daw79, Lemmas 4.1–4.3] and given their names by Pearl and Paz in [PP85]:

$$(\mathsf{I}, \emptyset|\mathsf{L}), \tag{Triviality}$$
$$(\mathsf{I}, \mathsf{J}|\mathsf{L}) \Rightarrow (\mathsf{J}, \mathsf{I}|\mathsf{L}), \tag{Symmetry}$$
$$(\mathsf{I}, \mathsf{JK}|\mathsf{L}) \Rightarrow (\mathsf{I}, \mathsf{J}|\mathsf{L}), \tag{Decomposition}$$
$$(\mathsf{I}, \mathsf{JK}|\mathsf{L}) \Rightarrow (\mathsf{I}, \mathsf{K}|\mathsf{JL}), \tag{Weak union}$$
$$(\mathsf{I}, \mathsf{J}|\mathsf{L}) \wedge (\mathsf{I}, \mathsf{K}|\mathsf{JL}) \Rightarrow (\mathsf{I}, \mathsf{JK}|\mathsf{L}). \tag{Contraction}$$

A *semigraphoid* is a set of CI statements which is closed under all of the above implications. Semigraphoids are a structured type of ternary relation on the powerset of N. By Dawid's results, the set of all true CI statements about a random vector form a semigraphoid; cf. [Stu05, Lemma 2.1] for a general measure-theoretic proof. Thus they may immediately be applied to any set of statements about the conditional independences among random variables to derive additional knowledge about the independence structure of the stochastic system. Writing this axiom system in a shorter, more symmetric way yields

$$(\mathsf{I}, \emptyset|\mathsf{L}), \tag{Triviality}$$
$$(\mathsf{I}, \mathsf{J}|\mathsf{L}) \Leftrightarrow (\mathsf{J}, \mathsf{I}|\mathsf{L}), \tag{Symmetry}$$
$$(\mathsf{I}, \mathsf{J}|\mathsf{L}) \wedge (\mathsf{I}, \mathsf{K}|\mathsf{JL}) \Leftrightarrow (\mathsf{I}, \mathsf{JK}|\mathsf{L}). \tag{Semigraphoidality}$$

Recall that the CI statement $(\mathsf{I}, \mathsf{J}|\mathsf{K})$ means that, given the outcome of ξ_{K}, revealing the value of ξ_{I} gives no further information about ξ_{J}. The previous sentence is inaccurate insofar as it assigns different roles to I and J: one of them is revealed and its effect on the distribution of the other is observed. In probability theory, conditional independence is the conditioning of a distribution (the revealing of K) followed by a check for stochastic independence, i.e., whether the distribution of IJ factors into the marginals of I and J, and this is a symmetric condition.

Example 1.1: Semigraphoids in topology. Although this thesis is not about topology, it is instructive on a first encounter with conditional independence to view a CI statement $(\mathsf{I}, \mathsf{J}|\mathsf{K})$ as an abstract *separation* statement in a topological space. Let each $\mathsf{i} \in \mathsf{N}$ stand for a path-connected subset A_{i} in some fixed ambient topological space and let $\mathsf{K} \subseteq \mathsf{N}$ stand for $A_{\mathsf{K}} = \bigcup_{\mathsf{k} \in \mathsf{K}} A_{\mathsf{k}}$. Then we say that $(\mathsf{I}, \mathsf{J}|\mathsf{K})$ holds if and only if every path from A_{I} to A_{J} intersects the point set A_{K}. Under this interpretation, which is in principle borrowed from Pearl's book [Pea88, Section 3.1.2], the semigraphoid axioms are plausible inference rules. Pearl considers $(\mathsf{I}, \mathsf{J}|\mathsf{K})$ more combinatorially as a statement about *cut sets* in a finite graph instead of a topological space with finitely many distinguished subsets, that is, I, J and K are subsets of the vertices of a graph and paths run along the edges of the graph. By viewing

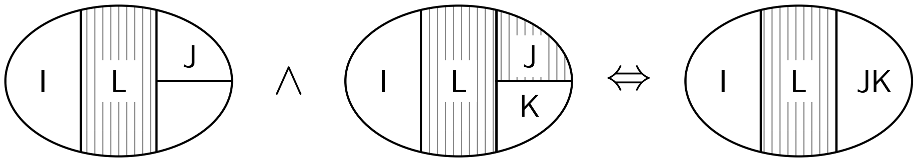

Figure 1.1: Interpretation of semigraphoidality $(\mathsf{I}, \mathsf{J}|\mathsf{L}) \wedge (\mathsf{I}, \mathsf{K}|\mathsf{JL}) \Leftrightarrow (\mathsf{I}, \mathsf{JK}|\mathsf{L})$ as path separation statements in a topological space. The shaded areas intercept any path between the other labeled areas. This depiction is inspired by [Pea88, Figure 3.1].

the graph as a CW complex [Hat02], Pearl's interpretation is subsumed by path separation in topological spaces. A topological sketch of semigraphoidality is contained in Figure 1.1. Information-theoretically, we imagine that information is exchanged in a topological space along paths, so the existence of a path between A_I and A_J is grounds for a dependence between I and J, whereas if all paths between A_I and A_J have to travel through A_K, then I and J are independent given K. △

Readers who verify Figure 1.1 will notice that one direction of the equivalence is particularly weak in topological spaces. If L separates I and JK, then surely it must separate I and J and I and K. The separation of I and K by L implies their separation by JL. This proof is so easy because separation in topological spaces is monotone with respect to the separating set; that is, it satisfies the additional property

$$(\mathsf{I}, \mathsf{J}|\mathsf{L}) \Rightarrow (\mathsf{I}, \mathsf{J}|\mathsf{KL}). \tag{Ascension}$$

The analogy of conditional independence to path separation is deceptive because this is *not* a universal property of probabilistic conditional independence. With random variables, it may happen that I and J are independent but when conditioning on a set K of random variables which depend on I and J, then K may **explain** a previously inaccessible dependence between I and J, therefore $\neg(\mathsf{I}, \mathsf{J}|\mathsf{K})$. This mismatch was understood and explained early on in probabilistic graphical modeling. The following concrete example taken from [Pea88, Section 3.1.3] clarifies the situation:

Example 1.2: Probabilistic CI is not ascending. Consider the following experiment: flip two independent coins and ring a bell if and only if the coins land with the same side up. This experiment involves three binary random variables: the two coins c_1, c_2 and the bell b. By assumption $[\mathsf{c}_1 \perp\!\!\!\perp \mathsf{c}_2]$, whereas the bell b is functionally dependent on the pair $(\mathsf{c}_1, \mathsf{c}_2)$. Given knowledge of b, the value of c_1 is uniquely determined by the outcome of c_2, and therefore $\neg[\mathsf{c}_1 \perp\!\!\!\perp \mathsf{c}_2 \mid \mathsf{b}]$. △

An early treatment of semigraphoids with the ascension property in [Mat92, Section 2] leads to a bijection with so-called weak families of connected sets, a notion chiefly inspired by connectedness in topological spaces.

1.2.2 Local semigraphoids and the hypercube. A significant reduction in the *syntax* of conditional independence was achieved by Matúš [Mat92]. It is based on the observation that semigraphoidality,

$$(\mathsf{I}, \mathsf{J}|\mathsf{L}) \wedge (\mathsf{I}, \mathsf{K}|\mathsf{JL}) \Leftrightarrow (\mathsf{I}, \mathsf{JK}|\mathsf{L}),$$

together with symmetry allows to recursively decompose variable sets I and JK in the statement $(\mathsf{I}, \mathsf{JK}|\mathsf{L})$ on the right until they are singletons $\mathsf{i}, \mathsf{j}, \mathsf{k}$, at the expense of increasing the number of CI statements and enlarging their conditioning sets. Since $(\mathsf{I}, \mathsf{JK}|\mathsf{L})$ is symmetric in J and K, one obtains the following exchange axiom expressed only over singleton variables $\mathsf{i}, \mathsf{j}, \mathsf{k} \in \mathsf{N}$ and arbitrary conditioning sets $\mathsf{L} \subseteq \mathsf{N}$:

$$(\mathsf{ij}|\mathsf{L}) \wedge (\mathsf{ik}|\mathsf{jL}) \Leftrightarrow (\mathsf{ik}|\mathsf{L}) \wedge (\mathsf{ij}|\mathsf{kL}). \tag{S}$$

Above and in the rest of this thesis, we deliberately write the first part of a local CI statement $(\mathsf{ij}|\mathsf{L})$ as a two-element set $\mathsf{ij} = \mathsf{ji}$, as per Matúš set notation, to suggest that the symmetry property holds and that we identify the two symmetric versions of the same statement.

Call the sets of CI statements $(\mathsf{I}, \mathsf{J}|\mathsf{K})$ which satisfy the semigraphoid axioms introduced originally *global semigraphoids* and the sets of $(\mathsf{ij}|\mathsf{K})$ statements which satsify the local version (S) of the global semigraphoid axioms *local semigraphoids*.

Localization of semigraphoids. The global and the local semigraphoids are in bijection, which is monotone with respect to inclusion, via the rule

$$(\mathsf{I}, \mathsf{J}|\mathsf{K}) \;\Leftrightarrow \bigwedge_{\substack{\mathsf{i}\in\mathsf{I},\mathsf{j}\in\mathsf{J},\\ \mathsf{K}\subseteq\mathsf{L}\subseteq\mathsf{IJK}\setminus\mathsf{ij}}} (\mathsf{ij}|\mathsf{L}). \tag{L}$$

The set of local CI statements over ground set N is $\mathcal{A}_\mathsf{N} := \{\, (\mathsf{ij}|\mathsf{K}) : \mathsf{ij} \in \binom{\mathsf{N}}{2}, \mathsf{K} \subseteq \mathsf{N}^{\mathsf{ij}} \,\}$. That the localization rule induces a bijection proves that there are at most $2^{\binom{n}{2}2^{n-2}}$ semigraphoids on n random variables — a bound which is not obvious from the definition of global semigraphoids. The construction in [BK20, Section 3] is suitable to prove a lower bound of $2^{\Theta(n2^n)}$. However, not all of these semigraphoids may be realizable by a probability distribution. The relatively small gap between these two bounds shows that localization gets closer to the "syntactic essence" of semigraphoids, i.e., how many bits of information are required to distinguish all n-semigraphoids.

Localization convention. The semigraphoid axioms are assumed throughout the text, which is reflected in the usage of local CI symbols.

After these reductions, we can give the definition of CI structure as it is used in this thesis:

Definition 1.3. A *CI structure* or *CI relation* is a pair $(\mathsf{N}, \mathcal{L})$ with $\mathcal{L} \subseteq \mathcal{A}_\mathsf{N}$. We also use the latter expression $\mathcal{L} \subseteq \mathcal{A}_\mathsf{N}$ as an equivalent way of specifying $(\mathsf{N}, \mathcal{L})$. The ground set N is often immaterial or implicitly given by context and its mention is omitted in these cases.

The passage to local CI symbols is not just a syntactic and quantitative relief. It opens up a new combinatorial interpretation of CI statements which was first proposed in [HMS+08]. Consider the unit (hyper)cube C_N in $\mathbb{R}^\mathsf{N}$ with its 2^n vertices in $\{\,0,1\,\}^\mathsf{N}$. Its vertices correspond to the subsets of N as indicator vectors. The set of CI statements $\mathcal{A}_\mathsf{N}$ is in bijection with the 2-dimensional faces of C_N, where $(\mathsf{ij}|\mathsf{K})$ uniquely identifies the face of dimension two extending in the directions of i and j in $\mathbb{R}^\mathsf{N}$ and compared to the "lowest" vertex $(0, 0, \ldots, 0)$ of C_N, it is shifted in all directions indicated by the set K. That is, the four vertices bounding the face are $\{\, e_\mathsf{K}, e_{\mathsf{iK}}, e_{\mathsf{jK}}, e_{\mathsf{ijK}} \,\}$ where e_S is the indicator vector of $\mathsf{S} \subseteq \mathsf{N}$. This makes C_N a useful ambient combinatorial object to study (local) CI structures in. A CI structure may be viewed as a subset of the 2-faces of C_N. Similarly, every couple of disjoint sets $(\mathsf{I}|\mathsf{L})$, where $\mathsf{I}, \mathsf{L} \subseteq \mathsf{N}$ and $\mathsf{I} \cap \mathsf{L} = \emptyset$ defines a unique $|\mathsf{I}|$-dimensional face of C_N.

Definition 1.4. A CI structure $\mathcal{L} \subseteq \mathcal{A}_\mathsf{N}$ is a *semigraphoid* if it fulfills the (local) semigraphoid axiom (S): $(\mathsf{ij}|\mathsf{L}) \wedge (\mathsf{ik}|\mathsf{jL}) \Rightarrow (\mathsf{ij}|\mathsf{kL}) \wedge (\mathsf{ik}|\mathsf{L})$ for all distinct $\mathsf{i}, \mathsf{j}, \mathsf{k} \in \mathsf{N}$ and $\mathsf{L} \subseteq \mathsf{N}^{\mathsf{ijk}}$.

The instances of the semigraphoid axiom schema are parametrized by a choice of $\mathsf{i}, \mathsf{j}, \mathsf{k}$ and L. This data may be viewed as an *oriented* 3-*face* of C_N, i.e., a 3-dimensional face $(\mathsf{ijk}|\mathsf{L})$ together with an ordering on its free dimensions ijk. The instance of the semigraphoid axiom

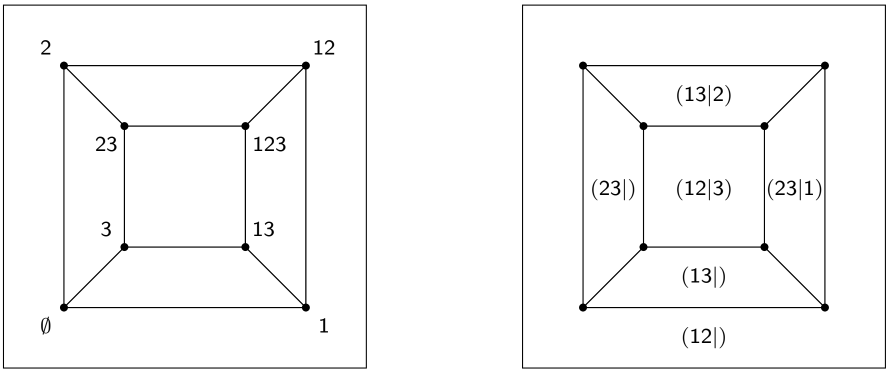

Figure 1.2: The vertices (left) and the 2-faces (right) of the 3-cube corresponding to the subsets and CI statements, respectively, of the 3-element ground set $\mathsf{N} = 123$. The 3-cube is displayed as a Schlegel diagram. $(\mathsf{ij}|)$ and $(\mathsf{ij}|\mathsf{k})$ are opposite faces because the latter is a translation of the former in direction k.

schema refers only to 2-faces in this fixed 3-face. Thus, when CI structures are viewed as sets of 2-faces of C_{N}, then semigraphoids are such sets with a prescribed pattern in every 3-face of C_{N}. Semigraphoidality is a property of CI structures which can be described "locally" in the cube.

1.2.3 CI structure theory. This section collects well-known and, from a statistical point of view, intuitive operations on general CI structures and, based on [Mat97] and [BK20], develops the interpretation of these operations on the cube C_{N}. They will be revisited in the special case of Gaussian random variables in Chapter 3. More advanced structure theory of general semigraphoids can be found in [Mat04].

Definition 1.5. Two CI structures $\mathcal{L} \subseteq \mathcal{A}_{\mathsf{N}}$ and $\mathcal{K} \subseteq \mathcal{A}_{\mathsf{M}}$ are *isomorphic* if there exists a bijection $\pi : \mathsf{N} \to \mathsf{M}$ such that $\pi(\mathcal{L}) = \mathcal{K}$, where the bijection acts element-wise and we set

$$\pi(\mathsf{ij}|\mathsf{K}) := (\pi(\mathsf{ij})|\pi(\mathsf{K})).$$

Isomorphy convention. All properties studied in this work are invariant under isomorphy, as it is just a relabeling of the ground set. Therefore we often identify CI structures up to isomorphy, which reduces the ground set part of $(\mathsf{N}, \mathcal{L})$ to the cardinality $n = |\mathsf{N}|$.

The probabilistic conditional independence relation of a vector of random variables is computed using a few primitives from probability theory. To check whether $(\xi_{\mathsf{i}})_{\mathsf{i}\in\mathsf{N}}$ satisfies $(\mathsf{ij}|\mathsf{K})$, first **marginalize** from N to ijK, then **condition** on K to obtain a bivariate vector on ij, and finally check for marginal, stochastic independence. Marginalization integrates out a subset of the random variables to obtain the probability distribution on the remaining ones. Conditioning, on the other hand, corresponds to a projection of the distribution given by assuming the outcome on a subset of the variables. The CI structures of marginalizations and conditionings of a random vector can be computed from its CI structure. Thus, we can define marginalization and conditioning of CI structures in general and in a purely formal manner.

Definition 1.6. For $\mathcal{L} \subseteq \mathcal{A}_{\mathsf{N}}$ and $\mathsf{k} \in \mathsf{N}$ we define

(i) The *marginalization* of $\mathcal{L}$ to $\mathsf{N} \setminus \mathsf{k}$ as $\mathcal{L} \setminus \mathsf{k} := \{ (\mathsf{ij}|\mathsf{L}) \in \mathcal{A}_{\mathsf{N}\setminus\mathsf{k}} : (\mathsf{ij}|\mathsf{L}) \in \mathcal{L} \}$.
(ii) The *conditioning* of $\mathcal{L}$ to $\mathsf{N} \setminus \mathsf{k}$ as $\mathcal{L} \,/\, \mathsf{k} := \{ (\mathsf{ij}|\mathsf{L}) \in \mathcal{A}_{\mathsf{N}\setminus\mathsf{k}} : (\mathsf{ij}|\mathsf{kL}) \in \mathcal{L} \}$.

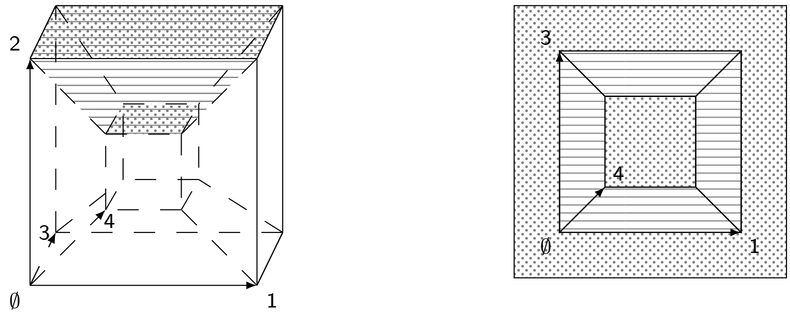

Figure 1.3: The minor of a CI structure corresponding to the 3-face $(134|2)$ of the 4-cube on the left. The ruled 2-faces around the 3-face are inside the CI structure, the dotted ones on the top and bottom are not — all other 2-faces do not matter for the minor operation. The minor on the right is a CI structure in the 3-dimensional cube labeled by 134.

Marginalization and conditioning can be defined for sets $\mathsf{K} \subseteq \mathsf{N}$ by consecutively performing the operation for all $\mathsf{k} \in \mathsf{K}$ (it is easy to see that the ordering of K does not matter).

Stochastically, these operations yield the "natural subconfigurations" of a random vector, comparable to the operations of restriction and contraction in graph or matroid theory. In analogy we define *minors* and an involution named *duality* which exchanges marginalization and conditioning:

Definition 1.7. A *minor* of $\mathcal{L} \subseteq \mathcal{A}_\mathsf{N}$ is any CI structure on a subset of N obtained by a sequence of marginalizations and conditionings. A minor of $\mathcal{L}$ on a k-element subset of N is a *k-minor*.

Definition 1.8. The *dual* of $\mathcal{L} \subseteq \mathcal{A}_\mathsf{N}$ is $\mathcal{L}^\rceil := \{\, (\mathsf{ij}|\mathsf{N}^{\mathsf{ij}} \setminus \mathsf{K}) \in \mathcal{A}_\mathsf{N} : (\mathsf{ij}|\mathsf{K}) \in \mathcal{L} \,\} \subseteq \mathcal{A}_\mathsf{N}$.

Lemma 1.9. For a CI structure $(\mathsf{N}, \mathcal{L})$ and $\mathsf{k}, \mathsf{m} \in \mathsf{N}$ distinct we have

$$(\mathcal{L}^\rceil)^\rceil = \mathcal{L}, \quad (\mathcal{L} \setminus \mathsf{k})^\rceil = \mathcal{L}^\rceil \,/\, \mathsf{k}, \quad (\mathcal{L} \,/\, \mathsf{k})^\rceil = \mathcal{L}^\rceil \setminus \mathsf{k},$$
$$(\mathcal{L} \setminus \mathsf{k}) \,/\, \mathsf{m} = (\mathcal{L} \,/\, \mathsf{m}) \setminus \mathsf{k}.$$

A further basic operation on random vectors is to take two of them, $(\xi_\mathsf{i})_{\mathsf{i}\in\mathsf{N}}$ and $(\eta_\mathsf{i})_{\mathsf{i}\in\mathsf{M}}$, and to join them to a new vector on NM. If the two random vectors are independent of each other or if they are completely dependent of each other, the CI structure of the joint vector can be determined:

Definition 1.10. Let $\mathcal{L} \subseteq \mathcal{A}_\mathsf{N}$ and $\mathcal{R} \subseteq \mathcal{A}_\mathsf{M}$ be CI structures and $\mathsf{N} \cap \mathsf{M} = \emptyset$. Their *dependent sum* is the set-theoretic union $\mathcal{L} \sqcup \mathcal{R} \subseteq \mathcal{A}_\mathsf{NM}$, stipulating that N and M are disjoint. Their *direct sum* is

$$\begin{aligned}\mathcal{L} \oplus \mathcal{R} := &\ \{\, (\mathsf{ij}|\mathsf{K}) \in \mathcal{A}_\mathsf{NM} : \mathsf{i} \in \mathsf{N}, \mathsf{j} \in \mathsf{M} \,\} \\ &\cup \{\, (\mathsf{ij}|\mathsf{KL}) \in \mathcal{A}_\mathsf{NM} : (\mathsf{ij}|\mathsf{K}) \in \mathcal{L}, \mathsf{L} \subseteq \mathsf{M} \,\} \\ &\cup \{\, (\mathsf{ij}|\mathsf{KL}) \in \mathcal{A}_\mathsf{NM} : (\mathsf{ij}|\mathsf{K}) \in \mathcal{R}, \mathsf{L} \subseteq \mathsf{N} \,\} \subseteq \mathcal{A}_\mathsf{NM}.\end{aligned}$$

Definition 1.11. A CI structure $\mathcal{L} \subseteq \mathcal{A}_\mathsf{N}$ is *connected* if it cannot be written as a direct sum of CI structures on strictly smaller ground sets. Every CI structure has a unique (up to ordering) decomposition as a direct sum of connected CI structures (see [Mat94]) which are its *connected components*.

Remark 1.12. By [Mat94] every simple matroid may be regarded as a CI structure and fulfills the semigraphoid axioms. The definitions of minor, duality, direct sum and connectedness are generalizations of the corresponding notions in matroid theory, which in turn generalize graph theory.

Viewing CI structures as sets of 2-faces of the N-cube C_{N}, minors are again very natural "subconfigurations". It is shown in [BK20] that the k-minors of $\mathcal{L} \subseteq \mathcal{A}_{\mathsf{N}}$ are precisely the restrictions of $\mathcal{L}$ to k-faces of C_{N} and viewed as a set of 2-faces on C_k, since every k-face of C_{N} is a k-dimensional cube C_k in its own right. This proves a fact observed by Matúš in [Mat97]: a CI structure is a semigraphoid if and only if all of its 3-minors are semigraphoids. There are precisely 22 semigraphoids on a 3-element ground set. This gives a finite minor-theoretic characterization of semigraphoids and an almost visual way to check the semigraphoid property: for a set of 2-faces on C_{N}, examine every 3-dimensional face and check if the set of 2-faces which lie on that face are among the 22 basic semigraphoids for $n = 3$.

The operations of isomorphy and duality can be understood as subgroups of the *hyperoctahedral group* $\mathfrak{B}_{\mathsf{N}}$, which is the group generated by reflection symmetries of C_{N}. Each group element induces an automorphism of the face lattice of C_{N} which permutes the faces of each dimension among each other. In particular, $\mathfrak{B}_{\mathsf{N}}$ induces an action on $\mathcal{A}_{\mathsf{N}}$ and therefore on CI structures. As an abstract group, $\mathfrak{B}_{\mathsf{N}}$ equals the semidirect product $(\mathbb{Z}/2)^{\mathsf{N}} \rtimes \mathfrak{S}_{\mathsf{N}}$, i.e., each of its elements can be uniquely written as a composition of a *swap* from $(\mathbb{Z}/2)^{\mathsf{N}}$ and a *permutation* from $\mathfrak{S}_{\mathsf{N}}$. Each vector in the group of swaps $(\mathbb{Z}/2)^{\mathsf{N}}$ is an indicator vector of a subset $\mathsf{Z} \subseteq \mathsf{N}$ and acts on a CI statement via

$$\mathsf{Z} \cdot (\mathsf{ij}|\mathsf{K}) := (\mathsf{ij}|\mathsf{K} \oplus \mathsf{Z}^{\mathsf{ij}}),$$

where $\oplus$ denotes the symmetric difference $\mathsf{K} \oplus \mathsf{L} = \mathsf{K} \setminus \mathsf{L} \cup \mathsf{L} \setminus \mathsf{K}$. Duality is then the special case of swapping everything, i.e., $\mathsf{Z} = \mathsf{N}$. The second constituent of the semidirect product are the permutations $\mathfrak{S}_{\mathsf{N}}$ of the axes of C_{N}, which is simply the isomorphy action. Due to its symmetry, the semigraphoid axiom schema is invariant under $\mathfrak{B}_{\mathsf{N}}$. That is, acting on one instance of the semigraphoid axioms by this group produces another instance of the schema.

Remark 1.13. Semigraphoids are closed under $\mathfrak{B}_{\mathsf{N}}$. Moreover, $g \cdot (\mathcal{L} \oplus \mathcal{R}) = (g \cdot \mathcal{L}) \oplus (g \cdot \mathcal{R})$ for all $g \in \mathfrak{B}_{\mathsf{N}}$. Thus, the connectedness of a CI structure is preserved by $\mathfrak{B}_{\mathsf{N}}$ and its subgroups, which includes isomorphy and duality.

1.3 Gaussian conditional independence

1.3.1 Algebraic statistics of Gaussian CI. The above theory of semigraphoids applies to the CI structure of every random vector. If the type of joint distribution is further constrained, more information about the CI structure can be deduced. The two most used types of distribution in the algebraic statistics literature are discrete distributions and Gaussians. Discrete random vectors are those where each component only takes a finite number of possible values and hence the vector takes finitely many values in the cartesian product. Therefore the joint distribution is completely specified by a (finite) tensor of real numbers. A Gaussian or *multivariate normal* distribution, on the other hand, is continuous. It has two sets of parameters: the *mean* $\mu \in \mathbb{R}^{\mathsf{N}}$ and the *covariance matrix* $\Sigma \in \mathrm{Sym}_{\mathsf{N}}(\mathbb{R})$. The density with respect to the Lebesgue measure on $\mathbb{R}^{\mathsf{N}}$ is then defined as

$$x \;\mapsto\; \frac{1}{\sqrt{(2\pi)^n \det \Sigma}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)\right).$$

The covariance matrix Σ must be positive-semidefinite. If it is even positive-definite, then the distribution is a *regular Gaussian*, otherwise it is *singular*. The probability mass of a singular distribution is concentrated on a proper subspace of $\mathbb{R}^{\mathsf{N}}$. This thesis is about the conditional independence structures of **regular** Gaussians and we omit this adjective from now on if there is no danger of confusion.

The reason why CI structures of discrete and Gaussian distributions are relatively well-studied is that the validity of a CI statement can be formulated as an algebraic equation in the respective parameters. This gives conditional independence an implicitly (algebro-)geometric character, which is usually more well-behaved in the Gaussian case. The starting point is the following observation in probability theory:

Lemma 1.14: [Sul18, Theorem 2.4.2 & Proposition 2.4.4]. Let ξ be distributed according to the (regular) Gaussian distribution with mean $\mu \in \mathbb{R}^{\mathsf{N}}$ and covariance $\Sigma \in \mathrm{PD}_{\mathsf{N}}$.

- — The marginal vector ξ_{K} is a regular Gaussian in $\mathbb{R}^{\mathsf{K}}$ with mean vector μ_{K} and covariance Σ_{K}, i.e., the $\mathsf{K} \times \mathsf{K}$ principal submatrix.
- — Let $y \in \mathbb{R}^{\mathsf{K}}$ and $\mathsf{J} = \mathsf{N} \setminus \mathsf{K}$. The conditional $\xi_{\mathsf{J}} \mid \xi_{\mathsf{K}} = y$ is a regular Gaussian in $\mathbb{R}^{\mathsf{J}}$ with mean vector $\mu_{\mathsf{J}} + \Sigma_{\mathsf{J},\mathsf{K}}\Sigma_{\mathsf{K}}^{-1}(y - \mu_{\mathsf{K}})$ and covariance $\Sigma_{\mathsf{J}} - \Sigma_{\mathsf{J},\mathsf{K}}\Sigma_{\mathsf{K}}^{-1}\Sigma_{\mathsf{K},\mathsf{J}}$, i.e., the Schur complement of Σ with respect to K.
- — Let a Gaussian over IJ be given with covariance $\Sigma \in \mathrm{PD}_{\mathsf{IJ}}$. Then $(\mathsf{I}, \mathsf{J}|)$ holds if and only if the submatrix $\Sigma_{\mathsf{I},\mathsf{J}} = 0$.

The general CI statement $(\mathsf{I}, \mathsf{J}|\mathsf{K})$ is the result of marginalizing a distribution to IJK, conditioning on K and then checking for independence of I and J. The previous lemma implies the following algebraic CI criterion for regular Gaussians

$$\begin{aligned}(\mathsf{I}, \mathsf{J}|\mathsf{K}) \;&\Leftrightarrow\; \left(\Sigma_{\mathsf{IJ}} - \Sigma_{\mathsf{IJ},\mathsf{K}}\Sigma_{\mathsf{K}}^{-1}\Sigma_{\mathsf{K},\mathsf{IJ}}\right)_{\mathsf{I},\mathsf{J}} = 0 \\ &\Leftrightarrow\; \Sigma_{\mathsf{I},\mathsf{J}} - \Sigma_{\mathsf{I},\mathsf{K}}\Sigma_{\mathsf{K}}^{-1}\Sigma_{\mathsf{K},\mathsf{J}} = 0 \\ &\Leftrightarrow\; \mathrm{rk}\,\Sigma_{\mathsf{IK},\mathsf{JK}} = |\mathsf{K}|. \end{aligned} \tag{$\perp\!\!\!\perp$}$$

The last equivalence follows from rank additivity of the Schur complement (see Section 3.3) and the observation that the $\mathsf{K} \times \mathsf{K}$ submatrix of Σ has full rank $|\mathsf{K}|$ because it is a principal submatrix of the positive-definite Σ and hence positive-definite as well. In particular, the truth of a conditional independence statement does not depend on the value on the K-subvector being conditioned on and it does not depend on the mean μ. Hence we identify regular Gaussians with their covariance matrices $\Sigma \in \mathrm{PD}_{\mathsf{N}}$ from now on.

The "$\geq$" part of the rank condition in ($\perp\!\!\!\perp$) always holds because the principal submatrix Σ_{K} has full rank $|\mathsf{K}|$. Then this minimal rank of $\Sigma_{\mathsf{IK},\mathsf{JK}}$ is attained if and only if all of its minors of size $|\mathsf{K}| + 1$ vanish. These minors correspond to local CI statements $(\mathsf{ij}|\mathsf{K})$ with $\mathsf{i} \in \mathsf{I}$ and $\mathsf{j} \in \mathsf{J}$. This proves a stronger version of the localization rule (L) for Gaussians (in fact, this rule is valid for all *compositional graphoids*; see below):

$$(\mathsf{I}, \mathsf{J}|\mathsf{K}) \;\Leftrightarrow\; \bigwedge_{\mathsf{i}\in\mathsf{I},\mathsf{j}\in\mathsf{J}} (\mathsf{ij}|\mathsf{K}).$$

Definition 1.15. The polynomial $\Sigma[\mathsf{ij}|\mathsf{K}] := \det \Sigma_{\mathsf{iK},\mathsf{jK}}$ is an *almost-principal minor*. The *CI structure* of $\Sigma \in \mathrm{PD}_{\mathsf{N}}$ is $[\![\Sigma]\!] := \{\, (\mathsf{ij}|\mathsf{K}) \in \mathcal{A}_{\mathsf{N}} : \Sigma[\mathsf{ij}|\mathsf{K}] = 0 \,\}$.

Example 1.16. Consider the statistical model on three jointly Gaussian random variables which satisfy the CI statements $(12|)$ and $(13|)$. These Gaussians have arbitrary mean vector and positive-definite covariance matrix $\Sigma = \left(\begin{smallmatrix} p & a & b \\ a & q & c \\ b & c & r \end{smallmatrix}\right)$ subject to the polynomial equations

$$0 \stackrel{!}{=} \Sigma[12|] = a, \qquad 0 \stackrel{!}{=} \Sigma[13|] = b.$$

Since $\Sigma[12|3] = pa - bc = 0$ and $\Sigma[13|2] = qb - ac = 0$ is implied by these equations, we see that the distributions on this statistical model also satisfy $(12|3)$ and $(13|2)$. △

Remark 1.17. The algebraic description of CI for a singular Gaussian is similar but combinatorially more involved. We refer to [Stu05, Appendix 8.3] and [Šim06b] and its references for details. In short, the Schur complement in the conditioning is replaced by a generalized Schur complement [Zha05, Section 1.6] which results in the following criterion:

$$(\mathsf{ij}|\mathsf{K}) \Leftrightarrow \det \Sigma_{\mathsf{iL},\mathsf{jL}} = 0,$$

where L is any inclusion-maximal subset of K such that $\det \Sigma_{\mathsf{L}}$ is invertible and one proves that the definition of CI does not depend on the choice of L. Thus, the polynomial in the entries of Σ to which a CI statement is interpreted depends on the vanishing and non-vanishing of principal minors of the positive-semidefinite matrix. These vanishings of principal minors themselves are non-trivially structured (see [HS07]) and this introduces complications which are avoided by restricting to regular Gaussians. Šimeček [Šim06b] compared to [LM07] gives examples of CI structures which are obtainable from singular but not from regular Gaussians.

Remark 1.18. The set of positive-semidefinite matrices forms a convex cone inside of $\mathrm{Sym}_{\mathsf{N}}(\mathbb{R})$ whose interior in the euclidean topology is precisely PD_{N}. Hence, the part of Gaussian CI theory which is specific to of singular Gaussians happens on the boundary of this cone only. Let $(\Sigma_k)_{k\in\mathbb{N}}$ be a sequence of positive-definite matrices which converges to a matrix Σ. If Σ is positive-definite and $(\mathsf{ij}|\mathsf{K}) \in [\![\Sigma_k]\!]$ for all sufficiently large k, then, since the almost-principal minor $[\mathsf{ij}|\mathsf{K}]$ is a continuous function, $(\mathsf{ij}|\mathsf{K}) \in [\![\Sigma]\!]$. This limit behavior may be expressed by $\bigcap_k [\![\Sigma_k]\!] \subseteq [\![\lim_k \Sigma_k]\!]$. This does **not** hold anymore if the limit is singular. The almost-principal minor $[\mathsf{ij}|\mathsf{K}]$ remains continuous, but the definition of CI in the semidefinite case may discontinuously attach the truth of a CI statement to a different almost-principal minor. An example of such a sequence is given in [Lau96, Example 3.11]:

$$\Sigma_k = \begin{pmatrix} 1 & 1/\sqrt{k} & 1/2 \\ 1/\sqrt{k} & 2/k & 1/\sqrt{k} \\ 1/2 & 1/\sqrt{k} & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 1/2 \\ 0 & 0 & 0 \\ 1/2 & 0 & 1 \end{pmatrix} = \Sigma.$$

Determinant calculations confirm that $[\![\Sigma_k]\!] = \{(\mathsf{13}|\mathsf{2})\}$ and $[\![\Sigma]\!] = \{(\mathsf{12}|), (\mathsf{23}|), (\mathsf{12}|\mathsf{3}), (\mathsf{23}|\mathsf{1})\}$. The additional statements come from additional zeros in the limit matrix, but the expected statement $(\mathsf{13}|\mathsf{2})$ is missing because the principal submatrix Σ_{2} is singular and thus $(\mathsf{13}|\mathsf{2})$ for Σ is determined by the (non-vanishing) almost-principal minor $\Sigma[\mathsf{13}|]$.

The CI structures of singular Gaussians do not fulfill the gaussoid axioms (to be introduced in the next section) which encode the most essential properties of regular Gaussians. For the reasons outlined above, this thesis deals almost exclusively with regular Gaussians.

1.3.2 Catalogues and axioms. This derivation leads from the statistical concept of conditional independence for regular Gaussian distributions to a family of polynomials in the entries of the covariance whose vanishing is decisive for their respective CI statement. Since the matrix is positive-definite, which can be expressed by requiring that all principal minors $\det \Sigma_{\mathsf{K}}$ be positive, Gaussian CI is a matter of **real algebraic geometry**. This thesis studies combinatorial, logical and geometric facets of this special family of polynomials.

The main task in the theory of CI structures is to characterize which structures arise from certain types of distributions, in this case regular Gaussians. This may either be done by compiling an extensive catalogue of the *realizable* structures or by finding all their *axioms*, i.e., inference rules such as the semigraphoid axioms (S) which are valid for all Gaussians and forbid every non-Gaussian structure. These two practices are equivalent because on every finite ground set N, there is only a finite number of realizable structures. These may be listed or they may be described by a boolean formula whose variables are CI statements.

Every such boolean formula can be transformed (via its conjunctive normal form) into a finite conjunction of CI axioms.

It was shown by Sullivant [Sul09] and independently Šimeček [Šim06a] that there is no finite list of axioms which completely describe Gaussian CI structures on arbitrarily many random variables. Lněnička and Matúš [LM07] compiled the lists of realizable structures on three and four variables. To accomplish this task, they combined the cataloguing and the axiomatic approach. Matúš [Mat05] first derived the following basic axioms:

Definition 1.19. A CI structure $\mathcal{G} \subseteq \mathcal{A}_\mathsf{N}$ is a *gaussoid* if it satisfies the *gaussoid axioms*:

$$(\mathsf{ij}|\mathsf{L}) \wedge (\mathsf{ik}|\mathsf{jL}) \Rightarrow (\mathsf{ik}|\mathsf{L}) \wedge (\mathsf{ij}|\mathsf{kL}), \tag{G.i}$$
$$(\mathsf{ij}|\mathsf{kL}) \wedge (\mathsf{ik}|\mathsf{jL}) \Rightarrow (\mathsf{ij}|\mathsf{L}) \wedge (\mathsf{ik}|\mathsf{L}), \tag{G.ii}$$
$$(\mathsf{ij}|\mathsf{L}) \wedge (\mathsf{ik}|\mathsf{L}) \Rightarrow (\mathsf{ij}|\mathsf{kL}) \wedge (\mathsf{ik}|\mathsf{jL}), \tag{G.iii}$$
$$(\mathsf{ij}|\mathsf{L}) \wedge (\mathsf{ij}|\mathsf{kL}) \Rightarrow (\mathsf{ik}|\mathsf{L}) \vee (\mathsf{jk}|\mathsf{L}), \tag{G.iv}$$

for all distinct $\mathsf{i}, \mathsf{j}, \mathsf{k} \in \mathsf{N}$ and $\mathsf{L} \subseteq \mathsf{N} \setminus \mathsf{ijk}$.

These axioms hold for all regular Gaussians. The first gaussoid axiom is the familiar semigraphoid axiom (S). The second and third axioms are known as *Intersection* and *Composition*, respectively. They are converses and duals of each other. The final axiom is *Weak transitivity* (sometimes called *Singleton transitivity*). A semigraphoid which satisfies Intersection is a *graphoid*. This notion is important in graphical modeling; see [PP85]. Thus, gaussoids may also be called singleton-transtive, compositional graphoids [Sad17, p. 7]. This terminology is taken from [Pea88, Chapter 3] who uses global CI symbols. This makes the axioms look substantially different. Our presentation of the axioms is localized through (L) and follows [LM07]. The two forms of the axioms are equivalent under the Localization convention.

Example 1.20. Consider the CI structure $\mathcal{L} = \{\, (12|), (12|345), (34|), (34|25) \,\} \subseteq \mathcal{A}_{12345}$. This structure satisfies the gaussoid axioms. Let $(1\ 3)$ be the cyclic permutation on $\mathsf{N} = 12345$ which exchanges 1 with 3 and leaves every other element fixed. The images of $(1\ 3)$, duality and swap by $\mathsf{Z} = 123$ on $\mathcal{L}$ are, respectively:

$$\begin{aligned} (1\ 3) \cdot \mathcal{L} &= \{(23|), (23|145), (14|), (14|25)\}, \\ \mathcal{L}^\rceil = 12345 \cdot \mathcal{L} &= \{(12|345), (12|), (34|125), (34|1)\}, \\ 123 \cdot \mathcal{L} &= \{(12|3), (12|45), (34|12), (34|15)\}. \end{aligned}$$

It is easy to see that the gaussoid axioms (G.i)–(G.iv) are invariant under the hyperoctahedral group. Thus, every CI structure obtained above from $\mathcal{L}$ by one of the group actions must be a gaussoid as well. △

Example 1.21. Consider $\mathcal{W} = \{\, (12|), (12|3) \,\} \subseteq \mathcal{A}_{123}$. This set is not a gaussoid because it is the antecedent set to the Weak transitivity axiom and does not contain any of its two consequents. This implies that $\mathcal{W}$ is not realizable by a Gaussian distribution. Indeed, if

$$\Sigma = \begin{pmatrix} p & a & b \\ a & q & c \\ b & c & r \end{pmatrix}$$

is any positive-definite matrix which satisfies $\mathcal{W}$, then we have $\Sigma[12|] = a = 0$ and $\Sigma[12|3] = pa - bc = 0$. Both equations imply $bc = pa = 0$ and thus $b = 0$ or $c = 0$. Since $b = \Sigma[13|]$ and $c = \Sigma[23|]$, this proves Weak transitivity and with it the non-realizable of $\mathcal{W}$. △

Remark 1.22. Singular Gaussians do not necessarily satisfy all of these axioms. Semigraphoidality, Composition and Weak transivity are proved in Studený's book [Stu05, Section 2.3.6], but Intersection may fail, as the positive-semidefinite matrix $\begin{pmatrix} 2 & -2 & 1 \\ -2 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix}$ with CI structure $\{\, (13|2), (23|1) \,\}$ shows. This matrix is a block of the model `M82` of [Šim06b].

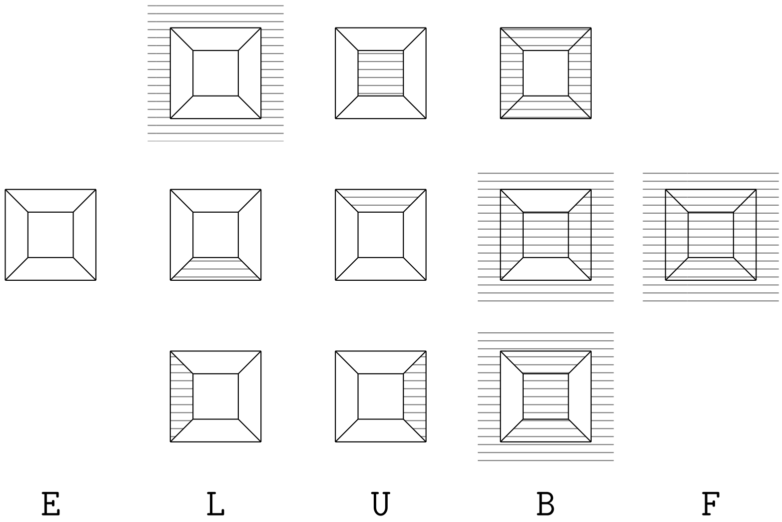

Figure 1.4: The eleven 3-gaussoids in five symmetry classes mod $\mathfrak{S}_3$ organized in columns. Ruled 2-faces of the cube correspond to CI statements inside the gaussoid.

Remark 1.23. Unlike semigraphoids and the CI structures of discrete distributions, gaussoids and Gaussian CI structures are not closed under intersection as subsets of $\mathcal{A}_\mathsf{N}$. Examples for this can be inferred from the Weak transitivity axiom which is the only gaussoid axiom that is not preserved under intersections.

Example 1.24: The 3-gaussoids. Of the $2^{\binom{n}{2}2^1} = 2^6 = 64$ possible subsets of $\mathcal{A}_{123}$ precisely eleven are gaussoids. They fall into five symmetry classes modulo $\mathfrak{S}_{123}$ pictured in Figure 1.4. We use the following abbreviations for them:

— The empty gaussoid $\mathtt{E} = \emptyset$.
— The lower singletons $\mathtt{L} = \{ (12|) \}$.
— The upper singletons $\mathtt{U} = \{ (12|3) \}$.
— The belts $\mathtt{B} = \{ (12|), (12|3), (13|), (13|2) \}$.
— The full gausssoid $\mathtt{F} = \mathcal{A}_{123}$.

In addition, L and U coincide Under the action of the hyperoctahedral group. △

Over a ground set of size three, the gaussoid axioms are not only necessary for being induced by a regular Gaussian distribution, they are also sufficient.

Theorem 1.25: [Mat05, Example 1]. Every 3-gaussoid is realizable.

As with semigraphoids, the gaussoid axioms are quantified over oriented 3-faces of the N-cube C_N and only refer to 2-faces in the given 3-face. A parsimonious graphical representation of the gaussoid axioms as inference rules in the 3-cube is shown in Figure 1.5.

Lemma 1.26: [BK20]. A CI structure $\mathcal{G} \subseteq \mathcal{A}_\mathsf{N}$ is a gaussoid if and only if each of its 3-minors is isomorphic to one of the 3-gaussoids $\{ \mathtt{E}, \mathtt{L}, \mathtt{U}, \mathtt{B}, \mathtt{F} \}$.

Of the $2^{24} = 16\,777\,216$ possible CI structures on a four-element ground set precisely 679 are gaussoids. This significant reduction in the search space allowed Lněnička and Matúš [LM07] to finish the classification of realizable structures. Of the 679 gaussoids 629 are realizable. The 50 non-realizable gaussoids come in five symmetry classes which are ruled

(G.i)—(G.iii): Any "knee" in the cube is completed to the unique "belt" which contains it.

(G.i)—(G.iii) ∘ (G.iv): Two opposite 2-faces are completed to (at least) one of the two belts which contain them.

Figure 1.5: The gaussoid axioms in the 3-cube. Premises of the axioms are dotted, conclusions are ruled. Conclusions appearing together have matching line patterns. The pictures encode the gaussoid axioms modulo $\mathfrak{B}_3$.

out by the additional *Lněnička-Matúš axioms*:

$$
\begin{aligned}
(\mathsf{ij}|\mathsf{M}) \wedge (\mathsf{kl}|\mathsf{M}) \wedge (\mathsf{ik}|\mathsf{jlM}) \wedge (\mathsf{jl}|\mathsf{ikM}) &\Rightarrow (\mathsf{ik}|\mathsf{M}), && \text{(LM.i)}\\
(\mathsf{ij}|\mathsf{M}) \wedge (\mathsf{kl}|\mathsf{iM}) \wedge (\mathsf{kl}|\mathsf{jM}) \wedge (\mathsf{ij}|\mathsf{klM}) &\Rightarrow (\mathsf{kl}|\mathsf{M}), && \text{(LM.ii)}\\
(\mathsf{ij}|\mathsf{M}) \wedge (\mathsf{jl}|\mathsf{kM}) \wedge (\mathsf{kl}|\mathsf{iM}) \wedge (\mathsf{ik}|\mathsf{jlM}) &\Rightarrow (\mathsf{ik}|\mathsf{M}), && \text{(LM.iii)}\\
(\mathsf{ij}|\mathsf{kM}) \wedge (\mathsf{ik}|\mathsf{lM}) \wedge (\mathsf{il}|\mathsf{jM}) &\Rightarrow (\mathsf{ij}|\mathsf{M}), && \text{(LM.iv)}\\
(\mathsf{ij}|\mathsf{kM}) \wedge (\mathsf{ik}|\mathsf{lM}) \wedge (\mathsf{jl}|\mathsf{iM}) \wedge (\mathsf{kl}|\mathsf{jM}) &\Rightarrow (\mathsf{ij}|\mathsf{M}), && \text{(LM.v)}
\end{aligned}
$$

for all distinct $\mathsf{i}, \mathsf{j}, \mathsf{k}, \mathsf{l} \in \mathsf{N}$ and $\mathsf{M} \subseteq \mathsf{N} \setminus \mathsf{ijkl}$. This result was confirmed by Drton and Xiao [DX10] who used this classification in their investigation of smoothness of Gaussian CI models motivated by the influence of model geometry on the performance of statistical tests. Computations based on the notion of *gaussoid orientability* from [BDKS19] quickly derive the Lněnička-Matúš axioms as necessary. The classification of CI structures on five random variables is not finished. There are 60 212 776 million gaussoids in 508 817 classes modulo isomorphy. The use of duality cuts this number roughly in half. This was determined in [BDKS19].

A gaussoid which satisfies (Ascension) is *graphic*. Matúš proved in [Mat97] that such gaussoids are precisely the CI structures derived from undirected graphs as in Example 1.1. These are also called *Markov networks* in the literature. There is extensive literature on the subject of Gaussian graphical models; see for instance [Lau96, Sul18] for an old and a new overview. Graphical models describe statistical models via different kinds of graphs and the goal is to relate the statistical properties of the model to the combinatorics of the graph. Next to Markov models, the most well-known graphical models are *Bayesian networks*; see also [Pea88, Stu05] for introductions. These models are **conditional independence models** and therefore their study, for Gaussian distributions, forms a special case of the CI theory pursued here. However, the questions which are of interest in the general Gaussian CI theory, such as:

The realizability problem Which CI structures are realizable by the given type of distribution (or graphical model)?

The inference problem Which CI statements are implied on every distribution of the specified type by a given set of statements?

Model geometry What is the topological or differential-geometric structure of CI models from this type?

are all solved for these types of graphical models. The combinatorics of the graph gives a complete solution to the inference problem and realizability problems. The statistical models have rational polynomial parametrizations [Sul18, Section 13.1 and 13.2], which imply the solutions to realizability and implication as well as certain facts about the model geometry.

In the general Gaussian CI theory, these questions are highly non-trivial. Chapter 2 collects general preliminary material. Afterwards, Chapter 3 develops the structure theory of Gaussian CI structures using matrix-algebraic tools and introduces the geometric view on the inference problem. Building on these tools, Chapter 4 deals with more complicated structure theory and their effect on the axioms of Gaussian CI. Chapter 5 proves that the inference and realizability problems for Gaussians are hard by multiple measures — complexity-theoretically, algebraically and topologically. Having established the hardness results, Chapter 6 develops approximations to the inference problem which are more tractable in practice. Chapter 7 closes with a summary of the results and discussion of open problems.

2

Context from logic, algebra and geometry

This chapter compiles well-established concepts and results from logic, complexity theory, algebra and geometry, which are used throughout the thesis, in a form which is most suitable for our applications.

2.1 Lattices and formal concept analysis

Lattices. A central theme of this thesis is the duality between the Gaussian realizability problem for CI structures and the CI inference problem for Gaussians. They both describe the same concept, the former via "objects" and the latter via "attributes". This theme reappears for each approximation to the inference problem and each time takes the form of a *Galois connection*. The related concepts from lattice theory are presented in this section and examples are given later in this chapter for concrete fields of mathematics. For more details see [Ber15, Chapter 6] or [Grä11, Chapter 1] or [Stu05, Section 5.4].

A set X is *partially ordered* by a binary relation $\le$ if it satisfies

$$x \le x, \tag{Reflexivity}$$
$$x \le x' \wedge x' \le x'' \Rightarrow x \le x'', \tag{Transitivity}$$
$$x \le x' \wedge x' \le x \Rightarrow x = x'. \tag{Antisymmetry}$$

"Partially ordered set" is abbreviated to "*poset*". A poset is a *lattice* if every finite subset $A \subseteq X$ has an *infimum* (or greatest lower bound) $\bigwedge A$ and a *supremum* (or least upper bound) $\bigvee A$. The lattice is *complete* if infima and suprema exist for all subsets, not necessarily just finite ones. The real numbers with their natural order are a (totally ordered) lattice but not a complete lattice. The *boolean lattice* of a set A consists of the powerset $\mathcal{P}(A)$ ordered by inclusion and is the canonical example of a complete lattice. To every lattice $(X, \le)$ there is a *dual* or *opposite* lattice $(X^{\mathsf{op}}, \ge)$ which is ordered in reverse: $x' \ge x :\Leftrightarrow x \le x'$. It is easily seen that the axioms of lattices are self-dual, so $\cdot^{\mathsf{op}}$ is an involution on the theory of lattices: every true statement about lattices (in the first-order language of posets) can be formally dualized to yield another true statement about (opposite) lattices.

Closure operators. A *closure operator* $\overline{\cdot} : X \to X$ is a map on a lattice X satisfying

$$\overline{x} \ge x, \tag{Extensivity}$$
$$x \le x' \Rightarrow \overline{x} \le \overline{x'}, \tag{Monotonicity}$$
$$\overline{\overline{x}} = \overline{x}. \tag{Idempotency}$$

The fixed points $\overline{x} = x$ are the *closed elements* of X with respect to $\overline{\cdot}$. By idempotency, the fixed points coincide with the image of $\overline{\cdot}$. The closed elements inherit a (complete) lattice structure from X. Closure operators appear in many areas of mathematics. Every topology is induced by a closure operator on the boolean lattice of the underlying space, which maps every set of points to the smallest closed set (in the topology) containing them. Similar operations of "closing" or "generating" of subgroups, linear subspaces, ideals, etc. are closure operators on suitable (semi)lattices.

The dual of a closure operator is an *interior operator*. In its definition, only the extensivity property is not self-dual and is replaced by the retraction property, so that the interior of an object is always lower in the lattice.

Galois connections. A *Galois connection* between two lattices X and Y consists of a pair of antitone maps $\alpha : X \to Y$ and $\omega : Y \to X$ such that $\omega\alpha : X \to X$ and $\alpha\omega : Y \to Y$ are closure operators.

Theorem: [Ber15, Lemma 6.5.1]. The restrictions of α and ω to each other's image establish a lattice antiisomorphism between the closed elements $\alpha(X)$ and $\omega(Y)$.

Galois connections between boolean lattices are conveniently specified by a binary relation. Consider a set $\mathcal{O}$ of "objects" and a set $\mathcal{A}$ of "attributes" and let $\diamond$ be a relation between $\mathcal{O}$ and $\mathcal{A}$. This defines a Galois connection between $\mathcal{P}(\mathcal{O})$ and $\mathcal{P}(\mathcal{A})$ via

$$\begin{aligned}\alpha(O) &\coloneqq \{\, a \in \mathcal{A} : o \diamond a \ \ \forall o \in O \,\}, \\ \omega(A) &\coloneqq \{\, o \in \mathcal{O} : o \diamond a \ \ \forall a \in A \,\}.\end{aligned}$$

That this construction gives indeed a Galois connection is shown in [Ber15, Lemma 6.5.1]. The closure operator on $\mathcal{O}$ *saturates* a set of objects with respect to all attributes it possesses. The other closure operator on $\mathcal{A}$ *infers* all attributes which are implied on $\mathcal{O}$ by a given set of attributes. Section 2.4 contains a number of examples of this construction.

2.2 Properties and propositions

Property lattice. For the purpose of accuracy, we take a formal approach to the notion of a "property" of CI structures. For example "being realizable by a regular Gaussian distribution" is a property of CI structures. It is invariant under isomorphy and we generally only care about isomorphy-invariant properties. Owing to the Isomorphy convention, we represent such a property by a subset $\mathfrak{p}$ of the countable direct product of boolean lattices

$$\mathfrak{P} \coloneqq \bigtimes_{n \geq 1} \mathcal{P}(\mathcal{P}(\mathcal{A}_n)),$$

where we identify any finite set N of size n with the distinguished set $[n] = \{1, \ldots, n\}$ (up to an isomorphy between CI structures over N and over $[n]$) and use the abbreviation $\mathcal{A}_n \coloneqq \mathcal{A}_{[n]}$. Each component $\mathfrak{p}(n)$ is a subset of $\mathcal{P}(\mathcal{A}_n)$ which contains all CI structures $([n], \mathcal{A})$ which satisfy the property under consideration. Realizability by a regular Gaussian is then the element $\mathfrak{g} \in \mathfrak{P}$ with $\mathcal{G} \in \mathfrak{g}(n)$ if and only if $\mathcal{G} = [\![\Sigma]\!]$ for some $\Sigma \in \mathrm{PD}_n$.

The set $\mathfrak{P}$ naturally has the structure of a lattice and is called the *property lattice*. Its order relation $\leq$ compares properties by generality. Let $\mathfrak{p} \leq \mathfrak{q}$ be properties. Thus for all $n \geq 1$ we have $\mathfrak{p}(n) \subseteq \mathfrak{q}(n)$ and $\mathfrak{p}$ is *sufficient* for $\mathfrak{q}$ and, equivalently, $\mathfrak{q}$ is *necessary* for $\mathfrak{p}$. This formalism is suitable to discuss properties of properties as well, such as minor-closedness of Gaussian realizability: this means that for every n and $\mathcal{G} \in \mathfrak{g}(n)$ all of its k-minors lie in $\mathfrak{g}(k)$.

Propositional calculus. A key object in this thesis are CI inference rules. These are boolean formulas whose variables are CI statements, for instance Weak transitivity (G.iv):

$$(\mathsf{ij}|\mathsf{L}) \wedge (\mathsf{ij}|\mathsf{kL}) \;\Rightarrow\; (\mathsf{ik}|\mathsf{L}) \vee (\mathsf{jk}|\mathsf{L}).$$

These inference rules express closure properties of elements in the property lattice of CI structures $\mathfrak{P}$. Since propositional logic is also embedded into first-order logic, we dedicate this section to the few elementary results on boolean formulas required in what follows.

It is well-known that the boolean lattice $(\mathcal{P}(X), \leq)$ is the same algebraic structure as the boolean algebra $(\{\bot, \top\}^X, \wedge, \vee)$ where each map $f : X \to \{\bot, \top\}$ is identified with $\{x \in X : f(x) = \top\} \in \mathcal{P}(X)$ and $\wedge$ and $\vee$ are the just the infimum and supremum operators of the lattice under this identification; this is discussed in detail in [Grä11, Chapter I]. We abbreviate $\mathbb{2} \coloneqq \{\bot, \top\}$.

A *boolean function* is any map $f : \mathbb{2}^X \to \mathbb{2}$. Such a function may be written in multiple ways as a formula in variables $(\xi_i)_{i \in X}$ and using the operators $\wedge$ and $\vee$ as well as the involution $\neg$ on $\mathbb{2}$ which swaps $\bot$ and $\top$. Each boolean function f may be written in *disjunctive normal form (DNF)* and *conjunctive normal form (CNF)*. The DNF is characterized as a disjunction over conjunctions of variables and negated variables, that is,

$$\bigvee_i \bigwedge_j \ell_{ij}, \text{ where } \ell_{ij} = \xi_{ij} \text{ or } \neg\xi_{ij}.$$

The expressions for negated or non-negated variables ℓ_{ij} are *literals*. Dually, the CNF is characterized by being a conjunction over disjunctions of literals:

$$\bigwedge_i \bigvee_j \ell_{ij}.$$

The inner disjunctions are referred to as *clauses*. Formulas in DNF and CNF exist for every boolean function but they are not unique. Their existence is easy to see. Let $f : \mathbb{2}^X \to \mathbb{2}$ and $A = \left\{ a \in \mathbb{2}^X : f(a) = \top \right\}$. Then a DNF for f is given by

$$\bigvee_{a \in A} \bigwedge_{i \in X} \ell_{ai}, \text{ with } \ell_{ai} = \begin{cases} \xi_i, & \text{if } a_i = \top, \\ \neg\xi_i, & \text{otherwise.} \end{cases}$$

Obviously, the only vectors $\xi \in \mathbb{2}^X$ which satisfy this formula are those in A, so the function described by the formula is precisely f. A CNF is constructed analogously as

$$\bigwedge_{a \in A^{\mathsf{c}}} \bigvee_{i \in X} \ell'_{ai}, \text{ with } \ell'_{ai} = \begin{cases} \neg\xi_i, & \text{if } a_i = \top, \\ \xi_i, & \text{otherwise.} \end{cases}$$

These constructions illustrate how a DNF can be thought of as an inner description of the function f, effectively listing — possibly in a compressed way — all the inputs to f which evaluate to $\top$, whereas a CNF describes f by a list of constraints which an input has to satisfy before it can be mapped to $\top$.

Consider a CNF clause $\bigvee_{i \in X_-} \neg\xi_i \vee \bigvee_{j \in X_+} \xi_j$ where X_+ and X_- index the positive and, respectively, negative literals. This clause may be written equivalently in *implication form*:

$$\bigwedge_{i \in X_-} \xi_i \;\Rightarrow\; \bigvee_{j \in X_+} \xi_j.$$

This justifies our focus on CI inference rules: when describing a property $\mathfrak{p} \in \mathfrak{P}$ of CI structures, it is sufficient for each ground set size n to list all implication formulas which are true on $\mathfrak{p}(n)$. Their conjunction is a CNF of the boolean function corresponding to $\mathfrak{p}(n) \in \mathcal{P}(\mathcal{A}_n)$ which completely describes this set.

2.3 Computational complexity

The basics of computational complexity theory presented in this section can be found in any textbook on the subject, for example [AB09] or [Sip06]. A very brief introduction is also contained in Schrijver's book on linear and integer programming [Sch98]. This section makes liberal use of the **Church–Turing thesis** [Sip06, Chapter 3] which supposes that algorithms described in sufficient detail in natural language can be formalized in the Turing machine model of computation.

Polynomial time. Computational complexity is about the resources, such as time and space, which are needed to solve certain computation tasks. These tasks are often *decision problems* — to give a "yes" or "no" answer to whether an input object has a property. A decision problem is represented by a set A of finite strings over a finite alphabet which contains *encodings* of all objects possessing the property A.

One is generally interested in the **intrinsic** complexity of the task of deciding A, that is, how well can the best algorithm for this task possibly perform when resource usage is measured in terms of the length of an input object? Such resource bounds cannot be determined in absolute terms. Firstly, the number of symbols in the finite alphabet and the chosen encoding of the objects both have an impact on the resources. For example the encoding of an integer in base 2 is longer than in base 256; or whether or not a number is even can be decided quickly if it is given in base 2 with the least significant bit first, while it takes more time if the number is given in base 3. These technical differences are uninteresting for complexity theory purposes and have to be blurred by considering algorithms modulo such implementation details. Secondly, placing any bound on the input size renders the resource usage "possibly large but constant". Therefore we care about the worst-case resource usage as a function of the input size only *asymptotically*. Landau notation is convenient to express asymptotic relationships between functions; see [AB09, Section 0.3].

These two points prompt the introduction of the complexity class **P** of all decision problems which can be solved by an algorithm whose running time is asymptotically bounded above by a polynomial in the input size. This class is defined by asymptotic time bounds and it is stable under common operations on algorithms, such as running two algorithms in sequence or calling another polynomial-time algorithm as a subroutine on polynomial-length data produced from the original input. The class is also robust against changes to the alphabet size and changes of the encoding of objects which can themselves be computed in polynomial time.

Reductions. Determining the asymptotic cost of the best algorithm for a computational problem is generally difficult, because the (asymptotically) best algorithm may not even be known. Therefore complexity theory compares problem costs by **reducing** one problem to the other.

Definition 2.1. A *Karp reduction from A to B* (also known as *many-one reduction*) is an algorithm f converting possible inputs for problem A to possible inputs for B such that $x \in A \Leftrightarrow f(x) \in B$ for all x in the problem domain of A.

Such a mapping receives an input to the decision problem associated with A and translates it to an **equivalent** input for the problem B. If a Karp reduction $A \to B$ exists, then A *Karp-reduces* to B, because the problem A can be effectively solved by solving B instead, up to a translation step carried out by f. If f can be computed in polynomial time (or *polytime* for short), it establishes a *polytime Karp reduction.* The relation of polytime Karp reducibility is transitive. If A reduces to B, then any algorithm for solving B can be turned into an algorithm for solving A with only polynomial overhead. In this sense A is at most as hard as B. This transitive relation is not a partial order because it is not antisymmetric. Instead, it induces an equivalence relation on decision problems: two decision problems are equivalent if they can be reduced to each other with just polynomial overhead. In this case, they are regarded as computationally equally hard.

There are more notions of reducibility which induce finer or coarser equivalence relations on problems. We shall also need the following:

Definition 2.2. A *Turing reduction from A to B* is an oracle-algorithm f which solves problem A. The algorithm has unlimited access to an *oracle* for the problem B. This is a blackbox device which solves problem B for an input generated by during the execution of f in constant time.

Every Karp reduction f is a Turing reduction by executing $f(x)$ and then in the last step making one query to the oracle about whether $f(x) \in B$, returning the oracle's answer unaltered. Thus Turing-reducibility is coarser than Karp-reducibility.

Completeness and `SAT`. Computational problems can be grouped into *complexity classes* which, by some measure, require the same resources to solve. Famous examples are the classes **P** and **NP** of decision problems which are solvable in deterministic and, respectively, non-deterministic polytime. These classes are downward-closed under Karp-reducibility. Thus whenever a decision problem B is in **NP** and A reduces to B in polytime, then A is in **NP**. The converse of this was investigated in the seminal paper by Cook [Coo71]: he proved that there exists a single problem denoted `SAT` such that **NP consists** of all decision problems with a polytime Karp-reduction to `SAT`. This makes `SAT` a representative of the hardest problems in **NP**. Such problems are **NP**-*complete*. A problem to which all problems in **NP** reduce but which may itself be too hard to lie in **NP** is **NP**-*hard*. Similar notions can be defined for other complexity classes.

`SAT` stands for the *boolean satisfiability problem*: given a boolean formula in conjunctive normal form, determine whether there exists an assignment of truth values to its variables which satisfies the formula. This is the prototypical **NP**-complete problem and many others have since been found. Because of its utility for modeling general problems, `SAT`-related tasks other than deciding and certifying a satisfying assignment are studied. In particular, we also make use of software for solving the `#SAT` and `AllSAT` problems which count the satisfying assignments or list all of them, respectively. Fast `SAT` solvers include `MiniSAT` [SE05] and `CaDiCaL` [Bie19]. For `#SAT` problems there is `DSHARP` [MMBH12], or the probabilistic model counter `GANAK` [SRSM19]. All satisfying assignments can be enumerated by the tools of Toda and Soh [TS16] based on `MiniSAT`.

The CNF input format for `SAT` is not a substantial restriction. Given any boolean formula it can be checked in polynomial time whether it is in CNF and an equisatisfiable CNF formula can be computed for it. The latter result is attributed to Tseitin who proved it in the introduction to [Tse83], citing his earlier work in Russian:

Tseitin transform. To every boolean formula an equisatisfiable boolean formula in CNF can be computed in polynomial time.

Proof sketch. Read the boolean formula into an abstract syntax tree whose internal nodes are operators and whose leaves are variables. This tree can be built in polynomial time and the number of nodes is linear in the formula length. To each internal node introduce a new variable. It is easy to find a CNF for each of the operations $\{\wedge, \vee, \neg\}$ that may appear. The following translation table shows how to enforce that the new variable Z introduced for an operation $X \odot Y$ in the syntax tree really takes the intended value $X \odot Y$, where X and Y are the variables introduced for the subexpressions in the syntax tree.

$Z = X \wedge Y$: $(\neg X \vee \neg Y \vee Z) \wedge (X \vee \neg Z) \wedge (Y \vee \neg Z)$.
$Z = X \vee Y$: $(X \vee Y \vee \neg Z) \wedge (\neg X \vee Z) \wedge (\neg Y \vee Z)$.
$Z = \neg X$: $(X \vee Z) \wedge (\neg X \vee \neg Z)$.

All of these CNFs are then concatenated and at the end the satisfiability of the entire formula is required by appending the clause R, where R is the new variable introduced for the root of the syntax tree. This formula is in CNF, can be constructed in polynomial time and is evidently equisatisfiable. $\square$

It should be noted that in the Tseitin transform each variable of the original boolean formula appears at most a constant number of times more often than in the original formula.

2.4 First-order theories in geometry

Model theory is a branch of mathematical logic which places the focus on *languages* and *structures* which interpret the symbols of the language. One writes axiom systems in the given language and studies the structures which fulfill them in their attached interpretation. Such structures are the *models* of the set of axioms. Conversely, given a specific model $\mathcal{M}$, one may ask which subsets of $\mathcal{M}$ are *definable* by formulas in the chosen language and which statements are true in $\mathcal{M}$. These statements comprise the *theory* of $\mathcal{M}$. For example, consider the quantified formula "$\exists x : x \cdot x = 1 + 1$". It is written in the first-order language of rings, which provides first-order variables like x which stand for *elements* of a structure in which the formula is interpreted; it also provides $\exists$ and $\forall$ quantifiers for those variables, operation symbols for addition, subtraction and multiplication as well as the constants 0 and 1. The obvious interpretation of this sentence in the real numbers is "true", in the rationals it is "false". A theory is *sound* if for every sentence φ in the language not both, φ and $\neg\varphi$, are contained in it, and it is *complete* if at least one of the two is contained.

The point of this approach is to deliberately and explicitly restrict, via the language, the kind of statements that may be considered about a mathematical structure. This restriction can result in desirable properties such as algorithmic decidability of the theory — while being sufficient to express the objects and theorems of interest. The exposition of the basic geometric theories used throughout this thesis is structured in this section through the model-theoretic lens. Each section follows the same pattern of introducing the language, the Galois connection of the theory, the definable and the closed sets and a quantifier elimination result. This pattern touches on the fundamental geometric results of each area in a unifying fashion.

The languages and theories discussed below are all based on classical first-order logic. Many textbooks on model theory explain this setup with great care to distinguish between syntax and semantics of formulas. In particular they give a recursive definition of well-formed formula in a given language and correspondingly a definition of truth in a structure. A quick but complete introduction can be found in [Mar02, Chapter 1]. However, all of these notions are introduced like an experienced mathematician would expect them to be.

2.4.1 Algebraic geometry. This work requires only elementary results from algebraic geometry and its first-order model theory. Proofs, often constructive, are generally found in [CLO15] or in [Har77, Chapter 1] and its references. For the model-theoretic parts, more background information is provided by [MT03, Chapters 1 and 2] or [Mar02, Chapter 3]. Facts and definitions concerning field theory are taken from [DF04, Chapter 13].

The *language of rings* contains the two constants 0 and 1, and operations symbols for addition, additive inverse and multiplication. Multiplicative inverse is not included in the language to avoid having to consider "0^{-1}" and the like as syntactically valid terms with ill-defined semantics in rings. Terms in this language are well-formed strings of operations of addition, subtraction and multiplication on the constants and variables. Formulas are quantified boolean combinations of equations and inequations. At this level, two terms are equal if and only if they coincide as strings modulo the universal laws of propositional calculus. The standard definition of *ring* (understood in this thesis always to be commutative and with unity) can be written down in this language, largely to the effect of introducing additional relations on the terms of this language, e.g., commutativity of multiplication or the distributive law:

$$\forall x, y : x \cdot y = y \cdot x,$$
$$\forall x, y, z : x \cdot (y + z) = x \cdot y + x \cdot z,$$

Zero divisors can be forbidden, turning commutative rings into integral domains:

$$\forall x, y : (x \cdot y = 0 \;\Rightarrow\; x = 0 \vee y = 0)\,.$$

So far, all axioms are universal, but to turn an integral domain into a field, an existential quantifier is necessary:

$$\forall x \exists y : (x \neq 0 \;\Rightarrow\; xy = 1)\,.$$

Every structure $\mathbb{K}$ of the language of rings which satisfies the field axioms must be a field. Due to the axioms, the interpretation of different terms, such as "$(x+y)+z$" and "$x+(y+z)$"', may be equivalent in $\mathbb{K}$. The equivalence classes are precisely the multivariate polynomials with integer coefficients in countably many variables, $\mathbb{Z}[x_1, x_2, \dots]$, where $\mathbb{Z}$ coefficients are identified with their image under the canonical ring homomorphism $\mathbb{Z} \to \mathbb{K}$. Formulas are quantified boolean combinations of polynomial equations and inequations. While there is no a priori bound on the number of variables in any formula, every formula and every proof in the first-order theory of fields consists of finitely many symbols and so for most concerns we work in a polynomial ring in finitely many variables. In this setting, the fundamental theorem in commutative algebra which opens the gates for applications of first-order logic in geometry is

Hilbert's Basis Theorem. A polynomial ring over a noetherian ring is noetherian. In particular, every ideal in $\mathbb{K}[x_1, \dots, x_n]$ is finitely generated.

