Matroid stratifications of hypergraph varieties, their realization spaces, and discrete conditional independence models

Oliver Clarke, Kevin Grace, Fatemeh Mohammadi, and Harshit J Motwani

Abstract

We study varieties associated to hypergraphs from the point of view of projective geometry and matroid theory. We describe their decompositions into matroid varieties, which may be reducible and can have arbitrary singularities by the Mnëv–Sturmfels universality theorem. We focus on various families of hypergraph varieties for which we explicitly compute an irredundant irreducible decomposition. Our main findings in this direction are threefold: (1) we describe minimal matroids of such hypergraphs; (2) we prove that the varieties of these matroids are irreducible and their union is the hypergraph variety; and (3) we show that every such matroid is realizable over real numbers. As corollaries, we give conceptual decompositions of various, previously-studied, varieties associated with graphs, hypergraphs, and adjacent minors of generic matrices. In particular, our decomposition strategy gives immediate matroid interpretations of the irreducible components of multiple families of varieties associated to conditional independence (CI) models in statistical theory and unravels their symmetric structures which hugely simplifies the computations.

Contents

1   Introduction

This work develops new connections between conditional independence (CI) models in statistics [Stu05, DSS09, Sul18], projective geometry [RG11, LV13], the theory of matroids [Oxl11, PW70] and their realization spaces [Mnë85, Mnë88, Stu89], and determinantal varieties in commutative algebra [BC03]. In particular, we study a family of varieties whose defining equations are indexed by the edges of some hypergraphs, and show that their irreducible decompositions lead to well-structured configurations of points and lines. Moreover, we apply our results to classes of *conditional independence* varieties which naturally arise in algebraic statistics. We also analyze well-known results from the literature [HHH$^+$10, Rau13, HS04, PS19, CMR20] and show, in these cases, that the irreducible components of their concerned varieties have concrete interpretations in terms of matroids.

Our main object of study is the *determinantal hypergraph variety* V_Δ which is associated to a hypergraph Δ; see Definition 2.1. The defining equations of V_Δ are the minors of a matrix X of indeterminates and the corresponding polynomial ideal I_Δ generalizes many familiar families of determinantal ideals. For example, given integers $t \leq d \leq n$ and the $d \times n$ matrix $X = (x_{ij})$ of indeterminates, the classical ideal $I_t(X)$ studied by Bruns and Conca [BC03] is the ideal of the hypergraph Δ whose edges are all t-subsets of $[n] := \{1, \ldots, n\}$. These ideals, along with many other families [Stu90, HHH$^+$10, EHHM13, MR18, PS19], are often studied using Gröbner bases. When it is possible to find them, Gröbner bases are a powerful tool for understanding ideals. For instance, if the initial ideal is square-free, then the original ideal is radical. However, when the hypergraph Δ lies outside of one of the particularly nice families in the list above, finding a Gröbner basis becomes exceedingly difficult. In this paper, we develop a general framework called the *decomposition strategy* for understanding the properties of arbitrary hypergraph ideals, without computing their Gröbner bases, using matroid varieties.

A *matroid* is a combinatorial axiomatization of linear dependence in a vector space. We refer to [Oxl11] for a complete introduction to matroids, however, it will now be useful for us to recall realizability and related notions. Given a finite collection of vectors in a fixed vector space, the collection of linearly dependent sets

of vectors defines a matroid. If this process can be reversed, i.e. for a given matroid M we can find such a collection of vectors, then we call these vectors a realization of M. We write Γ_M for the space of realizations of M. The *matroid variety* V_M of M is the Zariski closure of the realization space; see Definition 2.3. Our main tool to study hypergraph varieties is the Decomposition Theorem which provides a matroid stratification for V_Δ. One can think of this theorem as an altered version of the stratification of the Grassmannian by matroids, which was studied in [GGMS87].

Theorem (Theorem 3.1). *The variety associated to the hypergraph Δ is the union of matroid varieties. The union is taken over all realizable matroids M whose dependent sets contain the edges of Δ.*

Many of the matroid varieties appearing in the decomposition above are redundant. So, to refine this decomposition, we introduce the notion of *combinatorial closures* of matroids; see Definition 3.3. More precisely, the combinatorial closure V_M^{comb} of a given matroid M is the hypergraph variety associated to its set of circuits.

We recall that a *minimally dependent matroid* for a given hypergraph Δ is a realizable matroid M whose dependent sets contain Δ as a subset and is minimal among all such matroids. In other words, there does not exist a realizable matroid N such that $\Delta \subseteq \mathcal{D}(N) \subsetneq \mathcal{D}(M)$.

Proposition (Proposition 3.10). *The associated variety of a hypergraph Δ is the union of combinatorial closures of matroid varieties, where the union is taken over all minimally dependent matroids for Δ.*

We conjecture that combinatorial closures are sufficient to find the irredundant irreducible decomposition of any hypergraph variety. In fact, we show that this is the case for some of the well-known examples of hypergraph varieties such as those corresponding to binomial edge ideals [HHH+10, Rau13], conditional independence ideals with hidden variables [CMR20, CMM21] and ideals of adjacent minors [HS04]. In all of these cases, each combinatorial closure is equal to its *central component* which makes the computations easier. One of the families of hypergraphs we study is the *consecutive forest* hypergraphs; see Definition 6.1. We will see that the combinatorial closures have non-central components. To compute the irreducible decomposition of V_Δ, we apply the following method and note the corresponding sections for consecutive forest hypergraphs.

Decomposition Strategy §3.3.

(i) Identify minimally dependent matroids for Δ (§6.1).

(ii) For each minimally dependent matroid M, write the combinatorial closure V_M^{comb} as a minimal union of matroid varieties (§5).

(iii) For each matroid variety appearing in step (ii), show that it is irreducible (§4).

(iv) Determine redundancy of non-central components of combinatorial closures in the resulting decomposition and show the matroids are realizable (§6.2 and §6.3).

We are able to completely go through this strategy for various family of hypergraphs in §4-§7. In particular, for Step (iii) above, we show that the realization spaces of certain point and line configurations are irreducible; see §4. These are a family of matroids whose varieties appear as irreducible components of many examples of hypergraph varieties. We will prove Theorem 4.5 which allows us to inductively build up configurations with irreducible varieties, and use this to show that all configurations with at most 6 lines have irreducible varieties.

Theorem (Theorems 4.2 and 4.5). *Suppose that C' is a point and line configuration with irreducible realization space $\Gamma_{C'}$. If C is obtained from C' by adding a single line passing through at most two intersection points of C', then Γ_C is irreducible. In particular, the realization spaces of configurations with at most 6 lines are irreducible.*

The conclusion of the decomposition strategy is the following characterization of the minimal irreducible decomposition of the associated variety of each consecutive forest hypergraph.

Theorem (Theorems 5.14 and 6.20). *The irredundant irreducible components of the consecutive forest variety V_{Δ_G} are in one-to-one correspondence with the minimally dependent matroids for the hypergraph Δ_G. In particular, for each minimally dependent matroid, the non-central components of its combinatorial closure are completely characterized and are redundant in the irreducible decomposition.*

Our decomposition strategy gives immediate interpretations of the prime decompositions of families of ideals for which it is not feasible to calculate an explicit Gröbner basis due to the time complexity of the algorithm and the current available hardware. In [PS19], which is further explained in Example 3.12, the authors calculated the prime decomposition of a hypergraph ideal which contains 16 edges of size 3. This computation is at the limit of what is currently possible on available hardware. In this case, the prime components of this ideal have a straight-forward interpretation in terms of configurations of 12 points in the projective plane.

Moreover, in §7 we describe the hypergraphs $\Delta^{s,t}$ which arise from the study of conditional independence (CI) statements; see §8 and [Fin11, CMR20, CMM21]. We give a summary of the known cases of $V_{\Delta^{s,t}}$ in Remark 7.2 and Table 2. The family of varieties $V_{\Delta^{s,t}}$ for a fixed s, t can be studied effectively by understanding finite families of matroids. In Table 1, we count the number of combinatorial types of dependent matroids for $\Delta^{s,t}$. We show that as certain parameters increase, the number of combinatorial types eventually stabilizes which simplifies the computational task. The associated variety of $\Delta^{s,t}$ determines a statistical model corresponding to a collection of CI statements with hidden variables. The irreducible components of these varieties give information about additional constraints satisfied by distributions within the given CI model. We describe the irreducible components of these varieties in terms of so-called *grid matroids*. The first step to understanding such decompositions is to find the set of minimally dependent matroids for $\Delta^{s,t}$.

When the dimension of the ambient space d is low enough, we show that:

Theorem (Theorem 7.5). *If $t \leq d \leq s + t - 3$, then there is a unique minimal matroid for $\Delta^{s,t}$.*

In particular, the minimal matroid from the above theorem corresponds to the special irreducible component of the CI varieties studied in [CMR20, CMM21], which leads to the *intersection axiom* for CI models. This component is particularly important for the inference problem in statistics, as it gives information about additional constraints on distributions with *full support* which satisfy the given CI statements.

When there is no bound on d, we show that as the parameters for this hypergraph are allowed to become large, we obtain every matroid among the minimally dependent matroids for $\Delta^{s,t}$, up to a mild equivalence.

Theorem (Theorem 7.7). *For every matroid M, there exists a hypergraph of the form $\Delta^{s,t}$ and a dependent matroid M' for $\Delta^{s,t}$ such that a restriction of M' is isomorphic to M.*

We prove this theorem using an algorithm from [MF14] to explicitly construct the matroid M'. This theorem shows that understanding the varieties of a general CI model is very difficult since matroid varieties satisfy various *universality* results. Similarly, determining the realizability of a CI model is difficult as the matroids of its components might not be $\mathbb{R}$-realizable. More generally, we propose the following computational problem.

Question 1.1. *Find the irreducible components of $V_{\Delta^{s,t}}$ and determine whether they contain a rational point.*

We conclude by highlighting some of the difficulties arising in the study of Question 1.1, in particular in determining the irreducibility of matroid varieties. Following [BS89, RG99], consider a collection of vectors $v_1, \ldots, v_s$ in a finite dimensional vector space V over a field $\mathbb{K}$, which realizes a matroid M. They also define a hyperplane arrangement $H_{v_1}, \ldots, H_{v_s} \subset V^*$ in the dual vector space. The combinatorial type of this hyperplane arrangement is defined by M, thus, one can think of the realization space Γ_M as a parameter space of hyperplane arrangements of fixed combinatorial type. In [Ryb11], Rybnikov constructs a pair of combinatorially equivalent hyperplane arrangements whose complements have non-isomorphic fundamental groups. So, in general, it is not possible to study matroid varieties by picking a single generic point in the variety.

Unfortunately, very little can be said about the geometry of matroid varieties in general. The Mnëv–Sturmfels Universality Theorem shows that matroid varieties satisfy *Murphy's Law in Algebraic Geometry*. Specifically, given any singularity, appearing in a semi-algebraic set, there is a matroid variety with the same singularity up to a mild equivalence; see [Mnë85, Mnë88, Stu89, LV13]. In fact, point and line configurations already exhibit all singularities. More precisely, the equivalence above is defined on pointed schemes generated by $(X, p) \sim (Y, q)$ when there exists a smooth morphism $(X, p) \to (Y, q)$. Then, the universality theorem states that point and line configurations exhibit all equivalence classes of this relation called singularity types.

We also note that the defining equations of matroid varieties are in general very difficult to calculate. These equations often arise from geometric constraints satisfied by the matroid. Some of them can be interpreted as rank conditions on certain submatrices, however this is not true in general. Example 3.6 shows the smallest matroid for which *non-determinantal* equations appear. Proving that such a polynomial constraint holds is often achieved by finding an equivalent condition in the Grassmann-Cayley algebra [SJS17, STW21]. In the context of CI models, such conditions give further constraints on distributions satisfying the given CI statements.

Outline of paper. In §2, we introduce the key concepts used throughout the paper. In particular, we fix our notations for hypergraph, matroid varieties and introduce the dependence order on matroids. In §3, we prove the decomposition theorem and introduce the combinatorial closures of matroid varieties. In §4, we define point and line configurations and prove that, for certain families, their realization spaces are irreducible. In §5, we study the family of forest-like point and line configurations. We give a complete characterization of the irreducible components of their combinatorial closures using perturbation arguments. In §6, we apply the decomposition strategy to the family of consecutive forest hypergraphs. As a result, we prove Theorem 6.20 which gives an irredundant irreducible decomposition of the hypergraph variety. In §7, we introduce grid matroids and apply them to describe the irreducible components of the variety of $\Delta^{s,t}$; see Tables 1 and 2. We then proceed to use algorithmic procedures to prove Theorem 7.7. In §8, we show how the hypergraph $\Delta^{s,t}$ arises in algebraic statistics in the study of CI statements with hidden variables. We conclude by explaining how our results may shed light on a conjecture by Matúš in the context of CI models.

2 Preliminaries

2.1 Hypergraph varieties. Let $\mathbb{K}$ be a field, $d \leq n$ be two positive integers, $X = (x_{ij})$ be a $d \times n$ matrix of indeterminates and $R = \mathbb{K}[X]$ be the polynomial ring over $\mathbb{K}$ in the indeterminates x_{ij}. It is often convenient to write determinants of submatrices of X as $[I|J]_X$ where I and J are respectively the sets of rows and columns of the submatrix. If $I = [d]$, that is the submatrix covers all rows of X, then we write $[J]$ for $[I|J]_X$. We denote by x_i the ith column of X and by X_F the submatrix of X with columns indexed by $F \subseteq [n]$.

Definition 2.1. A (simple) hypergraph Δ on the vertex set $[n]$ is a subset of the power set $2^{[n]}$. We assume that no proper subset of an element of Δ is in Δ. The elements of Δ are called (hyper)edges.

- The *determinantal hypergraph ideal* of Δ is

$$I_\Delta = \left\langle [A|B]_X : A \subseteq [d], B \in \Delta, |A| = |B| \right\rangle \subset R.$$

- The variety of Δ is the zero set of I_Δ which is given by

$$V_\Delta = \{X \in \mathbb{C}^{d \times n} : \ \text{Rank}(X_F) < |F| \text{ for each } F \text{ in } \Delta\}.$$

The ideal I_Δ and its variety V_Δ depend on the value d, i.e. the dimension of the ambient space. However, to keep our notation concise, we suppress d from our notation. Unless otherwise stated, all results for hypergraph ideals and their varieties hold for all d as long as $d \geq \max\{|F| : F \in \Delta\}$.

Problem 2.2. *Find an irredundant irreducible decomposition of V_Δ for any hypergraph Δ.*

2.2 Matroid varieties. In this subsection, we recall the definitions of the realization space of a matroid and its associated variety. We refer the reader to [Oxl11] for basic definitions concerning matroids.

Definition 2.3. Let M be a matroid on $[n]$ of rank r and let $d \geq r$. If $\mathbb{K}$ is a field, a *realization* of M in $\mathbb{K}^d$ is a collection of vectors $X = \{x_1, \ldots, x_n\} \subset \mathbb{K}^d$ such that

$$\{x_{i_1}, \ldots, x_{i_p}\} \subset X \text{ is linearly dependent} \iff \{i_1, \ldots, i_p\} \text{ is a dependent set of } M.$$

If such a collection of vectors exists, we say that the matroid is *realizable* over $\mathbb{K}$. (Other words used interchangeably with *realizable* include *representable* and *linear*.) In this paper, if realizability is discussed without specifying $\mathbb{K}$, then $\mathbb{K} = \mathbb{C}$. The *realization space* of M in $\mathbb{C}^d$ is

$$\Gamma_M = \{X \subset \mathbb{C}^d : X \text{ is a realization of } M\}.$$

So each element of Γ_M is naturally identified with a $d \times n$ matrix over $\mathbb{C}$.

Definition 2.4. The *matroid variety* $V_M = \overline{\Gamma_M}$ is the Zariski closure of the realization space of M. We denote $I_M = I(V_M) \subseteq \mathbb{C}[X]$ for the corresponding ideal where $X = (x_{i,j})$ is a $d \times n$ matrix of indeterminates.

Note that I_M is a radical ideal. Similarly to hypergraph ideals and their varieties, we will always use d to denote the dimension of the ambient space. In order to simplify our notation, we suppress d from the notation of the realization space of matroids and their ideals and varieties.

2.3 Minimal matroids. We denote by $\leq$ the partial order on sets given by inclusion. We extend this notion to matroids by identifying them with their collection of dependent sets. We will denote the collection of dependent sets of a matroid M by $\mathcal{D}(M)$. So, given two matroids M_1 and M_2, if $\mathcal{D}(M_1) \subseteq \mathcal{D}(M_2)$ then we write $M_1 \leq M_2$. This is the *dependency order* on matroids. However, we caution the reader that this is precisely opposite of the *weak order* in the matroid literature [Oxl11, BS89]. Suppose that M is a matroid on ground set E. For any given collection of subsets $\mathcal{D}$ of E, we say that M is *dependent* for $\mathcal{D}$ if $\mathcal{D} \subseteq \mathcal{D}(M)$ and write $\mathcal{D} \leq M$. If there does not exist a matroid N on E such that $\mathcal{D} \subseteq \mathcal{D}(N) \subsetneq \mathcal{D}(M)$, then we say that M is *minimally dependent* for $\mathcal{D}$, i.e. M is a smallest matroid such that $\mathcal{D} \leq M$.

Example 2.5. Any matroid M is minimally dependent for its collection of circuits $\mathcal{C}$. In this case, M is the unique such matroid. In general, there can be many different minimally dependent matroids for a collection $\mathcal{D}$. For example, if $E = [5]$, and $\mathcal{D} = \{1234, 1235\}$, then the uniform matroid U_5^3, whose circuits are all 4-subsets of E, and the matroid with a single circuit 123 are both minimally dependent for $\mathcal{D}$; see Figure 1.

We will be interested in finding out the minimal matroids for a given collection of sets. In general, this is difficult. However, we are able to do this for some specific families; see §6.
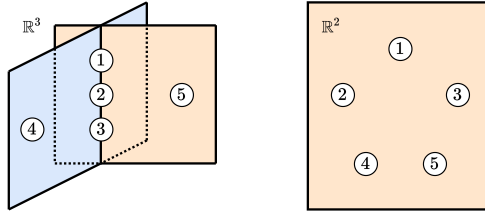
Figure 1: Geometric representations for two *minimally dependent matroids* for $\mathcal{D} = \{1234, 1235\}$.

Notation. We denote the corresponding simple hypergraph of any collection $\mathcal{D}$ of subsets of $[n]$ as:

$$\min(\mathcal{D}) = \{D \in \mathcal{D} : \text{ no set } D' \in \mathcal{D} \text{ exists such that } D' \subsetneq D\}. \tag{1}$$

Remark 2.6. Let Δ' be a simple hypergraph all of whose edges have size at most d, and let $\Delta = \min(\Delta' \cup \binom{[n]}{d+1})$. Note that $I_{\Delta'} = I_\Delta$ and that $V_{\Delta'} = V_\Delta$. Moreover, $\Gamma_M = \emptyset$ if M has rank greater than d. Therefore, finding the minimally dependent matroids of rank at most d for Δ is equivalent to finding those for Δ'.

3 Matroid stratifications of hypergraph varieties

3.1 The general decomposition theorem. We now use our notation from §2 to decompose any hypergraph variety into certain matroid varieties. In subsequent sections, we will investigate techniques which will allow us to strengthen this statement to an irredundant irreducible decomposition. The following theorem may be interpreted from the perspective of the stratification of Grassmannians by matroid varieties. See Remark 3.2.

Theorem 3.1 (Decomposition Theorem). *Let $\Delta \subseteq 2^{[n]}$ be a hypergraph on $[n]$ and let $\mathcal{M}(\Delta)$ be the collection of realizable matroids whose collections of dependent sets contain Δ. Then,*

$$V_\Delta = \bigcup_{M \in \mathcal{M}(\Delta)} V_M.$$

In particular, we have $\sqrt{I_\Delta} = \bigcap_{M \in \mathcal{M}(\Delta)} I_M$.

Proof. Let $M \in \mathcal{M}(\Delta)$. We begin by showing that $I_\Delta \subseteq I_M$. By definition, I_Δ is generated by minors, so let $f \in I_\Delta$ be any such minor. We may assume without loss of generality that f is a maximal minor of the submatrix X_F for some $F \in \Delta$. For each matroid $M \in \mathcal{M}(\Delta)$ and each point A in the realization space Γ_M, the columns of A_F are linearly dependent since F is a dependent set of M. Thus f vanishes on Γ_M, and by definition, f vanishes on its Zariski closure which is V_M. Thus, we have that $f \in I(V_M) = I_M$, so $I_\Delta \subseteq I_M$. Hence, $V_M \subseteq V_\Delta$. Let A be any point in the variety V_Δ. The matrix A is a realization of a matroid, which we denote M_A, on $[n]$ in $\mathbb{C}^d$. For each $F \in \Delta$, the ideal I_Δ contains all $|F|$-minors of X_F, so all $|F|$-minors of A_F vanish. Therefore, the columns of A_F are linearly dependent. Hence F is a dependent set in M_A, so $M_A \in \mathcal{M}(\Delta)$. □

3.2 Combinatorial closure. In Theorem 3.1, we decomposed the hypergraph variety V_Δ into matroid varieties. But this decomposition may have redundant components, because $V_{M'} \subset V_M$ for some $M, M' \in \mathcal{M}(\Delta)$. To solve Problem 2.2, we must determine the matroids that are necessary in the decomposition. It is clear that all matroids which are minimal with respect to the dependency order are necessary in the decomposition. One could conjecture that these are enough. Unfortunately this is *not* true, as we will see in Example 3.6. Nevertheless, we will define an alternative decomposition for which the minimal matroids correspond to irredundant parts of a decomposition. We do this by introducing the *combinatorial closure* of a matroid. This is a generalization of the *weak realization space* of a matroid [BS89, page 70] to an ambient space of dimension greater than the rank of the matroid.

In the following remark we connect our setting to that of the stratification of the Grassmannian by matroids.

Remark 3.2. The Grassmannian $\mathrm{Gr}(d, n+d)$ admits a stratification $\mathrm{Gr}(d, n+d) = \bigcup_M S_M$ by matroids [GGMS87, Stu89]. Here, $S_M := \{x \in \mathrm{Gr}(d, n+d) \mid \mathcal{M}_x = \mathcal{M}\}$ is the matroid stratum of $\mathcal{M}$, where $\mathcal{M}_x$ is the matroid associated to the point $x \in \mathrm{Gr}(d, n+d)$. Let us restrict to the standard affine open patch U_0 of $\mathrm{Gr}(d, n+d)$ where the matrices are of the form $[M \ I]$ such that M is a $d \times n$ matrix and I is the $d \times d$ identity matrix. We identify the ambient space $\mathbb{C}^{d \times n}$ with U_0 as they are isomorphic. After this identification, we can see $V(I_\Delta)$ as the closed embedding of U_0 given by the vanishing of the corresponding Plücker coordinates.