Algebraic closure. Many results in algebraic geometry require an algebraically closed field. A field $\mathbb{K}$ is *algebraically closed* if every non-constant univariate polynomial $f \in \mathbb{K}[x]$ factors into linear polynomials over this field. This property can be expressed using countably many axioms in the first-order language of rings:

$$\forall a_0, a_1, \dots, a_m \exists z : a_m z^m + \dots + a_1 z + a_0 = 0,$$

one for each degree m. Every field $\mathbb{K}$ is contained in an algebraically closed field $\mathbb{K}^*$. The subset of elements of $\mathbb{K}^*$ which are algebraic over $\mathbb{K}$ form an intermediate field between $\mathbb{K}$ and $\mathbb{K}^*$ which remains algebraically closed and is called an *algebraic closure* of $\mathbb{K}$. Since it is unique up to field isomorphisms, we speak of *the* algebraic closure and denote it by $\overline{\mathbb{K}}$. It is characterized by being an algebraic and algebraically closed extension of $\mathbb{K}$. For details, see [DF04, Section 13.4].

Characteristic. Every field contains a smallest subfield which is generated by 1 and called its *prime field*. The *characteristic* $\operatorname{char}\mathbb{K}$ of a field $\mathbb{K}$ is the cardinality of its prime field — except when the prime field is infinite, in which case the characteristic is *zero* by convention. There exists exactly one prime field (up to isomorphy) per characteristic, these are the finite fields $\mathbb{F}_p$ for every prime p and the field of rational numbers $\mathbb{Q}$ of characteristic zero. Let **ACF** be the theory of algebraically closed fields, i.e., all sentences in the first-order language of rings which are true in every algebraically closed field, and $\mathbf{ACF}_k$ the theory of algebraically closed fields of fixed characteristic k. These two theories differ, for example, by the truth of sentences such as $2 = 0$. This is true in $\mathbf{ACF}_2$ but there exist algebraically closed fields such as $\mathbb{C}$ where it is false, hence $2 = 0$ is not true in **ACF**. Its negation $2 \neq 0$ is also not true in **ACF** since $\overline{\mathbb{F}_2}$ falsifies it. This shows that **ACF** is not complete, because it contains neither $2 = 0$ nor $\neg(2 = 0)$.

Zariski topology. The point of departure into geometry is the following essential Galois connection between the points $a \in \mathbb{K}^n$ and polynomials $f \in \mathbb{K}[x_1, \dots, x_n]$:

$$a \diamond f \;:\Leftrightarrow\; f(a) = 0.$$

In the terminology of Section 2.1, points in the affine space $\mathbb{K}^n$ are the objects and polynomials f the attributes in this connection. Let $\mathcal{I} : \mathcal{P}(\mathbb{K}^n) \to \mathcal{P}(\mathbb{K}[x_1, \dots, x_n])$ and $\mathcal{V} : \mathcal{P}(\mathbb{K}[x_1, \dots, x_n]) \to \mathcal{P}(\mathbb{K}^n)$ denote the maps associated to the Galois connection. The induced closure operators correspond to a topological closure operator, namely the one of the *Zariski topology*, in which the closed sets of points are *varieties* and the closed sets of polynomials are characterized by the following theorem as the *radical ideals*:

Hilbert's Nullstellensatz. Let $\mathbb{K}$ be algebraically closed and $\mathcal{J}$ an ideal in $\mathbb{K}[x_1, \dots, x_n]$. Then $\mathcal{I}(\mathcal{V}(\mathcal{J})) = \sqrt{\mathcal{J}} \coloneqq \{\, f \in \mathbb{K}[x_1, \dots, x_n] : \exists m \geq 0 : f^m \in \mathcal{J} \,\}$.

By Hilbert's Basis Theorem, all ideals are finitely generated and in particular varieties can be dealt with by proxy of finite systems of polynomial equations which generate their associated radical ideal. Refer to the "Algebra-Geometry Dictionary" [CLO15, Chapter 4] for various set-theoretic operations on varieties and their ideal-theoretic counterparts.

Abstract solution sets. Following the exposition in [BS89, Chapter 4] we introduce the abstract solution space of a polynomial system over a **not necessarily** algebraically closed field $\mathbb{K}$ and derive another interpretation of Hilbert's Nullstellensatz. The *spectrum* $\operatorname{Spec}\mathcal{R}$ of a ring $\mathcal{R}$ is the set of all prime ideals in $\mathcal{R}$. If $\mathcal{R} = \mathbb{K}[x_1, \dots, x_n]/\mathcal{J}$ is the *coordinate ring* of an affine variety $V = \mathcal{V}(\mathcal{J})$, i.e., the equivalence classes of polynomial functions on V, then the homomorphisms $\mathcal{R} \to \mathbb{L}$ to field extensions $\mathbb{L}$ of $\mathbb{K}$ correspond to $\operatorname{Spec}\mathcal{R}$ via the construction that associates to each prime p the fraction field $\operatorname{Quot}(\mathcal{R}/p)$. In this sense, the spectrum may be seen as the set of all homomorphisms from the coordinate ring into extensions of the defining field $\mathbb{K}$. However, all such homomorphisms are evaluation homomorphisms by the universal property of polynomial rings and thus the spectrum is also the set of all points on the variety V in field extensions of $\mathbb{K}$. Such a point in a field extension $\mathbb{L}$ is an $\mathbb{L}$-*rational point*. Over an algebraically closed field $\mathbb{K}$ with $\mathcal{R} = \mathbb{K}[x_1, \dots, x_n]$, the spectrum consists of all irreducible varieties, which are viewed as "generic points" in the affine space $\mathbb{K}^n$ (in the sense that every affine variety is uniquely identified by a sample of one generic point per irreducible component), whereas the maximal ideals are the honest, geometric points in $\mathbb{K}^n$.

The spectrum is useful in tracking the points which satisfy polynomial equations and inequations in extensions of the defining field. The set of solutions to a system of polynomial equations $\{\, f_i = 0 \,\}$ is a variety and it is represented by its coordinate ring $\mathbb{K}[x_1, \dots, x_n]/\langle f_i \rangle$.

If inequations $g_j \neq 0$ are permitted, the solution set is no longer a variety in $\mathbb{K}^n$, but a *constructible set*, i.e., a finite union of locally closed sets in the Zariski topology. The constructible sets are precisely the ones definable over a field in the language of rings. The ring-theoretic operation to model inequations is *localization* $\mathcal{U}^{-1}\mathcal{R}$ of the coordinate ring $\mathcal{R}$ at the multiplicative monoid $\mathcal{U}$ generated by g; see [Kem11, Chapter 6].

Lemma 2.3: [Kem11, Lemma 1.22 and Theorem 6.5]. Let $\mathcal{R}$ be a commutative ring, $\mathcal{I}$ an ideal and $\mathcal{U}$ a multiplicatively closed set. Then

(i) $\operatorname{Spec}(\mathcal{R}/\mathcal{I}) = \{ p \in \operatorname{Spec}\mathcal{R} : p \supseteq \mathcal{I} \}$, and
(ii) $\operatorname{Spec}(\mathcal{U}^{-1}\mathcal{R}) = \{ p \in \operatorname{Spec}\mathcal{R} : p \cap \mathcal{U} = \emptyset \}$.

The *abstract solution set* to the polynomial system $\{ f_i = 0, g_j \neq 0 \}$ is the spectrum $\operatorname{Spec}(\mathcal{U}^{-1}(\mathbb{K}[x_1, \dots, x_n]/\mathcal{I}))$. It consists of all prime ideals $p \in \mathcal{R}$ which contain $\mathcal{I}$ and do not intersect $\mathcal{U}$. Thus, it contains the generalized points on which the f_i vanish and the g_j do not vanish. From this point of view, we get this other version of Hilbert's Nullstellensatz:

Alternatives in algebraic geometry. Let $\mathbb{K}$ be any field and $f_i, g_j \in \mathbb{K}[x_1, \dots, x_n]$. Then exactly one of the following two cases occurs:

(a) There exists a point $a \in \overline{\mathbb{K}}^n$ with $f_i(a) = 0$ and $g_j(a) \neq 0$ for all i and j.
(b) There exist $h_i \in \mathbb{K}[x_1, \dots, x_n]$ and $m \geq 0$ such that $\sum_i f_i h_i = \left(\prod_j g_j\right)^m$.

Proof. The two conditions obviously cannot occur simultaneously in the absence of zero divisors from $\mathbb{K}$. Suppose that (b) does not apply. Then the ideal $\mathcal{I}$ generated by the f_i and the multiplicative monoid $\mathcal{U}$ generated by the g_j do not intersect. The primary decomposition of $\mathcal{I}$ (see [CLO15, Section 4.8]) then provides a point in the abstract solution set of the polynomial system and hence in the algebraic closure of $\mathbb{K}$. □

The condition (b) in the above theorem is equivalent to $0 \in \mathcal{I} + \mathcal{U}$, i.e., the element-wise sum of the ideal $\mathcal{I}$ and the monoid $\mathcal{U}$ in $K[x_1, \dots, x_n]$. The point of this formulation is that a relation such as $0 \in \mathcal{I} + \mathcal{U}$ is **absurd** when the solution space is non-empty, for every point in the solution space evaluates to zero on every element of $\mathcal{I}$ and to non-zero on every element of $\mathcal{U}$, hence every element of $\mathcal{I} + \mathcal{U}$ must be non-zero. The condition $0 \in \mathcal{I} + \mathcal{U}$ may be equivalently stated as $\mathcal{I} \cap \mathcal{U} \neq \emptyset$ which relates it to the abstract solution set. This shows that either there exists a point in the solution set of the polynomial system all of whose coordinates are algebraic numbers over $\mathbb{K}$ or there exists an algebraic proof of the unsolvability of the system in the form of a polynomial in $\mathcal{I} \cap \mathcal{U}$ which has coefficients in $\mathbb{K}$. Thus, in particular, the solvability of a polynomial system with integer coefficients over an algebraically closed field of a given characteristic can be refuted over the prime field — if it can be refuted at all.

Quantifier elimination and the Lefschetz principle. The procedures introduced above for dealing with varieties via ideals are **effective** in the sense that they can in principle be carried out on a computer. Implementations are available in computer algebra systems such as `Mathematica` [WM] or `Macaulay2` [GS] based on *Gröbner bases*; see [CLO15, Chapter 2].

An important model-theoretic property of the theory **ACF** is that it admits *quantifier elimination* (which, again, can be implemented using Gröbner bases). This means that every formula in the language of rings is equivalent, modulo **ACF**, to a quantifier-free formula; see [MT03, Section 2.4]. In particular every sentence such as $\forall x : px = 0$ is equivalent to a quantifier-free formula ψ, but in a sentence every variable is bound by a quantifier, so ψ cannot contain any variables. It follows that any such ψ is a boolean formula over variable-less sentences in the language of rings. These are all of the form $m = n$ for integers m, n.

The sentence $\forall x : px = 0$, for example, is equivalent to $p = 0$. The boolean combination of these sentences points out exactly which **characteristics** make the sentence true.

Quantifier elimination in ACF. The theories **ACF** and $\mathbf{ACF}_k$, for all k, admit quantifier elimination in the language of rings. They all are decidable and the $\mathbf{ACF}_k$ are complete.

The fact that existential quantifiers can be eliminated in the language of rings, given the theory of algebraically closed fields, implies that the image of a coordinate projection of any constructible set is constructible:

Chevalley's Theorem. A projection of a constructible set over an algebraically closed field is constructible.

A corollary to quantifier elimination and the well-known Compactness theorem in first-order logic [Mar02, Lemma 2.1.14] yields the first-order Lefschetz principle [Mar02, Corollary 2.2.10]:

Lefschetz Principle. Let $F = \{ f_i = 0, g_j \neq 0 \}$ be a system of finitely many polynomial equations and inequations with integer coefficients, i.e., $f_i, g_j \in \mathbb{Z}[t_1, \dots, t_p]$.

— F has a solution over some algebraically closed field of characteristic k if and only if it has a solution over every such field, in particular the algebraic closure of the prime field.

— If F has a solution in characteristic zero, then it has a solution in the algebraic numbers $\overline{\mathbb{Q}}$. This is the case if and only if F has a solution in $\overline{\mathbb{F}_p}$ for all but finitely many primes p.

It follows that a polynomial system which has a solution over $\mathbb{C}$ must have a solution in $\overline{\mathbb{F}_p}$ for almost all p. Since each such solution consists of finitely many numbers of finite algebraic degree over $\mathbb{F}_p$, the solution is already found in a *finite field* extending $\mathbb{F}_p$.

Definition 2.4. The *characteristic set* $\chi(f_i, g_j)$ of a polynomial system $\{ f_i = 0, g_j \neq 0 \}$ is the set of all k such that the system has a solution over some field of characteristic k.

The characteristic set of a polynomial system can be computed with quantifier elimination. It is either finite and excludes zero or is a cofinite set of primes and includes zero.

2.4.2 Convex and polyhedral geometry. Let $\mathbb{R}^n$ be the set of objects and its dual space $(\mathbb{R}^n)^*$ the set of attributes of the following Galois connection:

$$x \diamond \alpha \;:\Leftrightarrow\; \alpha(x) \geq 0.$$

Standard convexity theory [HW20, Chapter 1] or [Zie95, Chapters 1 and 2] implies that the closed sets of points are the *closed convex cones* and the closed sets of attributes are also the closed convex cones in the dual space. See also [Stu93] for a self-contained introduction. The limitation to homogeneous linear inequalities and therefore cones is not substantial. An arbitrary convex set may be studied equivalently as the cone it generates in a higher-dimensional space by adding a homogenizing coordinate. The convex set $\operatorname{conv} A$ or convex cone $\operatorname{cone} A$ generated by another set A is the intersection of all convex sets or convex cones containing A. An equivalent (extensional) characterization of these sets is as the closure under *convex* and *conic* combinations:

$$\operatorname{conv} A := \left\{ \sum_i c_i a_i : c_i \geq 0 \text{ and } \sum_i c_i = 1 \text{ for finitely many } a_i \in A \right\},$$
$$\operatorname{cone} A := \left\{ \sum_i c_i a_i : c_i \geq 0 \text{ for finitely many } a_i \in A \right\}.$$

Polyhedra and spectrahedra. The above Galois connection does not have an analogue of Hilbert's Basis Theorem. There are convex cones which are not *polyhedral*, i.e., which are not describable by finitely many linear inequalities.

Definition 2.5. A *polyhedron* is a convex set described by finitely many affine-linear inequalities: $P = \{ Ax \le b \}$ for a matrix A and a fitting right-hand side vector b. A bounded polyhedron is a *polytope*. It is the convex hull of finitely many points.

By introducing slack variables and splitting each variable into a non-positive and a non-negative part, polyhedra may equivalently be studied in higher dimensions as solution sets to systems of the form $\{ Ax = b, x \ge 0 \}$.

Example 2.6. A prominent example of a non-polyhedral, closed convex cone is the set of positive-semidefinite matrices PSD_n. Consider the affine slice of this cone which sets all diagonal entries to one and furthermore the $(1,2)$-entry to zero: $\left(\begin{smallmatrix} 1 & 0 & x \\ 0 & 1 & y \\ x & y & 1 \end{smallmatrix}\right)$. The resulting 2-dimensional set of matrices is described by the inequality $x^2 + y^2 \le 1$ and is thus a disc of radius 1. This set is not polyhedral because it has infinitely many extreme points, but it was obtained by imposing only affine-linear equations on PSD_3. Therefore PSD_3, and in fact PSD_n, $n \ge 3$, is not polyhedral. △

Definition 2.7. A *spectrahedron* is an intersection of PSD_n with a linear subspace of the symmetric matrices Sym_n.

This is analogous to the definition of polyhedra as solutions to $\{ Ax = b, x \ge 0 \}$, except that x becomes a matrix variable, each row of A becomes a linear functional on Sym_n, and the non-negative orthant $\{ x \ge 0 \}$ is replaced by the positive-semidefinite cone. Spectrahedra generalize polyhedra: since **diagonal** matrices are positive-semidefinite if and only if their diagonal elements all lie in the non-negative orthant and affine-linear conditions may be imposed on them, the defining equations of any polyhedron can be encoded in a spectrahedron of diagonal matrices; see [MS21a, Proposition 12.3].

Model theory of rational polyhedra. A polyhedron P is *rational* if it can be defined as $P = \{ Ax = b, x \ge 0 \}$ where A and b have integer entries. This restriction makes it possible to study rational polyhedra as sets defined by a formal language, similar to the language of rings. The language of *ordered abelian groups* comes with a symbol $\le$ for the ordering, symbols $+$ and $-$ for the abelian group structure and the constant 0 for the identity element. The axioms for ordered groups are as expected: in addition to the separate axioms of an ordering and an abelian group, $\le$ should be monotone with respect to addition. An ordered abelian group is *divisible* if the solution $na = b$ can be solved for every b and every natural number $n \ge 1$ interpreted as the n-fold addition of a in the group. Clearly, divisibility can be formalized by countably many axioms in the language of ordered abelian groups. Let **ODAG** be the theory of non-trivial, ordered, divisible abelian groups. The smallest ordered, divisible abelian group is the trivial group $\{ 0 \}$. Any non-trivial such group must contain a non-zero element and hence, by the ordering and divisibility it must contain the additive ordered group $\mathbb{Q}$ of rational numbers. The language of ordered abelian groups may be used to express the defining equations and inequalities of rational polyhedra and **ODAG** admits quantifier elimination and is complete by [Mar02, Corollary 3.1.17]:

Quantifier elimination in ODAG. The theory **ODAG** admits quantifier elimination in the language of ordered abelian groups. It is decidable. Moreover, it is elementarily equivalent to the theory of $\mathbb{Q}$ as an ordered abelian (additive) group.

In geometric terms, the quantifier elimination and completeness results are known as **Fourier–Motzkin elimination** [Sch98, Section 12.2] and the **Farkas lemma** [Sch98, Section 7.3], respectively. The Farkas lemma is a theorem of the alternative similar to the

Alternatives in algebraic geometry. We avoid stating it here because it is the linear special case of the Positivstellensatz below. The above result implies that whenever a rational polyhedron has a real point, then it must have a rational point (since the satisfiability of the formula defining the polyhedron over any non-trivial, ordered, divisible abelian group is determined by the theory of $\mathbb{Q}$).

The face lattice. By the representation theorem [Zie95, Theorem 1.2] every polyhedron is the Minkowski sum of a polytope and a polyhedral cone. A polyhedral cone, in turn, is the direct sum of its lineality space with a pointed polyhedral cone, i.e., one where the zero vector is an extreme point. Every polytope or pointed cone comes with an associated combinatorial object — its face lattice (see [Zie95, Chapter 2]):

Definition 2.8. A subset F of a polyhedron P is a *face* if it is the entire set of maxima of some linear functional on P. Equivalently, it is the intersection of P with a supporting hyperplane. In particular, every face of a polyhedron is a polyhedron. The *face lattice* of P is the poset of its faces ordered by inclusion.

The question which finite lattices appear as the face lattices of polytopes is highly non-trivial; see [Ric97]. For a given polytope or pointed polyhedral cone, the face lattice can be explored by linear programming. To describe the face lattice, it is sufficient to find its *inference rules* of the form "if some subset of the defining inequalities are tight, which other inequalities become tight?" For a pointed cone $P = \{ Ax \geq 0 \}$, this directly translates to a (large) number of polyhedra, namely those where a subset of the inequalities are tightened to equalities and one other inequality is made strict. The strict inequality is implied by the tight inequalities if and only if the modified polyhedron is empty. This can be checked in practice by linear programming software such as `soplex` [GSW12, GSW15] (which is part of SCIP [GBE+18]) or `normaliz` [BISvdO]

2.4.3 Semialgebraic geometry. Consider again the Galois connection $x \diamond f :\Leftrightarrow f(x) = 0$ between points and polynomials from Section 2.4.1, which induces the Zariski topology on $\mathbb{R}^n$. Since $\mathbb{R}$ is not algebraically closed, Chevalley's Theorem does not hold: coordinate projections of varieties need not be describable by polynomial equations and inequations over $\mathbb{R}$. Consider for example the interval $[-1, 1]$ which is obtained as projection of the circle $x^2 + y^2 = 1$. The interval is not constructible because it is neither finite nor cofinite in $\mathbb{R}$. It follows that existential quantifiers cannot be eliminated in the language of rings for the theory of $\mathbb{R}$. To overcome this defect, the language is extended to that of *ordered rings*. The resulting theory is called **real algebra**. The primary sources of this section are [BCR98], [Mar08] and, for the model-theoretic parts, [MT03, Section 2.5] and [Mar02, Section 3.3].

Cones and ordered fields. In addition to the language of rings, the language of ordered rings features a symbol $\leq$ for the order relation. The axiom systems for rings and fields are extended by natural monotonicity properties for $\leq$ with respect to the algebraic structure and the requirement that the order be total. The order may be described by its *positive cone* $\mathcal{P} = \{ x \in \mathcal{R} : x \geq 0 \}$. The monotonicity axioms amount to the following closure properties: (i) $\mathcal{P} + \mathcal{P} \subseteq \mathcal{P}$, (ii) $\mathcal{P} \cdot \mathcal{P} \subseteq \mathcal{P}$, and (iii) $x \in \mathcal{R} \Rightarrow x^2 \in \mathcal{P}$. Any subset of a ring $\mathcal{R}$ satisfying these properties is a *cone*. A cone is *proper* if it is a strict subset of $\mathcal{R}$. This is equivalent to $-1 \notin \mathcal{P}$. The cones of $\mathcal{R}$ can be partially ordered by inclusion and there exists a smallest cone, the *sums of squares* $\sum \mathcal{R}^2$ which consist of all conic combinations of squares from $\mathcal{R}$. In the sequel we have the cases in mind where $\mathcal{R}$ is a field or a polynomial ring over a field.

Theorem 2.9: [BCR98, Theorem 1.1.8]. Let $\mathbb{K}$ be a field. The following are equivalent:

— $\mathbb{K}$ can be ordered,
— $\mathbb{K}$ has a proper cone,
— $\sum \mathbb{K}^2$ is a proper cone,
— $\sum_i x_i^2 = 0$ implies that $x_i = 0$ for all i.

In particular is the characteristic of any ordered field zero.

The fields $\mathbb{Q}$ and $\mathbb{R}$ are naturally ordered. The intermediate field $\mathbb{Q}(\sqrt{2})$ inherits a natural order from $\mathbb{R}$. On the other hand, $\mathbb{Q}(\sqrt{-1})$ cannot be ordered: we have $1^2 + \sqrt{-1}^2 = 0$ contradicting the last property in the above theorem.

Example 2.10. If $\mathbb{K}$ is an ordered field, then the field of rational functions $\mathbb{K}(\varepsilon)$ in an unknown ε can be ordered (in multiple ways). One order of interest in Section 4.1 makes ε *infinitesimally positive*, i.e., $0 < \varepsilon < x$ for all $x \in \mathbb{K}$. By the construction of the rational function field $\mathbb{K}(\varepsilon)$ and the properties of an order, this determines an extended order uniquely. By convention ε denotes a positive infinitesimal in the following. △

Every ordered field $\mathbb{K}$ comes with a positive cone $\mathcal{P}$. Because the ordering is total, this cone has the additional property that $\mathcal{P} \cup -\mathcal{P} = \mathbb{K}$. Conversely, every proper cone with this property defines an ordering of $\mathbb{K}$ via $x \le y :\Leftrightarrow y - x \in \mathcal{P}$.

In the more general case of rings, a *prime cone* is a proper cone $\mathcal{P}$ of $\mathcal{R}$ such that $xy \in \mathcal{P}$ implies $x \in \mathcal{P}$ or $-y \in \mathcal{P}$. By [BCR98, Proposition 4.3.2], these properties imply that $\mathcal{P} \cup -\mathcal{P} = \mathcal{R}$ and that $\mathcal{I}_{\mathcal{P}} \coloneqq \mathcal{P} \cap -\mathcal{P}$ is a prime ideal in $\mathcal{R}$. In this case, $\mathcal{R}/\mathcal{I}_{\mathcal{P}}$ is an integral domain and its field of fractions inherits an ordering from $\mathcal{P}$.

Theorem 2.11: [BCR98, Theorem 4.3.7]. For a ring $\mathcal{R}$ the following are equivalent:

— $\mathcal{R}$ has a proper cone,
— $\mathcal{R}$ has a prime cone,
— $\mathcal{R}$ has a homomorphism into an ordered field,
— $-1 \notin \sum \mathcal{R}^2$.

Real closure. Just like algebraically closed fields are maximal field extensions with the property of being algebraic over a ground field, we introduce *real-closed* fields as maximal ordered fields above a ground field. It turns out that this requirement makes the ordering accessible to the algebraic structure and even unique:

Theorem 2.12: [BCR98, Theorem 1.2.2]. For an ordered field $\mathbb{K}$ the following are equivalent:

— $\mathbb{K}$ is real-closed,
— $\sum \mathbb{K}^2$ is the unique ordering of $\mathbb{K}$ and every polynomial of odd degree in $\mathbb{K}[x]$ has a root in $\mathbb{K}$,
— $\mathbb{K}(\sqrt{-1})$ is algebraically closed.

Every ordered field $\mathbb{K}$ may be extended (in an order-preserving way) to a real-closed field. If this extension is even algebraic, then the real-closed field is unique up to unique isomorphism fixing the ordered ground field $\mathbb{K}$ [BCR98, Theorem 1.3.2]. This field is the *real closure* of $\widetilde{\mathbb{K}}$ of $\mathbb{K}$. Of course, the algebraic numbers in any real-closed extension $\mathbb{L}$ of $\mathbb{K}$ form a copy of $\widetilde{\mathbb{K}}$ in $\mathbb{L}$, so real closures always exist. The smallest real-closed field is the real-closure of the rationals $\widetilde{\mathbb{Q}}$. Denote the theory of real-closed fields in the language of (ordered) rings by **RCF**.

Quantifier elimination and Tarski's transfer principle. With these preparations, we are able to pinpoint the definable sets in the language of ordered rings over a real-closed field and give the analogue of Chevalley's Theorem in real algebraic geometry.

Definition 2.13. A set Z is *basic semialgebraic* if it has the form $Z = \{ f_i = 0, g_j \ge 0, h_k > 0 \}$ with families of polynomials $f_i, g_j, h_k \in \mathbb{Z}[x_1, \dots, x_n]$. It is *primary* if all inequalities are strict. A boolean combination of basic semialgebraic sets is just *semialgebraic*.

The following fundamental results are due to Tarski [Tar48] (see also [MT03, Section 2.5]):

Quantifier elimination in RCF. The theory **RCF** admits quantifier elimination in the language of ordered rings. It is complete and decidable.

Tarski–Seidenberg theorem. A projection of a semialgebraic set over a real-closed field is semialgebraic.

Tarski's transfer principle. Let $F = \{ f_i \bowtie_i 0 \}$ with $\bowtie_i \in \{ =, \ne, <, \le, \ge, > \}$ be a system of finitely many polynomial constraints with integer coefficients, i.e., $f_i \in \mathbb{Z}[x_1, \dots, x_n]$. F has a solution over some real-closed field if and only if it has a solution over every real-closed field. In particular, if a solution exists in $\mathbb{R}$, then a solution exists in a finite real extension of $\mathbb{Q}$.

Thus, all real-closed fields give the same answer to questions about the emptiness of semialgebraic sets, and first-order sentences in general. A field $\mathbb{K}$ is *archimedean* if for every $x \in \mathbb{K}$ there exists an integer n such that $x < n$ (where, as usual, n is interpreted as the n-fold summation of $1 \in \mathbb{K}$). Equivalently, there exist rational numbers in every neighborhood of zero. The property of being archimedean cannot be expressed in first-order logic with the language of ordered rings because $\widetilde{\mathbb{R}(\varepsilon)}$ and $\mathbb{R}$ are elementarily equivalent by Tarski's transfer principle but one is archimedean and the other is not. It follows that the integers are not definable in this theory because if the integers were a semialgebraic set, a first-order formula for archimedeanness would be obvious.

Another application of Tarski's transfer principle is the following observation about the abundance of algebraic numbers in semialgebraic sets. Let $Z \subseteq \mathbb{R}^n$ be semialgebraic. Then for every (small) rational $r > 0$ and $x_0 \in \mathbb{Q}^n$ the set $Z \cap B_r(x_0) = \{ x \in Z : \|x - x_0\|^2 \le r \}$ is semialgebraic. Any real point $x^* \in Z$ may be approximated by a sequence of rational points $x_i \in \mathbb{Q}^n$ in the euclidean topology and so $Z \cap B_{r_i}(x_i)$ is non-empty (containing x^*) for a suitable sequence of positive radii $r_i \to 0$. Tarski's transfer principle mandates that these semialgebraic sets are non-empty over $\widetilde{\mathbb{Q}}$ as well. This proves that real algebraic numbers are euclidean-dense in every semialgebraic set. In particular, all isolated points of a semialgebraic set must be algebraic. This is important for computer experiments, since real algebraic numbers are representable for example by their minimal polynomial and a root-isolating interval of rational numbers, which is a finite amount of data.

Stellensätze. The first-order gem of real algebra is the Positivstellensatz due to Krivine and Stengle which characterizes the polynomials which have a constant sign on a (basic closed) semialgebraic set (see [BCR98, Section 4.4] and [Mar08, Section 2.2]):

Positivstellensatz. Let $Z = \{ f_i = 0, g_j \ge 0 \}$ over a real-closed field $\mathbb{K}$ and let $\mathcal{P}$ be the cone generated by the g_j as well as f_i and $-f_i$. For any polynomial $f \in \mathbb{K}[x_1, \dots, x_n]$:

$$\begin{aligned} \boldsymbol{f|_Z = 0} &\Leftrightarrow \exists m \ge 0 : -f^{2m} \in \mathcal{P}. \\ \boldsymbol{f|_Z > 0} &\Leftrightarrow \exists p, q \in \mathcal{P} : pf = 1 + q. \\ \boldsymbol{f|_Z \ge 0} &\Leftrightarrow \exists m \ge 0 \, \exists p, q \in \mathcal{P} : pf = f^{2m} + q. \end{aligned}$$

In particular, $Z = \emptyset$ if and only if 1 vanishes on Z, which is equivalent to $-1 \in \mathcal{P}$.

Real Nullstellensatz. Let $\mathbb{K}$ be a real-closed field, $V = \{ f_i = 0 \}$ a real variety and $\mathcal{I} = \langle f_i \rangle$ the ideal of its vanishing conditions. Then f vanishes on V if and only if f belongs to the *real radical* $\sqrt[\mathbb{K}]{\mathcal{I}} := \{ f \in \mathbb{K}[x_1, \dots, x_n] : \exists m \geq 0 : f^{2m} + \sum \mathbb{K}^2 \in \mathcal{I} \}$.

The above theorems are stated in geometric terms over ordered fields. We shall have use for a more algebraic version of the Positivstellensatz which does not require real closedness in its premises and is proved in [BCR98, Proposition 4.4.1]:

Alternatives in real algebraic geometry. Let $\mathbb{K}$ be an ordered field and $Z = \{ f_i = 0, g_j \geq 0, h_k \neq 0 \}$ a semialgebraic set over $\mathbb{K}$. Let $\mathcal{I}$ denote the ideal generated by the f_i, $\mathcal{P}$ the cone generated by the g_j and $\mathcal{U}$ the multiplicative monoid generated by the h_k in $\mathbb{K}[x_1, \dots, x_n]$. Then exactly one of the following two cases occurs:

(a) There exists a point $a \in \widetilde{\mathbb{K}}^n$ with $f_i(a) = 0$, $g_j(a) \geq 0$ and $h_k(a) \neq 0$.
(b) There exist $f \in \mathcal{I}$, $g \in \mathcal{P}$ and $h \in \mathcal{U}$ such that $f + g + h^2 = 0$.

As with the Alternatives in algebraic geometry, the existence of f, g and h in the latter condition is a proof of the absurdity of the feasibility of the semialgebraic set Z. For any point $a \in Z$ must satisfy $(f + g + h^2)(a) = 0 + g(a) + h^2(a) > 0$. Positivity conditions $p_l > 0$ are split into $p_l \geq 0 \wedge p_l \neq 0$, so they contribute to both, the cone and the monoid. To decide the semialgebraic properties covered by the Positivstellensatz, it suffices to implement quantifier elimination. This is done via cylindrical algebraic decomposition [BPR06, Chapter 5] (but see also Chapters 11–14 there), which is available in `Mathematica` [WM].

2.5 Rank functions in information theory

The entropy region. Conditional independence in the setting of discrete random vectors is closely related to certain basic geometric properties of the *entropy region*, an object of study in information theory; see [Yeu05] for an introduction. The *entropy* of a discrete random variable ξ is the real quantity $H(\xi) = -\mathbb{E}[\log p]$, where p is the probability density associated to ξ.

Remark 2.14. The base of the logarithm may be either fixed or left unspecified. While the base does not matter for formula manipulations, in some applications and especially when comparing entropies, the base is crucial. Our convention for transparency is therefore to leave the base of the symbol log unspecified and to explicitly divide entropies by a factor of $\log b$ to denote that the logarithm in this case is to be understood with base b.

Entropy is a measure for the uncertainty or average information content of the random variable. For a vector $(\xi_i)_{i \in N}$ we define the *entropy vector* as the set function $h_\xi : \mathcal{P}(N) \to \mathbb{R}$ given by $K \mapsto H(\xi_K)$. This function is naturally identified with a vector in $\mathbb{R}^{\mathcal{P}(N)} \cong \mathbb{R}^{2^n}$. The set of all entropy vectors of discrete random vectors indexed by N forms the *entropy region* Γ^*_N. Certain linear functionals on entropy vectors establish the connection to CI theory. For not necessarily disjoint $I, J, K \subseteq N$ let

$$\triangle h(I, J|K) := h(IK) + h(JK) - h(IJK) - h(K).$$

It follows from the basic *Shannon information inequalities* derived in the seminal paper by Shannon [Sha48] that these functionals are non-negative on the entropic region. For a CI statement $(ij|K) \in \mathcal{A}_N$ we have a corresponding functional $\triangle(ij|K) := \triangle(iK, jK)$ which is the *conditional mutual information* between i and j given K and conditional independence $(ij|K)$ holds for ξ if and only if $\triangle h_\xi(ij|K) = 0$. Moreover, the functional dependence $(i|K)$ holds if and only if the *conditional entropy* $\triangle h_\xi(i|K) := \triangle h_\xi(ii|K) = h(iK) - h(K)$ vanishes.

Some supporting hyperplanes in the form of $\triangle(\mathsf{ij}|\mathsf{K})$ are known for Γ^*_{N} and the contact points on the boundary are related to conditional independence statements. However, Γ^*_{N} is not a convex cone and it is not a semialgebraic set (see [GR18, Theorem 2.2.3]). Its closure in the euclidean topology is a convex cone [ZY97] whose semialgebraicity is still open. It is known not to be polyhedral [Mat07b] and from this proof, a non-linear information inequality has been extracted [CG08].

Polymatroids. In light of these difficulties, the polyhedral approximation to Γ^*_{N} given by the Shannon information inequalities is studied. This approximation was observed by Fujishige [Fuj78] to be equivalent to the polymatroids introduced by Edmonds [Edm70]:

Definition 2.15. A *polymatroid* is a pair (N, h) where N is the usual finite ground set and $h : \mathcal{P}(\mathsf{N}) \to \mathbb{R}$ which satisfies the three properties

Normalization: $h(\emptyset) = 0$,
Monotonicity: $\triangle h(\mathsf{i}|\mathsf{K}) \geq 0$, for all $\mathsf{i} \in \mathsf{N}$ and $\mathsf{K} \subseteq \mathsf{N} \setminus \mathsf{i}$,
Submodularity: $\triangle h(\mathsf{ij}|\mathsf{K}) \geq 0$, for all $(\mathsf{ij}|\mathsf{K}) \in \mathcal{A}_{\mathsf{N}}$.

The collection H_{N} of all polymatroids in $\mathbb{R}^{\mathcal{P}(\mathsf{N})}$ is the *polymatroid cone*.

The normalization and monotonicity properties imply that H_{N} is contained in the non-negative orthant of $\mathbb{R}^{\mathcal{P}(\mathsf{N})}$. Hence, it is a pointed, rational, polyhedral cone. It is shown in [Whi08, Section 10.1] that the special cases $\triangle(\mathsf{ij}|\mathsf{K}) \geq 0$ are sufficient to imply the general submodularity law for set functions. The special cases $\triangle(\mathsf{i}|\mathsf{K})(h) \geq 0$ imply the general monotonicity law. Moreover, the collection of inequalities

$$\begin{aligned} \triangle(\mathsf{i}|\mathsf{N} \setminus \mathsf{i}) &\geq 0, \text{ for } \mathsf{i} \in \mathsf{N}, \\ \triangle(\mathsf{ij}|\mathsf{K}) &\geq 0, \text{ for } (\mathsf{ij}|\mathsf{K}) \in \mathcal{A}_{\mathsf{N}}, \end{aligned}$$

is known to be facet-defining [Stu21, Section III.A]. The proof writes each functional $\triangle(\mathsf{I}, \mathsf{J}|\mathsf{K})$ as a linear combination of these basic functionals, similar to the localization rule (L).

Two polymatroids are *isomorphic* if they coincide up to bijection between their ground sets. It is clear that restrictions of polymatroids to subsets are polymatroids on the smaller set.

Semimatroids. The inference problem for discrete CI structures can be formulated in terms of the entropy vector as follows: "given that an entropy vector achieves equality for some inequalities $\triangle(\mathsf{ij}|\mathsf{K}) \geq 0$, which other such inequalities become tight?" Since every entropy vector is a polymatroid, the entropic question may be approximated by asking it for polymatroids. This was the starting point for the series papers by Studený and Matúš [MS95, Mat95, Mat99a] characterizing the discrete CI structures on four random variables.

Each face of the polyhedral cone H_{N} can be identified with the set of facets it is contained in. Since the facets correspond to maximal functional dependence statements $(\mathsf{i}|\mathsf{N} \setminus \mathsf{i})$ and CI statements $(\mathsf{ij}|\mathsf{K})$, a face can be encoded by an *augmented CI structure* containing CI statements and functional dependence statements. These sets were called *semimatroids* by Matúš [Mat94]. The set of semimatroids is a lattice because, under the inclusion ordering, it is by definition antiisomorphic to the face lattice of H_{N}. Since this thesis deals with pure conditional independence statements, we would like to avoid having to consider the facets related to functional dependencies.

Definition 2.16. A polymatroid $h \in H_{\mathsf{N}}$ is *tight* if $\triangle h(\mathsf{i}|\mathsf{N} \setminus \mathsf{i}) = 0$ for all $\mathsf{i} \in \mathsf{N}$ and it is *modular* if $\triangle(\mathsf{ij}|\mathsf{K}) = 0$ for all $(\mathsf{ij}|\mathsf{K}) \in \mathcal{A}_{\mathsf{N}}$. The cones of tight and modular polymatroids are denoted by $H^{\mathsf{ti}}_{\mathsf{N}}$ and $H^{\mathsf{mod}}_{\mathsf{N}}$, respectively.

By [MC16], we have the direct decomposition $H_{\mathsf{N}} = H_{\mathsf{N}}^{\mathsf{ti}} \oplus H_{\mathsf{N}}^{\mathsf{mod}}$. Therefore, we focus on the tight part of the polymatroid cone in the following, which removes the undesirable facets of the polymatroid cone. Based on the characterization of conditional independence in terms of the entropy vector, we can finally give the definition of semimatroid as it is used in this thesis:

Definition 2.17. The *semimatroid* of a tight polymatroid $h \in H_{\mathsf{N}}^{\mathsf{ti}}$ is the CI structure

$$[\![h]\!] \coloneqq \{\, (\mathsf{ij}|\mathsf{K}) \in \mathcal{A}_{\mathsf{N}} : \triangle h(\mathsf{ij}|\mathsf{K}) = 0 \,\}.$$

Semimatroids by the above definition are precisely the intersections of the augmented semimatroids of [Mat94] with $\mathcal{A}_{\mathsf{N}}$. They form a lattice under inclusion which is antiisomorphic to the face lattice of the **tight** polymatroid cone. Since this face lattice is not simplicial, its structure reveals non-trivial CI inference properties. These can be discovered by linear programming as explained in Section 2.4.2. Denote the property of being a semigraphoid by $\mathfrak{sg}$ and the of being a semimatroid by $\mathfrak{sm}$.

Lemma 2.18. Semimatroids are semigraphoids: $\mathfrak{sm} \le \mathfrak{sg}$.

Proof. The proof due to [Mat97, Section 5] is a simple computation in the space $(\mathbb{R}^{\mathcal{P}(\mathsf{N})})^*$:

$$\begin{aligned}
&\triangle h(\mathsf{ij}|\mathsf{kL}) + \triangle h(\mathsf{ik}|\mathsf{L}) \\
&= h(\mathsf{ikL}) + h(\mathsf{jkL}) - h(\mathsf{ijkL}) - h(\mathsf{kL}) + h(\mathsf{iL}) + h(\mathsf{kL}) - h(\mathsf{ikL}) - h(\mathsf{L}) \\
&= h(\mathsf{jkL}) - h(\mathsf{ijkL}) + h(\mathsf{iL}) - h(\mathsf{L}) \\
&= h(\mathsf{iL}) + h(\mathsf{jL}) - h(\mathsf{ijL}) - h(\mathsf{L}) + h(\mathsf{ijL}) + h(\mathsf{jkL}) - h(\mathsf{ijkL}) - h(\mathsf{jL}) \\
&= \triangle h(\mathsf{ij}|\mathsf{L}) + \triangle h(\mathsf{ik}|\mathsf{jL}).
\end{aligned}$$

Since all functionals are non-negative on $H_{\mathsf{N}}^{\mathsf{ti}}$, the vanishing of the left-hand side is equivalent to the vanishing of the right-hand side. This proves the semigraphoid property (S). □

This gives many new (geometric) examples of semigraphoids via polymatroids. For example, the rank function of a subspace arrangement in a vector space is a polymatroid.

Matroids. A polymatroid h which is integer-valued and bounded by the cardinality map, i.e., $h(\mathsf{K}) \le |\mathsf{K}|$ is the rank function of a *matroid*. Matroids were conceived by Whitney [Whi35] in the 1930s as combinatorial abstractions of the common properties of independence relations in vector spaces and graphs. Today, matroid theory is a broad and active field of research with an extensive corpus of theorems and constructions and connections to many branches of mathematics. A standard introduction to matroids is Oxley [Oxl11].

The theory of matroids is a continuous source of inspiration in CI theory. This was especially true for the works of Matúš. Not only is the idea of synthetic geometry via matroids the prototype for studying statistics synthetically via semigraphoids, but the representability of a semigraphoid by discrete random variables properly generalizes the representability of a (poly)matroid by linear subspaces; see [Mat97, Lemma 10]. The Lefschetz Principle plays a crucial role in the proof by transferring a linear representation over any field to a representation over a finite field, over which a discrete distribution may be defined.

In this sense, synthetic statistics includes synthetic geometry (over fields) and matroid theory notions such as independent sets and flats translate in a meaningful manner to the conditional independence of discrete random variables. Let ξ be a discrete random vector which represents a matroid M with rank function r, i.e., there exists a constant $c > 0$ such that $h_\xi = c \cdot r$. If I is an independent set in M requires the subvector ξ_{I} to be completely stochastically independent; in other words $[\![\xi_{\mathsf{I}}]\!] = \mathcal{A}_{\mathsf{I}}$. Every dependence in the matroid results in a functional dependence of the random variables. The probabilistic representability of matroids is a challenging topic [Mat99b] which is also studied in coding theory and cryptography in the guise of ideal secret sharing schemes [Sey92, BBPT14] and almost-affine codes [SA98]. It is unknown whether this problem is decidable. Undecidability results for related problems have recently emerged in the work of Li [Li21b, Li21a].

Gaussian CI is **not** compatible with the structure of a matroid like discrete CI is. Consider a simple matroid of rank at least two on N. This includes all interesting matroids. Then every two-element set ij is independent and $(\mathsf{ij}|) \in [\![r]\!]$. Any gaussoid which contains all $(\mathsf{ij}|)$ must already be $\mathcal{A}_\mathsf{N}$ by the Composition axiom (G.iii). Nevertheless, many ideas from synthetic geometry carry over into our theory, particularly in Chapters 3 and 5.

3

Algebraic realization spaces and inference

This chapter introduces *algebraic Gaussians*, a relaxation of regular Gaussians which replaces positive definiteness of the covariance matrix by non-vanishing of its principal minors. This relaxation can be studied over every field. The naming "algebraic Gaussian" should not suggest any probabilistic content. Instead, the emphasis is on an algebraic treatment of principal and almost-principal minors of symmetric matrices as polynomials — a treatment of covariance matrices with a restriction to the language of algebraic geometry. The idea and the limitations of this approach are similar to how one first studies linear algebra before intersecting linear spaces with the non-negative orthant to study polyhedra. Basic closure properties of algebraic Gaussians are derived, but the emphasis in this chapter is on the algebra of realization spaces and the geometric formulation of the Gaussian conditional independence inference problem.

3.1 Algebraic and positive Gaussians

Definition 3.1. Let $\mathbb{K}$ be an ordered field. A matrix $\Sigma \in \mathrm{Sym}_{\mathsf{N}}(\mathbb{K})$ is *positive-definite* over $\mathbb{K}$ if $\Sigma[\mathsf{L}] > 0$ for all $\mathsf{L} \subseteq \mathsf{N}$. The set of positive-definite matrices is denoted by $\mathrm{PD}_{\mathsf{N}}(\mathbb{K})$. The *CI structure* of Σ is defined as

$$[\![\Sigma]\!] \coloneqq \{ (\mathsf{ij}|\mathsf{K}) \in \mathcal{A}_{\mathsf{N}} : \Sigma[\mathsf{ij}|\mathsf{K}] = 0 \} .$$

A CI structure of this form is a *positive Gaussian*. The property of being positively realizable over $\mathbb{K}$ is denoted by $\mathfrak{g}_{\mathbb{K}}^{+}$.

This most fundamental definition warrants a few remarks. First, we do not consider hermitian, positive-definite matrices over $\mathbb{C}$. While this is a valid and interesting direction — and nothing is known yet about the resulting CI theory —, its definition requires the complex conjugation and hence creates a special case, whereas the definition of symmetric matrices is field-agnostic. Second, there exist multiple characterizations of positive definiteness over the real numbers. Every course in linear algebra shows that the following are equivalent over $\mathbb{R}$:

- $\Sigma = A^{\mathsf{T}} A$ with an invertible matrix A,
- the bilinear form Σ is positive-definite,
- all eigenvalues of Σ exist and are positive,
- the principal minors of Σ are positive,
- the leading principal minors of Σ are positive.

When interpreted over a general ordered field, these properties are not all equivalent. For example, it is easy to find symmetric 2×2-matrices with rational entries and positive principal minors, whose characteristic polynomial does not factor over $\mathbb{Q}$, so the eigenvalue criterion fails. Other characterizations, like the two involving positivity of principal minors, are always equivalent. The proof consists of passing to the real closure, which does not change the principal minors as polynomials, and then noticing that the counterexamples to the claimed equivalence are a semialgebraic set ("all symmetric matrices with positive leading principal minors but at least one other principal minor non-positive"). Thus, the truth of the claim is decidable in the first-order theory of real-closed fields and by Tarski's transfer principle the textbook proofs over $\mathbb{R}$ suffice to establish the fact that there are no counterexamples over any ordered field.

The choice of the principal minor characterization to generalize positive definiteness is justified by the **application**: conditional independence concerns the vanishing of almost-principal minors of the matrix. It is to be expected that a formulation of positive definiteness — here serving as a regularity condition — which is most similar to almost-principal minors will be most useful. This expectation is substantiated in Section 3.2. Besides the application, the requirement of positivity of principal minors can be formulated **within** the field. The positivity-of-eigenvalues criterion tacitly requires that the characteristic polynomial factors which is not generally true unless the field is real-closed. This definition easily generalizes further to general fields (in fact, even to commutative rings):

Definition 3.2. Let $\mathbb{K}$ be a field. A matrix $\Gamma \in \mathrm{Sym}_{\mathsf{N}}(\mathbb{K})$ is *principally regular* if $\Gamma[\mathsf{L}] \neq 0$ for all $\mathsf{L} \subseteq \mathsf{N}$. The set of principally regular matrices is denoted by $\mathrm{PR}_{\mathsf{N}}(\mathbb{K})$. The CI structure $[\![\Gamma]\!]$ is defined just as in Definition 3.1. Structures of this form are *algebraic Gaussians* or *algebraically realizable* over $\mathbb{K}$. This property is denoted by $\mathfrak{g}^*_{\mathbb{K}}$.

Principally regular matrices share the closure properties with positive-definite matrices which make the development of a combinatorial theory of marginalization, conditioning, direct sums and more possible; cf. Section 3.3. In this setting, the non-vanishing of all principal minors is stronger than leading principal minors only. Examples of this abound, for instance $\left(\begin{smallmatrix}1&1\\1&0\end{smallmatrix}\right)$. Committed to the Isomorphy convention, we prefer the condition which is $\mathfrak{S}_{\mathsf{N}}$-invariant. In fact, for dealing with realizing matrices we establish the even stronger

Sign Convention. Changing the ordering of rows or columns in a matrix incurs a sign change of the determinant, which is significant when its positivity is concerned. According to the Isomorphy convention we view CI structures up to isomorphy, which corresponds to indexing the rows and columns of symmetric matrices with an unordered set. This is not problematic because simultaneous reordering of rows and columns by the same permutation leaves the sign of principal minors unchanged.

Thus the only convention to be established in addition is on the **pairing** of ground set elements, i.e., when writing down a submatrix to take its determinant, which row and column labels $\mathsf{r}, \mathsf{c} \in \mathsf{N}$ appear together in the k^{th} position from the top-left corner of the matrix? For principal minors $\Sigma[\mathsf{K}]$ and almost-principal minors $\Sigma[\mathsf{ij}|\mathsf{K}]$ there is a canonical choice: in the principal K-part, pair each $\mathsf{k} \in \mathsf{K}$ with itself, and in the almost-principal minor additionally pair row i with column j. For example

$$\Sigma = \begin{array}{cc} \begin{array}{ccc} \mathsf{i} & \mathsf{j} & \mathsf{k} \end{array} & \\ \begin{pmatrix} \sigma_{\mathsf{ii}} & \sigma_{\mathsf{ij}} & \sigma_{\mathsf{ik}} \\ \sigma_{\mathsf{ij}} & \sigma_{\mathsf{jj}} & \sigma_{\mathsf{jk}} \\ \sigma_{\mathsf{ik}} & \sigma_{\mathsf{jk}} & \sigma_{\mathsf{kk}} \end{pmatrix} & \begin{array}{c} \mathsf{i} \\ \mathsf{j} \\ \mathsf{k} \end{array} \end{array} \qquad \begin{aligned} \Sigma[\mathsf{ij}|\mathsf{k}] &= \det\left(\begin{smallmatrix}\sigma_{\mathsf{ij}} & \sigma_{\mathsf{ik}} \\ \sigma_{\mathsf{jk}} & \sigma_{\mathsf{kk}}\end{smallmatrix}\right) = \sigma_{\mathsf{ij}}\sigma_{\mathsf{kk}} - \sigma_{\mathsf{ik}}\sigma_{\mathsf{jk}}, \\ \Sigma[\mathsf{ik}|\mathsf{j}] &= \det\left(\begin{smallmatrix}\sigma_{\mathsf{ik}} & \sigma_{\mathsf{ij}} \\ \sigma_{\mathsf{jk}} & \sigma_{\mathsf{jj}}\end{smallmatrix}\right) = \sigma_{\mathsf{ik}}\sigma_{\mathsf{jj}} - \sigma_{\mathsf{ij}}\sigma_{\mathsf{jk}}, \\ \Sigma[\mathsf{jk}|\mathsf{i}] &= \det\left(\begin{smallmatrix}\sigma_{\mathsf{jk}} & \sigma_{\mathsf{ij}} \\ \sigma_{\mathsf{ik}} & \sigma_{\mathsf{ii}}\end{smallmatrix}\right) = \sigma_{\mathsf{jk}}\sigma_{\mathsf{ii}} - \sigma_{\mathsf{ij}}\sigma_{\mathsf{ik}}. \end{aligned}$$

The first ordering of the rows and columns is natural from the way Σ is written: the rows and columns indicated by $(\mathsf{ij}|\mathsf{k})$ are taken in the same order they appear on the left. But this is not true of the other two almost-principal minors. The pairing convention has the effect of making the **polynomials** corresponding to CI statements invariant under permutation of i, j and k. It is also consistent with marginalization: the sign of $\Sigma[\mathsf{ij}|\mathsf{K}]$ does not depend on whether it is evaluated in $\Sigma \in \mathrm{Sym}_{\mathsf{N}}$ or in the submatrix $\Sigma_{\mathsf{ijK}} \in \mathrm{Sym}_{\mathsf{ijK}}$.

If $\mathbb{K}$ is an ordered field, there are two notions of realizability over $\mathbb{K}$: positive realizability in the ordering of $\mathbb{K}$ or the algebraic realizability over $\mathbb{K}$ which forgets its ordering. Since positive elements are always non-zero, positive realizability is stronger than algebraic realizability, i.e., $\mathfrak{g}^+_{\mathbb{K}} \le \mathfrak{g}^*_{\mathbb{K}}$. The inclusion can be strict:

Example 3.3. Consider the CI structure $\check{\mathcal{S}}_4 := \{ (12|3), (13|4), (14|2) \}$ over $\mathsf{N} = 1234$. This CI structure is not positively realizable over any ordered field. To prove this, it suffices to show that it is non-realizable over $\mathbb{R}$. Consider a generic matrix which satisfies the CI equations of $\check{\mathcal{S}}_4$:

$$\Sigma = \begin{array}{cc} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} & \\ \begin{pmatrix} p & a & b & c \\ a & q & d & e \\ b & d & r & f \\ c & e & f & s \end{pmatrix} & \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} \end{array} \qquad ra = bd, \quad sb = cf, \quad qc = ae.$$

These equations imply $qsra = aefd$. Using that $a \neq 0$ on every realization of $\check{\mathcal{S}}_4$, we obtain $qrs = def$. But positive definiteness implies that $\Sigma[23] = qr - d^2$, $\Sigma[24] = qs - e^2$ and $\Sigma[34] = rs - f^2$ are all positive. Thus

$$q^2 r^2 s^2 = qr \cdot qs \cdot rs > d^2 \cdot e^2 \cdot f^2,$$

which is absurd given the equation derived earlier. However, an algebraic realization of $\check{\mathcal{S}}_4$ is

$$\begin{pmatrix} 1 & 2 & 20 & 1/5 \\ 2 & 1 & 1/10 & 1/10 \\ 20 & 1/10 & 1 & 100 \\ 1/5 & 1/10 & 100 & 1 \end{pmatrix},$$

which exists even over $\mathbb{Q}$. It is easy to check that all principal minors are non-zero and that the only almost-principal minors which vanish are those in $\check{\mathcal{S}}_4$. △

Remark 3.4. This is a smallest example which separates algebraic and positive realizability, in two ways. First, every $\mathbb{C}$-algebraically realizable CI structure over $\mathsf{N} = 123$ is $\mathbb{Q}$-positively realizable (see Remark 3.9), so a ground set of size four is needed. Second, as we prove in Section 4.6, every CI structure with at most two elements which is a gaussoid (this is necessary for algebraic realizability, by Proposition 3.8) is positively realizable over $\mathbb{Q}$, so at least three CI statements are needed.

3.2 Matúš's identity and the gaussoid axioms

The relaxation to principally regular matrices preserves some essential structural properties of their conditional independence relations: algebraic Gaussians are gaussoids. This was proved by Matúš in [Mat05, Corollary 1] even before gaussoids were defined. The lemma leading to this result is stated for arbitrary matrices over $\mathbb{C}$. In fact, the proof is written for (not necessarily symmetric) principally regular matrices and extends to all matrices by continuity. Because of its fundamental importance to the development of our theory, we give a full proof of a generalization of this lemma to arbitrary fields here:

Lemma 3.5: Matúš's identity. For $\Gamma \in \mathrm{Sym}_{\mathsf{N}}(\mathbb{K})$, the following identity holds:

$$\Gamma[\mathsf{kL}] \cdot \Gamma[\mathsf{ij}|\mathsf{L}] = \Gamma[\mathsf{L}] \cdot \Gamma[\mathsf{ij}|\mathsf{kL}] + \Gamma[\mathsf{ik}|\mathsf{L}] \cdot \Gamma[\mathsf{jk}|\mathsf{L}], \quad \text{for all } \mathsf{ijkL} \subseteq \mathsf{N}.$$

Proof. First suppose that $\Gamma \in \mathrm{PR}_{\mathsf{N}}(\mathbb{K})$. Let $\widetilde{\Gamma}$ be the $\mathsf{ik} \times \mathsf{jk}$ matrix with entries $\widetilde{\gamma}_{\mathsf{ab}} = \Gamma_{\mathsf{ab|L}} / \mathsf{L} = \gamma_{\mathsf{ab}} - \Gamma_{\mathsf{a,L}}\Gamma_{\mathsf{L}}^{-1}\Gamma_{\mathsf{L,b}}$. Using the Schur complement expansion of the determinant [Zha05, Theorem 1.1] and the formula for a 2×2 determinant, we see

$$\begin{aligned} \frac{\Gamma[\mathsf{ij|kL}]}{\Gamma[\mathsf{L}]} &= \det\left(\Gamma_{\mathsf{ij|k}} - \Gamma_{\mathsf{ik,L}}\Gamma_{\mathsf{L}}^{-1}\Gamma_{\mathsf{L,jk}}\right) \\ &= \det\widetilde{\Gamma} = \widetilde{\gamma}_{\mathsf{ij}}\widetilde{\gamma}_{\mathsf{kk}} - \widetilde{\gamma}_{\mathsf{ik}}\widetilde{\gamma}_{\mathsf{jk}} \\ &= \frac{\Gamma[\mathsf{kL}]\cdot\Gamma[\mathsf{ij|L}]}{\Gamma[\mathsf{L}]^2} - \frac{\Gamma[\mathsf{ik|L}]\cdot\Gamma[\mathsf{jk|L}]}{\Gamma[\mathsf{L}]^2}, \end{aligned}$$

which is equivalent to the claimed identity. This proves the formula for principally regular matrices. The subset $\mathcal{V}$ of the affine space $\mathrm{Sym}_{\mathsf{N}}(\mathbb{K})$ for which the identity holds is Zariski-closed. If $\mathbb{K}$ has characteristic zero, $\mathrm{Sym}_{\mathsf{N}}(\mathbb{K})$ is irreducible and then $\mathrm{PR}_{\mathsf{N}}(\mathbb{K})$ is a dense subset. But since $\mathrm{PR}_{\mathsf{N}}(\mathbb{K}) \subseteq \mathcal{V}$, it follows that $\mathcal{V} = \mathrm{Sym}_{\mathsf{N}}(\mathbb{K})$ and thus the identity holds on all symmetric matrices over characteristic zero.

Finally, consider that the Matúš identity as an integer polynomial in the entries of a generic symmetric matrix vanishes on the entire affine space $\mathrm{Sym}_{\mathsf{N}}(\mathbb{Q})$. This implies that it is the zero polynomial in $\mathbb{Z}[\Gamma] \subseteq \mathbb{Q}[\Gamma]$, which in turn shows that this identity holds in every commutative ring with unity, in particular every field independently of characteristic. □

Remark 3.6. In Matúš's original work [Mat05, Lemma 1] there is an additional sign which depends on the relative ordering of i, j and k. This sign is fixed to $+1$ by our Sign Convention.

The identity still holds when $\mathsf{i} = \mathsf{j}$ with nearly the same proof. This yields a special case of Dodgson condensation formula; see [Abe08] for historical remarks.

Lemma 3.7: Dodgson condensation. For $\Gamma \in \mathrm{Sym}_{\mathsf{N}}(\mathbb{K})$, the following identity holds:

$$\Gamma[\mathsf{ij|L}]^2 = \Gamma[\mathsf{iL}]\cdot\Gamma[\mathsf{jL}] - \Gamma[\mathsf{L}]\cdot\Gamma[\mathsf{ijL}], \quad \text{for all } \mathsf{ijL} \subseteq \mathsf{N}. \qquad \square$$

Proposition 3.8. If Γ is principally regular, then $[\![\Gamma]\!]$ is a gaussoid.

Proof. The gaussoid axioms are straightforward consequences of the Matúš identity, using that all principal minors are non-zero and that there are no zero divisors in the field $\mathbb{K}$.

(G.i) first requires the Matúš identity with ordered variables ikj:

$$\Gamma[\mathsf{jL}]\cdot\Gamma[\mathsf{ik|L}] = \Gamma[\mathsf{L}]\cdot\underbrace{\Gamma[\mathsf{ik|jL}]}_{=0} + \underbrace{\Gamma[\mathsf{ij|L}]}_{=0}\cdot\Gamma[\mathsf{jk|L}].$$

Hence $\Gamma[\mathsf{ik|L}] = 0$ and $(\mathsf{ik|L}) \in [\![\Gamma]\!]$. Use the variable order ijk in the Matúš identity and the fact $\Gamma[\mathsf{ik|L}] = 0$, which was just proved, to see $(\mathsf{ij|kL}) \in [\![\Gamma]\!]$.

(G.ii) is invariant under swapping j and k and since the Matúš identity is symmetric as well, it suffices to derive only one of the conclusions. The instances of the identity with variables ordered ijk and ikj show:

$$\begin{aligned} \Gamma[\mathsf{ij|L}] &= \Gamma[\mathsf{kL}]^{-1}\cdot\Gamma[\mathsf{ik|L}]\cdot\Gamma[\mathsf{jk|L}], \\ \Gamma[\mathsf{ik|L}] &= \Gamma[\mathsf{jL}]^{-1}\cdot\Gamma[\mathsf{ij|L}]\cdot\Gamma[\mathsf{jk|L}], \end{aligned}$$

and hence

$$\Gamma[\mathsf{ij|L}] = \frac{\Gamma[\mathsf{jk|L}]^2}{\Gamma[\mathsf{jL}]\cdot\Gamma[\mathsf{kL}]}\Gamma[\mathsf{ij|L}]$$

which implies $\Gamma[\mathsf{ij|L}] = 0$ by Lemma 3.7.

(G.iii) is the converse of (G.ii) and invariant as well. It follows from the ordering ijk:

$$\Gamma[\mathsf{kL}] \cdot \underbrace{\Gamma[\mathsf{ij|L}]}_{=0} = \Gamma[\mathsf{L}] \cdot \Gamma[\mathsf{ij|kL}] + \underbrace{\Gamma[\mathsf{ik|L}]}_{=0} \cdot \Gamma[\mathsf{jk|L}] \quad \Rightarrow \quad \Gamma[\mathsf{ij|kL}] = 0.$$

(G.iv) follows immediately from the ordering ijk:

$$\Gamma[\mathsf{kL}] \cdot \underbrace{\Gamma[\mathsf{ij|L}]}_{=0} = \Gamma[\mathsf{L}] \cdot \underbrace{\Gamma[\mathsf{ij|kL}]}_{=0} + \Gamma[\mathsf{ik|L}] \cdot \Gamma[\mathsf{jk|L}] \quad \Rightarrow \quad \Gamma[\mathsf{ik|L}] = 0 \text{ or } \Gamma[\mathsf{jk|L}] = 0. \qquad \square$$

The gaussoid axioms are **some** inference rules that follow from principal regularity. These axioms are "three-variate" in that they require three distinct ground set elements i, j and k, the conditioning set L being arbitrary and present universally in every CI statement in these axioms. The natural question at this point is whether there are more such rules to be derived from Matúš's identity or from principal regularity in general. The first question can be answered right here: the gaussoid axioms imply all CI axioms on three-variate algebraic (and positive) Gaussians. The more general question of which axioms hold for algebraic Gaussians is difficult. Section 4.5 shows that there is no finite set of axioms like (G.i)–(G.iv) from which all true CI inferences can be deduced as the ground set size grows. Section 5.5 further shows that the answer depends on the characteristic of the field. Here, we concentrate on characteristic zero with its prime field $\mathbb{Q}$.

Remark 3.9. Out of the $64 = 2^6$ subsets of $\mathcal{A}_{123}$, exactly eleven are gaussoids. Modulo $\mathfrak{S}_{123}$ there are five structures pictured in Figure 1.4 and they are all positively realizable over $\mathbb{Q}$:

$$\mathtt{E} = \left[\!\left[\begin{pmatrix} 8 & 1 & 1 \\ 1 & 8 & 1 \\ 1 & 1 & 8 \end{pmatrix} \right]\!\right], \quad \mathtt{L} = \left[\!\left[\begin{pmatrix} 8 & 0 & 1 \\ 0 & 8 & 1 \\ 1 & 1 & 8 \end{pmatrix} \right]\!\right], \quad \mathtt{U} = \left[\!\left[\begin{pmatrix} 8 & 1 & 2 \\ 1 & 8 & 4 \\ 2 & 4 & 8 \end{pmatrix} \right]\!\right],$$
$$\mathtt{B} = \left[\!\left[\begin{pmatrix} 8 & 0 & 0 \\ 0 & 8 & 1 \\ 0 & 1 & 8 \end{pmatrix} \right]\!\right], \quad \mathtt{F} = \left[\!\left[\begin{pmatrix} 8 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 8 \end{pmatrix} \right]\!\right].$$

Since all 3-gaussoids are realizable, there cannot be any three-variate CI axioms beyond what is implied by the gaussoid axioms. Any such axiom would have to contradict the realizability of one of the five isomorphy classes of gaussoids.

3.3 Minors, direct sums and symmetry

This section presents the fundamental matrix algebra results on which the algebraic development of regular Gaussian CI structures is based. The aim is not only to make this thesis more self-contained but also to collect in one place the linear algebra facts which drive the structure theory of Gaussian CI, for comparison and possible future generalization. As presented in Section 1.3, marginalization and conditioning of Gaussians correspond to taking principal submatrices and Schur complements. Moreover, duality corresponds to matrix inversion. Hence, the numerous lemmas about the interplay of these operations directly translate to combinatorial properties of Gaussians. It should be noted that principal regularity is precisely the condition on a matrix which ensures the existence of all Schur complements, i.e., "conditional distributions". The elementary combinatorial theory developed in this section takes its matrix-algebraic substance from [Zha05]. Indeed, a similar goal is already pursued in its Chapter 4, entitled "Closure Properties", but it is focused on special classes of matrices from the applications and closedness under Schur complement and inversion.

Minors and duality. Consider a matrix Γ whose rows and columns are indexed by $\mathsf{N} = \mathsf{KL}$:

$$\Sigma = \begin{array}{cc} \mathsf{K} & \mathsf{L} \end{array} \atop \begin{pmatrix} A & B \\ B^{\mathsf{T}} & D \end{pmatrix} \begin{array}{c} \mathsf{K} \\ \mathsf{L} \end{array}.$$

For each choice of such a block decomposition of Γ where A is invertible, there is a factorization of its determinant by the Schur complement of A in Γ:

$$\begin{aligned} \det \begin{pmatrix} A & B \\ B^{\mathsf{T}} & D \end{pmatrix} &= \det \left(\begin{pmatrix} \mathbb{1}_{\mathsf{K}} & 0 \\ -B^{\mathsf{T}} A^{-1} & \mathbb{1}_{\mathsf{L}} \end{pmatrix} \begin{pmatrix} A & B \\ B^{\mathsf{T}} & D \end{pmatrix} \begin{pmatrix} \mathbb{1}_{\mathsf{K}} & -A^{-1}B \\ 0 & \mathbb{1}_{\mathsf{L}} \end{pmatrix} \right) \\ &= \det \begin{pmatrix} A & 0 \\ 0 & D - B^{\mathsf{T}} A^{-1} B \end{pmatrix} = \det A \cdot \det(D - B^{\mathsf{T}} A^{-1} B) \\ &= \det(\Sigma_{\mathsf{K}}) \cdot \det(\Sigma \,/\, \mathsf{K}). \end{aligned}$$

It is obvious that Σ is regular if and only if both Σ_{K} and $\Sigma \,/\, \mathsf{K}$ are regular. Moreover, one can easily derive from the above computation that $\operatorname{rk} \Sigma = \operatorname{rk}(\Sigma_{\mathsf{K}}) + \operatorname{rk}(\Sigma \,/\, \mathsf{K})$. Given that Σ and its block $A = \Sigma_{\mathsf{K}}$ are invertible, the Schur complement of A must be invertible as well. Then the following formula can be confirmed by calculation:

$$\Sigma^{-1} = \begin{pmatrix} A & B \\ B^{\mathsf{T}} & D \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} + A^{-1}B(\Sigma \,/\, \mathsf{K})^{-1} B^{\mathsf{T}} A^{-1} & -A^{-1}B(\Sigma \,/\, \mathsf{K})^{-1} \\ -(\Sigma \,/\, \mathsf{K})^{-1} B^{\mathsf{T}} A^{-1} & (\Sigma \,/\, \mathsf{K})^{-1} \end{pmatrix}.$$

In particular we find $(\Sigma^{-1})_{\mathsf{K}^c} = (\Sigma/\mathsf{K})^{-1}$ and the fundamental combinatorial relations follow:

Proposition 3.10. Let $\Gamma \in \mathrm{PR}_{\mathsf{N}}(\mathbb{K})$. Every principal submatrix, every Schur complement and the inverse of Γ are principally regular. If $\mathbb{K}$ is ordered and Γ positive-definite, then this property is inherited by principal submatrices, Schur complements and inverse. Furthermore:

— $[\![\Gamma^{-1}]\!] = [\![\Gamma]\!]^{\rceil}$,
— $[\![\Gamma^{-1} \setminus \mathsf{K}]\!] = [\![\Gamma \,/\, \mathsf{K}]\!]^{\rceil}$,
— $[\![\Gamma \setminus \mathsf{K}]\!] = [\![\Gamma_{\mathsf{K}^c}]\!] = [\![\Gamma]\!] \setminus \mathsf{K}$,
— $[\![\Gamma \,/\, \mathsf{K}]\!] = [\![\Gamma]\!] \,/\, \mathsf{K}$. □

Corollary 3.11. $\mathfrak{g}^*$ and $\mathfrak{g}^+$ are closed under minors and duality, for all (ordered) fields. □

Lemma 3.12. For $\Sigma \in \mathrm{PR}_{\mathsf{N}}(\mathbb{K})$ over an ordered field $\mathbb{K}$ the following are equivalent:

— Σ is positive-definite,
— $\Sigma[\mathsf{K}] = \det(\Sigma_{\mathsf{K}}) > 0$ for all $\mathsf{K} \subseteq \mathsf{N}$,
— $\det(\Sigma \,/\, \mathsf{K}) > 0$ for all $\mathsf{K} \subseteq \mathsf{N}$,
— $\Sigma[\mathsf{K}] > 0$ for all K in a complete flag $\emptyset \subseteq 1 \subseteq 12 \subseteq \cdots \subseteq \mathsf{N}$,
— there exists $\mathsf{K} \subseteq \mathsf{N}$ with $\Sigma_{\mathsf{K}} \in \mathrm{PD}_{\mathsf{K}}(\mathbb{K})$ and $\Sigma \,/\, \mathsf{K} \in \mathrm{PD}_{\mathsf{K}^c}(\mathbb{K})$.

Proof. The only contentious implication is from the last to the first characterization. The others follow either by definition of by the factorization of the determinant into principal minor and Schur complement, and using that Σ is available as both, a principal submatrix and a Schur complement, via $\Sigma = \Sigma_{\mathsf{N}} = \Sigma \,/\, \emptyset$.

The assertion that positive definiteness of Σ_{K} and $\Sigma \,/\, \mathsf{K}$ for some $\mathsf{K} \subseteq \mathsf{N}$ imply that of Σ can be written as a sentence in the first-order language of ordered rings ("positivity of some polynomials implies positivity of others"). The counterexamples form a semialgebraic set. Since $\mathbb{K}$ with its order can be extended to a real-closed field which would include any counterexample that exists in $\mathbb{K}$, we may suppose that $\mathbb{K}$ is real-closed. But the truth of the assertion is decidable in the first-order theory, which means that, by Tarski's transfer principle, it suffices to treat $\mathbb{K} = \mathbb{R}$. In this well-known case, the non-existence of counterexamples follows from Sylvester's Law of Inertia; see [Zha05, Theorem 1.12]. □

The last condition does **not** generalize to a characterization of principal regularity:

Example 3.13. Consider the matrix

$$\Gamma = \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \left(\begin{array}{cc|ccc} 1 & -1 & 1/2 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ \hline 1/2 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right) \end{array} \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \in \mathrm{Sym}_{1\dots5}(\mathbb{C}).$$

Obviously $\Gamma[12] = 0$, so this matrix is not principally regular. But there exists the subset 345 with principally regular submatrix $\Gamma_{345} = \mathbb{1}_{345}$ and Schur complement $\Gamma/345 = \left(\begin{smallmatrix} 3/4 & -1 \\ -1 & 1 \end{smallmatrix}\right)$. △

Direct sums. The direct sum construction from Section 1.2.3 is valid as well for principally regular matrices:

Lemma 3.14. If $\Sigma \in \mathrm{PR}_\mathsf{N}(\mathbb{K})$ and $\Gamma \in \mathrm{PR}_\mathsf{M}(\mathbb{K})$ for disjoint N and M, then $\Sigma \oplus \Gamma \in \mathrm{PR}_{\mathsf{NM}}(\mathbb{K})$. The same holds when PR is replaced by PD everywhere. Furthermore, $[\![\Sigma \oplus \Gamma]\!] = [\![\Sigma]\!] \oplus [\![\Gamma]\!]$.

Proof. For any $\mathsf{K} \subseteq \mathsf{NM}$ we have $(\Sigma \oplus \Gamma)[\mathsf{K}] = \Sigma[\mathsf{K} \cap \mathsf{N}] \cdot \Gamma[\mathsf{K} \cap \mathsf{M}]$ by the block-diagonal structure of the direct sum. This shows that positive definiteness is preserved as well. Denote $\Phi = \Sigma \oplus \Gamma$ as well as $\mathsf{N}' = \mathsf{K} \cap \mathsf{N}$ and $\mathsf{M}' = \mathsf{K} \cap \mathsf{M}$. For a statement $(\mathsf{ij}|\mathsf{K}) \in \mathcal{A}_{\mathsf{NM}}$:

$$\begin{aligned} \Phi[\mathsf{ij}|\mathsf{K}] &= \det \begin{array}{c} \begin{array}{ccc} \mathsf{i} & \mathsf{N}' & \mathsf{M}' \end{array} \\ \begin{pmatrix} \phi_{\mathsf{ij}} & u_{\mathsf{N}'}^\mathsf{T} & u_{\mathsf{M}'}^\mathsf{T} \\ v_{\mathsf{N}'} & \Sigma_{\mathsf{N}'} & 0 \\ v_{\mathsf{M}'} & 0 & \Gamma_{\mathsf{M}'} \end{pmatrix} \end{array} \begin{array}{c} \\ \mathsf{j} \\ \mathsf{N}' \\ \mathsf{M}' \end{array} \\ &= \Sigma[\mathsf{N}'] \cdot \Gamma[\mathsf{M}'] \cdot \left(\phi_{\mathsf{ij}} - \begin{pmatrix} u_{\mathsf{N}'}^\mathsf{T} & u_{\mathsf{M}}^\mathsf{T} \end{pmatrix} \begin{pmatrix} \Sigma_{\mathsf{N}'} & 0 \\ 0 & \Gamma_{\mathsf{M}'} \end{pmatrix}^{-1} \begin{pmatrix} v_{\mathsf{N}'} \\ v_{\mathsf{M}'} \end{pmatrix} \right) \\ &= \Sigma[\mathsf{N}'] \cdot \Gamma[\mathsf{M}'] \cdot \left(\phi_{\mathsf{ij}} - u_{\mathsf{N}'}^\mathsf{T} \Sigma_{\mathsf{N}'}^{-1} v_{\mathsf{N}'} - u_{\mathsf{M}'}^\mathsf{T} \Gamma_{\mathsf{M}'}^{-1} v_{\mathsf{M}'} \right). \end{aligned}$$

If $\mathsf{ij} \subseteq \mathsf{N}$, then $u_{\mathsf{M}'}$ and $v_{\mathsf{M}'}$ are zero and we obtain a factorization

$$\begin{aligned} \Phi[\mathsf{ij}|\mathsf{K}] &= \Sigma[\mathsf{N}'] \cdot \Gamma[\mathsf{M}'] \cdot \left(\phi_{\mathsf{ij}} - u_{\mathsf{N}'}^\mathsf{T} \Sigma_{\mathsf{N}'}^{-1} v_{\mathsf{N}'} \right) \\ &= \Gamma[\mathsf{M}'] \cdot \Sigma[\mathsf{ij}|\mathsf{N}']. \end{aligned}$$

Thus $(\mathsf{ij}|\mathsf{K}) \in [\![\Phi]\!]$ if and only if $(\mathsf{ij}|\mathsf{N}') \in [\![\Sigma]\!]$. The case $\mathsf{ij} \subseteq \mathsf{M}$ is analogous. If $\mathsf{i} \in \mathsf{N}$ and $\mathsf{j} \in \mathsf{M}$ (or vice versa), then $\phi_{\mathsf{ij}} = 0$, $u_{\mathsf{N}'} = 0$ and $v_{\mathsf{M}'} = 0$, hence $\Phi[\mathsf{ij}|\mathsf{K}] = 0$. This proves $[\![\Phi]\!] = [\![\Sigma]\!] \oplus [\![\Gamma]\!]$. □

Corollary 3.15. $\mathfrak{g}^*$ and $\mathfrak{g}^+$ are closed under direct sums, for all (ordered) fields. □

Symmetries. The hyperoctahedral group $\mathfrak{B}_\mathsf{N}$ is generated by the reflection symmetries of the hypercube in $\mathbb{R}^\mathsf{N}$ and it is shown in Section 1.2.2 and Section 1.3 that semigraphoid and gaussoid axioms are invariant under this group action. This action can be extended from CI structures to principally regular matrices over every field, where it is a quotient of the group $(\mathbb{Z}/4)^\mathsf{N} \rtimes \mathfrak{S}_\mathsf{N}$. This group is, in turn, a discrete subgroup of the $\mathrm{SL}_2(\mathbb{K})^\mathsf{N} \rtimes \mathfrak{S}_\mathsf{N}$ action on the Lagrangian Grassmannian; cf. [HS07, BDKS19]. It turns out that positive definiteness is not preserved under the matrix action of $\mathfrak{B}_\mathsf{N}$ but principal regularity is. This is another important reason to study algebraic Gaussians. The $\mathfrak{B}_\mathsf{N}$ symmetry is a primary tool in the **positive** realizability results in Section 4.6.

As an abstract group, $\mathfrak{B}_\mathsf{N}$ is the semidirect product $(\mathbb{Z}/2)^\mathsf{N} \rtimes \mathfrak{S}_\mathsf{N}$ of the group of *swaps* $(\mathbb{Z}/2)^\mathsf{N}$ and the group of *permutations* $\mathfrak{S}_\mathsf{N}$. The group of swaps is generated by reflections over coordinate hyperplanes; the $\mathfrak{S}_\mathsf{N}$ factor acts by permuting the coordinate axes. Concretely, we obtain $\mathfrak{B}_\mathsf{N}$ as a quotient of the group $(\mathbb{Z}/4)^\mathsf{N} \rtimes \mathfrak{S}_\mathsf{N}$ acting on principally regular matrices. In the semidirect product, every group element can be written uniquely as the composition of an element of $\mathfrak{S}_\mathsf{N}$ and one of $(\mathbb{Z}/4)^\mathsf{N}$. The permutation part is just an orthogonal coordinate change, permuting rows and columns of the matrix, implementing isomorphy of the CI structure and merely permuting the set of principal minors. This action changes neither principal regularity nor positive definiteness and therefore we focus on the $(\mathbb{Z}/4)^\mathsf{N}$ part in the remainder of this section. Each $\mathbb{Z}/4$ factor in the N-fold product is represented by the four 2×2 matrices

$$\mathbb{Z}/4 = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \right\} \subseteq \mathrm{SL}_2(\mathbb{K}).$$

To each tuple $X = (X_\mathsf{i})_{\mathsf{i} \in \mathsf{N}} \in (\mathbb{Z}/4)^\mathsf{N}$ associate four $\mathsf{N} \times \mathsf{N}$ diagonal matrices A, B, C, D such that

$$X_\mathsf{i} = \begin{pmatrix} A_\mathsf{ii} & B_\mathsf{ii} \\ C_\mathsf{ii} & D_\mathsf{ii} \end{pmatrix}.$$

The image of a symmetric matrix Γ under $(X_\mathsf{i})_{\mathsf{i} \in \mathsf{N}}$ is $\Gamma' = (A + \Gamma C)^{-1}(B + \Gamma D)$. Γ' is again symmetric by [HS07, Lemma 13] and the following Proposition 3.16 describes its principal and almost-principal minors. To facilitate this description we use a parametrization of this group action. For any subset $\mathsf{Z} \subseteq \mathsf{N}$ and a tuple of signs $\delta \in \{\pm 1\}^\mathsf{N}$, choose the group element X where

$$X_\mathsf{i} = \delta_\mathsf{i} \begin{cases} \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right), & \mathsf{i} \not\in \mathsf{Z}, \\ \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right), & \mathsf{i} \in \mathsf{Z}. \end{cases}$$

Then the action can be written as $\mathcal{S}^\delta_\mathsf{Z}(\Gamma) \coloneqq (A + \Gamma C)^{-1}(B + \Gamma D)$ with

$$A_\mathsf{ii} = D_\mathsf{ii} = \begin{cases} \delta_\mathsf{i}, & \mathsf{i} \not\in \mathsf{Z}, \\ 0, & \mathsf{i} \in \mathsf{Z}, \end{cases} \qquad C_\mathsf{ii} = -B_\mathsf{ii} = \begin{cases} 0, & \mathsf{i} \not\in \mathsf{Z}, \\ \delta_\mathsf{i}, & \mathsf{i} \in \mathsf{Z}. \end{cases}$$

In expressing minors of $\mathcal{S}^\delta_\mathsf{Z}(\Gamma)$ in terms of Z, δ and Γ, it becomes necessary to recombine the involved subsets of N. Using the abbreviations $\mathsf{AB} = \mathsf{A} \cup \mathsf{B}$ and $\langle \mathsf{AB} \rangle = \mathsf{A} \cap \mathsf{B}$ as well as $\mathsf{A}^\mathsf{c} = \mathsf{N} \setminus \mathsf{A}$, any combination of interest can be efficiently written down in "disjunctive normal form". For example, $\langle \mathsf{Z}\mathsf{K}^\mathsf{c} \rangle \langle \mathsf{Z}^\mathsf{c}\mathsf{K} \rangle = (\mathsf{Z} \cap \mathsf{K}^\mathsf{c}) \cup (\mathsf{K} \cap \mathsf{Z}^\mathsf{c}) = (\mathsf{Z} \setminus \mathsf{K}) \cup (\mathsf{K} \setminus \mathsf{Z}) = \mathsf{Z} \oplus \mathsf{K}$.

Proposition 3.16. Let Γ be principally regular over $\mathbb{K}$ and $\mathsf{Z} \subseteq \mathsf{N}$ and $\delta \in \{\pm 1\}^\mathsf{N}$ be arbitrary. Then $\Gamma' = \mathcal{S}^\delta_\mathsf{Z}(\Gamma)$ is principally regular over $\mathbb{K}$. The gaussoid $[\![\Gamma']\!] = [\![\Gamma]\!]^\mathsf{Z} = \{ (\mathsf{ij}|\mathsf{Z}^\mathsf{ij} \oplus \mathsf{K}) : (\mathsf{ij}|\mathsf{K}) \in [\![\Gamma]\!] \}$. More precisely, we have the following formulas for the principal and almost-principal minors of Γ':

$$\Gamma'[\mathsf{K}] = (-1)^{\langle \mathsf{ZK} \rangle} \Gamma[\mathsf{Z}]^{-1} \cdot \Gamma[\mathsf{Z} \oplus \mathsf{K}],$$
$$\Gamma'[\mathsf{ij}|\mathsf{K}] = (-1)^{\langle \mathsf{ZK} \rangle} \Gamma[\mathsf{Z}]^{-1} \cdot \delta_\mathsf{i}\delta_\mathsf{j} \cdot \Gamma[\mathsf{ij}|\mathsf{Z}^\mathsf{ij} \oplus \mathsf{K}].$$

Proof. Γ' satisfies the matrix equation $(A + \Gamma C)\Gamma' = B + \Gamma D$ and hence its minors can be computed with a generalized Cramer's rule [GAE02]:

$$\det \Gamma'_{\mathsf{I},\mathsf{J}} = \det(A + \Gamma C)^{-1} \det \left[(A + \Gamma C)(\mathsf{I}, \mathsf{J} : B + \Gamma D) \right],$$

for sets $\mathsf{I}, \mathsf{J} \subseteq \mathsf{N}$ of the same size and where $X(\mathsf{I}, \mathsf{J} : Y)$ denotes the matrix X where the columns indexed by I are replaced by the columns of Y indexed by J. In this notation, we omit J when it equals I. In addition we use δX to denote the matrix X where the i^{th} column is scaled with δ_i.

By definition of A and C we have $A + \Gamma C = \delta\Gamma(\mathsf{Z}^{\mathsf{c}} : \mathbb{1}_{\mathsf{N}})$ and by Laplace expansion on the unit columns in Z^{c} we easily derive

$$\det(A + \Gamma C) = \prod_{\mathsf{k}} \delta_{\mathsf{k}} \cdot \Gamma[\mathsf{Z}].$$

To compute the principal minor for $\mathsf{K} \subseteq \mathsf{N}$, notice that

$$(A + \Gamma C)(\mathsf{K} : B + \Gamma D) = \delta[\Gamma(\mathsf{Z}^{\mathsf{c}} : \mathbb{1}_{\mathsf{N}})](\mathsf{K} : \Gamma(\mathsf{Z} : -\mathbb{1}_{\mathsf{N}})) =: \delta\Gamma''.$$

The columns of Γ'' are composed as follows:

$\langle \mathsf{ZK}^{\mathsf{c}} \rangle$ respective columns of Γ,
$\langle \mathsf{ZK} \rangle$ respective negative unit vectors,
$\langle \mathsf{Z}^{\mathsf{c}}\mathsf{K}^{\mathsf{c}} \rangle$ respective unit vectors,
$\langle \mathsf{Z}^{\mathsf{c}}\mathsf{K} \rangle$ respective columns of Γ.

By Laplace expansion of the unit vector columns, we obtain

$$\det \delta\Gamma'' = \prod_{\mathsf{k}} \delta_{\mathsf{k}} \cdot (-1)^{\langle \mathsf{ZK} \rangle} \cdot \Gamma[\langle \mathsf{ZK}^{\mathsf{c}} \rangle \langle \mathsf{Z}^{\mathsf{c}}\mathsf{K} \rangle],$$

and $\langle \mathsf{ZK}^{\mathsf{c}} \rangle \langle \mathsf{Z}^{\mathsf{c}}\mathsf{K} \rangle = \mathsf{Z} \oplus \mathsf{K}$ proves the principal minor formula. Notice that the Laplace expansions deleted the same rows and columns, so the Sign Convention is preserved.

For the almost-principal minor $(\mathsf{ij}|\mathsf{K})$, the same procedure yields the columns of Γ''

$\langle \mathsf{Z}(\mathsf{iK})^{\mathsf{c}} \rangle$ respective columns of Γ,
$\langle \mathsf{ZK} \rangle$ respective negative unit vectors,
$\langle \mathsf{Z}^{\mathsf{c}}(\mathsf{iK})^{\mathsf{c}} \rangle$ respective unit vectors,
$\langle \mathsf{Z}^{\mathsf{c}}\mathsf{K} \rangle$ respective columns of Γ,
i the j^{th} column of $\Gamma(\mathsf{Z} : -\mathbb{1}_{\mathsf{N}})$.

Pulling out the δ signs from the determinant, we get the sign $\prod_{\mathsf{k} \neq \mathsf{i},\mathsf{j}} \delta_{\mathsf{k}} \cdot \delta_{\mathsf{j}}^2 = \delta_{\mathsf{i}}\delta_{\mathsf{j}} \prod_{\mathsf{k}} \delta_{\mathsf{k}}$ because the i^{th} column's δ_{i} was replaced by a j^{th} column's δ_{j}. In addition, performing again Laplace expansion on the unit vector columns, the determinant so far is

$$\det \delta\Gamma'' = \delta_{\mathsf{i}}\delta_{\mathsf{j}} \prod_{\mathsf{k}} \delta_{\mathsf{k}} \cdot (-1)^{\langle \mathsf{ZK} \rangle} \cdot \Gamma''[\mathsf{i}\langle \mathsf{Z}(\mathsf{iK})^{\mathsf{c}} \rangle \langle \mathsf{Z}^{\mathsf{c}}\mathsf{K} \rangle]. \tag{$\diamond$}$$

The contents of column i of Γ'' depend on whether $\mathsf{j} \in \mathsf{Z}$ or not. If $\mathsf{j} \notin \mathsf{Z}$, then the i^{th} column of Γ'' is just the j^{th} column of Γ. Thus the remaining minor of Γ'' in ($\diamond$) is revealed to be the almost-principal minor of Γ with row indices $\mathsf{i}\langle \mathsf{Z}(\mathsf{iK})^{\mathsf{c}} \rangle \langle \mathsf{Z}^{\mathsf{c}}\mathsf{K} \rangle = \mathsf{i}\langle \mathsf{Z}(\mathsf{ijK})^{\mathsf{c}} \rangle \langle \mathsf{Z}^{\mathsf{c}}\mathsf{K} \rangle$ and column indices $\mathsf{j}\langle \mathsf{Z}(\mathsf{ijK})^{\mathsf{c}} \rangle \langle \mathsf{Z}^{\mathsf{c}}\mathsf{K} \rangle$. It is easy to see that the replacement of column i by column j leaves the rows and columns correctly paired according to the Sign Convention and it remains to compute the conditioning set as $\langle \mathsf{Z}(\mathsf{ijK})^{\mathsf{c}} \rangle \langle \mathsf{Z}^{\mathsf{c}}\mathsf{K} \rangle = \mathsf{Z}^{\mathsf{ij}} \oplus \mathsf{K}$.

If $\mathsf{j} \in \mathsf{Z}$, then the i^{th} column contains the negative j^{th} unit vector. Laplace expansion with respect to this column results in the column labeled i and the row labeled j to be removed and incurs a sign change which depends on the distance between these columns. By simultaneously reordering rows and columns, we can assume that rows and columns i and j are next to each other. In this case, the sign change is -1, which is compensated by the *entry* -1 in the eliminated column. The reordering ensures that rows and columns are properly paired after Laplace expansion. The remaining minor of Γ'' has row indices $\mathsf{i}\langle \mathsf{Z}(\mathsf{iK})^{\mathsf{c}} \rangle \langle \mathsf{Z}^{\mathsf{c}}\mathsf{K} \rangle \setminus \mathsf{j} = \mathsf{i}\langle \mathsf{Z}(\mathsf{ijK})^{\mathsf{c}} \rangle \langle \mathsf{Z}^{\mathsf{c}}\mathsf{K} \rangle$ and column indices $\langle \mathsf{Z}(\mathsf{iK})^{\mathsf{c}} \rangle \langle \mathsf{Z}^{\mathsf{c}}\mathsf{K} \rangle = \mathsf{j}\langle \mathsf{Z}(\mathsf{ijK})^{\mathsf{c}} \rangle \langle \mathsf{Z}^{\mathsf{c}}\mathsf{K} \rangle$ and is thus again the almost-principal minor $(\mathsf{ij}|\mathsf{Z}^{\mathsf{ij}} \oplus \mathsf{K})$ of Γ. $\square$

Remark 3.17. These formulas describe in particular all the entries of $\mathcal{S}_{\mathsf{Z}}^{\delta}(\Gamma)$ in terms of Γ, Z and δ. Remarkably, the choice of δ has no influence at all on the principal minors, and only changes the sign of almost-principal ones. Hence, by identifying in each $\mathbb{Z}/4$ factor the two matrices with opposite signs, we obtain a quotient group isomorphic to $(\mathbb{Z}/2)^{\mathsf{N}} \rtimes \mathfrak{S}_{\mathsf{N}}$ which faithfully implements the hyperoctahedral group on the CI structure over any field. The realizing matrix may not be well-defined but the quotient is conclusive about its positivity, over ordered fields, and thus can be used to certify positive realizability of hyperoctahedral images of gaussoids.

Remark 3.18. Since all symmetric matrices satisfy the Matúš identities (Lemma 3.5) and the principally regular matrices are closed under the hyperoctahedral group, it follows that all principally regular matrices (and hence, analogously to the proof of the Matúš identity, all symmetric matrices) satisfy **hyperoctahedral images** of the Matúš identity. These were called *edge trinomials* and *square trinomials* in [BDKS19, Section 2] and are further discussed in Chapter 6. For example, consider the identity $[3][12|] = [\emptyset][12|3] + [13|][23|]$. Its image under the swap $\mathsf{Z} = 2$ is $[23][12|] = [2][12|3] + [13|2][23|]$ and this holds for all symmetric matrices as well. It does not fit the pattern of the Matúš identity because the two almost-principal minors in the last product term have different degrees, but since swapping preserves PR_{N}, it follows from it.

Remark 3.19. Let $\Gamma \in \mathrm{PR}_{\mathsf{N}}(\mathbb{R})$ and consider its *signature*, i.e., the vector of length 2^n containing the signs of its principal minors. None of the signature coordinates is zero due to principal regularity. By Proposition 3.16 a set $\mathsf{Z} \subseteq \mathsf{N}$ furnishes the following map on signature vectors:

$$(\operatorname{sgn}\Gamma[\mathsf{K}])_{\mathsf{K}} \;\mapsto\; \left(\operatorname{sgn}(-1)^{\langle \mathsf{Z}\mathsf{K}\rangle} \cdot \operatorname{sgn}\Gamma[\mathsf{Z}] \cdot \operatorname{sgn}\Gamma[\mathsf{Z} \oplus \mathsf{K}]\right)_{\mathsf{K}}.$$

Fix the all-positive signature vector, corresponding to a positive-definite matrix, and consider its image under this map for varying sets $\mathsf{Z} \subseteq \mathsf{N}$. Identifying the signs $\{+1, -1\}$ with $\mathbb{F}_2$, this is just

$$\mathsf{Z} \mapsto (|\mathsf{Z} \cap \mathsf{K}| \bmod 2)_{\mathsf{K} \subseteq \mathsf{N}}.$$

Clearly this is injective and the image vectors are not only closed under swaps but also permutations, so there exist exactly 2^n out of 2^{2^n} sign vectors which permit a positive-definite matrix in the $\mathfrak{B}_{\mathsf{N}}$-orbit of a matrix carrying this signature. These 2^n signatures coincide with those obtained from the hyperoctahedral images of the identity matrix.

Corollary 3.20. $\mathfrak{g}^*$ is closed under $\mathfrak{B}_{\mathsf{N}}$ and $\mathfrak{g}^+$ is closed under the subgroup $\mathfrak{T}_{\mathsf{N}}$ generated by $\mathfrak{S}_{\mathsf{N}}$ and duality. □

Remark 3.21. Fix $\mathsf{N} = \mathsf{ijkl}$. There are 679 gaussoids in 58 classes modulo isomorphy, precisely 53 of which are positively realizable (even over $\mathbb{Q}$). This is the main result of [LM07]. The five non-realizable gaussoids still have positively realizable gaussoids in their **hyperoctahedral** orbits, so Corollary 3.20 implies that all 4-gaussoids are algebraically realizable over $\mathbb{Q}$. This can be seen by considering the Lněnička–Matúš axioms (LM.i)–(LM.v) which, together with the gaussoid axioms, characterize positive realizability on a 4-element ground set. The antecedent sets of these five axioms correspond to the five positively non-realizable gaussoids (up to isomorphy). For instance, swap ij in the antecedent set of (LM.i): $(\mathsf{ij}|) \wedge (\mathsf{kl}|) \wedge (\mathsf{ik}|\mathsf{jl}) \wedge (\mathsf{jl}|\mathsf{ik}) \Rightarrow (\mathsf{ik}|)$. This produces the gaussoid $\{(\mathsf{ij}|), (\mathsf{kl}|\mathsf{ij}), (\mathsf{ik}|\mathsf{l}), (\mathsf{jl}|\mathsf{k})\}$ which satisfies all Lněnička–Matúš axioms vacuously and therefore must be positively realizable. Its inverse image under the swapping is then at least algebraically realizable.

Example 3.22. Consider $\mathcal{L} = \{(12|5), (34|2)\}$ over $\mathsf{N} = 12345$. By swapping 25, we obtain $\{(12|), (34|5)\}$, which is a *dependent sum* of the statements $(12|)$ and $(34|5)$ over disjoint

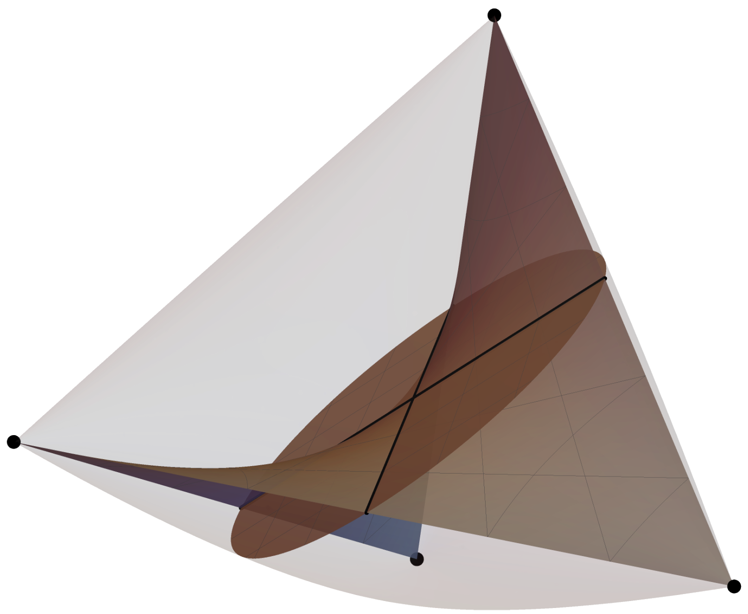

Figure 3.1: CI models pictured inside the *elliptope* (gray), which is the set of 3×3 correlation matrices parametrized by their off-diagonal entries in $\mathbb{R}^3$. The disc is the model of $(12|)$. The saddle-like surface is the model of $(12|3)$. It is a square twisted in 3-space, whose four vertices (black dots) coincide with those of the elliptope. The two black lines mark the intersection of both models, which is the weak realization space of $\mathcal{W} = \{\,(12|), (12|3)\,\}$. This model has two components which intersect in the identity matrix. They are the models of the two realizable $\mathtt{B}$ relations lying over $\mathcal{W}$.

ground sets. Dependent sums are introduced and studied further in Section 4.3. The disjointness (together with Remark 3.9) helps in deriving the algebraic realization of the image below on the left. Then, acting on this matrix with $\mathcal{S}_{25}$:

$$\begin{array}{cc} \begin{matrix} 1 & 2 & 3 & 4 & 5 \end{matrix} & \\ \left(\begin{array}{cc|ccc} 1 & 0 & 1/2 & 1/2 & 1/2 \\ 0 & 1 & 1/2 & 1/2 & 1/2 \\ \hline 1/2 & 1/2 & 8 & 1 & 2 \\ 1/2 & 1/2 & 1 & 8 & 4 \\ 1/2 & 1/2 & 2 & 4 & 8 \end{array}\right) & \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} \end{array} \xrightarrow{\mathcal{S}_{25}} \begin{array}{cc} \begin{matrix} 1 & 2 & 3 & 4 & 5 \end{matrix} & \\ \begin{pmatrix} 30/31 & 1/31 & 12/31 & 8/31 & -2/31 \\ 1/31 & -32/31 & -12/31 & -8/31 & 2/31 \\ 12/31 & -12/31 & 228/31 & -3/31 & -7/31 \\ 8/31 & -8/31 & -3/31 & 184/31 & -15/31 \\ -2/31 & 2/31 & -7/31 & -15/31 & -4/31 \end{pmatrix} & \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} \end{array}$$

indeed gives a principally regular realization of $\mathcal{L}$. The left matrix is positive-definite and therefore the right one is not, by Remark 3.19. △

3.4 Realization spaces

The set of all matrices which realize a given CI structure form its realization space. This is a constructible set and encodes inference information about algebraic Gaussians. The study of this connection is initiated in this section and occupies the rest of this chapter. There are two notions of realization space for a CI structure in the literature. We begin with a third one which simultaneously generalizes both of them and will be more useful in the discussions in Section 3.5 and throughout the remaining chapters.

Definition 3.23. A *CI constraint system* over ground set N is a two-sorted set $\mathcal{S} \subseteq \mathcal{A}_\mathsf{N} \cup \neg\mathcal{A}_\mathsf{N}$ which contains CI statements $(\mathsf{ij}|\mathsf{K})$ and negated CI statements $\neg(\mathsf{ij}|\mathsf{K})$. By $\neg\mathcal{S}$ we denote the element-wise negation of $\mathcal{S}$, where $\neg\neg(\mathsf{ij}|\mathsf{K}) = (\mathsf{ij}|\mathsf{K})$.

Definition 3.24. The *algebraic CI model* of a CI constraint system $\mathcal{S}$ over a field $\mathbb{K}$ is the set $\mathcal{R}^*_\mathbb{K}(\mathcal{S})$ of all $\Gamma \in \mathrm{PR}_\mathsf{N}(\mathbb{K})$ which satisfy

$$\begin{aligned} \Gamma[\mathsf{ij}|\mathsf{K}] &= 0 \ \text{ for all } \ \ (\mathsf{ij}|\mathsf{K}) \in \mathcal{S}, \\ \Gamma[\mathsf{ij}|\mathsf{K}] &\neq 0 \ \text{ for all } \ \neg(\mathsf{ij}|\mathsf{K}) \in \mathcal{S}. \end{aligned}$$

The *positive CI model* over an ordered field $\mathbb{K}$ is $\mathcal{R}^+_\mathbb{K}(\mathcal{S}) \coloneqq \mathcal{R}^*_\mathbb{K}(\mathcal{S}) \cap \mathrm{PD}_\mathsf{N}(\mathbb{K})$.

When a distinction is immaterial, the unqualified symbol $\mathcal{R}(\mathcal{S})$ is used to denote either kind of model. There are two ways to view a CI structure $\mathcal{L}$ as a constraint system: either one regards it as $\mathcal{L} \cup \neg(\mathcal{A}_\mathsf{N} \setminus \mathcal{L})$ so that everything which is not specified to vanish really must not vanish; or as $\mathcal{L}$ itself so that everything which is not specified to vanish is left unspecified. We distinguish these two kinds of realization space for subsets of $\mathcal{A}_\mathsf{N}$:

Definition 3.25. Let $\mathcal{L} \subseteq \mathcal{A}_\mathsf{N}$ and $\mathbb{K}$ a field. The *algebraic realization space* of $\mathcal{L}$ over $\mathbb{K}$ is the set of all principally regular matrices over $\mathbb{K}$ which realize $\mathcal{L}$:

$$\mathcal{R}^*_\mathbb{K}(\mathcal{L}) \coloneqq \{\, \Gamma \in \mathrm{PR}_\mathsf{N}(\mathbb{K}) : [\![\Gamma]\!] = \mathcal{L} \,\}.$$

If $\mathbb{K}$ is ordered, the *positive realization space* is $\mathcal{R}^+_\mathbb{K}(\mathcal{L}) \coloneqq \mathcal{R}^*_\mathbb{K}(\mathcal{L}) \cap \mathrm{PD}_\mathsf{N}$.

Definition 3.26. The *weak realization spaces* (algebraic or positive, respectively) of $\mathcal{L}$ over an (ordered) field are $\mathcal{V}^*_\mathbb{K}(\mathcal{L}) \coloneqq \{\, \Gamma \in \mathrm{PR}_\mathsf{N}(\mathbb{K}) : [\![\Gamma]\!] \supseteq \mathcal{L} \,\}$ and $\mathcal{V}^+_\mathbb{K}(\mathcal{L}) \coloneqq \mathcal{V}^*_\mathbb{K}(\mathcal{L}) \cap \mathrm{PD}_\mathsf{N}$.

Remark 3.27. For questions about realizability, it is natural to look at the realization spaces associated to a given CI structure. On the other hand, conditional independence models in algebraic statistics are usually specified by compulsory CI statements **without** forbidding others. This leads to the weak realization space.

CI structures on the same ground set (modulo the Isomorphy convention) are naturally ordered by inclusion. If $\mathcal{L} \subseteq \mathcal{M}$, then every CI statement which holds for $\mathcal{L}$ also holds for $\mathcal{M}$. Since CI statements are Zariski-closed conditions, this means that $\mathcal{M}$ is more constrained than $\mathcal{L}$, the random variables modeled by $\mathcal{M}$ are in a "more special position" than required by $\mathcal{L}$. This situation is analogous to the notion of *weak maps* in matroid theory, which induce a similar partial order [Whi08, Chapter 9].

Lemma 3.28. The realization space map $\mathcal{L} \mapsto \mathcal{R}(\mathcal{L})$ is inclusion-reversing. □

If $\mathcal{M} \supseteq \mathcal{L}$, then $\mathcal{M}$ *lies above* $\mathcal{L}$. Clearly $\mathcal{V}(\mathcal{L}) = \bigcup_{\mathcal{M} \supseteq \mathcal{L}} \mathcal{R}(\mathcal{M})$. Inside the quasiaffine variety $\mathrm{PR}_\mathsf{N}(\mathbb{K})$ with the induced Zariski topology from $\mathrm{Sym}_\mathsf{N}(\mathbb{K})$, the weak realization space $\mathcal{V}(\mathcal{L})$ is a closed set and $\mathcal{R}(\mathcal{L})$ is open inside of it. It is, however, not true that $\mathcal{V}(\mathcal{L})$ is the Zariski closure of $\mathcal{R}(\mathcal{L})$, even when the realization space is non-empty. An example on $\mathsf{N} = 1234$ is due to Drton and Xiao:

Example 3.29: [DX10, Example 4.1]. Let $\mathcal{L} = \{\, (12|), (12|34), (34|), (34|12) \,\}$. We study the minimal primes of the ideal defined by $\mathcal{L}$ in `Macaulay2`:

```
R = QQ[p,a,b,c, q,d,e, r,f, s];
-- p,q,r,s on the diagonal and a,b,c,d,e,f off-diagonal
X = genericSymmetricMatrix(R,p,4);
decompose ideal(
  det X_{0}^{1}, det X_{0,2,3}^{1,2,3},    -- (12|), (12|34)
  det X_{2}^{3}, det X_{2,0,1}^{3,0,1}     -- (34|), (34|12)
);
--> ⟨a, f, bcq + pde, cer + bds, −pe²r + b²qs, −c²qr + pd²s⟩ ∩ ⟨a, f, b, e⟩ ∩ ⟨a, f, c, d⟩
```

Clearly, the last two components are contained in $\mathcal{V}^*_{\mathbb{C}}(\mathcal{L})$, but $\mathcal{R}^*_{\mathbb{C}}(\mathcal{L})$ and its Zariski closure are entirely contained in the first component. △

CI models are invariant under an action of the algebraic torus $(\mathbb{K}^\times)^{\mathsf{N}}$ which scales the rows and columns of an $\mathsf{N} \times \mathsf{N}$ matrix simultaneously. The following lemma is easily inferred from the multilinearity of the determinant:

Lemma 3.30. Let $\Gamma \in \mathrm{PR}_{\mathsf{N}}(\mathbb{K})$ and $D = \mathrm{diag}(d_1, \dots, d_{\mathsf{n}})$ with $d_{\mathsf{i}} \in \mathbb{K}^\times$. Then

$$(D\Gamma D)[\mathsf{L}] = \prod_{\mathsf{l} \in \mathsf{L}} d_{\mathsf{l}}^2 \cdot \Gamma[\mathsf{L}],$$
$$(D\Gamma D)[\mathsf{ij}|\mathsf{L}] = d_{\mathsf{i}} d_{\mathsf{j}} \prod_{\mathsf{l} \in \mathsf{L}} d_{\mathsf{l}}^2 \cdot \Gamma[\mathsf{ij}|\mathsf{L}].$$

In particular $D\Gamma D \in \mathrm{PR}_{\mathsf{N}}(\mathbb{K})$ (or even in $\mathrm{PD}_{\mathsf{N}}(\mathbb{K})$ if $\Gamma \in \mathrm{PD}_{\mathsf{N}}(\mathbb{K})$) and $[\![D\Gamma D]\!] = \Gamma$. □

Remark 3.31. Let $\mathbb{K}$ be quadratically closed. Then a diagonal matrix D may be chosen for a principally regular Γ so that $D\Gamma D$ has unit diagonal. This eliminates n degrees of freedom from the polynomial system describing the realizability of a gaussoid. The same trick works for positive realizations over euclidean fields. In statistics, that is for regular Gaussian distributions over $\mathbb{R}$, the normalization of the diagonal of a covariance matrix produces the *correlation matrix* of the distribution. For example, since the complex numbers are quadratically closed, every algebraic Gaussian over $\mathbb{C}$ has a realization with unit diagonal. The real numbers are missing $\sqrt{-1}$, so algebraic realizations of gaussoids over $\mathbb{R}$ can only be assumed to have ± 1 entries on the diagonal. The restriction to positive realizations restores the expectation of a 1-diagonal. Finally, the rational numbers are missing many square roots. Over $\mathbb{Q}(\sqrt{-1})$, for example, one has to expect positive squarefree integers on the diagonal of an algebraic realization. Over $\mathbb{Q}$ they may also be negative.

Definition 3.32. The *characteristic set* of a constraint system $\mathcal{S}$ over $\mathcal{A}_{\mathsf{N}}$ is set $\chi(\mathcal{S})$ of all integers k for which there exists a field $\mathbb{K}$ of characteristic k with $\mathcal{R}^*_{\mathbb{K}}(\mathcal{S}) \neq \emptyset$. For a CI structure $\mathcal{L}$, the characteristic set tests for emptiness of the realization space $\mathcal{R}^*_{\mathbb{K}}(\mathcal{L})$.

This is justified because the weak realization space of any $\mathcal{L}$ contains the identity matrix over every field, so a characteristic set referring to $\mathcal{V}(\mathcal{L})$ would be trivial. Also note that ordered fields always have characteristic zero, so a positive version of the concept is not useful.

CI models, in particular realization spaces, are described by polynomial constraints with integer coefficients on symmetric matrices. The following statements are direct consequences of the Lefschetz Principle:

Theorem 3.33: Lefschetz principle for gaussoids. (1) The properties $\mathfrak{g}^*_k$ and $\bigcup_k \mathfrak{g}^*_k$ of being realizable over some field (of a given characteristic k) are decidable. (2) The characteristic set of a gaussoid is decidable. (3) If a gaussoid is realizable over some field, then it is realizable over a finite field. □

[BDKS19, Example 13] has a 5-gaussoid which is not realizable over $\mathbb{C}$. By Remarks 3.9 and 3.21 this is the smallest ground set where such a gaussoid can occur. We present this example here with a minor modification (the helpful observation that one of the equations in the original example is superfluous is due to Xiangying Chen) and show that it is not realizable over any field:

Example 3.34: A non-algebraic gaussoid. The gaussoid

$$\mathcal{G} = \{\, (12|), (13|4), (14|5), (23|5), (35|1), (45|2), (15|23), (34|12), (24|135) \,\}$$

is not algebraically realizable over any field. It is sufficient to check this over algebraically closed fields. Over these fields, we can impose a unit diagonal on a principally regular realization Γ, by Lemma 3.30. Then, $\mathcal{G}$ imposes the following easy equations on Γ:

$$\Gamma = \begin{pmatrix} 1 & 0 & b & c & d \\ 0 & 1 & e & f & g \\ b & e & 1 & h & i \\ c & f & h & 1 & j \\ d & g & i & j & 1 \end{pmatrix}, \qquad e = gi, \qquad \begin{matrix} i = bd, & c = dj, \\ b = ch, & j = fg. \end{matrix}$$

Using these variable substitutions, the longer equations, corresponding to CI statements with bigger conditioning sets, shrink. They are

$$0 = d[1 - d^2 f^2 g^2 h^2 (1 + d^2 g^2 - g^2)], \tag{15|23}$$
$$0 = h[1 - 2d^2 f^2 g^2], \tag{34|12}$$
$$0 = f[(1 - d^2)(1 + d^2 f^2 g^4 h^2 (1 + d^2) - g^2 (1 + d^2 h^2 (1 + f^2)))]. \tag{24|135}$$

Dividing by the non-zero variables d, f, and h yields a system in even powers of the variables. Replacing $d^2 = D$ and so on, we have:

$$1 = DFGH(1 + DG - G), \tag{a}$$
$$1 = 2DFG, \tag{b}$$
$$0 = (1 - D)(1 + DFG^2H + D^2FG^2H - G - DGH - DFGH). \tag{c}$$

In a principally regular Γ we have $1 - D = \Gamma[15] \neq 0$, so the last equation is equivalent to

$$0 = 1 + DFG^2H + D^2FG^2H - G - DGH - DFGH. \tag{c$'$}$$

By adding up (a) and (c$'$) and then using (b):

$$\begin{aligned} 0 &= [1 + DFG^2H + D^2FG^2H - G - DGH - DFGH] + \\ &\quad [DFGH + D^2FG^2H - DFG^2H - 1] \\ &= 2D^2FG^2H - DGH - G \\ &= -G, \end{aligned}$$

which is a contradiction to $(25|) \notin \mathcal{G}$. Notice that this contradiction was derived using only the non-vanishing of principal minors and division by variables known to be non-zero by the definition of $\mathcal{G}$. This shows that $\mathcal{G}$ is not algebraically realizable over any field. △

3.5 Geometry of inference

Consider the kind of polynomial constraints defining $\mathcal{R}(\mathcal{S})$ over the complex numbers: they are composed of equations for $(ij|K) \in \mathcal{S}$ and inequations for $\neg(ij|K) \in \mathcal{S}$ and principal regularity. **In general**, inequations are genericity conditions and they are almost always fulfilled. That is to say, the space on which at least one of k (irreducible, non-constant) polynomials vanishes is the union of k proper hypersurfaces, which has measure zero in the ambient affine space. Moreover, all polynomials in our system are homogeneous, because they are determinants, so by Lemma 3.30 we may regard the variety defined by the equations of the system as an intersection of hypersurfaces in projective space. **Generically**, intersections are complete in projective space over $\mathbb{C}$, so Bézout's theorem [MS21a, Theorem 2.16] tells us to expect a variety of codimension exactly the number of equations. Generic inequations will then only remove lower-dimensional parts of this variety.

But we have already seen that these expectations are not met with Gaussian CI models, for instance in Example 3.34: had the polynomials been generic, the realization space of $\mathcal{G}$ would have had dimension $10 - 9 = 1$, but it turned out empty instead. The polynomials

$$\begin{aligned}
[ij|] &= x_{ij} \\
[ij|k] &= x_{ij}x_{kk} - x_{ik}x_{jk} \\
[ij|kl] &= x_{ij}x_{kk}x_{ll} - x_{il}x_{jl}x_{kk} + x_{il}x_{jk}x_{kl} + x_{ik}x_{jl}x_{kl} - x_{ij}x_{kl}^2 - x_{ik}x_{jk}x_{ll} \\
[ij|klm] &= x_{ij}x_{kk}x_{ll}x_{mm} + x_{im}x_{jm}x_{kl}^2 - x_{im}x_{jl}x_{kl}x_{km} - x_{il}x_{jm}x_{kl}x_{km} + {} \\
&\quad x_{il}x_{jl}x_{km}^2 - x_{im}x_{jm}x_{kk}x_{ll} + x_{im}x_{jk}x_{km}x_{ll} + x_{ik}x_{jm}x_{km}x_{ll} - {} \\
&\quad x_{ij}x_{km}^2x_{ll} + x_{im}x_{jl}x_{kk}x_{lm} + x_{il}x_{jm}x_{kk}x_{lm} - x_{im}x_{jk}x_{kl}x_{lm} - {} \\
&\quad x_{ik}x_{jm}x_{kl}x_{lm} - x_{il}x_{jk}x_{km}x_{lm} - x_{ik}x_{jl}x_{km}x_{lm} + 2x_{ij}x_{kl}x_{km}x_{lm} + {} \\
&\quad x_{ik}x_{jk}x_{lm}^2 - x_{ij}x_{kk}x_{lm}^2 - x_{il}x_{jl}x_{kk}x_{mm} + x_{il}x_{jk}x_{kl}x_{mm} + {} \\
&\quad x_{ik}x_{jl}x_{kl}x_{mm} - x_{ij}x_{kl}^2x_{mm} - x_{ik}x_{jk}x_{ll}x_{mm}
\end{aligned}$$

$$\begin{aligned}
[i] &= x_{ii} \\
[ij] &= x_{ii}x_{jj} - x_{ij}^2 \\
[ijk] &= x_{ii}x_{jj}x_{kk} - x_{ik}^2x_{jj} + 2x_{ij}x_{ik}x_{jk} - x_{ii}x_{jk}^2 - x_{ij}^2x_{kk} \\
[ijkl] &= x_{ii}x_{jj}x_{kk}x_{ll} + x_{il}^2x_{jk}^2 - 2x_{ik}x_{il}x_{jk}x_{jl} + x_{ik}^2x_{jl}^2 - x_{il}^2x_{jj}x_{kk} + {} \\
&\quad 2x_{ij}x_{il}x_{jl}x_{kk} - x_{ii}x_{jl}^2x_{kk} + 2x_{ik}x_{il}x_{jj}x_{kl} - 2x_{ij}x_{il}x_{jk}x_{kl} - {} \\
&\quad 2x_{ij}x_{ik}x_{jl}x_{kl} + 2x_{ii}x_{jk}x_{jl}x_{kl} + x_{ij}^2x_{kl}^2 - x_{ii}x_{jj}x_{kl}^2 - x_{ik}^2x_{jj}x_{ll} + {} \\
&\quad 2x_{ij}x_{ik}x_{jk}x_{ll} - x_{ii}x_{jk}^2x_{ll} - x_{ij}^2x_{kk}x_{ll}
\end{aligned}$$

Figure 3.2: Very specific polynomials. Up to $\mathfrak{S}_N$ symmetry, there is only one almost-principal minor and one principal minor in every degree and the constraints defining Gaussian CI models are assembled from different instantiations of these polynomials.

which Gaussian CI is concerned with are not generic. They are **very specific** because they all are subdeterminants of a single symmetric matrix, some of which are listed in Figure 3.2. One instance of this specificity is Matúš's identity in Lemma 3.5. No such polynomial identity would hold on generic polynomials, by definition.

The point is that deciding if the model of a CI constraint system is empty or not is apparently hard. But this is a consequence of the high degree of structure in this problem, one manifestation of which was pointed out by Drton and Xiao [DX10, Section 2]. They consider the weak order on CI structures together with the maps $\mathcal{A}_N \supseteq \mathcal{L} \mapsto \mathcal{V}^+(\mathcal{L}) \subseteq \mathrm{PD}_N$ and $\mathrm{PD}_N \supseteq V \mapsto \bigcap_{\Sigma \in V} [\![\Sigma]\!]$. This map is a **Galois connection** likewise induced by the binary relation

$$(ij|K) \diamond \Gamma \;:\Leftrightarrow\; \Gamma[ij|K] = 0,$$

which is just a finite, specific version of the Zariski topology on PD_N, where the "attributes" of matrices, i.e., polynomials, are restricted to almost-principal minors. This construction furnishes a closure operator on CI structures.

Definition 3.35. Fix a field and a space of matrices, i.e., $\mathrm{PR}_N(\mathbb{K})$ or $\mathrm{PD}_N(\mathbb{K})$ if $\mathbb{K}$ is ordered. The closure operator defined above is called *completion* ($\mathfrak{g}^*_{\mathbb{K}}$*-completion* or $\mathfrak{g}^+_{\mathbb{K}}$*-completion* to be precise). The closed CI structures are also called *complete*.

The complete relations are in bijection with the weak realization spaces. The act of completion adds to a CI structure $\mathcal{L}$ all CI statements which hold on $\mathcal{V}(\mathcal{L})$. It should be noted at this point that even though the defining Galois connection is a coarsening of the connection which induces the Zariski topology, the complete subsets of $\mathcal{A}_N$ do **not** form the closed subsets of a finite topology on $\mathcal{A}_N$. This is because (given that the field is large enough) all singletons $\{ (ij|K) \}$ are complete (see Lemma 4.60) and thus if complete sets were closed under unions, the topology would be trivial — which contradicts the validity of the gaussoid axioms. However, Drton and Xiao pull back the unique decomposition theorem for closed sets in noetherian topologies from the Zariski topology on the space of matrices to this combinatorial version of it. Their proof remains valid over general fields and for principally regular matrices. To state it, we need the following lattice-theoretic notion:

Definition 3.36. A complete relation $\mathcal{L}$ is *irreducible* if it cannot be written as a non-trivial intersection of complete relations. Dually, the model of $\mathcal{L}$ cannot be written as a non-trivial union of weak realization spaces.

Theorem 3.37: [DX10, Theorems 2.1 and 2.2]. A CI structure $\mathcal{L}$ is complete if and only if it is an intersection of realizable relations. In this case it has a unique (up to order) decomposition $\mathcal{L} = \bigcap_i \mathcal{R}_i$ into realizable and irreducible relations $\mathcal{R}_i$. □

These notions are easily transferred to general properties instead of the realizability properties $\mathfrak{g} \in \{ \mathfrak{g}^*_{\mathbb{K}}, \mathfrak{g}^+_{\mathbb{K}} \}$. These completions of properties $\mathfrak{p}$ are sometimes easier to compute than $\mathfrak{p}$ itself and still useful if $\mathfrak{p}$ is necessary for realizability, as in Chapter 6.

Definition 3.38. Let $\mathfrak{p}$ be a property. The *completion* of $\mathfrak{p}$ is the property $\overline{\mathfrak{p}}$ which contains $\mathcal{L} \subseteq \mathcal{A}_\mathsf{N}$ if and only if it can be written as an intersection of relations from $\mathfrak{p}(\mathsf{N})$. A relation $\mathcal{L} \in \mathfrak{p}(\mathsf{N})$ is *irreducible* if it is not a non-trivial intersection of relations from $\mathfrak{p}(\mathsf{N})$.

Example 3.39. The relation $\mathcal{L} = \{ (12|), (12|3) \}$, whose model is pictured in Figure 3.1, is $\mathfrak{g}^+_{\mathbb{Q}}$-complete (non-realizable) and reducible. This is true in the Zariski sense (where it is the union of line segments) and in the CI sense (where it is the union of weak realization spaces of $\mathsf{B}_1 = \mathcal{L} \cup \{ (13|), (13|2) \}$ and $\mathsf{B}_2 = \mathcal{L} \cup \{ (23|), (23|1) \}$). Their intersection is equal to $\mathcal{L}$. Since B_1 and B_2 are realizable, this is the unique decomposition into irreducibles. △

Example 3.40. Recall Example 3.3 where it was shown that $\check{\mathcal{S}}_4 \coloneqq \{ (12|3), (13|4), (14|2) \}$ over $\mathsf{N} = 1234$ is not positively realizable. The proof derives a contradiction from the positive definiteness assumption and from $\neg(12|)$, i.e., $a \neq 0$. This proves the following inference rule for positive Gaussians: $(12|3) \wedge (13|4) \wedge (14|2) \;\Rightarrow\; (12|)$. But then the semigraphoid axioms further imply

$$
\begin{aligned}
(12|) \wedge (14|2) \;&\Rightarrow\; (12|4) \wedge (14|), \\
(14|) \wedge (13|4) \;&\Rightarrow\; (14|3) \wedge (13|).
\end{aligned}
$$

With $(12|)$, $(13|)$ and $(14|)$ certainly vanishing on the weak realization space of $\check{\mathcal{S}}_4$, there is a block-diagonal structure $\mathcal{V}^+(\check{\mathcal{S}}_4) \subseteq \mathrm{PD}_1 \oplus \mathrm{PD}_{234}$. This shows that the completion $\overline{\check{\mathcal{S}}_4} \supseteq \emptyset_1 \oplus \emptyset_{234}$, where $\emptyset_\mathsf{N}$ is the empty CI structure over N. Since this lower bound on the completion lies above $\check{\mathcal{S}}_4$ and is realizable, this is in fact an equality. △

This point of view suggests to view complete relations as "radical ideals" which correspond to weak realization spaces — the varieties in our setup — and realizable relations as their "minimal primes". The completion operator is a **specific** combinatorial version of ordinary affine geometry (inside PR or PD) tailored to Gaussian CI. This opens up the geometric perspective on the inference problem:

Lemma 3.41. Let $\mathcal{L} \subseteq \mathcal{A}_\mathsf{N}$ and $\overline{\mathcal{L}}$ its $\mathfrak{g}$-completion. Then $\bigwedge \mathcal{L} \Rightarrow \bigwedge \overline{\mathcal{L}}$ is valid for $\mathfrak{g}$. □

Thus computing the completion of a CI structure produces CI inference rules. Of course, not all valid inferences on Gaussians can be found in this way. The prototypical counterexample is weak transitivity Equation (G.iv) corresponding to the model from Example 3.39. The form of the weak transitivity axiom shows the problem: it implies a disjunction, and on the **intersection** of all realizable models above it implies nothing. Example 3.40 furthermore shows that the completion operators for $\mathfrak{g}^*$ and for $\mathfrak{g}^+$ can differ. The relation $\check{\mathcal{S}}_4$ is algebraically realizable over $\mathbb{Q}$, as shown in Example 3.3, and hence complete, but with respect to positive realizability over $\mathbb{Q}$ it is not even complete because the inference rules derived in Example 3.40 are *Horn clauses*. These clauses only have conjunctions on the right-hande side and thus the consequences hold on all realizable relations above $\check{\mathcal{S}}_4$, thus also on the completion and since $\check{\mathcal{S}}_4$ does not contain them, it cannot be complete.

The general Gaussian CI inference problem can be understood not by considering varieties but constructible sets inside of PR_N or PD_N. Namely, the inference rule $\bigwedge \mathcal{L} \Rightarrow \bigvee \mathcal{M}$ holds if and only if the the product $\prod_{(ij|K) \in \mathcal{M}} [ij|K]$ vanishes on the space $\mathcal{V}(\mathcal{L})$; because then every matrix which satisfies all statements in $\mathcal{L}$ satisfies at least one of those in $\mathcal{M}$. Whether a polynomial vanishes on a variety (inside of PR_N or PD_N) can be effectively decided by the results in Chapter 2. This is the essence of Matúš's fundamental paper [Mat05], but was only stated for complex numbers, presumably due to practical considerations about computer algebra. The following is an extension of this idea, linking models and inference:

Theorem 3.42. The set of counterexamples to the validity for $\mathfrak{g}^\bullet_{\mathbb{K}}$ of the inference rule $\varphi : \bigwedge \mathcal{L} \Rightarrow \bigvee \mathcal{M}$ coincides with the CI model $\mathcal{R}^\bullet_{\mathbb{K}}(\varphi) \coloneqq \mathcal{R}^\bullet_{\mathbb{K}}(\mathcal{L} \cup \neg\mathcal{M})$.

Proof. The counterexamples to φ are characterized by satisfying $\mathcal{L}$ but satisfying none of the statements in $\mathcal{M}$. This is precisely what is required in $\mathcal{R}(\varphi)$. The ambient set of matrices, $\mathrm{PR}_{\mathsf{N}}(\mathbb{K})$ or $\mathrm{PD}_{\mathsf{N}}(\mathbb{K})$, is the same in both cases. □

This means that the validity of inference rules over algebraically or real-closed fields is decidable by Hilbert's Nullstellensatz or, respectively, the Positivstellensatz.

Definition 3.43. Denote by $\mathtt{GR}^\bullet_{\mathbb{K}}$ the decision problem associated to the property $\mathfrak{g}^\bullet_{\mathbb{K}}$, i.e., given a CI structure $(\mathsf{N}, \mathcal{L})$ decide whether $\mathcal{L} \in \mathfrak{g}^\bullet_{\mathbb{K}}(\mathsf{N})$. Analogously let $\mathtt{GCI}^\bullet_{\mathbb{K}}$ the inference problem on $\mathfrak{g}^\bullet_{\mathbb{K}}$.

These two problems concern the extensional and the intensional description of the same object, namely the property $\mathfrak{g}^\bullet_{\mathbb{K}}$. Naturally, they are equivalent:

Theorem 3.44. GR and GCI are Turing-equivalent. They are decidable over algebraically or real-closed fields.

Proof. Decidability follows from reduction to the Nullstellensätze via Theorem 3.42. An input $(\mathsf{N}, \mathcal{L})$ is accepted by GR if and only if $\bigwedge \mathcal{L} \Rightarrow \bigvee \mathcal{L}^{\mathsf{c}}$ is an invalid inference rule, which can be checked by a single invocation of a GCI oracle. Conversely, to decide GCI with the help of a GR oracle, it suffices to list all realizable subsets of $\mathcal{A}_{\mathsf{N}}$ using finitely many queries and then to simply check whether an inference formula holds on all of them. □

Remark 3.45. Fix a characteristic k. A CI inference rule is valid for all fields of characteristic k if and only if it valid over the algebraic closure of the prime field. This is because the set of counterexamples to the validity is constructible and the Lefschetz Principle applies.

Example 3.46: Example 3.34, inference rule version. Manual algebraic manipulation of the equations and certain inequations in Example 3.34 showed that $\mathcal{G}$ is non-realizable. When written as an inference rule, the proof reads

$$\bigwedge \mathcal{G} \Rightarrow (15|) \vee (24|) \vee (25|) \vee (34|). \tag{$*$}$$

This is a valid CI inference rule for algebraic Gaussians over every field. Using a computer algebra system like Macaulay2, one can compute the (Zariski closure of the) algebraic realization space of $\mathcal{G}$ over fixed characteristic (here zero) and find even stronger inference rules:

```
R = QQ[p,a,b,c,d, q,e,f,g, r,h,i, s,j, t];
-- A generic symmetric matrix with 1-diagonal (since we work over ℂ).
X = genericSymmetricMatrix(R,p,5);
X = sub(X, { p=>1, q=>1, r=>1, s=>1, t=>1 });
I = ideal(  -- Ideal of CI relations:
  det X_{0}^{1},              -- (12|)
  det X_{0,3}^{2,3},          -- (13|4)
  det X_{0,4}^{3,4},          -- (14|5)
  det X_{1,4}^{2,4},          -- (23|5)
  det X_{2,0}^{4,0},          -- (35|1)
  det X_{3,1}^{4,1},          -- (45|2)
  det X_{0,1,2}^{4,1,2},      -- (15|23)
  det X_{2,0,1}^{3,0,1},      -- (34|12)
  det X_{1,0,2,4}^{3,0,2,4}   -- (24|135)
);
```

```
-- Saturate iteratively at the principal minors, which results in
-- a Zariski-closed superset of the algebraic realization space.
fold((I,f) -> I:f, radical I, subsets(numRows(X)) / (K -> det X_K^K))
--> ⟨a, b, c, d, e, f, h, i, j⟩
```

The result is the ideal of all off-diagonal entries of `X` except for `g`. By construction, this is the vanishing ideal of a Zariski-closed superset of $\mathcal{R}^*_{\mathbb{C}}(\mathcal{G})$. The vanishing relation is then a subset of the $\mathfrak{g}^*_{\mathbb{C}}$-completion of $\mathcal{G}$ and therefore

$$\bigwedge \mathcal{G} \;\Rightarrow\; \bigwedge \left(\mathcal{A}_{12345} \setminus \{\, (25|\mathsf{K}) : \mathsf{K} \subseteq \mathsf{N}^{25} \,\}\right)$$

is valid for algebraic Gaussians over characteristic zero. The union of antecedents and consequents is graphic and hence realizable by Theorem 4.6, which means that we have found the algebraic (and the positive) completion of $\mathcal{G}$ over characteristic zero and it is irreducible. △

3.6 The bracket ring and final polynomials

Ideas very similar to those in this chapter are presented in the work [BS89] by Bokowski and Sturmfels on "Computational synthetic geometry", where matroids are the underlying combinatorial structure. Matroid theory leads to another class of special polynomials: linear independence of n vectors in $\mathbb{K}^d$ can be studied by listing the full-rank $d \times d$ submatrices of a $d \times n$ matrix whose columns contain these vectors. Thus, matroid theory cares about the **maximal** minors of a general $d \times n$ matrix as polynomials (where the bases and non-bases of a matroid define the analogue of our CI constraint system), and those give rise to another combinatorial shadow of the Zariski topology. Its closure operator proves incidence theorems for point configurations, whereas we care about conditional independence theorems for Gaussian random variables. Both have been identified as special cases of the incidence of a polynomial to an ideal in (semi)algebraic geometry.

It is worthwhile to trace this parallel further and introduce *final polynomials*, non-realizability certificates for CI structures whose existence is guaranteed by the Nullstellensätze, but which can additionally be written in terms of principal and almost-principal minors alone. As such they are algebraic proofs of the **validity** of inferences rules, by Theorem 3.42. On the other hand, a proof of the **invalidity** of an inference rule φ is a point in the CI model $\mathcal{R}(\varphi)$ of counterexamples to φ. Since this set is defined by integer polynomials, the Lefschetz Principle or Tarski's transfer principle guarantee that a counterexample with (real) algebraic entries over a prime field can be found — if a counterexample exists at all. The objective of this section is to derive the definition of final polynomial in such a way that a **theorem of the alternative** holds: **either** an inference rule is invalid and an algebraic counterexample to it can be found, written down exactly and verified with standard computer algebra software, **or** the inference rule is valid and a proof for it can be compressed into a single final polynomial with coefficients in a prime field, written down exactly and verified with standard computer algebra software. This strong alternative makes the classification problem attached to the decision problems `GR` and `GCI` transparent: every proof or refutation for `GR` or `GCI` derived by (costly) computation can be turned into a certificate of finite size which everyone can independently check (much quicker).

The point of departure is a ring $\mathcal{R}_\mathsf{N}$ whose variables are formal *brackets*: by slight abuse of notation let $\mathcal{P}_\mathsf{N}$ denote the set of brackets $[\mathsf{K}]$ denoting principal minors and $\mathcal{A}_\mathsf{N}$ the set of brackets $[\mathsf{ij}|\mathsf{K}]$ for almost-principal minors. Throughout this section we study the ring $\mathcal{R} = \mathcal{R}_\mathsf{N} := \mathbb{K}[\mathcal{P}_\mathsf{N} \cup \mathcal{A}_\mathsf{N}]$ over a fixed field $\mathbb{K}$, in which brackets are the variables. Unless the potential for confusion arises, we avoid explicit mention of the fixed the ground set N. Consider the evaluation homomorphism $\psi = \psi_{\mathsf{N}} : \mathcal{R}_{\mathsf{N}} \to \mathbb{K}[\Gamma]$ which sends every bracket to the subdeterminant it represents, i.e., $[\mathsf{K}] \mapsto \Gamma[\mathsf{K}]$ and $[\mathsf{ij}|\mathsf{K}] \mapsto \Gamma[\mathsf{ij}|\mathsf{K}]$, which are polynomials in the entries of a generic symmetric matrix Γ over $\mathbb{K}$. This map is a surjection because every entry of Γ is either a principal minor $[\mathsf{k}]$ or an almost-principal minor $[\mathsf{ij}|]$. Since the image is an integral domain, the kernel $\ker\psi$ is a prime ideal. Note that the empty principal minor $[\emptyset]$, which evaluates to 1 under ψ, acts as a homogenizing variable in $\mathcal{R}$, so we may restrict attention to the homogeneous polynomials in the kernel. Let $\mathcal{J} = \mathcal{J}_{\mathsf{N}}$ denote this homogeneous prime ideal.

$\mathcal{J}$ contains the *universal relations* among principal and almost-principal minors of symmetric matrices over a field. Using sophisticated techniques from representation theory, a finite generating set of the homogeneous quadratic part of $\mathcal{J}$ was derived in [BDKS19, Theorem 5]. Among these generators are the familiar relations (quadratic in the ring $\mathcal{R}$ where each bracket is a variable of degree 1)

$$[\mathsf{kL}] \cdot [\mathsf{ij}|\mathsf{L}] = [\mathsf{L}] \cdot [\mathsf{ij}|\mathsf{kL}] + [\mathsf{ik}|\mathsf{L}] \cdot [\mathsf{jk}|\mathsf{L}],$$
$$[\mathsf{ij}|\mathsf{L}]^2 = [\mathsf{iL}] \cdot [\mathsf{jL}] - [\mathsf{L}] \cdot [\mathsf{ijL}],$$

proved in Lemmas 3.5 and 3.7. In addition we define the multiplicative monoid $\mathcal{U} = \mathcal{U}_{\mathsf{N}}$ which is generated by the principal minor brackets in $\mathcal{R}$. Together these two sets of polynomials yield a ring-theoretic version of the set PR_{N} as the localization $\mathcal{U}_{\mathsf{N}}^{-1}(\mathcal{R}/\mathcal{J})$.

Definition 3.47. The *CI ideal* of a constraint system $\mathcal{S}$ is the ideal $\mathcal{I}_{\mathcal{S}}$ generated by all vanishing conditions indicated by $\mathcal{S}$, i.e., $[\mathsf{ij}|\mathsf{K}]$ for $(\mathsf{ij}|\mathsf{K}) \in \mathcal{S}$. The *non-vanishing monoid* $\mathcal{U}_{\mathcal{S}}$ is the monoid generated by $\mathcal{U}_{\mathsf{N}}$ and the brackets $[\mathsf{ij}|\mathsf{K}]$ for $\neg(\mathsf{ij}|\mathsf{K}) \in \mathcal{S}$.

Abuse of bracket notation. Formally, brackets $[\mathsf{K}]$ and $[\mathsf{ij}|\mathsf{K}]$ are elements of degree 1 in the polynomial ring $\mathcal{R}$. In the following we often identify them with their image under the evaluation map ψ. Furthermore, since the generators of all of $\mathcal{U}$, $\mathcal{I}_{\mathcal{S}}$ and $\mathcal{U}_{\mathcal{S}}$ are polynomials with integer coefficients in the entries of a generic symmetric matrix, these sets of polynomials can naturally be interpreted over every field, hence no field is attached to the notation.

The choice of letters should suggest that $\mathcal{J}$ is not contained in $\mathcal{I}_{\mathcal{S}}$ but $\mathcal{U}$ is contained in $\mathcal{U}_{\mathcal{S}}$ by definition. This is for notational reasons which will become clear shortly.

Algebraic final polynomials. Final polynomials are sought as "obvious" proofs for the emptiness of the model $\mathcal{R}(\mathcal{S})$ of a CI constraint system $\mathcal{S}$. We first concentrate on algebraic realizability over a field $\mathbb{K}$. Recall from Section 2.4.1 how commutative algebraists solve polynomial systems:

Definition 3.48. The *abstract algebraic CI model* of a constraint system $\mathcal{S}$ over $\mathbb{K}$ is the spectrum $\operatorname{Spec}\mathcal{R}^*_{\mathcal{S}}$ of the *algebraic coordinate ring*

$$\mathcal{R}^*_{\mathcal{S}} \coloneqq \mathcal{U}_{\mathcal{S}}^{-1}(\mathcal{R}/(\mathcal{J} + \mathcal{I}_{\mathcal{S}})).$$

This is the set of all prime ideals in $\mathcal{R}$ lying above $\mathcal{J} + \mathcal{I}_{\mathcal{S}}$ and not intersecting the monoid $\mathcal{U}_{\mathcal{S}}$. By identifying $\mathcal{R}/(\mathcal{J}+\mathcal{I}_{\mathcal{S}})$ with $(\mathcal{R}/\mathcal{J})/\mathcal{I}_{\mathcal{S}} = \mathbb{K}[\Gamma]/\mathcal{I}_{\mathcal{S}}$ (tacitly identifying brackets and their images under ψ), this is a ring-theoretic version of the set of all symmetric matrices which satisfy the equations and the inequations in $\mathcal{S}$ plus principal regularity, thus $\mathcal{R}(\mathcal{S})$. There is always a fixed field $\mathbb{K}$ in the background; the elements of $\operatorname{Spec}(\mathcal{R}_{\mathcal{S}})$ correspond to realizations of $\mathcal{S}$ in algebraic extensions of $\mathbb{K}$. In particular, over an algebraically closed field, the maximal ideals in $\mathcal{R}_{\mathcal{S}}$ correspond exactly to the realizing matrices in $\mathcal{R}(\mathcal{S})$. No realization exists over $\overline{\mathbb{K}}$ if and only if $\operatorname{Spec}(\mathcal{R}_{\mathcal{S}})$ is empty, which happens if and only if $\mathcal{J} \cap (\mathcal{I}_{\mathcal{S}} + \mathcal{U}_{\mathcal{S}}) \neq \emptyset$.

Definition 3.49. A $\mathfrak{g}_k^*$-*final polynomial* for a CI constraint system $\mathcal{S}$ over $\mathbb{K}$ is a bracket polynomial $f \in \mathfrak{J} \cap (\mathfrak{I}_{\mathcal{S}} + \mathfrak{U}_{\mathcal{S}})$.

The way to read this definition is that a polynomial f is final for (the attempt to algebraically realize) $\mathcal{S}$ over $\mathbb{K}$ if it is a universal relation among the principal and almost-principal minors of a symmetric matrix, i.e., $f \in \mathfrak{J}$, so f evaluates to zero on every symmatric matrix, but on the model of $\mathcal{S}$ it also evaluates to the sum of zero ($\mathfrak{I}_{\mathcal{S}}$) and non-zero ($\mathfrak{U}_{\mathcal{S}}$), which is non-zero. More bluntly, a final polynomial infers that $0 = 1$ from the existence of an algebraic realization of $\mathcal{S}$. We apply this to CI structures and their realization spaces. The following result and its corollary are stated analogously for matroids in [BS89, Section 4.2]:

Theorem 3.50. Let $\mathcal{L}$ be a CI structure and $\mathbb{K}$ be a prime field. Then precisely one of the following occurs:

(a) $\mathcal{L}$ is algebraically realizable over $\overline{\mathbb{K}}$.
(b) There exists a final polynomial for $\mathcal{L}$ over $\mathbb{K}$.

Proof. Suppose that a final polynomial f exists and that Γ realizes $\mathcal{L}$ algebraically over $\overline{\mathbb{K}}$. Then $(\psi f)(\Gamma) = 0$ since $f \in \mathfrak{J}$ but also $(\psi f)(\Gamma) = h(\Gamma) + g(\Gamma) \neq 0$ since $\psi f = h + g \in \mathfrak{I}_{\mathcal{S}} + \mathfrak{U}_{\mathcal{S}}$, a contradiction. If $\mathcal{L}$ is algebraically realizable over $\overline{\mathbb{K}}$, then there is a maximal ideal $\mathfrak{m} \in \operatorname{Spec}(\mathcal{R}^*_{\mathcal{L}})$, i.e., $\mathfrak{m} \supseteq \mathfrak{J} + \mathfrak{I}_{\mathcal{L}}$ and $\mathfrak{m} \cap \mathfrak{U}_{\mathcal{L}} = \emptyset$. But then in particular $\mathfrak{J} \cap (\mathfrak{I}_{\mathcal{L}} + \mathfrak{U}_{\mathcal{L}}) = (\mathfrak{J} + \mathfrak{I}_{\mathcal{L}}) \cap \mathfrak{U}_{\mathcal{L}} \subseteq \mathfrak{m} \cap \mathfrak{U}_{\mathcal{L}} = \emptyset$ and there is no final polynomial. □

The polynomials generating $\mathfrak{I}_{\mathcal{S}}$ and $\mathfrak{U}_{\mathcal{S}}$ have integer coefficients, so the spectrum can also be considered over the ring $\mathbb{Z}$. Since $\mathbb{Z}$ injects into every field, we immediately obtain

Corollary 3.51. A constraint system $\mathcal{S}$ is not realizable over any field if and only if it has a final polynomial with integer coefficients. □

Corollary 3.52. Let φ be a CI inference formula and $\mathbb{F}_k$ the prime field of characteristic k. If φ is invalid for algebraic Gaussians in characteristic k, then there exists a counterexample covariance matrix with algebraic entries over $\mathbb{F}_k$. If φ is valid, there exists a final polynomial proof for it with coefficients from $\mathbb{F}_k$. □

Remark 3.53. A very accessible example of final polynomial proofs in matroid theory can be found in the introductory chapter of [Ric11]. The chapter is dedicated to different proofs of Pappus's theorem in the projective plane, and Section 1.3 takes the final polynomial perspective. Our ideal $\mathfrak{J}$ is replaced by the ideal of Grassmann–Plücker relations and brackets are used in synthetic geometry to denote maximal minors of a rectangular matrix.