For example, take $\Delta = \{12, 134\}$ and $d = 3$. Now $V(I_\Delta)$, is a closed subset of the standard affine open patch U_0 of Grassmannian $\mathrm{Gr}(3, 7)$. The equations of $V(I_\Delta)$ in terms of Plücker coordinates are given by

$P_{123} = 0, P_{124} = 0, P_{125} = 0, P_{126} = 0, P_{127} = 0, P_{134} = 0$. So, $V(I_\Delta)$ can be written in the following form:

$$V(I_\Delta) \cap \mathrm{Gr}(d, n+d) \cap U_0 = \bigcup_{\mathcal{M}} S_{\mathcal{M}} \cap V(I_\Delta) \cap U_0.$$

In this setting, we may restate Theorem 3.1 as $\bigcup_{\mathcal{M}} S_{\mathcal{M}} \cap V(I_\Delta) \cap U_0 = \bigcup_{M \in \mathcal{M}(\Delta)} V_M$.

Definition 3.3. We define the *combinatorial closure* V_M^{comb} of a matroid M to be the union of all matroid varieties $V_{M'}$ for which $M' \geq M$. In other words,

$$V_M^{\mathrm{comb}} = \bigcup_{M' \geq M} V_{M'}$$

We will denote the ideal of the combinatorial closure by I_M^{comb}.

Remark 3.4.

(i) Note that V_M^{comb} might not be a matroid variety itself. The inclusion $V_M \subseteq V_M^{\mathrm{comb}}$ holds in general, however the equality might not hold. (See Example 3.6 below.) We will call V_M the *central component* of V_M^{comb} and matroid varieties which intersect the complement will be called *non-central components*.

(ii) The combinatorial closure is the closure of the union of realization spaces, i.e. $V_M^{\mathrm{comb}} = \overline{\bigcup_{M' \geq M} \Gamma_{M'}}$. This follows from the topological fact that a finite union of closures coincides with the closure of a union of sets.

The ideal of the combinatorial closure of a matroid M can be seen as follows. The collection of circuits $C(M)$ of M can be considered as a hypergraph. So the hypergraph ideal $I_{C(M)}$ for a given d, is defined as:

$$I_{C(M)} = \left\langle [A|B]_X : A \subseteq [d], B \in C(M), |A| = |B| \right\rangle.$$

Lemma 3.5. *For any matroid M we have that*

$$V_M^{\mathrm{comb}} = V_{C(M)} \quad \text{or equivalently} \quad I_M^{\mathrm{comb}} = \sqrt{I_{C(M)}} .$$

Proof. Let $X \in V_M^{\mathrm{comb}}$ be a collection of vectors, then $X \in \Gamma_{M'}$ for some $M' \geq M$. So, any $B \in C(M)$ is a dependent set of M', and therefore $[A|B]_X = 0$ for any $A \subseteq [d]$ with $|A| = |B|$. This shows that $V_M^{\mathrm{comb}} \subseteq V_{C(M)}$.

Similarly, let X be a point in $V_{C(M)}$, and let M_X be the matroid represented by X. For any $B \in C(M)$ and any subset $A \subseteq [d]$ with $|A| = |B|$, we have $[A|B]_X = 0$. Therefore, any $B \in C(M)$ is also a dependent set of M_X, hence $M_X \geq M$. This shows that $V_{C(M)} \subseteq V_M^{\mathrm{comb}}$. $\square$

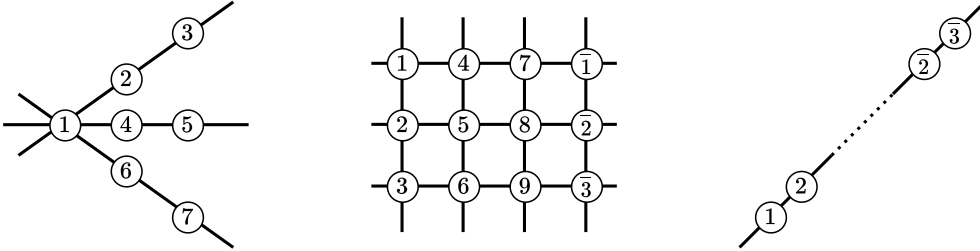In general, the combinatorial closures of matroid varieties are *reducible*.

Figure 2: (Left) A realization of M from Example 3.6. (Center and Right) Realizations of matroids corresponding to the prime components of I_Δ in Example 3.12.

Example 3.6. Let $d = 3$, $n = 7$ and $\Delta = \min(\{123, 145, 167\} \cup \binom{[7]}{4})$. It is easy to check that Δ is the collection of circuits for a matroid M, hence M is the unique minimal matroid for Δ. By Lemma 3.5, the ideal of the combinatorial closure is given by the radical ideal $I_{C(M)} = I_\Delta = \langle [123], [145], [167] \rangle \subseteq \mathbb{C}[x_{1,1}, \ldots, x_{3,7}]$. Using Macaulay2 [GS], we find that $I_{C(M)}$ has two prime components given by

$$I_{C(M)} = I_1 \cap I_2 = \langle x_{1,1}, x_{2,1}, x_{3,1} \rangle \cap \langle [123], [145], [167], [234][567] - [235][467] \rangle .$$

By inspection, the ideal I_1 is the ideal of the matroid M' that has a single circuit 1. In this case, we say that M' is obtained from M by setting 1 to be a loop. On the other hand, I_2 is the ideal of the matroid M and is called the central component of the combinatorial closure. The generator $[234][567] - [235][467]$ of I_2 is a geometric condition satisfied by realizations of M. More precisely, it is a condition satisfied by six generic points which lie on three lines that intersect at a common point, as shown in Figure 2.

In Lemma 5.13, we generalize this example to an arbitrary number of lines meeting at a point.

Remark 3.7. We note that Example 3.6 has been studied in [BS89, Theorem 4.28] to resolve a problem of White, by showing that there exists a matroid M and a dependent matroid N (Section 2.3) such that $\overline{\Gamma_M} \cap \Gamma_N = \emptyset$. In Section 6, we provide a family of matroids M and dependent matroids N such that $\overline{\Gamma_M} \cap \Gamma_N \neq \emptyset$. In particular, the perturbation procedure (Lemma 5.9) allows us to approximate realizations of N with realizations of M. This gives a way to generate families of dependent matroids that give a positive answer to White's problem.

We recall the definition of ideal quotients and saturation. See, e.g. [CLO15, §4] for more details.

Lemma 3.8. *Consider a polynomial ring $R = \mathbb{K}[x_1, \ldots, x_n]$. Let $I \subset R$ be a radical ideal and $J \subset R$ any ideal. Then $I : J = I : J^\infty$, where*

$$I : J = \{f \in R : fJ \subseteq I\} \quad and \quad I : J^\infty = \bigcup_{i \geq 1} (I : J^i).$$

For a matroid M, let us define its *bases ideal* J_M to be

$$J_M = \sqrt{\prod_{B \text{ basis of } M} \langle [A|B]_X : A \subseteq [d], |A| = |B| \rangle}.$$

If M has rank d, then J_M is generated by the product of all maximal minors of X whose columns are bases of M.

Proposition 3.9. *The ideal of the matroid variety can be obtained by saturating $I(V_M^{\text{comb}})$ with respect to J_M:*

$$I_M = I(V_M^{\text{comb}}) : J_M^\infty.$$

Proof. Geometrically, the saturation $I(V_M^{\text{comb}}) : J_M^\infty$ corresponds to the ideal of $\overline{V_M^{\text{comb}} \setminus V(J_M)}$; see [CLO15, §4, Theorem 10(iii)]. By Lemma 3.8, if I is a radical ideal, then $I : J = I : J^\infty$. So it is enough to show that the difference $V_M^{\text{comb}} \setminus V(J_M)$ is the realization space Γ_M. For this, notice that the variety $V(J_M)$ consists of collections of vectors for which at least one of the bases B of M is dependent. Therefore, the difference $V_M^{\text{comb}} \setminus V(J_M)$ consists of collections of vectors V_M^{comb} for which all bases B of M are independent, which is by definition the realization space Γ_M. □

The above observation is shown in [STW21, Proposition 2.1.3] under the identification of $\mathbb{C}^{d \times n}$ with an open affine patch of the Grassmannian, as explained in Remark 3.2.

Proposition 3.10. *Let $\Delta \subseteq 2^{[n]}$ be a hypergraph on $[n]$ and let $\mathcal{M}'(\Delta)$ be the collection of minimal realizable matroids whose dependent sets contain Δ. Then*

$$\sqrt{I_\Delta} = \bigcap_{M \in \mathcal{M}'(\Delta)} I_M^{\text{comb}} \quad or \; equivalently \quad V_\Delta = \bigcup_{M \in \mathcal{M}'(\Delta)} V_M^{\text{comb}}.$$

Proof. Following the notation of Theorem 3.1, let $\mathcal{M}(\Delta)$ be the collection of realizable matroids whose collections of dependent sets contain Δ. We have that $\mathcal{M}'(\Delta)$ consists of the minimal matroids in $\mathcal{M}(\Delta)$. We also have that if $M \leq M'$ then $I_M^{\text{comb}} \subseteq I_{M'}^{\text{comb}}$. So it follows that,

$$\bigcap_{M \in \mathcal{M}'(\Delta)} I_M^{\text{comb}} = \bigcap_{M \in \mathcal{M}(\Delta)} I_M^{\text{comb}}.$$

And so the proof is complete by Theorem 3.1. □

3.3 Decomposition strategy. We can now make precise the decomposition strategy detailed in §1. By Proposition 3.10, we have a decomposition of the hypergraph variety V_Δ into combinatorial closures of minimally dependent matroids for Δ. To upgrade this to an irreducible decomposition, there are several remaining steps.

First, one has to find the minimal collection of matroid varieties which cover V_M^{comb}. In particular, one has to show that the matroids appearing in this collection are realizable, since otherwise they are clearly redundant. In general, these matroid varieties might not be irreducible, so one has to either prove their irreducibility (which we will do in several examples) or find their irreducible decomposition. By going through these steps, we will obtain such decompositions for various families of hypergraphs. It still might be the case that some of the components are redundant. So to obtain the actual minimal decomposition, one has to check which ones are necessary. We remark that step (ii) of the decomposition strategy, i.e. writing V_M^{comb} as a minimal union of matroid varieties, involves showing that the matroid varieties whose union is V_M^{comb} are nonempty. In other words, we must show that the matroids are realizable. In particular, if $V_M^{\text{comb}} = V_M$ for each minimally dependent matroid M of Δ, then step (ii) can be accomplished by showing that all of the minimally dependent matroids for Δ are realizable.
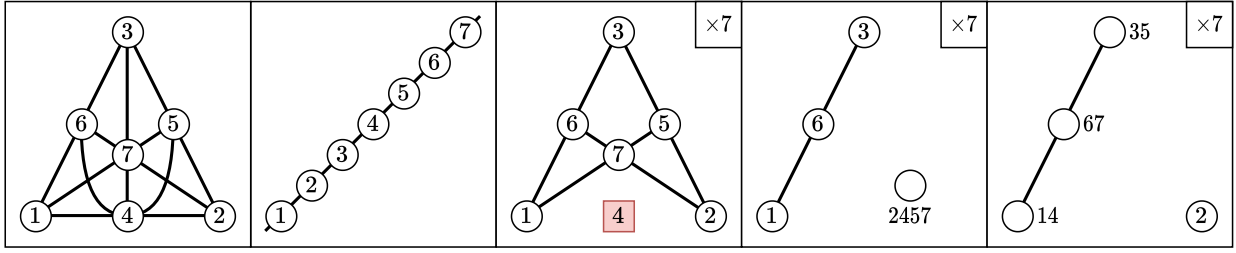
Figure 3: The Fano plane (left) and the configurations of the associated primes of its combinatorial closure.

Example 3.11 (Fano plane). Let $d = 3$ and consider the hypergraph $\Delta = \{124, 136, 157, 235, 267, 347, 456\}$. It is straightforward to check that Δ is the collection of circuits for a matroid M, which is called the Fano plane. The matroid M is not realizable over $\mathbb{C}$ so its realization space Γ_M is empty. However its combinatorial closure is non-empty and we can compute the associated prime ideals of I_Δ in Macaulay2 [GS]. We find that I_Δ has 22 associated primes which are all matroid varieties of 4 combinatorial types of point and line configurations; see Figure 3. The configurations in the figure, from left to right, are:

- A single line with 7 points. The ideal is a hypergraph ideal I_{Δ_0} where $\Delta = \binom{[7]}{3}$.

- A *quadrilateral set*, see [RG11, §8], together with a loop. There are 7 associated primes of I_Δ with this combinatorial type which are parametrised by the point of the Fano plane to be taken as the loop.

- A line with 3 points together with a free point with 4-labels. There are 7 associated primes of I_Δ with this combinatorial type which are parametrised by the lines of the Fano plane.

- A line with 3 points together with a free point. Each point on the line has two labels. There are 7 associated primes of I_Δ with this combinatorial type which are parametrised by the point of the Fano plane to be taken as the free point.

Example 3.12. Let $d = 3$ and $E = [9] \cup \{\bar{1}, \bar{2}, \bar{3}\}$. Let Δ be the hypergraph

$$\Delta = \{123, 456, 789, \bar{1}\bar{2}\bar{3}, 147, \bar{1}14, \bar{1}47, \bar{1}17, 258, \bar{2}58, \bar{2}28, \bar{2}25, 369, \bar{3}69, \bar{3}39, \bar{3}36\}$$

depicted in Figure 2. It is shown in [PS19, Theorem 4.1] that I_Δ has two prime components such that:

- The first component is generated by all 3-minors of a generic 3×12 matrix. Hence, it corresponds to the matroid M_0 with dependent sets,
  $$C(M_0) = \{\text{all 3-subsets of } E\}.$$

- The second component is generated by 44 polynomials of which 16 are the original generators of I_Δ and the remaining 28 generators are all homogeneous of degree 12.

Note that there exists a unique minimal matroid M for Δ. The circuits of M are elements of Δ and all subsets of size 4 of E, which do not contain an element of Δ. (For the proof of a more general statement, see Theorem 7.5.) Therefore, by Proposition 3.10, we have $\sqrt{I_\Delta} = I_M^{\text{comb}}$, so the result of [PS19] describes the decomposition of the combinatorial closure of M. This provides a geometric meaning of the second component of the decomposition. In particular, it can be obtained by saturating $\sqrt{I_\Delta}$ as in Proposition 3.9; see Figure 2.

Therefore, we can give a geometric interpretation of the second component described above. More precisely, the 28 non-determinantal generators are analogous to the geometric constraints of *quadrilateral sets*; see [RG11, §8]. More precisely, given a generic collection of 12 points arranged in a 3×4 grid in the projective plane, if we project this configuration onto a line then there are constraints on the distances between the projected points. The 28 polynomials describe precisely these constraints.

4 Irreducible point and line configurations

In this section, we will introduce point and line configurations. These are a family of matroids whose ideals appear as prime components of many examples of hypergraph ideals. We will prove Theorem 4.5 which allows us to inductively build up configurations with irreducible varieties. In particular, using this theorem, we will show that all configurations with at most 6 lines have irreducible varieties; see Theorem 4.2.

Notation. We write $C = (\mathcal{P}, \mathcal{L}, \mathcal{I})$ for a configuration of points and lines in the plane, or simply a configuration. The set $\mathcal{P} = \{p_1, \ldots, p_n\}$ is the collection of *points*, $\mathcal{L} = \{\ell_1, \ldots, \ell_k\}$ is the collection of *lines* and $\mathcal{I} \subseteq \mathcal{P} \times \mathcal{L}$ is the set of *incidences* of points and lines. If the point p lies on the line ℓ then $(p, \ell) \in \mathcal{I}$ and we say that p is incident to line ℓ, or ℓ passes through p. We assume that any two distinct lines meet in at most one point.

For each point $p \in \mathcal{P}$, we write $\mathcal{L}_p \subseteq \mathcal{L}$ for the set of lines which pass through the point p. With a slight abuse of notation, we will often identify the line $\ell \in \mathcal{L}$ with the set of points which lie on the line ℓ.

Definition 4.1. Let $C = (\mathcal{P}, \mathcal{L}, \mathcal{I})$ be a configuration. For a line $\ell \in \mathcal{L}$ we define $C \backslash \ell = (\mathcal{P}', \mathcal{L} \backslash \ell, \mathcal{I} \cap (\mathcal{P}' \times (\mathcal{L} \backslash \ell)))$ to be the configuration where $\mathcal{P}'$ is the collection of points of C which do not lie solely on ℓ.

For each configuration $C = (\mathcal{P}, \mathcal{L}, \mathcal{I})$, there is a simple matroid of rank at most 3 associated to it. The matroid has ground set $\mathcal{P}$, and is defined such that a 3-subset $D \subseteq \mathcal{P}$ is dependent if and only if there exists a line $\ell \in \mathcal{L}$ with $D \subseteq \ell$. The realization spaces for point and line configurations, are the realization spaces for the matroid; see Definition 2.3. In the projective plane, any two distinct points determine a line and so any pair of points in $\mathcal{P}$ can be taken to lie on a line. However, if there is not a third element of $\mathcal{P}$ on the line, then this line does not give rise to any dependent set in the matroid. Therefore, in the remainder of this section, we require that each line in $\mathcal{L}$ contain three or more points in $\mathcal{P}$.

We note that configurations C are purely combinatorial objects. In particular, they need not be realizable, i.e. we may have $\Gamma_C = \emptyset$. For example, the Fano plane, see Example 3.11, is the smallest non-realizable configuration over $\mathbb{C}$. In the following sections, the purpose of using configurations is to give a parametrization of the irreducible components of certain hypergraph ideals via their realization spaces Γ_C. For these cases, we may assume that the configurations are realizable. However, for families of configurations that we study directly, such as forest-like configurations, we will prove that they are indeed realizable.

Theorem 4.2. *For any configuration C with at most 6 lines, Γ_C is irreducible with respect to the Zariski topology.*

We prove this theorem using the Grassmann-Cayley algebra. We review some of its theory following [Stu08].

Definition 4.3. Let V be a vector space of dimension d over $\mathbb{C}$ and let $\bigwedge(V)$ be the exterior algebra of V. The Grassmann-Cayley algebra is the vector space $\bigwedge(V)$ together with the two operations $\wedge$ and $\vee$. We denote by $a_1 \vee a_2 \vee \cdots \vee a_k \in \bigwedge(V)$ the extensor of length k which is also referred to as the join of $a_1, a_2 \ldots, a_k$. The meet operation $\wedge$ is a binary operation on two extensors A and B of length j, k with $j + k \geq d$ as:

$$A \wedge B := \sum_{\sigma} \text{sign}(\sigma) [a_{\sigma(1)}, \ldots, a_{\sigma(d-k)}, b_1, \ldots, b_k] a_{\sigma(d-k+1)} \vee \cdots \vee a_{\sigma(j)},$$

where the sum is taken over all permutations σ of $\{1, 2, \ldots, j\}$.

Any extensor $A = a_1 \vee \cdots \vee a_k$ has an associated vector space $\overline{A} = \text{span}(a_1, \ldots, a_k)$ with following properties.

Lemma 4.4. *Let A and B be two extensors. Then the following hold:*

- *Any extensor A is uniquely determined from $\overline{A}$ up to a scalar multiple* [Stu08, Section 3.3].

- *The meet of two extensors is again an extensor* [Stu08, Theorem 3.3.2 (b)].

- *We have that $A \wedge B \neq 0$ if and only if $\overline{A} + \overline{B} = V$. In this case $\overline{A \wedge B} = \overline{A} \cap \overline{B}$* [Stu08, Theorem 3.3.2 (c)].

In order to prove Theorem 4.2, we will take cases on the possible configurations. We begin by proving the following theorem which allows us to build new irreducible configurations from old. The cases in the following theorem are depicted in Figure 4 and explained further in Example 4.6.

Theorem 4.5. *Let $C = (\mathcal{P}, \mathcal{L}, \mathcal{I})$ be a configuration and ℓ a line in $\mathcal{L}$. Let $S = \{p \in \ell : |\mathcal{L}_p| \geq 3\}$ be set of the points on ℓ which lie on at least 2 other lines. Suppose that $|S| \leq 2$. If $\Gamma_{C \backslash \ell}$ is irreducible then so is Γ_C.*

Proof. Let n be the number of points in $C \backslash \ell$ and $\ell_1, \ldots, \ell_m$ be the lines of $C \backslash \ell$ which intersect ℓ in $\mathcal{L}$ at a unique point. That is, if ℓ_i intersects ℓ at the point p, then there are no other lines through p, i.e. $\mathcal{L}_p = \{\ell, \ell_i\}$. For each line ℓ_i we fix two distinct points $p_{i,1}$ and $p_{i,2}$ in $\mathcal{P} \backslash \ell$ which lie on ℓ_i. For each realization $\gamma \in \Gamma_{C \backslash \ell}$ we write $\gamma_1, \ldots, \gamma_n \in \mathbb{C}^3$ for the points of γ and $\gamma_{\ell_1}, \ldots, \gamma_{\ell_m} \subseteq \mathbb{C}^3$ for the 2-dimensional linear subspaces corresponding to the lines. Note that the subspace γ_{ℓ_i} is the linear span of the points $\gamma_{p_{i,1}}$ and $\gamma_{p_{i,2}}$. We denote by r the number of points of C which lie on ℓ and no other lines. We proceed by taking cases on $|S| \in \{0, 1, 2\}$.

Case 1. Assume that $|S| = 0$. We define the space

$$X = \Gamma_{C \backslash \ell} \times \text{Hom}_{\mathbb{C}}(\mathbb{C}^2, \mathbb{C}^3) \times (\mathbb{C}^2)^r \times \mathbb{C}^m,$$

where $\text{Hom}_{\mathbb{C}}(\mathbb{C}^2, \mathbb{C}^3)$ denotes the set of linear maps from $\mathbb{C}^2$ to $\mathbb{C}^3$. Here, we think of a linear map $\phi \in \text{Hom}_{\mathbb{C}}(\mathbb{C}^2, \mathbb{C}^3)$ as a 3×2 matrix with entries in $\mathbb{C}$. We define the subset $X' \subseteq X$ to be the collection of $(\gamma, \phi, (y_1, \ldots, y_r), (\lambda_1, \ldots, \lambda_m)) \in X$ such that the following conditions hold:

9

(a) $\operatorname{rank}(\phi) = 2$,

(b) for each $i \in [n]$, the point γ_i does not lie in $\phi(\mathbb{C}^2) \subseteq \mathbb{C}^3$,

(c) for each $i \in [r]$, the point $\phi(y_i)$ does not lie in any of the subspaces $\gamma_{\ell_1}, \ldots, \gamma_{\ell_m}$,

(d) for any $i, j \in [r]$, if $\phi(y_i) = \phi(y_j)$ then $i = j$.

Since $\Gamma_{C \setminus \ell}$, $(\mathbb{C}^2)^r$ and $\operatorname{Hom}_{\mathbb{C}}(\mathbb{C}^2, \mathbb{C}^3)$ are irreducible, we have that X is a product of irreducible varieties over an algebraically closed field. Hence, X is also irreducible. By construction, each of the conditions for the subset $X' \subseteq X$ above is a rank constraint on certain submatrices. Explicitly we have that:

- $\operatorname{rank}(\phi) \neq 2$ if all of the 2-minors of ϕ are zero,

- Assuming that $\operatorname{rank}(\phi) = 2$, a point γ_i lies inside $\phi(\mathbb{C}^2)$ if the 3-minor $[\gamma_i, \phi(1,0), \phi(0,1)]$ is zero,

- A point $\phi(y_i)$ lies inside the subspace γ_{ℓ_i} if the 3-minor $[\phi(y_i), \gamma_{p_{i,1}}, \gamma_{p_{i,2}}]$ is zero,

- If $\operatorname{rank}(\phi) = 2$ then it follows that ϕ is injective.

Therefore $X' \subseteq X$ is an open subset which implies that X' is also irreducible.

Now, for each $(\gamma, \phi, (y_1, \ldots, y_r), (\lambda_1, \ldots, \lambda_m)) \in X'$ and for each $i \in [m]$ we fix a non-zero point $q_i \in \mathbb{C}^3$ which lies in the intersection of γ_{ℓ_i} and $\phi(\mathbb{C}^2)$. Note that the ambient space is $\mathbb{C}^3$, hence any pair of 2-dimensional linear subspaces intersect. By construction of X', we have that $\phi(\mathbb{C}^2)$ and γ_{ℓ_i} do not coincide. Hence, $\dim(\phi(\mathbb{C}^2) \cap \gamma_{\ell_i}) = 1$ and q_i is unique up to a non-zero scalar multiple. We give an explicit formula for q_i using the Grassmann-Cayley algebra as follows. Recall that γ_{ℓ_i} is the span of $a_1 := \gamma_{p_{i,1}}$, $a_2 := \gamma_{p_{i,2}}$ and $\phi(\mathbb{C}^2)$ is the span of $b_1 := \phi(1,0)$, $b_2 := \phi(0,1)$. Since the spaces γ_{ℓ_i} and $\phi(\mathbb{C}^2)$ do not coincide, they must span the entire space $\mathbb{C}^3$. Since $\dim(\phi(\mathbb{C}^2) \cap \gamma_{\ell_i}) = 1$, by Lemma 4.4, we have that $q_i := [a_1, b_1, b_2]a_2 - [a_2, b_1, b_2]a_1 \in \gamma_{\ell_i} \cap \phi(\mathbb{C}^2)$ is a non-zero vector lying in the intersection. We define the map

$$\psi : X' \to (\mathbb{C}^3)^{n+m+r} : (\gamma, \phi, (y_1, \ldots, y_r), (\lambda_1, \ldots, \lambda_m)) \mapsto (\gamma, \lambda_1 q_1, \ldots, \lambda_m q_m, \phi(y_1), \ldots, \phi(y_r)).$$

By construction, ϕ is a linear map and the coordinates of the points $q_i \in \mathbb{C}^3$ are polynomials in the entries of γ. Thus ψ is a polynomial map and so it is continuous.

It remains to show that Γ_C is an open subset of the image of ψ. It is easy to see that the image of ψ contains Γ_C since any realization of C can be viewed as a realization of $\Gamma_{C \setminus \ell}$ together with some additional points on ℓ. To show that Γ_C is open in the image of ψ, we note that the image of ψ is contained in the combinatorial closure V_C^{comb}. So Γ_C is obtained from the image of ψ by removing the vanishing locus of the bases ideal of J_C; see Proposition 3.9.

Case 2. Assume that $|S| = 1$ and write $S = \{s\}$ where s is the corresponding point of intersection in C and $C \setminus \ell$. For any $\gamma \in \Gamma_C$ we denote by $\gamma_s \in \mathbb{C}^3$ the coordinates of the point corresponding to s. We define the space

$$X = \{(\gamma, \phi, (y_1, \ldots, y_r), (\lambda_1, \ldots, \lambda_m)) : \phi(1,0) = \gamma_s\} \subseteq \Gamma_{C \setminus \ell} \times \operatorname{Hom}_{\mathbb{C}}(\mathbb{C}^2, \mathbb{C}^3) \times (\mathbb{C}^2)^r \times \mathbb{C}^m$$

where the condition $\phi(1,0) = \gamma_s$ on a matrix $\phi \in \operatorname{Hom}_{\mathbb{C}}(\mathbb{C}^2, \mathbb{C}^3)$ is equivalent to the condition that the first column of ϕ is equal to γ_s. So, we have that $X \cong \Gamma_C \times \mathbb{C}^3 \times (\mathbb{C}^2)^r \times \mathbb{C}^m$ which is irreducible. We define the subset $X' \subseteq X$ identically to Case 1, except that we allow $\gamma_s \in \phi(\mathbb{C}^2)$ in condition (b). Note that $\gamma_s \neq 0$ and so by the same argument, $X' \subseteq X$ is an open subset. The remainder of the argument from Case 1, including the construction of the points $q_1, \ldots, q_m \in \mathbb{C}^3$ and the map ψ, follows identically.

Case 3. Assume that $|S| = 2$ and write $S = \{s_1, s_2\}$ for the corresponding points of intersection lying in both C and $C \setminus \ell$. For any $\gamma \in \Gamma_C$ we denote by $\gamma_{s_1}, \gamma_{s_2} \in \mathbb{C}^3$ the coordinates of the points corresponding to s_1 and s_2, respectively. We define the space

$$X = \{(\gamma, \phi, (y_1, \ldots, y_r), (\lambda_1, \ldots, \lambda_m)) : \phi(1,0) = \gamma_{s_1}, \ \phi(0,1) = \gamma_{s_2}\} \subseteq \Gamma_{C \setminus \ell} \times \operatorname{Hom}_{\mathbb{C}}(\mathbb{C}^2, \mathbb{C}^3) \times (\mathbb{C}^2)^r \times \mathbb{C}^m$$

where the conditions $\phi(1,0) = \gamma_{s_1}$ and $\phi(0,1) = \gamma_{s_2}$ on a matrix $\phi \in \operatorname{Hom}_{\mathbb{C}}(\mathbb{C}^2, \mathbb{C}^3)$ are equivalent to the conditions that the first column of ϕ is equal to γ_{s_1} and the second column of ϕ is equal to γ_{s_2}, respectively. Thus $X \cong \Gamma_C \times (\mathbb{C}^2)^r \times \mathbb{C}^m$, which is irreducible. We define the subset $X' \subseteq X$ identically to Case 1, except that we allow $\gamma_{s_1}, \gamma_{s_2} \in \phi(\mathbb{C}^2)$ in condition (b). Note that s_1 and s_2 are distinct points of C so $\operatorname{rank}(\phi) = 2$. By the same argument as in Case 1, $X' \subseteq X$ is an open subset. The remainder of the argument from Case 1, including the construction of the points $q_1, \ldots, q_m \in \mathbb{C}^3$ and the map ψ, follows identically. $\square$
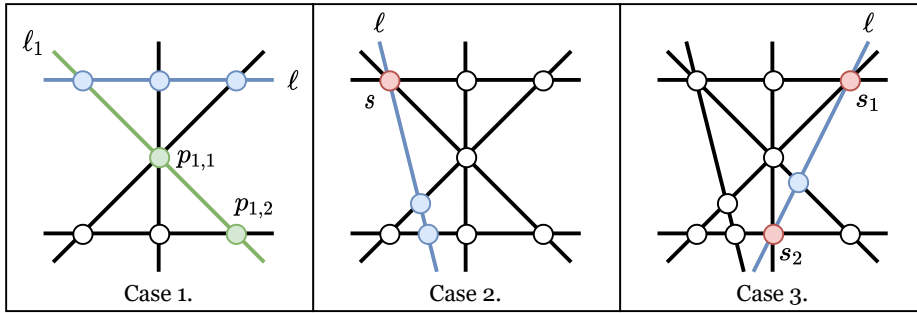
Figure 4: Configurations in Example 4.6. The left and middle figures are obtained from the figure immediately to their right by removing the line ℓ.

Example 4.6. Let $\mathcal{C}$ be the configuration in Case 1 of Figure 4 and ℓ the line indicated in the diagram. It is shown in [CMM21] that all configurations with at most 4 lines have irreducible realization spaces. In particular, $\Gamma_{C\backslash\ell}$ is irreducible. For each point p on ℓ we have that $|\mathcal{L}_p| = 2$, thus Γ_C is irreducible by Theorem 4.5. Note that, in the proof, the points $p_{1,1}, p_{1,2}$, as shown in the diagram, are used to give *coordinates* on the line ℓ_1, which allows us to give an explicit formula for the intersection point q_1 of ℓ and ℓ_1.

In the figure, the configurations labelled Case 2 and Case 3 show examples of configurations, and the line ℓ, corresponding to their respective cases in the proof of Theorem 4.5. The points which lie in the set S are labelled by s, s_1, s_2. In particular, each of the realization spaces for these configurations is irreducible.

The proof of Theorem 4.2 follows easily from the above theorem.

Proof of Theorem 4.2. We show that Γ_C is irreducible by taking cases on the number of lines in $\mathcal{C}$. If $\mathcal{C}$ contains at most 4 lines then it is irreducible by [CMM21, Corollary 4]. On the other hand, if $\mathcal{C}$ contains 5 or 6 lines, then we will show that Theorem 4.5 applies, and so we reduce to a line configuration with fewer lines.

Let ℓ be any line of $\mathcal{C}$ and let $S = \{p \in \ell : |\mathcal{L}_p| \geq 3\}$. We show that $|S| \leq 2$ by contradiction. Assume that $|S| \geq 3$. So there are three distinct points $p_1, p_2, p_3 \in \ell$ such that $|\mathcal{L}_{p_i}| \geq 3$ for each $i \in [3]$. Note that a pair of lines can intersect in at most one point. So if $\ell' \neq \ell$ is a line passing through p_1 then it does not pass through p_2 or p_3. Since $|\mathcal{L}_{p_i}| \geq 3$ for each i, the total number of distinct lines passing through p_1, p_2 and p_3 is at least 7, a contradiction. By induction, $\Gamma_{C\backslash\ell}$ is irreducible. So by Theorem 4.5 we have that Γ_C is irreducible. □

Remark 4.7. The proof of Theorem 4.2 shows that Theorem 4.5 applies to all line configurations with at most 6 lines. However, for configurations with 7 lines, consider the Fano plane depicted in Figure 3. This configuration contains 7 points each of which belongs to 3 lines. Therefore, Theorem 4.5 does not apply.

5 Combinatorial closures of configurations

In this section, we will focus on point and line configurations whose underlying graph is a forest. As in the previous section, if $\mathcal{C} = (\mathcal{P}, \mathcal{L}, \mathcal{I})$ is a point and line configuration, then each line in $\mathcal{L}$ contains at least three points in $\mathcal{P}$, unless stated otherwise.

We will give a complete description of their combinatorial closures by describing their components. Our main tool for this section is the *perturbation argument* which gives a way to determine when a particular point in the combinatorial closure belongs to a specific realization space. Let us begin by making precise our notation.

Let $\mathcal{C} = (\mathcal{P}, \mathcal{L}, \mathcal{I})$ be a point and line configuration with the ordered set of points $\mathcal{P} = \{p_1, \ldots, p_n\}$. Let G_C be a graph with vertex set $\mathcal{P}$ and edges $\{p_i, p_j\}$ such that

- both points lie on the same line $\ell \in \mathcal{L}$ (which is necessarily unique) and

- there is no point p_k lying on ℓ with $i < k < j$ (or $j < k < i$).

Intuitively, G_C is obtained from $\mathcal{C}$ by making points into vertices and lines into paths. Moreover, every edge of G_C is in exactly one of these paths. Each pair of paths intersects in at most one vertex, and this occurs if and only if the corresponding lines intersect in a point of $\mathcal{P}$.

Definition 5.1. A configuration $\mathcal{C}$ is *forest-like* if the corresponding graph G_C is a forest, i.e. it has no cycle.

Note that the graph G_C depends on the choice of ordering of the points in $\mathcal{P}$. However, the condition for G_C to be a forest is independent of this choice. Essentially, swapping the order of two points on a line corresponds to simply "flipping" an edge of G_C. We now make this more precise.

Lemma 5.2. *If C is forest-like, then any ordering of its points $\mathcal{P}$ makes G_C a forest.*

Proof. Let $\mathcal{P} = \{p_1, \ldots, p_n\}$ be an ordering of $\mathcal{P}$ such that G_C is a forest. Inductively, it suffices to show that the order $\{p_1, \ldots, p_{i-1}, p_{i+1}, p_i, p_{i+2}, \ldots, p_n\}$ gives rise to a forest.

If p_i and p_{i+1} do not lie on the same line, then there is no change to G_C. So it suffices to consider the case where p_i and p_{i+1} lie on a line $\ell \in \mathcal{L}$. Then $\{p_i, p_{i+1}\}$ is an edge in G_C. Let us call these vertices of the graph v and w. Let $A(v)$ be the set of vertices of G_C, adjacent to v, whose corresponding points in C lie on some line $\ell' \neq \ell$, and define $A(w)$ similarly. Then swapping the order of i and $i+1$ corresponds to replacing the edges in $\{\{x, v\} : x \in A(v)\} \cup \{\{x, w\} : x \in A(w)\}$ with the edges in $\{\{x, v\} : x \in A(w)\} \cup \{\{x, w\} : x \in A(v)\}$. One can see that this resulting graph is still a forest if G_C was a forest. $\square$

Because of Lemma 5.2, Definition 5.1 is well-defined for point and line configurations with unordered sets of vertices. So, we say that C is forest-like if there exists an ordering of its vertices such that the resulting configuration is forest-like.

The next result will also be helpful.

Lemma 5.3. *If $C = (\mathcal{P}, \mathcal{L}, \mathcal{I})$ is a forest-like configuration with $\mathcal{L} \neq \emptyset$, then there exists a line ℓ which intersects $C \setminus \ell$ in at most one point of $\mathcal{P}$.*

Proof. Because two lines intersect in at most one point, it will suffice to show that there exists a line ℓ which intersects $C \setminus \ell$ (at a point of $\mathcal{P}$) in at most one line. Suppose otherwise for a contradiction. Then every line intersects with at least two other lines. Let $\ell_1, \ell_2, \ldots$ be a sequence of lines in C such that ℓ_i intersects both ℓ_{i-1} and ℓ_{i+1}. Since there are only finitely many lines, eventually there will be a repeated element ℓ_k of the sequence.

Since these lines correspond to paths in G_C, there is a cycle in G_C whose set of edges includes at least one edge from each path corresponding to the lines in the sequence $\ell_k, \ell_{k+1}, \ldots, \ell_k$. This cycle contradicts the assumption that G_C is a forest. $\square$

It is perhaps worth noting that forest-likeness is a rather restrictive condition for a matroid of rank 3. However, forest-like configurations are not regular (realizable over every field) matroids. For example, no line with four or more points is a regular matroid. Each forest-like configuration is realizable over all infinite fields. To show this, it will be useful to recall the matroid-theoretic notion of *freely adding* an element to a flat of a matroid. Intuitively, this operation takes a new element of the ground set and adds it to a flat as "freely" as possible, that is, we keep as many sets independent as possible; see [Cra65] and [Oxl11, Section 7.2]. We will also use this notion in Section 6.3.

Definition 5.4. Let F_0 be a flat of a matroid M with rank function r_M. We say that M' is the single-element extension obtained by *freely adding* e to F_0 if the flats of M' fall into the following disjoint classes:

- flats F of M that do not contain F_0,

- sets $F \cup e$ where F is a flat of M that contains F_0, and

- sets $F \cup e$ where F is a flat of M that does not contain F_0, and there is no flat F' of M of rank $r_M(F) + 1$ such that $F \subseteq F'$ and $F_0 \subseteq F'$.

The following results are fairly well-known. In particular, Lemma 5.5 follows from a result of Piff and Welsh [PW70]; see also [Oxl11, Proposition 11.2.16]. We provide a more direct proof to keep the paper self-contained.

Lemma 5.5. *Let $\mathbb{F}$ be an infinite field, and let M' be a matroid obtained by freely adding an element to a flat F of a matroid M. Then M' is $\mathbb{F}$-realizable if and only if M is $\mathbb{F}$-realizable.*

Proof. Realizability over a field is closed under deletion. Therefore, if M' is $\mathbb{F}$-realizable, then so is M.

For the converse, let M be $\mathbb{F}$-realizable, and let r be the rank function of M. A flat F of M corresponds to a linear subspace of $\mathbb{F}^{r(M)}$ of dimension $r(F)$. To construct a representation of M' we must find a point p in the subspace corresponding to F that is not contained in any subspace corresponding to a flat F' of M that does not contain F. For any such flat F', the rank of $F \cap F'$ is strictly less than $r(F)$. Therefore, since $\mathbb{F}$ is an infinite field, there are infinitely many such points p. $\square$

Lemma 5.6. *Let M' be obtained by adding a coloop to an $\mathbb{F}$-realizable matroid M. Then M' is $\mathbb{F}$-realizable.*

Proof. Let A be a matrix whose columns form a realization of M. We construct a matrix realizing M' by adding a row and column to A such that all entries of this row and column are 0 except for the entry contained in both the row and column. $\square$

Proposition 5.7. *Let $C = (\mathcal{P}, \mathcal{L}, \mathcal{I})$ be a forest-like point and line configuration, and let M be the simple matroid of rank at most 3 associated with it. Then M is realizable over all infinite fields.*

Proof. We proceed by induction on the number of lines in $\mathcal{L}$. If $|\mathcal{L}| = 0$, then M is the uniform matroid $U_{r,|\mathcal{P}|}$, where $r = \min\{3, |\mathcal{P}|\}$. For all positive integers n, it is clear that $U_{n,n}$ is realizable over all fields. If $n > 3$, then $U_{3,n}$ is obtained from $U_{3,3}$ by repeatedly freely adding elements to the rank-3 flat of the matroid. Therefore, by Lemma 5.5, $U_{r,|\mathcal{P}|}$ is realizable over all infinite fields for all values of n.

Now, suppose that we have shown that the result holds for all forest-like configurations with at most $k-1$ lines and we wish to prove the result for $C = (\mathcal{P}, \mathcal{L}, \mathcal{I})$, where $|\mathcal{L}| = k$. By Lemma 5.3, there is a line ℓ in $\mathcal{L}$ that intersects with at most one of the points in $C \setminus \ell$. By the induction hypothesis, the matroid associated with $C \setminus \ell$ is realizable over all infinite fields.

First, we consider the case where ℓ intersects with a point v in $C \setminus \ell$. Let w be an additional point in ℓ, and let $T = \ell \setminus \{v, w\}$. The matroid $M \setminus T$ is obtained from $C \setminus \ell$ either by adding w as a coloop or by freely adding w to the unique rank-3 flat (depending on the rank of $C \setminus \ell$). Then M is obtained from $M \setminus T$ by freely adding the points in T to the rank-2 flat defined by v and w.

Now, we consider the case where ℓ and $C \setminus \ell$ have no points in common. Let w_1 and w_2 be points in ℓ, and let $T = \ell \setminus \{w_1, w_2\}$. The matroid $M \setminus (T \cup \{w_2\})$ is obtained from $C \setminus \ell$ either by adding w_1 as a coloop or by freely adding w_1 to the unique rank-3 flat (depending on the rank of $C \setminus \ell$). Then $M \setminus T$ is obtained from $M \setminus (T \cup \{w_2\})$ by freely adding w_2 to the unique rank-3 flat. Then M is obtained from $M \setminus T$ by freely adding the points in T to the rank-2 flat defined by w_1 and w_2.

In either case, Lemmas 5.5 and 5.6 imply that M is realizable over all infinite fields. □

One can see that the last line ℓ to be added to a forest-like configuration in the inductive process described in the proof of Proposition 5.7 has the property that $|\{p \in \ell : |\mathcal{L}_p| \geq 3\}| \leq 1$. In particular, the only point $p \in \ell$ that might have the property that $|\mathcal{L}_p| \geq 3$ is the one point that was on at least one of the other $|\mathcal{L}| - 1$ lines in $\mathcal{L}$.

So, we have the following straightforward corollary of Theorem 4.5.

Corollary 5.8. *The realization space of a forest-like configuration is irreducible with respect to Zariski topology.*

The following proof introduces the *perturbation procedure* which we will use throughout this section.

Lemma 5.9 (Perturbation procedure)**.** *Let C be a forest-like configuration with realization space $\Gamma_C \subseteq \mathbb{C}^d$. Assume that $|\mathcal{L}_p| \leq 2$ for every $p \in \mathcal{P}$. For every $\epsilon > 0$ and for any $A \in V_C^{\mathrm{comb}} \setminus \Gamma_C$ there exists $A' \in \Gamma_C$ such that $||A - A'|| < \epsilon$, where $|| \cdot ||$ is the Euclidean norm on $\mathbb{C}^{d \times n}$.*

Proof. We think of A as a realization of a configuration C_A. As $A \in V_C^{\mathrm{comb}} \setminus \Gamma_C$ and the rank of the corresponding matroid is at most 3, the dependencies satisfied by the configuration C_A that are not in C are the following:

- A point of C may be a loop in C_A,

- Two distinct points of C may coincide in C_A,

- Two distinct lines in C may coincide in C_A,

- A triple of non-collinear points in C may lie on a common line in C_A.

We now construct the realization A' by induction on the number of lines. For the base case, assume that the configuration C_A has no lines. Let us form A' by going through the points in order $p_1, \ldots, p_r$ and perturbing the corresponding vector in A by at most ϵ/r. For each $i \in [r]$, the vector corresponding to point p_i is perturbed such that it forms no dependencies, listed above, with subsets of points from $\{p_j : 1 \leq j \leq i\}$.

For the induction step, let us assume that for any forest-like configuration $C' = (\mathcal{P}', \mathcal{L}', \mathcal{I}')$ with at most $n-1$ lines, we have that for all $\epsilon > 0$ and any point $x \in V_{C'}^{\mathrm{comb}}$ there exists $x' \in \Gamma_{C'}$ such that $||x - x'|| < \epsilon$. Since the graph G_C is a forest, we may take $\ell \in \mathcal{L}$ to be a line which intersects $C \setminus \ell$ in at most one point. By the inductive hypothesis, we can perturb vectors in A corresponding to the configuration $C \setminus \ell$ so that no additional dependencies are satisfied. It remains to show that the points on the line ℓ may be perturbed to remove all additional dependencies. We proceed by applying some or all of the following steps in order. In particular, after each step we ensure that no new dependencies are introduced which would be removed in a previous step.

(a) Suppose that the line ℓ intersects a point p in C_A which does not belong to ℓ in C. We may perturb the line ℓ in A to produce A' so that ℓ does not pass through the point p. In particular, if ℓ contains k points which are not p, then we perturb each point by at most ϵ/k.