Example 3.54: Example 3.34, final polynomial version. Example 3.34 is not realizable over any field. The manual proof can be converted into a final polynomial with the assistance of `Macaulay2`. We start over characteristic zero. To fix the notation, let the generic symmetric matrix be

$$\begin{array}{cc} \begin{matrix} 1 & 2 & 3 & 4 & 5 \end{matrix} & \\ \begin{pmatrix} p & a & b & c & d \\ a & q & e & f & g \\ b & e & r & h & i \\ c & f & h & s & j \\ d & g & i & j & t \end{pmatrix} & \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} \end{array}$$

In the manual calculation, it was helpful to suppose that the diagonal elements are all equal to one, but to produce a proper final polynomial, we make no such assumption. Recall that the non-realizability of $\mathcal{G}$ in Example 3.34 was proved using all equations encoded in $\mathcal{G}$ but only some of the inequations: the vanishing of g is implied by the non-vanishing of all diagonals p, q, r, s, t, the off-diagonals d, f, h as well as the principal minor $pt - d^2$. This suggests that the product of g with all these non-vanishing polynomials, which lies in the non-vanishing monoid $\mathfrak{U}_{\mathcal{G}}$, might also lie in $\mathfrak{I}_{\mathcal{G}}$. This can be confirmed with a computer algebra system, which can then also return the product as a linear combination of the generators of $\mathfrak{I}_{\mathcal{G}}$:

```
-- Setup from the inference version (without unit diagonals)
R = QQ[p,a,b,c,d, q,e,f,g, r,h,i, s,j, t];
X = genericSymmetricMatrix(R,p,5);
I = ideal(
  det X_{0}^{1}, det X_{0,3}^{2,3}, det X_{0,4}^{3,4},
  det X_{1,4}^{2,4}, det X_{2,0}^{4,0}, det X_{3,1}^{4,1},
  det X_{0,1,2}^{4,1,2}, det X_{2,0,1}^{3,0,1},
  det X_{1,0,2,4}^{3,0,2,4}
);
-- Test our hypothesis
g*d*f*h*p*q*r*s*t*(p*t-d^2) % I --> 0, meaning 𝔘_𝒱 ∩ 𝔍_𝒱 ≠ ∅ in ℚ[X].
-- Indeed, even the shorter element of 𝔘_𝒱 lies in 𝔍_𝒱:
U = g*h*p*q*r*(p*t-d^2);
U % I --> 0
-- Get a proof that U is in I:
G = gens I; -- the equations generating 𝔍_𝒱
H = U // G; -- linear combinators for U from G
U == G*H --> true
```

The variable `H` harbors the following equation which is a final polynomial for $\mathcal{G}$:

$$
\begin{gathered}
[25|][34|] \cdot [1][2][3][15] = \\
\Big(cd^2egr + bd^2fgr - ad^2grh - 2cd^2e^2i - 2bd^2efi - 2pdfgri + 2ad^2ehi + 2pdefi^2 - 2pdqhi^2 + 2pcqi^3 + \\
2pdqrij - 2pbqi^2j - pcegrt + pbfgrt + pagrht + 2pce^2it - 2pcqrit + 2pbqhit - 2paehit\Big) \cdot [12|] + \\
\Big(pdqer + pbqgr - 2pbqei\Big) \cdot [14|5] - \Big(pcdqr + p^2fgr - 2pbcqi + 2pb^2qj - 2p^2qrj\Big) \cdot [23|5] + \\
\Big(cdqgr - 2cdqei + 2pqghi - 2pqfi^2 - pqgrj + 2pqeij - 2pe^2ft + 2pqfrt\Big) \cdot [35|1] + \\
\Big(pd^2er - 2pbdei + p^2gri + 2pb^2et - 2p^2ert\Big) \cdot [45|2] - \Big(2pdfi - 2pbft\Big) \cdot [15|23] - \\
\Big(d^2gr - 2d^2ei - pgrt + 2peit\Big) \cdot [34|12] - 2pqi \cdot [24|135].
\end{gathered}
$$

Here we have avoided to convert all variables in the linear combinators to their bracket form to save some space. This single polynomial in the bracket ring is a certificate for the non-realizability of $\mathcal{G}$. Despite its size, it is a **straightforward** symbolic computation to expand the brackets and verify that this equation lies in $\mathcal{J}_5$. The left-hand side must not vanish on any realization of $\mathcal{G}$ while the right-hand side must vanish on every realization. Thus, there is no realization. Since all coefficients are integral, this polynomial is final over every field. $\triangle$

The power of the final polynomial method lies in the ease of verification. Every non-realizable CI structure has a final polynomial with coefficients from a prime field. This polynomial can be stored exactly in a database. The verification of this certificate by a user of the database is fast, exact and, unlike the process of obtaining it, requires no ingenuity.

Example 3.55. The gaussoid axioms are necessary for algebraic realizability by Proposition 3.8, thus if a CI structure is not a gaussoid, there must be a final polynomial for it. Consider for example $\mathcal{L} = \{\, (12|), (23|1) \,\}$. The gaussoid axioms are derived from the Matúš identity $[kL][ij|L] = [L][ij|kL] + [ik|L][jk|L]$, all instances of which are elements of $\mathcal{J}$, as shown in Lemma 3.5. There is one instance of the identity, $[1][23|] = [\emptyset][23|1] + [12|][13|]$, which serves as a final polynomial already: $[1][23|] \in \mathfrak{U}_{\mathcal{L}}$ on the left-hand side is also contained in $\mathfrak{J}_{\mathcal{L}}$ by the right-hand side. $\triangle$

It is no accident that a failure of the gaussoid axioms results in a final Matúš identity. In a sense, gaussoids are combinatorial shadows of the Matúš identity in precisely the right way to align with final polynomial proofs. This is made more precise in Section 6.2.

Positive final polynomials. The definition for positive realizability is similar but more involved, as real algebra always is. The positive model $\mathcal{R}^+_{\mathbb{K}}(\mathcal{S})$ over an ordered field $\mathbb{K}$ is defined by equations, inequations and the strict inequalities for positive definiteness. The *positive cone* $\mathcal{P} = \mathcal{P}_{\mathsf{N}}$ is generatd in $\mathcal{R}$ by the principal minor brackets. The following definition is directly inspired by the Positivstellensatz:

Definition 3.56. A $\mathfrak{g}^+$*-final polynomial* for a CI constraint system $\mathcal{S}$ over an ordered field $\mathbb{K}$ is a bracket polynomial $f \in \mathcal{J} \cap (\mathcal{I}_{\mathcal{S}} + \mathcal{U}_{\mathcal{S}}^2 + \mathcal{P})$.

Theorem 3.57. Let $\mathcal{L}$ be a CI structure. Then precisely one of the following occurs:

(a) $\mathcal{L}$ is positively realizable over $\widetilde{\mathbb{Q}}$.
(b) There exists a final polynomial for $\mathcal{L}$ over $\mathbb{Q}$.

Proof. Let f be a final polynomial and suppose $\Sigma \in \mathrm{PD}$ realizes $\mathcal{L}$. By Tarski's transfer principle we may assume that Σ has entries in $\widetilde{\mathbb{Q}}$. Since $f \in \mathcal{J}$ we have $(\psi f)(\Gamma) = 0$, but also $\psi f = h + g^2 + p \in \mathcal{I}_{\mathcal{S}} + \mathcal{U}_{\mathcal{S}}^2 + \mathcal{P}$ and so $(\psi f)(\Gamma) > 0$, a contradiction. Conversely, suppose that the positive realization space of $\mathcal{L}$ is empty. By the Alternatives in real algebraic geometry this implies the existence of $p \in \mathcal{P}$, $g \in \mathcal{U}_{\mathcal{S}}$ and $h \in \mathcal{I}_{\mathcal{S}}$ such that $h + g^2 + p = 0$. The inverse image of $f = h + g^2 + p$ under ψ lies in $\mathcal{J}$ and this is the final polynomial. □

Corollary 3.58. Let φ be a CI inference formula. If φ is invalid for positive Gaussians, then there exists a counterexample covariance matrix with real algebraic entries. If φ is valid, there exists a final polynomial proof for it with rational coefficients. □

Remark 3.59. Over ordered fields, the notion of CI constraint system may be broadened to include **sign constraints** on almost-principal minors. Statistically, this means specifying for a CI statement $(\mathsf{ij}|\mathsf{K})$ of Gaussian random variables whether independence holds ($[\mathsf{ij}|\mathsf{K}] = 0$), or whether the variables i and j are positively ($[\mathsf{ij}|\mathsf{K}] > 0$) or negatively ($[\mathsf{ij}|\mathsf{K}] < 0$) correlated, given K. The framework of real algebra offers an obvious analogue of final polynomials for these *oriented constraint systems* and a corresponding theorem of the alternative is proved in the same way.

In general, polynomials which are "obviously" non-zero on the realization space despite lying in $\mathcal{I}_{\mathcal{S}}$ are final polynomials. For example, if a polynomial of the form $g(p+q)$ is found with $g \in \mathcal{U}_{\mathcal{L}}$, $p \in \mathcal{U}_{\mathcal{L}} \cap \mathcal{P}$ and $q \in \mathcal{P}$ and if it can be shown that $g(p+q) \in \mathcal{I}_{\mathcal{L}}$, then this is a final polynomial in disguise. Despite lying in $\mathcal{I}_{\mathcal{L}}$, when evaluated on the realization space $g(p+q)$ is non-zero times strictly-positive and thus it cannot vanish. This ought to be a final polynomial. To derive its required form, use the bracket polynomial $h \in \mathcal{I}_{\mathcal{L}}$ such that $h + g(p+q) = 0$ and write $0 = g(p+q)h + g^2(p+q)^2 = g(p+q)h + (gp)^2 + [2g^2pq + g^2q^2] \in \mathcal{I}_{\mathcal{L}} + \mathcal{U}_{\mathcal{L}}^2 + \mathcal{P}$. In our setting the condition that $p \in \mathcal{U}_{\mathcal{L}} \cap \mathcal{P}$ is common because all generators of $\mathcal{P}$ are also in $\mathcal{U}$. We denote the set of such $p + q$ by $\mathcal{P}_{\mathcal{L}}^+$ in the example below.

Example 3.60. Inspection of the proof of the Lněnička–Matúš axioms (LM.i)–(LM.v) from [LM07, Lemma 10] yields final polynomial proofs for the positive non-realizability of their antecedent sets. The given proofs are often **almost** final polynomials, in that a relation $gf \in \mathcal{I}_{\mathcal{L}}$ is derived where $g \in \mathcal{U}_{\mathcal{L}}$ and f is (geometrically) positive on the realization space. This is a priori weaker than the algebraic positivity that is proved by writing f as an element of $\mathcal{P} + \mathcal{U}_{\mathcal{L}}^2$. This is all that remains to be done.

$\boldsymbol{\mathcal{L} = \{\ (12|3), (13|4), (14|2)\ \}}$**:** By repeated substitution under the quadratic binomials which generate $\mathcal{I}_\mathcal{L}$, one finds $a(q^2r^2s^2 - d^2e^2f^2) \in \mathcal{I}_\mathcal{L}$. This is not a final polynomial because the term $q^2r^2s^2 - d^2e^2f^2$ is not apparently an element of the cone $\mathcal{P}$. However, Lněnička and Matúš argue that geometrically its positivity on the realization space follows from the principal minors $[23] = qr - d^2 > 0$, $[24] = qs - e^2 > 0$ and $[34] = rs - f^2 > 0$. **By chance**, asking `Macaulay2` for a $\mathbb{Q}[\Gamma]$-linear combination of the above difference in terms of these principal minors turns up the **positive** linear combination: $q^2r^2s^2 - d^2e^2f^2 = e^2f^2[23] + qr^2s[24] + qe^2r[34] \in P_\mathcal{L}$. This is an algebraic proof of the positivity of $q^2r^2s^2 - d^2e^2f^2$ on the realization space. Thus we have the polynomial

$$a(q^2r^2s^2 - d^2e^2f^2) = a(e^2f^2[23] + qr^2s[24] + qe^2r[34]) \in \mathcal{I}_\mathcal{L} \cap \mathcal{U}_\mathcal{L}\mathcal{P}^+_\mathcal{L},$$

which can easily be transformed into an element of $\mathcal{J} \cap (\mathcal{I}_\mathcal{L} + \mathcal{U}^2_\mathcal{S} + \mathcal{P})$.

$\boldsymbol{\mathcal{L} = \{\ (12|3), (13|4), (24|1), (34|2)\ \}}$**:** Using the same ideas as in the previous case, we obtain the relation $a(pqrs - c^2d^2) = a(qr[14] + c^2[23]) \in \mathcal{I}_\mathcal{L} \cap \mathcal{U}_\mathcal{L}\mathcal{P}^+_\mathcal{L}$.

In the other cases, a final polynomial can be assemled from the geometric arguments provided by Lněnička and Matúš, working around the difference between their proof which uses a unit diagonal and our requirement for an element in the homogeneous ideal $\mathcal{J}$:

$\boldsymbol{\mathcal{L} = \{\ (23|), (14|2), (14|3), (23|14)\ \}}$**:**

$$e^2f^2[123] + b^2f^2q^2s \in \mathcal{I}_\mathcal{L} \cap (\mathcal{U}^2_\mathcal{L}\mathcal{P}^+_\mathcal{L} + \mathcal{P}).$$

$\boldsymbol{\mathcal{L} = \{\ (13|), (14|2), (24|3), (23|14)\ \}}$**:**

$$d(pq^2r[34] + a^2f^2[23]) \in \mathcal{I}_\mathcal{L} \cap \mathcal{U}_\mathcal{L}\mathcal{P}^+_\mathcal{L}.$$

$\boldsymbol{\mathcal{L} = \{\ (13|), (24|), (14|23), (23|14)\ \}}$**:**

$$(c^2qr + d^2ps)[1234] + qs(acr + dfp)^2 + pr(ads + cfq)^2 \in \mathcal{I}_\mathcal{L} \cap (\mathcal{U}^2_\mathcal{L}\mathcal{P}^+_\mathcal{L} + \mathcal{P}).$$

Note how the sums of square help to conduct an element of $\mathcal{U}_\mathcal{L}$ into $\mathcal{I}_\mathcal{L}$. Indeed, by Remark 3.21, there is no algebraic final polynomial for any of the above gaussoids, so the sums of squares are **necessary**. Once these terms are found, `Macaulay2` can be used as in Example 3.54 to compute the linear combinators for $\mathcal{I}_\mathcal{L}$ and thus the complete final polynomial. This routine step is omitted here for brevity. △

It should be noted that the above example makes the paper [LM07] by no means redundant because it also constructs sample realizations for the $58 - 5 = 53$ isomorphism types of positive 4-Gaussians. But with the final polynomials derived above, the algebraic and positive Gaussians on ground sets of size 3 and 4 are now classified in a computer-checkable way: for each gaussoid there is either a rational realizing matrix or a rational final polynomial.

In this example, final polynomials could be derived from preexisting non-realizability proofs. In general, one would employ semidefinite programming methods to search for the required sums of squares combinators for a candidate final polynomial, as described in [Par03]. The currently available SDP tools are numeric and may not produce the exact results we desire. The recent work [MP21] in the context of codes in the real projective plane developed criteria for turning approximate certificates into exact ones.

4

Structure of Gaussians and their axioms

The algebraic point of view taken in the previous chapter is developed further in this chapter. It is fruitful especially when fields of rational functions are considered and the concerns of satisfying the algebraic CI equations and of ensuring positive definiteness can effectively be treated separately. The arising methods are used to investigate closure properties of algebraic Gaussians and their conditional independence axioms. We prove that algebraic realizability over infinite fields cannot be characterized by a finite collection of forbidden minors, or equivalently a finite list of CI inference axioms, in Theorem 4.47. This result was known for positive Gaussians but the available proof does not generalize. Furthermore, the gaussoid axioms resurface as a complete characterization of all valid CI inference formulas for algebraic and positive Gaussians with up to two antecendents, by Theorem 4.58. The most difficult positive-definite realizability proofs step through the setting of algebraic Gaussians over rational function fields.

4.1 A rational transfer principle and its consequences

The basic device for proving the structural properties of Gaussians over *infinite* fields is the following lemma. It allows constructions to step through the field of rational functions over the base field, where some genericity conditions are more easily enforced. This technique is used extensively in the remainder of this chapter.

Lemma 4.1. Consider the following two situations:

(1) $\mathbb{K}$ is an infinite field and $\mathbb{L} = \mathbb{K}(x_1, \dots, x_p)$ the field of rational functions in $x_1, \dots, x_p$.

(2) $\mathbb{K}$ is an ordered field and $\mathbb{L} = \mathbb{K}(\varepsilon_1, \dots, \varepsilon_p)$ the ordered field of rational functions in infinitesimals $0 < \varepsilon_1 < \dots < \varepsilon_p$.

In both situations, a gaussoid is realizable over $\mathbb{L}$ if and only if it is already realizable over $\mathbb{K}$.

Lemma 4.2. Let $\mathbb{K}$ be an ordered field and $\mathbb{L} = \mathbb{K}(\varepsilon)$ with a positive infinitesimal ε. If $f \in \mathbb{K}[\varepsilon]$ is positive in $\mathbb{L}$, then there exists $a^* > 0$ such that $f(a) > 0$ in $\mathbb{K}$ for all $0 < a < a^*$.

Proof. By dividing out powers of $\varepsilon > 0$, we can assume that $f = \sum_i c_i \varepsilon^i$ with $c_0 \neq 0$. By the definition of the ordering of $\mathbb{L}$, $f > 0$ means $c_0 > 0$. Without loss of generality, assume that all the other coefficients are negative. For $0 < a < 1$ we have $a^i \le a$ and thus we have to satisfy:

$$\sum_{i \ge 1} (-c_i) a^i \le a \cdot \sum_{i \ge 1} (-c_i) =: a\gamma \overset{!}{<} c_0.$$

It suffices to pick $a^* = \min\{1, c_0/\gamma\}$, which is positive. □

Proof of Lemma 4.1. One inclusion is obvious by the inclusion of fields. In the other direction, it suffices to show how to adjoin one variable x or one infinitesimal ε, so the proof proceeds by induction on p.

(1) Let $\mathbb{K}$ be an infinite field and Γ principally regular over $\mathbb{L}$. The CI structure $[\![\Gamma]\!]$ is defined by vanishing and non-vanishing constraints on principal and almost-principal minors of Γ. These are polynomials in the entries of Γ and therefore rational functions over $\mathbb{K}$. If a rational function $f \in \mathbb{L}$ is zero in $\mathbb{L}$, then every evaluation $f(a)$ for $a \in \mathbb{K}$ is zero. Otherwise, when f is non-zero in $\mathbb{L}$, then the function f on $\mathbb{K}$ has finitely many poles and outside of these finitely many zeros, because the numerator and denominator of f are univariate non-zero polynomials over the field $\mathbb{K}$. Since $\mathbb{K}$ is infinite, one can find a point $a \in \mathbb{K}$ avoiding all the undesirable poles and zeros and such that $[\![\Gamma(a)]\!] = [\![\Gamma]\!]$, where $\Gamma(a)$ is now a principally regular matrix over $\mathbb{K}$.

(2) Suppose $\mathbb{K}$ is ordered. This implies that its characteristic is zero and in particular that it is infinite. Let Γ positively realize a CI structure over $\mathbb{L}$. Again we seek a positive realization of $[\![\Gamma]\!]$ over $\mathbb{K}$ by plugging in elements of $\mathbb{K}$ for ε. By the previous part of the proof, the "algebraic part" of positive realizability, i.e., the vanishing and non-vanishing conditions of almost-principal minors, is satisfied on all but finitely many points. It remains to find infinitely many points of $\mathbb{K}$ on which all principal minors of Γ evaluate to positive elements of $\mathbb{K}$. By the hypothesis, the principal minors of Γ are positive in the ordering of $\mathbb{L}$. We can assume that numerators and denominators are both positive. Let a^* be the minimum of the numbers guaranteed to exist by Lemma 4.2 for all the (numerators and denominators of) principal minors of Γ. Since $a^* > 0$, the interval $(0, a^*)$ is infinite, and all but finitely many evaluations of Γ on this interval yield a positive realization of $[\![\Gamma]\!]$ over $\mathbb{K}$. □

Remark 4.3. This result applies in greater generality:

(1) Given a concrete gaussoid and a realizing matrix over $\mathbb{K}(x_1, \ldots, x_p)$, the proof works for all finite fields of sufficient size. A lower bound can be given based on the size of the given matrix and the maximal degree of its entries. Less constructively, it follows from the Lefschetz Principle that for fixed N all gaussoids realizable in the algebraic closure of $\mathbb{K}$ are realizable in $\mathbb{K}$ if $\mathbb{K}$ is a sufficiently large extension of its prime field.

(2) The proof exploits only the finiteness of the polynomial system whose solutions are the realizations of a gaussoid. It shows that a constructible set (inside a suitable algebraically closed field) has a $\mathbb{K}$-rational point if and only if it has a $\mathbb{K}(x_1, \ldots, x_p)$-rational point. The same is true for semialgebraic sets in the ordered setting. In particular, this lemma can be used to formally construct realizations of matroids and oriented matroids in the same manner as displayed in the rest of this chapter. However, the author is not aware of any such applications.

This is a "transfer principle" like the Lefschetz Principle or Tarski's transfer principle, but it is restricted to field extensions that take the form of rational function fields. Their great advantage over the other transfer principles is that the base field $\mathbb{K}$ does not have to be enlarged (to its algebraic or real closure) to recover a solution from the function field.

For geometric intuition, suppose that $\mathbb{K} = \mathbb{R}$ for the moment. Inspection of the previous proof then paints the following picture of this transfer technique: we define a space of real matrices parametrized by rational functions in variables $\varepsilon_1, \ldots, \varepsilon_p$. In fact, we can replace all infinitesimals by powers of a single infinitesimal and imagine a *curve segment* of matrices parametrized by ε. By the defining rational functions, we control the algebraically realized CI structure on this curve, and as ε tends to zero, the matrices may approach a limit matrix whose principal regularity or positive definiteness carries over to them by continuity. In this way a certificate for algebraic or positive realizability of the CI structure on the curve over the base field $\mathbb{K}$ is obtained. The appeal to continuity and limits can be avoided by an easy sufficient condition, which is the subject of the next definition and is justified by

Lemma 4.4. Let $\Gamma \in \mathrm{Sym}_{\mathsf{N}}(\mathbb{K}(x_1, \dots, x_p))$ and f a polynomial in the entries of a generic symmetric $\mathsf{N} \times \mathsf{N}$ matrix with coefficients in $\mathbb{K}$. If there exist $a_1, \dots, a_p \in \mathbb{K}$ such that the entry-wise evaluation $\Gamma(a_1, \dots, a_p)$ is a well-defined matrix over $\mathbb{K}$ and $f(\Gamma(a_1, \dots, a_p)) \neq 0$ in $\mathbb{K}$, then $f(\Gamma) \neq 0 \in \mathbb{K}(x_1, \dots, x_p)$.

Proof. The value $f(\Gamma)$ is a rational function in $\mathbb{K}(x_1, \dots, x_p)$. Because the evaluation of the polynomial f commutes with the evaluation of rational functions inside of Γ, we see that $f(\Gamma(a_1, \dots, a_p)) = (f(\Gamma))(a_1, \dots, a_p)$ and this proves the lemma because a rational function $f(\Gamma)$ taking a non-zero value cannot be the zero element in $\mathbb{K}(x_1, \dots, x_p)$. □

More concretely, suppose that $\Gamma \in \mathrm{Sym}_{\mathsf{N}}(\mathbb{L})$ with $\mathbb{L}$ as in Lemma 4.1 and that the denominators of all of its entries have a non-zero constant term. Then the evaluation $\Gamma^\circ := \Gamma(0, \dots, 0)$ is a matrix over $\mathbb{K}$. Notice that each minor of Γ° is the constant term of the corresponding minor of Γ and thus principal regularity and positive definiteness of Γ° over $\mathbb{K}$ imply that of Γ over $\mathbb{L}$, allowing application of Lemma 4.1 to see that $[\![\Gamma]\!]$ is a $\mathbb{K}$-algebraic gaussoid.

Definition 4.5. Let $\mathbb{K}$ be a field and Γ_0 a principally regular matrix. A gaussoid $\mathcal{G}$ is *realizable near* Γ_0 if there exists Γ over $\mathbb{K}(x_1, \dots, x_p)$ such that $\mathcal{G} = [\![\Gamma]\!]$ and $\Gamma^\circ = \Gamma_0$. If such a Γ can be chosen over $\mathbb{Q}(x_1, \dots, x_p)$, we add the adverb *rationally*.

Realizability near a positive-definite matrix immediately implies positive realizability by Lemma 4.1. What this definition achieves is an algebraization of positive realizability: to show that a gaussoid $\mathcal{G}$ is positively realizable over an ordered field $\mathbb{K}$, it suffices to find an *algebraic* realization Γ of $\mathcal{G}$ in the rational functions $\mathbb{K}(x_1, \dots, x_p)$ and point out that the evaluation Γ° is positive-definite in $\mathbb{K}$.

Lněnička and Matúš systematically use this technique but do not formally introduce it. It is on vivid display in their collection of realizations of 4-gaussoids in [LM07, Table 1]. Moreover, they give a particularly useful application of this technique in their Theorem 1 about graphic gaussoids (recall Section 1.3), which we reprove here as an example:

Theorem 4.6. A graphic gaussoid is algebraic over every infinite field and rationally realizable near the identity matrix.

Proof. Let $\mathcal{G} = [\![G]\!]$ for an undirected graph $G = (\mathsf{N}, E)$. Consider the matrix Γ with entries

$$\gamma_{\mathsf{ij}} := \begin{cases} 1, & \text{if } \mathsf{i} = \mathsf{j}, \\ 0, & \text{if } \mathsf{ij} \notin E, \\ \varepsilon_{\mathsf{ij}}, & \text{if } \mathsf{ij} \in E. \end{cases}$$

This is an adjacency matrix of G with implicit loops on every vertex and generic, independent weights $\varepsilon_{\mathsf{ij}}$ on every edge. Γ is principally regular (and positive-definite in the ordered setting) over $\mathbb{K}(\varepsilon_{\mathsf{ij}})$ since $\Gamma^\circ = \mathbb{1}_{\mathsf{N}}$. By construction $(\mathsf{ij}|\mathsf{N}^{\mathsf{ij}}) \in [\![G]\!] \Rightarrow (\mathsf{ij}|\mathsf{N}^{\mathsf{ij}}) \in [\![\Gamma]\!]^{\rceil}$. This implies $[\![G]\!] \subseteq [\![\Gamma]\!]^{\rceil}$ by [LM07, Lemma 3]. This lemma requires only the pseudographoid property on $[\![\Gamma]\!]$ and proceeds axiomatically.

The other inclusion is of more interest regarding Lemma 4.1. In general, we have the following expression for an almost-principal minor of an adjacency matrix Γ, which follows after marginalization and inversion from the Jones–West formula [JW05] (see also the version in [BKKR21, Proposition 3.19]):

$$\pm\Gamma[\mathsf{ij}|\mathsf{L}] = \sum_{\substack{p \text{ path} \\ \mathsf{i} \to \mathsf{j} \text{ in } G_{\mathsf{ijL}}}} \pm\Gamma[\mathsf{L} \setminus p] \cdot \gamma^p, \tag{$*$}$$

where the sign of each summand depends on p (but is immaterial for our argument), $\Gamma[\mathsf{L} \setminus p]$ is the principal minor for the set of vertices not on p and γ^p is the product of entries of Γ over all edges on p. Suppose that $(\mathsf{ij}|\mathsf{K}) \notin [\![G]\!]$. Then there exists a path p^* from i to j which does not pass through K. This is a path in G_{ijL} with $\mathsf{L} = \mathsf{N}^{\mathsf{ij}} \setminus \mathsf{K}$. This means that at least one summand in ($*$) is non-zero. The principal minor $\Gamma[\mathsf{L} \setminus p^*]$ has constant term 1 by the definition of the diagonal of Γ, thus the monomial γ^{p^*} appears in $\Gamma[\mathsf{ij}|\mathsf{L}]$. The monomial γ^p uniquely identifies its path p, so γ^{p^*} cannot be canceled in the summation and $\Gamma[\mathsf{ij}|\mathsf{L}]$ is non-zero in $\mathbb{K}[\varepsilon_{\mathsf{ij}}]$, which proves $(\mathsf{ij}|\mathsf{N}^{\mathsf{ij}} \setminus \mathsf{K}) \notin [\![\Gamma]\!]$ and thus $[\![G]\!] = [\![\Gamma]\!]^{\rceil}$. □

Remark 4.7. The adjacency matrix construction in the preceding proof, when generic instead of infinitesimal values are used, provides even a parametrization of the Gaussian graphical model. The same technique works for directed acyclic graphs and their associated Bayesian network models as well; see [Sul18, Section 13.2].

Example 4.8. Consider the 4-gaussoid $\mathcal{L}_5 \coloneqq \{\,(12|4), (13|2), (24|3), (34|1)\,\}$. With the ordering $\mathsf{ijkl} = 1324$, this is the antecedent set to (LM.v). Hence, it is not realizable by a positive-definite matrix over $\mathbb{R}$ (or any ordered field, by Tarski's transfer principle). We examine the algebraic realizability of $\mathcal{L}_5$ over small fields and in $\overline{\mathbb{F}_2}$.

(1) An exhaustive search (or Theorem 4.50) reveals that $\mathcal{L}_5$ has no principally regular realization in $\mathbb{F}_2$ or $\mathbb{F}_3$, but it is realizable over $\mathbb{F}_5$ by

$$\begin{pmatrix} 1 & 1 & 1 & 2 \\ 1 & 3 & 2 & 1 \\ 1 & 2 & 2 & 2 \\ 2 & 1 & 2 & 1 \end{pmatrix}.$$

This is the smallest prime field over which $\mathcal{L}_5$ is realizable.

(2) The realization space of $\mathcal{L}_5$ over $\overline{\mathbb{F}_2}$ can be studied with `Macaulay2`:

```
R = ZZ/2[p,q,r,s, a,b,c,d,e,f]
X = matrix{{p,a,b,c}, {a,q,d,e}, {b,d,r,f}, {c,e,f,s}}
I = radical ideal(
    det X_{0,3}^{1,3},   -- (12|4)
    det X_{0,1}^{2,1},   -- (13|2)
    det X_{1,2}^{3,2},   -- (24|3)
    det X_{2,0}^{3,0}    -- (34|1)
  )
J = saturate(I, product gens R)  -- necessary non-vanishings
--> ⟨re + df, bc + pf, qb + ad, sa + ce, pqrs + c^2d^2⟩
```

The following parametrization of the variety over $\overline{\mathbb{F}_2}$ is easily read off from the quadratic generators of this prime ideal and it also satisfies the quartic:

$$(a, c, d, p, q, r) \quad\mapsto\quad b = \frac{ad}{q}, \quad e = \frac{acd^2}{pqr}, \quad f = \frac{acd}{pq}, \quad s = \frac{c^2d^2}{pqr}.$$

This parametrization by rational functions defines the generic realization over infinite fields of characteristic two:

$$\begin{pmatrix} p & a & \frac{ad}{q} & c \\ a & q & d & \frac{acd^2}{pqr} \\ \frac{ad}{q} & d & r & \frac{acd}{pq} \\ c & \frac{acd^2}{pqr} & \frac{acd}{pq} & \frac{c^2d^2}{pqr} \end{pmatrix}.$$

This matrix is indeed principally regular over $\mathbb{F}_2(a, c, d, p, q, r)$ and its CI structure is exactly $\mathcal{L}_5$. This shows that $\mathcal{L}_5$ is realizable over every infinite field of characteristic two, in particular $\overline{\mathbb{F}_2}$.

(3) Finally, realizability over $\overline{\mathbb{F}_2}$ implies realizability over some finite extension of $\mathbb{F}_2$. As noted above, $\mathbb{F}_2$ is too small. An exhaustive search of the $(\mathbb{F}_4^\times)^6 = 729$ possible matrices in the image of the parametrization yields no principally regular matrix over $\mathbb{F}_4$. However, over $\mathbb{F}_8 = \mathbb{F}_2(\alpha)$ with primitive element α, one finds

$$\begin{pmatrix} 1 & \alpha & \alpha^2 & 1 \\ \alpha & 1 & \alpha & \alpha+1 \\ \alpha^2 & \alpha & 1 & \alpha^2 \\ 1 & \alpha+1 & \alpha^2 & \alpha^2 \end{pmatrix},$$

which is principally regular and realizes $\mathcal{L}_5$. △

4.2 Long minimally valid inference rules

The structure $\mathcal{L}_5$ from Example 4.8 is also the instance on $n = 4$ of the family used by Sullivant to prove that positive Gaussians over $\mathbb{R}$ have minimally valid CI inference rules with arbitrarily many antecedents [Sul09]. An inference is *minimally valid* if its antecedent and consequent sets are inclusion-minimal with the property that the formula is valid (for all Gaussians). The topic of finite axiomatization is further discussed in Section 4.4. In this section, we give an analogous construction of arbitrarily long, minimally valid inference rules of Horn type for algebraic Gaussians. The construction depends on the characteristic of the field and the proofs all rely on Lemma 4.1.

For his proof in the positive real case, Sullivant computes the minimal primes of the CI ideals for his family. For each ground set, there are two of them: one which yields the desired minimal inference rule and another one whose variety does not intersect PD_n. The non-positive component can have principally regular points over $\mathbb{C}$ and therefore in positive characteristic as well: an example of this is $\mathcal{L}_5$ from Example 4.8. Thus, the inference rules for algebraic Gaussians require a different family of antecedents.

Let $n \geq 0$ and $\mathsf{N} = \{1, 2, \dots, \mathsf{n}\}$ as usual. We add to the ground set under consideration two distinguished elements i and j and define the CI structure

$$\begin{aligned} \mathcal{L}_n \coloneqq\ & \{ (\mathsf{ij}|\mathsf{N}) \} && (\mathcal{L}.\mathrm{i}) \\ \cup\ & \{ (\mathsf{ij}|\mathsf{k}) : \mathsf{k} \in \mathsf{N} \} && (\mathcal{L}.\mathrm{ii}) \\ \cup\ & \{ (\mathsf{kl}|) : \mathsf{k}, \mathsf{l} \in \mathsf{N} \text{ distinct} \}. && (\mathcal{L}.\mathrm{iii}) \end{aligned}$$

To understand the meaning of these structures, consider a matrix which satisfies all the relations given in $\mathcal{L}_n$:

$$\Phi = \begin{array}{c} \begin{array}{cccccc} \mathsf{i} & \mathsf{j} & \mathsf{1} & \mathsf{2} & \cdots & \mathsf{n} \end{array} \\ \left(\begin{array}{cc|cccc} p_{\mathsf{i}} & z & a_1 & a_2 & \cdots & a_{\mathsf{n}} \\ z & p_{\mathsf{j}} & b_1 & b_2 & \cdots & b_{\mathsf{n}} \\ \hline a_1 & b_1 & p_1 & 0 & \cdots & 0 \\ a_2 & b_2 & 0 & p_2 & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & 0 \\ a_{\mathsf{n}} & b_{\mathsf{n}} & 0 & \cdots & 0 & p_{\mathsf{n}} \end{array}\right) \end{array} \begin{array}{c} \\ \mathsf{i} \\ \mathsf{j} \\ \mathsf{1} \\ \mathsf{2} \\ \vdots \\ \mathsf{n} \end{array},$$

where the zeros in the $\mathsf{N} \times \mathsf{N}$ block are imposed by ($\mathcal{L}$.iii). To satisfy the $(\mathsf{ij}|\mathsf{k})$ relations, it is necessary that

$$z = \frac{a_1 b_1}{p_1} = \frac{a_2 b_2}{p_2} = \cdots = \frac{a_{\mathsf{n}} b_{\mathsf{n}}}{p_{\mathsf{n}}}$$

and furthermore $(\mathsf{ij}|\mathsf{N})$ is then equivalent to

$$z = \begin{pmatrix} a_1 & \cdots & a_{\mathsf{n}} \end{pmatrix} (\Phi_{\mathsf{N}})^{-1} \begin{pmatrix} b_1 \\ \vdots \\ b_{\mathsf{n}} \end{pmatrix} = \sum_{\mathsf{k} \in \mathsf{N}} \frac{a_{\mathsf{k}} b_{\mathsf{k}}}{p_{\mathsf{k}}} = nz, \qquad (*)$$

where $n \in \mathbb{N}$ is, as usual, identified with its image in $\mathbb{K}$ under the natural ring homomorphism $\mathbb{Z} \to \mathbb{K}$. This analysis holds true over every field $\mathbb{K}$. Unless $(\mathsf{ij}|)$ holds, i.e., $z = 0$, the equation $1 = n$ is implied by $\mathcal{L}_n$. This non-trivial assertion about the characteristic of $\mathbb{K}$ is the main idea to obtain the following theorem:

Theorem 4.9. Let $\mathbb{K}$ be infinite. There exist infinitely many n, depending on $\operatorname{char}\mathbb{K}$, for which $\bigwedge \mathcal{L}_n \Rightarrow (\mathsf{ij}|)$ is a minimal valid inference rule for $\mathfrak{g}^*_{\mathbb{K}}$.

Proof. To make the implication of $(\mathsf{ij}|)$ valid, it is necessary that $1 \neq n \in \mathbb{K}$. This can be achieved for infinitely many ground set sizes n in every characteristic. As n grows, so does $|\mathcal{L}_n|$, giving valid inference rules with arbitrarily many antecedents. Since the set of consequents cannot be reduced further, it remains to show that the antecedent set $\mathcal{L}_n$ is minimal. This is done in a series of lemmas, which take different kinds of statements out of $\mathcal{L}_n$ and prove that the subsets do not imply $(\mathsf{ij}|)$ anymore. The removal of a statement of type ($\mathcal{L}$.i) or ($\mathcal{L}$.ii) is dealt with by Lemma 4.12 and Lemma 4.13, respectively. Both work in every characteristic and impose only $n \neq 1 \in \mathbb{K}$. The remaining type ($\mathcal{L}$.iii) is subject to a case distinction:

— If $\operatorname{char}\mathbb{K} = 2$, then Lemma 4.15 applies. This lemma only works in characteristic two but imposes just $n \neq 1 \in \mathbb{K}$, which is satisfied by infinitely many ground set sizes modulo two.
— If $\operatorname{char}\mathbb{K} > 2$, then Lemma 4.16 ensures minimality. This construction requires $n \neq 1$ and $n = 3$ in $\mathbb{K}$. There are infinitely many such n since $\operatorname{char}\mathbb{K}$ is finite but not two.
— If $\operatorname{char}\mathbb{K} = 0$, then Lemma 4.17 proves minimality under the assumption that $n \notin \{1, 2, 3\} \subseteq \mathbb{K}$. Again, there are infinitely many such n in this characteristic.

In each case, there are infinitely many $n \neq 1 \in \mathbb{K}$ so that the inference rules with unbounded number of antecedents are valid and minimal. $\square$

Remark 4.10. This family of CI structures $\mathcal{L}_n$ or variants of it appear in manifold examples. The structure $\mathcal{L}_{\mathsf{ij}} = \{ (\mathsf{ij}|), (\mathsf{ij}|\mathsf{N}) \}$ over ijN is realizable by Lemma 4.63. In [BKKR21, Examples 2.17 and 4.2] it is shown, based on [DX10, Proposition 4.2], that the Zariski closure of its realization space is singular. Its singular locus is described by adding to $\mathcal{L}_{\mathsf{ij}}$ the marginals $(\mathsf{ik}|)$ and $(\mathsf{jk}|)$ for all $\mathsf{k} \in \mathsf{N}$. Since $(\mathsf{ij}|\mathsf{N})$ is implied by $(\mathsf{ij}|)$, $(\mathsf{ik}|)$ and $(\mathsf{jk}|)$, this singular locus is, in fact, a linear space of codimension $2n+1$. A variant of this family appears in Section 6.2.2 to give an infinite set of forbidden minors for gaussoid orientability.

Remark 4.11. The idea to model the equation $(n-1)z = 0$ also appears in matroid theory in the study of characteristic sets. Matúš [Mat99b] points back to [BK80, Section 24.A5, p. 108].

Lemma 4.12. Let $\mathcal{L}' = \mathcal{L}_n \setminus (\mathsf{ij}|\mathsf{N})$. When $n \neq 1 \in \mathbb{K}$, then there exists a principally regular matrix over $\mathbb{K}$ whose CI structure includes $\mathcal{L}'$ but which does not include $(\mathsf{ij}|)$.

Proof. The relation $(\mathsf{ij}|\mathsf{N})$ which is removed is the crucial equation in the above reasoning for forcing $z = 0$. Over an infinite field $\mathbb{K}$, Lemma 4.1 can be applied to construct a realizable gaussoid in a suitable rational function field over $\mathbb{K}$ which contains $\mathcal{L}'$ but does not contain $(\mathsf{ij}|)$:

$$\begin{array}{c@{}l}
\begin{array}{cccccc} \mathsf{i} & \mathsf{j} & \mathsf{1} & \mathsf{2} & \cdots & \mathsf{n} \end{array} & \\
\begin{pmatrix}
1 & xy & x & x & \cdots & x \\
xy & 1 & y & y & \cdots & y \\
x & y & & & & \\
x & y & & & & \\
\vdots & \vdots & & \mathbb{1}_{\mathsf{12\cdots n}} & & \\
x & y & & & &
\end{pmatrix} & \begin{array}{l} \mathsf{i} \\ \mathsf{j} \\ \mathsf{1} \\ \mathsf{2} \\ \vdots \\ \mathsf{n} \end{array}
\end{array}.$$

This matrix is clearly principally regular over $\mathbb{K}(x, y)$ because of Lemma 4.4 with $x = y = 0$. It contains $\mathcal{L}'$ but not $(\mathsf{ij}|)$ and the CI statement $(\mathsf{ij}|\mathsf{1}\cdots\mathsf{n})$ does not hold unless $n = 1 \in \mathbb{K}$. □

Lemma 4.13. Let $\mathcal{L}' = \mathcal{L}_n \setminus (\mathsf{ij}|\mathsf{k})$ for some $\mathsf{k} \in \mathsf{N}$. When $n \neq 1 \in \mathbb{K}$, then there exits a principally regular matrix over $\mathbb{K}$ whose CI structure includes $\mathcal{L}'$ but which does not include $(\mathsf{ij}|)$.

Proof. Without $(\mathsf{ij}|\mathsf{k})$, we do not enforce $zp_{\mathsf{k}} = a_{\mathsf{k}}b_{\mathsf{k}}$ anymore and $(\mathsf{ij}|\mathsf{N})$ is equivalent to

$$0 = (n-2)z + \frac{a_{\mathsf{k}}b_{\mathsf{k}}}{p_{\mathsf{k}}}.$$

We distinguish two cases. If $n = 2 \in \mathbb{K}$, consider the following matrix:

$$\begin{array}{c@{}l}
\begin{array}{cccccccc} \mathsf{i} & \mathsf{j} & \mathsf{1} & \mathsf{2} & \cdots & \mathsf{k} & \cdots & \mathsf{n} \end{array} & \\
\begin{pmatrix}
1 & xy & x & x & \cdots & 0 & \cdots & x \\
xy & 1 & y & y & \cdots & 0 & \cdots & y \\
x & y & & & & & & \\
x & y & & & & & & \\
\vdots & \vdots & & & & & & \\
0 & 0 & & & \mathbb{1}_{\mathsf{12\cdots n}} & & & \\
\vdots & \vdots & & & & & & \\
x & y & & & & & &
\end{pmatrix} & \begin{array}{l} \mathsf{i} \\ \mathsf{j} \\ \mathsf{1} \\ \mathsf{2} \\ \vdots \\ \mathsf{k} \\ \vdots \\ \mathsf{n} \end{array}
\end{array}.$$

Otherwise we have an equation to solve for z which yields the following matrix:

$$\begin{array}{cc}
\begin{array}{cccccccc} \mathsf{i} & \mathsf{j} & \mathsf{1} & \mathsf{2} & \cdots & \mathsf{k} & \cdots & \mathsf{n} \end{array} & \\
\begin{pmatrix}
1 & \frac{xy}{2-n} & x & x & \cdots & x & \cdots & x \\
\frac{xy}{2-n} & 1 & \frac{y}{2-n} & \frac{y}{2-n} & \cdots & y & \cdots & \frac{y}{2-n} \\
x & \frac{y}{2-n} & & & & & & \\
x & \frac{y}{2-n} & & & & & & \\
\vdots & \vdots & & & & & & \\
x & y & & & \mathbb{1}_{\mathsf{12\cdots n}} & & & \\
\vdots & \vdots & & & & & & \\
x & \frac{y}{2-n} & & & & & &
\end{pmatrix} & \begin{array}{c} \mathsf{i} \\ \mathsf{j} \\ \mathsf{1} \\ \mathsf{2} \\ \vdots \\ \mathsf{k} \\ \vdots \\ \mathsf{n} \end{array}
\end{array}.$$

Both matrices are principally regular over $\mathbb{K}(x, y)$ by Lemma 4.4 with $x = y = 0$, they contain $\mathcal{L}'$ but not $(\mathsf{ij}|)$ since $2 - n \neq 1 \in \mathbb{K}$. □

The final case is the removal of $(\mathsf{kl}|)$ of type ($\mathcal{L}$.iii). Without $(\mathsf{kl}|)$, the kl-entry of Φ_N is no longer required to vanish. This situation only affects the equation for $(\mathsf{ij}|\mathsf{N})$ which involves the inverse of Φ_N.

Lemma 4.14. Let $\mathbb{K}$ be a field and Γ a symmetric matrix over $\mathbb{K}$ with diagonal elements $p_1, \ldots, p_\mathsf{n} \neq 0$ and a single non-zero off-diagonal $\gamma_\mathsf{kl} = \gamma_\mathsf{lk} = \gamma$. Let $\Delta := \prod_{\mathsf{t}\in\mathsf{N}} p_\mathsf{t}$. Then:

(a) $\det \Gamma = \delta := \frac{\Delta}{p_\mathsf{k} p_\mathsf{l}} \left(p_\mathsf{k} p_\mathsf{l} - \gamma^2\right)$.
(b) The entries of the adjugate matrix of Γ are:

— the diagonals in place $\mathsf{i} \neq \mathsf{k}, \mathsf{l}$: δ/p_i.
— the diagonals in place $\mathsf{i} = \mathsf{k}$ or $\mathsf{i} = \mathsf{l}$: Δ/p_i.
— the kl- and lk-entry: $-\gamma \frac{\Delta}{p_\mathsf{k} p_\mathsf{l}}$.
— all other entries are zero.

In particular Γ is principally regular if and only if $\gamma^2 \neq p_\mathsf{k} p_\mathsf{l}$.

Proof. By applying a simultaneous row and column permutation under the Sign Convention, we can assume that $\mathsf{kl} = \mathsf{12}$. Thus Γ is block-diagonal with a 2×2 general symmetric block and an $(n - 2) \times (n - 2)$ diagonal block. The assertions follow by straightforward calculation. Principal regularity follows from the determinant formula and the fact that principal submatrices of Γ are either diagonal (which is trivial), or have the same structure as Γ, so the same argument applies recursively. □

The analysis of this case is subdivided into three lemmas, depending on the characteristic of the field. In all cases below, we suppose $\phi_\mathsf{ij} = z = xy$ for independent variables x, y, and also $a_\mathsf{t} = x$ and $b_\mathsf{t} = y$ for all $\mathsf{t} \in \mathsf{N} \setminus \mathsf{kl}$. Additionally set $b_\mathsf{t} = xyp_\mathsf{t} \cdot a_\mathsf{t}^{-1}$ for $\mathsf{t} \in \mathsf{kl}$. The concrete values of a_k and a_l are subject to the a case distinction further below. These settings are valid in a rational function field and the $(\mathsf{ij}|\mathsf{t})$ relations all hold. Furthermore, denote $\gamma := \phi_\mathsf{kl}$. A straightforward calculation based on applying Lemma 4.14 to Φ_N (assuming that it is invertible, i.e., $\gamma^2 \neq p_\mathsf{k} p_\mathsf{l}$) shows that

$$\begin{aligned}
(\mathsf{ij}|\mathsf{N}) \;\Leftrightarrow\; 0 &= p_\mathsf{k} p_\mathsf{l} (n-1)xy - \gamma(a_\mathsf{k} b_\mathsf{l} + a_\mathsf{l} b_\mathsf{k}) - \gamma^2 (n-3)xy \\
&= p_\mathsf{k} p_\mathsf{l} (n-1)xy - \gamma xy \frac{p_\mathsf{l} a_\mathsf{k}^2 + p_\mathsf{k} a_\mathsf{l}^2}{a_\mathsf{k} a_\mathsf{l}} - \gamma^2 (n-3)xy.
\end{aligned}$$

Since $xy \neq 0$ is required, we may work with the equivalent equation

$$0 = p_\mathsf{k} p_\mathsf{l} (n-1) - \gamma \frac{p_\mathsf{l} a_\mathsf{k}^2 + p_\mathsf{k} a_\mathsf{l}^2}{a_\mathsf{k} a_\mathsf{l}} - \gamma^2 (n-3). \tag{$*$}$$

Lemma 4.15. Let $\mathcal{L}' = \mathcal{L}_n \setminus (\mathsf{kl}|)$ for distinct $\mathsf{k}, \mathsf{l} \in \mathsf{N}$. Suppose that $\operatorname{char} \mathbb{K} = 2$ and $n \neq 1 \in \mathbb{K}$. Then there exists a principally regular matrix over $\mathbb{K}$ whose CI structure includes $\mathcal{L}'$ but which does not include $(\mathsf{ij}|)$.

Proof. Since $n \neq 1 \in \mathbb{K}$, but n is an element of the prime field of $\mathbb{K}$, we must have $n = 0 \in \mathbb{K}$ and hence $n - 1 = n - 3 = 1$. Suppose that $p_\mathsf{k} = p_\mathsf{l} = p$. Let ω be a new variable and set $a_\mathsf{k} = \omega$ and $a_\mathsf{l} = \omega + 1$. Using $xy \neq 0$, equation $(*)$ simplifies to

$$0 = p^2 + \frac{\gamma}{\omega(\omega+1)} p + \gamma^2.$$

One can verify that $p = \frac{\omega\gamma}{\omega+1}$ solves this quadratic equation over $\mathbb{K}(\gamma, \omega)$. Thus the matrix

$$\begin{array}{c@{}c}
\begin{array}{ccccccccc} \mathsf{i} & \mathsf{j} & \mathsf{1} & \mathsf{2} & \cdots & \mathsf{k} & \mathsf{l} & \cdots & \mathsf{n} \end{array} & \\
\begin{pmatrix}
1 & xy & x & x & \cdots & \omega & \omega+1 & \cdots & x \\
xy & 1 & y & y & \cdots & \frac{\gamma xy}{\omega+1} & \frac{\omega\gamma xy}{(\omega+1)^2} & \cdots & y \\
x & y & 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
x & y & 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & & \ddots & & & & \vdots \\
\omega & \frac{\gamma xy}{\omega+1} & 0 & & & \frac{\omega\gamma}{\omega+1} & \gamma & & 0 \\
\omega+1 & \frac{\omega\gamma xy}{(\omega+1)^2} & 0 & & & \gamma & \frac{\omega\gamma}{\omega+1} & & 0 \\
\vdots & \vdots & \vdots & & & & & \ddots & \vdots \\
x & y & 0 & 0 & \cdots & 0 & 0 & \cdots & 1
\end{pmatrix} &
\begin{array}{c} \mathsf{i} \\ \mathsf{j} \\ \mathsf{1} \\ \mathsf{2} \\ \vdots \\ \mathsf{k} \\ \mathsf{l} \\ \vdots \\ \mathsf{n} \end{array}
\end{array}$$

satisfies $\mathcal{L}'$ but not $(\mathsf{ij}|)$. It remains to show that it is principally regular over characteristic two. To simplify the treatment set $x = y = 0$ via Lemma 4.4. By Lemma 4.14, the $\mathsf{N} \times \mathsf{N}$ block is principally regular if only $\gamma \neq 0$, which we impose anyway. jN and ijN have a block-diagonal structure which allows the elimination of j, when $x = y = 0$, so the proof is completed by computing the iN-principal minor using the Schur complement with respect to i:

$$\det \begin{array}{c@{}c}
\begin{array}{cccccccc} \mathsf{i} & \mathsf{1} & \mathsf{2} & \cdots & \mathsf{k} & \mathsf{l} & \cdots & \mathsf{n} \end{array} & \\
\left(\begin{array}{c|ccccccc}
1 & 0 & 0 & \cdots & \omega & \omega+1 & \cdots & 0 \\ \hline
0 & 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
0 & 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & & \ddots & & & & \vdots \\
\omega & 0 & & & \frac{\omega\gamma}{\omega+1} & \gamma & & 0 \\
\omega+1 & 0 & & & \gamma & \frac{\omega\gamma}{\omega+1} & & 0 \\
\vdots & \vdots & & & & & \ddots & \vdots \\
0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 1
\end{array}\right) &
\begin{array}{c} \mathsf{i} \\ \mathsf{1} \\ \mathsf{2} \\ \vdots \\ \mathsf{k} \\ \mathsf{l} \\ \vdots \\ \mathsf{n} \end{array}
\end{array} = 1 + \frac{\omega(\omega+1)}{\gamma},$$

which is generically non-zero since γ and ω are independent variables. □

Lemma 4.16. Let $\mathcal{L}' = \mathcal{L}_n \setminus (\mathsf{kl}|)$ for distinct $\mathsf{k}, \mathsf{l} \in \mathsf{N}$. Suppose that $\operatorname{char} \mathbb{K} > 2$ and $n = 3 \neq 1 \in \mathbb{K}$. Then there exists a principally regular matrix over $\mathbb{K}$ whose CI structure includes $\mathcal{L}'$ but which does not include $(\mathsf{ij}|)$.

Proof. Set $p_\mathsf{i} = p_\mathsf{j} = 1$. With $n = 3 \in \mathbb{K}$, equation $(*)$ turns into a linear equation which can be solved for $\gamma = \frac{(n-1)a_\mathsf{k}a_\mathsf{l}}{a_\mathsf{k}^2 + a_\mathsf{l}^2}$. Introduce a new variable ω and set $a_\mathsf{k} = x\omega$ and $a_\mathsf{l} = x\omega^{-1}$, hence $b_\mathsf{k} = y\omega^{-1}$ and $b_\mathsf{l} = y\omega$. The value of γ becomes

$$\gamma = \frac{n-1}{\omega^2 + \omega^{-2}},$$

which is non-zero over $\mathbb{K}(x, y, \omega)$. It follows that this matrix satisfies the CI constraints:

$$
\begin{array}{cc}
\begin{matrix} \mathsf{i} & \mathsf{j} & \mathsf{1} & \mathsf{2} & \cdots & \mathsf{k} & \mathsf{l} & \cdots & \mathsf{n} \end{matrix} & \\
\begin{pmatrix}
1 & xy & x & x & \cdots & x\omega & x\omega^{-1} & \cdots & x \\
xy & 1 & y & y & \cdots & y\omega^{-1} & y\omega & \cdots & y \\
x & y & 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
x & y & 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & & \ddots & & & & \vdots \\
x\omega & y\omega^{-1} & 0 & & & 1 & \frac{n-1}{\omega^2+\omega^{-2}} & & 0 \\
x\omega^{-1} & y\omega & 0 & & & \frac{n-1}{\omega^2+\omega^{-2}} & 1 & & 0 \\
\vdots & \vdots & \vdots & & & & & \ddots & \vdots \\
x & y & 0 & 0 & \cdots & 0 & 0 & \cdots & 1
\end{pmatrix} &
\begin{matrix} \mathsf{i} \\ \mathsf{j} \\ \mathsf{1} \\ \mathsf{2} \\ \vdots \\ \mathsf{k} \\ \mathsf{l} \\ \vdots \\ \mathsf{n} \end{matrix}
\end{array}.
$$

Substitution of $x = y = 0$ gives a matrix over $\mathbb{K}(\omega)$ whose only non-zero off-diagonal entry is again γ. Since $\gamma^2 \neq 1 \in \mathbb{K}(\omega)$, this matrix, by Lemma 4.14, and hence the original, by Lemma 4.4, are principally regular. ☐

Lemma 4.17. Let $\mathcal{L}' = \mathcal{L}_n \setminus (\mathsf{kl}|)$ for distinct $\mathsf{k}, \mathsf{l} \in \mathsf{N}$. Suppose that $\operatorname{char} \mathbb{K} = 0$ and $n \neq 1, 2, 3 \in \mathbb{K}$. Then there exists a principally regular matrix over $\mathbb{K}$ whose CI structure includes $\mathcal{L}'$ but which does not include $(\mathsf{ij}|)$.

Proof. With $p_\mathsf{i} = p_\mathsf{j} = 1$ and $a_\mathsf{k} = a_\mathsf{l} = x$ and $b_\mathsf{k} = b_\mathsf{l} = y$, equation $(*)$ reduces to a quadratic one in γ:

$$0 = (n-1) - 2\gamma - \gamma^2(n-3).$$

Solving this equation yields $\gamma = -\frac{n-1}{n-3}$ (the other solution is identically 1 and would contradict principal regularity). We obtain the following matrix:

$$
\begin{array}{cc}
\begin{matrix} \mathsf{i} & \mathsf{j} & \mathsf{1} & \mathsf{2} & \cdots & \mathsf{k} & \mathsf{l} & \cdots & \mathsf{n} \end{matrix} & \\
\begin{pmatrix}
1 & xy & x & x & \cdots & x & x & \cdots & x \\
xy & 1 & y & y & \cdots & y & y & \cdots & y \\
x & y & 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
x & y & 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & & \ddots & & & & \vdots \\
x & y & 0 & & & 1 & -\frac{n-1}{n-3} & & 0 \\
x & y & 0 & & & -\frac{n-1}{n-3} & 1 & & 0 \\
\vdots & \vdots & \vdots & & & & & \ddots & \vdots \\
x & y & 0 & 0 & \cdots & 0 & 0 & \cdots & 1
\end{pmatrix} &
\begin{matrix} \mathsf{i} \\ \mathsf{j} \\ \mathsf{1} \\ \mathsf{2} \\ \vdots \\ \mathsf{k} \\ \mathsf{l} \\ \vdots \\ \mathsf{n} \end{matrix}
\end{array}.
$$

After substituting $x = y = 0$, one obtains a matrix over $\mathbb{K}$ which has only one non-zero entry γ. Lemma 4.14 shows that this matrix is principally regular when $\gamma^2 \neq 1$, which holds when $n \neq 2$. By construction it fulfills the CI constraints in the lemma. ☐

4.3 Embedding, dependent sum and isolation

From now on let $\mathfrak{g}^\bullet$ denote the class of either of the two cases to which Lemma 4.1 applies: either an infinite field or an ordered field. The underlying field is fixed and denoted by $\mathbb{K}$. All closure properties in this section are derived from this lemma as a black box. The differences between algebraic and positive realizations are immaterial. What matters is that $\mathbb{K}$ is large enough to satisfy the genericity requirements of *embedding* and *dependent sum*.

Let $\mathcal{F}^{\mathsf{N}}$ denote the face lattice of the N-cube and $\mathcal{F}^{\mathsf{N}}_k$ the set of its k-dimensional faces.

Definition 4.18. Let $\mathcal{G}$ be a CI structure over L and $\mathsf{m} \notin \mathsf{L}$. The *unmarginalization* of $\mathcal{G}$ with respect to m is the CI structure $\mathcal{G} \subseteq \mathcal{A}_{\mathsf{mL}}$. Dually, the *unconditioning* is $\{ (\mathsf{ij}|\mathsf{mK}) : (\mathsf{ij}|\mathsf{K}) \in \mathcal{G} \} \subseteq \mathcal{A}_{\mathsf{mL}}$. Every combination of these operations is an *embedding* of $\mathcal{G}$ into a larger ground set $\mathsf{N} \supseteq \mathsf{L}$. The embedding can be described by a face $F = (\mathsf{L}|\mathsf{M}) \in \mathcal{F}^{\mathsf{N}}$ as

$$\mathcal{G}{\uparrow}F \coloneqq \{ (\mathsf{ij}|\mathsf{K}) \in \mathcal{A}_{\mathsf{N}} : (\mathsf{ij}|\mathsf{K}){\downarrow}F \in \mathcal{G} \} \subseteq \mathcal{A}_{\mathsf{N}}$$

where L is the ground set of $\mathcal{G}$, M is the set of unconditionings and $\mathsf{N} \setminus \mathsf{LM}$ is the set of unmarginalizations performed, where $\mathcal{G}{\downarrow}F$ denotes the minor of $\mathcal{G}$ given by face F.

Definition 4.19. Let $\mathcal{G}$ and $\mathcal{H}$ be CI structures on disjoint ground sets N and M, respectively. Their *dependent sum* is the union of sets $\mathcal{G} \sqcup \mathcal{H}$ as a CI structure on NM.

Embedding is the opposite operation of taking a minor. Mirroring the minor operation, embedding decomposes into unmarginalizations which add a new element to the ground set and do not modify the CI structure and unconditionings which add a new element to the ground set to every conditioning set in the CI structure. The embedding $\mathcal{G}{\uparrow}F$ is the most generic CI structure such that $\mathcal{G}$ is its F-minor — "generic" in the sense that no further CI relations hold except those necessitated by $\mathcal{G}$. Since CI relations are algebraic equations, this use of genericity is compatible with its use in algebraic geometry. Likewise, the dependent sum is the most generic CI structure joining its summands, which live on disjoint ground sets. The direct sum models combinatorially the joining of two random vectors in an independent manner, introducing no dependencies between the two vectors. The dependent sum does the opposite: it avoids any mixed independencies between the joined vectors.

Lemma 4.20. Let $\mathcal{G} = [\![\Sigma]\!] \in \mathfrak{g}^\bullet(\mathsf{N})$ and $\mathcal{H} = [\![\Gamma]\!] \in \mathfrak{g}^\bullet(\mathsf{M})$ with $\mathsf{N} \cap \mathsf{M} = \emptyset$. Then $\mathcal{G} \sqcup \mathcal{H}$ is realizable near $\Sigma \oplus \Gamma$ and hence belongs to $\mathfrak{g}^\bullet(\mathsf{NM})$.

Proof. Define an $(\mathsf{NM} \times \mathsf{NM})$-matrix

$$\Phi = \begin{array}{cc} \begin{array}{cc} \mathsf{N} & \mathsf{M} \end{array} & \\ \begin{pmatrix} \Sigma & \boldsymbol{\varepsilon} \\ \boldsymbol{\varepsilon}^{\mathsf{T}} & \Gamma \end{pmatrix} & \begin{array}{c} \mathsf{N} \\ \mathsf{M} \end{array} \end{array},$$

where $\boldsymbol{\varepsilon} = (\varepsilon_{\mathsf{ij}})_{\mathsf{ij} \in \mathsf{N} \times \mathsf{M}}$ consists of independent variables. Obviously Φ° is the aforementioned block-diagonal matrix $\Sigma \oplus \Gamma$ and the entries of Φ are polynomials over $\mathbb{K}$. Φ defines a realizable gaussoid $[\![\Phi]\!] = \mathcal{U}$ which restricts to $\mathcal{G}$ on N and to $\mathcal{H}$ on M. It remains to see that $\mathcal{U}$ contains no other CI statement $(\mathsf{ij}|\mathsf{K})$ where we decompose $\mathsf{K} = \mathsf{N}'\mathsf{M}'$ with $\mathsf{N}' \subseteq \mathsf{N}$, $\mathsf{M}' \subseteq \mathsf{M}$. We apply Lemma 4.1 and show that $\Phi[\mathsf{ij}|\mathsf{K}]$ vanishes in $\mathbb{K}(\varepsilon_{\mathsf{ij}})$ only when $\mathcal{G}$ or $\mathcal{H}$ mandate it.

First assume $\mathsf{ij} \subseteq \mathsf{N}$ and $\mathsf{M}' \neq \emptyset$ (as $\mathsf{M}' = \emptyset$ would have already $\Phi[\mathsf{ij}|\mathsf{K}] = \Sigma[\mathsf{ij}|\mathsf{N}']$ whose vanishing depends only on $\mathcal{G}$). Then:

$$\begin{aligned}
\Phi[\mathsf{ij}|\mathsf{K}] &= \det \begin{array}{cc} \begin{array}{ccc} \mathsf{j} & \mathsf{N}' & \mathsf{M}' \end{array} & \\ \begin{pmatrix} & & \varepsilon_{\mathsf{i},\mathsf{M}'} \\ & \Sigma_{\mathsf{ij}|\mathsf{N}'} & \varepsilon_{\mathsf{N}',\mathsf{M}'} \\ \varepsilon_{\mathsf{M}',\mathsf{j}} & \varepsilon_{\mathsf{M}',\mathsf{N}'} & \Gamma_{\mathsf{M}'} \end{pmatrix} & \begin{array}{c} \mathsf{i} \\ \mathsf{N}' \\ \mathsf{M}' \end{array} \end{array} \\
&= \Gamma[\mathsf{M}'] \det \left(\Sigma_{\mathsf{ij}|\mathsf{N}'} - \begin{pmatrix} \varepsilon_{\mathsf{i},\mathsf{M}'} \\ \varepsilon_{\mathsf{N}',\mathsf{M}'} \end{pmatrix} \Gamma_{\mathsf{M}'}^{-1} \begin{pmatrix} \varepsilon_{\mathsf{M}',\mathsf{j}} & \varepsilon_{\mathsf{M}',\mathsf{N}'} \end{pmatrix} \right).
\end{aligned}$$

The first determinant is a principal minor of Γ and hence non-zero. It suffices to show that the determinant of the Schur complement expression is not the zero polynomial. For row $\mathsf{a} \in \mathsf{iN}'$ and column $\mathsf{b} \in \mathsf{jN}'$ the corresponding entry of the Schur complement is, by the Leibniz formula,

$$\sigma_{\mathsf{ab}} - \sum_{\mathsf{k},\mathsf{l} \in \mathsf{M}'} (\Gamma_{\mathsf{M}'}^{-1})_{\mathsf{kl}} \varepsilon_{\mathsf{al}} \varepsilon_{\mathsf{bk}}$$

and hence the determinant equals

$$\sum_{\substack{\tau:\mathsf{iN'}\to\mathsf{jN'}\\ \text{bijective}}} \operatorname{sgn}(\tau) \prod_{\mathsf{a}\in\mathsf{iN'}} \left(\sigma_{\mathsf{a}\tau(\mathsf{a})} - \sum_{\mathsf{k},\mathsf{l}\in\mathsf{M'}} (\Gamma_{\mathsf{M'}}^{-1})_{\mathsf{kl}} \varepsilon_{\mathsf{al}} \varepsilon_{\tau(\mathsf{a})\mathsf{k}} \right).$$

By our assumption there exists $\mathsf{m} \in \mathsf{M'}$. We find that in this multivariate polynomial the coefficient of $\varepsilon_{\mathsf{im}}\varepsilon_{\mathsf{jm}}$ is

$$\pm \sum_{\tau\in\mathfrak{S}_{\mathsf{N'}}} \operatorname{sgn}(\tau)(\Gamma_{\mathsf{M'}}^{-1})_{\mathsf{mm}} \prod_{\mathsf{a}\in\mathsf{N'}} \Sigma_{\mathsf{a}\tau(\mathsf{a})} = \pm(\Gamma_{\mathsf{M'}}^{-1})_{\mathsf{mm}} \Sigma[\mathsf{N'}] \neq 0,$$

which shows that $(\mathsf{ij}|\mathsf{K}) \notin \mathcal{U}$ by Lemma 4.1 in case $\mathsf{ij} \subseteq \mathsf{N}$ and $\mathsf{M'} \neq \emptyset$. The same proof applies to the symmetric case when N and M are exchanged.

The remaining case has $\mathsf{i} \in \mathsf{N}$ and $\mathsf{j} \in \mathsf{M}$. Then, by Laplace expansion

$$\Phi[\mathsf{ij}|\mathsf{K}] = \det \begin{pmatrix} \varepsilon_{\mathsf{ij}} & \cdots \\ \vdots & \Phi_{\mathsf{K}} \end{pmatrix} = \varepsilon_{\mathsf{ij}} \Phi[\mathsf{K}] \mp \ldots,$$

where $\Phi[\mathsf{K}]$ is non-zero and all other terms do not involve the variable $\varepsilon_{\mathsf{ij}}$. □

Over an infinite field, the empty structure is realizable near the identity matrix: $\emptyset = [\![\mathbb{1}_{\mathsf{N}} + (\varepsilon_{\mathsf{ij}})_{\mathsf{i},\mathsf{j}\in\mathsf{N}}]\!] \in \mathfrak{g}^{\bullet}(\mathsf{N})$ with pairwise independent variables. In fact, the same reasoning shows that $\emptyset$ is realizable near *every* principally regular matrix Γ by considering $\Gamma + (\varepsilon_{\mathsf{ij}})_{\mathsf{i},\mathsf{j}\in\mathsf{N}}$.

Lemma 4.21. Let $\mathfrak{a}$ be a property of CI structures which contains the empty structure on every ground set and is closed under duality and dependent sum. Then it is closed under embeddings.

Proof. It suffices to treat unmarginalizations and unconditionings. Let $\mathcal{G} \in \mathfrak{a}(\mathsf{L})$ and $\mathsf{m} \notin \mathsf{L}$. Then $\mathcal{G} \in \mathfrak{a}(\mathsf{mL})$ follows directly from Lemma 4.20 and the realizability of the empty set over m. For unconditioning, we make use of duality:

$$(\mathsf{ij}|\mathsf{K}) \in \mathcal{G} \;\Leftrightarrow\; (\mathsf{ij}|\mathsf{N}^{\mathsf{ij}} \setminus \mathsf{K}) \in \mathcal{G}^{\rceil} \;\Leftrightarrow\; (\mathsf{ij}|(\mathsf{mN})^{\mathsf{ij}} \setminus \mathsf{mK}) \in \mathcal{G}' := \mathcal{G}^{\rceil} \subseteq \mathcal{A}_{\mathsf{mL}} \;\Leftrightarrow\; (\mathsf{ij}|\mathsf{mK}) \in \mathcal{G}'^{\rceil}.$$

The structure $\mathcal{G}'^{\rceil}$ arises from operations on $\mathcal{G}$ under which $\mathfrak{a}$ is closed by the hypotheses and the first part of the proof, thus $\mathcal{G}'^{\rceil} \in \mathfrak{a}(\mathsf{mL})$. No other CI statements over mL arise because duality and embedding preserve also cardinality, so $\mathcal{G}'^{\rceil}$ is indeed the unconditioning of $\mathcal{G}$ with respect to m. □

Remark 4.22. The two primitive operations yielding Lemma 4.21 are dependent sum with the empty structure and duality. These operations also transport realizability near chosen principally regular matrices. It is easy to see that an embedding of $[\![\Sigma]\!]$ is realizable near $\Sigma \oplus \Gamma$ for every principally regular Γ.

Remark 4.23. Closedness under dependent sum and existence of the empty structure allows to view every $\mathcal{A} \in \mathfrak{a}(\mathsf{L})$ as a structure in $\mathfrak{a}(\mathsf{N})$, for every $\mathsf{N} \supseteq \mathsf{L}$; in other words, there is an inclusion $\mathfrak{a}(\mathsf{L}) \hookrightarrow \mathfrak{a}(\mathsf{N})$ given by unmarginalization. Under the additional hypothesis of duality-closedness, this is only one of many inclusions which are the embeddings of CI structures given by the choice of an $|\mathsf{L}|$-face in the N-cube. All of these inclusions preserve the cardinality of the independence structure. Swapping in this construction the empty for the full structure and the dependent for the direct sum, one obtains a complementary set of inclusions $\mathfrak{a}(\mathsf{L}) \hookrightarrow \mathfrak{a}(\mathsf{N})$ which preserve the cardinality of the corresponding *dependence structures* $\mathfrak{d}(\mathsf{L}) = \{\, \mathcal{A}_{\mathsf{L}} \setminus \mathcal{A} : \mathcal{A} \in \mathfrak{a}(\mathsf{L}) \,\}$.

Definition 4.24. For $\mathcal{A} \subseteq \mathcal{A}_\mathsf{N}$ and $F \in \mathcal{F}_k^\mathsf{N}$ the *isolation* $\mathcal{A}\restriction_F$ *of* $\mathcal{A}$ *to* F consists of all elements of $\mathcal{A}$ which lie on F. If $k < 2$, then $\mathcal{A}\restriction_F = \emptyset$. A property $\mathfrak{a}$ of CI structures is *isolation-closed* if for every ground set N and $\mathcal{A} \in \mathfrak{a}(\mathsf{N})$ all $\mathcal{A}\restriction_F$, for $F \in \mathcal{F}^\mathsf{N}$, satisfy $\mathfrak{a}$.

The operation of isolation to F mimics the effect of taking the minor with respect to F and then embedding the result back into F. Thus it is a form of mixed marginalization and conditioning which does not reduce the ground set, but instead turns any independencies outside of the face into dependencies.

Lemma 4.25. Let $\mathfrak{a}$ be a property of CI structures which is closed under minors and embeddings. Then $\mathfrak{a}$ is closed under isolations. □

We summarize the developments of this and the previous sections in:

Theorem 4.26. Over an infinite or ordered field, $\mathfrak{g}^\bullet$ is closed under isomorphy, duality, minors, direct and dependent sums, embeddings and isolation. □

As shown in Section 3.3, closedness under isomorphy and duality (and, in the algebraic case, the entire hyperoctahedral symmetry), minors and direct sums holds for finite fields as well. The remaining properties were derived for infinite fields based on appeals to genericity, mainly using Lemma 4.1. This lemma and its consequences do not hold in general for finite fields. One easily shows that the only principally regular matrix over $\mathbb{F}_2$ is $\mathbb{1}_\mathsf{N}$. Thus $\mathfrak{g}^*_{\mathbb{F}_2}(\mathsf{N}) = \{ \mathcal{A}_\mathsf{N} \}$ and this property is not closed under embeddings, dependent sums or isolation.

4.4 About finite axiomatizability

The following sections focus on the the axioms, the outer description, of the properties $\mathfrak{g}^\bullet$. In particular, we first wish to prove a lower complexity bound on the structure of these axioms: there is no finite complete set of axioms from which all the axioms follow, as the ground set size grows. The term "finite axiomatization" is used in the literature of CI structures to mean a number of different things. Works of Matúš [Mat97] and Šimeček [Šim06a] take the approach which is also popular in matroid theory: a finite axiomatization of a property is a finite list of forbidden minors.