(b) Suppose that the line ℓ coincides with another line ℓ' in C_A. Since ℓ intersects $C \setminus \ell$ in at most one point, we may rotate the line ℓ in A (i.e. perturb each point along the line) by small amount to obtain A' so that ℓ does not coincide with any other line. In particular, if ℓ contains k points which do not lie on the intersection of ℓ and ℓ' in C, then we perturb each point by at most ϵ/k.
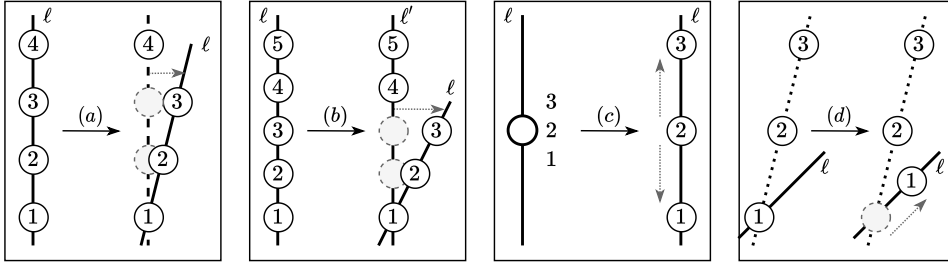
13

Figure 5: Depiction of the steps in the perturbation procedure: (a) line ℓ is perturbed away from point 4, (b) points $2, 3 \in \ell'$ are perturbed off ℓ, (c) coincident points $1, 2, 3$ are perturbed away from each other along ℓ and (d) point 1 is perturbed along ℓ away from the intersection.

(c) Suppose that two points in ℓ coincide or some of the points are loops. We may perturb these points away from each other so that they remain on the line ℓ. Similarly if any of the points are loops, then we can perturb these points away from zero by the same method. In particular, if a point p in C_A is a loop then we recall that p is incident to at most two lines. If it is incident to exactly two lines ℓ, ℓ' in C, then we may perturb p along $\ell \cap \ell'$. Otherwise if p is incident only to ℓ in C then we may perturb it to some non-zero point on ℓ.

(d) Suppose that for some point p_i in ℓ, there are two distinct points p_j, p_k in $C \backslash \ell$ such that p_i, p_j and p_k lie on a line ℓ' in C_A but do not lie on a line in C. By the step (a) of the procedure, we have that ℓ and ℓ' are distinct lines. For this case, it is useful to consider all lines of C_A, including those which contain only two points. The points of intersection between ℓ and all other distinct lines of C_A is a finite set. Since we work over the infinite field $\mathbb{C}$ and the line is homeomorphic to $\mathbb{C}^1$ with respect to Euclidean topology, therefore we may perturb p_i along ℓ so that it lies only on ℓ and no other line.

As a result of the above procedure, we have constructed a realization A' of C with $||A - A'|| < \epsilon$. $\qquad \square$

Remark 5.10. By perturbation procedure, the corresponding configurations are realizable over any subfield of $\mathbb{C}$.

Let C be the configuration with two lines 123 and 345. Consider the following points in V_C^{comb}:

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} \quad \text{and} \quad A' = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & \epsilon & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}. \tag{2}$$

Note that $A \notin \Gamma_C$ is not a realization of C because 3 is a loop in C_A. See Figure 6. Following the perturbation argument, we perturb A to A' which corresponds to moving point 3 so that it is non-zero and lies on the intersection of the planes spanned by 1, 2 and 4, 5.
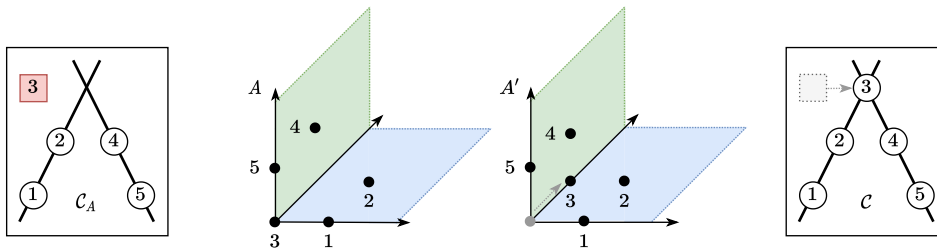

Figure 6: On the right a configuration C and on the left the configuration C_A for the matrix A in (2), together with a depiction of the perturbation of point 3 taking $A \in V_C^{\mathrm{comb}} \backslash \Gamma_C$ to $A' \in \Gamma_C$. The loop 3 in C_A is shown in a square and the shaded planes in $\mathbb{C}^3$ correspond to the lines in C_A and C.

Theorem 5.11. *Let $C = (\mathcal{P}, \mathcal{L}, \mathcal{I})$ be a forest-like configuration with $|\mathcal{L}_p| \le 2$ for every $p \in \mathcal{P}$. Then the combinatorial closure V_C^{comb} coincides with the matroid variety V_C and in particular is irreducible.*

Proof. Recall that $V_C^{\mathrm{comb}} = V_{C(C)}$ where $C(C)$ denotes the circuits of the matroid corresponding to C. So it remains to show that $V_{C(C)} = V_C$. It is clear that $V_C \subseteq V_{C(C)}$. For the opposite inclusion, by the perturbation procedure each $A \in V_{C(C)}$ is a limit point of the configuration space Γ_C with respect to the Euclidean topology. This implies $V_{C(C)} \subseteq V_C$, since the Zariski topology is coarser than the Euclidean topology. $\qquad \square$

Definition 5.12. Let C be a point and line configuration and p a point. Let $\mathcal{D}$ be the dependent sets of the matroid associated to C on ground set $\mathcal{P}$. By *setting $P \subseteq \mathcal{P}$ to be loops*, we mean the matroid M whose dependent sets are given by $\mathcal{D} \cup \{D \subseteq \mathcal{P} : P \cap D \neq \emptyset\}$. In other words the circuits of M are: the circuits of the matroid associated to C which do not meet P; and P thought of as a set of loops.

Before proving an analogous result for a general forest-like point and line configuration, we consider the following simple configuration which has at most one point that lies on more than two lines.

Lemma 5.13. *Let $C = (\mathcal{P}, \mathcal{L}, \mathcal{I})$ be a forest-like configuration which has a unique point $p := p_i$ such that $|\mathcal{L}_p| > 2$ and for all other points p_j we have $|\mathcal{L}_{p_j}| \leq 2$. Then the combinatorial closure V_C^{comb} has exactly two irreducible components: one is the central component and the other arises by setting the intersection point p to be a loop.*

Proof. For each element $A \in V_C^{\mathrm{comb}}$, we write $A_p \in \mathbb{C}^3$ for the column vector corresponding to the intersection point p. We have that V_C is the central component of V_C^{comb} and we write $V_0 = \{A \in V_C^{\mathrm{comb}} : A_p = \underline{0}\}$ for the collection of elements of the combinatorial closure where the intersection point is zero. Let C_0 be the configuration obtained from C by setting the intersection point to be a loop. Clearly, we have $V_{C_0} \cup V_C \subseteq V_0 \cup V_C \subseteq V_C^{\mathrm{comb}}$. Since Γ_{C_0} and Γ_C are irreducible by Corollary 5.8, the lemma follows from showing that $V_{C_0} \cup V_C = V_C^{\mathrm{comb}}$ by proving the opposite inclusion.

Take any element $A \in V_C^{\mathrm{comb}}$ and fix $\epsilon > 0$. We will show that there exists $A' \in \Gamma_{C_0} \cup \Gamma_C$ such that $||A - A'|| < \epsilon$ by applying the perturbation procedure. Let us take cases on whether A_p is zero.

Case 1. Assume that $A_p = \underline{0}$. Since C_0 is a configuration that contains only points p_j with $|\mathcal{L}_{p_j}| \leq 2$, we may apply the perturbation procedure to $C_0 \backslash p$. As a result we have $A' \in \Gamma_{C_0}$.

Case 2. Assume that $A_p \neq \underline{0}$. We may now apply steps (a) to (d) in the proof of the perturbation procedure to construct a point $A' \in \Gamma_C$ with $||A - A'|| < \epsilon$. In the procedure, by assumption we have that $A_p \neq \underline{0}$. This assumption guarantees that if p, p_1, p_2 are collinear points in C then the corresponding points are collinear in C_A, and all lines arising in this way pass through the common point p. $\square$

We are now ready to state our main result in this section for a general forest-like point and line configuration.

Theorem 5.14. *Let $C = (\mathcal{P}, \mathcal{L}, \mathcal{I})$ be a forest-like configuration. Let $S = \{p \in \mathcal{P} : |\mathcal{L}_p| \geq 3\}$ be the collection of points contained in at least 3 lines. Then the combinatorial closure V_C^{comb} has at most $2^{|S|}$ irredundant irreducible components. Moreover, these components can be obtained from C by setting a subset of S to be loops.*

Proof. For each element $A \in V_C^{\mathrm{comb}}$ and point $p \in \mathcal{P}$, we write $A_p \in \mathbb{C}^3$ for the column vector corresponding to the point p. For each subset $J \subseteq S$, we write C_J for the configuration obtained from C by setting the points in J to be loops. We will show that $\bigcup_{J \subseteq S} V_{C_J} = V_C^{\mathrm{comb}}$. Note that by Corollary 5.8, each variety V_{C_J} is irreducible. By construction, it is immediately clear that $V_{C_J} \subseteq V_C^{\mathrm{comb}}$ for each J. So, to prove the theorem, it remains to show the opposite inclusion, i.e. $V_C^{\mathrm{comb}} \subseteq \bigcup_{J \subseteq S} V_{C_J}$. Fix $A \in V_C^{\mathrm{comb}}$ and $\epsilon > 0$. We will construct $A' \in \Gamma_{C_J}$ for some $J \subseteq S$ such that $||A - A'|| < \epsilon$.

We proceed by induction on $s := |S|$. If $s = 1$, then the result follows by Lemma 5.13. So let us assume that $s > 1$. Since G_C is a forest, we can find a point $p \in S$ such that the path (if it exists) between any other pair of points in S does not pass through p. Let $P \subseteq \mathcal{P}$ be the collection of points $q \in \mathcal{P}$ lying in the same connected component of G_C as p, such that for all $p' \in S \backslash \{p\}$, if there is a path from q to p', then the path passes through p. By convention, we assume $p \in P$.

Consider the configuration $C' = (\mathcal{P}', \mathcal{L}', \mathcal{I}')$ obtained from C by removing points P. We have $S' := \{p \in \mathcal{P}' : |\mathcal{L}'| \geq 3\} = S \backslash \{p\}$. Clearly, C' is a forest-like configuration and $|S'| = s - 1$. So, by induction, we can perturb the points in $\mathcal{P}'$ so that they lie in a realization space $\Gamma_{C'_{J'}}$ for some $J' \subseteq S'$. More precisely, there exists a subset $J' \subseteq S \backslash \{p\}$ and a realization $A'_{\mathcal{P}'} \in \Gamma_{C'_{J'}}$ such that $||A_{\mathcal{P}'} - A'_{\mathcal{P}'}|| < \epsilon/2$, where $A_{\mathcal{P}'}$ is the set of vectors obtained from A by removing the points P. We proceed by taking cases on whether A_p is zero.

Case 1. Assume that $A_p = \underline{0}$. Let $J = J' \cup \{p\}$, we construct $A' \in \Gamma_{C_J}$ from $A'_{\mathcal{P}'}$ by taking the vectors in A for the points in $P \backslash \{p\}$ and applying the perturbation procedure. By the perturbation procedure we have ensured that $A' \in \Gamma_{C_J}$ and by perturbing each point $p \in P$ by a distance of at most $\epsilon/2|P|$, we have that $||A - A'|| \leq ||A_{\mathcal{P}'} - A'_{\mathcal{P}'}|| + (|P| - 1)(\epsilon/2|P|) < \epsilon$.

Case 2. Assume that $A_p \neq \underline{0}$. Let $J = J'$ and construct $A' \in \Gamma_{C_J}$ from $A'_{\mathcal{P}'}$ by taking the vectors in A for points in P and applying steps (a) to (d) in the proof of the perturbation procedure. In the procedure, by assumption we have that $A_p \neq \underline{0}$. This assumption guarantees that all lines of C passing through p are contained in a line of C_A passing through p. As a result we obtain a realization $A' \in \Gamma_{C_J}$ and by perturbing each point $p \in P$ by a distance of at most $\epsilon/2|P|$, we have that $||A - A'|| \leq ||A_{\mathcal{P}'} - A'_{\mathcal{P}'}|| + |P|(\epsilon/2|P|) < \epsilon$. $\square$

Remark 5.15. It is not hard (but a bit tedious) to classify all irreducible components of V_C^{comb} for forest-like configurations C. This classification follows from the same proof of Theorem 5.14 by taking into account that components may become redundant; see Example 5.16.
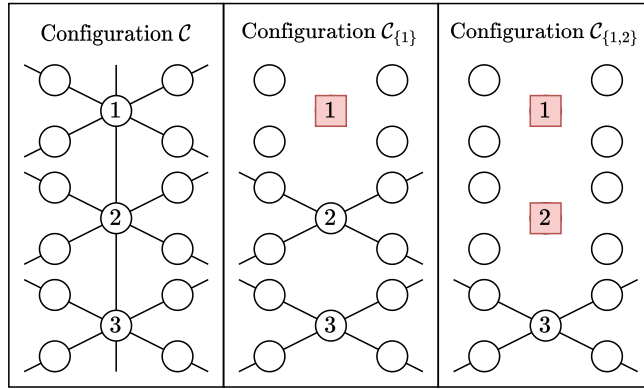
Figure 7: Depiction of $\mathcal{C}$, $\mathcal{C}_{\{1\}}$ and $\mathcal{C}_{\{1,2\}}$ in Example 5.16. Shaded squares represent loops of the configuration.

Example 5.16. Consider the configuration $\mathcal{C}$ in Figure 7. Note that the points $1, 2$ and 3 have degree 3 in the graph $G_\mathcal{C}$. So by Theorem 5.14, $V_\mathcal{C}^{\mathrm{comb}}$ has at most 8 irredundant irreducible components which arise from setting the intersection points: $1, 2,$ and 3 to be loops in some combination. When points are set to loops, we remove the lines which contain at most two non-loop points, as they do not contribute any new dependencies. The variety of $\mathcal{C}_{\{1\}}$ appears as an irredundant component of $V_\mathcal{C}^{\mathrm{comb}}$. However $\mathcal{C}_{\{1,2\}}$ gives rise to a redundant component since its variety is contained the variety of $\mathcal{C}_{\{1\}}$. The irreducible decomposition of $V_\mathcal{C}^{\mathrm{comb}}$ has 4 irredundant components which are: the central component $V_\mathcal{C}$, and the non-central components $V_{\mathcal{C}_{\{1\}}}, V_{\mathcal{C}_{\{2\}}}, V_{\mathcal{C}_{\{3\}}}$.

6 Consecutive forest hypergraphs Δ_G

We fix a $d \times n$ matrix $X = (x_{i,j})$ of indeterminates and the polynomial ring $R = \mathbb{C}[X]$. Throughout this section, we fix a forest G (acyclic graph) with vertices labeled $1, \dots, n$.

Definition 6.1. We define the *consecutive forest hypergraph of G* as

$$\Delta_G = \min\left(P_2(G) \cup \binom{[n]}{4}\right) \quad \text{where} \quad P_2(G) = \{P \subseteq V(G) : P \text{ is a 2-path in } G\}.$$

We recall that a k-path P in G is a subgraph of G whose vertex set is a non-repeating sequence $p_1, \dots, p_{k+1} \subseteq V(G)$ and whose edge set is $\{\{p_i, p_{i+1}\} : 1 \le i \le k\}$. We identify P with its vertex set, and simply write $C \subseteq P$ instead of $C \subseteq V(P)$. If Q is another path, then $P \cup Q$ is the set of all vertices of G lying in P and Q.

Remark 6.2. The hypergraph ideal I_{Δ_G} of the n-path G is exactly the ideal $I_{3n}(3)$ of adjacent minors studied in [HS04]. We note that our construction generalises this particular family by allowing G to be an arbitrary graph. In [HS04], the ideals $I_{mn}(m)$ are studied for general $m \ge 3$ which can be thought of as the consecutive forest hypergraph of the n-path except with higher order minors. These ideals can be studied using positroid varieties as described in [KLS13]. It is not too difficult to show that each of the prime components of $I_{mn}(m)$ is a positroid variety. For instance, one can use the characterisation of positroids via their excluded minors [Oh09, Theorem 16] or the decomposition of the Grassmannian into positroid varieties [KLS13, Section 5.2]. In particular, we note that the defining ideals of positroid varieties are generated by minors [KLS13, Theorem 5.15]. However, this is no longer true for the consecutive forest hypergraph varieties. More precisely, the varieties arising in the irreducible decomposition of V_{Δ_G} are, in general, not positroid varieties; for instance see Example 6.21.

6.1 Minimal matroids. In this section, we will define so-called prime collections which are collections of subsets of vertices of G. To each prime collection $\mathcal{S}$ we will associate a unique matroid $M_\mathcal{S}$. We will then see that such matroids are realizable over the real numbers and have irreducible realization spaces. Moreover, we prove that such matroids appear as minimal matroids for consecutive forest hypergraphs.

Definition 6.3. Let $\mathcal{S}$ be a collection of singleton subsets of $[n]$. We say $\mathcal{S}$ is a *prime collection of singletons* for G if $\mathcal{S}$ satisfies the following inductive definition.

- The empty set $\mathcal{S} = \emptyset$ is a prime collection of singletons.
- If $|\mathcal{S}| \ge 1$, then $\mathcal{S} = \{\{s_1\}, \dots, \{s_t\}\}$ is a prime collection of singletons if, for each natural number i with $1 \le i \le t$, $\mathcal{S} \backslash \{s_i\}$ is a prime collection of singletons and s_i satisfies the following two rules:
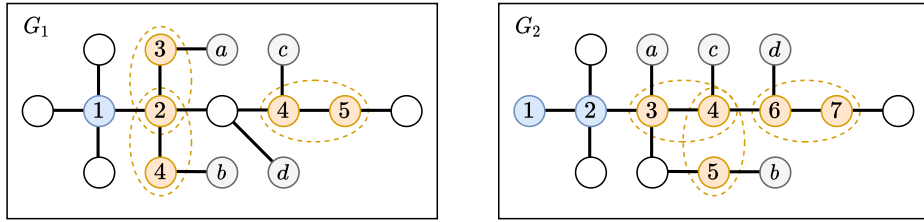
Figure 8: Graphs in Examples 6.5 and 6.7. (Left) G_1 with prime collection $\mathcal{S}_1 = \{1, 23, 24, 45\}$ and clouds 234 and 45. (Right) G_2 has a valid collection $\mathcal{S}_2 = \{1, 2, 34, 45, 67\}$ which is not prime. The clouds are 345 and 67.

1. s_i is not a leaf or isolated vertex of G',

2. If s_i is adjacent to a leaf of G' then it has degree at least 3 in G',

where G' is the induced subgraph of G obtained by deleting the vertices in the set $\{s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_t\}$.

For ease of notation and when it is clear, we denote a collection of singletons $\mathcal{S} = \{\{s_1\}, \ldots, \{s_t\}\}$ as $\{s_1, \ldots, s_t\}$.

Definition 6.4. Let $\mathcal{S}$ be a collection of singleton subsets $\text{sing}(\mathcal{S})$ and 2-subsets of $V(G)$. Let G' be the induced subgraph of G obtained by removing all vertices in $\text{sing}(\mathcal{S})$.

- We say $\mathcal{S}$ is a *prime collection* for G if:

  - The $\text{sing}(\mathcal{S})$ is a prime collection of singletons for G.
  - The 2-subsets in $\mathcal{S}$ are a subset of edges of G' such that for every vertex $v \in V(G')$ which is incident to an edge in $\mathcal{S}$, there exists an edge $\{v, w\}$ of G' which is not in $\mathcal{S}$.

- We say $\mathcal{S}$ is a *valid collection* for G if each 2-subset is disjoint from each singleton set.

Note that every prime collection is a valid collection, but the converse is not true.

Example 6.5. Consider the graph G_1 in Figure 8. Let $\mathcal{S}_1 = \{1, 23, 24, 45\}$. We see that the singletons, i.e. $\{1\}$, form a prime collection of singletons for G_1 because 1 is adjacent to a leaf of G_1 and has degree at least three. Then we consider the induced subgraph G' of G_1 obtained by removing the vertex 1. The vertices which lie in some 2-subset in $\mathcal{S}$ are shaded yellow and vertices in G' which do not lie in any 2-subset are white. Since every yellow vertex is adjacent to a white vertex, we have that the 2-subsets form a prime collection of 2-subsets for G'. Therefore, $\mathcal{S}_2$ is a prime collection for G_1.

On the other hand, in the graph G_2, the set $\mathcal{S}_2 = \{1, 2, 34, 45, 67\}$ is a valid collection since each 2-subset in $\mathcal{S}_2$ does not contain any of the singleton subsets. However $\mathcal{S}$ is not a prime collection because 1 is a leaf and so the singletons do not form a prime collection of singletons. Additionally $45 \in \mathcal{S}$ is not an edge of G_2.

In the following, we show how a valid collection for G can be extended to the set of circuits of a dependent matroid for the consecutive forest hypergraph Δ_G.

Definition 6.6. Let $\mathcal{S}$ be a valid collection.

- Let $G(\mathcal{S})$ be the graph whose vertex set is given by the union of all 2-subsets of $\mathcal{S}$ and whose edges are the 2-subsets of $\mathcal{S}$. A *cloud of $\mathcal{S}$* is the set of vertices of a connected component of $G(\mathcal{S})$.

- We say that a path $P : v_0, v_1, \ldots, v_t$ in G *crosses* the cloud containing v_j and v_k if there exists $i < j < k < \ell$ such that some cloud of $\mathcal{S}$ contains $\{v_j, v_k\}$ but no cloud of $\mathcal{S}$ contains either $\{v_i, v_j\}$ or $\{v_k, v_\ell\}$. (This may mean that v_i or v_ℓ does not lie in any cloud.)

- We say that the subset $A \subset V(G)$ is *blocked* by $\mathcal{S}$ if there exist $v, w \in A$ such that one of the following conditions holds:

  (i) The vertices v and w lie in different connected components of G.

  (ii) The vertices v and w lie in the same connected component of G and the unique path from v to w in G crosses a cloud or contains a vertex $i \in V(G) \backslash \{v, w\}$ such that $\{i\}$ is a singleton in $\mathcal{S}$.

We note that if a set $A \subseteq V(G)$ is not blocked by some valid set $\mathcal{S}$ then it follows that A is contained within a connected component of G.

17

Example 6.7. In Figure 8, we have illustrated a prime collection $\mathcal{S}_1$ for the graph G_1 which contains two clouds 234 and 45. The path from a to b crosses the cloud 234 whereas the path from c to d does not cross any cloud. Therefore any set of vertices containing a and b is blocked by $\mathcal{S}$, but $\{4, c, d\}$ is not blocked by $\mathcal{S}$.

We also illustrate a non-prime valid collection $\mathcal{S}_2$ for the graph G_2 which contains two clouds. The path from a to b crosses the cloud 345 since the path passes through both 3 and 5. However the path from c to d does not cross any clouds of $\mathcal{S}_2$

We now give the construction of a matroid whose circuits include $\mathcal{S}$ for a given valid collection $\mathcal{S}$.

Proposition 6.8. *Let $\mathcal{S}$ be a valid collection. Let $\mathcal{C}$ be the collection of subsets of $[n]$ consisting of the*

1. *singletons in $\mathcal{S}$,*

2. *2-subsets of clouds of $\mathcal{S}$,*

3. *3-subsets of $[n]$ which are not blocked and do not contain any set in $\mathcal{S}$, and*

4. *4-subsets of $[n]$ containing none of the sets listed in 1, 2, or 3 above.*

Then $\mathcal{C}$ is the collection of circuits of a matroid.

Proof. We need to check that $\mathcal{C}$ satisfies the circuit elimination axiom. That is, if $C_1, C_2 \in \mathcal{C}$ and $x \in C_1 \cap C_2$, we must check that there is a set in $\mathcal{C}$ contained in $(C_1 \cup C_2) \backslash \{x\}$. This is clear if either C_1 or C_2 is a set of size 1 or 4. If C_1 and C_2 both have size 2, then C_1 and C_2 are contained in the same cloud. Thus, $(C_1 \cup C_2) \backslash \{x\}$ is contained in the same cloud, implying that $(C_1 \cup C_2) \backslash \{x\} \in \mathcal{C}$.

Suppose $|C_1| = 2$ and $|C_2| = 3$. Let $C_1 = \{v, x\}$ and $C_2 = \{x, y, z\}$.

Since C_2 is not blocked, it is contained in a connected component of G.

Let P_1 be the path from x to y; let P_2 be the path from v to y; and let P_3 be the path from v to x. Since $\{x, y, z\} \in \mathcal{C}$, we have $\{x, y\} \notin \mathcal{C}$. Therefore, y is not contained in the same cloud that contains v and x. Since a cloud contains every vertex on the path between two vertices in the cloud, we see then that either $P_3 \subseteq P_1$, or $P_3 \subseteq P_2$, or $P_3 = P_1 \cap P_2$. In either case, the fact that $\{x, y\}$ is not blocked implies that $\{v, y\}$ is not blocked. Similarly, $\{v, z\}$ is not blocked. We already know that $\{y, z\}$ is not blocked because $\{y, z\} \subseteq \{x, y, z\}$. Therefore, we conclude that $\{v, y, z\}$ is not blocked and contains a member of $\mathcal{C}$.

Finally, we consider the case where $|C_1| = |C_2| = 3$. If $|C_1 \cap C_2| = 1$, then $(C_1 \cup C_2) \backslash \{x\}$ has size 4 and therefore contains a member of $\mathcal{C}$. If $|C_1 \cap C_2| = 2$, let $C_1 = \{v, x, y\}$ and $C_2 = \{w, x, y\}$. We will show that $(C_1 \cup C_2) \backslash \{x\} = \{v, w, y\}$ contains a member of $\mathcal{C}$. Since $\{v, x, y\}$ and $\{w, x, y\}$ are not blocked, the pairs $\{v, y\}$ and $\{w, y\}$ are not blocked. Suppose for a contradiction that the pair $\{v, w\}$ is blocked. Recalling that there is a unique path between any pair of vertices in a connected component of a forest, one can see that either $\{v, x\}$ or $\{w, x\}$ is blocked, a contradiction, or x is contained in the cloud that blocks $\{v, w\}$. So we deduce that x is contained in the cloud that blocks $\{v, w\}$. But then, either $\{v, y\}$ or $\{w, y\}$ is blocked, or x and y are contained in the same cloud, a contradiction. $\square$

Notation. Given a valid collection $\mathcal{S}$, we denote by $M_{\mathcal{S}}$ its corresponding matroid from Proposition 6.8.

Remark 6.9 (Flats of $M_{\mathcal{S}}$). The set of loops L of $M_{\mathcal{S}}$ is the set of singletons in $\mathcal{S}$. Note that the loops of a matroid are contained in every flat of the matroid. The flats of $M_{\mathcal{S}}$ of rank 1 are either of the form $B \cup L$, where B is a cloud, or of the form $\{x\} \cup L$, where $x \in [n]$ is a non-loop element contained in no cloud. The flats of rank 2 are either of the form $F \cup L$, where F is an inclusion-wise maximal set that is not blocked, or of the form $F_1 \cup F_2$, where F_1 and F_2 are flats of rank 1 such that $\{x_1, x_2\}$ is blocked for every pair $\{x_1, x_2\} \in F_1 \cup F_2$. Since, all 4-subsets of $[n]$ are dependent in $M_{\mathcal{S}}$, the matroid has rank at most 3.

We now show that the matroids $M_{\mathcal{S}}$ associated to prime collections $\mathcal{S}$ in Definition 6.4 are precisely the minimal matroids for Δ_G. Firstly, we observe that the simplification of $M_{\mathcal{S}}$ is the matroid associated to a forest-like configuration.

Proposition 6.10. *Let $\mathcal{S}$ be a prime collection. Then the simplification of $M_{\mathcal{S}}$ is the matroid of a forest-like configuration.*

Proof. By definition, the simplification $(M_{\mathcal{S}})^s$ is the matroid obtained from $M_{\mathcal{S}}$ by deleting all its loops and deleting elements from parallel classes such that each parallel class contains only one element. Since $M_{\mathcal{S}}$ has rank at most 3, it follows that $(M_{\mathcal{S}})^s$ is completely determined by its dependent rank-2 flats (that is, the flats of rank 2 with at least 3 elements). We define the point and line configuration $\mathcal{C}$ whose points are the elements of the ground set of $(M_{\mathcal{S}})^s$ and whose lines are the dependent rank-2 flats of $(M_{\mathcal{S}})^s$. Clearly, we have that the matroid associated to $\mathcal{C}$ is equal to $(M_{\mathcal{S}})^s$. It remains to show that $\mathcal{C}$ is forest-like.

We proceed by induction on $|\mathcal{S}|$. If $\mathcal{S} = \emptyset$, then we have that the matroid associated to $\mathcal{C}$ is equal to $M_{\mathcal{S}}$. For any labelling of the points in $\mathcal{C}$, we show that $G_{\mathcal{C}}$ is a disjoint union of paths. By Definition 6.6,

the 3-subsets of $V(G)$ contained in a connected component of G are exactly those which are not blocked by $\mathcal{S}$. So, by the definition of the matroid $M_{\mathcal{S}}$, its dependent rank-2 flats are the connected components of G. It follows immediately from the definition of the graph G_C that its connected components are paths. Hence $\mathcal{C}$ is forest-like.

Assume that $|\mathcal{S}| \geq 1$. Then, either $\mathcal{S}$ contains a singleton of $V(G)$ or an edge of G.

Case 1. Let $\{v\} \in \mathcal{S}$ be a singleton. By definition, we have that $\mathcal{S} \setminus \{v\}$ is a prime collection. By induction, the simplification $(M_{\mathcal{S} \setminus \{v\}})^s$ is the matroid of a forest-like configuration C'. Since v is not contained in any cloud of $\mathcal{S}$, it follows that v is contained in a unique dependent rank-2 flat $F = \{v, f_1, \ldots, f_k\}$ of $(M_{\mathcal{S} \setminus \{v\}})^s$. The dependent rank-2 flats of $(M_{\mathcal{S}})^s$ are obtained from the dependent rank-2 flats of $(M_{\mathcal{S} \setminus \{v\}})^s$ by replacing F with a disjoint collection of flats $F_1, \ldots, F_d$ which partitions $\{f_1, \ldots, f_k\}$. Each flat F_i corresponds to a neighbor of v in G. To see this, let $A \subseteq V(G) \setminus v$ be a 3-subset and suppose that A is not blocked by $\mathcal{S} \setminus \{v\}$. Then A is not blocked by $\mathcal{S}$ if and only if for any pair of vertices $x, y \in A$ we have that the path from x to y in G does not pass through v. Therefore the point and line configuration $\mathcal{C}$ is obtained from C' by replacing the line containing v with d non-intersecting lines. Since C' is forest-like, it follows that $\mathcal{C}$ is also forest-like.

Case 2. Let $\{v, w\} \in \mathcal{S}$ be an edge of G. Since $\mathcal{S}$ is a prime collection, the induced subgraph G' of G on the vertices of the cloud of $\mathcal{S}$ containing $\{v, w\}$ is connected. By assumption G is a forest, hence G' is forest. So, without loss of generality, we may assume that v is a leaf in G'. By the definition of a prime collection, it follows that $\mathcal{S} \setminus \{v, w\}$ is a prime collection. By induction, the simplification $(M_{\mathcal{S} \setminus \{v, w\}})^s$ is the matroid of a forest-like configuration C'. Let F be the unique dependent flat of rank 2 of $(M_{\mathcal{S} \setminus \{v, w\}})^s$ that contains v. The dependent rank-2 flats of $(M_{\mathcal{S}})^s$ are obtained from the dependent rank-2 flats of $(M_{\mathcal{S} \setminus \{v, w\}})^s$ by replacing F with a collection of flats $F_1, \ldots, F_d$ such that $F_i \cap F_j = \{p\}$ where p is the element of the ground set of $(M_{\mathcal{S}})^s$ that corresponds to the cloud of $\mathcal{S}$ containing v and w. To see this, let $A \subseteq V(G)$ be a 3-subset and suppose that A is not blocked by $\mathcal{S} \setminus \{v, w\}$. Then A is blocked by $\mathcal{S}$ if and only if there exist $x, y \in A$ such that the path from x to y in G crosses a cloud of $\mathcal{S}$ via the edge $\{v, w\}$. Therefore, the point and line configuration C' is obtained from $\mathcal{C}$ by removing the line containing v and replacing it with d lines that pass through a common vertex. Since C' is forest-like, it follows that $\mathcal{C}$ is also forest-like. $\square$

We now state and prove our main result, and then the auxiliary results used in this proof.

Theorem 6.11. *The minimal matroids for* Δ_G *are* $\mathcal{M}_G = \{M_{\mathcal{S}} : \mathcal{S}$ *is a prime collection*$\}$.

Proof. On the one hand, if we take a minimal matroid M for Δ_G, then by Lemma 6.17 we have that $M = M_{\mathcal{S}}$ for some prime collection $\mathcal{S}$. On the other hand, let $\mathcal{S}$ be a prime collection. First, we observe that, for every prime collection $\mathcal{S}$, every member of Δ_G is indeed a dependent set in $M_{\mathcal{S}}$. More precisely, $\Delta_G \subseteq \mathcal{D}(M_{\mathcal{S}})$ because every path in G with three vertices either contains a singleton set in $\mathcal{S}$, contains two elements from a cloud of $\mathcal{S}$, or is not blocked. Since $\Delta_G \subseteq \mathcal{D}(M_{\mathcal{S}})$, Lemma 6.17 implies that there is a prime collection $\mathcal{T}$ such that $\Delta_G \subseteq \mathcal{D}(M_{\mathcal{T}}) \subseteq \mathcal{D}(M_{\mathcal{S}})$ and such that $M_{\mathcal{T}}$ is a minimal matroid for Δ_G. But Lemma 6.18 implies that $\mathcal{S} = \mathcal{T}$. Therefore, $M_{\mathcal{S}}$ is a minimal matroid for Δ_G. $\square$

Example 6.12. Let G be the graph on vertex set $[7]$ with edges $E(G) = \{12, 23, 34, 45, 56, 47\}$. The consecutive forest hypergraph for G is given by $\Delta = \{123, 234, 345, 347, 456, 457\}$. In Figure 9, we write down the prime collections $\mathcal{S}$ for G. For each $\mathcal{S}$, the matroid $M_{\mathcal{S}}$ is the matroid of a point and line configuration which we illustrate in the figure. For instance if $\mathcal{S} = \emptyset$, then $M_{\mathcal{S}}$ is the matroid of the point and line configuration with all seven points lying on a single line. If $\mathcal{S} = \{34, 45\}$, then the corresponding configuration has a line containing the points $1, 2, 3, 4, 5$ and two free points 6 and 7. The three points $3, 4, 5$ coincide since they are a cloud of $\mathcal{S}$.

We devote the rest of this subsection to prove the lemmas used in the proof of Theorem 6.11. We first prove a matroid result that we will need later.

Lemma 6.13. *Let M be a matroid, and let $\{a, b, c, d\}$ be a subset of the ground set of M such that $\{a, d\}$ is an independent set. If $\{a, b, d\}$ and $\{a, c, d\}$ are dependent sets, then so is $\{a, b, c\}$.*

Proof. Since $\{a, b, d\}$ is dependent, the rank of $\{a, b, d\}$ is at most 2. Since $\{a, c, d\}$ is dependent and $\{a, d\}$ is independent, c is in the closure of $\{a, d\}$. Therefore, c is in the closure of $\{a, b, d\}$. Since the rank of $\{a, b, d\}$ is at most 2 and c is in the closure of $\{a, b, d\}$, it follows that the rank of $\{a, b, c, d\}$ is at most 2. Therefore, $\{a, b, c\}$ is a dependent set. $\square$

Now we consider the loops of a minimal matroid and give an algorithmic formulation of Definition 6.3. Given a collection $\mathcal{S}$ of singletons, if $\mathcal{S}$ *passes* the following algorithm, it is a good candidate to be a prime collection of singletons for G. While every prime collection of singletons will pass Algorithm 1, not all sets which pass the algorithm are prime collections of singletons. In order to guarantee that a set $\mathcal{S}$ is a prime collection of singletons, we require that $\mathcal{S}$ passes regardless of the ordering of its elements.
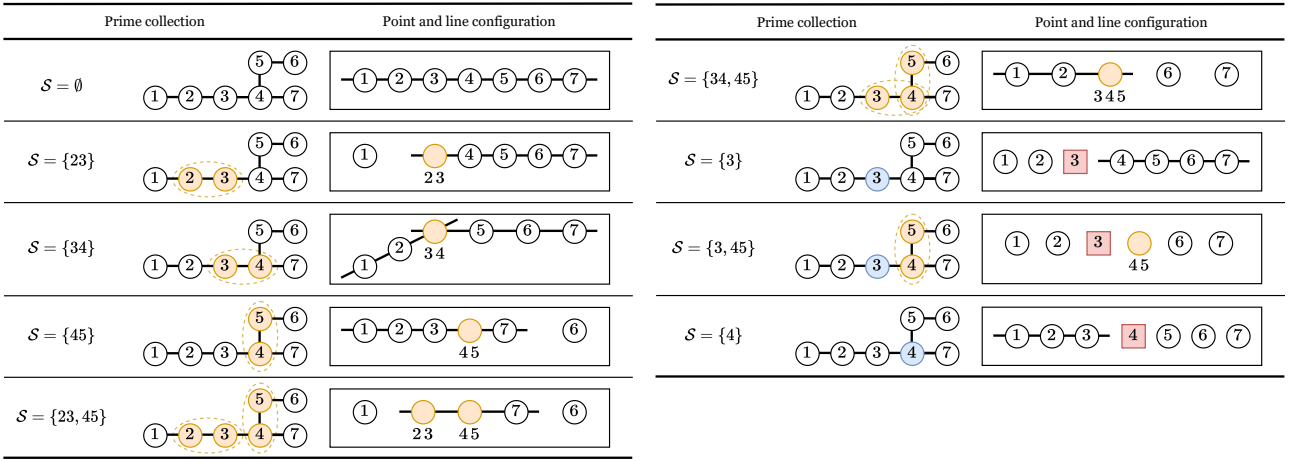
Figure 9: Prime collections $\mathcal{S}$ and their matroid $M_{\mathcal{S}}$ as a point and line configuration from Example 6.12. The dotted pairs of vertices are the two subsets of $\mathcal{S}$ and the square points in the configurations represent loops.

Algorithm 1: Prime collection test

Input: $\mathcal{S} = (s_1, \ldots, s_t)$ an ordered list of vertices of G

Output: pass or **fail**. (If fail then $\mathcal{S}$ is not a prime collection of singletons for G.)

Initialize: $H \leftarrow G$

for $i \leftarrow 1$ **to** t **do**
    if s_i *is a leaf or isolated vertex of H* **then**
        ⌊ **return fail**
    if s_i *is adjacent to a leaf of H* **and** s_i *has degree 2 in H* **then**
        ⌊ **return fail**
    $H \leftarrow$ induced subgraph of H obtained by deleting vertex s_i

return pass

Proposition 6.14. *Let $\mathcal{S} = \{s_1, \ldots, s_t\}$ be a collection of singleton sets of vertices of G. Then $\mathcal{S}$ is a prime collection of singletons for G if and only if $\mathcal{S}$ passes Algorithm 1 for any ordering of its elements.*

Proof. Suppose that $\sigma \in S_t$ is a permutation such that the algorithm fails with the input $(s_{\sigma(1)}, \ldots, s_{\sigma(t)})$. In particular, assume that the algorithm fails at step $p \in \{1, \ldots, t\}$. Then it follows that $\{s_{\sigma(1)}, \ldots, s_{\sigma(p)}\}$ is not a prime collection of singletons. So, by Definition 6.3, the set $\mathcal{S}$ is not a prime collection of singletons. Conversely, if $\mathcal{S}$ is not a prime collection of singletons then, by definition, there exists $p \in \{1, \ldots, t\}$ and $j \in \{1, \ldots, p\}$ such that $\{s_1, \ldots, s_{p-1}\}$ is a prime collection of singletons and either: s_j is a leaf or isolated vertex in G', or s_j is adjacent to a leaf of G' and has degree two, where G' is the induced subgraph of G obtained by deleting the vertices $\{s_1, \ldots, s_{j-1}, s_{j+1}, \ldots, s_p\}$. Now we order the elements of $\mathcal{S}$ as: $(s_1, \ldots, s_{j-1}, s_{j+1}, \ldots, s_p, s_j, s_{p+1}, \ldots, s_t)$. By construction, we have that Algorithm 1 fails in the for-loop when checking vertex s_j. $\square$

Lemma 6.15. *Let M be a minimal matroid for Δ_G. Then the loops of M are a prime collection of singletons for G.*

Proof. We follow the algorithmic description of the construction of a prime collection of singletons. Let $\mathcal{S}$ be the loops of M. Suppose for a contradiction that $\mathcal{S}$ fails to be a prime collection of singletons for G. Then there exists an ordering of the loops, say $\mathcal{S} = (s_1, s_2, \ldots, s_p)$, such that at step i we are unable to choose s_i for the prime collection. Let G' be the induced subgraph of G obtained by removing vertices $s_1, \ldots, s_{i-1}$. In order for us to be unable to choose s_i for the prime collection there are two cases:

1. s_i is a leaf or isolated vertex of G',

2. s_i is adjacent to a leaf and has degree two in G'.

Case 1. We construct a new matroid M' on $[n]$ as follows. First construct the matroid $M \backslash s_i$ by deleting s_i.

If s_i is a leaf of G', denote by x the unique vertex adjacent to s_i. Then, let M' be the matroid obtained from $M \backslash s_i$ by adding to the ground set the element s_i in the same parallel class as x.

If s_i is an isolated vertex of G', note that every neighbor of s_i in G is a loop of M. Let M' be the matroid obtained from $M \backslash s_i$ by freely adding s_i to the ground set of $M \backslash s_i$.

In either case, note that the collection of dependent sets of M' contains Δ_G because every path with three vertices containing s_i contains either x or some loop of M'.

We now show that $M' \preceq M$. Take any dependent set D of M'. If $s_i \in D$, then clearly D is a dependent set of M since s_i is a loop of M. On the other hand, if $s_i \notin D$, then D is a dependent set of $M' \backslash s_i = M \backslash s_i$. Hence D is a dependent set of M. To show that $M \neq M'$ we simply note that s_i is a loop in M but is not a loop in M'. So we have shown $\Delta_G \subseteq \mathcal{D}(M') \subsetneq \mathcal{D}(M)$. Hence M is not a minimal matroid for Δ_G, a contradiction.

Case 2. Let s_i be a vertex of degree two which is adjacent to a leaf of G'. Let x and y be the vertices adjacent to s_i, with x being a leaf. We construct a new matroid M' by first deleting s_i to form $M \backslash s_i$. Then we define M' to be the matroid obtained by adding s_i to the same parallel class as y in $M \backslash s_i$. Note that $\mathcal{D}(M')$ must contain Δ_G because any path with three vertices containing s_i (and not containing any loop of M') must also contain y. To show that $M' \preceq M$ we use the same argument as Case 1, so we have that $\Delta_G \subseteq \mathcal{D}(M') \subsetneq \mathcal{D}(M)$. Therefore M is not a minimal matroid for Δ_G, a contradiction. $\qquad\square$

We now characterize matroids which are minimal among all the matroids whose circuits of size one and two are exactly the members of $\mathcal{S}$. One can think of this as a constrained notion of minimality. In the following lemma, note that M is fixed, and $\mathcal{S}$ is obtained from M. In a matroid, every circuit of size two is disjoint from every circuit of size one. Therefore, $\mathcal{S}$ is a valid collection for G, and $M_{\mathcal{S}}$ is well-defined.

Lemma 6.16 (Constrained minimal matroids). *Let M be a matroid with $\Delta_G \subseteq \mathcal{D}(M)$ whose circuits of size one and two are the members of $\mathcal{S}$. Then $M_{\mathcal{S}} \preceq M$.*

Proof. By assumption, all of the circuits of $M_{\mathcal{S}}$ of size one and two are also circuits of M. All 4-subsets of $[n]$ are dependent in both M and $M_{\mathcal{S}}$. Therefore, to show that $M_{\mathcal{S}} \preceq M$, it suffices to show that all 3-circuits of $M_{\mathcal{S}}$ are dependent in M. Let $D = \{a, b, c\}$ be such a circuit of $M_{\mathcal{S}}$. Let G' be the induced subgraph of G obtained by removing those vertices v for which $\{v\} \in \mathcal{S}$. Now since D is a circuit in $M_{\mathcal{S}}$, we have that for any pair of elements $x, y \in D$, there is a path in G' from x to y and it is not blocked by $\mathcal{S}$. We denote by $\mathrm{Conv}(x, y)$ the collection of vertices on the path from x to y. Similarly we define the convex hull of a set of vertices:

$$\mathrm{Conv}(A) = \bigcup_{x,y \in A} \mathrm{Conv}(x, y).$$

By abuse of notation, we identify $\mathrm{Conv}(A)$ with the induced subgraph of G' whose vertices are $\mathrm{Conv}(A)$.

To complete the proof of the theorem, it suffices to prove the following claim. Since, as a result we have that $\{a, b, c\}$ is a dependent set in M. Therefore, every dependent set in $M_{\mathcal{S}}$ is also dependent in M, as desired.