Definition 4.27. A property $\mathfrak{f}$ of CI structures has a *finite set of forbidden minors in dimension* k if it is minor-closed and if for all ground sets N with $|\mathsf{N}| \geq k$ it holds that $\mathcal{A} \subseteq \mathcal{A}_\mathsf{N}$ belongs to $\mathfrak{f}(\mathsf{N})$ if and only if all k-minors of $\mathcal{A}$ belong to $\mathfrak{f}(\mathsf{K})$.

The assumption of minor-closedness guarantees **soundness** of the forbidden minor characterization, that is, when a structure does not appear in $\mathfrak{p}(\mathsf{N})$, it can be labeled as forbidden and must never appear in a structure from $\mathfrak{p}(\mathsf{M})$, $\mathsf{M} \supseteq \mathsf{N}$. The second condition ensures that for every structure not having property $\mathfrak{p}$, the reason for that is found in the presence of a forbidden minor; thus it is the **completeness** of the characterization.

Matúš writes a justification for this point of view which appeals to the applications of CI inference in expert systems [Mat97, Section 7] and emphasizes the role of minors as natural subconfigurations of CI structures:

> It is a nice archievement [sic!] if a set of the features is melted into an easily understandable axiom on CI-relations [...] But, we find it not natural if a class of CI-relations pretending to model conditional thinking of a human expert in a restricted domain is defined by a greater number of axioms. A greater number of small configurations seems to be, if unavoidable, more acceptable.

On the other hand, Studený [Stu92, SV98] and Sullivant [Sul09] prove theorems about the necessity of appearance of the great number of axioms mentioned by Matúš. Their statements take a form such as: "there exist valid inference rules with arbitrarily many antecedents which are not implied by shorter rules". It follows that the property cannot be described by a finite set of inference rules, since their number of antecedents needs to be bounded.

There are seemingly two issues with the inference rule approach to the notion of finite axiomatizability. The purpose of the present section is to clarify these issues and resolve them, leading to a definition of finite axiomatizability which, under mild assumptions about the property, is equivalent to the forbidden-minor approach. The ingredients to that can be found already, for discrete random variables, in the literature.

The first issue is that there is a distinction to be made between inference rules such as $(12|) \wedge (13|2) \Rightarrow (13|)$ over the ground set $\mathsf{N} = 123$ and inference rule **schemes** such as $(\mathsf{ij}|) \wedge (\mathsf{ik}|\mathsf{j}) \Rightarrow (\mathsf{ik}|)$ for all distinct $\mathsf{i}, \mathsf{j}, \mathsf{k} \in \mathsf{N}$ for all N, of which the former is a specific instance. A result about a property $\mathfrak{p}$, which is an infinite object indexed by all finite ground sets N, not having a description by means of finitely many inference rules (in the former sense) would not be surprising. The theorems of Studený and Sullivant concern instead the latter inference schemes. Such schemes generate, as N grows, an **unbounded** number of inference rules, but still, because the number of schemes is finite and the number of antecedents in each instance is fixed by the schemes, theorems of the kind that Studený and Sullivant proved imply that CI structures appear on large enough N which satisfy all instances of any finite number of valid schemes but not have the property $\mathfrak{p}$ in question. This definition of finite axiomatization based on inference rule schemes was given by Studený under the name *quasi-axiomatic characterization* [Stu05, p. 51 f.]. See also the careful presentation of the axiomatic framework behind Studený's non-axiomatizability result for discrete CI [Stu92] in the work [SV98]. A considerably stronger logical framework — the *monadic second-order logic for matroids* — was investigated in matroid theory in [MNW14, MNW18] to rectify technical shortcomings in the famous work of Vámos [Vá78]. Developing such a theory for conditional independence is beyond the scope of this chapter. Hannula, Kontinen and collaborators work on related topics in the guise of probabilistic team semantics in model theory; see e.g., [HK16] and [HHK+19].

In light of the structure theory developed so far, we find it more natural to think about inference rule schemes not as formulas in a logical language, as in Studený's quasi-axiomatic approach, but as specific inference rules up to a number of transformations.

Lemma 4.28. Let N be a fixed ground set and $\varphi : \bigwedge \mathcal{L} \Rightarrow \bigvee \mathcal{M}$ be an inference formula over N. If $f : \mathcal{A}_\mathsf{N} \to \mathcal{A}_\mathsf{M}$ is any injection which preserves $\mathfrak{p}$ in both directions, then φ is valid for $\mathfrak{p}(\mathsf{N})$ if and only if $f(\varphi) : \bigwedge f(\mathcal{L}) \Rightarrow \bigvee f(\mathcal{M})$ (applied element-wise) is valid for $\mathfrak{p}(\mathsf{M})$.

Proof. Suppose there exists $\mathcal{N} \in \mathfrak{p}(\mathsf{N})$ violating φ. Thus $\mathcal{N}$ contains $\mathcal{L}$ but is disjoint from $\mathcal{M}$. Its image under f is a CI structure over M which contains $f(\mathcal{L})$ but is disjoint from $f(\mathcal{M})$ and hence a counterexample to the validity of $f(\varphi)$ for $\mathfrak{p}(\mathsf{M})$. These arguments can be reversed to prove the other direction of the if-and-only-if claim. $\square$

The closure properties proved in Theorem 4.26 give rise to such operations where the image and preimage of every Gaussian is a Gaussian over the respective ground set. For example, each N-face $(\mathsf{I}|\mathsf{K})$ of the M-cube gives rise to one (injective) embedding map $\iota_{(\mathsf{I}|\mathsf{K})}$ which preserves realizability. The statement that realizability is also transferred from the image to the preimage is minor-closedness. With these embedding maps and the $\mathfrak{S}_\mathsf{M}$ action, Lemma 4.28 shows that the gaussoid axioms on 3×3 principally regular matrices, which are trivial to prove algebraically, imply that the gaussoid axioms hold for principally regular matrices of all sizes.

Example 4.29. Consider the semigraphoid axiom and the $\mathfrak{g}^\bullet$-preserving maps,

$$
\begin{array}{rcl}
 & (12|) \wedge (13|2) \Rightarrow (13|) & \text{on } 123, \\
\iota_{(123\mathsf{M}|\mathsf{L})} \Rightarrow & (12|\mathsf{L}) \wedge (13|2\mathsf{L}) \Rightarrow (13|\mathsf{L}) & \text{on } 123\mathsf{LM}, \\
\mathfrak{S}_{123\mathsf{M}} \Rightarrow & (\mathsf{ij}|\mathsf{L}) \wedge (\mathsf{ik}|\mathsf{jL}) \Rightarrow (\mathsf{ik}|\mathsf{L}) & \text{on } 123\mathsf{LM}.
\end{array}
$$

This proof of the (first half of the) general semigraphoid scheme does not require confirming Lemma 3.5 for arbitrarily large symmetric matrices, as our original proof of the gaussoid axioms did. Instead, a calculation on 3×3 matrices suffices, together with universal closure properties of Gaussians. △

The second, more subtle, issue with inference rule (scheme) axiomatizations of an entire infinitary property $\mathfrak{p}$ is that the inferences are often stated without reference to a ground set. To reflect the practice in the literature, for example [Stu92] and [Sul09], we wish to impose a regularity condition on properties $\mathfrak{p}$ which allows the definition of "finitely axiomatizable" to disregard the ground set which should formally be attached to an axiom. The embeddings $\iota_{(\mathsf{N}|)}$ in particular allows us to view every inference rule (scheme) which is valid over N as a scheme over $\mathsf{M} \supseteq \mathsf{N}$. This justifies the practice of leaving the ground set out, implying that the smallest possible ground set should be assumed, but it has to be proved that the interpretation of an inference formula φ which is valid for $\mathfrak{p}(\mathsf{N})$ over any $\mathsf{M} \supseteq \mathsf{N}$ never becomes invalid for $\mathfrak{p}$. Lemma 4.28 shows that this is indeed the case, provided that $\mathfrak{p}$ has certain closure properties. In particular, the closedness under marginalization and unmarginalization is sufficient.

Definition 4.30. A valid inference rule $\varphi : \bigwedge \mathcal{L} \Rightarrow \bigvee \mathcal{M}$ over N for a property $\mathfrak{p}$ is *minimal* if removing one element from either $\mathcal{L}$ or $\mathcal{M}$ results in an invalid inference formula. The *minimal axioms* $\mathfrak{A}(\mathfrak{p})$ of $\mathfrak{p}$ consist of all minimal valid inference rules for $\mathfrak{p}$ stratified by ground set.

Definition 4.31. The *context* of a CI statement $(\mathsf{ij}|\mathsf{K})$ is the set $[(\mathsf{ij}|\mathsf{K})] \coloneqq \mathsf{ijK}$. This definition is extended to CI structures and to inference formulas. For a formula $\varphi : \bigwedge \mathcal{L} \Rightarrow \bigvee \mathcal{M}$ the *antecedental* and *consequential* contexts, $[\mathcal{L}]$ and $[\mathcal{M}]$, respectively, are distinguished. The formula φ is *context-preserving* if $[\mathcal{M}] \subseteq [\mathcal{L}]$. A property $\mathfrak{p}$ is *context-complete* if every minimal axiom $\varphi \in \mathfrak{A}(\mathfrak{p})$ is context-preserving and its unmarginalization to any $\mathsf{N} \supseteq [\varphi]$ remains valid and minimal.

The context of an inference rule is the smallest ground set over which it can be stated. Context-completeness requires the minimal valid inference rules to be recognizable over their contexts. In particular, every element of N appearing in the consequents of a minimal valid inference rule must appear in the antecedents. Its minimal valid inference rules are sufficient to characterize a property $\mathfrak{p}$. Context-completeness allows for a stratified enumeration of $\mathfrak{A}(\mathfrak{p})$ by (context) ground sets: first find the minimal axioms for $n = 1$, then for $n = 2$ and so on. All of the found axioms are required to rule out certain forbidden minors on their respective context, and they never come invalid.

Clearly, non-minimal valid inference rules can fail to be context-preserving even for properties which are context-complete. Consider the valid and context-preserving semigraphoid instance $(12|) \wedge (13|2) \Rightarrow (13|)$ for algebraic Gaussians (which are context-complete by Corollary 4.33). Adding (disjunctively) a redundant consequent $(45|678)$ produces a valid inference rule which violates context preservation.

Lemma 4.32. A property $\mathfrak{p}$ which is closed under marginalization and unmarginalization and contains $\mathcal{A}_\mathsf{N} \in \mathfrak{p}(\mathsf{N})$ is context-complete.

Proof. Let $\varphi : \bigwedge \mathcal{L} \Rightarrow \bigvee \mathcal{M}$ be a minimal axiom for $\mathfrak{p}$. Suppose that $[\mathcal{M}] \not\subseteq [\mathcal{L}]$. Since $\mathcal{A}_{[\mathcal{L}]} \in \mathfrak{p}([\mathcal{L}])$, unmarginalize it to $\mathsf{M} = [\varphi]$. This structure contains $\mathcal{L}$ but does not contain every consequent of M, contradicting either the validity of φ or its minimality. That any unmarginalization to a set $\mathsf{N} \supseteq [\varphi]$ remains valid and minimal follows from Lemma 4.28. □

Corollary 4.33. Algebraic (positive) Gaussians over infinite (ordered) fields are context-complete. □

Remark 4.34. In the preparatory work for his main result in [Stu92], Studený proves in Lemma 2 that the property $\mathfrak{d}$ of being realizable by discrete random variables is preserved under intersections of CI structures. This result together with other constructions in Lemmas 3 and 4 implies that $\mathfrak{d}$ is context-complete.

Combining the two strands of the present section, finally leads to the following definition of finite axiomatizability. Inference rule schemes are formalized as inference rules over a finite ground set N. The instantiation of a scheme over a larger ground set M corresponds to applying the validity-preserving embedding operation. By context-completeness, the validity of any such image of an inference formula depends only on the validity over its context. This process generates a countably infinite set of inference rules from a finite set of axiom schemes and any CI structure which satisfies all of them, disregarding ground sets, has the property $\mathfrak{p}$ and vice versa.

Definition 4.35. A context-complete property $\mathfrak{p}$ is *finitely axiomatizable* if there is a finite set N such that $\mathfrak{A}(\mathfrak{p}(\mathsf{N}))$ generates by logical deduction $\mathfrak{A}(\mathfrak{p}(\mathsf{M}))$ for all $\mathsf{M} \supseteq \mathsf{N}$ under the embeddings $\iota_{(\mathsf{I}|\mathsf{K})} : \mathcal{A}_\mathsf{N} \to \mathcal{A}_\mathsf{M}$ in the sense of Lemma 4.28.

Proposition 4.36. Let $\mathcal{L}_n$ be a sequence of CI structures which is defined for sufficiently large n. Suppose that $|\mathcal{L}_n| > n$ and that $\mathcal{L}_n$ does not have the property $\mathfrak{p}$. If additionally all n-subsets $\mathcal{L}' \subsetneq \mathcal{L}_n$ have property $\mathfrak{p}$, then $\mathfrak{p}$ has no finite axiomatization.

Proof. Suppose $\varphi_1, \dots, \varphi_s$ is a finite axiomatization of $\mathfrak{p}$ and pick n large enough so that all φ_i have at most n antecedents. Then $\mathcal{L}_n$ fulfills each φ_i:

(a) If $\mathcal{L}_n$ does not contain the antecedents of φ_i, then $\mathcal{L}_n \models \varphi_i$ vacuously.
(b) If $\mathcal{L}_n$ contains the antecedents of φ_i, then there is a strict subset $\mathcal{L}'$ of size at most n which contains the antecedents. By assumption $\mathcal{L}'$ has $\mathfrak{p}$ and thus satisfies φ_i. But φ_i is an **inference** rule and since $\mathcal{L}' \subseteq \mathcal{L}$, we necessarily have $\mathcal{L} \models \varphi_i$.

Thus $\mathcal{L}_n \models \varphi_i$ but $\mathcal{L}_n \notin \mathfrak{p}$ by hypothesis, contradicting finite axiomatization. □

Remark 4.37. The proof of Studený for the non-axiomatizability of discrete CI structures in [Stu92] uses Proposition 4.36 as explained in [SV98, Section 5.2.2]. Sullivant's paper about positive real Gaussians [Sul09] fails to show that strict subsets of a chosen size of the antecedents are realizable. Example 4.38 below shows that Proposition 4.36 cannot be strengthened to obtain non-axiomatizability from just the construction of arbitrarily long minimal axioms.

Example 4.38. This example gives an infinite family of minimally valid inference rules with arbitrarily many antecedents for the semigraphoid property. Since semigraphoids have a finite axiomatization, this proves that finding long minimal axioms does not suffice for a non-axiomatizability proof. Consider the sequence of formulas:

$$\begin{aligned} (12|) \wedge (13|2) &\Rightarrow (12|3), \\ (12|) \wedge (13|2) \wedge (14|23) &\Rightarrow (12|34), \\ (12|) \wedge (13|2) \wedge (14|23) \wedge (15|234) &\Rightarrow (12|345), \\ &\vdots \end{aligned}$$

The first formula is a semigraphoid axiom and each successive formula reduces to a semigraphoid axiom modulo the previous formula. Thus, all formulas are valid for semigraphoids. It is easy to see inductively that these inference rules are minimally: in the rule with $n+1$ antecedents, removing any of the first n makes the previous rule invalid; removing the $(n+1)^{\text{st}}$ makes the rule invalid because of context-completeness of the semigraphoid property. △

In particular, the construction in Section 4.2 does not constitute a proof of non-axiomatizability. This result is proved in the next section based on the following

Theorem 4.39. Let $\mathfrak{p}$ be a minor- and embedding-closed property (in particular context-complete). Then $\mathfrak{p}$ has a finite forbidden-minor characterization if and only if it has a finite axiomatization.

Proof. Let $F_1, \dots, F_t$ be a complete set of forbidden minors for $\mathfrak{p}$ defined over some common finite ground set N. Thus, $F_1, \dots, F_t$ are the non-$\mathfrak{p}$ structures in $\mathcal{A}_{\mathsf{N}}$. Their complement $\mathfrak{p}(\mathsf{N})$ is defined by a boolean formula in conjunctive normal form whose clauses $\varphi_1, \dots, \varphi_s$ are interpreted as inference rules. Clearly, a CI structure $\mathcal{N}$ over $\mathsf{M} \supseteq \mathsf{N}$ does not have any of the forbidden minors in its N-face $(\mathsf{I}|\mathsf{K})$ if and only if $\iota_{(\mathsf{I}|\mathsf{K})}(\varphi_i)$, $i \in [s]$, hold for $\mathcal{N}$. Thus, the φ_i are a finite axiomatization.

Conversely, suppose a finite axiomatization $\varphi_1, \dots, \varphi_s$ is given. Let $\mathsf{N} = \bigcup_i [\varphi_i]$ denote their common context. Let $F_1, \dots, F_t$ be the non-$\mathfrak{p}$ structures in $\mathcal{A}_{\mathsf{N}}$. We claim that this the required finite axiomatization. Let $\mathcal{N} \subseteq \mathcal{A}_{\mathsf{M}}$, $\mathsf{M} \supseteq \mathsf{N}$, be given. By context-completeness, $\mathcal{N}$ satisfies the axioms $\iota_{(\mathsf{I}|\mathsf{K})}(\varphi_i)$ if and only if $\mathcal{N}{\downarrow}(\mathsf{I}|\mathsf{K})$ is not a forbidden minor. Thus $F_1, \dots, F_t$ is a forbidden-minor characterization. □

4.5 Infinitely many forbidden minors over infinite fields

In this section, the above closure properties are applied to derive a clear combinatorial obstruction to the existence of a finite forbidden-minor characterization of algebraic and positive Gaussians over infinite fields. This was proved by Šimeček for positive real Gaussians in [Šim06a]. Indeed, he gives forbidden minors which are non-realizable over $\mathbb{R}$ and whose proper minors are rationally realizable near the identity matrix. Consequently, the same proof works for all ordered fields. But his family of forbidden minors is only an obstruction for positive realizability:

Example 4.40. Let $\mathcal{S}_n := \{(\mathsf{01}|\mathsf{2}), (\mathsf{02}|\mathsf{3}), \dots, (\mathsf{0n}|\mathsf{1})\} \subseteq \mathcal{A}_{\mathsf{01\cdots n}}$, or in other words, $\mathcal{S}_n$ is generated by the CI statement $(\mathsf{01}|\mathsf{2})$ under the n-cycle $\tau = (\mathsf{1}\ \mathsf{2}\ \mathsf{3}\ \cdots\ \mathsf{n})$ with fixed point $\mathsf{0}$. This is the same family that Studený used in [Stu92] to prove the non-axiomatizability of discrete CI. Supposing an algebraically closed field, we can assume that a realization, if it exists, has a unit diagonal and off-diagonal entries x_{ij}. Then each CI statement is equivalent to a variable substitution

$$(\mathsf{0i}|\mathsf{j}) \;\Leftrightarrow\; x_{\mathsf{0i}} = x_{\mathsf{0j}} x_{\mathsf{ij}}. \qquad (*)$$

The circularity of $\mathcal{S}_n$ leads to a circular substitution and eventually to the equalities

$$x_{\mathsf{0i}} = x_{\mathsf{0i}} \prod_{k=1}^{n-1} x_{\tau^k(\mathsf{12})}, \quad \mathsf{i} \in [\mathsf{n}],$$

which cannot be satisfied in a positive-definite matrix, as all off-diagonal entries must be strictly smaller than 1 in absolute value, unless of course $x_{\mathsf{0i}} = 0$.

The realizability of $\mathcal{S}_n$ hinges on the equalities $(*)$ and the implied extra condition $\prod_k x_{\tau^k(12)} = 1$. Over an algebraically closed (and hence infinite) field, these equations are easily satisfied by a principally regular matrix over a suitable function field such as

$$
\begin{array}{c}
\begin{array}{cccccc} \mathsf{0} & \mathsf{1} & \mathsf{2} & \mathsf{3} & \cdots & \mathsf{n} \end{array} \\
\left(\begin{array}{c|ccccc}
1 & y_1 & y_2 & y_3 & \cdots & y_n \\ \hline
y_1 & 1 & y_1/y_2 & z_{13} & \cdots & y_n/y_1 \\
y_2 & y_1/y_2 & 1 & y_2/y_3 & \cdots & z_{2n} \\
y_3 & z_{13} & y_2/y_3 & 1 & \cdots & z_{3n} \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
y_n & y_n/y_1 & z_{2n} & z_{3n} & \cdots & 1
\end{array}\right)
\begin{array}{c} \mathsf{0} \\ \mathsf{1} \\ \mathsf{2} \\ \mathsf{3} \\ \vdots \\ \mathsf{n} \end{array}
\end{array}.
$$

It seems plausible that this matrix is principally regular and realizes $\mathcal{S}_n$ over every algebraically closed field. For the purpose of motivating this section, it suffices to confirm this by computation over $\mathbb{C}$ for $n = 4, 5, 6$:

```
-- CI statements, principal and almost-principal minors
CIstmts = N -> flatten(subsets(N, 2) / (ij ->
  subsets(set(N) - ij) / (L -> (ij#0, ij#1, toList(L)))
));
pr  = (X, K) -> det X_K^K;
apr = (X, ijK) -> (
  I := flatten({ijK#0, ijK#2});
  J := flatten({ijK#1, ijK#2});
  det X_I^J
);
CI = X -> (
  N := toList(0 .. numRows(X)-1);
  if 0 != X - transpose(X) then
  then error "X is not symmetric";
  if select(subsets(N), K -> pr(X, K) == 0) != {}
  then error "X is not principally regular";
  select(CIstmts(N), ijK -> apr(X, ijK) == 0)
);

-- S_4:
R = QQ[y1,y2,y3];
X = matrix{{1,y1,y2,y3}, {y1,1,y1/y2,y3/y1},
     {y2,y1/y2,1,y2/y3}, {y3,y3/y1,y2/y3,1}};
CI X --> { (01|2), (02|3), (03|1) }

-- S_5:
R = QQ[y1,y2,y3,y4, z13,z24];
X = matrix{{1,y1,y2,y3,y4}, {y1,1,y1/y2,z13,y4/y1},
  {y2,y1/y2,1,y2/y3,z24}, {y3,z13,y2/y3,1,y3/y4},{y4,y4/y1,z24,y3/y4,1}};
CI X --> { (01|2), (02|3), (03|4), (04|1) }

--  S_6:
R = QQ[y1,y2,y3,y4,y5, z13,z14,z24,z25,z34,z35];
X = matrix{{1,y1,y2,y3,y4,y5}, {y1,1,y1/y2,z13,z14,y5/y1},
   {y2,y1/y2,1,y2/y3,z24,z25}, {y3,z13,y2/y3,1,y3/y4,z35},
   {y4,z14,z24,y3/y4,1,y4/y5}, {y5,y5/y1,z25,z35,y4/y5,1}};
CI X --> { (01|2), (02|3), (03|4), (04|5), (05|1) }
```
$\triangle$

Hence, this family cannot be reused to obtain infinitely many forbidden minors for algebraic realizability. The proof technique for this section was developed in [BK20] to give a doubly exponential lower bound on the number of gaussoids, more precisely the bound $2^{\Theta(n2^n)}$, but applies to a wider range of properties: namely those which have a finite forbidden-minor characterization and are closed under embedding and isolation. Assuming that there are at least two compulsory minors, the lower bound obtained is of the form $2^{\Theta(2^n/f(n))}$ where f is a polynomial in n which depends exponentially on the ground set size of the forbidden minors but is polynomial in n. Then, a result in matroid theory implies that the number of elements of $\mathfrak{g}^\bullet$ has a single exponential upper bound in n. Therefore, this property which is embedding- and isolation-closed cannot also have a finite forbidden-minor characterization.

The double exponential construction takes disjoint unions — not dependent sums — of embeddings of compulsory minors. Such unions need not belong to the class described by its forbidden minors. A valid inference rule with multiple antecedents can easily be violated by a disjoint union of singleton structures, even though the singletons themselves are all realizable. Thus, the proof requires a detailed understanding of the incidence structure of low-dimensional faces in hypercubes. This is encoded in the following undirected graph:

Definition 4.41. Let $Q(n, k, p, q)$, for $n \geq k \geq p \geq q$, be the undirected simple graph with vertex set $\mathcal{F}_k^n$ and an edge between $D, F \in \mathcal{F}_k^n$ if and only if there is a p-face S such that $\dim(D \cap S) \geq q$ and $\dim(F \cap S) \geq q$.

In this section, we use the convention that N is a set of cardinality n and K a subset of cardinality k. The main theorem about $Q(n, k, p, q)$ was proved in [BK20]:

Theorem 4.42. The graph $Q(n, k, p, q)$ is transitive, hence regular. It is complete if and only if $n + q \leq p + k$. The degree of any vertex can be calculated as follows:

$$\deg Q(n, k, p, q) = -1 + \sum_{m,j\ (\dagger)} \binom{k}{j} 2^{k-j} \binom{n-k}{k-j} \binom{n-2k+j}{m},$$

where the sum extends over $0 \leq m \leq n-k$ and $0 \leq j \leq k$ which satisfy the feasibility and connectivity conditions

$$n - 2k + j \geq m \quad \wedge \quad p \geq m + 2q - \min\{q, j\}. \tag{$\dagger$}$$

Lemma 4.43. Let $\mathfrak{f}$ be a property which is defined by a finite set of forbidden minors in dimension k and suppose that it is closed under embedding and isolation. Fix an independent set $\mathcal{F}$ in $Q(n, k, k, 2)$. For every function $\mathcal{F} \to \mathfrak{f}(\mathsf{K})$ there is a unique element of $\mathfrak{f}(\mathsf{N})$.

Proof. The proof is analogous to [BK20, Proposition 3.9]. We prove that for each assignment $\alpha : \mathcal{F} \to \mathfrak{f}(\mathsf{K})$ there is a unique CI structure in $\mathcal{A}_\mathsf{N}$ defined by α all whose k-minors belong to $\mathfrak{f}(\mathsf{K})$. Define $\mathcal{A}_\alpha := \bigsqcup_{D \in \mathcal{F}} \alpha(D){\uparrow}D \subseteq \mathcal{A}_\mathsf{N}$. By independence of $\mathcal{F}$, this union is disjoint and $\mathcal{A}_\alpha$ is a well-defined subset of $\mathcal{A}_\mathsf{N}$. Clearly, the map $\alpha \mapsto \mathcal{A}_\alpha$ is injective.

It remains to see that $\mathcal{A}_\alpha$ belongs to the set of structures generated by $\mathfrak{f}(\mathsf{K})$. Pick any k-minor $\mathcal{A}_\alpha{\downarrow}D$. If $D \in \mathcal{F}$, then the minor is $\alpha(D) \in \mathfrak{f}(\mathsf{K})$ by construction. Otherwise it may overlap with a face from $\mathcal{F}$ in $2 \leq l \leq k$ dimensions. Since $\mathcal{F}$ is independent in $Q(n, k, k, 2)$ there is at most one such face F. If it exists, this k-minor is an isolation $\alpha(D){\upharpoonright}_{D \cap F}$, which belongs to $\mathfrak{f}(\mathsf{K})$ by assumption. If D shares only at most one-dimensional faces with every element of $\mathcal{F}$, then $\mathcal{A}_\alpha{\downarrow}D$ is empty and hence belongs to $\mathfrak{f}(\mathsf{K})$ as a special case of isolation-closedness. ☐

Lemma 4.44. For sufficiently big n, the vertex degree of $Q(n, k, k, 2)$ is a polynomial in n of degree $2k - 4$.

Proof. By Theorem 4.42, the degree is exactly

$$-1 + \sum_{j,m} \binom{k}{j} 2^{k-j} \binom{n-k}{k-j} \binom{n-2k+j}{m},$$

where the sum extends over all $0 \le j \le k$ and $0 \le m \le n - 2k + j$ which also satisfy $m \le k - 4 + \min\{2, j\}$. When n is sufficiently large, then the second upper bound on m dominates the former. The summands in the degree formula are polynomials in n of degree $k - j + m$. This quantity is bounded by $2k - 4 + \min\{2, j\} - j$ which is maximized for $j \le 2$ giving the degree $2k - 4$. The leading coefficient of each polynomial summand is $\frac{1}{(k-j)!m!}\binom{k}{j}2^{k-j} > 0$, so no degree-reducing cancellations occur. □

By bounding the degree of the regular graph $Q(n, k, k, 2)$ from above, we obtain a good lower bound for its independence number. Clearly, a proper coloring of the graph with $\deg Q(n, k, k, 2)$ colors exists and each color class is an independent set. There must exist a color class of size at least that of an average color class, which is

$$\frac{|\mathcal{F}_k^n|}{\deg Q(n,k,k,2)} \in \Omega_k\left(\frac{n^k 2^{n-k}}{n^{2k-4}}\right) = \Omega_k\left(n^{4-k}2^n\right),$$

where Ω_k hides a factor depending on k (which is the ground set size on which the supposed forbidden-minor characterization is defined) but not on n.

Proposition 4.45. Let $\mathfrak{f}$ have a finite set of forbidden minors in dimension k with at least two compulsory minors and suppose that $\mathfrak{f}$ is closed under embedding and isolation. Then the asymptotic number of structures admitted by $\mathfrak{f}$ is at least

$$\log|\mathfrak{f}(\mathsf{N})| \in \Omega_k(n^{4-k}2^n),$$

which is double exponential in n. □

The proposition needs only two compulsory minors. In fact, Theorem 4.6 implies that $\mathfrak{g}^\bullet(\mathsf{K})$ has at least $2^{\binom{n}{2}}$ elements, but this only improves the polynomial part of the bound from n^{4-k} to n^{2-k}.

On the other hand, a recent result by Nelson [Nel18] about the asymptotic number of linear matroids provides a universal upper bound on the number of algebraically (and hence also positively) realizable gaussoids over any field — that is, the **union** of algebraic Gaussians over all fields, not only a fixed field.

Proposition 4.46. The number of gaussoids which are algebraically realizable over any field is bounded as $\log|\mathfrak{g}^\bullet(\mathsf{N})| \in \mathcal{O}(n^3)$.

Proof. By [Nel18, Theorem 1.1] the number of linear matroids on an n-element ground set over any field is at most $2^{\mathcal{O}(n^3)}$. The reduction of algebraic Gaussians to linear matroids has already been sketched in [BK20, Remark 3.11]. It is an easy observation for geometers who are familiar with the *Lagrangian Grassmannian*.

Let $\mathcal{G} = [\![\Gamma]\!]$ for an $\mathsf{N} \times \mathsf{N}$ matrix Γ over $\mathbb{K}$. Let N^* be a disjoint copy of N generated by the involution $\cdot^*$: $\mathsf{i}^* \in \mathsf{N}^* \Leftrightarrow \mathsf{i} \in \mathsf{N}$ and $\mathsf{i}^{**} = \mathsf{i}$. Consider the $\mathsf{N} \times \mathsf{N}\mathsf{N}^*$ matrix

$$\Phi = \begin{array}{cc} \begin{array}{cc} \mathsf{N} & \mathsf{N}^* \end{array} & \\ \left(\mathbb{1}_{\mathsf{N}} \;\middle|\; \Gamma\right) & \mathsf{N} \end{array}.$$

The columns of this matrix determine a vector matroid of rank n. For any $(\mathsf{ij}|\mathsf{K}) \in \mathcal{A}_{\mathsf{N}}$, pick the set $\mathsf{L} = (\mathsf{iK})^{\mathsf{c}} \cup (\mathsf{jK})^* \subseteq \mathsf{N}\mathsf{N}^*$ and a Laplace expansion of the unit columns shows that

$$\det \Phi_{\mathsf{N},\mathsf{L}} = \pm\Gamma[\mathsf{ij}|\mathsf{K}].$$

This proves that the **matroid** of the image matrix of the map $\Gamma \mapsto \Phi$ uniquely identifies the **gaussoid** of the preimage. Indeed $(\mathsf{ij}|\mathsf{K})$ holds for Γ if and only if L is a non-basis of Φ. Hence this is an injection of the N-gaussoids which are algebraically realizable over $\mathbb{K}$ into the linear rank-n matroids over $\mathbb{K}$ on ground set $\mathsf{N}\mathsf{N}^*$. Doubling the ground set size for this injection does not change the asymptotic upper bound. □

Combining Proposition 4.45 and Proposition 4.46 we obtain the main result:

Theorem 4.47. For every infinite (ordered) field $\mathbb{K}$, the property $\mathfrak{g}^\bullet_{\mathbb{K}}$ has infinitely many forbidden minors. It is not finitely axiomatizable. □

Nelson's upper bound applies to the set $\bigcup_k \mathfrak{g}^*_k$ of gaussoids which are realizable over **any** field. Since $\emptyset_{\mathsf{N}}$ and $\mathcal{A}_{\mathsf{N}}$ are realizable over every infinite field, the doubly exponential lower bound holds independently of the field. This even allows the following conclusion:

Corollary 4.48. There exists a single infinite family of simultaneous forbidden minors for all $\mathfrak{g}^*_{\mathbb{K}}$ where $\mathbb{K}$ is an infinite field. □

An analogue of Rota's conjecture. The realizability problem and the combinatorial theory of minors and duality for gaussoids and matroids are alike. Theorem 4.47 shows that realizability over infinite fields has infinitely many forbidden minors. The same result exists in matroid theory (see [Oxl11, Theorem 6.5.17]) and according to Oxley was originally proved by Geelen and Whittle. Their proof is explicit and constructive, while ours exploits a recent result of Nelson about the number of linear matroids.

The converse of Theorem 4.47 is known in matroid theory as *Rota's conjecture* [Rot71]. It alleges that linearity of matroids over a fixed finite field has a finite forbidden-minor characterization — that is, the number of forbidden minors is finite if and only if the field is so. It is natural to ask the same about gaussoids:

Question 4.49: Rota's conjecture for gaussoids. Does $\mathfrak{g}^*_{\mathbb{F}_q}$ have a finite forbidden-minor characterization for every prime power q?

We can give support of this conjecture for the two smallest fields. Over these fields, principal regularity, which is an algebraic genericity condition, is particularly restrictive. $\mathbb{F}_2$ only has the identity matrix, hence $\mathfrak{g}^*_{\mathbb{F}_2}$ is characterized by the compulsory 3-minor $\mathtt{F}$. For $\mathbb{F}_3$ the situation is completely understood but becomes combinatorially more interesting:

Theorem 4.50. A gaussoid $\mathcal{G}$ is realizable over $\mathbb{F}_3$ if and only if each of its 3-minors is in $\{\mathtt{B}, \mathtt{F}\}$. In particular, it is graphic.

Proof. One easily checks by a symbolic argument or brute enumeration that every 3×3 principally regular matrix Γ over $\mathbb{F}_3$ has all diagonal entries non-zero and at least two zero off-diagonal entries. If $\gamma_{\mathsf{ij}} \neq 0$, then $\gamma_{\mathsf{ii}} = -\gamma_{\mathsf{jj}} \neq 0$ is necessary and sufficient for $\Gamma[\mathsf{ij}] \neq 0$. This shows $\mathfrak{g}^*_{\mathbb{F}_3}(\mathsf{ijk}) = \{\mathtt{B}, \mathtt{F}\}$.

Consequently, $\mathbb{F}_3$-algebraic Gaussians are $\mathtt{BF}$-gaussoids, which are ascending and descending. It follows that they are completely described by their marginal independence statements $(\mathsf{ij}|)$ and for each triple ijk at least two of the three marginal statements holds. Given any such CI structure, it is clear how to make a compatible $\mathbb{F}_3$-matrix Γ: put $\gamma_{\mathsf{ij}} = 0$ whenever $(\mathsf{ij}|)$ holds. If not, then $\gamma_{\mathsf{ij}} = 1$ and $\gamma_{\mathsf{ii}} = 1 = -\gamma_{\mathsf{jj}}$. This is consistent because in each row and column there is at most one non-zero off-diagonal entry. Set all other diagonals to 1. To see that this matrix is principally regular, invoke the Sign Convention to order the ground set so that i and j are adjacent whenever $\neg(\mathsf{ij}|)$ holds. Then Γ is a block-diagonal matrix with blocks of size 1×1 or 2×2, all of which are regular. This proof works for every principal submatrix of Γ. Thus $[\![\Gamma]\!]$ is an $\mathbb{F}_3$-algebraic Gaussian, hence a $\mathtt{BF}$-gaussoid. Its marginal independence statements and therefore all of $[\![\Gamma]\!]$ coincide with the given $\mathtt{BF}$-gaussoid. □

Remark 4.51. Classes of gaussoids with prescribed (isomorphy-closed) 3-minors have been completely classified in [BK20]. It follows from these results that $\mathbb{F}_3$-algebraic Gaussians on N are in bijection with the involutions in $\mathfrak{S}_{\mathsf{N}}$.

The following question poses a generalization of Rota's conjecture. In matroid theory, it is well-known that even if two classes of matroids have finitely many forbidden minors, their union need not. This is known as the *intertwining problem*; see [Oxl11, Section 14.5].

Question 4.52. Let T be a finite set of primes. Is the class of gaussoids $\mathcal{G}$ with $T \subseteq \chi(\mathcal{G})$ characterized by finitely many forbidden minors?

Many regular semimatroids. The combinatorial argument of this section does not apply to discrete probability distributions. Studený showed that there exist doubly exponentially many discrete CI structures.

The single exponential bound on algebraic Gaussians is inherited from matroid theory. To construct a double exponential number of distinct discrete semigraphoids, we start again in matroid theory, with *regular matroids*. These matroids are realizable over every field and therefore have many kinds of discrete probabilistic realizations. The double exponential growth comes from the fact that the discrete semigraphoids even contain all **intersections** (as CI structures, corresponding to **sums** of rank functions) of these regular matroids.

The construction is due to [Stu05, Corollary 2.7], which, however, does not emphasize the regularity of the matroids. The proof is reproduced here in matroid language.

Lemma 4.53. Let $\mathsf{L} \subseteq \mathsf{N}$ and let C_{L} be the graphic matroid of the undirected graph with edges indexed by N such that the edges in L form a simple cycle and all other edges form a path which does not cross the cycle. The semimatroid $\mathcal{C}_{\mathsf{L}} = [\![C_{\mathsf{L}}]\!]$ is regular and for every $(\mathsf{ij}|\mathsf{K}) \in \mathcal{A}_{\mathsf{N}}$ with $|\mathsf{ijK}| = |\mathsf{L}|$ it holds that $(\mathsf{ij}|\mathsf{K}) \in \mathcal{C}_{\mathsf{L}}$ if and only if $\mathsf{ijK} \neq \mathsf{L}$.

Proof. Every graphic matroid is regular by [Oxl11, Proposition 5.1.2] and this is the same notion as regularity for its associated semimatroid. Let $(\mathsf{ij}|\mathsf{K}) \in \mathcal{A}_{\mathsf{N}}$ be arbitrary with $|\mathsf{ijK}| = |\mathsf{L}|$ and denote the rank function of C_{L} by r. Then $(\mathsf{ij}|\mathsf{K}) \in \mathcal{C}_{\mathsf{L}}$ holds if and only if $\triangle r(\mathsf{ij}|\mathsf{K}) = r(\mathsf{iK}) + r(\mathsf{jK}) - r(\mathsf{ijK}) - r(\mathsf{K}) = 0$. L is the unique circuit in C_{L}, so $\mathsf{ijK} \neq \mathsf{L}$ implies that ijK is independent and hence $\triangle r(\mathsf{ij}|\mathsf{K}) = 0$. Otherwise $\mathsf{ijK} = \mathsf{L}$ and $\triangle r(\mathsf{ij}|\mathsf{K}) = 1$. ☐

Lemma 4.54. There exist asymptotically $2^{\Omega(2^n/\sqrt{n})}$ regular semimatroids.

Proof. Choose any set $\mathcal{L}$ of $\lfloor \frac{n}{2} \rfloor$-element subsets of N. The intersection of the corresponding semimatroids constructed in Lemma 4.53, $\mathcal{C} = \bigcap_{L \in \mathcal{L}} \mathcal{C}_L$, is a regular semimatroid by [Mat97, Lemma 7] and since for every $(\mathsf{ij}|\mathsf{K}) \in \mathcal{A}_{\mathsf{N}}$ with $|\mathsf{ijK}| = \lfloor \frac{n}{2} \rfloor$ we have

$$(\mathsf{ij}|\mathsf{K}) \in \mathcal{C} \;\Leftrightarrow\; \mathsf{ijK} \notin \mathcal{L},$$

this semimatroid identifies the set $\mathcal{L}$ uniquely. It follows that there exist at least as many regular semimatroids as there are choices for $\mathcal{L}$. Using Stirling's formula, one obtains an asymptotically tight estimate for the number of $\lfloor n/2 \rfloor$-element subsets of N as $\Theta(2^n/\sqrt{n})$. ☐

4.6 Two-antecedental completeness of the gaussoid axioms

As outlined in the Introduction, the research into CI structures of discrete random variables in the late 1980s was in part driven by the conjecture of Pearl and Paz that the semigraphoid axioms were complete for the theory of discrete CI structures. This conjecture was refuted by Studený in [Stu92] as discussed in Section 4.4. Studený's inference rules naturally require a growing number of random variables but they also use more and more antecedents. A finite complete list of valid inference rules being impossible to obtain, the conjecture was revised to state that the semigraphoid axioms are complete for only those inference rules of discrete CI which have at most two antecedents, as the semigraphoid axioms themselves do. This was in turn resolved positively by Studený [Stu94]. In this sense, the semigraphoid axioms are the most fundamental laws of CI on discrete random vectors: they logically imply all the non-trivial valid inference rules with the lowest number of antecedents, or premises — those which one would expect to be most acceptable in human reasoning. The goal of this section is to prove the analogous result for Gaussians and the gaussoid axioms.

By context-completeness (Corollary 4.33) the minimal axioms for Gaussians can be discovered peu à peu together with the compulsory minors by growing the ground set under consideration. On each fixed ground set the realizability of a CI structure or the validity of an inference formula is an algebraic problem. The gaussoid axioms were discovered as a **complete** set of axioms for 3-variate (positive or algebraic) Gaussians in [Mat05]. Later investigations by Lněnička and Matúš [LM07] produced the complete axioms (LM.i)–(LM.v) for 4-variate positive Gaussians. With the solution of the realizability problem on each successive ground set size, a better approximation of Gaussianity is obtained, but the approximations apparently become more difficult in each step. This effect is obvious for the partial forbidden-minor characterization obtained, because of minor-closedness but it is not clear why the minimal axioms should become more complicated with growing ground set. The following definition makes this complexity measure more precise:

Definition 4.55. Let $\mathfrak{p} \geq \mathfrak{p}^*$ be properties of CI structures. $\mathfrak{p}$ is a *k-antecedental approximation* of $\mathfrak{p}^*$ if on every ground set N every inference form φ with at most k antecedents and variables in $\mathcal{A}_{\mathsf{N}}$ which is valid for $\mathfrak{p}^*$ is also valid for $\mathfrak{p}$.

We imagine $\mathfrak{p}$ to be a simpler, necessary property approximating $\mathfrak{p}^*$. Because of the inclusions $\mathfrak{p}^* \leq \mathfrak{p}$, every inference rule which is valid for $\mathfrak{p}$ also holds for $\mathfrak{p}^*$. The definition above concerns a degree k to which the converse holds. In this section, the role of $\mathfrak{p}$ is played by gaussoids and that of $\mathfrak{p}^*$ by realizable gaussoids, for different notions of realizability. From an axiomatic point of view, one may also say that the axioms for $\mathfrak{p}$ are *k-antecedentally complete* for the chosen notion of realizabillity $\mathfrak{p}^*$. We prove

Theorem 4.56. Gaussoids are two-antecedental approximations of algebraic and of positive Gaussian conditional independence structures over characteristic zero.

Our proof of the two-antecedental approximation property relies on a general principle which was also used in Studený's proof for discrete CI. A *minimal* $\mathfrak{p}$*-extension* of a CI structure $\mathcal{A}$ is a CI structure $\mathcal{A}'$ which is inclusion-minimal with the properties that $\mathcal{A}' \supseteq \mathcal{A}$ and $\mathcal{A}' \in \mathfrak{p}$. In the world of discrete CI and semigraphoids, this minimal extension is unique, because both sets are closed under intersection, but for Gaussians and gaussoids it is not.

Lemma 4.57. Let $\mathfrak{p} \geq \mathfrak{p}^*$ be properties of CI structures. Then $\mathfrak{p}$ is a k-antecedental approximation of $\mathfrak{p}^*$ if for every N every minimal $\mathfrak{p}$-extension of every subset of $\mathcal{A}_{\mathsf{N}}$ of cardinality at most k belongs to $\mathfrak{p}^*$.

Proof. Let $\varphi : \bigwedge \mathcal{L} \Rightarrow \bigvee \mathcal{M}$ be a valid inference rule for $\mathfrak{p}^*$ with $|\mathcal{L}| \leq k$. We have to show that φ is valid for $\mathfrak{p}$. Equivalently, letting $\mathfrak{p} \wedge \varphi$ denote the largest subproperty of $\mathfrak{p}$ which additionally satisfies φ, we show that $\mathfrak{p} \leq \mathfrak{p} \wedge \varphi$.

Consider any $\mathcal{A} \in \mathfrak{p}(\mathsf{N})$. If the antecedents $\mathcal{L}$ of φ are not contained in $\mathcal{A}$, then it vacuously satisfies $\mathfrak{p} \wedge \varphi$. On the other hand, if it contains $\mathcal{L}$, then it also contains a minimal $\mathfrak{p}$-extension $\mathcal{L}'$ of $\mathcal{L}$. Since $|\mathcal{L}| \leq k$, the structure $\mathcal{L}'$ belongs to $\mathfrak{p}^*$ by assumption. Hence $\mathcal{L}'$ satisfies φ, which means that $\mathcal{L}' \cap \mathcal{M} \neq \emptyset$. Then $\mathcal{A}$, containing $\mathcal{L}'$, also satisfies φ. □

The main theorem is a consequence of this technical refinement:

Theorem 4.58. Over every ground set, every minimal gaussoid extension of at most two CI statements is realizable by a positive-definite matrix with rational entries, which can be picked arbitrarily close to the identity matrix.

Among all realizabilities over characteristic zero, the positive one over $\mathbb{Q}$ is the strongest result. The closeness to the identity matrix is a yet stronger topological property which proves helpful in the proof details. This result rests again on our rational transfer principle Lemma 4.1 which allows the recovery of positive realizations from an algebraic construction and a continuity-style argument. A further prominent role is played by the hyperoctahedral group acting on algebraic Gaussians. We turn the problem of finding positive-definite realizations for the given gaussoids around and instead find algebraic realizations of **one** easy hyperoctahedral representative of each gaussoid orbit converging to **every** hyperoctahedral image of the identity matrix. Then, given any gaussoid in the representative's orbit, we apply the inverse group action to the right algebraic realization so that it is transformed into a **near-identity** and hence rational positive realization.

The hyperoctahedral orbit of the identity matrix. It easily follows from Proposition 3.16 that the quotient action $\mathcal{S}_{\mathsf{Z}}$ on the $(\mathbb{Z}/4)^{\mathsf{N}}$-orbit of the identity matrix, where components X_{i} with different signs δ_{i} are identified, produces well-defined matrices, independent of the choice δ_{i} of representatives. This orbit consists of all 2^n diagonal matrices with entries $(\pm 1, \dots, \pm 1)$. For any matrix J in this orbit the action $\mathcal{S}_{\mathsf{Z}}(J)$ flips the signs of the diagonal entries of J indicated by Z. The action of $\mathfrak{S}_{\mathsf{N}}$ does not leave this set of matrices either, so it constitutes an orbit under the hyperoctahedral group $\mathfrak{B}_{\mathsf{N}}$.

Realizability near the identity matrix or its hyperoctahedral images is a well-behaved notion in the theory of CI structures: it is closed under minors, embeddings, isolation, direct and dependent sum as well as symmetries. For example, take a principally regular matrix Γ over $\mathbb{K}(\varepsilon_1, \dots, \varepsilon_p)$ with Γ° in this orbit. By Proposition 3.16, the hyperoctahedral action produces a principally regular matrix Δ over the same field such that Δ° belongs to the hyperoctahedral orbit of the identity as well. The dependent sum in Lemma 4.20, duality, marginalization and conditioning and their reversals in Lemma 4.21 of algebraic Gaussians preserve realizability near a hyperoctahedral image of the identity over their respective ground sets. We use these facts in the following proposition which will be the main realizability tool for the most complicated cases in Section 4.6.

Proposition 4.59. If a gaussoid $\mathcal{G}$ is (rationally) realizable near every one of the 2^n hyperoctahedral image of the identity matrix, then every hyperoctahedral image of $\mathcal{G}$ is (rationally) near-identity realizable, in particular positive (over $\mathbb{Q}$).

Proof. Let $\mathcal{H}$ be in $\mathcal{G}$'s hyperoctahedral orbit, arising from $\mathcal{G}$ by a swap and a permutation. $\mathcal{H}$ is realizable near the identity if and only if $\mathcal{G}$ is realizable near the matrix which is obtained from the identity by permuting and swapping in reverse. These operations result in a hyperoctahedral sibling of the identity near which $\mathcal{G}$ is realizable by assumption. The hyperoctahedral action which transports this curve of realizations back to realize $\mathcal{H}$ near the identity does not change the field, so rationality is preserved. □

Rational realizability proofs. We can now give the proof that every minimal gaussoid extension of at most two CI statements over $\mathcal{A}_{\mathsf{N}}$ for any N is rationally realizable near the identity in a series of lemmas.

Lemma 4.60. All CI structures with at most one element are rationally realizable near the identity.

Proof. The empty structure is realized by a symmetric matrix with 1-diagonal and independent variables in the off-diagonal entries. Clearly, none of the almost-principal minors of this matrix vanish as polynomials. Every singleton subset of $\mathcal{A}_{\mathsf{N}}$ is vacuously a gaussoid. The singleton gaussoids form a single orbit under the action of the hyperoctahedral group.

While this action does not in general preserve positive realizability, we can emulate it using Lemma 4.21 in a way that shows that it *is* preserved in this case. Given any singleton $(\mathsf{ij}|\mathsf{K})$, first permute it to $(12|\mathsf{K}')$, marginalize to $12\mathsf{K}'$ and then contract K' to arrive at the singleton $(12|)$ over the ground set $\mathsf{N} = 12$. These transformations preserve rational realizability near the identity and their inverses do as well. Thus we can transform every singleton into every other singleton while preserving realizability and it remains to see that $\{ (12|) \}$ is rationally realizable near the identity, for which $\mathbb{1}_{12}$ itself is a witness. □

From now on we consider two-element sets $\{ (\mathsf{ij}|\mathsf{N}), (\mathsf{kl}|\mathsf{M}) \}$ and their minimal gaussoid extensions. Using the fact that marginalizations and conditionings of Gaussians are Gaussians (over the same field) and that we can undo these operations generically via Lemma 4.21, we can assume that we work over the ground set ijklNM and that $\mathsf{N} \cap \mathsf{M} = \emptyset$.

The gaussoid axioms have two antecedents. Every antecedent set of a gaussoid axiom is therefore not a gaussoid. The following lemma deals with this type:

Lemma 4.61. If $\mathcal{B} = \{ (\mathsf{ij}|\mathsf{N}), (\mathsf{kl}|\mathsf{M}) \}$ is not a gaussoid, then each of its minimal gaussoid extensions has cardinality four and is rationally realizable near the identity.

Proof. $\mathcal{B}$ not being a gaussoid requires that $(\mathsf{ij}|\mathsf{N})$ and $(\mathsf{kl}|\mathsf{M})$ are distinct and form the antecedent set of a gaussoid axiom. Thus the two CI statements lie in a 3-face of the ambient ijklNM-cube. We can therefore reduce the study of gaussoid extensions of $\mathcal{B}$ to this 3-face and hence, after conditioning, to a 3-element ground set. Every gaussoid closure of $\mathcal{B}$ is thus a 3-gaussoid which is placed in a 3-face of the ijklNM-cube. With two generators, each closure has exactly four elements. The 3-gaussoids are all realizable as undirected graphical models or their duals. Rational near-identity realizations have been constructed in [LM07, Theorem 1] and those are embedded back into the ijklNM-cube via Lemma 4.21. □

The remaining type of gaussoids is comprised of pairs of so-called *inferenceless* generators with respect to the gaussoid axioms: two-element subsets of $\mathcal{A}_\mathsf{N}$ which are vacuously gaussoids. We expect this type to be the hardest to realize. The gaussoid axioms, as the previous proof shows, govern inferences of two CI statements in a common 3-face of the hypercube. The realizabilities of inferenceless pairs prove that there are no valid two-antecedental inference rules for Gaussian CI whose antecedents lie further apart in the hypercube than in a common 3-face.

We continue to assume that the ground set is ijklNM and that $\mathsf{N} \cap \mathsf{M} = \emptyset$. In addition, the assumption of inferenceless generators can be expressed as

$$|\mathsf{N} \oplus \mathsf{M}| \geq |\mathsf{ij} \cap \mathsf{kl}|. \tag{†}$$

This type splits into a number of cases depending on how "entangled" ij, kl, N and M are, as these entanglements influence the form of a potential realizing matrix. Up to the group $\mathbb{Z}/2 \times \mathfrak{S}_\mathsf{N}$ of duality and isomorphy, which preserves rational positive realizability, and symmetries in the roles of ij and kl, there are seven cases:

	1	2	3	4	5	6	7
$\mathsf{ij} \cap \mathsf{kl}$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$	i	i	ij
$\mathsf{ij} \cap \mathsf{M}$	$\emptyset$	i	i	ij	$\emptyset$	j	$\emptyset$
$\mathsf{kl} \cap \mathsf{N}$	$\emptyset$	$\emptyset$	k	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$

The hyperoctahedral group has a considerably wider reach. Every gaussoid $\{ (\mathsf{ij}|\mathsf{N}), (\mathsf{kl}|\mathsf{M}) \}$ can be transformed into $\{ (\mathsf{ij}|), (\mathsf{kl}|\mathsf{M}') \}$, where $\mathsf{ij} \cap \mathsf{M}' = \emptyset$, by swapping out $\mathsf{N} \cup (\mathsf{M} \cap \mathsf{ij})$. This reduces the table above to only three cases:

— $\{\,(\mathsf{ij}|), (\mathsf{kl}|\mathsf{M})\,\}$ on ijklM with $\mathsf{ij} \cap \mathsf{klM} = \emptyset$,
— $\{\,(\mathsf{ij}|), (\mathsf{ik}|\mathsf{M})\,\}$ on ijkM with $\mathsf{ij} \cap \mathsf{kM} = \emptyset$,
— $\{\,(\mathsf{ij}|), (\mathsf{ij}|\mathsf{M})\,\}$ on ijM with $\mathsf{ij} \cap \mathsf{M} = \emptyset$.

These cases are fewer and easier because $(\mathsf{ij}|)$ only mandates that a specific entry of the realizing matrix be zero. This reduction comes at the cost of not preserving positive realizability. Using Proposition 4.59, to obtain rational positive realizability of the entire orbit, we realize the above three case representatives rationally near all the matrices which are equivalent to the identity under the hyperoctahedral action.

The first case is the union of gaussoids $\{\,(\mathsf{ij}|)\,\}$ and $\{\,(\mathsf{kl}|\mathsf{M})\,\}$ over disjoint ground sets ij and klM and is already settled as an instance of Lemma 4.20 with the rational near-identity realizations constructed for singleton gaussoids in Lemma 4.60. The remaining two cases are settled by Lemmas 4.62 and 4.63 below.

Lemma 4.62. The gaussoid $\{\,(\mathsf{ij}|), (\mathsf{ik}|\mathsf{M})\,\}$ on the ground set ijkM is rationally realizable near all hyperoctahedral images of the identity.

Proof. We introduce the following notation:

$$\Phi = \begin{array}{c} \begin{array}{cccc} \mathsf{i} & \mathsf{j} & \mathsf{k} & \mathsf{M} \end{array} \\ \begin{pmatrix} \pm 1 & 0 & \xi & \boldsymbol{u}^\mathsf{T} \\ 0 & \pm 1 & \eta & \boldsymbol{v}^\mathsf{T} \\ \xi & \eta & \pm 1 & \boldsymbol{w}^\mathsf{T} \\ \boldsymbol{u} & \boldsymbol{v} & \boldsymbol{w} & \Sigma \end{pmatrix} \end{array} \begin{array}{c} \mathsf{i} \\ \mathsf{j} \\ \mathsf{k} \\ \mathsf{M} \end{array},$$

where $(\mathsf{ij}|)$ is already fulfilled by imposing the zero entry. The statement $(\mathsf{ik}|\mathsf{M})$ is equivalent to the following relation, after a Schur complement:

$$\xi = \boldsymbol{u}^\mathsf{T} \Sigma^{-1} \boldsymbol{w}.$$

Thus we impose this relation on ξ. All other appearing symbols are supposed to be generic, i.e., η is a variable, the vectors have independent variable entries u_m, v_m, w_m, for $\mathsf{m} \in \mathsf{M}$, and Σ is a generic symmetric matrix with ± 1-diagonal and independent $\varepsilon_{\mathsf{mn}}$ off-diagonals. The signs of the diagonal elements of Φ are arbitrary but fixed. Φ is a matrix over $\mathbb{Q}(\eta, u_\mathsf{m}, v_\mathsf{m}, w_\mathsf{m}, \varepsilon_\mathsf{mn})$ whose off-diagonal entries tend to zero with the infinitesimal variables and thus it approaches any hyperoctahedral image of the identity matrix. The only denominator appears in ξ and is the principal minor $\det \Sigma$ with constant term ± 1, which is infinitesimally non-zero.

By construction, $(\mathsf{ij}|)$ and $(\mathsf{ik}|\mathsf{M})$ hold for Φ. It is clear that the only interesting almost-principal minors are those involving ξ. For any $\mathsf{N} \subsetneq \mathsf{M}$, the statement $(\mathsf{ik}|\mathsf{N})$ surely does not hold because it is equivalent to

$$\boldsymbol{u}^\mathsf{T} \Sigma^{-1} \boldsymbol{w} = \boldsymbol{u}_\mathsf{N}^\mathsf{T} \Sigma_\mathsf{N}^{-1} \boldsymbol{w}_\mathsf{N},$$

where the variables in $\boldsymbol{u}, \boldsymbol{w}, \Sigma$ are all independent. There are four remaining cases: $(\mathsf{ik}|\mathsf{jN})$, $(\mathsf{il}|\mathsf{kN})$, $(\mathsf{kl}|\mathsf{iN})$ and $(\mathsf{lm}|\mathsf{ikN})$, for relevant choices of l, m and $\mathsf{N} \subseteq \mathsf{M}$.

The almost-principal minor $(\mathsf{ik}|\mathsf{jN})$ is rewritten using Schur complement to

$$\Phi[\mathsf{ik}|\mathsf{jN}] = \Phi[\mathsf{jN}] \left(\xi - \begin{pmatrix} 0 & \boldsymbol{u}_\mathsf{N}^\mathsf{T} \end{pmatrix} \Phi_{\mathsf{jN}}^{-1} \begin{pmatrix} \eta \\ \boldsymbol{w}_\mathsf{N}^\mathsf{T} \end{pmatrix} \right),$$

which vanishes if and only if the parenthesized factor vanishes as a rational function. Numerator and denominator of ξ do not involve the variable η, so it suffices to show that there is a monomial divisible by η with non-zero coefficient in the "bilinear term" inside the parentheses. All terms involving η are:

$$\eta \sum_{\mathsf{n}\in\mathsf{N}} u_{\mathsf{n}} \left(\Phi_{\mathsf{jN}}^{-1}\right)_{\mathsf{jn}}$$

Each of these summands has a unique variable u_{n} which does not appear in Φ_{jN}. When $\mathsf{N} \neq \emptyset$, this ensures that the η terms do not cancel and that the almost-principal minor does not vanish. In case $\mathsf{N} = \emptyset$, it is sufficient to remark that $\phi_{\mathsf{ik}} = \xi \neq 0$ because $\mathsf{M} \neq \emptyset$ due to the assumption of inferenceless generators (†).

For the case $(\mathsf{il}|\mathsf{kN})$ first assume that $\mathsf{l} \neq \mathsf{j}$. Laplace expansion on the first row of the almost principal minor gives a sum

$$\det \begin{pmatrix} u_{\mathsf{l}} & \xi & \boldsymbol{u}_{\mathsf{N}}^{\mathsf{T}} \\ w_{\mathsf{l}} & \pm 1 & \boldsymbol{w}_{\mathsf{N}} \\ \Sigma_{\mathsf{N},\mathsf{l}} & \boldsymbol{w}_{\mathsf{N}}^{\mathsf{T}} & \Sigma_{\mathsf{N}} \end{pmatrix} = u_{\mathsf{l}} \Phi[\mathsf{kN}] \mp \dots$$

of which the omitted terms are not divisible by u_{l}. Since the constant term of $\Phi[\mathsf{kN}]$ is ± 1, the monomial u_{l} arises in the sum and cannot be canceled by other terms. If $\mathsf{l} = \mathsf{j}$, then $(\mathsf{ij}|\mathsf{kN})$ is equivalent to

$$0 = \begin{pmatrix} \xi & \boldsymbol{u}_{\mathsf{N}}^{\mathsf{T}} \end{pmatrix} \Phi_{\mathsf{kN}}^{-1} \begin{pmatrix} \eta \\ \boldsymbol{v}_{\mathsf{N}}^{\mathsf{T}} \end{pmatrix}.$$

Again we investigate the terms divisible by η:

$$\eta \left(\xi (\Phi_{\mathsf{kN}}^{-1})_{\mathsf{kk}} + \sum_{\mathsf{n}\in\mathsf{N}} u_{\mathsf{n}} (\Phi_{\mathsf{kN}}^{-1})_{\mathsf{kn}} \right).$$

Since $\xi \neq 0$ and $(\Phi_{\mathsf{kN}}^{-1})_{\mathsf{kk}}$ has constant term ± 1, we find the monomial $\eta u_{\mathsf{m}} w_{\mathsf{m}}$ for some $\mathsf{m} \in \mathsf{M}$ in the numerator of this almost-principal minor.

The case $(\mathsf{kl}|\mathsf{iN})$ for $\mathsf{l} \neq \mathsf{j}$ is completely analogous to the previous $(\mathsf{il}|\mathsf{kN})$ one. In fact, the involved matrices are identical up to exchanging the places of the generic vectors $\boldsymbol{u}$ and $\boldsymbol{w}$, which already play symmetric roles in the definition of ξ. The matrix for $(\mathsf{kj}|\mathsf{iN})$ is

$$\begin{pmatrix} \eta & \xi & \boldsymbol{w}_{\mathsf{N}}^{\mathsf{T}} \\ 0 & \pm 1 & \boldsymbol{u}_{\mathsf{N}}^{\mathsf{T}} \\ \boldsymbol{v}_{\mathsf{N}} & \boldsymbol{u}_{\mathsf{N}} & \Sigma_{\mathsf{N}} \end{pmatrix}$$

and again Laplace expansion can be used to see that η survives as a monomial of degree one.

The last case is $(\mathsf{lm}|\mathsf{ikN})$. When $\mathsf{j} \notin \mathsf{lm}$, the almost-principal minor of

$$\begin{pmatrix} \varepsilon_{\mathsf{lm}} & u_{\mathsf{l}} & w_{\mathsf{l}} & \Sigma_{\mathsf{l},\mathsf{N}} \\ u_{\mathsf{m}} & \pm 1 & \xi & \boldsymbol{u}_{\mathsf{N}}^{\mathsf{T}} \\ w_{\mathsf{m}} & \xi & \pm 1 & \boldsymbol{w}_{\mathsf{N}}^{\mathsf{T}} \\ \Sigma_{\mathsf{N},\mathsf{m}} & \boldsymbol{u}_{\mathsf{N}} & \boldsymbol{w}_{\mathsf{N}} & \Sigma_{\mathsf{N}} \end{pmatrix}$$

has a monomial $\varepsilon_{\mathsf{lm}}$ via Laplace expansion in the first row. The numerators of other summands in this expansion are not divisible by $\varepsilon_{\mathsf{lm}}$, making it impossible to cancel this degree-1 monomial. Otherwise, without loss of generality, $\mathsf{m} = \mathsf{j}$ and the matrix

$$\begin{pmatrix} v_{\mathsf{l}} & u_{\mathsf{l}} & w_{\mathsf{l}} & \Sigma_{\mathsf{l},\mathsf{N}} \\ 0 & \pm 1 & \xi & \boldsymbol{u}_{\mathsf{N}}^{\mathsf{T}} \\ \eta & \xi & \pm 1 & \boldsymbol{w}_{\mathsf{N}}^{\mathsf{T}} \\ \boldsymbol{v}_{\mathsf{N}} & \boldsymbol{u}_{\mathsf{N}} & \boldsymbol{w}_{\mathsf{N}} & \Sigma_{\mathsf{N}} \end{pmatrix}$$

is susceptible to the same Laplace expansion proof yielding a monomial v_{l}. $\square$

Lemma 4.63. The gaussoid $\{\,(\mathsf{ij}|), (\mathsf{ij}|\mathsf{M})\,\}$ on the ground set ijM is rationally realizable near all hyperoctahedral images of the identity.

Proof. We use the matrix pattern

$$\Phi = \begin{array}{c} \\ \\ \\ \\ \end{array}\begin{array}{ccc} \mathsf{i} & \mathsf{j} & \mathsf{M} \end{array}$$

$$\Phi = \begin{pmatrix} \pm 1 & 0 & \boldsymbol{u}^\mathsf{T} \\ 0 & \pm 1 & \boldsymbol{v}^\mathsf{T} \\ \boldsymbol{u} & \boldsymbol{v} & \Sigma \end{pmatrix}\begin{array}{l} \mathsf{i} \\ \mathsf{j} \\ \mathsf{M} \end{array}$$

with column vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ and a generic matrix Σ with ± 1-diagonal and independent $\varepsilon_{\mathsf{mn}}$ off-diagonals. Again, $(\mathsf{ij}|)$ is imposed already by a zero entry. Unlike the situation of Lemma 4.62, we cannot make $(\mathsf{ij}|\mathsf{M})$ hold by imposing a relation on ϕ_{ij} which is already set to zero. The equation for $(\mathsf{ij}|\mathsf{M})$ is equivalent to

$$0 = \boldsymbol{u}^\mathsf{T} \operatorname{adj}(\Sigma)\, \boldsymbol{v},$$

that is, $\boldsymbol{u}$ and $\boldsymbol{v}$ are orthogonal with respect to the (infinitesimally principally regular) adjoint of Σ. Equivalently we could have used Σ^{-1} but prefer not to introduce denominators into Φ needlessly. To enforce this relation, we define $\boldsymbol{u}$ and $\boldsymbol{v}$ via the Gram–Schmidt process on vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ of mutually independent variables indexed by M, as follows:

$$\begin{aligned} u_\mathsf{k} &= x_\mathsf{k}, \\ v_\mathsf{k} &= \alpha_\mathsf{M} y_\mathsf{k} - \beta_\mathsf{M} x_\mathsf{k}, \end{aligned}$$

with the inner products $\alpha_\mathsf{L} = \boldsymbol{x}_\mathsf{L}^\mathsf{T} \operatorname{adj}(\Sigma_\mathsf{L}) \boldsymbol{x}_\mathsf{L}$ and $\beta_\mathsf{L} = \boldsymbol{x}_\mathsf{L}^\mathsf{T} \operatorname{adj}(\Sigma_\mathsf{L}) \boldsymbol{y}_\mathsf{L}$ for any $\mathsf{L} \subseteq \mathsf{M}$. This completes the definition of Φ, which is a matrix over $\mathbb{Q}(x_\mathsf{m}, y_\mathsf{m}, \varepsilon_{\mathsf{mn}})$ whose off-diagonal entries tend to zero with the infinitesimal variables, and clearly $[\![\Phi]\!]$ contains $(\mathsf{ij}|)$ and $(\mathsf{ij}|\mathsf{M})$. Evidently $(\mathsf{kl}|\mathsf{N}) \notin [\![\Phi]\!]$ whenever $\mathsf{j} \notin \mathsf{klN}$ because the almost-principal submatrix is generic in this case. The remainder of the proof treats CI statements of the forms $(\mathsf{ij}|\mathsf{N})$, $(\mathsf{jk}|\mathsf{N})$ and $(\mathsf{kl}|\mathsf{jN})$ each for all suitable k, l and $\mathsf{N} \subseteq \mathsf{M}$.

When N is any non-empty subset of M, the almost-principal minor $(\mathsf{ij}|\mathsf{N})$ becomes

$$\begin{aligned} \Phi[\mathsf{ij}|\mathsf{N}] &= \boldsymbol{u}_\mathsf{N}^\mathsf{T} \operatorname{adj}(\Sigma_\mathsf{N})\, \boldsymbol{v}_\mathsf{N} \\ &= \alpha_\mathsf{M} \boldsymbol{x}_\mathsf{N}^\mathsf{T} \operatorname{adj}(\Sigma_\mathsf{N})\, \boldsymbol{y}_\mathsf{N} - \beta_\mathsf{M} \boldsymbol{x}_\mathsf{N}^\mathsf{T} \operatorname{adj}(\Sigma_\mathsf{N})\, \boldsymbol{x}_\mathsf{N} \\ &= \alpha_\mathsf{M} \beta_\mathsf{N} - \alpha_\mathsf{N} \beta_\mathsf{M}. \end{aligned}$$

When $\mathsf{N} \neq \mathsf{M}$ it suffices to find a monomial in $\alpha_\mathsf{M}\beta_\mathsf{N}$ which does not appear in $\alpha_\mathsf{N}\beta_\mathsf{M}$. Given $\mathsf{k} \in \mathsf{N}$ and $\mathsf{m} \in \mathsf{M} \setminus \mathsf{N}$, and using that x_m^2 only appears in α_M via the constant term ± 1 in the cofactor $(\operatorname{adj}\Sigma)_{\mathsf{mm}}$, such a monomial is $x_\mathsf{m}^2 x_\mathsf{k} y_\mathsf{k}$.

Next consider type $(\mathsf{jk}|\mathsf{N})$ with

$$\Phi[\mathsf{jk}|\mathsf{N}] = \Phi[\mathsf{N}]\,(\alpha_\mathsf{M} y_\mathsf{k} - \beta_\mathsf{M} x_\mathsf{k}) - \Phi_{\mathsf{j},\mathsf{N}} \operatorname{adj}(\Phi_\mathsf{N})\, \Phi_{\mathsf{N},\mathsf{k}}.$$

By the assumption of inferenceless generators (†), $|\mathsf{M}| \geq 2$, so there exists $\mathsf{m} \in \mathsf{M} \setminus \mathsf{k}$. The expansion of $\alpha_\mathsf{M} y_\mathsf{k}$ produces the monomial $x_\mathsf{m}^2 y_\mathsf{k}$ which does not appear in $\beta_\mathsf{M} x_\mathsf{k}$. Thus this monomial arises from the product term. The remaining term is a bilinear form with respect to $\operatorname{adj}(\Phi_\mathsf{N})$. Expanding the $\Phi_{\mathsf{j},\mathsf{N}}$ vector with the convention $x_\mathsf{i} = y_\mathsf{i} = 0$ in case $\mathsf{N} \ni \mathsf{i}$, we find

$$\Phi_{\mathsf{j},\mathsf{N}} \operatorname{adj}(\Phi_\mathsf{N})\, \Phi_{\mathsf{N},\mathsf{k}} = \alpha_\mathsf{M} \boldsymbol{y}_\mathsf{N}^\mathsf{T} \operatorname{adj}(\Phi_\mathsf{N})\, \Phi_{\mathsf{k},\mathsf{N}} - \beta_\mathsf{M} \boldsymbol{x}_\mathsf{N}^\mathsf{T} \operatorname{adj}(\Phi_\mathsf{N})\, \Phi_{\mathsf{k},\mathsf{N}}.$$

Each monomial in α_M or β_M has total degree at least 2; y_n, x_n and Φ_{kn} are variables or zero if $\mathsf{n} = \mathsf{i} \in \mathsf{N}$. Under no circumstance does any monomial of total degree 3 arise. This proves that $x_\mathsf{m}^2 y_\mathsf{k}$ is a monomial with non-zero coefficient in the expansion of $\Phi[\mathsf{jk}|\mathsf{N}]$, hence $(\mathsf{jk}|\mathsf{N}) \notin [\![\Phi]\!]$.