Claim. *All 3-subsets of $\mathrm{Conv}(a, b, c)$ are dependent in M.*

We must show that every 3-subset $\{a', b', c'\} \subseteq \mathrm{Conv}(a, b, c)$ is dependent in M. We proceed by induction on $|\mathrm{Conv}(a', b', c')|$. For the base case, suppose that $|\mathrm{Conv}(a', b', c')| = 3$. Then $\mathrm{Conv}(a', b', c')$ is a path in G with three vertices. By definition of Δ_G, we have that $\{a', b', c'\} \in \Delta_G$. Since $\Delta_G \subseteq \mathcal{D}(M)$, we have that $\{a', b', c'\}$ is dependent in M. For the inductive step, assume that all 3-subsets of $\mathrm{Conv}(a, b, c)$ whose convex hull has size at most $k \geq 3$ are dependent and $|\mathrm{Conv}(a', b', c')| = k + 1$. Since $|\mathrm{Conv}(a', b', c')| \geq 4$, there is a vertex $d \in \mathrm{Conv}(a', b', c') \backslash \{a', b', c'\}$. Note that neither a', b', c', nor d can be a loop of M because then it would be a loop of $M_{\mathcal{S}}$. This implies either that $\{a, b, c\}$ is blocked or that an element of $\{a, b, c\}$ is a loop of $M_{\mathcal{S}}$. This is impossible since $\{a, b, c\}$ is a circuit of $M_{\mathcal{S}}$.

By induction, $\{a', b', d\}$, $\{a', c', d\}$, and $\{b', c', d\}$, are all dependent since their respective convex hulls lie strictly inside $\mathrm{Conv}(a', b', c')$. Since $\{a', b', d\}$ and $\{a', c', d\}$ are dependent, Lemma 6.13 implies that either $\{a', b', c'\}$ or $\{a', d\}$ is dependent. Similarly, if $\{a', b', c'\}$ is not dependent, then $\{b', d\}$ and $\{c', d\}$ are dependent. Therefore, since d is not a loop, we may assume that the rank of $\{a', b', c', d\}$ is at most one, implying that $\{a', b', c'\}$ is dependent. This completes the proof of the claim. $\qquad\square$

Lemma 6.17. *If M is a minimal matroid for Δ_G, then $M = M_{\mathcal{S}}$ for some prime collection $\mathcal{S}$.*

Proof. Let $\mathcal{S}$ be the collection of all circuits of M of size one and two. Since $\Delta_G \subseteq \mathcal{D}(M)$, Lemma 6.16 implies that $M_{\mathcal{S}} \preceq M$. However, M is a minimal matroid. Since $\Delta_G \subseteq \mathcal{D}(M_{\mathcal{S}})$, the minimality of M implies that $M = M_{\mathcal{S}}$. So it suffices to show that $\mathcal{S}$ is a prime collection.

By Lemma 6.15, we have that the loops of M form a prime collection of singletons for G. The collection of 2-element circuits of a matroid does not intersect the set of loops, so $\mathcal{S}$ is a valid collection of subsets.

Let G' be the induced subgraph of G obtained by deleting the vertices that are singleton sets in $\mathcal{S}$. Suppose for a contradiction that $\mathcal{S}$ is not a prime collection. Then it follows that there exists a vertex v in G' such that, for every vertex w adjacent to v in G', we have that $\{v, w\}$ is a member of $\mathcal{S}$.

Let w be a vertex adjacent to v, and consider the valid collection $\mathcal{S}' = \mathcal{S} \backslash \{\{v, w\}\}$. A set is a singleton in $\mathcal{S}$ if and only if it is a singleton in $\mathcal{S}'$. Since every vertex adjacent to v is in the same cloud of $\mathcal{S}$ as v, a 3-subset of $[n]$ is blocked by $\mathcal{S}$ if and only if it is blocked by $\mathcal{S}'$. We also have $\Delta_G \subseteq \mathcal{D}(M_{\mathcal{S}'})$.

Since the collection of 2-subsets in $\mathcal{S}'$ is properly contained in the the collection of 2-subsets in $\mathcal{S}$, we have that $M_{\mathcal{S}'} \preceq M_{\mathcal{S}}$, a contradiction. Therefore, we conclude that $\mathcal{S}$ is a prime collection. $\qquad\square$

The last ingredient in the proof of Theorem 6.11 is to show that $M_{\mathcal{S}}$ is minimal for any prime collection $\mathcal{S}$.

Lemma 6.18. *If $\mathcal{S}$ and $\mathcal{T}$ are prime collections with $M_{\mathcal{S}} \geq M_{\mathcal{T}}$, then $\mathcal{S} = \mathcal{T}$.*

Proof. Let us begin by considering the loops of $M_{\mathcal{S}}$ and $M_{\mathcal{T}}$. Since $M_{\mathcal{S}} \geq M_{\mathcal{T}}$, we have that every loop of $M_{\mathcal{T}}$ is a loop of $M_{\mathcal{S}}$. Suppose for a contradiction that $M_{\mathcal{S}}$ has strictly more loops than $M_{\mathcal{T}}$. Let $G_{\mathcal{S}}$ be the induced subgraph of G obtained by removing the loops of $M_{\mathcal{S}}$. We define $G_{\mathcal{T}}$ similarly. Pick any loop v in $M_{\mathcal{S}}$ that is not a loop in $M_{\mathcal{T}}$. We now show that $\deg(v) \geq 2$ in $G_{\mathcal{T}}$. Suppose otherwise. Then all but at most one of the neighbors of v are loops in $M_{\mathcal{T}}$. Since $M_{\mathcal{S}} \geq M_{\mathcal{T}}$, all but at most one of the neighbors of v are loops in $M_{\mathcal{S}}$ also. But this fact, with the assumption that v is a loop in $M_{\mathcal{S}}$, contradicts the assumption that $\mathcal{S}$ is a prime collection. Therefore, $\deg(v) \geq 2$ in $G_{\mathcal{T}}$. Similar reasoning shows that each vertex adjacent to v in $G_{\mathcal{T}}$ must be adjacent to a vertex in $G_{\mathcal{T}}$ that is not a loop in $M_{\mathcal{S}}$.

Since $\mathcal{T}$ is a prime collection, there is at least one vertex x adjacent to v in $G_{\mathcal{T}}$ such that $\{x, v\} \notin \mathcal{T}$. Since $\deg(v) \geq 2$ in $G_{\mathcal{T}}$, let $y \neq x$ be another neighbor of v in $G_{\mathcal{T}}$. We have already observed that there must be a vertex adjacent to x in $G_{\mathcal{T}}$ that is not a loop in $M_{\mathcal{S}}$. Call this vertex x'. We may choose x' so that $\{x, x'\} \notin \mathcal{S}$ because, if there is no such vertex, then x has no neighbor outside the cloud containing x in $G_{\mathcal{S}}$, which implies that $\mathcal{S}$ is not a prime collection. Similarly, we let y' be any vertex adjacent to y such that y is not a loop in $M_{\mathcal{S}}$ and such that $\{y, y'\} \notin \mathcal{S}$. For $X \subseteq [n]$, let $\mathrm{cl}(X)$ denote the closure of X in $M_{\mathcal{T}}$. Since $\Delta_G \subseteq \mathcal{D}(M_{\mathcal{T}})$, we have that $\{x, x', v\}$ and $\{x, y, v\}$ are both dependent sets in $M_{\mathcal{T}}$. By assumption, $\{x, x'\}$ is independent in $M_{\mathcal{S}}$, implying that it is independent in $M_{\mathcal{T}}$ also. Since $\{x, x', v\}$ is dependent, we have $v \in \mathrm{cl}(\{x, x'\})$. Thus $\mathrm{cl}(\{x, x'\}) = \mathrm{cl}(\{x, x', v\})$. Similarly, since $\{x, v\}$ is independent in $M_{\mathcal{T}}$ but $\{x, y, v\}$ is dependent, we have that $y \in \mathrm{cl}(\{x, v\}) \subseteq \mathrm{cl}(\{x, x', v\}) = \mathrm{cl}(\{x, x'\})$. Since $y \in \mathrm{cl}(\{x, x'\})$, we have that $\{x, x', y\}$ is a dependent set in $M_{\mathcal{T}}$. However, by assumption, we have that $\{x, x'\}$ is independent in $M_{\mathcal{S}}$. Hence $\{x, x', y\}$ is independent in $M_{\mathcal{S}}$ because it is blocked by v. So we have found a dependent set in $M_{\mathcal{T}}$ which is independent in $M_{\mathcal{S}}$, a contradiction. Therefore, we deduce that $M_{\mathcal{S}}$ and $M_{\mathcal{T}}$ have the same loops.

To show that $\mathcal{S} = \mathcal{T}$, it remains to show that the 2-subsets in $\mathcal{S}$ are the same as the 2-subsets in $\mathcal{T}$. Since $M_{\mathcal{S}}$ and $M_{\mathcal{T}}$ have the same loops and $M_{\mathcal{S}} \geq M_{\mathcal{T}}$, it follows that every 2-subset in $\mathcal{T}$ is also in $\mathcal{S}$. Suppose for a contradiction that $\mathcal{S}$ has strictly more 2-subsets than $\mathcal{T}$. Therefore, either $\mathcal{S}$ has more clouds than $\mathcal{T}$, or there is a cloud of $\mathcal{T}$ that is properly contained in a cloud of $\mathcal{S}$. In either case, since every vertex in the cloud is adjacent to a vertex not in the cloud, there is a 3-element set that is blocked by $\mathcal{S}$ but not by $\mathcal{T}$. This contradicts the assumption that $M_{\mathcal{S}} \geq M_{\mathcal{T}}$. $\square$

6.2 Irreducible matroid varieties. In this subsection, we combine the results of §6.1 to obtain a classification of the irreducible components of V_{Δ_G}. In particular, we will prove Theorem 6.20. This gives a straightforward description of the irreducible components.

Proposition 6.19. *Let G be a forest and $\mathcal{S}$ be a prime collection for G as in Definition 6.4. Then for any non-central component V_N of $V_{M_{\mathcal{S}}}^{\mathrm{comb}}$ there exists another prime collection $\mathcal{S}'$ for G, such that $V_N \subseteq V_{M_{\mathcal{S}'}}$.*

Proof. Let J be the ideal associated to the matroid N and for notation we write N as M_J. By Proposition 6.10, the simplification of $M_{\mathcal{S}}$ is forest-like. So, by Theorem 5.14,

the matroid M_J is obtained from $M_{\mathcal{S}}$ by setting all elements inside certain clouds of $\mathcal{S}$ to be loops, see Definition 5.12. Denote these clouds by $C_1, \ldots, C_s$. More precisely, the circuits of M_J are the singleton subsets of $C_1 \cup \cdots \cup C_s$ along with the collection of circuits of $M_{\mathcal{S}} \backslash (C_1 \cup \cdots \cup C_s)$. Note that $M_{\mathcal{S}} < M_J$.

Let $(\mathcal{P}, \mathcal{L}, \mathcal{I})$ be the point and line configuration whose matroid is the simplification of $M_{\mathcal{S}}$. By Theorem 5.11, if a non-central component J of $I_{M_{\mathcal{S}}}^{\mathrm{comb}}$ exists, there is a point in $\mathcal{P}$ contained in at least three lines in $\mathcal{L}$. This implies that a cloud of $\mathcal{S}$ contains at least three elements.

Let C be a cloud of $\mathcal{S}$ that is not one of $C_1, \ldots C_s$. Let $p_i \in \mathcal{P}$ be the corresponding point in the simplification of $M_{\mathcal{S}}$. We define the *star of* p_i to be $S(p_i) = \bigcup_{\ell \in \mathcal{L}_i} \ell$, where ℓ is identified with the set of points it passes through. By Lemma 5.13, we see that all generators of $I_{M_{\mathcal{S}}}$ constructed using points in $S(p_i)$ are also generators of J. However, we also see that $I_{M_{\mathcal{S}}} \nsubseteq J$ because the elements of $C_1 \cup \cdots \cup C_s$ are loops of M_J. Therefore, it is enough to show that there exists a prime collection $\mathcal{S}'$ such that:

- $M_{\mathcal{S}'} < M_J$ and

- every cloud of $\mathcal{S}'$ contained in any of $C_1, \ldots, C_s$ has size at most two.

We will construct $\mathcal{S}'$ from $\mathcal{S}$ by an inductive procedure modifying the clouds $C_1, \ldots, C_s$. Let $\mathcal{S}_0 = \mathcal{S}$. At each step of the procedure, we construct $\mathcal{S}_{i+1}$ from $\mathcal{S}_i$ by modifying a cloud C of $\mathcal{S}_i$ with $|C| \geq 3$ such that $C \subseteq C_1 \cup \cdots \cup C_s$. By induction, we assume $\mathcal{S}_i$ is a prime collection. At each step of this procedure, we construct a prime collection $\mathcal{S}_{i+1}$ such that every set that is dependent in $M_{\mathcal{S}_{i+1}}$ but not in $M_{\mathcal{S}_i}$ is dependent in M_J. Since $M_{\mathcal{S}} < M_J$, this implies that $M_{\mathcal{S}'} < M_J$.

Let G' be the induced subgraph of G obtained by deleting the singletons in $\mathcal{S}_i$. Let $G[C]$ be the induced subgraph of G with vertex set C, and let ℓ be a leaf of $G[C]$. Since $|C| \geq 3$, the neighbor x of ℓ in $G[C]$ is not

a leaf of $G[C]$. In particular, $\deg(x) \geq 3$ in G', so we can add x to the collection of singletons of $\mathcal{S}_i$ to result in a new prime collection of singletons. Therefore, we modify $\mathcal{S}_i$ as follows. Make x a singleton set in $\mathcal{S}_{i+1}$ and remove all 2-subsets in $\mathcal{S}_i$ which contain x. We now consider the neighbors of x which lie outside C. For each such neighbor y of x, if y belongs to a cloud $C' \neq C$ and x is the only neighbor of y not contained in C', then we remove y from the cloud C'. This ensures that the new collection is a prime collection. Let $\mathcal{S}_{i+1}$ be this new prime collection. Note that this procedure does not add any new 3-circuits because all paths crossing C' that contain y necessarily pass through x.

As a result of this procedure, the cloud C no longer contains x or ℓ and might also be split into multiple smaller clouds. Note that all other clouds affected by this procedure have decreased in size. Thus, the only subset of $[n]$ that is dependent in $\mathcal{S}_{i+1}$ but not in $\mathcal{S}_i$ is the singleton set $\{x\}$, which is a loop of M_J. Therefore, $M_{\mathcal{S}_{i+1}} < M_J$ as desired. We apply this procedure inductively until we obtain a prime collection $\mathcal{S}'$ such that every cloud of $\mathcal{S}'$ contained in $C_1 \cup \cdots \cup C_s$ has size 2. $\qquad\square$

Theorem 6.20. *Let G be a forest and let Δ_G be the corresponding consecutive forest hypergraph. Then,*

$$V_{\Delta_G} = \bigcup_{\mathcal{S}} V_{M_\mathcal{S}},$$

is an irredundant irreducible decomposition of V_{Δ_G}, where the union is taken over all prime collections $\mathcal{S}$ of G. In particular, a minimal prime decomposition of $\sqrt{I_{\Delta_G}}$ is given as $\sqrt{I_{\Delta_G}} = \bigcap_{\mathcal{S}} I_{M_\mathcal{S}}$.

Proof. By Theorem 6.11, the set of all minimal matroids of Δ_G is $\{M_\mathcal{S} : \mathcal{S}$ is a prime collection$\}$. It follows, by Proposition 3.10, that $V_{\Delta_G} = \bigcup_{\mathcal{S}} V_{M_\mathcal{S}}^{\text{comb}}$. It is easy to see that, the simplification of the matroid $M_\mathcal{S}$ is a forest-like configuration. So, the irreducible components of $V_{M_\mathcal{S}}^{\text{comb}}$ are in one-to-one correspondence with the irreducible components of the combinatorial closure of the matroid variety of this configuration, which are given by Theorem 5.14. By Proposition 6.19, for every non-central component V_{M_J} of $V_{M_\mathcal{S}}^{\text{comb}}$, there exists a prime collection $\mathcal{S}'$ such that $V_{M_J} \subseteq V_{M_{\mathcal{S}'}}$. Moreover, by Theorem 6.22, each of the matroids $M_\mathcal{S}$ is realizable. Hence, the irredundant irreducible decomposition of V_{Δ_G} is given by $V_{\Delta_G} = \bigcup_{\mathcal{S}} V_{M_\mathcal{S}}$, as desired. $\qquad\square$

In the example below we show that the irreducible components of V_{Δ_G} are not necessarily positroid varieties.

Example 6.21. Let G be the graph on the vertex set $\{1, 2, \ldots, 9\}$ and edge set $\{12, 23, 14, 45, 56, 47, 78, 89\}$. Let $\mathcal{S} = \{14, 47\}$ be a prime collection. By Remark 6.9, the matroid $M_\mathcal{S}$ has three rank 2 flats given by 12347, 14567 and 14789. So the simplification of $M_\mathcal{S}$ is the matroid of the configuration given by three lines passing through a point. See Example 3.6. In particular, the ideal of $M_\mathcal{S}$ is not generated by determinants. So by [KLS13, Theorem 5.15], the matroid $M_\mathcal{S}$ is not a positroid.

6.3 Realizability.
We are now ready to consider the minimal matroids for Δ_G. Recall that, if $\mathcal{S}$ is a valid collection for G, then the matroid $M_\mathcal{S}$ arising from Proposition 6.8 is minimal with respect to Δ_G. Here, we show that $M_\mathcal{S}$ is realizable over the real numbers hence realizable over $\mathbb{C}$.

Theorem 6.22. *If $\mathcal{S}$ is a valid collection, then $M_\mathcal{S}$ is $\mathbb{R}$-realizable.*

Proof. Let r be the rank function of $M_\mathcal{S}$. The elements of $M_\mathcal{S}$ are the vertices of the forest G. Let L be the set of loops of $M_\mathcal{S}$, so L consists of the singletons in $\mathcal{S}$. We will proceed inductively by considering subgraphs of G with increasing numbers of vertices. For the base case, we begin with the set L of loops of M. The restriction $M_\mathcal{S}|L$ can be represented by the $3 \times |L|$ zero matrix.

For the inductive step, suppose that we have shown that $M_\mathcal{S}|V(G_0)$ is realizable over $\mathbb{R}$ for some subgraph G_0 of G containing the vertices in L. Let $v \in V(G) \backslash V(G_0)$. We will show that the restriction $M_\mathcal{S}|(V(G_0) \cup \{v\})$ is obtained from $M_\mathcal{S}|(V(G_0))$ either by adding v as a coloop or by freely adding v to a flat. Thus, Lemma 5.5 or Lemma 5.6 implies that $M_\mathcal{S}|(V(G_0) \cup \{v\})$ is realizable over the reals. We consider the following cases.

(1) There is a vertex $w \in V(G_0)$ such that v and w are in the same cloud.

(2) Case (1) does not hold, but there are vertices $w, x \in V(G_0) \backslash L$, not in the same cloud, such that $\{v, w, x\}$ is an unblocked set.

(3) Neither Case (1) nor Case (2) holds.

For Case (1), note that $C \cup L$ is a flat of rank 1 for every cloud C. Therefore, if C is the cloud containing v and w, then we obtain $M_\mathcal{S}|(V(G_0) \cup \{v\})$ by freely adding v to the flat $C \cup L$.

In Case (2), the fact that w and x are not in the same cloud and the fact that neither w nor x is in L imply that $r(\{w, x\}) = 2$. The fact that $\{v, w, x\}$ is an unblocked set implies that $r(\{v, w, x\}) = 2$ also. Thus, v is in the closure of $\{w, x\}$ in the matroid $M_\mathcal{S}$. Since Case (1) does not hold, $M_\mathcal{S}|(V(G_0) \cup \{v\})$ is obtained by adding v freely to the closure of $\{w, x\}$.

Now we consider Case (3). Since Case (1) does not hold, Proposition 6.8 implies that $M_{\mathcal{S}}$ has no circuit of size 2 consisting of v and some vertex of G_0. Since Case (2) does not hold, Proposition 6.8 implies that $M_{\mathcal{S}}$ has no circuit of size 3 consisting of v and two vertices of G_0. Therefore, the only circuits of $M_{\mathcal{S}}$ containing v and some vertex of G_0 are sets of size 4. Thus, the matroid $M_{\mathcal{S}}|(V(G_0) \cup \{v\})$ is obtained either by freely adding v to the ground set $V(G_0)$ (if $r(V(G_0)) = 3$) or by adding v as a coloop (if $r(V(G_0)) < 3$). □

7 Hypergraphs $\Delta^{s,t}$

Here, we study hypergraph ideals which arise naturally in the study of conditional independence statements. We begin with notation and the general setup of the problem. We denote the $k \times \ell$ matrix of integers

$$\mathcal{Y} = (\mathcal{Y}_{i,j})_{i,j} = \begin{bmatrix} 1 & k+1 & \dots & (\ell-1)k+1 \\ 2 & k+2 & \dots & (\ell-1)k+2 \\ \vdots & \vdots & \ddots & \vdots \\ k & 2k & \dots & \ell k \end{bmatrix}. \tag{3}$$

For each $i \in [k]$ and $j \in [\ell]$, the rows and columns of $\mathcal{Y}$ are denoted

$$R_i = \{\mathcal{Y}_{i,1}, \mathcal{Y}_{i,2}, \dots, \mathcal{Y}_{i,\ell}\} = \{i, k+i, \dots, (\ell-1)k+i\},$$

$$C_j = \{\mathcal{Y}_{1,j}, \mathcal{Y}_{2,j}, \dots, \mathcal{Y}_{k,j}\} = \{(j-1)k+1, (j-1)k+2, \dots, (j-1)k+k\}.$$

For each s and t with $s \leq k$ and $t \leq \ell$, we define $\Delta^{s,t}$ be the following collection of subsets of $[k\ell]$,

$$\Delta^{s,t} = \bigcup_{1 \leq i \leq k} \binom{R_i}{t} \cup \bigcup_{1 \leq j \leq \ell} \binom{C_j}{s}.$$

If the values of s and t have been fixed, then we simply write Δ for $\Delta^{s,t}$. We are interested in studying the ideals $I_{\Delta^{s,t}}$ for $d \geq \max\{s,t\}$ as these are examples of conditional independence ideals with hidden variables.