The last type $(\mathsf{kl}|\mathsf{jN})$ splits into two cases, depending on whether $\mathsf{i} \in \mathsf{kl}$ or not. The proofs are similar, so suppose first that $\mathsf{i} \notin \mathsf{kl}$. Then

$$\Phi[\mathsf{kl}|\mathsf{jN}] = \Phi[\mathsf{jN}]\varepsilon_{\mathsf{kl}} - \Phi_{\mathsf{k},\mathsf{jN}} \operatorname{adj}(\Phi_{\mathsf{jN}})\, \Phi_{\mathsf{jN},\mathsf{l}}.$$

Because $\Phi[\mathsf{jN}]$ has constant term ± 1, the monomial $\varepsilon_{\mathsf{kl}}$ appears in the above expansion of the almost-principal minor. It suffices to show that the bilinear form term does not produce this monomial. Indeed,

$$\Phi_{\mathsf{k},\mathsf{jN}} \operatorname{adj}(\Phi_{\mathsf{jN}})\, \Phi_{\mathsf{jN},\mathsf{l}} = \sum_{\mathsf{a},\mathsf{b} \in \mathsf{jN}} \phi_{\mathsf{ak}}\phi_{\mathsf{bl}}(\operatorname{adj} \Phi_{\mathsf{jN}})_{\mathsf{ab}}$$

sums over products of three polynomials of which at most the entry of the adjoint may have a non-zero constant term. Analogously to above, this sum has no monomial of total degree 1, so the $\varepsilon_{\mathsf{kl}}$ monomial cannot be canceled. Finally, when $\mathsf{l} = \mathsf{i}$, the $\varepsilon_{\mathsf{kl}}$ term in the calculation above becomes x_{k} instead and one of the coordinates in the bilinear form term is the 0 in ϕ_{ij}. However, this does not interfere with the argument. □

This proves that on every ground set N, a boolean formula in inference form $\varphi : \bigwedge \mathcal{L} \Rightarrow \bigvee \mathcal{M}$ in variables $\mathcal{A}_{\mathsf{N}}$ and with $|\mathcal{L}| \le 2$ is valid for all regular Gaussian distributions if and only if the gaussoid axioms on N logically imply φ. The same holds for algebraic (positive) realizability over all (ordered) fields of characteristic zero. Theorem 4.58 does not hold in general in positive characteristic. For example, the only principally regular matrix over $\mathbb{F}_2$ is the identity matrix, so $\mathfrak{g}^*_{\mathbb{F}_2}$ satisfies many inference rules which are not implied by the gaussoid axioms. The proof strategy begins to fail over finite fields with the genericity requirements of Lemma 4.1. The result does not generalize to more antecedents either: a valid *three*-antecedental inference rule for Gaussians which is not implied by the gaussoid axioms was found by Lněnička and Matúš in [LM07, Lemma 10, (20)]. The offending gaussoid is $\{\, (12|3), (13|4), (14|2) \,\}$ over $N = 1234$, which is the instance in dimension 4 of the family in Example 4.40 used by Studený and Šimeček for their non-axiomatizability results.

In particular, our proof strategy shows that all minimal gaussoid extensions of at most two CI statements are rationally realizable near the identity matrix. This is not true anymore for gaussoids with *three* elements:

Example 4.64: A non-near-identity realizable gaussoid. Entry № 20 in [LM07, Table 1] contains the curve of matrices

$$\begin{pmatrix} 1 & 2-\delta^{-2} & \delta & \delta \\ 2-\delta^{-2} & 1 & 0 & \delta \\ \delta & 0 & 1 & \delta^2 \\ \delta & \delta & \delta^2 & 1 \end{pmatrix} \quad \longrightarrow \quad \begin{pmatrix} 1 & 2/9 & 3/4 & 3/4 \\ 2/9 & 1 & 0 & 3/4 \\ 3/4 & 0 & 1 & 9/16 \\ 3/4 & 3/4 & 9/16 & 1 \end{pmatrix}, \text{ as } \delta \to 3/4.$$

The gaussoid $\mathcal{G} = \{(13|24), (23|), (34|1)\}$ realized by this matrix is the algebraic Gaussian in our sense over $\mathbb{Q}(\varepsilon)$ where $3/4 + \varepsilon$ is substituted for δ.

Consider the slice of the positive realization space over $\mathbb{R}$ of this gaussoid on the affine-linear space of symmetric matrices

$$\Sigma = \begin{pmatrix} 1 & a & b & c \\ a & 1 & 0 & e \\ b & 0 & 1 & f \\ c & e & f & 1 \end{pmatrix}.$$

This slice is the intersection of the algebraic realization space of the gaussoid with the elliptope. The identity matrix lies in the center of the elliptope and we wish to show that the realization space of $\mathcal{G}$, although non-empty, does not approach this center. In particular

$\mathcal{G}$ is a gaussoid with *three* elements, rationally positively realizable, but not realizable near the identity. On the given slice, these two equations hold:

$$f = bc, \tag{34|1}$$
$$b + aef = cf + be^2. \tag{13|24}$$

Substituting the first into the second equality and canceling the non-zero factor b in every term we find

$$1 + ace = c^2 + e^2.$$

This equation cannot be satisfied if a, c and e all tend to zero. It can be shown that the realization space of $\mathcal{G}$ decomposes into eight *reorientation classes* which are identical up to an orthogonal transformation; for an explanation of reorientation, see Section 6.2.2. This transformation preserves euclidean distances and it fixes the center of the elliptope, thus all of these components have the same distance to the identity matrix. Focusing on one of them, we can assume that a, c and e are all positive and then the euclidean distance of a realization of $\mathcal{G}$ to the identity is

$$\sqrt{2}\sqrt{a^2+b^2+c^2+e^2+f^2} = \sqrt{2}\sqrt{1+ace+a^2+b^2+f^2} \geq \sqrt{2}.$$

By allowing e to converge to one while a, b and c converge to zero, one can find positive-definite realizations of $\mathcal{G}$ which approach this lower bound. Thus, the elliptope slice of the realization space of $\mathcal{G}$ has distance $\sqrt{2}$ to the identity matrix. △

Example 4.65: Near-identity realizability not preserved under $\mathfrak{B}_\mathsf{N}$. The bracketed self-dual gaussoid $\{(12|), (12|34), (34|1), (34|2)\}$ in item 4_{12} in [BDKS19, p. 15] is not positively realizable over $\mathbb{R}$. However, in its hyperoctahedral orbit is the likewise self-dual $\{(12|3), (12|4), (34|1), (34|2)\}$ and this gaussoid is even realizable rationally near the identity matrix — it is № 30 in [LM07, Table 1]. △

In private correspondence, Milan Studený kindly pointed out that Theorem 4.58 is not a complete analogue to [Stu94] because it concerns, per the convention established in the beginning of this thesis, only CI inference forms over **local** CI statements. As pointed out in Section 1.2.2, a general theory of gaussoids can forgo global CI statements $(\mathsf{I}, \mathsf{J}|\mathsf{K})$ since local and global semigraphoids are in bijection. However, for inference rules with a bound on the number of antecedents, there is a major difference in allowing global statements. For example, the single global CI statement $(12, 34|5)$ corresponds to the local statements $(13|5)$, $(14|5)$, $(23|5)$ and $(24|5)$, which have a unique minimal gaussoid extension with 16 elements. This gaussoid is realizable, hence $(12, 34|5)$ is not the antecedent set of a non-trivial valid global inference rule for Gaussians, but this is not covered by our proof. Hence we have

Conjecture 4.66. All minimal gaussoid extensions of at most two global CI statements are realizable.

5

Geometry and complexity of CI models

The previous chapter investigated the structure of CI inference axioms as a totality and showed that there does not exist a finite axiomatization of all true inferences. We now turn to the complexity of the implication problem: how hard is it to decide for a given inference formula whether it is valid or not? This question has an obvious *complexity-theoretic* interpretation, but we also care about *algebraic* and *topological* complexity measures. If an inference rule is invalid, how complicated, in terms of the field extension degree over $\mathbb{Q}$, can the easiest witness matrix be? How does the space of all counterexamples to an invalid inference formula look like? Can every point on a CI model be continuously deformed into every other? How bad can the singularities of these models be? In this chapter a number of *universality theorems* are proved which give precise answers to these questions, which may be summarized as "it is as complicated as possible".

5.1 About universality theorems

As a general idea, a universality theorem states that some **object or construction** exhausts all the features it could possibly have, confined to **obvious limits** and perhaps modulo a notion of **equivalence** which blurs the concrete object but not the complexity of the features under consideration. The expression "Murphy's law" was used by Vakil [Vak06] for universality theorems in algebraic geometry, in the sense that for every bad trait the kind of object under consideration could possibly have, there exists one object which exhibits that trait. Take for example the complex projective realization space $\mathcal{R}$ of a matroid M of rank d on an n-element ground set. This is a geometric object defined by some combinatorial data: every point in $\mathcal{R}$ is an equivalence class of points on the Grassmannian $\mathrm{Gr}(n, d)$ of which the matroid M specifies which entries in its Plücker vector must vanish and which must not. Hence $\mathcal{R}$ can be viewed as a constructible subset of the affine space of $d \times n$-matrices defined by integer polynomials. Sturmfels and Mnëv independently proved that for every affine algebraic variety V over $\mathbb{C}$ there exists a matroid whose projective realization space is birationally equivalent to V [BS89, Theorem 4.30]. This is a universality theorem, where the objects are **realization spaces of matroids**, the obvious complexity bound is being first-order definable over $\mathbb{C}$ and therefore a **variety up to birational equivalence**. The theorem says that indeed every birational type of variety appears in the study of these *a priori special* types given by matroids.

5.1.1 A blueprint for universality. Matroid theory has more such universality theorems, e.g., concerning the computational complexity of testing whether a polynomial system has a solution [BS89, Theorem 2.2], or the stable equivalence type of primary basic semialgebraic sets [BLS+99, Theorem 8.6.6], [Ric97]. In the theory of games, a remarkable universality

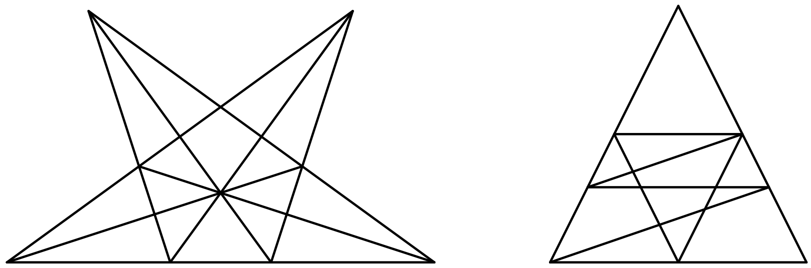

Figure 5.1: The Perles configuration on the left requires $\sqrt{5}$ to be realized over characteristic zero [Grü03, Section 5.5.3]. The affine configuration on the right requires $\sqrt{2}$ [Har00, Example 14.4.2].

result by Datta [Dat03] states that the set of totally mixed Nash equilibria of a three-person game can assume the stable isomorphy type of any real algebraic variety; see [Stu02, Chapter 6] for an introduction.

All of the mentioned theorems share the same core strategy: the act of solving a system of polynomial constraints is encoded in the combinatorial structure (of a matroid or a CI inference formula or oriented versions of these) in such a way that the solution space to the system is "equivalent" to the realization space. The kind of universality result obtained, and its strength, depends on the nature of the equivalence and the technical finesse of its proof. Homotopy equivalence of the solution set to the polynomial system and the realization space requires more effort and attention than proving that one is non-empty if and only if the other is. In this chapter, the technique of encoding of polynomial systems is applied to the Gaussian CI inference problem and multiple universality theorems are derived from it. The construction is divided into three steps:

1. An arbitrary polynomial system is brought into a normal form. The normal form should be simpler to encode but equivalent to the original system for the scope of the universality result. For our purposes variants of the *Shor normal form* are most useful.

2. The translation of the simpler polynomial system into CI constraints takes a detour through projective geometry. The classical *von Staudt constructions* described in Section 5.4.1 show how to transform a normalized system into an (oriented) incidence relation in the projective plane over a field $\mathbb{K}$. The space of geometric realizations of this incidence structure over $\mathbb{K}$ is equivalent in all desired ways to the space of $\mathbb{K}$-rational solutions to the normalized system and hence of the original system. Figure 5.1 shows (geometric realizations of) incidence relations which require a solution to $x^2 = 5$ or $x^2 = 2$ from a field to be realizable over it.

3. It remains to model *incidence relations* in the projective plane via Gaussian CI constraints. That is, the homogeneous coordinates of points and lines must be encoded into a symmetric matrix such that CI statements express which points lie on which lines (or left or right of which lines, in the oriented version). This is done in Section 5.4.2.

The Gaussian CI inference problem boils down to whether the semialgebraic set of counterexamples to an inference formula is empty or not, hence to ETR. If a counterexample exists, then a real-algebraic one exists by Tarski's transfer principle. This gives three upper bounds on the complexity of the inference problem. The subsequent Sections 5.5–5.7 obtain algebraic, complexity-theoretic and topological universality results for Gaussian CI as corollaries to the construction described above, proving that the inference problem is *hard* in multiple ways. The main results of this chapter are:

Theorem 5.34. For every real number field $\mathbb{K}$ there exists a CI model which has an $\mathbb{L}$-rational point, for any real algebraic extension $\mathbb{L}$ of $\mathbb{Q}$, if and only if $\mathbb{K} \subseteq \mathbb{L}$.

Theorem 5.38. The decision problem $\mathtt{GR}^+_{\mathbb{R}}$ is $\exists\mathbb{R}$-complete.

Theorem 5.43. For every primary basic semialgebraic set there is an **oriented** CI model over $\mathbb{R}$ which is stably equivalent to it.

This chapter focuses on positive realizability over fields between $\mathbb{Q}$ and $\mathbb{R}$ for its direct connection to statistics. The encoding of von Staudt constructions into CI constraints in Section 5.4 is carried out over general infinite fields because the proof is general enough and it allows to obtain a corollary about characteristic sets of algebraic Gaussians. The technique can even be observed to work for singular Gaussian models (see Remark 5.27).

5.2 The existential theory of the reals and varieties

A fundamental decision problem in computational geometry (and algebraic statistics) is $\mathtt{ETR}$: given an existential first-order sentence in the language of ordered rings, decide if it is satisfiable over $\mathbb{R}$. The question of whether a Gaussian CI model (oriented or not), and in fact any statistical model definable by polynomial constraints on real parameters or probabilities, is empty or not is a special case of $\mathtt{ETR}$, as are many decision problems in geometry. Following [SŠ17, Section 4], the equivalence class of $\mathtt{ETR}$ modulo polytime Karp-reductions is denoted by $\exists\mathbb{R}$. To complete this definition, the coding length of a formula has to be explained: it is simply the number of symbols required to write down the first-order sentence in the language of ordered rings. We assume that variables are numbered consecutively $x_1, \ldots, x_n$ — a convention which can easily be established with a polytime transformation of any given formula. Moreover, each variable may be counted as a single symbol or it may count as $\lceil \log n \rceil$ symbols — since a variable, if it occurs in the formula, contributes to its length, the $\lceil \log n \rceil$ convention incurs only polynomial overhead. By the same reasoning, we may omit the existential quantifiers in front of the formula and assume that all occuring variables are free. This reduces $\exists\mathbb{R}$ to the question of whether a boolean combination of polynomial constraints with integer coefficients has a solution in a real-closed field.

Remark 5.1. This definition of the length of a system of polynomial constraints via its representation in the language of ordered rings is standard in complexity theory, although it is not always given in these words. Other common representations of polynomials include *straight line programs* and *algebraic circuits* [AB09, Chapter 16]. It is easy to see that all these representations are polytime-equivalent if enough "temporary variables" are introduced.

It should be noted that, despite only having constants 0 and 1 in our language, coefficients of polynomials need not be written in the naïve unary encoding as $c = (\cdots(1+1)+\cdots)+1$, which is wasteful (and algorithms with polynomial runtime in the unary coding length of numbers are termed *pseudo-polynomial* and not considered efficient). By introducing more variables, the coefficients can be written in binary: if $c = \sum_{i=0}^{n} b_i 2^i$, for $b_i \in \{0, 1\}$, then

$$t_2 = 1+1, \;\; t_4 = t_2 \cdot t_2, \;\; t_8 = t_4 \cdot t_2, \;\; \ldots, \;\; t_{2^n} = t_{2^{n-1}} \cdot t_2,$$
$$c = b_0 t_0 + b_1 t_1 + \cdots + b_n t_n$$

makes the coefficient c available to other equations in the system and requires an amount of symbols to write down which is of polynomial order in $\log_2 c$. This allows for an efficient encoding of polynomial systems (even taking into account the sparsity of the system) which is close to the representation of polynomials in computer algebra systems in practice.

Remark 5.2. The complexity of practically-minded algorithms for semialgebraic sets in real algebraic geometry [BPR06, Chapter 8] are **not** measured in terms of the length of a formula which specifies the set. Instead, characteristic quantities of the formula, like the number of variables, the number of polynomials and their maximum degree are used to give more precise runtime bounds. These bounds reveal that the complexity of available algorithms is influenced by these quantities on different orders of magnitude: the feasibility of a system of s equations of degree at most d in k variables can be checked in time $(sd)^{\mathcal{O}(k)}$ [BPR06, Theorem 13.13] (assuming that arithmetic on the coefficients has no cost). So, adding many new variables has disastrous consequences in practice; adding more low-degree polynomials less so. It has been shown, however, that when measuring `ETR` inputs with their formula length, the problem is in **PSPACE** [Can88].

The following lemma shows that $\exists\mathbb{R}$ is really the feasibility problem for real algebraic varieties, the question of whether a set of integer polynomials has a common real root.

Lemma 5.3. The special case of `ETR` for conjunctions of equations is $\exists\mathbb{R}$-complete.

Proof. By applying the Tseitin transform, the formula can be put into conjunctive normal form with only polynomial overhead. The algorithm in Section 2.3 is stated only for formulas on boolean variables. The boolean variables are replaced by relations on polynomial expressions in this case. The new boolean variables Z introduced in the process can be modeled by introducing new real variables z and using the predicates $z = 0$ for boolean variables. Since each variable is replicated by Tseitin only a constant number of times, the transformation remains polytime. Then, using the following transformations, general polynomial constraints are replaced by equations:

— replace $f \neq 0$ by $yf = 1$.
— replace $\pm f > 0$ by $\pm y^2 f = 1$.
— replace $\pm f \geq 0$ by $\pm f = y^2$.

These rules introduce one new variable y per constraint and can be performed in polynomial time on the input formula. The previous step also removes negations from the CNF. Finally, the disjunctive terms can be dissolved into conjunctions of equations via

— replace $\bigvee_i [f_i = 0]$ by $\bigwedge_i [y_i = f_i] \wedge [\prod_i y_i = 0]$.

This introduces more variables to represent the product of the polynomials f_i in polynomial space and time. The general formula has been turned into a conjunction of equations without changing its satisfiability, which finishes the proof. □

Remark 5.4. This construction works for arbitrary **euclidean** ordered fields. The two inequality transformations require that non-negativity can be expressed algebraically as being a square. A variant reduces the existential theory of any unordered field $\mathbb{K}$ in polynomial time to the question of whether a $\mathbb{Z}$-defined variety has a $\mathbb{K}$-rational point.

5.3 Stable equivalence and Shor's normal form

Stable equivalence. The other notion of equivalence is geometric: it is *stable equivalence* of semialgebraic sets. This exposition is guided by Section 2.5 in Richter-Gebert's book [Ric97] on universality theorems for realization spaces of polytopes. It shows that for polytopes in dimension $d \geq 4$, they can attain the stable equivalence type of any primary basic semialgebraic set. According to a comment in [BLS+99, Appendix A.1], it was for some time unclear what the right definition of stable equivalence should be. Richter-Gebert eventually gave a natural definition which was stronger than previous versions and still applicable — yielding the desired consequences of the "ideal" definition of stable equivalence. Unfortunately, we have to deviate from Richter-Gebert's version again in this work for technical reasons (see Remarks 5.9 and 5.11). What is essentially wanted from a stable equivalence is that equivalent semialgebraic sets should have the same topological and algebraic features, more precisely the same homotopy type and the same answers to questions about existence of $\mathbb{K}$-rational points for all real number fields $\mathbb{K}$.

Definition 5.5. Two semialgebraic sets V and W are *rationally equivalent* if there is an inverse pair of (euclidean) homeomorphisms between them which are effected by rational functions with rational coefficients.

Definition 5.6. Let $W \subseteq \mathbb{R}^{n+m}$ and $\pi(W) = V \subseteq \mathbb{R}^n$ be its projection onto the first n coordinates. This is a *stable projection* if every fiber $\pi^{-1}(v)$ for $v \in V$ is a convex semialgebraic set which can be written as

$$\pi^{-1}(v) = \left\{ (v, w', w'') \in \mathbb{R}^{n+(m'+m'')} : \phi_i(v, w') > 0 \text{ and } w''_j = \psi_j(v) \right\},$$

where $\phi_i \in \mathbb{Q}(v, w')$ and $\psi_j \in \mathbb{Q}(v)$ are two sets of rational maps which are well-defined on V.

Remark 5.7. If $V = \pi(W)$ is a stable projection with fiber-defining maps ϕ_i and ψ_j, then this implies a description of $W = \{ (v, w', w'') \in \mathbb{R}^{n+(m'+m'')} : v \in V, \phi_i(v, w') > 0, w''_j = \psi_j(v) \}$. Thus, after clearing denominators, W is a (primary) basic semialgebraic set if V is. The same holds for rational equivalences.

Definition 5.8. Two semialgebraic sets V and W are *stably equivalent* if they are in the same class of the equivalence relation generated by rational equivalences and stable projections.

Remark 5.9. Richter-Gebert's version [Ric97, p. 21] differs in the notion of stable projection. While our fiber-defining equations $w''_j = \psi_j(v)$ are direct assignments to some components w'' which are projected away as a rational function of the image coordinates v, Richter-Gebert allows general constraints $\psi_j(v, w) = 0$ for polynomials $\psi_j \in \mathbb{Q}[v, w]$ which are **linear** in the w-coordinates. He requires additionally that the inequalities $\phi_i(v, w) > 0$ are linear in w as well, and so all fibers $\pi^{-1}(v)$ are relatively open, rational polyhedra whose inequalities vary polynomially in the image point v. Polyhedra are too restrictive for the projections we have to apply in Section 5.7, but relative interiors of **spectrahedra** are sufficient (recall Section 2.4.2). Both are convex semialgebraic sets whose defining equations are affine-linear.

The purpose of stable equivalence is to serve as a common refinement of multiple other equivalence notions, while also having a definition which is reasonably simple to check. The desired properties of a stable equivalence presented in [Ric97, Lemma 2.5.2] hold:

Definition 5.10. (1) Two semialgebraic sets V and W have the same *homotopy type* if there are continuous maps $f : V \to W$ and $g : W \to V$ such that fg and gf are homotopic to the respective identity maps on V and W. That is, there exists a continuous function $h : V \times [0, 1] \to V$ such that $h(\cdot, 0) = fg$ and $h(\cdot, 1) = \mathrm{id}_V$, and analogously for gf on W. (2) V and W have the same *algebraic number type* if for every real algebraic number field $\mathbb{K}$ one set has a $\mathbb{K}$-rational point if and only if the other does.

Remark 5.11: On the definition in Richter-Gebert's book. The definition of stable projection presented in [Ric97] is erroneous because his Lemmma 2.5.2 claims (without proof and indeed incorrectly) that it preserves homotopy equivalence. The following simple counterexample was pointed out by Andreas Kretschmer: consider the union of the w-axis with the real hyperbola in $\mathbb{R}^2$, i.e., $W = \left\{ (v, w) \in \mathbb{R}^2 : v(vw - 1) = 0 \right\}$. The projection down to the v-coordinate is stable in the sense of [Ric97, p. 21], because the fibers are described by the equation $v^2 w = v$ which is affine-linear in w. However, W has three path-connected components in $\mathbb{R}^2$ whereas the projection $\mathbb{R}^1$ only has one. This contradicts homotopy equivalence. The reason for this mismatch seems to lie in allowing inhomogeneous linear conditions on w. Other sources, for example the joint report of Richter-Gebert with Ziegler [RZ95] or the presentation in [BLS+99, Appendix A.1] uses homogeneous linear functionals. Other notions of stable equivalence also require the linear spaces in the fiber definition to be equidimensional [BS89, Section 6.3]. Equidimensionality is also featured in the equivalence in the main Theorem of [Gün96] where a smooth manifold appears in the direct product. Notice that in the hyperbola example the rank of the fiber-defining linear system $v^2 w = v$ increases in every neighborhood of the point $v = 0$. Definition 5.6 uses the combination of affine equalities with a **zero-dimensional** linear space for the fibers. This is sufficient for our applications and it makes the proof of homotopy equivalence relatively easy to repair.

We also remark that the proof of [Ric97, Lemma 2.5.2 (ii)] is wrong in claiming that for every $v \in \pi(W)$ under a stable projection π a point $(v, w) \in \pi^{-1}(v)$ with $w \in \mathbb{Q}^m$ can be found. The reasoning assumes that the affine-linear functions are homogeneous. A counterexample to the proof is the stable projection of $W = \{ (\sqrt{2}, \sqrt{2}) \in \mathbb{R}^2 \}$ whose fibers are defined by the affine equation $v = w$. However, it is easy to see that, although the w-coordinates need not be rational, their algebraic number type is bounded by that of v.

Lemma 5.12. Let $V = \pi(W)$ be a stable projection of basic semialgebraic sets. Then there exists a continuous section $s : V \hookrightarrow W$ of π.

Proof. Let ϕ_i and ψ_j be the fiber-defining rational maps from the definition of stable projection, so that $\pi^{-1}(v) = \{ \phi_i(v, w') > 0, w''_j = \psi_j(v) \}$. Fix a $v \in V$ and pick a solution $w_v = (w'_v, w''_v)$ to the fiber-defining system. Since the w'' depend rationally on v and the inequalities $\phi_i(v, w') > 0$ are strict, this particular solution can be extended to a continuous local section s_v of π on an open subset U_v of $\mathbb{R}^n$. This defines an open cover of V by $(U_v)_{v \in V}$ of local sections. By [Bre93, Section I.12] and because basic semialgebraic sets are paracompact, there exists a partition of unity subordinate to this cover of V, i.e., continuous functions $\rho_v : V \to [0, 1]$ such that

— $\operatorname{supp} \rho_v \subseteq U_v$,
— for every $\widetilde{v} \in V$ all but finitely many $\rho_v(\widetilde{v})$ vanish,
— $\sum_{v \in V} \rho_v(\widetilde{v}) = 1$ for all $\widetilde{v} \in V$;

see also [Spi65, Theorem 3-11]. A global section $s : V \to W$ is obtained by

$$s(\widetilde{v}) \coloneqq \sum_{v \in V} \rho_v(\widetilde{v}) s_v(\widetilde{v}).$$

This expression is well-defined because the summation is finite for every $\widetilde{v}$ and outside of the domain of s_v, which is U_v, the function ρ_v vanishes. The function is continuous because ρ_v and s_v are continuous. Lastly, the sum is a convex combination of $s_v(\widetilde{v})$ over finitely many $v \in V$. By definition s_v satisfies $s_v(\widetilde{v}) \in \pi^{-1}(\widetilde{v})$ and $\pi(s_v(\widetilde{v})) = \widetilde{v}$. Since the fiber $\pi^{-1}(\widetilde{v})$ is convex, this convex combination remains inside it. Thus s is a section of π. □

Lemma 5.13. Stable equivalence preserves the homotopy and algebraic number type.

Proof. A rational equivalence f is by definition a homeomorphism, so in particular the homotopy type is preserved. Since f is a rational function with rational coefficients, every algebraic number α over $\mathbb{Q}$ satisfies $\mathbb{Q}(f(\alpha)) \subseteq \mathbb{Q}(\alpha)$. Since f^{-1} is a rational equivalence as well, this is indeed an equality of field extensions, which implies invariance of the algebraic number type between V and W. This shows the claim for every step in a chain of stable equivalence maps which is a rational equivalence. Now assume that we have a stable projection $\pi : \mathbb{R}^{n+(m'+m'')} \supseteq W \twoheadrightarrow V \subseteq \mathbb{R}^n$ with the fiber-defining maps $\phi_i \in \mathbb{Q}(v, w')$ and $\psi_j \in \mathbb{Q}(v)$:

Homotopy type: Let s be the section from Lemma 5.12. Then $\pi \circ s = \mathrm{id}_V$ proves one part of the homotopy equivalence. For the other part consider $s \circ \pi$ mapping (v, w) (we write $w = (w', w'')$) to $(v, w(v))$ with some continuous function $w(v)$ mapping into the fiber $\pi^{-1}(v)$. Since each fiber is convex, there is a canonical uniform contraction of $\pi^{-1}(v)$ to the single point $\{(v, w(v))\}$ by moving each point (v, w) on a straight line to $(v, w(v))$. This defines the homotopy $h(v, w, t) := (v, (1 - t)w + tw(v))$ which is id_W at $t = 0$ and $s \circ \pi$ at $t = 1$. In fact, the embedding of $s(V) \subseteq W$ is a strong deformation retract of W.

Algebraic number type: One direction is obvious: if V contains no $\mathbb{K}$-rational point, then W with additional coordinates cannot contain one either. In the opposite direction, recall that the equations ψ_j directly assign to w'' some $\mathbb{Q}$-defined rational functions of v. As in the beginning of the proof, this implies that w''_j belong to the field extension of $\mathbb{Q}$ generated by the components of v. The remaining variables w' are subject to strict inequalities, which have a solution since the fiber is non-empty. Thus, there is an open w'-ball around this solution in the fiber and we may pick a solution $w' \in \mathbb{Q}^{m'}$. $\square$

Remark 5.14. Homotopy type in particular preserves the number of connected components of a topological space. A universality theorem that recovers stable equivalence types therefore shows that the universal spaces can have arbitrarily many connected components. **Ringel's isotopy question** [Rin56] in computational geometry asked whether every two geometric realizations of a combinatorial type of pseudoline arrangement can be continuously transformed into each other. This is equivalent to the realization space of the combinatorial arrangement to be path-connected. Mnëv's universality theorem [Mnë88] for oriented matroids provides a strong negative answer to this question. See also the explanation and a simple counterexample in [Ric99b].

Shor normal form. Lemma 5.3 showed how to reduce the general decision problem for existential first-order formulas over a euclidean ordered field to the special case of whether a variety has a point. A similar construction works for existential first-order sentences in the language of rings over fields. Every polynomial equation $f = 0$ with $f \in \mathbb{Z}[t_1, \ldots, t_k]$ can be successively decomposed into a system of very easy polynomial equations, namely addition constraints $x = y + z$ and multiplication constraints $x = y \cdot z$. This procedure introduces many auxiliary variables and requires a distinguished variable containing the constant 1. For example, the equation $x^2 = 2$ might be written as the system

$$\begin{aligned} t_1 &= 1, & t_1 &= t_1 + t_0, && \text{(define 1 and 0)} \\ t_2 &= t_1 + t_1, & t_{x^2} &= t_x \cdot t_x, && \text{(write 2 and } x^2\text{)} \\ t_{x^2} &= t_{x^2-2} + t_2, & t_{x^2-2} &= t_0 + t_0. && \text{(force } x^2 = 2\text{)} \end{aligned}$$

This construction preserves the algorithmic complexity of the problem (modulo polytime reductions) and also the algebraic complexity of its solution set (the algebraic number type), but it but it does not preserve topological properties of the solution set. For example, the inequality $x^2 \geq -1$ would be lifted to $\exists z : x^2 + 1 = z^2$. The solution set $\{\, z = \pm\sqrt{x^2 + 1}\,\}$ has two connected components which both project to the solution set $\mathbb{R}^1$ of the original inequality.

To preserve topological properties, a more refined normal form is necessary. Shor, in reviewing Mnëv's universality theorem in [Sho91], introduced a suitable one:

Theorem 5.15: Shor normal form. For every primary basic semialgebraic set $Z = \{\, f_i = 0, g_j > 0 \,\}$ over $\mathbb{R}$ there is a polynomial system (in possibly more variables) of the form

$$t_{k_p} = t_{i_p} + t_{j_p} \quad \text{or} \quad t_{k_p} = t_{i_p} \cdot t_{j_p} \quad \text{for finitely many } p,$$
$$1 = t_1 < t_2 < \cdots < t_n,$$

whose solution set is stably equivalent to Z. This system is a *Shor normal form* of Z and it can be computed in polytime from the f_i and g_j.

Proof sketch. See [Sho91, Section 4] for a complete proof. The idea is to introduce variables for every polynomial term which appears in the primary basic polynomial system. This reduces the system to addition and multiplication inequalities and the comparison of variables. In the second step, a new variable $a \gg 0$ is introduced and all variables are formally shifted by $t_i \mapsto t_i + a$ to ensure that they are strictly greater than 1. In the third phase, another variable $b \gg 0$ is introduced and required to be large enough that by replacing $t_i \mapsto t_i + b^i$ the variables become a priori totally ordered. In both of these phases, the equations can be rewritten according to the variable substitutions. These transformations introduce many variables which are polynomial functions of preexisting variables, and the two variables a and b which are bounded from below. Projecting them away results in fibers which are essentially right-infinite intervals (a_0, ∞) and (b_0, ∞), respectively, which can be defined by strict polynomial inequalities in the original variables. □

In a Shor normal form every equation is a direct addition or multiplication, as before, but strict inequalities are preserved in the form of a total order on the variables. In summary, the easier "variety normal form" from Lemma 5.3 is sufficient to capture the algorithmic and algebraic properties of polynomial systems. To capture the topological properties, Shor's normal form including a particular type of strict inequality has to be employed.

5.4 Solving equations with CI constraints

5.4.1 Polynomial systems as ruler constructions. The universality theorems for matroids rest on an encoding of arbitrary polynomial systems in the bases and non-bases of a matroid. This was first demonstrated by MacLane [Mac36] using the *von Staudt constructions* in projective geometry (see [Ric11, Section 5.6]). We first treat the case of polynomial equation systems only. This gives algorithmic and algebraic equivalence. The finer treatment of topological universal via the Shor normal form is contained in Section 5.7. Throughout this section $\mathbb{K}$ denotes any field and $\mathbb{PK}^2 \cong (\mathbb{K}^3 \setminus \{0\}) / \mathbb{K}^\times$ its projective plane; see [Ric11].

Definition 5.16. The *standard projective basis* consists of the infinite point on the x-axis $\boldsymbol{\infty}_\mathsf{x} = [1:0:0]$, the infinite point on the y-axis $\boldsymbol{\infty}_\mathsf{y} = [0:1:0]$, the origin $\mathbf{0} = [0:0:1]$ and the point of units $\mathbf{1} = [1:1:1]$.

From these points, the x- and y-axes ℓ_x and ℓ_y, unit points on the axes $\mathbf{1}_\mathsf{x}$ and $\mathbf{1}_\mathsf{y}$ and the line at infinity ℓ_∞ can be constructed, which complete the framework in which ruler constructions are carried out. The standard basis has favorable properties for the constructions in the next section, notably its shape can be prescribed easily using CI constraints, which is why we insist on it.

Definition 5.17. A *ruler construction* over a field $\mathbb{K}$ is a finite list of instructions which constructs a set of points and lines in $\mathbb{PK}^2$ from a given set of points including the standard projective basis using the computational primitives of (a) *joining* two already constructed, distinct points to form the line through them, and (b) *meeting* two already constructed, distinct lines to form their intersection point. The construction algorithm may receive parameters in the form of *indeterminate points* which are placed on the x-axis ℓ_x.

Ruler constructions are required to be deterministic: by stipulating the distinctness of joined points and met lines in $\mathbb{PK}^2$, the resulting line or point is uniquely defined as the one-dimensional space of solutions to two independent linear equations in $\mathbb{K}^3$. For instance,

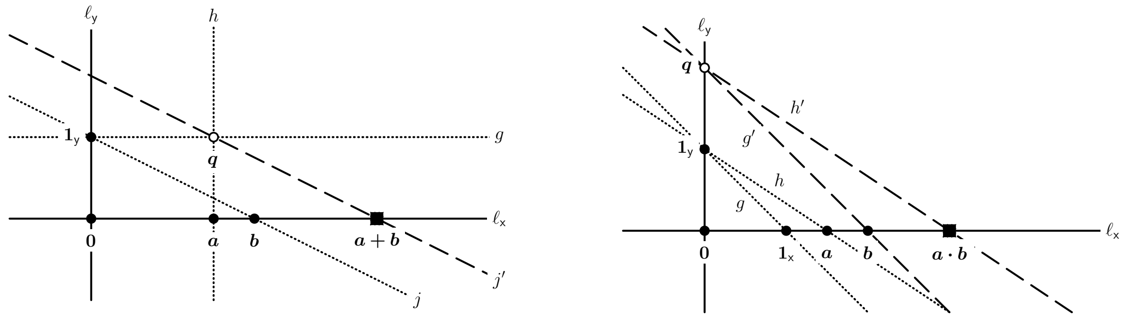

Figure 5.2: Von Staudt constructions in two affine pictures. The solid points are given, the hollow ones are helper points in the construction of the square target points. The axes are displayed as solid lines, helper lines are dotted and the dashed lines, which are parallel to the dotted ones, yield the target points.

the line ℓ through two distinct points p, p' is given by $\langle p, \ell \rangle = 0$ and $\langle p', \ell \rangle = 0$. Usage of the indeterminate points and all objects constructed from them is permitted as long as all joins and meets are provably between distinct objects in every instantiation of the indeterminates. In this case, the join ℓ of the distinct points p, p' can be immediately computed by the cross product

$$[p^{\mathsf{x}} : p^{\mathsf{y}} : p^{\mathsf{z}}] \times [p'^{\mathsf{x}} : p'^{\mathsf{y}} : p'^{\mathsf{z}}] := \left[\det\begin{pmatrix} p^{\mathsf{y}} & p'^{\mathsf{y}} \\ p^{\mathsf{z}} & p'^{\mathsf{z}} \end{pmatrix} : -\det\begin{pmatrix} p^{\mathsf{x}} & p'^{\mathsf{x}} \\ p^{\mathsf{z}} & p'^{\mathsf{z}} \end{pmatrix} : \det\begin{pmatrix} p^{\mathsf{x}} & p'^{\mathsf{x}} \\ p^{\mathsf{y}} & p'^{\mathsf{y}} \end{pmatrix}\right].$$

The same operation computes, dually, the coordinates of the meet of two distinct lines.

Construction Problem. Given a system $\{f_1, \ldots, f_r\}$ of polynomials $f_i \in \mathbb{Z}[t_1, \ldots, t_k]$, construct with a ruler, starting from the standard projective basis and an indeterminate point $\boldsymbol{t_i} = [t_i : 0 : 1]$ for each unknown t_i, the points $\boldsymbol{f_i} = [f_i(t_1, \ldots, t_k) : 0 : 1]$.

The von Staudt constructions implement precisely addition and multiplication of points on the x-axis, which is enough to model arbitrary integer polynomial expressions. Before describing these algorithms, we construct a larger projective framework out of the standard basis, containing points and lines which are used in both:

Framework:

1. $\ell_{\mathsf{x}} := \mathbf{0} \times \infty_{\mathsf{x}} = [0 : 1 : 0]$
2. $\ell_{\mathsf{y}} := \mathbf{0} \times \infty_{\mathsf{y}} = [1 : 0 : 0]$
3. $\ell_\infty := \infty_{\mathsf{x}} \times \infty_{\mathsf{y}} = [0 : 0 : 1]$
4. $\ell_{1\mathsf{x}} := \mathbf{1} \times \infty_{\mathsf{x}} = [0 : -1 : 1]$
5. $\ell_{1\mathsf{y}} := \mathbf{1} \times \infty_{\mathsf{y}} = [-1 : 0 : 1]$
6. $\mathbf{1}_{\mathsf{x}} := \ell_{1\mathsf{y}} \times \ell_{\mathsf{x}} = [1 : 0 : 1]$
7. $\mathbf{1}_{\mathsf{y}} := \ell_{1\mathsf{x}} \times \ell_{\mathsf{y}} = [0 : 1 : 1]$

Figure 5.2 contains pictures of the von Staudt constructions for addition and multiplication of indeterminate points in the affine xy-plane by projective ruler constructions from the standard basis. The pictures join points, meet lines and construct the parallel to a line through another point. This last affine operation can be performed by the projective ruler using the line at infinity not pictured here. Full descriptions of these classical constructions are given in [Ric11, Section 5.6] and with emphasis on matroids (over skew fields) in [KPY20]. Here we give the algorithms with indeterminates $\boldsymbol{a} = [a : 0 : 1]$ and $\boldsymbol{b} = [b : 0 : 1]$ using cross products and using the same notation as in Figure 5.2:

Addition:

1. $g \coloneqq \mathbf{1}_{\mathsf{y}} \times \boldsymbol{\infty}_{\mathsf{x}} = [0 : -1 : 1]$
2. $h \coloneqq \boldsymbol{a} \times \boldsymbol{\infty}_{\mathsf{y}} = [-1 : 0 : a]$
3. $\boldsymbol{q} \coloneqq g \times h = [a : 1 : 1]$
4. $j \coloneqq \boldsymbol{b} \times \mathbf{1}_{\mathsf{y}} = [-1 : -b : b]$
5. $\boldsymbol{\infty}_j \coloneqq j \times \ell_\infty = [-b : 1 : 0]$
6. $j' \coloneqq \boldsymbol{q} \times \boldsymbol{\infty}_j = [-1 : -b : a + b]$
7. $\boldsymbol{a} + \boldsymbol{b} \coloneqq j' \times \ell_{\mathsf{x}} = [a + b : 0 : 1]$

Multiplication:

1. $g \coloneqq \mathbf{1}_{\mathsf{x}} \times \mathbf{1}_{\mathsf{y}} = [-1 : -1 : 1]$
2. $h \coloneqq \boldsymbol{a} \times \mathbf{1}_{\mathsf{y}} = [-1 : -a : a]$
3. $\boldsymbol{\infty}_g \coloneqq g \times \ell_\infty = [-1 : 1 : 0]$
4. $\boldsymbol{\infty}_h \coloneqq h \times \ell_\infty = [-a : 1 : 0]$
5. $g' \coloneqq \boldsymbol{\infty}_g \times \boldsymbol{b} = [1 : 1 : -b]$
6. $\boldsymbol{q} \coloneqq g' \times \ell_{\mathsf{y}} = [0 : b : 1]$
7. $h' \coloneqq \boldsymbol{q} \times \boldsymbol{\infty}_h = [1 : a : -a \cdot b]$
8. $\boldsymbol{a} \cdot \boldsymbol{b} \coloneqq h' \times \ell_{\mathsf{x}} = [a \cdot b : 0 : 1]$

Lemma 5.18. Given the standard basis, the von Staudt constructions solve the Construction Problem. □

This very analytic treatment of the construction is required to observe the following subtle point which will be important in Section 5.4.2:

Lemma 5.19. All meet and join operations in the von Staudt construction are between distinct points and lines, independently of the positions of the indeterminates $\boldsymbol{t}_i$ on the x-axis. Moreover, for every point and line needed in the construction, one homogeneous coordinate can be given which is non-zero, also independently of the indeterminates. □

Remark 5.20. The von Staudt constructions show how to model polynomial expressions using incidence geometry in the projective plane over a field. Once the value of a polynomial $f(t_1, \dots, t_k)$ is constructed at a point $\boldsymbol{f}$ on the x-axis, the condition that $t_1, \dots, t_k$ are a root of f can be imposed by constraining the point $\boldsymbol{f}$ to also lie on ℓ_{y}. This shows that the ability to set variables to 0 or 1 (for the projective basis) and to require the vanishing of the inner product on $\mathbb{K}^3$ for arbitrary pairs of triples of variables, which is a single homogeneous quadratic trinomial in six variables, is enough to reach the full algebraic complexity of varieties over $\mathbb{K}$. In the case of rank-3 matroids, the variables only denote coordinates of points and not lines. Matroids express collinearity of points instead of point-line incidences directly. In non-basis constraints imposed on a point configuration by a matroid, the lines are constructed ad hoc: the generic 3×3 determinant decomposes into the inner product of one of the points and the cross product of the other two:

$$\det \begin{pmatrix} p^{\mathsf{x}} & q^{\mathsf{x}} & r^{\mathsf{x}} \\ p^{\mathsf{y}} & q^{\mathsf{y}} & r^{\mathsf{y}} \\ p^{\mathsf{z}} & q^{\mathsf{z}} & r^{\mathsf{z}} \end{pmatrix} = p^{\mathsf{x}} \det\begin{pmatrix} q^{\mathsf{y}} & r^{\mathsf{y}} \\ q^{\mathsf{z}} & r^{\mathsf{z}} \end{pmatrix} - p^{\mathsf{y}} \det\begin{pmatrix} q^{\mathsf{x}} & r^{\mathsf{x}} \\ q^{\mathsf{y}} & r^{\mathsf{y}} \end{pmatrix} + p^{\mathsf{z}} \det\begin{pmatrix} q^{\mathsf{x}} & r^{\mathsf{x}} \\ q^{\mathsf{y}} & r^{\mathsf{y}} \end{pmatrix} = \langle p, q \times r \rangle.$$

Instead of a quadratic trinomial, this is a determinant of degree 3, which also yields algebraic universality. The (usually algorithmic) complexity of sets expressible using combinations of specific functions is studied in the theory of **constraint satisfaction problems (CSPs)**; see [MS21b] for geometric CSPs.

5.4.2 Ruler constructions as CI constraints. To model a ruler construction as CI constraints, we work over a ground set $\mathsf{N} = \mathsf{PLE}$, which decomposes into sets $\mathsf{P} = \{ \mathsf{p}_1, \mathsf{p}_2, \dots \}$ and $\mathsf{L} = \{ \mathsf{l}_1, \mathsf{l}_2, \dots \}$ for labeling the points and lines which are used during the algorithm, respectively, and $\mathsf{E} = \{ \mathsf{x}, \mathsf{y}, \mathsf{z} \}$ which indexes the homogeneous coordinates of the points and lines. Instead of implementing the join and meet primitives via collinearity of points, as matroids do, or by the cross product, we use the following scalar product interpretation of almost-principal minors obtained by a Schur complement with respect to the E-block:

$$\Sigma[\mathsf{pl}|\mathsf{E}] = \Sigma[\mathsf{E}] \left(\Sigma_{\mathsf{pl}} - \Sigma_{\mathsf{p,E}} \Sigma_{\mathsf{E}}^{-1} \Sigma_{\mathsf{E,l}} \right) \stackrel{!}{=} 0. \qquad (\measuredangle)$$

Namely, vanishing of this almost-principal minor is equivalent to assigning the value of $\Sigma_{\mathsf{p,E}} \Sigma_{\mathsf{E}}^{-1} \Sigma_{\mathsf{E,l}}$, which is a non-degenerate symmetric bilinear form in the vectors $\Sigma_{\mathsf{p,E}}$ and $\Sigma_{\mathsf{l,E}}$, to the **entry** Σ_{pl}. Given homogeneous coordinates of a point $p = \Sigma_{\mathsf{p,E}}$ and of a line $\ell = \Sigma_{\mathsf{l,E}}$, the incidence $p \in \ell^\perp$ is equivalent to the vanishing of the standard scalar product $\langle p, \ell \rangle$ (which appears in the expression in ($\measuredangle$) if $\Sigma_{\mathsf{E}} = \mathbb{1}_{\mathsf{E}}$). To construct the line ℓ_{pq} joining two already constructed, distinct points p, q labeled by $\mathsf{p}, \mathsf{q} \in \mathsf{P}$, introduce a new variable $\mathsf{l_{pq}}$ into the set L and require the incidence of the points p and q to the new line $\mathsf{l_{pq}}$.

The complete encoding of the von Staudt constructions of a polynomial system into a set of CI constraints is given in Definition 5.21 below. Up to some implementation details, the basic ideas behind this definition are:

- The homogeneous coordinates of the point indexed by $\mathsf{p} \in \mathsf{P}$ are stored in the entries Σ_{pe} of a model matrix Σ, for $\mathsf{e} \in \mathsf{E} = \{\mathsf{x}, \mathsf{y}, \mathsf{z}\}$. Likewise for lines $\mathsf{l} \in \mathsf{L}$.
- For every pair of distinct points and/or lines $\mathsf{a}, \mathsf{b} \in \mathsf{PL}$, we impose the CI statement $(\mathsf{ab}|\mathsf{xyz})$ in order to store the scalar product of their homogeneous coordinates in the entry Σ_{ab}. This scalar product is with respect to the inverse block matrix Σ_{E}^{-1}, according to ($\measuredangle$).
- The desired orthogonalities between $\mathsf{p} \in \mathsf{P}$ and $\mathsf{l} \in \mathsf{L}$ which assert incidence relationships can then be prescribed with CI constraints $(\mathsf{pl}|)$.

Definition 5.21. Let $F = \{f_1, \dots, f_r\} \subseteq \mathbb{Z}[t_1, \dots, t_k]$. Consider the von Staudt construction of these polynomials making reference to points labeled $\mathsf{P} = \{\mathsf{t}_1, \dots, \mathsf{t}_k, \mathsf{f}_1, \dots, \mathsf{f}_r, \mathsf{p}_1, \dots, \mathsf{p}_n\}$ and lines labeled $\mathsf{L} = \{\mathsf{l}_1, \dots, \mathsf{l}_m\}$, where the t_i represent the indeterminate points and f_i represent the values of the f_i in the construction. Define a set of CI constraints $\widetilde{\mathcal{I}}(F)$ over the ground set PLE consisting of:

($\mathcal{I}$.i) $(\mathsf{pe}|)$ or $\neg(\mathsf{pe}|)$ for all points p corresponding to the standard projective basis and $\mathsf{e} \in \mathsf{E}$, depending on whether the e-coordinate of the point is zero or not.

($\mathcal{I}$.ii) $(\mathsf{pq}|)$ or $\neg(\mathsf{pq}|)$ for all points p, q of the standard projective basis depending on whether $\langle p, q \rangle = 0$ or not, respectively.

($\mathcal{I}$.iii) $(\mathsf{ty}|)$ and $\neg(\mathsf{tz}|)$ for indeterminate points $\mathsf{t} = \mathsf{t}_1, \dots, \mathsf{t}_k$.

($\mathcal{I}$.iv) $\neg(\mathsf{ae}|)$ for each $\mathsf{a} \in \mathsf{PL}$ and one of the coordinates $\mathsf{e} \in \mathsf{E}$ on which the point or line labeled a is non-zero, which can be deduced by Lemma 5.19.

($\mathcal{I}$.v) $(\mathsf{ab}|\mathsf{xyz})$ for all distinct $\mathsf{a}, \mathsf{b} \in \mathsf{PL}$.

($\mathcal{I}$.vi) $(\mathsf{pl}|)$ for any incidence relationship between $\mathsf{p} \in \mathsf{P}$ and $\mathsf{l} \in \mathsf{L}$ which is required to express a join or meet operation of the construction.

Let $\widetilde{\mathcal{R}}^\bullet_{\mathbb{K}}(F) \coloneqq \mathcal{R}^\bullet_{\mathbb{K}}(\widetilde{\mathcal{I}}(F))$ be the model (either algebraic or positive) of the above constraints over the (ordered) field $\mathbb{K}$.

Figure 5.3 shows the generic matrix satisfying constraint type ($\mathcal{I}$.v). The constraints $\widetilde{\mathcal{I}}(F)$ emulate the incidence relations behind the von Staudt construction. Every matrix which satisfies $\widetilde{\mathcal{I}}(F)$ gives values to the parameters $t_1, \dots, t_k$ and all other points and lines such that the same incidence relations hold, which forces the f_i to assume the evaluation of $f_i(t_1, \dots, t_k)$ up to various scalings. The caveat, however, is that each model starts the construction with coordinates of points which are not necessarily the standard basis and executes the ruler construction with a possibly non-standard notion of "incidence" which comes from ($\measuredangle$) and constraint type ($\mathcal{I}$.v): the linear system defining incidence $p \in \ell^\perp$ switches from $\langle p, \ell \rangle = 0$ to $\langle\!\langle p, \ell \rangle\!\rangle = 0$, where $\langle\!\langle \cdot, \cdot \rangle\!\rangle$ is a non-degenerate symmetric bilinear form defined by the inverse of Σ_{E} in the matrix Σ thought of as executing the ruler construction.

$$
\begin{array}{cc}
\begin{array}{ccccccccc}
\mathsf{p}_1 & \cdots & \mathsf{p}_n & \mathsf{l}_1 & \cdots & \mathsf{l}_m & \mathsf{x} & \mathsf{y} & \mathsf{z}
\end{array} & \\
\left(\begin{array}{ccc|ccc|ccc}
p_1^* & & \langle\!\langle p, p' \rangle\!\rangle & & & & p_1^{\mathsf{x}} & p_1^{\mathsf{y}} & p_1^{\mathsf{z}} \\
& \ddots & & & \langle\!\langle p, \ell \rangle\!\rangle & & & \vdots & \\
\langle\!\langle p', p \rangle\!\rangle & & p_n^* & & & & p_n^{\mathsf{x}} & p_n^{\mathsf{y}} & p_n^{\mathsf{z}} \\
\hline
& & & \ell_1^* & & \langle\!\langle \ell, \ell' \rangle\!\rangle & \ell_1^{\mathsf{x}} & \ell_1^{\mathsf{y}} & \ell_1^{\mathsf{z}} \\
& \langle\!\langle \ell, p \rangle\!\rangle & & & \ddots & & & \vdots & \\
& & & \langle\!\langle \ell', \ell \rangle\!\rangle & & \ell_m^* & \ell_m^{\mathsf{x}} & \ell_m^{\mathsf{y}} & \ell_m^{\mathsf{z}} \\
\hline
p_1^{\mathsf{x}} & & p_n^{\mathsf{x}} & \ell_1^{\mathsf{x}} & & \ell_m^{\mathsf{x}} & & & \\
p_1^{\mathsf{y}} & \cdots & p_n^{\mathsf{y}} & \ell_1^{\mathsf{y}} & \cdots & \ell_m^{\mathsf{y}} & & \Sigma_{\mathsf{E}} & \\
p_1^{\mathsf{z}} & & p_n^{\mathsf{z}} & \ell_1^{\mathsf{z}} & & \ell_m^{\mathsf{z}} & & &
\end{array}\right) &
\begin{array}{c}
\mathsf{p}_1 \\ \vdots \\ \mathsf{p}_n \\ \mathsf{l}_1 \\ \vdots \\ \mathsf{l}_m \\ \mathsf{x} \\ \mathsf{y} \\ \mathsf{z}
\end{array}
\end{array}
$$

Figure 5.3: The generic matrix satisfying the encoding of incidence relations among points $\mathsf{p}_1, \dots, \mathsf{p}_n$ and lines $\mathsf{l}_1, \dots, \mathsf{l}_m$ in the projective plane, according to Definition 5.21. The scalar product $\langle\!\langle \cdot, \cdot \rangle\!\rangle$ is given by the inverse of the Σ_{E} block.

Lemma 5.22. Let $\Sigma \in \widetilde{\mathcal{R}}^{\bullet}_{\mathbb{K}}(F)$ in the notation of Definition 5.21. Then:

(1) Σ contains the homogeneous coordinates of points $\boldsymbol{t}_i$, $\boldsymbol{f}_i$, p_i and lines ℓ_i in $\mathbb{PK}^2$ in the entries $\mathsf{P} \times \mathsf{E}$ and $\mathsf{L} \times \mathsf{E}$. The points corresponding to the projective basis coincide with the standard projective basis, except for $\widetilde{\mathbf{1}} = [s_{\mathsf{x}} : s_{\mathsf{y}} : 1]$, which may be different from $\mathbf{1}$. The x-axis, the y-axis and the line at infinity are the same as with the standard projective basis. The points $\boldsymbol{t}_i$ and $\boldsymbol{f}_i$ lie on the x-axis.

(2) Σ_{E} is a diagonal matrix. With the non-degenerate symmetric bilinear form $\langle\!\langle \cdot, \cdot \rangle\!\rangle$ defined by Σ_{E}^{-1}, an incidence $\mathsf{p}_i \in \mathsf{l}_j^{\perp}$ imposed in the construction implies $\langle\!\langle p_i, \ell_j \rangle\!\rangle = 0$, i.e., $p_i \in \Sigma_{\mathsf{E}}(\ell_j^{\perp})$.

(3) The $\boldsymbol{f}_i$, p_i and ℓ_i are uniquely determined as points in $\mathbb{PK}^2$ by the points $\boldsymbol{t}_i$, the scalings s_{x}, s_{y} and the Σ_{E} block. All other off-diagonal entries of Σ are functions of these homogeneous coordinates.

Proof. (1) By the relations $(\mathcal{I}.\mathrm{i})$, we have $\widetilde{\boldsymbol{\infty}}_x = [1 : 0 : 0]$, $\widetilde{\boldsymbol{\infty}}_y = [0 : 1 : 0]$, $\widetilde{\mathbf{0}} = [0 : 0 : 1]$ and $\widetilde{\mathbf{1}} = [s_{\mathsf{x}} : s_{\mathsf{y}} : 1]$ as points in the projective plane, with $s_{\mathsf{x}}, s_{\mathsf{y}} \neq 0$. This is still a projective basis and the x-axis, the y-axis and the line at infinity remain the same. This is consistent with constraints $(\mathcal{I}.\mathrm{iii})$, proving that indeterminate points are on the x-axis. The $\boldsymbol{f}_i$ are constructed by von Staudt as intersection points with ℓ_{x}, so they remain on the x-axis. Because of constraints $(\mathcal{I}.\mathrm{iv})$ all homogeneous coordinate vectors are non-zero and hence valid points/lines in $\mathbb{PK}^2$.

(2) Denote by $\langle\!\langle v, w \rangle\!\rangle := v^{\mathsf{T}}\, \Sigma_{\mathsf{E}}^{-1}\, w$ the non-degenerate symmetric bilinear form defined by $\Sigma_{\mathsf{E}}^{-1} = \Sigma_{\mathsf{xyz}}^{-1}$. The relations $(\mathsf{p}_i \mathsf{l}_j | \mathsf{xyz})$ of type $(\mathcal{I}.\mathrm{v})$ are equivalent to

$$
\Sigma_{\mathsf{p}_i \mathsf{l}_j} \overset{!}{=} \Sigma_{\mathsf{p}_i, \mathsf{xyz}}\, \Sigma_{\mathsf{xyz}}^{-1}\, \Sigma_{\mathsf{xyz}, \mathsf{l}_j} = \langle\!\langle p_i, \ell_j \rangle\!\rangle
$$

and then type $(\mathcal{I}.\mathrm{vi})$ makes this scalar product vanish, for every relation $\mathsf{p}_i \in \mathsf{l}_j^{\perp}$ requested. Applying this to constraints of type $(\mathcal{I}.\mathrm{ii})$ with the three unit vectors $\widetilde{\boldsymbol{\infty}}_{\mathsf{x}}$, $\widetilde{\boldsymbol{\infty}}_{\mathsf{y}}$ and $\widetilde{\mathbf{0}}$ shows that all off-diagonal entries of Σ_{E} vanish.

(3) Since all points and lines are valid objects in $\mathbb{PK}^2$ — in particular due to type $(\mathcal{I}.\mathrm{iv})$, the zero vector is never permissible as a vector of homogeneous coordinates (even though

as a vector in $\mathbb{K}^3$ it satisfies all incidence relations it may be involved in), thus the construction never degenerates —, the uniqueness of the result of the von Staudt construction in Lemma 5.19 proves that all points and lines are uniquely determined by the starting points, which are the projective basis and the indeterminates, as well as the definition of incidence. Relations ($\mathcal{I}$.v) then fix all off-diagonal entries on $\mathsf{PL} \times \mathsf{PL}$ as functions of the homogeneous coordinates, as per ($\measuredangle$). ☐

Lemma 5.22 shows that constraints ($\mathcal{I}$.i) for the standard projective basis fix the *standard* projective basis in every model up to a scaling of the x- and y-axis by non-zero quantities s_{x} and s_{y}. The points $\boldsymbol{f}_i$ which correspond to the evaluations of polynomials f_i end up on the x-axis and their location in $\mathbb{PK}^2$ is uniquely determined by the scalings, the bilinear form and the locations of $\boldsymbol{t}_i$. The next lemma makes this more precise:

Lemma 5.23. Let $\Sigma \in \widetilde{\mathcal{R}}^{\bullet}_{\mathbb{K}}(F)$. Denote by $\boldsymbol{t}_i = [t_i : 0 : 1]$ and $\boldsymbol{f}_i = [f_i^{\mathsf{x}} : 0 : 1]$ the points in $\mathbb{PK}^2$ determined by Σ according to Lemma 5.22. Then $f_i^{\mathsf{x}} = s_{\mathsf{x}} f_i(t_1/s_{\mathsf{x}}, \ldots, t_k/s_{\mathsf{x}})$.

Proof. Let $S = \left(\begin{smallmatrix} s_{\mathsf{x}} & 0 & 0 \\ 0 & s_{\mathsf{y}} & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right)$ be the scaling which maps the standard projective basis to the one contained in Σ. Use p' and ℓ' to refer to points constructed with the von Staudt algorithm and the standard basis and use p and ℓ for the same objects constructed by Σ with the scaled basis and non-standard scalar product. Consider also the points $\boldsymbol{t}'_i = [t_i/s_{\mathsf{x}} : 0 : 1]$. Then we obtain the projective basis and the indeterminates of Σ as images under S of the standard basis and the indeterminates $\boldsymbol{t}'_i$. It is straightforward to show by induction on the steps of the ruler construction using Lemma 5.22 (2) that:

— If ℓ is constructed from p_1 and p_2 with $p_i = Sp'_i$, then $\ell = \Sigma_{\mathsf{E}} S^{-1} \ell'$.
— If q is constructed in turn from ℓ_1 and ℓ_2 with $\ell_i = \Sigma_{\mathsf{E}} S^{-1} \ell'_i$, then $q = Sq'$.

Thus, $\boldsymbol{f}_i = S\boldsymbol{f}'_i$ and $f_i^{\mathsf{x}} = s_{\mathsf{x}} f_i'^{\mathsf{x}} = s_{\mathsf{x}} f_i(t_1/s_{\mathsf{x}}, \ldots, t_k/s_{\mathsf{x}})$ by Lemma 5.18. ☐

Lemma 5.24. Let $a_1, \ldots, a_k \in \mathbb{K}$ be arbitrary and suppose that $\mathbb{K}$ is infinite or ordered. Then $\widetilde{\mathcal{R}}^{\bullet}_{\mathbb{K}}(F)$ has a model which gives indeterminates the value $\boldsymbol{t}_i = [a_i : 0 : 1]$ and evaluations $\boldsymbol{f}_i = [f_i(a_1, \ldots, a_k) : 0 : 1]$.

Proof. A model Σ can be constructed based on Figure 5.3. Fill the coordinates of points of P corresponding to the standard projective basis with the actual standard projective basis, set the indeterminate points $\boldsymbol{t}_i$ as required and finally set Σ_{E} to be the identity matrix. Then for every incidence $\mathsf{p} \in \mathsf{l}^{\perp}$ demanded by the von Staudt construction we have, by Lemma 5.22 (2) with $\langle\!\langle \cdot, \cdot \rangle\!\rangle = \langle \cdot, \cdot \rangle$ that indeed $\langle p, \ell \rangle = 0$, i.e., $p \in \ell^{\perp}$. Execute the von Staudt construction from these settings and all off-diagonal entries will be filled to satisfy the constraints.

This gives values to all entries of Σ except for the diagonals p_i^* and ℓ_j^* as shown in Figure 5.3. These diagonals are used to make Σ principally regular or positive-definite. Viewing the matrix constructed in the previous paragraph as an element of $\mathbb{K}(p_i^*, \ell_j^*)$ shows that each principal minor is a non-zero polynomial (the product of all diagonal elements in the submatrix arises as a monomial), hence the matrix is generically principally regular over the infinite field $\mathbb{K}$ by Lemma 4.1. In the case of an ordered field, observe that the block Σ_{E} is the positive-definite identity matrix, whose Schur complement is $\Sigma_{\mathsf{PL}} - \Sigma_{\mathsf{PL},\mathsf{E}} \Sigma_{\mathsf{E},\mathsf{PL}}$. The diagonal entries appear only in the left summand. Clearly, they can be chosen large enough to make this difference diagonally dominant and hence positive-definite. Lemma 3.12 implies that such a Σ with Schur-complementary positive-definite blocks is positive-definite. ☐

Definition 5.25. Let $F = \{f_1, \dots, f_r\} \subseteq \mathbb{Z}[t_1, \dots, t_k]$. Denote by $\mathcal{I}(F)$ (and $\mathcal{R}^\bullet_{\mathbb{K}}(F)$) the extension of $\widetilde{\mathcal{I}}(F)$ (respectively $\widetilde{\mathcal{R}}^\bullet_{\mathbb{K}}(F)$) by the constraints

($\mathcal{I}$.vii) $(\mathsf{fx}|)$ for all polynomial value symbols $\mathsf{f} = \mathsf{f}_1, \dots, \mathsf{f}_r$.

Proposition 5.26. The model $\mathcal{R}^\bullet_{\mathbb{K}}(F)$ of von Staudt CI constraints is non-empty for an infinite (or ordered) field $\mathbb{K}$ if and only if the variety of F (over the algebraic or real closure of $\mathbb{K}$) has a $\mathbb{K}$-rational point.

Proof. Suppose that there exists a model $\Sigma \in \mathcal{R}^\bullet_{\mathbb{K}}(F)$ which contains points $\widetilde{\mathbf{1}} = [s_{\mathsf{x}} : s_{\mathsf{y}} : 1]$ as well as $\boldsymbol{t}_i = [t_i : 0 : 1]$ and $\boldsymbol{f}_i = [f_i^{\mathsf{x}} : 0 : 1]$. The homogeneous coordinates are unique up to a scalar from $\mathbb{K}$. Then by Lemma 5.23 and constraint ($\mathcal{I}$.vii), we have $0 = f_i^{\mathsf{x}} = s_{\mathsf{x}} f_i(t_1/s_{\mathsf{x}}, \dots, t_k/s_{\mathsf{x}})$. Since $s_{\mathsf{x}} \neq 0$ and t_i/s_{x} are in $\mathbb{K}$, this gives a solution to F in $\mathbb{K}$.

Conversely, let $a_1, \dots, a_k \in \mathbb{K}$ be a solution to F. Then they define a matrix in the model $\widetilde{\mathcal{R}}^\bullet_{\mathbb{K}}(F)$ by Lemma 5.24. Moreover, since the a_i are roots of the polynomials, the constraints ($\mathcal{I}$.vii) are satisfied by Lemma 5.23. $\square$

Remark 5.27. In the case of singular Gaussian distributions, recall from Section 1.3 that the truth of a CI statement $(\mathsf{ij}|\mathsf{K})$ is determined by the vanishing of the almost-principal minor $[\mathsf{ij}|\mathsf{L}]$ where L is any inclusion-maximal subset of K such that $[\mathsf{L}] > 0$. The CI statements appearing in the constraints in Definition 5.21 are either of the form $(\mathsf{ij}|)$ or $(\mathsf{ij}|\mathsf{xyz})$. As shown in Lemma 5.22, the 3×3 block Σ_{E} is diagonal by constraints ($\mathcal{I}$.ii) applied to pairs of $\{\widetilde{\boldsymbol{\infty}}_{\mathsf{x}}, \widetilde{\boldsymbol{\infty}}_{\mathsf{y}}, \widetilde{\mathbf{0}}\}$ in the projective basis. Moreover, these same constraints applied to one of the previous vectors with the fourth basis vector $\widetilde{\mathbf{1}}$ shows that the diagonals of Σ_{E} are all non-zero. Hence Σ_{E} is always positive-definite, even in the positive-semidefinite Gaussian model. Since this is the only non-trivial conditioning set, all CI statements in the singular case are interpreted exactly as in the regular case and therefore the algebra behind the proofs in this section holds unaltered. This implies Proposition 5.26 for models of singular Gaussians.

Remark 5.28. It would be desirable to extend this result to arbitrary symmetric matrices over arbitrary (infinite) fields, in particular because principal regularity is part of the definition of a CI model just to ensure the well-definedness of the CI statement $(\mathsf{ij}|\mathsf{K})$. Afterwards, this property is almost trivially enforced in Lemma 5.24 and plays no role for encoding of point and line configurations. The well-definedness of $(\mathsf{ij}|\mathsf{K})$ in the singular Gaussian case is a feature of positive-semidefinite matrices and does not hold in general for symmetric ones. Consider for example the not semidefinite matrix

$$\Sigma = \begin{array}{c} \begin{array}{ccccc} \mathsf{i} & \mathsf{j} & \mathsf{x} & \mathsf{y} & \mathsf{z} \end{array} \\ \begin{pmatrix} 1 & 1 & -1 & 0 & 0 \\ 1 & -1 & -1 & 0 & 0 \\ -1 & -1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix} \end{array} \begin{array}{c} \mathsf{i} \\ \mathsf{j} \\ \mathsf{x} \\ \mathsf{y} \\ \mathsf{z} \end{array} \qquad \begin{array}{c} \Sigma[\mathsf{xyz}] = 0, \\ \Sigma[\mathsf{xy}] = 2 \neq 0, \;\; \Sigma[\mathsf{yz}] = 2 \neq 0, \\ \Sigma[\mathsf{ij}|\mathsf{xy}] = 2 - 2 = 0, \;\; \Sigma[\mathsf{ij}|\mathsf{yz}] = 2 \neq 0. \end{array}$$

As seen on the right, the interpretation of the CI symbol $(\mathsf{ij}|\mathsf{xyz})$ according to singular Gaussians is inconsistent: it depends on the choice of full-rank subset of xyz. Thus, it is not clear how to define an interpretation of $(\mathsf{ij}|\mathsf{K})$ for general symmetric matrices which is consistent with semidefinite matrices and which allows Proposition 5.26 to be proved in generality.

5.5 Algebraic universality and characteristic sets

Characteristic sets. A direct corollary to Proposition 5.26 concerns characteristic sets of CI constraints which were implicitly studied in Section 4.2 already:

Definition 5.29. The *characteristic set* $\chi(\mathcal{I})$ of a set of CI constraints is the set of characteristics over which $\mathcal{I}$ has a non-empty model. This is the characteristic set $\chi(F)$ of the associated polynomial system $F = \{ \Gamma[\mathsf{K}] \neq 0, \Gamma[\mathsf{ij|K}] \bowtie_{\mathsf{ij|K}} 0 \}$ where each $\bowtie_{\mathsf{ij|K}} \in \{ =, \neq \}$ is determined by $\mathcal{I}$.

The Lefschetz Principle implies that $\chi(\mathcal{I})$ is either a finite set of primes or a cofinite set of primes which additionally contains 0. These conditions completely characterize the characteristic sets of algebraic Gaussians. See [Oxl11, Section 6.8] for the history of the same problem (and the same resolution) in matroid theory.

Theorem 5.30. Let T be any possible characteristic set. There exists a set of CI constraints $\mathcal{I}$ such that $\chi(\mathcal{I}) = T$.

Proof. By Proposition 5.26 it suffices to find a general integer polynomial system F with characteristic set T. These instances are well-known. If S is a finite set of primes (and not including zero), then the arithmetic statement $\prod_{p \in S} p = 0$ is satisfiable precisely over characteristics in S. Otherwise, consider a cofinite set of primes T also containing zero. Let S be its finite complement in the set of primes. Then one uses the fact [Oxl11, Lemma 6.8.6] that $\overline{\mathbb{F}_p}$ contains a k^{th} root of unity if and only if p does not divide k. Thus the cyclotomic polynomial of order $k = \prod_{p \in S} p$ yields characteristic set T. □

Remark 5.31. One of the two classes of equations used in the proof above, $\prod_{p \in S} p = 0$, has no variables in it, so it is not sensible to think about its solution set and whether it is empty or not. Nevertheless, there is a ruler construction behind this equation, which constructs the product by repeated von Staudt multiplication and then requires that the resulting point lie on the x- and the y-axes. This incidence configuration is only realizable — and the corresponding CI constraints satisfiable — in planes over fields of the appropriate characteristic.

Algebraic degree and a question of Šimeček. In this section we apply Proposition 5.26 to prove a universality result for Gaussian CI constraints and field extensions which can be used to answer a question of Šimeček [Šim06b] about rational points on Gaussian CI models. The central notion is that of *positive algebraic degree*, which measures the algebraic complexity of CI models:

Definition 5.32. The *positive algebraic degree* of $\mathcal{I}$ is the minimal extension degree of a real field over $\mathbb{Q}$ which is required to satisfy $\mathcal{I}$:

$$\operatorname{posdeg} \mathcal{I} := \min_{\Sigma \in \mathcal{R}^+_{\mathbb{R}}(\mathcal{I})} \max_{ij} \deg \left(\mathbb{Q}(\sigma_{ij}) \, / \, \mathbb{Q} \right).$$

By Tarski's transfer principle, the model $\mathcal{R}^+_{\mathbb{R}}$ contains a real-algebraic point if it contains a point at all. Therefore, the positive algebraic degree of any feasible set of CI constraints is finite; otherwise it is infinite by convention. The reference to $\mathbb{R}$ in this definition is natural in the context of statistics, but the notion remains the same if $\mathbb{R}$ is replaced by any real-closed field such as $\widetilde{\mathbb{Q}}$. Notice that both, conditional independence and dependence statements, are necessary to make this notion interesting: without dependence statements, the identity matrix satisfies the constraints, whereas without independence statements, any generic rational positive-definite matrix does.

Šimeček's Question. If the positive algebraic degree of a model is finite, must it be one?

The question was motivated by his work [Šim06b] on the catalogue of (singular) Gaussian CI structures on four random variables. To obtain realizability certificates, he sampled positive-semidefinite **integer** matrices. For the purpose of determining CI structures, this is equivalent to sampling rational matrices. For all **but one** of the structures for which he did not find a non-realizability proof, his program found a rational realization. For the remaining one, called `M85`, he found an irrational one by hand.