Example 7.1. Let $k = 4, \ell = 7, s = 2$ and $t = 3$. We have

$$\mathcal{Y} = \begin{bmatrix} 1 & 5 & 9 & 13 & 17 & 21 & 25 \\ 2 & 6 & 10 & 14 & 18 & 22 & 26 \\ 3 & 7 & 11 & 15 & 19 & 23 & 27 \\ 4 & 8 & 12 & 16 & 20 & 24 & 28 \end{bmatrix} \quad \text{and} \quad \Delta^{2,3} = \left\{ \binom{\{1,5,9,13,17,21,25\}}{3} \cup \dots \cup \binom{\{25,26,27,28\}}{2} \right\}.$$

Calculating the dependent matroids for $\Delta^{2,3}$, we find that all such matroids are point and line configurations. In Table 1, there are 10 combinatorial types of configurations which appear as these matroids. Explicitly, these are the point and line configurations which have at most 4 lines and at most 7 points.

k \\ ℓ	3	4	5	6	7	8	9	10	11	12	$\cdots$	$\cdots$	∞
2	2	2	3	**4**	4	4	4	4	4	4	4	4	4
3	2	2	3	5	7	8	**9**	9	9	9	9	9	9
4	2	2	3	6	10	13	20	23	24	**25**	25	25	25

Table 1: The number of *combinatorial types* of configurations appearing among dependent matroids for $\Delta^{2,3}$. Note that increasing ℓ leads to having arbitrarily many irreducible components for the variety $V_{\Delta^{2,3}}$.

Remark 7.2. The minimal prime decomposition of $I_{\Delta^{s,t}}$ has been extensively studied in [HHH+10, Rau13, CMR20, PS19]. In each case, we find a matroidal description of the prime components. In Table 2 we give a unified perspective on these results where the minimally dependent matroids can be uniquely identified by their loops. In the case $k = s = 2$, it is possible to generalize the prime components of $I_{\Delta^{s,t}}$, described in [CMM21] for the $t = 3$, to all $t \geq 3$. A complete description of this can be found in our forthcoming work. In this case the minimally dependent matroids are given by configurations of points lying in $(t-1)$-dimensional affine subspaces. When $s = 2$, $t = 3$ and k, ℓ are arbitrary, the dependent matroids for $\Delta^{s,t}$ are given by configurations with at most k points and ℓ lines. The specific case with $d = k = s = t = 3$ and $\ell = 4$ is given in Example 3.12. The ideal $I_{\Delta^{s,t}}$ has two components both of which are matroid varieties corresponding to the configurations in Figure 2.

Example 7.3. Let $s = 2$ and $t = 3$. For each $2 \leq k \leq 4$ and $3 \leq \ell$ we calculate the number of possible combinatorial types of points and line configurations which appear among dependent matroids for $\Delta^{s,t}$. These calculations are displayed in Table 1. Note that with 12 points, we observe all combinatorial types of configurations with at most 4 lines. So, in order to determine the components of the hypergraph variety $V_{\Delta^{s,t}}$, we need only show that a finite number of point and line configurations have irreducible varieties.

(k,ℓ,s,t)	Minimally dependent matroids $M_{\mathcal{S}}$ for $\Delta^{s,t}$	Simplification
$(k,\ell,2,2)$	The parallel classes of $M_{\mathcal{S}}$ are the connected components of $\Delta^{s,t}\backslash\mathcal{S}$. $M_{\mathcal{S}}$ is minimally dependent if for any $\mathcal{T}\subseteq\mathcal{S}$, there exists a parallel class of $M_{\mathcal{T}}$ which contains at least two distinct parallel classes of $M_{\mathcal{S}}$.	$M_{\mathcal{S}}$: Uniform matroid on the ground set of parallel classes of $M_{\mathcal{S}}$.
$(k,\ell,2,\ell)$	$\mathcal{S}=\emptyset$: $M_{\emptyset}$ is the uniform matroid on $[k\ell]$ of rank $\ell-1$. $\mathcal{S}\neq\emptyset$: $M_{\mathcal{S}}$ is minimally dependent if $|\mathcal{S}\cap R_i|=1$ for each row R_i and $\mathcal{S}\cap C_j\neq\emptyset$ for at least two distinct columns C_j of $\mathcal{Y}$. The parallel classes of $M_{\mathcal{S}}$ are $C_j\backslash\mathcal{S}$ for each column C_j.	$M_{\emptyset}$: Already simplified. $M_{\mathcal{S}}$: Uniform matroid on the ground set of parallel classes of $M_{\mathcal{S}}$.
$(2,\ell,2,3)$	$\mathcal{S}$: a *minimal set* from [CMM21, Definition 3.15]. $M_{\mathcal{S}}$ is the matroid of a point and line configuration. The parallel classes of $M_{\mathcal{S}}$ are: $C=\bigcup C_j$ where the union is taken over all columns C_j of $\mathcal{Y}$ such that $C_j\cap\mathcal{S}=\emptyset$, and the sets $C_j\backslash\mathcal{S}$ where $C_j\cap\mathcal{S}\neq\emptyset$. The circuits of size 3 in $M_{\mathcal{S}}$ are given by 3-subsets of the rows $R_1\backslash\mathcal{S}$ and $R_2\backslash\mathcal{S}$ respectively.	$M_{\mathcal{S}}$: Matroid of a point and line configuration with at most two lines intersecting at C (possibly empty set).

Table 2: The minimally dependent matroids for $\Delta^{s,t}$, which are uniquely determined by their loops $\mathcal{S}\subseteq[k\ell]$.

7.1 Grid matroids of small rank. Although the minimal matroids for $\Delta^{s,t}$ are known for some specific values of s and t, it seems difficult to determine the minimal matroids for arbitrary values of s and t. However, there is no difficulty if d is sufficiently small. We first recall the notion of affine matroids.

Definition 7.4. Let $\mathbb{F}$ be a field, and let $S=\{v_1,\ldots,v_k\}$ be a collection of (not necessarily distinct) vectors in $\mathbb{F}^{d-1}$. Let $v_i'\in\mathbb{F}^d$ be the vector whose first coordinate is 1 and whose other coordinates are those of v_i.

- The collection S is *affinely dependent* if $k>0$ and there are elements $a_1,\ldots,a_k\in\mathbb{F}$ that are not all 0 with $\sum_{i=1}^k a_i v_i=\mathbf{0}$ and $\sum_{i=1}^k a_i=0$. This is equivalent to the condition that $\{v_1',\ldots,v_k'\}$ is linearly dependent.

- A matroid M on the set $[n]$ is *affine* over the field $\mathbb{F}$ if there is a function $\phi:[n]\rightarrow\mathbb{F}^{d-1}$ such that $X\subseteq[n]$ is an independent set in M if and only if $\phi(X)$ is affinely independent. In this case, M has rank at most d and can be realized by a $d\times n$ matrix whose columns are $v_1',\ldots,v_n'$. An affine matroid must be loopless, and it is simple if and only if ϕ is injective. If F is a flat of M of rank t, then there is a $(t-1)$-dimensional affine subspace of $\mathbb{F}^{d-1}$ whose intersection with $[n]$ is $\phi(F)$.

Theorem 7.5. *Let s,t,k,ℓ,d be positive integers such that $3\leq s\leq t\leq\ell$, $s\leq k$, and $t\leq d\leq s+t-3$. Then $C=\min(\Delta^{s,t}\cup\binom{[k\ell]}{d+1})$ is the collection of circuits of an $\mathbb{R}$-realizable matroid on $[k\ell]$ of rank d. This is the unique minimal matroid for $\Delta^{s,t}$ in this case.*

Proof. One can check fairly easily that C satisfies circuit elimination and therefore is the collection of circuits of a matroid. However, this will be unnecessary because we will prove this theorem by constructing a realization for an affine matroid over $\mathbb{R}$ whose circuits are C. Because C is itself the collection of circuits of a matroid M, there can be no matroid $N\neq M$ with rank at most d and ground set $[k\ell]$ such that $\Delta^{s,t}\subseteq\mathcal{D}(N)\subseteq\mathcal{D}(M)$. Thus, M is the unique minimal matroid for $\Delta^{s,t}$ with $d\leq s+t-3$.

To construct a realization for an affine matroid over $\mathbb{R}$ whose circuits are C, we must define a function $\phi:[k\ell]\rightarrow\mathbb{R}^{d-1}$. In such a matroid, each R_i must be a flat of rank $t-1$ and each C_j must be a flat of rank $s-1$. A flat of rank $d-m$ in an affine matroid over $\mathbb{R}$ corresponds to an affine subspace of $\mathbb{R}^{d-1}$ of dimension $d-m-1$. For each such subspace, there are m distinct affine hyperplanes in $\mathbb{R}^{d-1}$ whose intersection is the subspace. Therefore, each $\phi(R_i)$ must be defined by the intersection of a collection $\mathcal{H}_{R_i}$ of $d-t+1$ affine hyperplanes. These hyperplanes are defined by the following equations, with the matrix $A_i=[a_{p,q}]$ having full row rank.

$$
\begin{array}{ccccccc}
a_{1,1}x_1 & + & \cdots & + & a_{1,d-1}x_{d-1} & = & c_1 \\
& & & & & \vdots & \\
a_{d-t+1,1}x_1 & + & \cdots & + & a_{d-t+1,d-t+1}x_{d-1} & = & c_{d-t+1}
\end{array}
$$

We choose k such collections of hyperplanes $\mathcal{H}_{R_1},\mathcal{H}_{R_2},\ldots,\mathcal{H}_{R_k}$ such that the rank-$(t-1)$ subspaces they define are in "general position". That is, for every $n<t$, the intersection of every collection of n such subspaces is a subspace of rank $t-n$. Similarly, each $\phi(C_j)$ must be defined by the intersection of a collection $\mathcal{H}_{C_j}$ of $d-s+1$ hyperplanes defined by the following equations, with the matrix $B_j=[b_{p,q}]$ having full row rank.

$$
\begin{array}{ccccccc}
b_{1,1}x_1 & + & \cdots & + & b_{1,d-1}x_{d-1} & = & c_{d-t+2} \\
& & & & & \vdots & \\
b_{d-s+1,1}x_1 & + & \cdots & + & b_{d-s+1,r-1}x_{d-1} & = & c_{2d-s-t+2}
\end{array}
$$

Choosing the subspaces defined by $\mathcal{H}_{R_1}, \mathcal{H}_{R_2}, \ldots, \mathcal{H}_{R_k}, \mathcal{H}_{C_1}, \mathcal{H}_{C_2}, \ldots, \mathcal{H}_{C_\ell}$ will ensure that every subset of $\phi([k\ell])$ that should be affinely dependent is indeed so. However, we must also ensure that every subset of $\phi([k\ell])$ that should be affinely independent is so. That is, for every subset X with $|X| \le d$ such that no t-subset of any R_i or s-subset of any C_j is contained in X, $\phi(X)$ must be affinely independent. To do this, we will choose the collections $\mathcal{H}_{C_1}, \mathcal{H}_{C_2}, \ldots, \mathcal{H}_{C_\ell}$ successively. Suppose we have chosen the collections $\mathcal{H}_{C_1}, \ldots, \mathcal{H}_{C_{j-1}}$ and now must choose $\mathcal{H}_{C_j}$. If X is a subset of $C_1 \cup \cdots \cup C_j$ such that $\phi(X)$ is to be affinely independent, then $\phi(X - C_j)$ is a basis for a subspace S of $\mathbb{R}^{d-1}$. We must choose $\mathcal{H}_{C_j}$ so that, for each $i \le k$, the intersection of all hyperplanes in $\mathcal{H}_{R_i} \cup \mathcal{H}_{C_j}$ is a subspace of rank $d - (d - s + 1) - (d - t + 1) = d - (2d - s - t + 2) = s + t - d - 2$ that avoids all such subspaces S.

The requirement that this subspace has rank $s + t - d - 2$ is equivalent to the matrix $\left[\dfrac{A_i}{B_j} \right]$ having full row rank. This subspace must be nonempty since $R_i \cap C_j \ne \emptyset$. Thus, we must have $s + t - d - 2 \ge 1$, which is true because $d \le s + t - 3$. The requirement of avoiding all subspaces defined by independent sets can be achieved as there are only finitely many such subspaces to avoid but infinitely many collections to choose to be $\mathcal{H}_{C_j}$. $\square$

7.2 General grid matroids. Given a collection $\mathcal{D}$ of subsets of a ground set E, we wish to find the matroids that are minimally dependent for $\mathcal{D}$. However, for our purposes here, it is necessary to find the matroids minimally dependent for $\mathcal{D}$ among the class of realizable matroids.

Some work toward these ideas has been done by Martí-Farré [MF14]. Below, we recall the algorithmic procedure used in [MF14] to obtain minimally dependent matroids. This algorithm is not quite completely satisfactory from our perspective for two reasons. First, the algorithm does not take realizability into account at all. Martí-Farré and de Mier worked with realizability in [MFdM15], but as they said there, "The problems under consideration are far from being solved." Second, although the algorithm is guaranteed to give all of the minimally dependent matroids, it may also give additional matroids that are not minimally dependent. One must compare the matroids given by the algorithm to determine which ones are minimally dependent.

In fact, we will use the algorithm of Martí-Farré to prove a result that illustrates the difficulties of determining the minimal matroids for $\Delta^{s,t}$ in general. First, we recall some terminology and notation, much of it coming from [MF14]. A *clutter* on a set Ω is a collection of subsets of Ω such that no set is contained in another. (The term *clutter* is another word for what we have been calling a *simple hypergraph*.) If Λ is a clutter on Ω, let $\Lambda^+ = \{A \subseteq \Omega : A_0 \subseteq A \text{ for some } A_0 \in \Lambda\}$. If Υ is any collection of subsets of Ω, then recall from Section 2 that $\min(\Upsilon)$ denotes the clutter of inclusion-wise minimal sets in Υ. For $B \subseteq \Omega$, define

$$I_\Lambda(B) = \bigcap_{A \in \Lambda, A \subseteq B} A.$$

It follows from the circuit elimination axiom that Λ is the collection of circuits of a matroid if and only if $I_\Lambda(A_1 \cup A_2) = \emptyset$ for all $A_1, A_2 \in \Lambda$ with $A_1 \ne A_2$. To describe the algorithm, we use the notation of Martí-Farré and de Mier in [MFdM17]. Let Λ be a clutter with $A_1, A_2 \in \Lambda$. An α_1-transformation of Λ is

$$\alpha_1(\Lambda; A_1, A_2) = \begin{cases} \min(\Lambda \cup \{A_1 \cap A_2\}) & \text{if } I_\Lambda(A_1 \cup A_2) \ne \emptyset, \\ \Lambda & \text{otherwise.} \end{cases}$$

The α_2- and α_3-transformations of Λ are

$$\alpha_2(\Lambda) = \min(\Lambda \cup \{(A_1 \cup A_2)\backslash\{x\} : A_1, A_2 \in \Lambda, A_1 \ne A_2, x \in A_1 \cap A_2\}) \text{ and}$$

$$\alpha_3(\Lambda) = \min(\Lambda \cup \{(A_1 \cup A_2)\backslash I_\Lambda(A_1 \cup A_2) : A_1, A_2 \in \Lambda, A_1 \ne A_2\}).$$

The following result is proved in [MF14, Theorem 13].

Theorem 7.6. *Let $\Lambda \ne \{\emptyset\}$ be a clutter on a finite set Ω and let M be a minimally dependent matroid for Λ whose collection of circuits is $C(M)$. There is a sequence of clutters $\Lambda = \Lambda_0, \Lambda_1, \ldots, \Lambda_r = C(M)$ such that, for each $i \ge 1$, we have $\Lambda_{i-1}^+ \subsetneq \Lambda_i^+$ and such that Λ_i is either an α_1-, α_2-, or α_3-transformation of Λ_{i-1}.*

So, in order to obtain all minimally dependent matroids for a clutter, it suffices to perform every possible combination of α_1-, α_2-, and α_3-transformations on the clutter. Then, one must compare the resulting matroids and discard any matroid whose collection of dependent sets strictly contains the collection of dependent sets of another matroid obtained from the algorithm. Although the elements of the ground set of $\Delta^{s,t}$ are the integers $1, 2, \ldots, \ell k$, it will be convenient to think of each element as an ordered pair (i, j), where (i, j) is the unique element of $R_i \cap C_j$. In the remainder of this section, we show that every matroid can be obtained as a restriction of a matroid obtained from some $\Delta^{s,t}$ by α_1- and α_2-transformations. We prove the following result.

Theorem 7.7. *Let M be a matroid on ground set $[n]$. There are positive integers s, t, k, ℓ and a sequence of clutters $\Delta^{s,t} = \Lambda_0, \Lambda_1, \ldots, \Lambda_r$ on the set $[k\ell]$ such that, for each positive integer $i \le r$, Λ_i is either an α_1- or an α_2-transformation of Λ_{i-1} and Λ_r is the collection of circuits of a matroid with a restriction isomorphic to M.*

We will prove Theorem 7.7 after giving some definitions and proving some lemmas below. We remark that these lemmas and Theorem 7.7, as well as their proofs, are true if one replaces α_2-transformations with α_3-transformations. We will use α_1- and α_2-transformations to obtain a matroid whose restriction to $\{(1, j) : 1 \le j \le n\}$ is isomorphic to M. That isomorphism will be $j \to (1, j)$ for each $j \in [n]$. Thus, to simplify notation, we identify the element $(1, j) \in R_1$ with the element $j \in [\ell]$. (Similarly, we identify subsets of R_1 with subsets of $[\ell]$.)

Definition 7.8. Let $t \ge 3$, and let M be a matroid of rank $t - 1$ on $[n]$. Let c be the number of loops of M. Let $c' = \max\{c, t - 1\}$ and $n' = n + c' - c$. Let M' be the matroid on $[n']$ obtained from M by adding $c' - c$ loops. Let $\ell = n' + 2(t - 2)$ and M^+ be the matroid on $[\ell]$ obtained from M' by freely adding $\ell - n'$ elements to the flat $[n']$.

Note that c' is the number of loops of both M' and M^+. Also note that, since the rank of M' is $t - 1$, there are at least $t - 1$ non-loop elements of M'. Therefore, $n' \ge t - 1 + c' \ge 2(t - 1)$.

Definition 7.9. Let M be a matroid of rank $t - 1$ on ground set $[n]$, where $t \ge 3$. Let $\ell = n' + 2(t - 2)$, while $s = 3$ and $k = 5$. For a non-negative integer $p \le t - 1$, let $\Lambda_p(M)$ be the clutter on $[k\ell]$ consisting of the inclusion-wise minimal sets among

(1) 3-subsets of C_j, where $1 \le j \le \ell$,

(2) t-subsets of R_i, where $1 \le i \le k$,

(3) subsets of $[n'] \subseteq R_1$ of size at least $(t - p)$ that are circuits of M',

(4) sets of the form $A \cup \{x\} \subseteq R_1$, where $x \in \{\ell - 2p + 1, \ell - 2p + 2\}$ and A is a $(t - p - 1)$-subset of $[\ell - 2p]$ containing exactly one circuit of M', and

(5) sets of the form $H_i \cup K_j$ where $H_i \subseteq R_i \backslash C_j$ and $K_j \subseteq C_j \backslash R_i$, with $|H_i| = t - 1$ and $|K_j| = 2$.

We define a *cross set* to be a set of the form described in (5).

Note that $\Lambda_0(M) = \alpha_2(\Delta^{s,t}) = \alpha_3(\Delta^{s,t})$ and the members of $\Lambda_p(M)$ must be inclusion-wise minimal. Hence, not all sets listed under (1)-(5) above are necessarily members of $\Lambda_p(M)$. Therefore, we need the following result.

Lemma 7.10. *Every subset of R_1 of size at least $t - p$ that is a circuit of M' is a member of $\Lambda_p(M)$.*

Proof. It suffices to show that no member of $\Lambda_p(M)$ is properly contained in a circuit of M'. Other than circuits of M', the only members of $\Lambda_p(M)$ that are contained in R_1 are t-subsets of R_1 and the sets described in (4) of Definition 7.9. No circuit of M' has size greater than t. Therefore, no t-subset of R_1 is properly contained in a circuit of M'. If $p = t - 1$, then $t - p - 1 = 0$. Since the empty set is never a circuit of a matroid, none of the sets described in (4) of Definition 7.9 exist. If $p \le t - 2$, then $\ell - 2p + 1 \ge \ell - 2(t - 2) + 1 = n' + 2(t - 2) - 2(t - 2) + 1 > n'$. Since $[n']$ is the ground set of M', no set described in (4) is contained in a circuit of M'. $\qquad\square$

Lemma 7.11. *Let M be a matroid of rank $t - 1$, where $t \ge 3$. Let $1 \le p \le t - 1$. There is a sequence of clutters $\Lambda_{p-1}(M) = \Lambda_{p,0}, \Lambda_{p,1}, \ldots, \Lambda_{p,q} = \Lambda_p(M)$, where $\Lambda_{p,i}$ is an α_1-transformation of $\Lambda_{p,i-1}$.*

Proof. Note that $\Lambda_p(M) = \min(\Lambda_{p-1}(M) \cup \Gamma)$, where Γ is the collection of

- $(t - p)$-subsets of $[n] \subseteq R_1$ that are circuits of M' and

- sets of the form $A \cup \{x\}$, where $x \in \{\ell - 2p + 1, \ell - 2p + 2\}$ and A is a $(t - p - 1)$-subset of $[\ell - 2p]$ containing exactly one circuit of M'.

By induction, assume $\Lambda_{p,i-1} = \min(\Lambda_{p-1}(M) \cup \Gamma')$, where $\Gamma' \subseteq \Gamma$. Let $X \in \Gamma \backslash \Gamma'$. We proceed by taking cases on p.

Case 1. Assume that $p = 1$. Let $m_1, m_2 \in [\ell] \backslash X$. (This is possible because $\ell \ge t + 1$ if and only if $n' \ge -t + 5$. We have $t \ge 3$, implying that $3t \ge 9$. This implies $n' \ge 2t - 2 \ge -t + 7$.) Let $K_{m_1} = \{(2, m_1), (3, m_1)\}$ and $K_{m_2} = \{(4, m_2), (5, m_2)\}$. Note that $X \cup K_{m_1}$ and $X \cup K_{m_2}$ are cross sets and therefore members of $\Lambda_{p,0}$. Also note that $X \cup K_{m_1}$ and $X \cup K_{m_2}$ are the only members of $\Lambda_{p,i-1}$ contained in $X \cup K_{m_1} \cup K_{m_2}$. Therefore, $\alpha_1(\Lambda_{p,i-1}; X \cup K_{m_1}, X \cup K_{m_2}) = \min(\Lambda_{p,i-1} \cup \{X\})$. By iterating this process, we obtain $\Lambda_1(M)$.