Example 5.33: Šimeček's model № 85. The CI structure in question is

$$\texttt{M85} = \{\, (12|4), (14|3), (14|23), (24|3), (24|13), (34|12) \,\}$$

over $\mathsf{N} = 1234$. The irrational realization given by Šimeček is on the left below. This is the natural first candidate to investigate the failure of rationality for.

$$\begin{array}{cc} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} & \\ \begin{pmatrix} 1 & 3/\alpha\gamma & 100/\beta\gamma & 10/\beta\gamma \\ 3/\alpha\gamma & 1 & 3/4 & 3/40 \\ 100/\beta\gamma & 3/4 & 1 & 1/10 \\ 10/\beta\gamma & 3/40 & 1/10 & 1 \end{pmatrix} & \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} \end{array} \qquad \begin{array}{cc} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} & \\ \begin{pmatrix} 1 & -1/17 & -49/51 & -7/17 \\ -1/17 & 1 & 1/3 & 1/7 \\ -49/51 & 1/3 & 1 & 3/7 \\ -7/17 & 1/7 & 3/7 & 1 \end{pmatrix} & \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} \end{array}$$

with $\alpha = 632\,836$, $\beta = 158\,209$ and $\gamma = \sqrt{2\alpha - \beta}$.

On the right is a **rational** positive-semidefinite matrix which also realizes `M85`. Its vanishing principal minors are $\Sigma[123]$ and $\Sigma[1234]$, which do not affect the interpretation of **any** CI statement in $\mathcal{A}_{1234}$ as an almost-principal minor. This matrix was found (quickly) by `Mathematica` [WM] after optimistically imposing the vanishings of the above principal minors and simplifying the resulting polynomial system by hand. △

Hence, Šimeček's Question has an affirmative answer for at most four random variables. Using the von Staudt constructions of Proposition 5.26, we are able to recreate a proof of MacLane [Mac36] in matroid theory, which strongly implies a negative answer in general.

Theorem 5.34. For every real number field $\mathbb{K}$ there exists a CI model which has an $\mathbb{L}$-rational point, for any real algebraic extension $\mathbb{L}$ of $\mathbb{Q}$, if and only if $\mathbb{K} \subseteq \mathbb{L}$.

Proof. The prime field $\mathbb{Q}$ is perfect and hence by the primitive element theorem the finite extension $\mathbb{K}$ has a primitive element α over $\mathbb{Q}$ with minimal polynomial $f \in \mathbb{Z}[t]$. Application of Proposition 5.26 to $F = \{\, f \,\}$ produces a constraint set $\mathcal{I}(f)$ which has a positive model over $\mathbb{L}$ if and only if $\mathbb{L}$ contains a root of f. By standard facts about field extensions, this implies that $\mathbb{K}$ is contained in any field $\mathbb{L}$ which has a model of $\mathcal{I}(f)$. By Remark 5.27, this argument is valid for regular as well as singular CI models. □

Corollary 5.35. No proper subfield of $\widetilde{\mathbb{Q}}$ is sufficient to witness the non-emptiness of all (regular or singular) Gaussian CI models. □

Remark 5.36. The whole argument is based on Proposition 5.26 and works for general infinite fields, too. For each proper extension $\mathbb{L} \,/\, \mathbb{K}$ of an infinite field, there exists a **gaussoid** which shows $\mathfrak{g}^*_{\mathbb{K}} \subsetneq \mathfrak{g}^*_{\mathbb{L}}$. This gaussoid is obtained as a weak image over $\mathbb{L}$ (recall Definition 3.26) of the constraint system $\mathcal{I}(f)$ for the minimal polynomial f of a generator of $\mathbb{L}$ over $\mathbb{K}$. This is analogous to [Whi87, Remark 1.7.4] after White's presentation of MacLane's theorem.

Remark 5.37. The decision problem associated to Šimeček's Question, namely to decide when a CI model is non-empty over the rational numbers, is equivalent to the general problem of deciding whether a polynomial system has a rational solution. The decidability of this problem is famously open: it is Hilbert's 10^{th} problem over $\mathbb{Q}$; see [Maz92, Maz95].

5.6 Hardness of the implication problem

In this section we fix the setting of statistics: we consider positive, real Gaussians and determine the **computational** complexity of the inference problem, i.e., how difficult is it to decide whether a proposed CI inference rule is valid for all regular Gaussians?

Recall from Section 3.5 that an inference rule $\varphi : \bigwedge \mathcal{L} \Rightarrow \bigvee \mathcal{M}$ is valid for $\mathfrak{g}_{\mathbb{R}}^{+}$ if and only if the model $\mathcal{R}_{\mathbb{R}}^{+}(\mathcal{L} \cup \neg\mathcal{M})$ is empty. By context-completeness (Corollary 4.33), it suffices to treat these questions over the smallest ground set $\mathsf{N} = [\varphi]$ over which the formula φ can be stated. This makes the question a finite one. The emptiness of the model as a semialgebraic subset of $\mathrm{Sym}_{\mathsf{N}}(\mathbb{R})$ is an existential first-order problem over the reals and thus $\mathtt{ETR}$ is a natural upper bound for this problem. In Section 5.4 we have cast incidence geometry in the plane into Gaussian CI constraints. Reasoning about incidence statements in the plane is hard because the valid incidence theorems are the axioms of linear rank-3 matroids, and so we expect that reasoning about Gaussian CI inference is hard as well. This is the main result of this section:

Theorem 5.38. The decision problem $\mathtt{GR}_{\mathbb{R}}^{+}$ is $\exists\mathbb{R}$-complete.

Proof. There are two polytime reducibilities to prove:

$\mathtt{GR}_{\mathbb{R}}^{+} \le \mathtt{ETR}$**:** Let $\mathcal{I}$ be a set of CI constraints, N be its context and Σ a generic symmetric $\mathsf{N} \times \mathsf{N}$ matrix. $\mathtt{GR}$ asks about the existence of a real solution to a system of polynomial constraints specified by $\mathcal{I}$. This is already an instance of $\mathtt{ETR}$. It remains to show how to describe the system in polynomial time. Naïvely writing out the determinant of a $\mathsf{iK} \times \mathsf{jK}$ matrix requires $(|\mathsf{K}|+1)!$ terms, which is not polynomial in the input length $|\mathsf{ijK}| = |\mathsf{K}|+2$ of that statement. It is more economic to write down a **system** of polynomials in more variables which is equivalent to the CI constraint but uses only polynomially many characters in $|\mathsf{ijK}|$. Using the Schur complement, we write $\Sigma[\mathsf{ij}|\mathsf{K}] = \Sigma[\mathsf{K}]\left(\sigma_{\mathsf{ij}} - \Sigma_{\mathsf{i,K}}\Sigma_{\mathsf{K}}^{-1}\Sigma_{\mathsf{K,j}}\right)$. The entries of Σ_{K}^{-1} can be written down efficiently as the solutions $\check{\sigma}_{\mathsf{kl}}$ to a quadratic system $\Sigma_{\mathsf{K}} \cdot \check{\Sigma}_{\mathsf{K}} = \mathbb{1}_{\mathsf{K}}$ of size $\mathcal{O}(|\mathsf{K}|^3)$. Thus the parenthesized term in the Schur complement may be defined by a polynomially sized system of equations. On the set of positive-definite matrices, the vanishing or non-vanishing of this term is sufficient to enforce the respective CI constraint.

The positive definiteness of Σ is enforced in a similar fashion. By the definition of context, the ground set size N is polynomially bounded by the input length of $\mathcal{I}$. Thus, matrix transformations on Σ which are polytime in N are permissible in our reduction. Consider the matrix equation which constitutes one step of the diagonalization of the symmetric bilinear form Σ in the middle on the left:

$$
\left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline -\frac{a_2}{p} & & & \\ \vdots & & \mathbb{1}_{2\cdots n} & \\ -\frac{a_n}{p} & & & \end{array}\right)
\begin{array}{cc} \begin{array}{cccc} 1 & 2 & \cdots & n \end{array} & \\ \left(\begin{array}{c|ccc} p & a_2 & \cdots & a_n \\ \hline a_2 & & & \\ \vdots & & \Sigma_{2\cdots n} & \\ a_n & & & \end{array}\right) & \begin{array}{c} 1 \\ 2 \\ \vdots \\ n \end{array} \end{array}
\left(\begin{array}{c|ccc} 1 & -\frac{a_2}{p} & \cdots & -\frac{a_n}{p} \\ \hline 0 & & & \\ \vdots & & \mathbb{1}_{2\cdots n} & \\ 0 & & & \end{array}\right)
=
\begin{array}{cc} \begin{array}{cccc} 1 & 2 & \cdots & n \end{array} & \\ \left(\begin{array}{c|ccc} p & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & \Sigma' & \\ 0 & & & \end{array}\right) & \begin{array}{c} 1 \\ 2 \\ \vdots \\ n \end{array} \end{array}
$$

Σ is positive-definite if and only if $p > 0$ and Σ' is positive-definite. Writing down the necessary equations for this step can be done in polynomial time and since Σ' is one dimension smaller, only linearly many steps are needed to obtain a complete characterization of positive definiteness for Σ. Introducing new variables for the entries of Σ' in each step prevents expressions in the polynomial system from becoming too large to manipulate in polynomial time. For example, the step above imposes the equation $p\sigma'_{22} = p\sigma_{22} - a_2^2$. The next elimination step uses the variable σ'_{22} instead of its definition to avoid doubling the number of terms in every step.

$\texttt{ETR} \le \texttt{GR}^+_{\mathbb{R}}$**:** By Lemma 5.3 we may assume that we are given a finite set of polynomial equations. We may additionally assume, by introducing more temporary variables, that all equations are additions $x = y + z$ or multiplications $x = y \cdot z$ of variables, as described in the context of Shor's normal form in Section 5.3. These two simplifications of the given first-order formula to a system of elementary equations F can be performed in polynomial time. It remains to notice that the von Staudt construction of such a polynomial system requires only polynomially many ruler construction steps in its length (which is bounded polynomially by the number of additions and multiplications performed). Finally, the encoding of a ruler construction into the constraint system $\mathcal{I}(F)$ described in Section 5.4.2 takes polynomial time in the number of steps. This is all in all polynomial time and the correctness of the reduction is proved in Proposition 5.26. □

Remark 5.39. The computation of determinants is of course a well-studied problem in the complexity theory of linear algebra. An efficient algorithm is described, for example, in [BCS97, Section 16.4]. However, these algorithms suppose that a concrete matrix is given with entries in a field, whereas our situation in the above proof requires to assert the positivity of the determinant of a **generic** matrix such that the formula (or arithmetic circuit) is well-formed. The main obstacle is that the genericity of the problem makes it pointless to do pivoting in the Gaussian elimination algorithm which is normally used to quickly compute determinants, because every entry is generically non-zero, but we still have to avoid dividing by zero in every concrete instantiation of our circuit. The assumption of positive definiteness makes pivoting unnecessary because it is part of the property to describe that the leading principal minors are non-zero and thus can be divided by.

In complexity theory, a decision problem (the "yes" answers) and its complementary decision problem (the "no" answers) are distinguished because the certificates for these answers may be substantially different. This is a consequence of the definition of Karp reduction. Nevertheless, from a practical point of view, the two problems are equally difficult. The complement of the existential theory is the universal theory of the reals and analogously has a complete problem denoted by $\forall\mathbb{R}$.

Corollary 5.40. The Gaussian CI implication problem $\texttt{GCI}$ is $\forall\mathbb{R}$-complete. □

It is remarkable that the upper bound of 3 can be placed on the size of conditioning sets K in every CI symbol $(\mathsf{ij}|\mathsf{K})$ required in the reduction of $\texttt{ETR}$ to $\texttt{GR}$ (see Definition 5.21). On the other hand, the construction requires an unbounded number of each, negated and non-negated statements, which correspond to antecedents and consequents in the inference rule version. The prior research into infinite families of Gaussian inference rules has usually targeted single-consequent formulas with many antecedents; see [Sul09, Šim06a] and Section 4.2. The unbounded number of consequents arises in our construction from constraints of type ($\mathcal{I}$.iv) which ensure that all points and lines are valid objects in projective space.

Question 5.41. Is there a polynomial-time reduction of $\texttt{ETR}$ to $\texttt{GR}$ for which there is a universal upper bound on the number of consequents in the constructed inference formulas?

5.7 Topological universality of oriented CI models

The previous two sections dealt with algebraic and algorithmic complexity of the set of counterexamples to a CI inference formula for Gaussians. The algebraic complexity is a pointwise and therefore very local property of these sets which is summarized in the positive algebraic degree. The emptiness of the model is a rather global property. What is missing in the middle is more detailed information about the *shape* of CI models. This section is

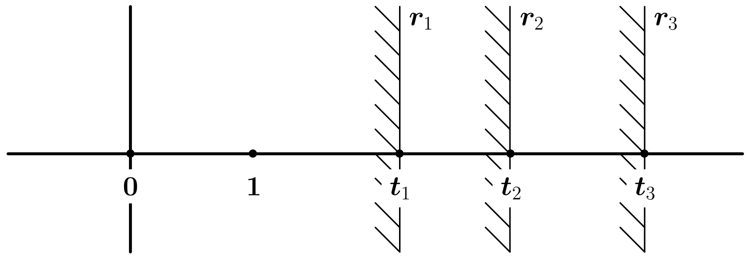

Figure 5.4: Points on the x-axis can be ordered by specifying their orientation with respect to lines passing through the other points. The shading on the side of each line indicates the "positive halfplane" with respect to its orientation.

devoted to an extension of the construction presented in Section 5.4 to *primary oriented CI models*. This is a proper extension of the CI models studied so far which allow to specify the **sign** of the almost-principal minors instead of just whether they vanish or not. This is equivalent to specifying the sign of the *partial correlations* $[ij|K]/[iK][jK]$ of the Gaussian distribution. Conditional independence in Gaussians coincides with the vanishing of the partial correlation, but the signs of partial correlations are a sound concept to study in synthetic probability theory as well. Its combinatorial approximation, *oriented gaussoids*, are further studied in Section 6.2.2.

Definition 5.42. Let $\mathsf{S} := \{ \mathsf{0}, \mathsf{+}, \mathsf{-} \}$ denote the set of signs. A *primary oriented CI constraint system* over ground set N is a three-sorted set $\mathcal{S} \subseteq \mathsf{0}\mathcal{A}_\mathsf{N} \cup \mathsf{+}\mathcal{A}_\mathsf{N} \cup \mathsf{-}\mathcal{A}_\mathsf{N}$ which specifies a collection of CI statements, each with a sign attached.

The model $\mathcal{R}^{\mathsf{S}}_{\mathbb{K}}(\mathcal{S})$ is defined as usual: all principal minors are positive and the CI statements in $\mathcal{S}$ must have the indicated sign. This makes sense only for ordered fields $\mathbb{K}$ and we fix $\mathbb{K} = \mathbb{R}$ in this section. The model of a primary oriented constraint system is a primary basic semialgebraic set. The reason for introducing this extension is that, as explained in Section 5.3, reducing a semialgebraic set to a system of *equations* only preserves the algebraic and algorithmic complexity, but this is not sufficient to capture its topology. The Shor normal form retains a few elementary inequalities from the description of a primary basic semialgebraic set. To model these inequalities, CI constraints are not enough. We have to be able to specify the signs of almost-principal minors. Our main result is a universality theorem for primary oriented CI models up to stable equivalence:

Theorem 5.43. For every primary basic semialgebraic set there is an **oriented** CI model over $\mathbb{R}$ which is stably equivalent to it.

This implies that the homotopy type of any primary basic semialgebraic set is attained by primary oriented CI models. For statistics this means that the set of counterexamples to a wrong inference rule may be fairly complicated, in particular it can have any finite number of connected components. To prove this theorem, we start with a Shor normal form of the given primary basic semialgebraic set $\{ f_i = 0,\ 1 < t_1 < \cdots < t_k \}$ with polynomial equations $f_i = 0$ expressing addition or multiplication of variables. A small refinement of the construction in Definition 5.21 allows the incorporation of the total order of variables and prove stable equivalence. Figure 5.4 shows the idea for how to impose the ordering geometrically: let a perpendicular line r_j fall on each variable t_j on the x-axis. This line is oriented towards positive y-infinity. Analytically, this means that it has coordinates $r_j = [-1 : 0 : t_j]$ up to a **positive** scalar. Then, for all $i < j$ we require $\langle \boldsymbol{t}_i, r_j \rangle = -t_i + t_j > 0$, which puts $\boldsymbol{t}_i$ on the left half-plane of r_j in its chosen orientation and equivalently imposes $t_i < t_j$ on the x-axis. Since the **sign** of inner products can effectively be prescribed by primary oriented CI constraints, the variables can be ordered.

Definition 5.44. Given a Shor normal form $F = \{ f_i = 0,\ 1 < t_1 < \cdots < t_k \}$ of a primary basic semialgebraic set, consider its von Staudt construction making reference to points P, lines L and coordinates $\mathsf{E} = \{ \mathsf{x}, \mathsf{y}, \mathsf{z} \}$. Among the points are f_i for the values of the polynomial expressions f_i, t_j for the unknowns t_j, the four points of a projective basis and the ordering lines r_j constructed by connecting t_j to ∞_y. The primary oriented CI constraint system $\mathcal{S}(F)$ corresponding to its von Staudt construction is:

($\mathcal{S}$.i) $\mathsf{0}(\mathsf{pe}|)$ or $\mathsf{+}(\mathsf{pe}|)$ for all points p corresponding to the standard projective basis and $\mathsf{e} \in \mathsf{E}$, depending on whether the e-coordinate of the point is zero or not (in which case it is positive).

($\mathcal{S}$.ii) $\mathsf{s}(\mathsf{pq}|)$ for all points p, q of the standard projective basis with the sign $\mathsf{s} = \operatorname{sgn}\langle p, q \rangle$.

($\mathcal{S}$.iii) $\mathsf{+}(\mathsf{tx}|)$, $\mathsf{0}(\mathsf{ty}|)$ and $\mathsf{+}(\mathsf{tz}|)$ for indeterminate points $\mathsf{t} = \mathsf{t}_1, \dots, \mathsf{t}_k$.
$\mathsf{-}(\mathsf{rx}|)$, $\mathsf{0}(\mathsf{ry}|)$, $\mathsf{+}(\mathsf{rz}|)$ for their corresponding ordering lines $\mathsf{r} = \mathsf{r}_1, \dots, \mathsf{r}_k$.

($\mathcal{S}$.iv) $\mathsf{+}(\mathsf{ae}|)$ for each $\mathsf{a} \in \mathsf{PL}$ and one of the coordinates $\mathsf{e} \in \mathsf{E}$ on which the point or line labeled a is non-zero, which can be deduced by Lemma 5.19.

($\mathcal{S}$.v) $\mathsf{0}(\mathsf{ab}|\mathsf{xyz})$ for all distinct $\mathsf{a}, \mathsf{b} \in \mathsf{PL}$.

($\mathcal{S}$.vi) $\mathsf{0}(\mathsf{pl}|)$ for any incidence relationship between $\mathsf{p} \in \mathsf{P}$ and $\mathsf{l} \in \mathsf{L}$ which is required to express a join or meet operation of the construction.

($\mathcal{S}$.vii) $\mathsf{+}(\mathsf{r_k t_j}|)$ for pairs of ordering lines and indeterminates with $j < k$.

($\mathcal{S}$.viii) $(\mathsf{fx}|)$ for all polynomial value symbols $\mathsf{f} = \mathsf{f}_1, \dots, \mathsf{f}_s$.

The proof that these constraints capture the topology of the semialgebraic set consists of multiple steps of stable projections and rational equivalences. The following routine will be useful:

Lemma 5.45. Let $V \subseteq \mathbb{R}^n$ be a semialgebraic set and $T = \{ \phi_i(w) > 0 \} \subseteq \mathbb{R}^m$ a convex semialgebraic set defined by rational polynomials. Consider any $\mathbb{Q}$-defined rational map $\varphi : T \to \mathrm{GL}(\mathbb{R}^n)$. Then V is stably equivalent to

$$V^\varphi = \{ \varphi(t) \cdot v \in \mathbb{R}^n : t \in T,\, v \in V \}.$$

Proof. The definition of V^φ is as the coordinate projection of the image of $V \times T$ under the rational map $(v, t) \mapsto (\varphi(t) \cdot v, t)$. This map has a rational inverse $(v', t) \mapsto (\varphi(t)^{-1} \cdot v', t)$ by Cramer's rule and clearly both are continuous. This shows that $V^\varphi \times T$ is rationally equivalent to $V \times T$. The two product spaces project down to V^φ and V, respectively, and stably so, since each fiber is isomorphic to T which has all the required properties. □

Proof of Theorem 5.43. Since Definition 5.44 is a refinement of Definition 5.21, the matrices in the model look like Figure 5.3 and Lemma 5.22 still applies. It is easy to see based on the new constraints in ($\mathcal{S}$.iii) and ($\mathcal{S}$.vii) and the previous proofs of Lemmas 5.23 and 5.24 that every solution to the Shor normal form system appears in some model matrix and, conversely, that every model matrix encodes an oriented point and line configuration which gives a solution to the system. The proof of stable equivalence proceeds in multiple steps of stable projections and applications of Lemma 5.45 starting with the model $\mathcal{R}^{\mathsf{S}}(\mathcal{S})$ and ending with the x-coordinates of the $\boldsymbol{t}_i$ points encoded in the matrices. These coordinates are the solution set to the Shor normal form system.

(1) First project away the $\mathsf{PL} \times \mathsf{PLE}$ part of the matrices, indicated in red:

$$
\begin{array}{cc}
\begin{array}{ccc|ccc|ccc} \mathsf{p}_1 & \cdots & \mathsf{p}_n & \mathsf{l}_1 & \cdots & \mathsf{l}_m & \mathsf{x} & \mathsf{y} & \mathsf{z} \end{array} & \\
\left(\begin{array}{ccc|ccc|ccc}
p_1^* & & \langle\!\langle p,p'\rangle\!\rangle & & & & p_1^{\mathsf{x}} & p_1^{\mathsf{y}} & p_1^{\mathsf{z}} \\
 & \ddots & & & \langle\!\langle p,\ell\rangle\!\rangle & & & \vdots & \\
\langle\!\langle p',p\rangle\!\rangle & & p_n^* & & & & p_n^{\mathsf{x}} & p_n^{\mathsf{y}} & p_n^{\mathsf{z}} \\
\hline
 & & & \ell_1^* & & \langle\!\langle \ell,\ell'\rangle\!\rangle & \ell_1^{\mathsf{x}} & \ell_1^{\mathsf{y}} & \ell_1^{\mathsf{z}} \\
 & \langle\!\langle \ell,p\rangle\!\rangle & & & \ddots & & & \vdots & \\
 & & & \langle\!\langle \ell',\ell\rangle\!\rangle & & \ell_m^* & \ell_m^{\mathsf{x}} & \ell_m^{\mathsf{y}} & \ell_m^{\mathsf{z}} \\
\hline
p_1^{\mathsf{x}} & & p_n^{\mathsf{x}} & \ell_1^{\mathsf{x}} & & \ell_m^{\mathsf{x}} & & & \\
p_1^{\mathsf{y}} & \cdots & p_n^{\mathsf{y}} & \ell_1^{\mathsf{y}} & \cdots & \ell_m^{\mathsf{y}} & & \Sigma_{\mathsf{E}} & \\
p_1^{\mathsf{z}} & & p_n^{\mathsf{z}} & \ell_1^{\mathsf{z}} & & \ell_m^{\mathsf{z}} & & &
\end{array}\right) &
\begin{array}{c} \mathsf{p}_1 \\ \vdots \\ \mathsf{p}_n \\ \mathsf{l}_1 \\ \vdots \\ \mathsf{l}_m \\ \mathsf{x} \\ \mathsf{y} \\ \mathsf{z} \end{array}
\end{array}
$$

This is a stable projection because $\mathsf{PL} \times \mathsf{E}$ is redundant by symmetry of the matrices, the off-diagonal entries in $\mathsf{PL} \times \mathsf{PL}$ are given by the assignments $\sigma_{\mathsf{pl}} = \Sigma_{\mathsf{p},\mathsf{E}} \Sigma_{\mathsf{E}}^{-1} \Sigma_{\mathsf{E},\mathsf{l}}$ which are rational functions in the surviving coordinates, and the diagonal entries in $\mathsf{PL} \times \mathsf{PL}$ are subject to the strict inequalities demanding positive definiteness. All in all, the fiber of any matrix under this projection is a **spectrahedron** (together with the array of fixed coordinates in $\mathsf{PL} \times \mathsf{E}$) and therefore convex. This shows that the projection is stable.

(2) This leaves an $\mathsf{E} \times \mathsf{PL}$ matrix whose columns are points and lines in the projective plane, and the $\mathsf{E} \times \mathsf{E}$ block holding the inner product. The open convex set $\mathbb{R}_{>0}^{\mathsf{PLE}}$ acts on all of these column vectors in $\mathbb{R}^3$ by scalar multiplication. This is an instance of Lemma 5.45. By Item ($\mathcal{S}$.iv), there exists one distinguished positive coordinate in each column and thus for each matrix in the model there is a unique element of $\mathbb{R}_{>0}^{\mathsf{PLE}}$ which sends these distinguished coordinates to 1; call the obtained matrices *dehomogenized*. Lemma 5.45 implies that the set of dehomogenized matrices is stably equivalent to the full model. From now on, we work with the dehomogenized model. In particular this turns the $\mathsf{E} \times \mathsf{E}$ block into the identity matrix which can be stably projected away. Moreover, the first three projective basis vectors $\widetilde{\boldsymbol{\infty}}_{\mathsf{x}} = [1:0:0]$, $\widetilde{\boldsymbol{\infty}}_{\mathsf{y}} = [0:1:0]$ and $\widetilde{\mathbf{0}} = [0:0:1]$ inside the matrix attain exactly the given unit coordinates as vectors in $\mathbb{R}^3$ and may be projected away as well.

$$
\begin{array}{cc}
\begin{array}{c|ccc|ccc|ccc|ccc|ccc} \mathbf{1} & \mathsf{t}_1 & \cdots & \mathsf{t}_k & \mathsf{f}_1 & \cdots & \mathsf{f}_s & \mathsf{p}_1 & \cdots & \mathsf{p}_n & \mathsf{r}_1 & \cdots & \mathsf{r}_k & \mathsf{l}_1 & \cdots & \mathsf{l}_m \end{array} & \\
\left(\begin{array}{c|ccc|ccc|ccc|ccc|ccc}
s_{\mathsf{x}} & t_1 & \cdots & t_k & \widetilde{f}_1 & \cdots & \widetilde{f}_s & p_1^{\mathsf{x}} & \cdots & p_n^{\mathsf{x}} & -1 & \cdots & -1 & 1 & \cdots & 1 \\
s_{\mathsf{y}} & 0 & \cdots & 0 & 0 & \cdots & 0 & p_1^{\mathsf{y}} & \cdots & p_n^{\mathsf{y}} & 0 & \cdots & 0 & \ell_1^{\mathsf{y}} & \cdots & \ell_m^{\mathsf{y}} \\
1 & 1 & \cdots & 1 & 1 & \cdots & 1 & 1 & \cdots & 1 & t_1 & \cdots & t_k & \ell_1^{\mathsf{z}} & \cdots & \ell_m^{\mathsf{z}}
\end{array}\right) &
\begin{array}{c} \mathsf{x} \\ \mathsf{y} \\ \mathsf{z} \end{array}
\end{array}
$$

(3) The result of dehomogenization and projections is the above matrix. The evaluation of the polynomial expressions f_i in the points $\boldsymbol{f}_i$ is deformed by s_{x} according to Lemma 5.23:

$$
\widetilde{f}_i = s_{\mathsf{x}} f(t_i/s_{\mathsf{x}}, \ldots, t_k/s_{\mathsf{x}}).
$$

The coordinates of other points and lines constructed in this matrix are also functions of these initial scalings of the x- and y-axis. Acting with $\mathbb{R}_{>0}^2$ via Lemma 5.45 as before removes these degrees of freedom and maps the fourth (dehomogenized) projective basis vector $\mathbf{1} = [s_{\mathsf{x}} : s_{\mathsf{y}} : 1]$ to the standard one $[1:1:1]$. These coordinates can be projected away. This is a rescaling of the finite affine chart of the projective plane in which the $\boldsymbol{t}_i$ and $\boldsymbol{f}_i$ lie. The rescaling results in $\boldsymbol{t}_i = [t_i/s_{\mathsf{x}} : 0 : 1]$ and $\boldsymbol{f}_i = [f(t_i/s_{\mathsf{x}}) : 0 : 1]$ with s_{x} normalized to 1.

(4) The remaining columns of the matrix contain exactly those dehomogenized coordinates of the points and lines in the von Staudt ruler construction, as if it had been initialized with the standard projective basis. These coordinates are uniquely defined as vectors in $\mathbb{R}^3$. Therefore, tracing the ruler construction backwards, all constructed points and lines can be successively stably projected away, because they are given by polynomial functions in the previously constructed points and lines. What remains is a matrix with columns indexed by the indeterminates:

$$\begin{array}{cc} \begin{array}{ccc} \mathsf{t}_1 & \cdots & \mathsf{t}_k \end{array} & \\ \begin{pmatrix} t_1 & \cdots & t_k \\ 0 & \cdots & 0 \\ 1 & \cdots & 1 \end{pmatrix} & \begin{array}{c} \mathsf{x} \\ \mathsf{y} \\ \mathsf{z} \end{array} \end{array}$$

The projection down to the x-row is stable. This yields the solution set of the Shor normal form. □

6

Approximations to the inference problem

As shown in the previous chapter, the Gaussian CI inference problem is hard in general. It can be used to decide whether any given boolean combination of integer polynomial equations and inequalities has a solution or not. Still, Gaussian CI concerns the relations of only **very specific** polynomials. The goal in this chapter is to exploit the special finite structure of principal and almost-principal minors to devise tractable approximations to the inference problem. These approximations find some but not all valid inferences, but they work faster than a general method. All computational results reported in this chapter were obtained with the still-experimental software package `CInet tools` (`https://conditional-independence.net`) developed by the author and bundling the `SAT` solvers `CaDiCaL` [Bie19], `GANAK` [SRSM19] and Toda and Soh's `AllSAT` solvers [TS16] as well as the LP solver `soplex` [GBE+18, GSW12, GSW15].

6.1 The Gaussian CI configuration space

In this chapter, we pick up again the themes of Sections 3.5 and 3.6. Gaussian CI is concerned with (semi)algebraic constraints on the principal and almost-principal minors of a symmetric matrix. These are special polynomials, so the study of their algebraic relationships leads to a special version of the Zariski topology on PR_{N} or PD_{N}. This finite version of the Zariski closure operator (it is not a topology) where the closed sets are the complete relations (see Definition 3.35) necessarily has a particular combinatorial flavor. Picking other sets of polynomials, such as the maximal minors of a $d \times n$ matrix, leads to a different combinatorial flavor, namely **matroid theory**. In both of these combinatorial shadows of the Zariski topology, some geometric theorems still hold. The Alternatives in algebraic geometry and Alternatives in real algebraic geometry for arbitrary polynomial systems have been shown to hold for the systems formulated in terms of maximal minors or principal and almost-principal minors only — meaning that the non-realizability certificates can be written in terms of the restricted set of special polynomials as well. These theorems are the **existence of final polynomials** proved for matroids in [BS89, Section 4.2] and for gaussoids in Section 3.6.

The key to these results is the bracket ring introduced in Chapter 3: given a field $\mathbb{K}$ (which we assume to be of characteristic zero throughout) and a fixed ground set N, let $\mathcal{R} = \mathbb{K}[\mathcal{P}_{\mathsf{N}}, \mathcal{A}_{\mathsf{N}}]$ be the ring generated by one variable per principal and almost-principal minor. Inside of it, we have the prime ideal $\mathcal{J}$ which is the kernel of the evaluation homomorphism $\mathcal{R} \to \mathbb{K}[\Sigma]$ of brackets into subdeterminants of a generic symmetric matrix. The ideal $\mathcal{J}$ contains the universal relations among principal and almost-principal minors. The realizable Gaussian CI theory happens in the image $\mathbb{K}[\Sigma]$, or rather its attached affine space $\mathrm{Sym}_{\mathsf{N}}(\mathbb{K})$, but to reason about the relations and design approximations to them, it is advantageous to view this image as the coordinate ring $\mathcal{R} \,/\, \mathcal{J}$.

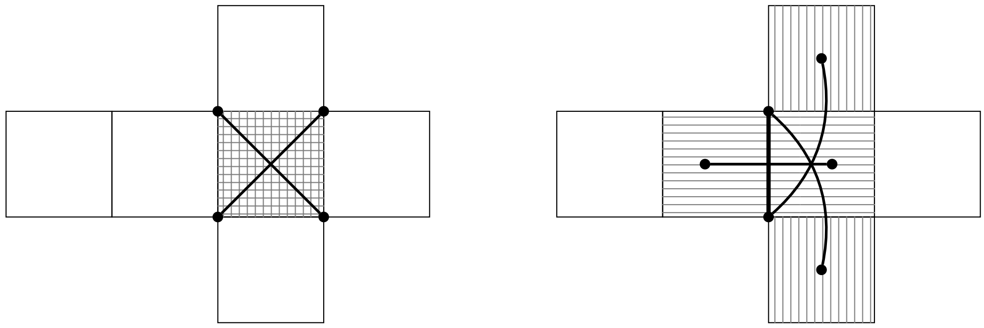

Figure 6.1: The two homogeneous trinomials associated to an edge and a square in a 3-cube. On the left is a square trinomial which involves the indicated 2-face (almost-principal minor) and the products of the coupled vertices (principal minors) with opposite signs. On the right is the trinomial corresponding to the marked edge. This trinomial is a combination of the two 2-faces incident to the edge and products of incident vertices and opposite 2-faces.

Definition 6.1. The *Gaussian CI configuration vector* of $\Sigma \in \mathrm{Sym}_N(\mathbb{K})$ is the vector in $\mathbb{K}^{\mathcal{P}_\mathsf{N} \cup \mathcal{A}_\mathsf{N}}$ containing all the principal and almost-principal minors of Σ. The *Gaussian CI configuration space* $\mathrm{Ga}^\bullet_\mathbb{K}(\mathsf{N})$ is the set of configuration vectors over all matrices in $\mathrm{PR}_\mathsf{N}(\mathbb{K})$ or $\mathrm{PD}_\mathsf{N}(\mathbb{K})$, as usual depending on $\bullet \in \{\,*, +\,\}$.

Configuration vectors are very wasteful representations of symmetric matrices. Since every **entry** of Σ is either a principal or almost-principal minor of degree one, the configuration vector contains the entire matrix and on top some polynomials in these entries. This redundancy has the benefit of producing a geometric object on which the principal regularity, positive definiteness and the CI structure of a matrix can be easily read off. The universal relations in $\mathcal{J}$ are polynomial relations on the entries of configuration vectors and hence, over characteristic zero, the defining equations of Ga^*, and Ga^+ is just the intersection of Ga^* over the same field with the polyhedron $\mathbb{K}^{\mathcal{P}_\mathsf{N}}_{>0} \times \mathbb{K}^{\mathcal{A}_\mathsf{N}}$.

This approach is used in matroid theory as well and leads to a configuration space known as the *Grassmannian* $\mathrm{Gr}(d, n)$; see [BS89, Section 1.2] and [Stu08, Chapter 3]. The *Lagrangian Grassmannian* $\mathrm{LGr}(n, 2n)$, which is parametrized by **all** minors of a symmetric $n \times n$ matrix, has already appeared in the proof of Proposition 4.46. Our configuration space $\mathrm{Ga}^*(n)$ is a projection of $\mathrm{LGr}(n, 2n)$ to the principal and almost-principal minors only; see the discussion in [BDKS19, Section 1] and [HS07]. The vanishing ideal $\mathcal{J}$ of this geometric object was studied in [BDKS19] via its symmetry group $\mathrm{SL}_2(\mathbb{K})^\mathsf{N} \rtimes \mathfrak{S}_\mathsf{N}$ using methods of representation theory. As explained in Section 3.3, the hyperoctahedral group is obtained as a discrete subgroup of this symmetry. The main result [BDKS19, Theorem 5] gives a complete description of the homogeneous quadratic part of $\mathcal{J}$ in terms of four families of relations and transformations on them, which eventually yield generators of the quadratic part. The shortest polynomials among these four families are the *edge trinomials* and *square trinomials*:

$$[\mathsf{kL}] \cdot [\mathsf{ij}|\mathsf{L}] = [\mathsf{L}] \cdot [\mathsf{ij}|\mathsf{kL}] + [\mathsf{ik}|\mathsf{L}] \cdot [\mathsf{jk}|\mathsf{L}], \tag{T.i}$$

$$[\mathsf{ij}|\mathsf{L}]^2 = [\mathsf{iL}] \cdot [\mathsf{jL}] - [\mathsf{L}] \cdot [\mathsf{ijL}], \tag{T.ii}$$

These relations are similar to the two cases of the Matúš identity (Lemma 3.5): one where ijk are distinct and the other where two of them coincide. However, here they appear with a larger symmetry group: there is one edge trinomial for every 1-face in every 3-face of the cube, and one square trinomial for every 2-face in every 3-face. These are the images of the Matúš identity under the hyperoctahedral group as explained in Remark 3.18; see also Figure 6.1 for a description of how to enumerate all of these trinomials.

The remainder of this chapter is dedicated to different, broadly applicable frameworks for turning geometric information such as the above relations on the configuration space into valid inference rules for Gaussian CI and to show how these inference rules can be discovered, in particular, by software using combinatorial and polyhedral methods, which are more efficient than a general search for Positivstellensatz certificates. Gaussoids show up again as one combinatorial shadow of the edge trinomials, but their discovery through this framework suggests to study refinements, which may be called **gaussoids with coefficients**. We introduce oriented gaussoids and show that orientability is a `SAT` problem which gives additional valid inference rules. Combinatorialization of the square trinomials leads to semimatroids and linear programming which connects to preexisting methods for CI inference of discrete random variables [BHLS10].

6.2 Gaussoids with coefficients

6.2.1 Hyperfields: algebra under uncertainty. A *hyperfield* $\mathtt{H}$ is a field-like algebraic structure. It has two operations which obey very similar axioms to the ones for fields. The difference is that, while the non-zero elements form an ordinary multiplicative group, the (hyper-)addition operation is **multivalued**. Formally, addition is a map $\boxplus : \mathtt{H} \times \mathtt{H} \to \mathcal{P}(\mathtt{H}) \setminus \{ \emptyset \}$ which carries pairs of elements to non-empty subsets. In expressions consisting of more than two terms, addition and multiplication are extended element-wise to subsets of $\mathtt{H}$, so that in general the value of any polynomial expression over a hyperfield is a set of **possible values**. The intention behind hyperfields is to model field operations with this uncertainty attached to the result of addition. The field axioms are therefore slightly amended, e.g.,

— Zero introduces no uncertainty: $0 \boxplus x = \{ x \}$ and $0 \cdot x = 0$.
— The additive inverse of x is a **unique** element $-x$ such that $0 \in x \boxplus -x$.
— The distributive law is exact: $a(x \boxplus y) = ax \boxplus ay \subseteq \mathtt{H}$.

A hyperfield with univalued addition is a field. [BB17] as well as [Vir10] contain an introduction to hyperstructures with many relevant examples and further references. Baker and Bowler's work on matroids over hyperfields appears in more general form in [BB19]. In this section, the following two hyperfields are of interest:

Example 6.2: The Krasner hyperfield. The *Krasner hyperfield* $\mathtt{K}$ is supported on a two-element set $\mathtt{K} = \{ \mathtt{0}, * \}$. The hyperfield axioms mandate the result of addition with the additively netural element $\mathtt{0}$. **Unlike** the field with two elements, where $* \boxplus *$ would equal $\mathtt{0}$, the Krasner hyperfield requires $* \boxplus * = \{ \mathtt{0}, * \}$. The multiplication is given by $* \cdot * = *$.

To give a familiar model of this hyperfield, consider a partition of the real numbers $\mathbb{R}$ into the set $\mathtt{0} := \{ 0 \}$ and the set $* = \mathbb{R} \setminus \mathtt{0}$. These two sets together with element-wise addition yield the Krasner hyperfield, where the convention $* \boxplus * = \mathbb{R} = \{ \mathtt{0}, * \}$ is understood. △

Example 6.3: The hyperfield of signs. The supporting set $\mathtt{S} = \{ \mathtt{0}, \mathtt{+}, \mathtt{-} \}$ of the hyperfield of signs was already introduced in Definition 5.42 anticipating a more combinatorial development of oriented CI structures. The addition of the hyperfield of signs is best explained by providing again a faithful model of this hyperfield through a partition of the real numbers. Let $\mathtt{0} = \{ 0 \}$, $\mathtt{+} = \{ x > 0 \}$ and $\mathtt{-} = \{ x < 0 \}$. Then we have in particular the familiar rules for multiplying signs $\mathtt{-} \cdot \mathtt{-} = \mathtt{+} \cdot \mathtt{+} = \mathtt{+}$ and $\mathtt{-} \cdot \mathtt{+} = \mathtt{-}$, and the uncertainty about the sign of a sum: $\mathtt{+} \boxplus \mathtt{-} = \{ \mathtt{0}, \mathtt{+}, \mathtt{-} \}$. △

A *morphism* of hyperfields is a map $f : \mathtt{H} \to \mathtt{H}'$ which is a homomorphism of the multiplicative groups, sends 0 to 0 and satisfies the inclusion $f(x \boxplus y) \subseteq f(x) \boxplus f(y)$. Baker and Bowler remark in [BB17, Section 4.1] the connection between morphisms from commutative rings into $\mathtt{K}$ and $\mathtt{S}$ and the spectrum, respectively real spectrum, of the ring. The real numbers surject onto the Krasner hyperfield, so the geometry of $\mathtt{K}$ can be seen as a (crude) coarsening of real geometry. However, the Krasner hyperfield provides a particularly bad resolution of this geometry because it is a final object in the category of hyperfields — it coarsens *every* hyperfield and in particular every field. The hyperfield of signs coarsens every ordered field.

Arithmetic in $\mathtt{K}$ or $\mathtt{S}$ may be seen as evaluating polynomial expressions over $\mathbb{R}$ with incomplete information about the value assigned to the unknowns. $\mathtt{K}$ reveals only whether an unknown is zero or not, whereas $\mathtt{S}$ reveals its sign. Elements $a_1, \dots, a_k \in \mathtt{H}$ are a *root* of a polynomial $f \in \mathtt{H}[x_1, \dots, x_k]$ if $0 \in f(a_1, \dots, a_k)$. The set of common roots of a set of polynomials F is the $\mathtt{H}$-variety of F. The notion of $\mathtt{K}$- or $\mathtt{S}$-variety is an approximation from above to real varieties. It contains the hyperfield image of every point which could possibly be in the real variety, based on the limited information revealed by the hyperfield. Naturally, the more arithmetic expressions occur in the defining polynomials, the more likely it is that $\mathtt{0}$ is among the possible outcomes of an expression. Hence, the hyperfield coarsening approach tends to be less useful for dense polynomial systems. For example, evaluating a determinant of a matrix with Krasner entries is unlikely to produce anything but $\{\mathtt{0}, *\}$ as the value. Nevertheless, the notion of hyperfield variety fits well with the short quadratic trinomials which hold on the Gaussian CI configuration space. Being the shortest and lowest-degree relations, they encode the most essential geometric information about the configuration space from the hyperfield point of view and this yields useful approximations. The following result was proved in [BDKS19, Theorem 1] using the notion of combinatorial compatibility, which is equivalent to being the root of a polynomial over $\mathtt{K}$ (cf. [BB17, Example 2.20]):

Theorem 6.4. Gaussoids over N are the points in the $\mathtt{K}$-variety of the edge and square trinomials which satisfy $[\mathsf{K}] = *$ for all $\mathsf{K} \subseteq \mathsf{N}$. □

Using hyperfields to coarsen a field's arithmetic is a general technique to obtain (combinatorial) approximations of difficult geometric objects. These ideas go back to works of Dress and Wenzel on **matroids with coefficients** in a *fuzzy ring* [DW87, DW91]. The treatment of Baker and Bowler [BB17] shows that the points in $\mathtt{K}^{\binom{n}{r}}$ which are non-zero, alternating (as functions $\binom{\mathsf{N}}{r} \to \mathtt{K}$) and roots of the 3-term Grassmann–Plücker relations are precisely the matroids of rank r on N. The hyperfield of signs yields a similar characterization of **oriented** matroids; and tropical geometry [MS15] has its own hyperfield.

The proof of Theorem 6.4 is essentially sketched already in Example 3.55. The edge trinomials, as elements in $\mathcal{J}$, can be used as final polynomials. The derivation of the gaussoid axioms in Proposition 3.8 is based on analyzing the implications of the vanishing of certain bracket variables in these final polynomials. This is the same as determining their $\mathtt{K}$-variety.

6.2.2 Orientability of gaussoids. In analogy to Theorem 6.4 we can define gaussoids with coefficients in other hyperfields, as the points in the variety of the Matúš identity.

Definition 6.5. An *oriented CI structure* on ground set N is a map $\mathcal{O} : \mathcal{A}_{\mathsf{N}} \to \mathtt{S}$. It is an *oriented gaussoid* if it is a common root of all edge and square trinomials over $\mathtt{S}$ (where in addition to the value $\mathcal{O}[\mathsf{ij}|\mathsf{K}]$ prescribed by the map we also set $\mathcal{O}[\mathsf{K}] = +$ for all $\mathsf{K} \subseteq \mathsf{N}$).

Remark 6.6. The square trinomials are trivially satisfied by oriented gaussoids (and ordinary gaussoids over $\mathtt{K}$) because the right-hand side of (T.ii) always contains the term $+ - +$ which equals $\{\mathtt{0}, +, -\}$ in $\mathtt{S}$. For larger hyperfields, the square trinomials become significant and we therefore include them in the definition.

The value of $\mathcal{O}$ at $(\mathsf{ij}|\mathsf{K})$ is written as $\mathcal{O}[\mathsf{ij}|\mathsf{K}]$ instead of $\mathcal{O}(\mathsf{ij}|\mathsf{K})$ to suggest the similarity to $\Sigma[\mathsf{ij}|\mathsf{K}]$ in the realizable case. The *support* of an oriented gaussoid is the map $\mathcal{L} : \mathcal{P}_\mathsf{N} \cup \mathcal{A}_\mathsf{N} \to \mathtt{K}$ induced by $\mathcal{O}$ and the morphism $\mathtt{S} \to \mathtt{K}$. The support is naturally identified with a subset of $\mathcal{A}_\mathsf{N}$ by collecting all $(\mathsf{ij}|\mathsf{K})$ which map to $\mathtt{0}$. The support of any oriented gaussoid is a gaussoid because the hyperfield of signs refines the Krasner hyperfield.

Definition 6.7. A gaussoid is *orientable* if it is the support of an oriented gaussoid. The property of being an orientable gaussoid is denoted by $\mathfrak{o}$.

Theorem 6.8. The properties of being positively realizable, orientable and a gaussoid form a chain of properties: $\mathfrak{g}^+ \le \mathfrak{o} \le \mathfrak{g}$.

Proof. Every positively realizable gaussoid is orientable because a realization $\Sigma \in \mathrm{PD}_\mathsf{N}(\mathbb{K})$ defines a *realizable oriented gaussoid* via $\mathcal{O}[\mathsf{ij}|\mathsf{K}] = \operatorname{sgn} \Sigma[\mathsf{ij}|\mathsf{K}]$. Since the configuration vector of Σ satisfies the square and edge trinomials over $\mathbb{R}$ and $\mathbb{R}$ refines $\mathtt{S}$, it follows that $\mathcal{O}$ is an oriented gaussoid whose support obviously coincides with $[\![\Sigma]\!]$. $\square$

There are 51 oriented 3-gaussoids. The count of oriented 3-gaussoids breaks down into the **isomorphy classes** of their supporting gaussoids as follows:

E	L	U	B	F
20	4	4	2	1

Since there are three isomorphic versions of each of L, U and B, this gives a total count of 51. The symmetry in numbers between L and U is explained by the symmetries of oriented gaussoids. In addition to isomorphy via $\mathfrak{S}_\mathsf{N}$, oriented gaussoids are closed under *oriented duality* and *reorientation*. The oriented duality notion is inherited, just as in the unoriented case, from the inversion map on PD_N:

Definition 6.9. For an oriented CI structure $\mathcal{O} : \mathcal{A}_\mathsf{N} \to \mathtt{S}$ the *oriented dual* is $\mathcal{O}^\rceil : \mathcal{A}_\mathsf{N} \to \mathtt{S}$ defined by $\mathcal{O}^\rceil[\mathsf{ij}|\mathsf{K}] = -\mathcal{O}[\mathsf{ij}|\mathsf{K}]^\rceil$.

It is easy to see that if $\mathcal{O}$ satisfies all square and edge trinomials, then $\mathcal{O}^\rceil$ does as well because duality exchanges the terms $[\mathsf{kL}][\mathsf{ij}|\mathsf{L}]$ and $[\mathsf{L}][\mathsf{ij}|\mathsf{kL}]$ which have opposite signs in (T.i). The *reorientation group* is $\mathfrak{R}_\mathsf{N} := (\mathbb{Z}/2)^\mathsf{N}$ acting by $(\mathsf{Z} \cdot \mathcal{O})[\mathsf{ij}|\mathsf{K}] = (-1)^{|\mathsf{ij} \cap \mathsf{Z}|} \cdot \mathcal{O}[\mathsf{ij}|\mathsf{K}]$; see [BDKS19, Section 5]. This is the combinatorial effect of the action $\Sigma \to D\Sigma D$ on almost-principal minors, where D is a diagonal matrix with entries ± 1 whose signs are determined by the indicator vector of $\mathsf{Z} \subseteq \mathsf{N}$. It should be noted that the symmetry group $\mathfrak{R}_\mathsf{N} \rtimes \mathfrak{S}_\mathsf{N}$ of oriented gaussoids is supported on the same set as the hyperoctahedral group $\mathfrak{B}_\mathsf{N}$, but the two groups do not act in the same way since reorientation does not change the supporting CI structure but only the signs of non-zero CI statements. Both groups are obtained from the $(\mathbb{Z}/4)^\mathsf{N}$ subgroup of the $\mathrm{SL}_2(\mathbb{R})^\mathsf{N}$ action discussed in Section 3.3, but $\mathfrak{R}_\mathsf{N}$ is a subgroup while the swap subgroup of $\mathfrak{B}_\mathsf{N}$ is a quotient.

Up to reorientation and isomorphy, there are $3+1+1+1+1 = 7$ oriented 3-gaussoids displayed in the table below as vectors in $\mathtt{S}^6$. The six components correspond to CI statements in $\mathcal{A}_3 = \{ (12|), (12|3), (13|), (13|2), (23|), (23|1) \}$ in that ordering.

E	L	U	B	F
++++++	0-----	+0++++	0000++	000000
+-++++				
--++++				

The notion of forbidden minors and characterization by means of them can be easily transferred from CI structures to oriented CI structures. We have:

Proposition 6.10. Oriented gaussoids have a finite forbidden-minor characterization given equivalently by the 51 compulsory 3-minors.

Proof. The proof is analogous to Lemma 1.26: the square and edge trinomials reference only CI statements in a fixed 3-face of the N-cube. Therefore, an oriented CI structure is an oriented gaussoid if and only if each of its 3-minors is an oriented 3-gaussoid. □

Example 6.11. By Theorem 6.8 and the realizability proofs in Remark 3.9, every 3-gaussoid is orientable. This is not true anymore for larger ground sets. Consider the CI structure $\mathcal{S} := \{\, (12|3), (13|4), (14|2) \,\}$. To see that this gaussoid is non-orientable, it suffices to consider the 3-minors of its possible orientations. These are oriented 3-gaussoids and their supports are the corresponding minors of $\mathcal{S}$. To prove non-orientability, one enumerates all possible orientations of these minors and shows that any combination would assign **incompatible** signs. It follows that no oriented 4-gaussoid exists whose support is $\mathcal{S}$. Of particular importance in this example are the orientations of `L` and `U`. The complete list, without factoring out the reorientation group, is:

L	U
0-----	+0----
0+++--	-0++--
0-++++	-0--++
0+--++	+0++++

By reorientation we can assume that, if an orientation of $\mathcal{S}$ exists, there exists one which assigns + to each of $(12|)$, $(13|)$ and $(14|)$. This in fact reduces the number of possible orientations of the minors $(123|)$, $(124|)$ and $(134|)$ (which are all `U`) from four to just one. Each group of columns in the following table corresponds to the signs assigned to CI statements in the order $(\mathsf{ij}|), (\mathsf{ij}|\mathsf{k}), (\mathsf{ij}|\mathsf{l}), (\mathsf{ij}|\mathsf{kl})$ for each ij as indicated.

	$(12\mid\cdots)$				$(13\mid\cdots)$				$(14\mid\cdots)$				$(23\mid\cdots)$				$(24\mid\cdots)$				$(34\mid\cdots)$			
$(123\mid)$	+	0	␣	␣	+	+	␣	␣	␣	␣	␣	␣	+	+	␣	␣	␣	␣	␣	␣	␣	␣	␣	␣
$(124\mid)$	+	␣	+	␣	␣	␣	␣	␣	+	0	␣	␣	␣	␣	␣	␣	+	+	␣	␣	␣	␣	␣	␣
$(134\mid)$	␣	␣	␣	␣	+	␣	0	␣	+	␣	+	␣	␣	␣	␣	␣	␣	␣	␣	␣	+	+	␣	␣

This (up to reorientation uniquely determined) partial oriented gaussoid then reduces the possibilities for the orientation of each of the `L`-supported minors $(123|4)$, $(124|3)$ and $(134|2)$, which in fact successively determine each other uniquely. To see this first notice that in every orientation of `L` the opposite faces have the same sign (unless one of them is `0`). The sign opposite to `0` determines whether the two other pairs of opposite faces have the same or different signs. The signs which determine the orientation are printed in blue, while the conclusions are in red:

	$(12\mid\cdots)$				$(13\mid\cdots)$				$(14\mid\cdots)$				$(23\mid\cdots)$				$(24\mid\cdots)$				$(34\mid\cdots)$			
(above)	+	0	+	␣	+	+	0	␣	+	0	+	␣	+	+	␣	␣	+	+	␣	␣	+	+	␣	␣
$(123\mid 4)$	␣	␣	+	+	␣	␣	0	+	␣	␣	␣	␣	␣	␣	−	−	␣	␣	␣	␣	␣	␣	␣	␣
$(124\mid 3)$	␣	0	␣	+	␣	␣	␣	␣	␣	␣	+	+	␣	␣	␣	␣	␣	␣	−	−	␣	␣	␣	␣
$(134\mid 2)$	␣	␣	␣	␣	␣	+	␣	+	␣	0	␣	+	␣	␣	␣	␣	␣	␣	␣	␣	␣	␣	−	−

Thus, a unique oriented CI structure is determined (up to reorientation) by having support $\mathcal{S}$, but this is not an oriented gaussoid because its $(234|)$-minor is `+-+-+-` which is not an oriented 3-gaussoid because its value under the trinomial corresponding to the edge $(2|3)$ in the $(234|)$-cube, $[3] \cdot [34|] - [23] \cdot [34|2] - [23|] \cdot [24|3] = + \cdot + - + \cdot - - + \cdot - = +$, does not contain `0`.

This edge trinomial is one which is not a Matúš identity, but which can be obtained from one by an application of the hyperoctahedral group (cf. Figure 6.1). This example shows that orientability is a strictly better approximation to realizability than being a gaussoid. △

By Proposition 6.10, oriented gaussoids have a finite axiomatization. However, their supports, the orientable gaussoids, do **not**, as shown below. The following result gives a surprising axiomatization of the class of *positively oriented gaussoids*, which are oriented gaussoids whose image in $\mathtt{S}$ is contained in $\{\, \mathtt{0}, \mathtt{+} \,\}$:

Proposition 6.12. The positively orientable gaussoids, the gaussoids which satisfy the ascension property $(\mathsf{ij}|\mathsf{L}) \Rightarrow (\mathsf{ij}|\mathsf{kL})$, and the graphic gaussoids all coincide.

The connection between positively orientable gaussoids and graphic gaussoids was made in [BDKS19, Theorem 4], which also proves that their realization spaces are homeomorphic to open euclidean balls.

Proof. The proof consists of noticing that all three properties have a finite axiomatization by their 3-minors (for graphic gaussoids this is proved in [Mat97, Proposition 2]) and that their 3-minors are all the same. In particular $\mathtt{L}$ is the only 3-gaussoid which is not positively orientable — precisely because it is not ascending. □

No finite axiomatization for orientability. The goal of this subsection is to prove the following theorem:

Theorem 6.13. The property $\mathtt{o}$ has no finite forbidden-minor characterization.

The non-axiomatizability result in Section 4.5 was bad news for realizable Gaussians because it ruled out one way of solving the inference problem. In this case, however, non-axiomatizability is good news, because orientability axioms are relatively easy to derive and this result means that new, independent axioms which further constrain the possibly realizable gaussoids will always be derivable using this method.

We explicitly construct infinitely many forbidden minors. These are gaussoids which are not orientable but all whose proper minors are rationally realizable near the identity matrix, which implies orientability. The family of gaussoids is familiar from Section 4.2 but appears here slightly changed: $\mathcal{G}_n \coloneqq \{\, (\mathsf{ij}|\mathsf{N}) \,\} \cup \{\, (\mathsf{ij}|\mathsf{k}) : \mathsf{k} \in \mathsf{N} \,\} \cup \mathcal{A}_\mathsf{N}$. The difference to Section 4.2 is that the marginal independencies $(\mathsf{kl}|)$ for all $\mathsf{k}, \mathsf{l} \in \mathsf{N}$ have been extended to the entire $\mathcal{A}_\mathsf{N}$. This is because the gaussoid axioms imply this, which was no concern in Section 4.2.

Lemma 6.14. $\mathcal{G}_n$ is a gaussoid for all $n \geq 3$.

Proof. By Lemma 1.26 it suffices to check its 3-minors. This involves a case distinction, exploiting the symmetries in the definition of $\mathcal{G}_n$. For all suitable $\mathsf{klmM} \subseteq \mathsf{N}$:

$$\textbf{(ijk|M):} \begin{cases} \mathtt{U}, & \text{if } |\mathsf{M}| = 0 \text{ or } n-1, \\ \mathtt{L}, & \text{if } |\mathsf{M}| = 1, \\ \mathtt{E}, & \text{otherwise,} \end{cases} \qquad \begin{aligned} &\textbf{(ikl|M):}\ \mathtt{L}, \\ &\textbf{(ikl|jM):}\ \mathtt{E}, \\ &\textbf{(klm|M):}\ \mathtt{F}. \end{aligned}$$

All of these minors are gaussoids, so $\mathcal{G}_n$ is a gaussoid. The minor $(\mathsf{ijk}|\mathsf{m}) = \mathtt{L}$ requires $|\mathsf{N}| \geq 3$ since otherwise $(\mathsf{ij}|\mathsf{k})$ and $(\mathsf{ij}|\mathsf{N})$ can appear in the same 3-minor $(\mathsf{ijk}|)$ or $(\mathsf{ijl}|\mathsf{k})$ and $\mathcal{G}_n$ would not be a gaussoid. □

Lemma 6.15. $\mathcal{G}_n$ is not orientable for any $n \geq 3$.

Proof. Suppose there is an orientation $\mathcal{O}$. Examine the 5-minor $(\mathsf{ijklm}|)$ of $\mathcal{G}_n$ for arbitrary $\mathsf{klm} \subseteq \mathsf{N}$. If $n = 3$, then this is the whole structure and has no orientation. For all $n > 3$, this is a uniquely determined 5-gaussoid with sixteen orientations, one of which must be the corresponding minor of the assumed orientation $\mathcal{O}$. These sixteen orientations are printed below using grouping and group-internal ordering familiar from Example 6.11.

$(ij\|\cdots)$	$(ik\|\cdots)$	$(il\|\cdots)$	$(im\|\cdots)$	$(jk\|\cdots)$	$(jl\|\cdots)$	$(jm\|\cdots)$	$(kl\|\cdots)$	$(km\|\cdots)$	$(lm\|\cdots)$
+000----	--------	--------	--------	--------	--------	--------	0--0----	0--0----	0--0----
-000++++	++++++++	++++++++	++++++++	--------	--------	--------	0--0----	0--0----	0--0----
-000++++	--------	--------	--------	++++++++	++++++++	++++++++	0--0----	0--0----	0--0----
+000----	++++++++	++++++++	++++++++	++++++++	++++++++	++++++++	0--0----	0--0----	0--0----
+000----	--------	++++++++	++++++++	--------	++++++++	++++++++	0++0++++	0++0++++	0--0----
-000++++	--------	++++++++	++++++++	++++++++	--------	--------	0++0++++	0++0++++	0--0----
-000++++	++++++++	--------	--------	--------	++++++++	++++++++	0++0++++	0++0++++	0--0----
+000----	++++++++	--------	--------	++++++++	--------	--------	0++0++++	0++0++++	0--0----
+000----	--------	++++++++	--------	--------	++++++++	--------	0++0++++	0--0----	0++0++++
-000++++	++++++++	--------	++++++++	--------	++++++++	--------	0++0++++	0--0----	0++0++++
-000++++	--------	++++++++	--------	++++++++	--------	++++++++	0++0++++	0--0----	0++0++++
+000----	++++++++	--------	++++++++	++++++++	--------	++++++++	0++0++++	0--0----	0++0++++
+000----	--------	--------	++++++++	--------	--------	++++++++	0--0----	0++0++++	0++0++++
-000++++	--------	--------	++++++++	++++++++	++++++++	--------	0--0----	0++0++++	0++0++++
-000++++	++++++++	++++++++	--------	--------	--------	++++++++	0--0----	0++0++++	0++0++++
+000----	++++++++	++++++++	--------	++++++++	++++++++	--------	0--0----	0++0++++	0++0++++

This list can be confirmed with `SAT` solvers as explained in Section 6.2.3. The following relations hold for $L = kl$ on every orientation:

— $\mathcal{O}[ij|] = -\mathcal{O}[ij|L] = -\mathcal{O}[ij|Lm]$,
— $\mathcal{O}[im|] = \mathcal{O}[im|L]$ and $\mathcal{O}[jm|] = \mathcal{O}[jm|L]$,
— $\mathcal{O}[im|] \cdot \mathcal{O}[jm|] = \mathcal{O}[ij|]$.

We prove these properties for growing sets L inductively. Suppose they hold for all L of a given size and $m \notin L$. Let $n \in N \setminus Lm$ be given. By the edge trinomial

$$\underbrace{\mathcal{O}[in|L]}_{=\mathcal{O}[in|]} = \mathcal{O}[in|Lm] + \mathcal{O}[im|L] \cdot \underbrace{\mathcal{O}[nm|L]}_{=0}$$

it follows that $\mathcal{O}[in|Lm] = \mathcal{O}[in|]$. The same holds for $\mathcal{O}[jn|Lm] = \mathcal{O}[jn|]$. Then by another edge trinomial

$$\underbrace{\mathcal{O}[ij|Lm]}_{=-\mathcal{O}[ij|]} = \mathcal{O}[ij|Lmn] + \underbrace{\underbrace{\mathcal{O}[in|Lm]}_{=\mathcal{O}[in|]} \cdot \underbrace{\mathcal{O}[jn|Lm]}_{=\mathcal{O}[jn|]}}_{=\mathcal{O}[ij|]}$$

which again implies $\mathcal{O}[ij|Lmn] = -\mathcal{O}[ij|]$ as elements of S. Inductively, this shows $0 = \mathcal{O}[ij|N] = -\mathcal{O}[ij|] \neq 0$ which is absurd. □

It remains to show that $\mathcal{G}_n$ are minor-minimal with the property of being non-orientable. Since orientability is minor-closed, it suffices to show that the maximal proper minors are orientable. Indeed, we can even show that they are rationally realizable near the identity matrix. The first two cases are trivial. The other two are treated by the lemmas below. We have these maximal proper minors:

$\mathcal{G} \setminus \mathbf{i} =$ $\mathcal{A}_N$ and realizable by the identity matrix.
$\mathcal{G} \,/\, \mathbf{i} =$ $\emptyset$ and evidently realizable near the identity matrix.
$\mathcal{G} \setminus \mathbf{k} =$ $\{\, (ij|l) : l \in N \setminus k \,\} \cup \mathcal{A}_{N \setminus k}$.
$\mathcal{G} \,/\, \mathbf{k} =$ $\{\, (ij|), (ij|N \setminus k) \,\} \cup \mathcal{A}_{N \setminus k}$.

Lemma 6.16. $\{\,(ij|k) : k \in N\,\} \cup \mathcal{A}_N$ over ijN is rationally realizable near the identity matrix.

Proof. This case is essentially Lemma 4.12, except that we have to show here that the CI structure is realizable instead of just showing that it contains some statements and not others. Take the matrix

$$\begin{array}{cc}
\begin{array}{cccccc} i & j & 1 & 2 & \cdots & n \end{array} & \\
\begin{pmatrix} 1 & xy & x & x & \cdots & x \\ xy & 1 & y & y & \cdots & y \\ x & y & & & & \\ x & y & & & & \\ \vdots & \vdots & & \mathbb{1}_{12\cdots n} & & \\ x & y & & & & \end{pmatrix} & \begin{array}{c} i \\ j \\ 1 \\ 2 \\ \vdots \\ n \end{array}
\end{array}.$$

This is a rational realization of some gaussoid near the identity matrix. The realized gaussoid certainly contains the subject structure. Showing that no other CI statements hold generically is easy. The following list contains all CI statements of interest and a short computation of the almost-principal minor. Let $k, l \in N$ and $L \subseteq N$ be arbitrary:

$$\begin{aligned}
[\mathbf{ij}|\mathbf{L}] &= xy - |L|xy \neq 0 \text{ unless } L = k, \\
[\mathbf{ik}|\mathbf{L}] &= x \neq 0, \\
[\mathbf{kl}|\mathbf{iL}] &= \det \begin{pmatrix} 0 & x \\ x & 1 - |L|x^2 \end{pmatrix} = -x^2 \neq 0,
\end{aligned}$$

Every proof writes the almost-principal minor using Schur complement expansion with respect to the embedded $L \times L$ identity matrix. The cases $(ik|jL)$ and $(kl|ijL)$ have the cases $(ik|L)$ and $(kl|iL)$, respectively, as submatrices. It follows from a Laplace expansion exploiting the unit diagonals that in both cases the non-vanishing certificate of the subcase appears as a term in the larger case as well. □

Lemma 6.17. $\{\,(ij|), (ij|N)\,\} \cup \mathcal{A}_N$ over ijN is rationally realizable near the identity matrix.

Proof. This case is essentially Lemma 4.63, except that we have an $N \times N$ identity matrix in place of the a generic matrix. We use the same construction:

$$\begin{array}{cc}
\begin{array}{ccc} i & j & N \end{array} & \\
\begin{pmatrix} 1 & 0 & \boldsymbol{u}^\mathsf{T} \\ 0 & 1 & \boldsymbol{v}^\mathsf{T} \\ \boldsymbol{u} & \boldsymbol{v} & \mathbb{1}_N \end{pmatrix} & \begin{array}{c} i \\ j \\ N \end{array}
\end{array},$$

where the Gram–Schmidt process with respect to $\mathbb{1}_N$ defines $\boldsymbol{u}$ and $\boldsymbol{v}$ as in Lemma 4.63. The definition is simpler here because the $N \times N$ block is the identity. In the nomenclature of Lemma 4.63, we have $\alpha_L = \sum_{p \in L} x_p^2$ and $\beta_L = \sum_{p \in L} x_p y_p$ and then

$$\begin{aligned}
u_k &= x_k, \\
v_k &= y_k \sum_{p \in N} x_p^2 - x_k \sum_{p \in N} x_p y_p.
\end{aligned}$$

This matrix is a rational realization near the identity matrix and its CI structure includes the subject structure of this lemma. It remains to check the other almost-principal minors as in Lemma 4.63:

$[\mathbf{ij}|\mathbf{L}] = \sum_{p \in N} \left[x_p^2 \sum_{k \in L} x_k y_k - x_p y_p \sum_{k \in L} x_k^2\right] \neq 0$ unless $L = \emptyset$ or N, as witnessed by the monomial $x_n^2 x_l y_l$, for $n \in N \setminus L$ and $l \in L$, which appears in the first sum and cannot be canceled by the second one,

$$\begin{aligned}
[\mathsf{jk}|\mathsf{L}] &= v_{\mathsf{k}} \neq 0, \\
[\mathsf{kl}|\mathsf{jL}] &= \det \begin{pmatrix} 0 & v_{\mathsf{k}} \\ v_{\mathsf{l}} & 1 - \sum_{\mathsf{p} \in \mathsf{L}} v_{\mathsf{p}}^2 \end{pmatrix} = v_{\mathsf{k}} v_{\mathsf{l}} \neq 0,
\end{aligned}$$

again using Schur complements of the $\mathsf{L} \times \mathsf{L}$ identity blocks. □

6.2.3 Applications of orientability. The kind of combinatorial matching and propagation performed by hand in Example 6.11 is well suited to the `SAT` solvers mentioned in Section 2.3. Because the hyperfields `K` and `S` are finite, gaussoids and oriented gaussoids can be described by a finite set of boolean variables obeying certain axioms. Once the defining formula is written down, the inference problem can be solved directly by `SAT` solvers. The gaussoid axioms are already written as boolean formulas; the axioms for oriented gaussoids can be extracted from the 51 oriented 3-gaussoids. To obtain the formula on any ground set N, the axioms are replicated for every oriented 3-face of the N cube as explained in Section 1.3.2.

Since in an oriented gaussoid every CI statement has three possible states, at least two boolean variables have to be employed per CI statement. For **orientability testing** of a given gaussoid it is advantageous to allocate one variable $Z_{\mathsf{ij|K}}$ to decide whether $[\mathsf{ij}|\mathsf{K}]$ is `0` or not and another variable $S_{\mathsf{ij|K}}$ to decide its sign in case $\neg Z_{\mathsf{ij|K}}$ holds. This allows orientable gaussoids to be described as a **projection** of the satisfying assignments of the oriented gaussoid formula. In order to write a formula for which a `#SAT` solver gives correct counts and an `AllSAT` solver does not enumerate the same oriented gaussoid twice, one of the four resulting states of each pair of Z and S variables is forbidden.

Algorithm 1 Orientability testing

function is-orientable($\mathsf{N}, \mathcal{G}$)
  $A \leftarrow$ oriented-axioms(N)
  $A \leftarrow A \wedge Z_{\mathsf{ij|K}}$ **for** $(\mathsf{ij}|\mathsf{K}) \in \mathcal{G}$
  return `SAT`(A) ▷ or `AllSAT`(A) to get all orientations
end function

The **orientable completion** $\overline{\mathfrak{o}}$ (see Definition 3.38) of a gaussoid is easy to compute as well using multiple invocations of the solver. In practice, `SAT` solvers such as `CaDiCaL` [Bie19] allow a formula to be stored and the `SAT` problem on it to be solved multiple times with different assumptions about the variables. This would reduce the redundant enumerations of orientation axioms in this algorithm:

Algorithm 2 Orientable completion

function orientable-completion($\mathsf{N}, \mathcal{G}$)
  $\overline{\mathcal{G}} \leftarrow \mathcal{G}$
  for all $(\mathsf{ij}|\mathsf{K}) \in \mathcal{A}_{\mathsf{N}} \setminus \mathcal{G}$ **do**
    $A \leftarrow$ oriented-axioms(N)
    $A \leftarrow A \wedge Z_{\mathsf{kl|M}}$ **for** $(\mathsf{kl}|\mathsf{M}) \in \mathcal{G}$
    $A \leftarrow A \wedge \neg Z_{\mathsf{ij|K}}$
    $\overline{\mathcal{G}} \leftarrow \overline{\mathcal{G}} \cup (\mathsf{ij}|\mathsf{K})$ **if** `SAT`$(A) =$ **false**
  end for
  return $\overline{\mathcal{G}}$
end function

Based on these algorithms the orientable 4- and 5-gaussoids can be enumerated. We report on the results of these computations:

Computation 6.18. A 4-gaussoid is realizable if and only if it is orientable. Thus, the orientability axioms on $n = 4$ are, modulo the gaussoid axioms and deductive closure, the Lněnička–Matúš axioms (LM.i)–(LM.v).

Corollary 6.19. The property $\mathfrak{o}$ is not closed under $\mathfrak{B}_\mathsf{N}$. □

Computation 6.20. Out of the 60 212 776 gaussoids on $n = 5$, exactly 20 584 290 are orientable. There are 175 215 classes modulo isomorphy and 87 834 classes modulo isomorphy and duality.