Case 2. Assume that $p > 1$. By Definition 7.9, we see that $X \cup \{\ell - 2(p - 1) + 1\} \in \Lambda_{p-1}(M)$ and $X \cup \{\ell - 2(p - 1) + 2\} \in \Lambda_{p-1}(M)$. Also note that $X \cup \{\ell - 2(p - 1) + 1\}$ and $X \cup \{\ell - 2(p - 1) + 2\}$ are the only members of $\Lambda_{p,i-1}$ contained in $X \cup \{\ell - 2(p - 1) + 1, \ell - 2(p - 1) + 2\}$. This follows from the fact that all members of $\Lambda_{p,i-1}$ are inclusion-wise minimal (so $X \cup \{\ell - 2(p - 1) + 1, \ell - 2(p - 1) + 2\}$ is not a member of $\Lambda_{p,i-1}$) and the fact that $t \ge 3$ (so $\{\ell - 2(p - 1) + 1, \ell - 2(p - 1) + 2\}$ is not a member of $\Lambda_{p,i-1}$). Therefore, $\alpha_1(\Lambda_{p,i-1}; X \cup \{\ell - 2(p - 1) + 1\}, X \cup \{\ell - 2(p - 1) + 2\}) = \min(\Lambda_{p,i-1} \cup \{X\})$. By iterating this process, we obtain $\Lambda_p(M)$. $\qquad\square$

Proof of Theorem 7.7. First we consider the case where the rank of M is at most 1. Let c be the number of loops of M. Let $s = t = k = 2$ and $\ell = \max\{c+1, n\}$. For $i \leq c$, let $\Lambda_i = \alpha_1(\Lambda_{i-1}; \{(1,i),(1,\ell)\}, \{(1,i),(2,i)\}) = \Lambda_{i-1} \cup \{(1,i)\}$. The result is that $\{(1,j)\} \in \Lambda_c$ for every $j \leq c$. Now, let $\Lambda_{c+1} = \alpha_2(\Lambda_c)$. This is the union of Λ_c with the collection of 2-subsets of $\{(2,j) : 1 \leq j \leq \ell\} \cup \{(1,j) : c+1 \leq j \leq \ell\}$. This is the collection of circuits of a matroid whose restriction to $\{(1,j) : 1 \leq j \leq n\}$ is isomorphic to M.

Now let M be a matroid on $[n]$ of rank $t-1$, where $t \geq 3$. Let c, c', n', M', ℓ, and M^+ be as in Definition 7.8, while $s = 3$ and $k = 5$. As noted above, $\alpha_2(\Delta^{s,t}) = \Lambda_0(M)$. By repeated use of Lemma 7.11, we can obtain $\Lambda_{t-1}(M)$ from $\Lambda_0(M)$ by a sequence of α_1-transformations.

Now, let Γ be the clutter consisting of all 2-subsets of columns C_j for $1 \leq j \leq \ell$ other than the subsets containing a loop of M'. We will construct a sequence of clutters $\Lambda_{t-1}(M) = \Lambda'_0, \Lambda'_1, \ldots, \Lambda'_q = \min(\Lambda_{t-1}(M) \cup \Gamma)$ such that Λ'_i is an α_1-transformation of Λ'_{i-1}. By induction, assume $\Lambda'_{i-1} = \min(\Lambda_{t-1}(M) \cup \Gamma')$, where $\Gamma' \subseteq \Gamma$. Let $X \in \Gamma \backslash \Gamma'$. Then there are positive integers j, m_1, m_2 such that $X = \{(m_1, j), (m_2, j)\}$. Let $\{m_3, m_4\} \subseteq \{2, 3, 4, 5\} \backslash \{m_1, m_2\}$. Let $K_{m_3} \subseteq R_{m_3} \backslash C_j$ and $K_{m_4} \subseteq R_{m_4} \backslash C_j$ with $|K_{m_3}| = |K_{m_4}| = t-1$. Moreover, we can choose K_{m_3} and K_{m_4} so that no column contains an element of both. (This is possible because $\ell \geq 2(t-1)+1$ if and only if $n' \geq 3$. This is true because $n' \geq 2(t-1) \geq 4$.) Note that the only members of Λ'_{i-1} contained in $X \cup K_{m_3} \cup K_{m_4}$ are the cross sets $X \cup K_{m_3}$ and $X \cup K_{m_4}$. Therefore, $\alpha_1(\Lambda'_{i-1}; X \cup K_{m_3}, X \cup K_{m_4}) = \min(\Lambda'_{i-1} \cup \{X\})$. By iterating this process, we obtain $\min(\Lambda_{t-1}(M) \cup \Gamma)$. Let $\Lambda'' = \min(\Lambda_{t-1}(M) \cup \Gamma)$. By Lemma 7.10, the circuits of M' are members of $\Lambda_{t-1}(M)$. The circuits of M^+ that are not circuits of M' are the t-subsets of R_1 not containing a circuit of M'. Thus, the members of Λ'' are:

- 2-subsets of C_j, where $1 \leq j \leq \ell$, other than the subsets containing a loop of M',

- subsets of R_1 that are circuits of M^+, and

- t-subsets of R_i, where $2 \leq i \leq k$.

Recall that c is the number of loops of M and c' is the number of loops of M'. Without loss of generality, assume that the loops of M' are $\{n - c + 1, n - c + 2, \ldots, n'\}$.

Let Ψ be the clutter consisting of singleton sets $\{(i,j)\}$, where $2 \leq i \leq 5$ and either $1 \leq j \leq n - c$ or $n' + 1 \leq j \leq \ell$. (So j is not a loop of M^+.) We will construct a sequence of clutters $\Lambda'' = \Lambda''_0, \Lambda''_1, \ldots, \Lambda''_z = \min(\Lambda'' \cup \Psi)$ such that Λ''_i is an α_1-transformation of Λ''_{i-1}. By induction, assume $\Lambda''_{i-1} = \min(\Lambda'' \cup \Psi')$, where $\Psi' \subseteq \Psi$. Let $\{(i,j)\} \in \Psi \backslash \Psi'$. Let $A_1 = \{(i,j)\} \cup \{(i,m) : n - c + 1 \leq m \leq n - c + t - 1\}$ and let $A_2 = \{(i,j), (1,j)\}$. Since $n' \geq n - c + t - 1$, no element of $A_1 \cup A_2$ is a singleton set in Ψ. Thus, A_1 and A_2 are the only members of Λ''_{i-1} contained in $A_1 \cup A_2$ and $\alpha_1(\Lambda''_{i-1}; A_1, A_2) = \min(\Lambda'' \cup \{(i,j)\})$. By iterating this process, we obtain $\min(\Lambda'' \cup \Psi)$.

Let $\Lambda''' = \min(\Lambda'' \cup \Psi)$. The members of Λ''' are

- singleton sets $\{(i,j)\}$ where $2 \leq i \leq 5$ and either $1 \leq j \leq n - c$ or $n' + 1 \leq j \leq \ell$,

- 2-subsets of $C_j \backslash \{(1,j)\}$, where $n - c + 1 \leq j \leq n'$,

- t-subsets of $\{(i,j) : n - c + 1 \leq j \leq n'\}$ for some fixed i with $2 \leq i \leq 5$, and

- subsets of R_1 that are circuits of M^+.

Note that each member of Λ''' is either contained in R_1 or disjoint from R_1. For every clutter, a process of repeated α_2-transformations on the clutter must eventually terminate in a clutter that is the collection of circuits of a matroid. Therefore, repeated α_2-transformations on Λ''' will result in a clutter that is the collection of circuits of a matroid N whose circuits are either contained in R_1 or disjoint from R_1. Since the members of Λ''' contained in R_1 are the circuits of M^+, circuit elimination implies that the circuits of N contained in R_1 are precisely the circuits of M^+. In particular, the restriction of N to $\{(1,j) : 1 \leq j \leq n\}$ is isomorphic to M. $\qquad \square$

Theorem 7.7 applies to all matroids, regardless of rank. However, in the case where the matroid is simple and has rank at most 3, the result can be illustrated more clearly using a different procedure from the one used in the proof of the theorem. We use point and line configurations. Consider the Fano matroid whose point and line configuration is the Fano plane given in Figure 3. It is well-known that the Fano matroid is not realizable over any field of characteristic other than 2; therefore, it is not $\mathbb{C}$-realizable. In the case where $n = 7$ and

$$\mathcal{D} = \{L_1 = \{1,2,4\}, L_2 = \{1,3,6\}, L_3 = \{1,5,7\}, L_4 = \{2,3,5\}, L_5 = \{2,6,7\}, L_6 = \{3,4,7\}, L_7 = \{4,5,6\}\} \cup \binom{[n]}{4},$$

the only minimally dependent matroid with respect to $\min(\mathcal{D})$ is the Fano matroid, which is not relevant for our purposes because it is not $\mathbb{C}$-realizable. However, this can arise as the simplification of a matroid coming from a point and line configuration with $s = 2$ and $t = 3$.

Indeed, consider the case $k = \ell = 7$. Let (i,j) be the unique element contained in $R_i \cap C_j$. Iteratively performing α_1-transformations (defined earlier in this subsection), we can obtain from the clutter $\Delta^{2,3}$ a clutter that contains all of the singletons in R_i except for the elements (i,j) such that $j \in L_i$. In particular, the members of this clutter are the following:

- all singletons except for the elements (i,j) such that $j \in L_i$,

- 3-subsets of the form $\{(i,j) \in R_i : j \in L_i\}$, and

- 2-subsets of C_j of the form $\{(i_1,j), (i_2,j)\}$, where $j \in L_{i_1} \cap L_{i_2}$.

Then, by performing two transformations (both α_2-transformations, both α_3-transformations, or one of each), we obtain a clutter that is the collection of circuits of a matroid whose simplification is the Fano matroid and whose rank-1 flats each have three non-loop elements. The second transformation is needed to include the Fano matroid's circuits of size 4. (This works because, for each 4-element circuit C of the Fano matroid, there is a fifth element x such that $C \cup \{x\}$ is the union of two 3-element circuits whose intersection is x.)

8 Conditional independence models

Conditional independence (CI) models play an important role in algebraic statistics [Stu05]. Given a collection of random variables and knowledge of the conditional dependencies, or independencies, among them, we can ask what are the distributions that satisfy them. In a more general setting, some of the random variables appearing in a CI model can be prescribed as unobserved (or hidden). Our goal is to determine when certain constraints on the observed variables arise from conditions on the hidden variables [SA15]. This problem can be restated algebraically by noting that probability distributions satisfying CI statements are the solutions of certain polynomial equations [DSS09, Sul18] which generate the so-called *CI ideal*. The distributions satisfying a given collection of CI statements can be recovered by intersecting the CI ideal with the probability simplex. When there are no hidden variables, these polynomials are binomials and their associated ideals are well-studied; see e.g. [Fin11, HHH⁺10, Rau13, ST13]. However, in the presence of hidden variables, the polynomials become far more complicated of arbitrarily high degrees and very difficult to calculate; see e.g. [PS19, CMR20].

Let X, Y_1, Y_2 be observed and H_1, H_2 be hidden random variables taking values in the finite sets $\mathcal{X}, \mathcal{Y}_1, \mathcal{Y}_2, \mathcal{H}_1, \mathcal{H}_2$ of cardinalities $|\mathcal{X}| = d$, $|\mathcal{Y}_1| = k$, $|\mathcal{Y}_2| = \ell$, $|\mathcal{H}_1| = s-1$, $|\mathcal{H}_2| = t-1$. Consider the CI model given by:

$$\mathcal{C}: \quad X \perp\!\!\!\perp Y_1 \mid \{Y_2, H_1\} \quad \text{and} \quad X \perp\!\!\!\perp Y_2 \mid \{Y_1, H_2\}. \tag{4}$$

Then the CI ideal associated to $\mathcal{C}$ is precisely the hypergraph ideal $I_{\Delta^{s,t}}$ in §7; see [CMM21, CMR20].

Example 8.1. Let $d = k = s = t = 3$ and $\ell = 4$. The joint distribution of Y_1 and Y_2 has state space $\mathcal{Y} = \mathcal{Y}_1 \times \mathcal{Y}_2$ which is identified with the 3×4 matrix $\mathcal{Y}$ with values in the set $[12]$ as in (3). In this case, the CI ideal $I_\Delta \subseteq \mathbb{C}[P] := \mathbb{C}[p_{x,y} : x \in \mathcal{X}, y \in \mathcal{Y}]$ is the hypergraph ideal from Example 3.12. The ideal I_Δ has two prime components. One component, associated to the matroid M_0 from Example 3.12, is generated by all 3-minors of the matrix of variables P. This is the ideal associated to the CI statements $X \perp\!\!\!\perp \{Y_1, Y_2\} \mid H_1$ and $X \perp\!\!\!\perp \{Y_1, Y_2\} \mid H_2$. Note that H_1 and H_2 are hidden random variables taking the same number of values, and so the ideals do not distinguish them. The other prime component I_M is the ideal of the configuration described in Example 3.12 consisting of 12 points and 7 lines. Besides the rank constraints given by the original CI model, the ideal I_M contains geometric constraints which are satisfied by all distributions which do not lie in V_{M_0}.

We note that for large values of d, understanding the algebraic properties of the ideal $I_{\Delta^{s,t}}$ and its (primary) decomposition is hard. In particular, Theorem 7.7 shows that, up to simplification, any matroid can appear among the dependent matroids for $\Delta^{s,t}$. And, in the example above, some high-degree polynomials may appear in the generating sets of the primary components of $I_{\Delta^{s,t}}$. However, for $d \leq s+t-3$ we have shown in Theorem 7.5 that $\Delta^{s,t}$ has a unique minimally dependent matroid M. So, for a generic point p in the variety $V_{\Delta^{s,t}}$, the additional polynomial constraints on p arise from geometric constraints on realizations of M.

Remark 8.2. Suppose that H_1 and H_2 are both constant. The *intersection axiom* states that any distribution satisfying $\mathcal{C}$ in (4) *generically* satisfies $X \perp\!\!\!\perp \{Y_1, Y_2\}$. Here, genericity means that the distributions have non-zero probabilities. The intersection axiom has been studied for various CI models; see e.g. [HHH⁺10, Rau13, CMR20, CMM21, PS19]. We note that our family of hypergraph varieties include all these cases as examples. In particular, the corresponding ideal $I_{\Delta^{s,t}}$ has a distinguished prime component with a particular statistical significance, since the distributions which lie inside do not contain any structural zeros [Stu02]. Importantly, if we assume, without loss of generality, that $s \leq t$ then this prime ideal can be realized as the CI ideal of $X \perp\!\!\!\perp \{Y_1, Y_2\} \mid H_2$. We may therefore deduce a hidden variable version of the intersection axiom as follows:

$$\mathcal{C} = \{X \perp\!\!\!\perp Y_1 \mid \{H_1, Y_2\}, \ X \perp\!\!\!\perp Y_2 \mid \{Y_1, H_2\}\} \implies X \perp\!\!\!\perp \{Y_1, Y_2\} \mid H_2.$$

Finally, we give connections of our work to an interesting conjecture by Matúš; see [Mat99].

Conjecture 8.3 ([Mat99]). *For any discrete conditional independence model $\mathcal{C}$, there exists a distribution $p \in V(J_\mathcal{C})$ such that all joint probabilities of p are rational.*

In Theorem 7.7, we have seen that for large enough s, t, k, ℓ, any matroid may appear among the dependent matroids for $\Delta^{s,t}$. A natural approach is to carefully choose additional conditional independence and dependence statements for the model (4), in order to guarantee that any distribution p satisfying C is a realization of a given, realizable, matroid. Note that there exist matroids that are not realizable over the rationals but are realizable over a real field extension. Hence, this might lead to a characterization of CI models with hidden variables for which Conjecture 8.3 does not hold.

References

[BC03]   Winfried Bruns and Aldo Conca. "Gröbner bases and determinantal ideals." *Commutative Algebra, Singularities and Computer Algebra* 179 (2003): 9–66.

[BS89]   Jürgen Bokowski and Bernd Sturmfels. "Computational synthetic geometry." *Lecture Notes in Math.* 1355 (1989).

[CLO15]   David A. Cox, John Little, and Donal O'Shea. "Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra (4th edition)." *Undergrad. Texts Math.* (2015).

[CMM21]   Oliver Clarke, Fatemeh Mohammadi, and Harshit J Motwani. "Conditional probabilities via line arrangements and point configurations." *Linear Multilinear Algebra* (2021): 1–33.

[CMR20]   Oliver Clarke, Fatemeh Mohammadi, and Johannes Rauh. "Conditional independence ideals with hidden variables." *Adv. in Appl. Math.* 117 (2020): 102029.

[Cra65]   Henry H. Crapo. "Single-element extensions of matroids." *Journal of Research of the National Bureau of Standards, Section B*, 69B (1965): 55–65.

[DSS09]   Mathias Drton, Bernd Sturmfels, and Seth Sullivant. "Lectures on Algebraic Statistics." *Oberwolfach Semin.* 39 (2009).

[EHHM13] Viviana Ene, Jürgen Herzog, Takayuki Hibi, and Fatemeh Mohammadi. "Determinantal facet ideals." *Michigan Math. J.* 62 (2013): 39–57.

[Fin11]   Alex Fink. "The binomial ideal of the intersection axiom for conditional probabilities." *J. Algebraic Combin.* 33 (2011): 455–463.

[GGMS87] Israel M. Gelfand, R. Mark Goresky, Robert D. MacPherson, and Vera V. Serganova. "Combinatorial geometries, convex polyhedra, and schubert cells." *Adv. Math.* 63 no. 3 (1987): 301–316.

[GS]   Daniel R. Grayson and Michael E. Stillman. "Macaulay2." A software system for research in algebraic geometry. Available at https://faculty.math.illinois.edu/Macaulay2/.

[Har13]   Robin Hartshorne. "Algebraic geometry." *Grad. Texts in Math.* 52 (2013).

[HHH+10] Jürgen Herzog, Takayuki Hibi, Freyja Hreinsdóttir, Thomas Kahle, and Johannes Rauh. "Binomial edge ideals and conditional independence statements." *Adv. in Appl. Math.* 45 no. 3 (2010): 317–333.

[HS04]   Serkan Hoşten and Seth Sullivant. "Ideals of adjacent minors." *J. Algebra*, 277 no. 2 (2004): 615–642.

[KLS13]   Allen Knutson, Thomas Lam and David E. Speyer. "Positroid varieties: juggling and geometry." *Compos. Math* 149 no. 10 (2013): 1710–1752.

[LV13]   Seok Hyeong Lee and Ravi Vakil. "Mnëv-Sturmfels universality for schemes." *Clay Mathematics Proceedings* 18, A celebration of algebraic geometry, (2013): 457–468.

[Mat99]   František Matúš. "Conditional independences among four random variables III: Final conclusion." *Combin. Probab. Comput.* 8 no. 3 (1999): 269–276.

[MF14]   Jaume Martí-Farré. "From clutters to matroids." *Electron. J. Combin.*, 21 no. 1 (2014): P1.11.

[MFdM15] Jaume Martí-Farré and Anna de Mier. "Completion and decomposition of a clutter into representable matroids." *Linear Algebra Appl.* 472 (2015): 31–47.

[MFdM17] Jaume Martí-Farré and Anna de Mier. "Transformation and decomposition of clutters into matroids." *Adv. Math.* 312 (2017): 286–314.

[Mnë85] Nikolai E. Mnëv. "On manifolds of combinatorial types of projective configurations and convex polyhedra." *Soviet Math. Doklady* 32 (1985): 335–337.

[Mnë88] Nikolai E. Mnëv. "The universality theorems on the classification problem of configuration varieties and convex polytopes varieties." *Lecture Notes in Math.* 1346 *Topology and geometry—Rohlin seminar* (1988): 527–543.

[MR18] Fatemeh Mohammadi and Johannes Rauh. "Prime splittings of determinantal ideals." *Comm. Algebra* 46 no. 5 (2018): 2278–2296.

[Oh09] Suho Oh. "Combinatorics of Positroids." *Discrete Math. Theor. Comput. Sci.* DMTCS Proceedings vol. AK, 21st International Conference on Formal Power Series and Algebraic Combinatorics (2009).

[Oxl11] James Oxley. "Matroid Theory." Second edition. *Oxf. Grad. Texts Math* (2011).

[PS19] Gerhard Pfister and Andreas Steenpass. "On the primary decomposition of some determinantal hyperedge ideal." *J. Symbolic Comput.* 103 (2019): 14–21.

[PW70] Mike J. Piff and Dominic J.A. Welsh. "On the vector representation of matroids." *J. Lond. Math. Soc. (2)* 2 (1970): 284–288.

[Rau13] Johannes Rauh. "Generalized binomial edge ideals." *Adv. in Appl. Math.* 50 no. 3 (2013): 409–414.

[RG99] Jürgen Richter-Gebert. "The universality theorems for oriented matroids and polytopes. *Contemp. Math.* 223 (1999): 269–292.

[RG11] Jürgen Richter-Gebert. "Perspectives on Projective Geometry: A Guided Tour Through Real and Complex Geometry." Springer Berlin Heidelberg, 2011.

[Ryb11] Grigori L. Rybnikov. "On the fundamental group of the complement of a complex hyperplane arrangement." *Funct. Anal. Appl.* 45 no. 2 (2011): 137–148.

[SA15] Bastian Steudel and Nihat Ay. "Information-theoretic inference of common ancestors." *Entropy* 17 no. 4 (2015): 2304–2327.

[SJS17] Meera Sitharam, Audrey St. John, and Jessica Sidman. "Handbook of Geometric Constraint Systems Principles (1st edition)." *Chapman & Hall/CRC Math.* (2017).

[ST13] Irena Swanson and Amelia Taylor. "Minimal primes of ideals arising from conditional independence statements." *J. Algebra* 392 (2013): 299–314.

[Stu89] Bernd Sturmfels. "On the matroid stratification of Grassmann varieties, specialization of coordinates, and a problem of N. White." *Adv. Math.* 75 no. 2 (1989): 202–211.

[Stu90] Bernd Sturmfels. "Gröbner bases and Stanley decompositions of determinantal rings." *Math. Z.* 205 no. 1 (1990): 137–144.

[Stu02] Bernd Sturmfels. "Solving systems of polynomial equations." *CBMS Reg. Conf. Ser. Math.* 97 (2002).

[Stu05] Milan Studený. "Probabilistic conditional independence structures." *Information Science and Statistics* Springer, London, (2005).

[Stu08] Bernd Sturmfels. "Algorithms in invariant theory." *Texts Monogr. Symbol. Comput.* (2008).

[STW21] Jessica Sidman, Will Traves, and Ashley Wheeler. "Geometric equations for matroid varieties." *J. Combin. Theory Ser. A* 178 (2021): 105360.

[Sul18] Seth Sullivant. "Algebraic Statistics." *Grad. Stud. Math.* (2018).

Authors' addresses:

Department of Mathematics, University of Bristol, Bristol, UK
E-mail address: `oliver.clarke@bristol.ac.uk`

School of Mathematics, University of Bristol, Bristol, UK
Heilbronn Institute for Mathematical Research, Bristol, UK
Current address: Department of Mathematics, Vanderbilt University, Nashville, Tennessee, USA
E-mail address: `kevin.m.grace@vanderbilt.edu`

Department of Computer Science, KU Leuven, Celestijnenlaan 200A, B-3001 Leuven, Belgium
Department of Mathematics, KU Leuven, Celestijnenlaan 200B, B-3001 Leuven, Belgium
Department of Mathematics and Statistics, UiT – The Arctic University of Norway, 9037 Tromsø, Norway
E-mail address: `fatemeh.mohammadi@kuleuven.be`

Department of Mathematics, Ghent University, Gent, Belgium
E-mail address: `harshitjitendra.motwani@ugent.be`