Example 6.21. The CI structure

$$\mathcal{G} = \{\, (12|), (13|4), (14|5), (23|5), (35|1), (45|2), (15|23), (34|12), (24|135) \,\}$$

from Example 3.34 is a gaussoid but it not orientable. Its orientable completion equals $\mathcal{A}_5 \setminus \{\, (25|\mathsf{K}) : \mathsf{K} \subseteq 134 \,\}$. Thus, orientability derives 63 CI statements which hold on the positive model of $\mathcal{G}$. Moreover, the orientable completion is an ascending gaussoid and therefore graphic by Proposition 6.12 and hence positively realizable by Theorem 4.6. In particular, $\mathcal{G}$ does not imply any more CI statements than found by its orientable completion. △

6.3 Semimatroids and information inequalities

The connection between Gaussian CI and semimatroids is established by the square trinomials (T.ii):

$$[\mathsf{ij}|\mathsf{L}]^2 = [\mathsf{iL}] \cdot [\mathsf{jL}] - [\mathsf{L}] \cdot [\mathsf{ijL}].$$

They show that the vanishing of almost-principal minors (however, not their signs) is already determined by polynomials in the principal minors. In this section, we therefore restrict the configuration space to the principal minor part. As with oriented gaussoids, we work over ordered fields only. Then, the fact that the left-hand side is a square means that the right-hand side is non-negative and thus $[\mathsf{iL}][\mathsf{jL}] \geq [\mathsf{L}][\mathsf{ijL}]$ holds. This is also known as the Koteljanskii inequality [JB93]. Conditional independence is characterized by the tightness of this inequality, or equivalently the equality in

$$\log[\mathsf{iL}] + \log[\mathsf{jL}] - \log[\mathsf{L}] - \log[\mathsf{ijL}] \geq 0.$$

With the theory and notation of poly- and semimatroids from Section 2.5, this condition can be recognized as the submodular inequality

$$\triangle \log \mathrm{pr}(\mathsf{ij}|\mathsf{L}) \geq 0,$$

where pr is the *principal minor map* which sends a symmetric (positive-definite) matrix to its vector of principal minors in $\mathbb{R}^{\mathcal{P}(\mathsf{N})}$. The image of this map was studied by Holtz and Sturmfels [HS07] and recently by Ahmadieh and Vinzant [AV21] who found a complete set-theoretic semialgebraic characterization of it over $\mathbb{R}$ and many other rings.

The log-principal minor vector $h = \log \mathrm{pr}\, \Sigma$ of a positive-definite matrix Σ satisfies the normalization condition $h(\emptyset) = \log 1 = 0$ and the submodular inequality. This vector has special significance in information theory because it appears in the *differential entropy* of a Gaussian distribution [Rao73, Section 8a.6],

$$H(\Sigma) = \frac{1}{2}\Big(|\mathsf{N}| \log(2\pi e) + \log\det\Sigma\Big).$$

For many purposes, including conditional independence, the scalar $1/2$ and the additive $|\mathsf{N}|\log(2\pi e)$ term do not matter because they cancel in all the difference expressions $\triangle(\mathsf{ij}|\mathsf{K})$. Therefore, $\log \operatorname{pr} \Sigma$ is essentially the differential entropy vector of the Gaussian. However, this vector is not generally a polymatroid because it need not be isotone with respect to inclusion. Antitone examples can be constructed from the Fischer inequality [JB93]: consider a *correlation* matrix Σ, i.e., $\Sigma \in \mathrm{PD}_{\mathsf{N}}$ with $\sigma_{\mathsf{ii}} = 1$ for all $\mathsf{i} \in \mathsf{N}$, then

$$\Sigma[\mathsf{iK}] \leq \underbrace{\Sigma[\mathsf{i}]}_{=1} \Sigma[\mathsf{K}] = \Sigma[\mathsf{K}].$$

However, by making σ_{ii} large enough, the inequality between $\Sigma[\mathsf{iK}]$ and $\Sigma[\mathsf{K}]$ can be reversed. This phenomenon was explained in the broader context of information inequalities by Chan [Cha03] who showed that an information inequality such as $h(\mathsf{iK}) - h(\mathsf{K}) \geq 0$ is valid for all continuous distributions only if it is *balanced*, i.e., for each $\mathsf{i} \in \mathsf{N}$, the coefficients of all terms in which i appears sum to zero. This is not the case for monotonicity and indeed it does not hold for the continuous Gaussian distributions.

The objective of this section is to prove that positive Gaussians are semimatroids and to extract valid inference rules from the face lattice of the polymatroid cone using linear programming. For further details on the connection between information theory and determinantal inequalities, see [Cha11] and [CGY12].

6.3.1 The multiinformation region. The log-principal minor map of a positive-definite matrix is submodular but not monotone. Thus, it is not clear whether positive Gaussians satisfy the definition of semimatroids. Picking the **correlation** matrix of a distribution — which does not change the CI structure, by Remark 3.31, but generally requires a euclidean ordered field — gives a canonical equivalent distribution whose log-principal minor vector is normalized, submodular and antitone. These are properties enjoyed by the (negative of the) *multiinformation function* of a distribution. In fact, by [Stu05, Corollary 2.6], the log-principal minor vector of a correlation matrix is almost exactly the multiinformation vector of the Gaussian distribution: they differ by a factor of $-1/2$. The polyhedral approximation to multiinformation functions, corresponding to the role polymatroids play for discrete entropies, are the ℓ-standardized supermodular functions in the terminology of [Stu05, Section 2.4].

Definition 6.22. An *ℓ-standardized supermodular function* on N is an $m \in \mathbb{R}^{\mathcal{P}(\mathsf{N})}$ satisfying

Standardization: $m(\emptyset) = 0$ and $m(\mathsf{k}) = 0$ for all $\mathsf{k} \in \mathsf{N}$,
Supermodularity: $\triangle m(\mathsf{ij}|\mathsf{K}) \leq 0$, for all $(\mathsf{ij}|\mathsf{K}) \in \mathcal{A}_{\mathsf{N}}$.

The set of all ℓ-standardized supermodular functions on N is evidently a pointed, rational, closed convex cone denoted M_{N}.

Corollary 6.23. Let $\Sigma \in \mathrm{PD}_{\mathsf{N}}(\mathbb{K})$ be a correlation matrix. Then $-\log \operatorname{pr} \Sigma$ is ℓ-standardized and supermodular. □

The following result seems to be folklore, but no reference is available:

Lemma 6.24. The linear map $f : h \mapsto m$ of $\mathbb{R}^{\mathcal{P}(\mathsf{N})}$ given by $m(\mathsf{K}) := \sum_{\mathsf{k}} h(\mathsf{k}) - h(\mathsf{K})$ establishes a bijection between $H^{\mathrm{ti}}_{\mathsf{N}}$ and M_{N}. This map negates the value of all functionals $\triangle(\mathsf{ij}|\mathsf{K})$, i.e., $\triangle f(h)(\mathsf{ij}|\mathsf{K}) = -\triangle h(\mathsf{ij}|\mathsf{K})$.

Proof. This construction is essentially given in [SV98, Section 2.5]. The multiinformation of a multivariate distribution is defined as its conditional entropy with respect to the product of its marginals, which directly inspires the given map. It is an easy calculation to see that $f(H^{\mathrm{ti}}_{\mathsf{N}}) \subseteq M_{\mathsf{N}}$ and that $\triangle f(h)(\mathsf{ij}|\mathsf{K}) = -\triangle h(\mathsf{ij}|\mathsf{K})$. To check that f is a bijection between the cones, we construct its inverse g. For a vector $m \in \mathbb{R}^{\mathcal{P}(\mathsf{N})}$ define its image $h = g(m)$ via $h(\emptyset) = 0$, $h(\mathsf{k}) = m(\mathsf{N}) - m(\mathsf{N} \setminus \mathsf{k})$ and $h(\mathsf{K}) = \sum_{\mathsf{k} \in \mathsf{K}} h(\mathsf{k}) - m(\mathsf{K})$. This map sends M_{N} to $H_{\mathsf{N}}^{\mathsf{ti}}$ and composes with f on both sides to the identity map. □

Theorem 6.25. Positive Gaussians are semimatroids: $\mathfrak{g}^+ \le \mathfrak{sm}$.

Proof. Let $\Sigma \in \mathrm{PD}_{\mathsf{N}}(\mathbb{K})$. By passing to the real closure of $\mathbb{K}$, we can assume that Σ is a correlation matrix without changing the realized gaussoid. The log-principal minor vector $\log \mathrm{pr}\, \Sigma$ is (up to a scalar) an ℓ-standardized supermodular function and thus $[\![\Sigma]\!]$ corresponds to a face of M_{N}. Since M_{N} and $H_{\mathsf{N}}^{\mathsf{ti}}$ have isomorphic face lattices by Lemma 6.24, it follows that $[\![\Sigma]\!]$ corresponds to a face of $H_{\mathsf{N}}^{\mathsf{ti}}$ and thus satisfies all the all axioms of the property $\mathfrak{sm}$ which are encoded in this lattice. Hence, $[\![\Sigma]\!]$ is a semimatroid. □

Matúš proved in [Mat97, Proposition 4] that semimatroids have no finite forbidden-minor characterization. A closer look at the proof even shows

Theorem 6.26. The property $\mathfrak{g} \wedge \mathfrak{sm}$ has no finite forbidden-minor characterization.

Proof. One of the two families of CI structures employed by Matúš is the same as the one in Example 4.40 which shows the non-finite axiomatizability of discrete and Gaussian CI. It is clear that these structures are gaussoids and therefore all their proper minors are gaussoids. This gives an example of an infinite family of minor-minimal CI structures which do not belong to $\mathfrak{g} \wedge \mathfrak{sm}$. □

6.3.2 Applications of information inequalities. Theorem 6.25 makes it possible to use the face lattice of the tight polymatroid cone to deduce valid inference rules for Gaussian CI over ordered fields. Since the cone is rational, all obtainable inference rules are valid for the smallest positive property $\mathfrak{g}^+_{\mathbb{Q}}$. This also means that linear programming with integer matrices can be employed to find relatively interior points on faces of the cone (corresponding to invalidity of an inference rule) or exact, rational Farkas certificates which are the linear version of final polynomials as validity proofs for an inference rule. Such proofs take the form of linear combinations of the facet defining inequalities. We give two examples here:

Example 6.27. The proof of Lemma 2.18 used the equality

$$\triangle h(\mathsf{ij}|\mathsf{kL}) + \triangle h(\mathsf{ik}|\mathsf{L}) = \triangle h(\mathsf{ij}|\mathsf{L}) + \triangle h(\mathsf{ik}|\mathsf{jL}).$$

This linear “final polynomial” is a statement about the face lattice of the cone of tight polymatroids: if a point lies on the facets $(\mathsf{ij}|\mathsf{kL})$ and $(\mathsf{ik}|\mathsf{L})$ simultaneously, then it lies on $(\mathsf{ij}|\mathsf{L})$ and $(\mathsf{ik}|\mathsf{jL})$ as well. This is an inference rule for semimatroids (in fact, it is the semigraphoid axiom). △

Example 6.28. That the family of CI structures used in Example 4.40 and Theorem 6.26 are not semimatroids follows from the equality

$$\triangle(01|2) + \triangle(02|3) \cdots + \triangle(0\mathsf{n}|1) = \triangle(02|1) + \triangle(03|2) \cdots + \triangle(01|\mathsf{n}). \qquad \triangle$$

In general, whenever an information inequality of the form

$$\sum_{(\mathsf{ij}|\mathsf{K}) \in \mathcal{L}} \alpha_{(\mathsf{ij}|\mathsf{K})} \triangle(\mathsf{ij}|\mathsf{K}) \ge \sum_{(\mathsf{kl}|\mathsf{M}) \in \mathcal{M}} \beta_{(\mathsf{kl}|\mathsf{M})} \triangle(\mathsf{kl}|\mathsf{M}), \quad \text{with } \alpha, \beta > 0, \qquad (\triangle\Rightarrow)$$

is valid for all log-principal minor vectors of positive-definite matrices, then this proves the inference rule $\bigwedge \mathcal{L} \Rightarrow \bigwedge \mathcal{M}$. Verifying such inequalities for **outer polyhedral approximations** to the region of log-principal minor vectors or, equivalently, Gaussian multiinformation functions, is a task for linear programming.

Algorithm 3 Semimatroid completion

function semimatroid-completion($\mathsf{N}, \mathcal{L}$)
    $\overline{\mathcal{L}} = \mathcal{L}$
    for all $(\mathsf{ij}|\mathsf{K}) \in \mathcal{A}_\mathsf{N} \setminus \mathcal{L}$ **do**
        $P \leftarrow$ tight-polymatroids(N)
        $P \leftarrow P \wedge \triangle(\mathsf{kl}|\mathsf{M}) = 0$ **for** $(\mathsf{kl}|\mathsf{M}) \in \mathcal{L}$
        $P \leftarrow P \wedge \triangle(\mathsf{ij}|\mathsf{K}) > 0$
        $\overline{\mathcal{L}} \leftarrow \overline{\mathcal{L}} \cup (\mathsf{ij}|\mathsf{K})$ **if** polyhedron$(P) = \emptyset$
    end for
    return $\overline{\mathcal{L}}$
end function

Since semimatroids are closed under intersection, the completion $\overline{\mathfrak{sm}}$ is exactly equal to $\mathfrak{sm}$. Thus, the above algorithm computes the uniquely determined semimatroid *closure* of a given CI structure. It can be short-circuited to serve as a semimatroid test as well.

Computation 6.29. A 4-gaussoid is realizable if and only if it is a semimatroid. Out of the 508 817 isomorphy classes of 5-gaussoids, exactly 336 838 are semimatroids.

6.4 Structural selfadhesivity

In 2007, Matúš introduced the notion of selfadhesive polymatroids [Mat07a] to mimic a curious property of entropy vectors of discrete random variables, but notions of adhesivity and amalgamation are classic topics in matroid theory [Oxl11, Section 11.4]. The picture to have in mind is that of gluing together two geometric objects along a shared subconfiguration by identifying the two copies of that configuration. The underlying concept — and a gem of synthetic probability theory — is the *Copy lemma* [ZY98, Kac13]: given jointly distributed discrete random variables ijk there exists a further discrete random variable k' such that ik' has the same distribution as ik and the conditional independence $(\mathsf{k}', \mathsf{jk}|\mathsf{i})$ holds. In very rough and informal analogy to a geometric construction, the idea is to obtain a copy k' of the variable k which is "glued" to i just like k is but "orthogonal" to jk given i. This result follows from a construction known as the *conditional product* [Stu05, Section 2.3.3].

This section transfers the selfadhesivity concept to Gaussians and then to properties of CI structures in general. Let $\mathsf{L} \subseteq \mathsf{N}$ be fixed. An L-*copy of* N is a finite set M of the same size as N with $\mathsf{N} \cap \mathsf{M} = \mathsf{L}$. Polymatroids on M are naturally identified with those on N.

Definition 6.30. A polymatroid (N, h) is *selfadhesive at* L if for an L-copy M of N there exists a polymatroid $(\mathsf{NM}, \overline{h})$ such that

— $\overline{h}|_\mathsf{N} = h$ and $\overline{h}|_\mathsf{M} = h$ up to isomorphy, and
— $\triangle\overline{h}(\mathsf{N}, \mathsf{M}) = 0$.

(N, h) is *selfadhesive* if it is selfadhesive at all $\mathsf{L} \subseteq \mathsf{N}$.

The second condition is information-theoretic: it means that the two copies N and M are independent given their overlap L, or that $(\mathsf{N}, \mathsf{M}|\mathsf{L}) \in [\![\overline{h}]\!]$ (modulo the localization rule (L)). Before we reinterpret and extend this definition of selfadhesivity to arbitrary properties of CI structures, we state and prove the analogous selfadhesivity result for positive-definite matrices, which even features the uniqueness of the distribution:

Theorem 6.31. Let $\mathbb{K}$ be an ordered field. For every $\Sigma \in \mathrm{PD}_{\mathsf{N}}(\mathbb{K})$ and every $\mathsf{L} \subseteq \mathsf{N}$ together with an L-copy M of N there exists a unique $\Phi \in \mathrm{PD}_{\mathsf{NM}}(\mathbb{K})$ such that:

— $\Phi_{\mathsf{N}} = \Sigma = \Phi_{\mathsf{M}}$, and
— $\operatorname{rk} \Phi_{\mathsf{N},\mathsf{M}} = |\mathsf{N} \cap \mathsf{M}| = |\mathsf{L}|$.

Proof. Define $\mathsf{N}' = \mathsf{N} \setminus \mathsf{L}$, $\mathsf{M}' = \mathsf{M} \setminus \mathsf{L}$ and consider the matrix

$$\Phi = \begin{array}{c@{}c} \begin{array}{ccc} \mathsf{L} & \mathsf{N}' & \mathsf{M}' \end{array} & \\ \begin{pmatrix} \Sigma_{\mathsf{L}} & \Sigma_{\mathsf{L},\mathsf{N}'} & \Sigma_{\mathsf{L},\mathsf{N}'} \\ \Sigma_{\mathsf{N}',\mathsf{L}} & \Sigma_{\mathsf{N}'} & \Lambda \\ \Sigma_{\mathsf{N}',\mathsf{L}} & \Lambda & \Sigma_{\mathsf{N}'} \end{pmatrix} & \begin{array}{l} \mathsf{L} \\ \mathsf{N}' \\ \mathsf{M}' \end{array} \end{array},$$

where Λ will be determined shortly. Its restrictions to N and M are clearly equal to Σ.

By the rank additivity formula for Schur complements,

$$|\mathsf{L}| \overset{!}{=} \operatorname{rk} \Phi_{\mathsf{N},\mathsf{M}} = \operatorname{rk} \begin{array}{c@{}c} \begin{array}{cc} \mathsf{L} & \mathsf{M}' \end{array} & \\ \begin{pmatrix} \Sigma_{\mathsf{L}} & \Sigma_{\mathsf{L},\mathsf{N}'} \\ \Sigma_{\mathsf{N}',\mathsf{L}} & \Lambda \end{pmatrix} & \begin{array}{l} \mathsf{L} \\ \mathsf{N}' \end{array} \end{array} = \underbrace{\operatorname{rk}(\Sigma_{\mathsf{L}})}_{=|\mathsf{L}|} + \operatorname{rk}(\Lambda - \Sigma_{\mathsf{N}',\mathsf{L}} \Sigma_{\mathsf{L}}^{-1} \Sigma_{\mathsf{L},\mathsf{N}'}),$$

the rank requirement necessitates $\Lambda = \Sigma_{\mathsf{N}',\mathsf{L}} \Sigma_{\mathsf{L}}^{-1} \Sigma_{\mathsf{L},\mathsf{N}'} = \Sigma_{\mathsf{L}} - \Sigma \,/\, \mathsf{L}$. Thus, Φ is uniquely determined by Σ via the two conditions. It remains to show its positive definiteness. We have the block LDU decomposition

$$\begin{aligned} P^{\mathsf{T}} \Phi P &= \begin{pmatrix} \mathbb{1}_{\mathsf{L}} & 0 & 0 \\ -\Sigma_{\mathsf{L}}^{-1} \Sigma_{\mathsf{L},\mathsf{N}'} & \mathbb{1}_{\mathsf{N}'} & 0 \\ -\Sigma_{\mathsf{L}}^{-1} \Sigma_{\mathsf{L},\mathsf{N}'} & 0 & \mathbb{1}_{\mathsf{M}'} \end{pmatrix} \begin{pmatrix} \Sigma_{\mathsf{L}} & \Sigma_{\mathsf{L},\mathsf{N}'} & \Sigma_{\mathsf{L},\mathsf{N}'} \\ \Sigma_{\mathsf{N}',\mathsf{L}} & \Sigma_{\mathsf{N}'} & \Lambda \\ \Sigma_{\mathsf{N}',\mathsf{L}} & \Lambda & \Sigma_{\mathsf{N}'} \end{pmatrix} \begin{pmatrix} \mathbb{1}_{\mathsf{L}} & -\Sigma_{\mathsf{L}}^{-1} \Sigma_{\mathsf{L},\mathsf{N}'} & -\Sigma_{\mathsf{L}}^{-1} \Sigma_{\mathsf{L},\mathsf{N}'} \\ 0 & \mathbb{1}_{\mathsf{N}'} & 0 \\ 0 & 0 & \mathbb{1}_{\mathsf{M}'} \end{pmatrix} \\ &= \begin{pmatrix} \Sigma_{\mathsf{L}} & 0 & 0 \\ 0 & \Sigma \,/\, \mathsf{L} & 0 \\ 0 & 0 & \Sigma \,/\, \mathsf{L} \end{pmatrix} = \Sigma_{\mathsf{L}} \oplus (\Sigma \,/\, \mathsf{L}) \oplus (\Sigma \,/\, \mathsf{L}). \end{aligned}$$

The matrix P is invertible and thus the transformation preserves positive definiteness (this holds over every ordered field as is easily seen via Tarski's transfer principle similarly to Lemma 3.12). The matrix on the right is clearly positive definite and then so is Φ. □

Remark 6.32. The Zhang–Yeung inequality [ZY98] was historically the first discovered valid information inequality which does not follow from the Shannon inequalities. It has the same significance for the theory of discrete CI structures as the Ingleton inequality [Ing71] has for matroid theory. Matúš [Mat07a, Section 4] writes this inequality as follows:

$$\triangledown_{\mathsf{i},\mathsf{j}|\mathsf{kl}} := \triangle(\mathsf{kl}|\mathsf{i}) + \triangle(\mathsf{kl}|\mathsf{j}) + \triangle(\mathsf{ij}|) - \triangle(\mathsf{kl}|) + \triangle(\mathsf{ik}|\mathsf{l}) + \triangle(\mathsf{il}|\mathsf{k}) + \triangle(\mathsf{kl}|\mathsf{i}) \overset{!}{\geq} 0$$

and proves that over ground sets of size four the selfadhesive polymatroids are precisely those which satisfy all six versions of the Zhang–Yeung inequality. Since the above expression is written in $\triangle$ terms, it makes sense to interpret it on multiinformation functions. Then, as a corollary to Theorem 6.31 we obtain that Gaussian multiinformations satisfy the Zhang–Yeung inequality. This is one half of the result proved by Lněnička [Lně03].

The proof requires positive-definite matrices. It fails to produce a selfadhesive extension which is principally regular if Σ is only principally regular:

Example 6.33. Consider the following principally regular realization of the gaussoid $\mathcal{G} := (12|*) \cup (23|*) \cup (24|*) \cup \{ (14|), (14|3) \}$ on $\mathsf{N} = 1234$ (where $(ij|*) := \{ (ij|K) : K \subseteq N^{ij} \}$):

$$\begin{array}{c c}
\begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} & \\
\left(\begin{array}{cc|cc}
1 & 0 & \frac{7}{8} & 0 \\
0 & 1 & 0 & 0 \\
\hline
\frac{7}{8} & 0 & 1 & -\frac{\sqrt{1695}}{64} \\
0 & 0 & -\frac{\sqrt{1695}}{64} & 1
\end{array}\right) & \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array}
\end{array}$$

and fix $\mathsf{L} = 12$. By the proof of Theorem 6.31, the submatrix and rank conditions uniquely determine an L-selfadhesive extension for this matrix over the ground set $12343'4'$. The candidate matrix is

$$\begin{array}{c c}
\begin{array}{cccccc} 1 & 2 & 3 & 4 & 3' & 4' \end{array} & \\
\left(\begin{array}{cc|cc|cc}
1 & 0 & \frac{7}{8} & 0 & \frac{7}{8} & 0 \\
0 & 1 & 0 & 0 & 0 & 0 \\
\hline
\frac{7}{8} & 0 & 1 & -\frac{\sqrt{1695}}{64} & \frac{49}{64} & 0 \\
0 & 0 & -\frac{\sqrt{1695}}{64} & 1 & 0 & 0 \\
\hline
\frac{7}{8} & 0 & \frac{49}{64} & 0 & 1 & -\frac{\sqrt{1695}}{64} \\
0 & 0 & 0 & 0 & -\frac{\sqrt{1695}}{64} & 1
\end{array}\right) & \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 3' \\ 4' \end{array}
\end{array}.$$

But this matrix is not principally regular, as the $343'$-principal minor is zero. On the other hand, $\mathcal{G}^{\rceil}$ is ascending and therefore graphical. Thus it is positively realizable over $\mathbb{Q}$. This positive realization is a selfadhesive algebraic realization by the theorem and thus $\mathcal{G}$ does belong to $(\mathfrak{g}^*_{\mathbb{Q}})^{\mathsf{sa}}$, even though the explicit algebraic realization above does not reveal this. With positive realizability this confusion does not arise. **Every** positive realization has all selfadhesive extensions, by Theorem 6.31. △

The selfadhesivity of positive-definite matrices induces similar properties on their CI structures. On the CI level, we use the term *structural selfadhesivity* to emphasize that it is generally a weaker notion than what is proved for realizations in Theorem 6.31.

Definition 6.34. Let $\mathfrak{p}$ be a property of CI structures. Define the *selfadhesion* $\mathfrak{p}^{\mathsf{sa}}(\mathsf{N})$ of $\mathfrak{p}$ as the set of CI structures $\mathcal{L}$ such that for every $\mathsf{L} \subseteq \mathsf{N}$ together with an L-copy M of N there exists $\overline{\mathcal{L}} \in \mathfrak{p}(\mathsf{NM})$ satisfying the two conditions:

— $\overline{\mathcal{L}}|_{\mathsf{N}} = \mathcal{L} = \overline{\mathcal{L}}|_{\mathsf{M}}$, and
— $(\mathsf{N}, \mathsf{M}|\mathsf{L}) \in \overline{\mathcal{L}}$ in the sense of (L).

A property is *selfadhesive* if $\mathfrak{p} = \mathfrak{p}^{\mathsf{sa}}$.

Lemma 6.35. The operator $\cdot^{\mathsf{sa}}$ is recessive and monotone on the property lattice $\mathfrak{P}$.

Proof. Let $\mathfrak{p}$ be a property and pick $\mathsf{L} = \mathsf{N}$. Let $\mathcal{L} \in \mathfrak{p}^{\mathsf{sa}}(\mathsf{N})$. In particular $\mathcal{L}$ is selfadhesive with respect to $\mathfrak{p}$ at $\mathsf{L} = \mathsf{N}$. In this case the L-copy M of N in the definition must be $\mathsf{M} = \mathsf{N}$ and it follows that $\mathcal{L} \in \mathfrak{p}(\mathsf{NM}) = \mathfrak{p}(\mathsf{N})$. This proves recessiveness $\mathfrak{p}^{\mathsf{sa}} \le \mathfrak{p}$. For monotonicity, let $\mathfrak{p} \le \mathfrak{q}$ and $\mathcal{L}$ in $\mathfrak{p}^{\mathsf{sa}}(\mathsf{N})$. Then for every L with L-copy M of N there exist a certificate for the existence of $\mathcal{L}$ in $\mathfrak{p}^{\mathsf{sa}}$. This certificate lives in $\mathfrak{p}(\mathsf{NM}) \subseteq \mathfrak{q}(\mathsf{NM})$ which proves $\mathcal{L} \in \mathfrak{q}^{\mathsf{sa}}(\mathsf{N})$. □

Lemma 6.36. If $\mathfrak{g}^+ \le \mathfrak{p}$, then $\mathfrak{g}^+ \le \mathfrak{p}^{\mathsf{sa}}$.

Proof. This follows from monotonicity of $\cdot^{\mathsf{sa}}$ and the fact that $\mathfrak{g}^+$ is a fixed point. □

Theorem 6.31 shows that every positive-definite matrix is selfadhesive as a matrix. This is of course much stronger than required for structural selfadhesivity of $\mathfrak{g}^+$ over ordered fields. The same does not hold for principally regular matrices, as shown in Example 6.33. For $\mathfrak{g}^*$ to be selfadhesive, every algebraic Gaussian only needs for each L **some** matrix in its realization space which has a selfadhesive extension at L.

Conjecture 6.37. The property $\mathfrak{g}^*_{\mathbb{C}}$ is not selfadhesive.

Question 6.38. Does $\cdot^{\mathsf{sa}}$ stabilize after the first application to "well-behaved" properties like semigraphoidality? Under which additional assumptions is is an interior operator?

6.4.1 Applications of selfadhesivity. Selfadhesivity can be applied to any of the necessary properties for positive realizability derived in this thesis to potentially improve them, Lemma 6.36. To check whether a CI structure on N is in $\mathfrak{p}^{\mathsf{sa}}(n)$ for some property $\mathfrak{p}$, any blackbox algorithm for deciding $\mathfrak{p}$ on partially defined structures may be used and run 2^n times. Given the speed of contemporary `SAT` and LP solvers, the following computations are feasible:

Computation 6.39. Out of the 1 512 isomorphy classes of 4-semigraphoids, precisely 1 352 are selfadhesive. All 4-gaussoids are selfadhesive. 485 727 out of 508 817 isomorphy classes of 5-gaussoids are selfadhesive.

Computation 6.40. There are 168 010 selfadhesive orientable 5-gaussoids and 335 047 gaussoids which are selfadhesive semimatroids, modulo isomorphy. The property $\mathfrak{o} \wedge \mathfrak{sm}$ of being an orientable gaussoid as well as a semimatroid has 175 139 isomorphy classes, its selfadhesion $(\mathfrak{o} \wedge \mathfrak{sm})^{\mathsf{sa}}$ only 167 989.

7

Summary and open problems

This thesis dealt with Gaussian conditional independence structures and the inference problem from a combinatorial, logical and geometric point of view. The main results are:

> **Chapter 3:** The Gaussian CI inference problem is an essentially geometric problem which can be studied over every field. Proofs and refutations for the validity of any inference rule can be effectively computed.
>
> **Chapter 4:** The set of valid inference rules for Gaussians over infinite fields has no finite characterization, but all inference rules with at most two antecedents follow from the gaussoid axioms.
>
> **Chapter 5:** The inference problem is hard in terms of its computational complexity, the field extensions involved in writing down invalidity proofs and the topology of the set of counterexamples.
>
> **Chapter 6:** The relations on the Gaussian CI configuration space can be used to design approximations to the inference problem which allow the finding of certain types of valid inference more efficiently, despite the problem being generally hard.

In addition to the questions and conjectures posed in the relevant sections themselves, the following paragraphs indicate directions for further research.

7.1 Special geometry. Much of the algebraic and geometric treatment of Gaussian CI structures is owed to matroid theory, in particular the concept of final polynomials in Section 3.6 and the universality results in Chapter 5. Interestingly, matroids and gaussoids both concern certain **special** types of subdeterminants of matrices which induce a combinatorial version of the Zariski topology. The existence of final polynomials as universal non-realizability certificates follows in both cases from the general theorems of the alternative in algebraic or semialgebraic geometry. It would be interesting to pursue the reverse mathematics of such **special geometries** further.

Concerning the parallel development of gaussoids over hyperfields presented in Section 6.2, the following concrete conjecture about whether the square and edge trinomials truly play the same role for this theory as the 3-term Grassmann–Plücker relations do for matroids were presented in [BDKS19] and remain unsolved:

Conjecture 7.1. Every gaussoid is a root over $\mathtt{K}$ of all quadrics in $\mathcal{J}$, not just (T.i) and (T.ii).

Conjecture 7.2. The ideal $\mathcal{J}$ is generated by its quadrics.

7.2 Asymptotics of properties. During the proof of the existence of infinitely many forbidden minors for Gaussians over infinite fields in Section 4.5, it is shown that the number of realizable gaussoids is bounded from above by $\log|\mathfrak{g}^\bullet(n)| \in \mathcal{O}(n^3)$. On the other hand, graphic gaussoids, via Theorem 4.6, give a lower bound of $\Omega(n^2)$.

Conjecture 7.3. Asymptotically almost every realizable gaussoid is not graphic. More precisely, there exists $\varepsilon > 0$ such that $\log|\mathfrak{g}^\bullet(n)| \in \Omega(n^{2+\varepsilon})$ for all infinite fields.

Question 7.4. Given that not all realizable gaussoids are rationally realizable (Theorem 5.34), what is the asymptotic growth of $\mathfrak{g}^+_\mathbb{Q}$ compared to $\mathfrak{g}^+_\mathbb{R}$? What about $\mathfrak{g}^+_\mathbb{R}$ compared to $\mathfrak{g}^*_\mathbb{C}$?

Question 7.5. Is asymptotically almost every realizable gaussoid realizable in every ε-ball around the identity matrix?

7.3 More universalities. The universality theorems in Chapter 5 cover the algorithmic, algebraic and topological complexity of the (oriented) inference problem for Gaussian CI. In the literature on matroids, there are results about the preservance of differential geometric structure [Gün96] and, in the algebraic case, the birational type of varieties [BS89, Theorem 4.30] which are not immediate corollaries of our treatment. Geometric constraint satisfaction problems [MS21b] offer a general framework to classify the **kinds** of polynomials required to attain such universality results. It would be especially interesting to study matrix-subdeterminant constraint languages in greather generality, matroid and gaussoid theory being two instances which attain universality.

The universality results for Gaussian CI models likely have implications for the worst-case complexity of standard optimization problems over these models in statistics such as maximum likelihood estimation.

Question 7.6. How hard is it to check, in the worst case with a rational sample, whether two points lie in the same log-Voronoi cell of a Gaussian CI model? (Note that the maximum likelihood estimators of **mixtures** of Gaussians are known to be transcendental in the worst case [ADS16].)

Finally, in the context of Question 5.41 about the unbounded number of consequents in the inference rules produced by the von Staudt constructions, it is surprising that no example for the following statement about the axioms of Gaussians has been found yet:

Conjecture 7.7. There is a minimal valid inference rule for $\mathfrak{g}^+_\mathbb{R}$ which is not implied by a set of Horn clauses and Weak transitivity (G.iv).

7.4 The catalogue of realizable 5-gaussoids. Since the classification of positively realizable gaussoids on four elements was achieved by Lněnička and Matúš [LM07], Bernd Sturmfels repeatedly asked for the classification on five random variables, including at the session in which the author was first exposed to Gaussian conditional independence structures and in the joint paper [BDKS19, Challenge 1]. This classification is still incomplete. The main practical result of Chapter 6 is the reduction of the 254 826 classes of gaussoids modulo isomorphy and duality to 84 434 classes in $(\mathfrak{o} \wedge \mathfrak{sm})^{\mathsf{sa}}$. The classification task seems to be achievable, however the extension to six random variables is likely hopeless since attempts to even count the number of 6-gaussoids using `GANAK` [SRSM19] have failed. The large symmetry group which cannot be factored out by this `#SAT` solver contributes to the problem and an *orderly algorithm* [MR08, Section 4] to enumerate gaussoids modulo isomorphy may resolve this issue.

Multiple refinements to the techniques presented in this thesis are possible:

(1) Gaussoids with coefficients in the *tropical hyperfield* have already been introduced as *valuated gaussoids* in [BDKS19] but their study was not continued so far. Unlike `K` and `S`, the

tropical hyperfield T is not finite, thus the combinatorial methods employing `SAT` solvers are of no use. The resulting coarsening of the geometry of the Gaussian CI configuration space is through the lens of a valuation. In connection with our algebraic realizability technique Lemma 4.1 and the *Puiseux series field*, this is a promising direction for obtaining more refined inference rules.

(2) The results of Ahmadieh and Vinzant [AV21] imply that the image of the PD cone under the principal minor map is cut out of the positive orthant in $\mathbb{R}^{\mathcal{P}(\mathsf{N})}$ (set-theoretically) by the Cayley $2\times2\times2$ hyperdeterminant under the action of the group $\mathrm{SL}_2(\mathbb{R})^{\mathsf{N}} \rtimes \mathfrak{S}_{\mathsf{N}}$. In light of Section 6.3, the tropicalization of these polynomials yields information inequalities, whose usefulness is yet to be evaluated.

(3) Algebraic realization spaces of gaussoids or Zariski-closed supersets of them are sometimes quickly computable and may prove non-realizability if they are empty. This routine has not been systematically used on the remaining gaussoids yet.

This thesis focused on finding inference rules for Gaussians, which are equivalent to non-realizability proofs, thus approximating the realizable structures from above. This is only half of the work, the other being to find realizability certificates for a subset of the remaining gaussoids. Apart from recognizing a gaussoid, or its completion under orientability and semimatroidality axioms, as a graphical model, no easy sufficient realizability criteria are known. For the development of the `CInet tools` software package and publication of the final database of Gaussian CI structures at `https://conditional-independence.net`, **certifiability** of all assertions about realizability is most important.

7.5 Discrete vs. Gaussian realizability. A fundamental question for the theory of Gaussian CI was posed in Studený's book [Stu05, Question 3] and remains unsolved:

Studený's question. Is every positive Gaussian CI structure realizable by discrete or even by positive binary random variables?

The thesis of his student Šimeček contains the following remarkable observations suggesting an affirmative answer [Šim07, Section 1.2.3]. A binary random vector on ground set N is specified by its 2^{N} atomic probabilities which must sum to one. Alternatively, it is specified by the 2^{N} (square-free) moments where the zeroth moment equals 1. Šimeček derives the conditional independence equations for conditioning sets of size ≤ 2 in terms of the moments. If the state space of each variable is $\{\pm 1\}$ and all moments except the ones of second-order e_{ij} are bound by obvious necessary equations, then the CI equations coincide with those of the regular Gaussians when written in the matrix entries σ_{ij} of a correlation matrix. This correspondence of CI equations breaks down for conditioning sets of size 3, where the ideal of binary equations properly contains the Gaussian almost-principal minor. This is a worthwhile direction to follow.

On the other hand, the answer to the generalization of Studený's question to algebraic Gaussians is negative even over $\mathbb{Q}$ and even for $n = 4$. This follows from the characterization of discretely representable CI structures on four random variables [Mat99a, Šim06c] and the observation that a 4-gaussoid is positively realizable if and only if it is discretely representable — however, there are algebraic 4-Gaussians which are not positive and hence not discrete. This may suggest a negative answer because the algebraic structure of Gaussian CI is seen in these cases to escape the flexibility even of arbitrary discrete distributions. But positive definiteness is a strong additional regularity condition for the induced CI structures, as exemplified by Example 6.33, and algebraic Gaussians may not even be semimatroids (Computation 6.29). The natural development on this side of the question would be to finish the classification of positive 5-Gaussians and check the realizable ones for discrete representability.

A counterexample to Studený's question is a Gaussian distribution which is not discretely realizable. In other words, it is an inference rule which is valid for discrete but not regular Gaussian random variables. The search for such properties of discrete CI was the original motivation behind studying semimatroids and structural selfadhesivity in the Gaussian setting. Namely, as a means of strengthening any property which is necessary for *discrete* representability to obtain one which is perhaps not necessary for Gaussians. This approach is proven futile by the selfadhesivity of Gaussians in Theorem 6.31: a property which is not already a counterexample to Studený's question cannot become a counterexample after applying $\cdot^{\mathsf{sa}}$ to it. Furthermore, CI inference rules which are derived from linear information inequalities of the type ($\triangle\Rightarrow$) cannot give a negative answer to Studený's question because these inequalities are balanced. By the results of Chan [Cha03], a balanced inequality is valid for discrete entropies if and only if it is valid for differential entropies of all continuous distributions, including Gaussians.

7.6 Near-identity realizability. Rational realizability near the identity matrix or its hyperoctahedral images is one of the main ingredients in the proof of Theorem 4.58. Denote this property by $\mathfrak{id}$. This notion is stable under minors, direct and dependent sum, embeddings and symmetries. The tools of Section 4.5 imply that this property has no finite axiomatization. This means that it is a well-behaved and non-trivial sufficient property for $\mathfrak{g}^*$.

Question 7.8. Does $\mathfrak{id}$ imply realizability in every characteristic?

7.7 More forms of realizability. Matroids in projective geometry have been studied over skew fields as well; see, e.g., [KPY20]. The notion of *quasideterminants* [GGRW05] provides a way to study non-commutative analogues of the principal and almost-principal minors of a symmetric matrix. Once the algebra and combinatorics of these objects is worked out, similarly to the early sections in Chapter 3, it might be possible to reuse the von Staudt constructions from Chapter 5 to obtain a more general universality theorem.

The fact that algebraic Gaussians over $\mathbb{C}$ are represented by symmetric principally regular matrices has its roots in the paper of Matúš [Mat05]. Matúš's reason for this is likely grounded in the use of Gröbner bases to obtain inference rules, since the symmetric matrices are just an affine space. Another natural idea, which has not been pursued in the context of Gaussian CI theory, is to study hermitian positive-definite matrices over $\mathbb{C}$. The definition of hermitian matrices requires complex conjugation and this changes the first-order theory away from routine algebraic geometry. Specifically, the real numbers can be defined in theory of the complex numbers with conjugation via $x = \overline{x}$. This makes the theory and hence perhaps the realizability problem for Gaussian CI structures more complicated than **ACF**.

7.8 Complexity of orientability. Gaussoid orientability has no finite axiomatization by Theorem 6.13. In the theory of matroids, orientability testing of rank-3 matroids is even known to be **NP**-complete [Ric99a, Tsc01]. The proof is based on combinatorics of pseudoline arrangements in the projective plane — a setting which gaussoid theory can approach via the constructions in Chapter 5. However, a gaussoid analogue of the Folkman–Lawrence representation theorem for oriented matroids [FL78, BMS01] is not known.

To state complexity-theoretic questions about gaussoids, their input format and its length must be defined. The obvious encoding as a binary string indexed by $\mathcal{A}_{\mathsf{N}}$ has exponential length in the ground set size $n = |\mathsf{N}|$ and is unsuitable. On the other hand, matroids

with a fixed rank have a polynomial-size encoding in their ground set size and the **NP**-completeness result is meaningful.

Since the encoding of projective ruler constructions in Chapter 5 required only conditioning sets of size up to three, a polynomial coding length can be achieved for "truncated" gaussoids $\mathcal{G} \cap \{\, (\mathsf{ij}|\mathsf{K}) \in \mathcal{A}_\mathsf{N} : |\mathsf{K}| \le 3 \,\}$ and perhaps this suffices for a construction. This truncation is meaningful from an information-theoretic perspective, because it limits the number of variables which can interact in any given CI statement.

Conjecture 7.9. The notion of orientability for a truncated gaussoid is meaningful and testing orientability of truncated gaussoids with conditioning sets of size at most three is **NP**-complete.

Bibliography

[AB09] Sanjeev Arora and Boaz Barak. *Computational complexity.* Cambridge University Press, 2009.

[Abe08] Francine F. Abeles. Dodgson condensation: The historical and mathematical development of an experimental method. *Linear Algebra Appl.*, 429(2-3):429–438, 2008.

[ADS16] Carlos Améndola, Mathias Drton, and Bernd Sturmfels. Maximum likelihood estimates for Gaussian mixtures are transcendental. In *Mathematical aspects of computer and information sciences. 6th international conference, MACIS 2015, Berlin, Germany, November 11–13, 2015. Revised selected papers*, pages 579–590. Springer, 2016.

[AH21] Yulia Alexandr and Alexander Heaton. Logarithmic Voronoi cells. *Algebr. Stat.*, 12(1):75–95, 2021.

[AV21] Abeer Al Ahmadieh and Cynthia Vinzant. Characterizing principal minors of symmetric matrices via determinantal multiaffine polynomials, 2021. arXiv:2105.13444.

[BB17] Matthew Baker and Nathan Bowler. Matroids over hyperfields, 2017. arXiv:1601.01204.

[BB19] Matthew Baker and Nathan Bowler. Matroids over partial hyperstructures. *Adv. Math.*, 343:821–863, 2019.

[BBPT14] Amos Beimel, Aner Ben-Efraim, Carles Padró, and Ilya Tyomkin. Multi-linear secret-sharing schemes. In *Theory of cryptography. Proceedings of the 11th theory of cryptography conference, TCC 2014, San Diego, CA, USA, February 24–26, 2014*, pages 394–418. Springer, 2014.

[BCE+21] Tobias Boege, Jane Ivy Coons, Christopher Eur, Aida Maraj, and Frank Röttger. Reciprocal maximum likelihood degrees of brownian motion tree models, 2021.

[BCR98] Jacek Bochnak, Michel Coste, and Marie-Françoise Roy. *Real algebraic geometry*, volume 36 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge.* Springer, 1998.

[BCS97] Peter Bürgisser, Michael Clausen, and Mohammad Amin Shokrollahi. *Algebraic complexity theory. With the collaboration of Thomas Lickteig*, volume 315 of *Grundlehren der Mathematischen Wissenschaften.* Springer, 1997.

[BDKS19] Tobias Boege, Alessio D'Alì, Thomas Kahle, and Bernd Sturmfels. The Geometry of Gaussoids. *Found. Comput. Math.*, 19(4):775–812, 2019.

[Ber15] George M. Bergman. *An invitation to general algebra and universal constructions.* Universitext. Springer, 2nd edition, 2015.

[BHLS10] Remco Bouckaert, Raymond Hemmecke, Silvia Lindner, and Milan Studený. Efficient algorithms for conditional independence inference. *J. Mach. Learn. Res.*, 11:3453–3479, 2010.

[Bie19] Armin Biere. CaDiCaL at the SAT Race 2019. In Marijn Heule, Matti Järvisalo, and Martin Suda, editors, *Proc. of SAT Race 2019 – Solver and Benchmark Descriptions*, volume B-2019-1 of *Department of Computer Science Series of Publications B*, pages 8–9. University of Helsinki, 2019.

[BISvdO] Winfried Bruns, Bogdan Ichim, Christof Söger, and Ulrich von der Ohe. Normaliz. Algorithms for rational cones and affine monoids. Available at `https://www.normaliz.uni-osnabrueck.de`.

[BK80] Thomas H. Brylawski and Douglas G. Kelly. *Matroids and combinatorial geometries*, volume 8 of *Carolina Lecture Series.* University of North Carolina, Department of Mathematics, 1980.

[BK20] Tobias Boege and Thomas Kahle. Construction Methods for Gaussoids. *Kybernetika*, 56(6):1045–1062, 2020.

[BKKR21] Tobias Boege, Thomas Kahle, Andreas Kretschmer, and Frank Röttger. The geometry of gaussian double markovian distributions, 2021. `arXiv:2107.00134v2`.

[BLS+99] Anders Björner, Michel Las Vergnas, Bernd Sturmfels, Neil White, and Günter Ziegler. *Oriented matroids*, volume 46 of *Encyclopedia of Mathematics and Its Applications.* Cambridge University Press, 2nd edition, 1999.

[BMS01] Jürgen Bokowski, Susanne Mock, and Ileana Streinu. On the Folkman-Lawrence topological representation theorem for oriented matroids of rank 3. *Eur. J. Comb.*, 22(5):601–615, 2001.

[Boe20] Tobias Boege. Gaussoids are two-antecedental approximations of Gaussian conditional independence structures, 2020. `arXiv:2010.11914`.

[Boe21] Tobias Boege. Incidence geometry in the projective plane via almost-principal minors of symmetric matrices, 2021. `arXiv:2103.02589`.

[BPR06] Saugata Basu, Richard Pollack, and Marie-Françoise Roy. *Algorithms in real algebraic geometry*, volume 10 of *Algorithms and Computation in Mathematics.* Springer, 2nd edition, 2006.

[Bre93] Glen E. Bredon. *Topology and geometry*, volume 139 of *Graduate Texts in Mathematics.* Springer, 1993.

[BS89] Jürgen Bokowski and Bernd Sturmfels. *Computational synthetic geometry*, volume 1355 of *Lecture Notes in Mathematics.* Springer, 1989.

[Can88] John Canny. Some algebraic and geometric computations in PSPACE. In *Proceedings of the Twentieth Annual ACM Symposium on Theory of Computing*, STOC '88, pages 460–467. Association for Computing Machinery (ACM), 1988.

[CG08] Terence H. Chan and Alex Grant. Non-linear information inequalities. *Entropy*, 10(4):765–775, 2008.

[CGY12] Terence H. Chan, Dongning Guo, and Raymond W. Yeung. Entropy functions and determinant inequalities. In *2012 IEEE International Symposium on Information Theory Proceedings*, pages 1251–1255. Institute of Electrical and Electronics Engineers (IEEE), 2012.

[Cha03] Terence H. Chan. Balanced information inequalities. *IEEE Transactions on Information Theory*, 49(12):3261–3267, 2003.

[Cha11] Terence H. Chan. Recent progresses in characterising information inequalities. *Entropy*, 13(2):379–401, 2011.

[CLO15] David A. Cox, John Little, and Donal O'Shea. *Ideals, varieties, and algorithms.* Undergraduate Texts in Mathematics. Springer, 4th edition, 2015.

[Coo71] Stephen A. Cook. The Complexity of Theorem-Proving Procedures. In *Proceedings of the Third Annual ACM Symposium on Theory of Computing*, STOC '71, pages 151–158. Association for Computing Machinery (ACM), 1971.

[Dat03] Ruchira S. Datta. Universality of Nash equilibria. *Math. Oper. Res.*, 28(3):424–432, 2003.

[Daw79] Alexander P. Dawid. Conditional independence in statistical theory. *J. Roy. Statist. Soc. Ser. B*, 41(1):1–31, 1979.

[DF04] David S. Dummit and Richard M. Foote. *Abstract algebra.* Wiley, 3rd edition, 2004.

[DSS09] Mathias Drton, Bernd Sturmfels, and Seth Sullivant. *Lectures on algebraic statistics*, volume 39 of *Oberwolfach Seminars.* Birkhäuser, 2009.

[DW87] Andreas Dress and Walter Wenzel. Matroids with coefficients. *Sémin. Lothar. Comb.*, 18, 1987.

[DW91] Andreas Dress and Walter Wenzel. Grassmann-Plücker relations and matroids with coefficients. *Adv. Math.*, 86(1):68–110, 1991.

[DX10] Mathias Drton and Han Xiao. Smoothness of Gaussian conditional independence models. In *Algebraic methods in statistics and probability II*, volume 516 of *Contemporary Mathematics*, pages 155–177. American Mathematical Society (AMS), 2010.

[Edm70] Jack Edmonds. Submodular functions, matroids, and certain polyhedra. In *Proceedings of the International Conference on Combinatorial Structures and Their Applications*, pages 69–87, 1970.

[Fag77] Ronald Fagin. Multivalued dependencies and a new normal form for relational databases. *ACM Trans. Database Syst.*, 2(3):262–278, 1977.

[FL78] Jon Folkman and Jim Lawrence. Oriented matroids. *J. Comb. Theory, Ser. B*, 25:199–236, 1978.

[Fuj78] Satoru Fujishige. Polymatroidal dependence structure of a set of random variables. *Inf. Control*, 39(1):55–72, 1978.

[GAE02] Zhiming Gong, Mohammad Aldeen, and Ludwig Elsner. A note on a generalized Cramer's rule. *Linear Algebra Appl.*, 340(1):253–254, 2002.

[GBE+18] Ambros Gleixner, Michael Bastubbe, Leon Eifler, Tristan Gally, Gerald Gamrath, Robert Lion Gottwald, Gregor Hendel, Christopher Hojny, Thorsten Koch, Marco Lübbecke, Stephen J. Maher, Matthias Miltenberger, Benjamin Müller, Marc Pfetsch, Christian Puchert, Daniel Rehfeldt, Franziska Schlösser, Christoph Schubert, Felipe Serrano, Yuji Shinano, Jan Merlin Viernickel, Matthias Walter, Fabian Wegscheider, Jonas T. Witt, and Jakob Witzig. The scip optimization suite 6.0. Technical Report 18-26, ZIB, Takustr. 7, 14195 Berlin, 2018.

[Gee99] James F. Geelen. On matroids without a non-fano minor. *Discrete Math.*, 203(1):279–285, 1999.

[GGRW05] Israel Gelfand, Sergei Gelfand, Vladimir Retakh, and Robert L. Wilson. Quasideterminants. *Adv. Math.*, 193(1):56–141, 2005.

[GH09] Steven Givant and Paul Halmos. *Introduction to Boolean algebras*. Undergraduate Texts in Mathematics. Springer, 2009.

[GP93] Dan Geiger and Judea Pearl. Logical and algorithmic properties of conditional independence and graphical models. *Ann. Statist.*, 21(4):2001–2021, 1993.

[GR18] Arley Ramsés Gómez Ríos. *On the theory of polynomial information inequalities*. Phd thesis, Departamento de Matemáticas, Universidad Nacional de Colombia, Bogotá, 2018.

[Grä11] George Grätzer. *Lattice theory: Foundation*. Birkhäuser, 2011.

[Grü03] Branko Grünbaum. *Convex polytopes*, volume 221 of *Graduate Texts in Mathematics*. Springer, 2nd edition, 2003. Prepared by Volker Kaibel, Victor Klee, and Günter M. Ziegler.

[GS] Daniel R. Grayson and Michael E. Stillman. Macaulay2, a software system for research in algebraic geometry. Available at http://www.math.uiuc.edu/Macaulay2/.

[GSW12] Ambros Gleixner, Daniel Steffy, and Kati Wolter. Improving the accuracy of linear programming solvers with iterative refinement. Technical Report 12-19, ZIB, Takustr. 7, 14195 Berlin, 2012.

[GSW15] Ambros Gleixner, Daniel Steffy, and Kati Wolter. Iterative refinement for linear programming. Technical Report 15-15, ZIB, Takustr. 7, 14195 Berlin, 2015.

[Gün96] Harald Günzel. The universal partition theorem for oriented matroids. *Discrete Comput. Geom.*, 15(2):121–145, 1996.

[Har77] Robin Hartshorne. *Algebraic geometry*, volume 52 of *Graduate Texts in Mathematics*. Springer, 1977.

[Har00] Robin Hartshorne. *Geometry: Euclid and beyond*. Undergraduate Texts in Mathematics. Springer, 2000.

[Hat02] Allen Hatcher. *Algebraic topology*. Cambridge University Press, 2002.

[HHK+19] Miika Hannula, Åsa Hirvonen, Juha Kontinen, Vadim Kulikov, and Jonni Virtema. Facets of distribution identities in probabilistic team semantics. In Francesco Calimeri, Nicola Leone, and Marco Manna, editors, *Logics in Artificial Intelligence*, volume 11468 of *Lecture Notes in Computer Science*, pages 304–320. Springer, 2019. The 16th European Conference on Logics in Artificial Intelligence (JELIA 2019).

[HK71] Kenneth Hoffman and Ray Kunze. *Linear algebra.* Prentice-Hall, 2nd edition, 1971.

[HK16] Miika Hannula and Juha Kontinen. A finite axiomatization of conditional independence and inclusion dependencies. *Inf. Comput.*, 249:121–137, 2016.

[HMS+08] Raymond Hemmecke, Jason Morton, Anne Shiu, Bernd Sturmfels, and Oliver Wienand. Three counter-examples on semi-graphoids. *Comb. Probab. Comput.*, 17(2):239–257, 2008.

[HS07] Olga Holtz and Bernd Sturmfels. Hyperdeterminantal relations among symmetric principal minors. *J. Algebra*, 316(2), 2007.

[HS11] Babak Hassibi and Sormeh Shadbakht. On the entropy region of gaussian random variables, 2011. arXiv:1112.0061.

[HW20] Daniel Hug and Wolfgang Weil. *Lectures on convex geometry*, volume 286 of *Graduate Texts in Mathematics*. Springer, 2020.

[Ing71] Aubrey W. Ingleton. Representation of matroids. In Dominic J. A. Welsh, editor, *Combinatorial mathematics and its applications. Proceedings, Oxford, 1969*, pages 149–167. Academic Press, 1971.

[ITK83] Minoru Ito, Kenichi Taniguchi, and Tadao Kasami. Membership problem for embedded multivalued dependencies under some restricted conditions. *Theor. Comput. Sci.*, 22:175–193, 1983.

[JB93] Charles R. Johnson and Wayne W. Barrett. Determinantal inequalities for positive definite matrices. *Discrete Math.*, 119(1-3):97–106, 1993.

[Jun08] Dieter Jungnickel. *Optimierungsmethoden. Eine Einführung.* Springer-Lehrbuch. Springer, 2nd edition, 2008.

[Jun21] Jaiung Jun. Geometry of hyperfields. *J. Algebra*, 569:220–257, 2021.

[JW05] Beatrix Jones and Mike West. Covariance decomposition in undirected Gaussian graphical models. *Biometrika*, 92(4):779–786, 2005.

[Kac13] Tarik Kaced. Equivalence of two proof techniques for non-shannon-type inequalities. In *2013 IEEE International Symposium on Information Theory*, pages 236–240. Institute of Electrical and Electronics Engineers (IEEE), 2013.

[Kem11] Gregor Kemper. *A course in commutative algebra*, volume 256 of *Graduate Texts in Mathematics*. Springer, 2011.

[KMP+08] Lutz Kettner, Kurt Mehlhorn, Sylvain Pion, Stefan Schirra, and Chee Yap. Classroom examples of robustness problems in geometric computations. *Comput. Geom.*, 40(1):61–78, 2008.

[Knu92] Donald E. Knuth. *Axioms and hulls*, volume 606 of *Lecture Notes in Computer Science*. Springer, 1992.

[KPY20] Lukas Kühne, Rudi Pendavingh, and Geva Yashfe. Von Staudt Constructions for Skew-Linear and Multilinear Matroids, 2020. arXiv:2012.07361.

[Lau96] Steffen Lauritzen. *Graphical models.*, volume 17 of *Oxford Statistical Science Series*. Oxford University Press, 1996.

[LB96] Michael Lundquist and Wayne Barrett. Rank inequalities for positive semidefinite matrices. *Linear Algebra Appl.*, 248:91–100, 1996.

[Li21a] Cheuk Ting Li. First-order theory of probabilistic independence and single-letter characterizations of capacity regions, 2021. arXiv:2108.07324.

[Li21b] Cheuk Ting Li. The undecidability of conditional affine information inequalities and conditional independence implication with a binary constraint, 2021. arXiv:2104.05634.

[LM07] Radim Lněnička and František Matúš. On Gaussian conditional independence structures. *Kybernetika*, 43(3):327–342, 2007.

[Lně03] Radim Lněnička. On the tightness of the Zhang-Yeung inequality for Gaussian vectors. *Commun. Inf. Syst.*, 3(1):41–46, 2003.

[LS18] Steffen Lauritzen and Kayvan Sadeghi. Unifying Markov properties for graphical models. *Ann. Statist.*, 46(5):2251–2278, 2018.

[Mac36] Saunders MacLane. Some interpretations of abstract linear dependence in terms of projective geometry. *Amer. J. Math.*, 58(1):236–240, 1936.

[Mar02] David Marker. *Model theory: An introduction*, volume 217 of *Graduate Texts in Mathematics*. Springer, 2002.

[Mar08] Murray Marshall. *Positive polynomials and sums of squares*, volume 146 of *Math. Surv. Monogr.* American Mathematical Society (AMS), 2008.

[Mat92] František Matúš. Ascending and descending conditional independence relations. In *Transactions of the 11th Prague Conference on Information Theory, Statistical Decision Functions and Random Processes*, volume B, pages 189–200, 1992.

[Mat94] František Matúš. Probabilistic conditional independence structures and matroid theory: background. *Int. J. Gen. Syst.*, 22:185–196, 1994.

[Mat95] František Matúš. Conditional independences among four random variables. II. *Combin. Probab. Comput.*, 4(4):407–417, 1995.

[Mat97] František Matúš. Conditional independence structures examined via minors. *Ann. Math. Artif. Intell.*, 21(1):99–30, 1997.

[Mat99a] František Matúš. Conditional independences among four random variables. III. Final conclusion. *Combin. Probab. Comput.*, 8(3):269–276, 1999.

[Mat99b] František Matúš. Matroid representations by partitions. *Discrete Math.*, 203(1-3):169–194, 1999.

[Mat04] František Matúš. Towards classification of semigraphoids. *Discrete Math.*, 277(1):115–145, 2004.

[Mat05] František Matúš. Conditional independences in gaussian vectors and rings of polynomials. In Gabriele Kern-Isberner, Wilhelm Rödder, and Friedhelm Kulmann, editors, *Conditionals, Information, and Inference*, pages 152–161. Springer, 2005.

[Mat07a] František Matúš. Adhesivity of polymatroids. *Discrete Math.*, 307(21):2464–2477, 2007.

[Mat07b] František Matúš. Infinitely many information inequalities. In *Proceedings of the IEEE ISIT 2007*, pages 41–44, 2007.

[Maz92] Barry Mazur. The topology of rational points. *Exp. Math.*, 1(1):35–45, 1992.

[Maz95] Barry Mazur. Speculations about the topology of rational points: An up-date. In *Columbia University number theory seminar, New York, 1992*, pages 165–181. Société Mathématique de France, 1995.

[MC16] František Matúš and Lászlo Csirmaz. Entropy region and convolution. *IEEE Trans. Inf. Theory*, 62(11):6007–6018, 2016.

[MMBH12] Christian Muise, Sheila A. McIlraith, J. Christopher Beck, and Eric I. Hsu. Dsharp: Fast d-DNNF Compilation with sharpSAT. In Leila Kosseim and Diana Inkpen, editors, *Advances in Artificial Intelligence*, pages 356–361. Springer, 2012.

[Mnë88] Nikolai E. Mnëv. The universality theorems on the classification problem of configuration varieties and convex polytopes varieties. In Oleg Y. Viro, editor, *Topology and geometry — Rohlin Seminar 1984–1986*, volume 1346 of *Lecture Notes in Mathematics*, pages 527–543. Springer, 1988.

[Mnë91] Nikolai E. Mnëv. The universality theorem on the oriented matroid stratification of the space of real matrices. In *Discrete and computational geometry: Papers from the DIMACS special year. Proceedings of workshops, held at DIMACS from September 1989 through August 1990*, pages 237–243. American Mathematical Society (AMS) and Association for Computing Machinery (ACM), 1991.

[MNW14] Dillon Mayhew, Mike Newman, and Geoff Whittle. Is the missing axiom of matroid theory lost forever? *Q. J. Math.*, 65(4):1397–1415, 2014.

[MNW18] Dillon Mayhew, Mike Newman, and Geoff Whittle. Yes, the "missing axiom" of matroid theory is lost forever. *Trans. Am. Math. Soc.*, 370(8):5907–5929, 2018.

[MP21] Dustin G. Mixon and Hans Parshall. Globally optimizing small codes in real projective spaces. *SIAM J. Discrete Math.*, 35(1):234–249, 2021.

[MR08] Dillon Mayhew and Gordon F. Royle. Matroids with nine elements. *J. Comb. Theory, Ser. B*, 98(2):415–431, 2008.

[MS95] František Matúš and Milan Studený. Conditional independences among four random variables. I. *Combin. Probab. Comput.*, 4(3):269–278, 1995.

[MS15] Diane Maclagan and Bernd Sturmfels. *Introduction to tropical geometry*, volume 161 of *Graduate Studies in Mathematics*. American Mathematical Society (AMS), 2015.

[MS21a] Mateusz Michałek and Bernd Sturmfels. *Invitation to nonlinear algebra*, volume 211 of *Graduate Studies in Mathematics*. American Mathematical Society (AMS), 2021.

[MS21b] Tillmann Miltzow and Reinier F. Schmiermann. On classifying continuous constraint satisfaction problems, 2021. arXiv:2106.02397.

[MT03] Annalisa Marcja and Carlo Toffalori. *A guide to classical and modern model theory*, volume 19 of *Trends in Logic—Studia Logica Library*. Kluwer, 2003.

[Nel18] Peter Nelson. Almost all matroids are non-representable. *Bull. Lond. Math. Soc.*, 50:245–248, 2018.

[Oxl11] James Oxley. *Matroid theory*, volume 21 of *Oxford Graduate Texts in Mathematics*. Oxford University Press, 2nd edition, 2011.

[Par03] Pablo A. Parrilo. Semidefinite programming relaxations for semialgebraic problems. *Math. Program.*, 96(2 (B)):293–320, 2003.

[Pea88] Judea Pearl. *Probabilistic reasoning in intelligent systems: networks of plausible inference*. The Morgan Kaufmann Series in Representation and Reasoning. Morgan Kaufmann, San Mateo, CA, 1988.

[PP85] Judea Pearl and Azaria Paz. GRAPHOIDS: A graph-based logic for reasoning about relevance relations, or When would x tell you more about y if you already know z. Technical Report CSD-850038, UCLA Computer Science Department, 1985.

[PS03] Pablo A. Parrilo and Bernd Sturmfels. Minimizing polynomial functions. In *Algorithmic and quantitative real algebraic geometry. Proceedings of the DIMACS workshop on algorithmic and quantitative aspects of real algebraic geometry in mathematics and computer science, DIMACS Center, March 12–16, 2001*, pages 83–99. American Mathematical Society (AMS), 2003.

[Rao73] Calyampudi Radhakrishna Rao. *Linear statistical inference and its applications. 2nd ed.* Wiley, 1973.

[Ric97] Jürgen Richter-Gebert. *Realization spaces of polytopes*, volume 1643 of *Lecture Notes in Mathematics*. Springer, 1997.

[Ric99a] Jürgen Richter-Gebert. Testing orientability for matroids is NP-complete. *Adv. Appl. Math.*, 23(1):78–90, 1999.

[Ric99b] Jürgen Richter-Gebert. The universality theorems for oriented matroids and polytopes. In *Advances in discrete and computational geometry. Proceedings of the 1996 AMS-IMS-SIAM joint summer research conference on discrete and computational geometry: ten years later, South Hadley, MA, USA, July 14–18, 1996*, pages 269–292. American Mathematical Society (AMS), 1999.

[Ric11] Jürgen Richter-Gebert. *Perspectives on projective geometry. A guided tour through real and complex geometry*. Springer, 2011.

[Rin56] Gerhard Ringel. Teilungen der Ebenen durch Geraden oder topologische Geraden. *Math. Z.*, 64:79–102, 1956.

[Rot71] Gian-Carlo Rota. Combinatorial theory, old and new. In *Actes du Congrès International des Mathématiciens (Nice, 1970), Tome 3*, pages 229–233. 1971.

[RZ95] Jürgen Richter-Gebert and Günter M. Ziegler. Realization spaces of 4-polytopes are universal. *Bull. Am. Math. Soc., New Ser.*, 32(4):403–412, 1995.

[SA98] Juriaan Simonis and Alexei Ashikhmin. Almost affine codes. *Des. Codes Cryptography*, 14(2):179–197, 1998.

[Sad17] Kayvan Sadeghi. Faithfulness of probability distributions and graphs. *J. Mach. Learn. Res.*, 18:1–29, 2017.

[Sch98] Alexander Schrijver. *Theory of linear and integer programming.* Wiley, 1998.

[Sch00] Stefan Schirra. Robustness and precision issues in geometric computation. In *Handbook of computational geometry*, pages 597–632. Elsevier, 2000.

[SE05] Niklas Sörensson and Niklas Een. MiniSat v1.13 - A SAT Solver with Conflict-Clause Minimization, 2005.

[Sey92] Paul D. Seymour. On secret-sharing matroids. *J. Comb. Theory, Ser. B*, 56(1):69–73, 1992.

[Sha48] Claude E. Shannon. A mathematical theory of communication. *Bell Syst. Tech. J.*, 27:379–423, 623–656, 1948.

[Shl11] Alexandra Shlapentokh. Defining integers. *Bull. Symb. Log.*, 17(2):230–251, 2011.

[Sho91] Peter W. Shor. Stretchability of pseudolines is NP-hard. In Peter Gritzmann and Bernd Sturmfels, editors, *Applied geometry and discrete mathematics. The Victor Klee Festschrift on the occasion of Professor Klee's 65th birthday*, volume 4 of *DIMACS. Series in Discrete Mathematics and Theoretical Computer Science*, pages 531–554. American Mathematical Society (AMS), 1991.

[Šim06a] Petr Šimecek. Classes of Gaussian, discrete and binary representable independence models have no finite characterization. In *Proceedings of Prague Stochastics*, volume 400, pages 622–631, 2006.

[Šim06b] Petr Šimeček. Gaussian representation of independence models over four random variables. In *COMPSTAT conference*, 2006.

[Šim06c] Petr Šimeček. A short note on discrete representability of independence models. In *Proccedings of the European Workshop on Probabilistic Graphical Models*, pages 287–292, 2006.

[Šim07] Petr Šimeček. *Nezávislostní modely.* Dissertation (in czech), Charles University, Prague, Czech Republic, 2007.

[Sip06] Michael Sipser. *Introduction to the theory of computation.* Thompson, 2006.

[Spi65] Michael Spivak. *Calculus on manifolds.* Mathematics Monograph Series. Westview Press, 1965.

[SRSM19] Shubham Sharma, Subhajit Roy, Mate Soos, and Kuldeep S. Meel. GANAK: A Scalable Probabilistic Exact Model Counter. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI-19*, pages 1169–1176. International Joint Conferences on Artificial Intelligence Organization, 7 2019.

[SŠ17] Marcus Schaefer and Daniel Štefankovič. Fixed points, Nash equilibria, and the existential theory of the reals. *Theory Comput. Syst.*, 60(2):172–193, 2017.

[Stu92] Milan Studený. Conditional independence relations have no finite complete characterization. In *Information Theory, Statistical Decision Functions and Random Processes*, volume B, pages 377–396. Kluwer, 1992.

[Stu93] Milan Studený. Convex cones in finite-dimensional real vector spaces. *Kybernetika*, 29(2):180–200, 1993.

[Stu94] Milan Studený. Semigraphoids are Two-Antecedental Approximations of Stochastic Conditional Independence Models. In Ramon Lopez de Mantaras and David Poole, editors, *Uncertainty Proceedings*, pages 546–552. Morgan Kaufmann, 1994.

[Stu02] Bernd Sturmfels. *Solving systems of polynomial equations*, volume 97 of *CBMS Regional Conference Series in Mathematics*. American Mathematical Society (AMS), 2002.

[Stu05] Milan Studený. *Probabilistic Conditional Independence Structures*. Information Science and Statistics. Springer, 2005.

[Stu08] Bernd Sturmfels. *Algorithms in invariant theory*. Texts & Monographs in Symbolic Computation. Springer, 2008.

[Stu19] Milan Studený. Conditional independence and basic Markov properties. In Marloes Maathuis, Mathias Drton, Steffen Lauritzen, and Martin Wainwright, editors, *Handbook of graphical models*, Chapman & Hall/CRC Handbooks of Modern Statistical Methods, pages 3–38. CRC Press, 2019.

[Stu20] Milan Studený. Conditional independence structures over four discrete random variables revisited: conditional Ingleton inequalities, 2020. arXiv:2012.04092.

[Stu21] Milan Studený. Conditional independence structures over four discrete random variables revisited: conditional ingleton inequalities. *IEEE Trans. Inf. Theory*, 67(11):7030–7049, 2021.

[Sul09] Seth Sullivant. Gaussian conditional independence relations have no finite complete characterization. *J. Pure Appl. Algebra*, 213(8):1502–1506, 2009.

[Sul18] Seth Sullivant. *Algebraic Statistics*, volume 194 of *Graduate Studies in Mathematics*. American Mathematical Society (AMS), 2018.

[SV98] Milan Studený and Jiřina Vejnarová. The multi-information function as a tool for measuring stochastic dependence. In *Learning in graphical models*, pages 261–297. Kluwer, 1998.

[Tar48] Alfred Tarski. A decision method for elementary algebra and geometry. US Air Force Project Rand, R-109. Santa Monica, California: The Rand Corporation. III, 60 p. (1948)., 1948. Prepared for publication by J. C. C. McKinsey.

[Thu06] Marc Thurley. sharpSAT – Counting Models with Advanced Component Caching and Implicit BCP. In Armin Biere and Carla P. Gomes, editors, *Theory and Applications of Satisfiability Testing - SAT 2006*, pages 424–429. Springer Berlin, 2006.

[TS16] Takahisa Toda and Takehide Soh. Implementing efficient all solutions SAT solvers. *Journal of Experimental Algorithmics*, 21:1.12:1–1.12:44, 2016.

[Tsc01] Falk Tschirschnitz. Testing extendability for partial chirotopes is NP-complete. In *Proceedings of the 13th Canadian Conference on Computational Geometry, University of Waterloo, Ontario, Canada, August 13-15, 2001*, pages 165–168, 2001.

[Tse83] Grigorij S. Tseĭtin. On the complexity of proof in propositional calculus. In Jörg H. Siekmann and Graham Wrightson, editors, *Automation of Reasoning*, volume 2, pages 466–483. Springer, 1983.

[Vak06] Ravi Vakil. Murphy's law in algebraic geometry: Badly-behaved deformation spaces. *Invent. Math.*, 164(3):569–590, 2006.

[Vir10] Oleg Y. Viro. Hyperfields for Tropical Geometry I. Hyperfields and dequantization, 2010. arXiv:1006.3034.

[Vá78] Peter Vámos. The missing axiom of matroid theory is lost forever. *J. Lond. Math. Soc., II. Ser.*, 18:403–408, 1978.

[Wel76] Dominic J. A. Welsh. *Matroid theory*, volume 8 of *London Mathematical Society Monographs*. Academic Press, 1976.

[Whi35] Hassler Whitney. On the abstract properties of linear dependence. *Am. J. Math.*, 57:509–533, 1935.

[Whi87] Neil White, editor. *Combinatorial geometries*, volume 29 of *Encyclopedia of Mathematics and Its Applications*. Cambridge University Press, 1987.

[Whi08] Neil White, editor. *Theory of matroids*, volume 26 of *Encyclopedia of Mathematics and Its Applications*. Cambridge University Press, 2008.

[WM] Wolfram Research, Inc. Mathematica, Version 11.3. Champaign, IL, 2018.

[Yeu05] Raymond W. Yeung. *A first course in information theory*. Information Technology: Transmission, Processing and Storage. Springer, 2005.

[Zha05] Fuzhen Zhang, editor. *The Schur complement and its applications*. Springer, 2005.

[Zie95] Günter M. Ziegler. *Lectures on polytopes*, volume 152 of *Graduate Texts in Mathematics*. Springer, 1995.

[ZY97] Zhen Zhang and Raymond W. Yeung. A non-Shannon-type conditional inequality of information quantities. *IEEE Trans. Inf. Theory*, 43(6):1982–1986, 1997.

[ZY98] Zhen Zhang and Raymond W. Yeung. On characterization of entropy function via information inequalities. *IEEE Trans. Inf. Theory*, 44(4):1440–1452, 1998.