Maximal Semigraphoids

Laura Silverstein and Michael O'Sullivan
Department of Mathematics and Statistics, San Diego State University

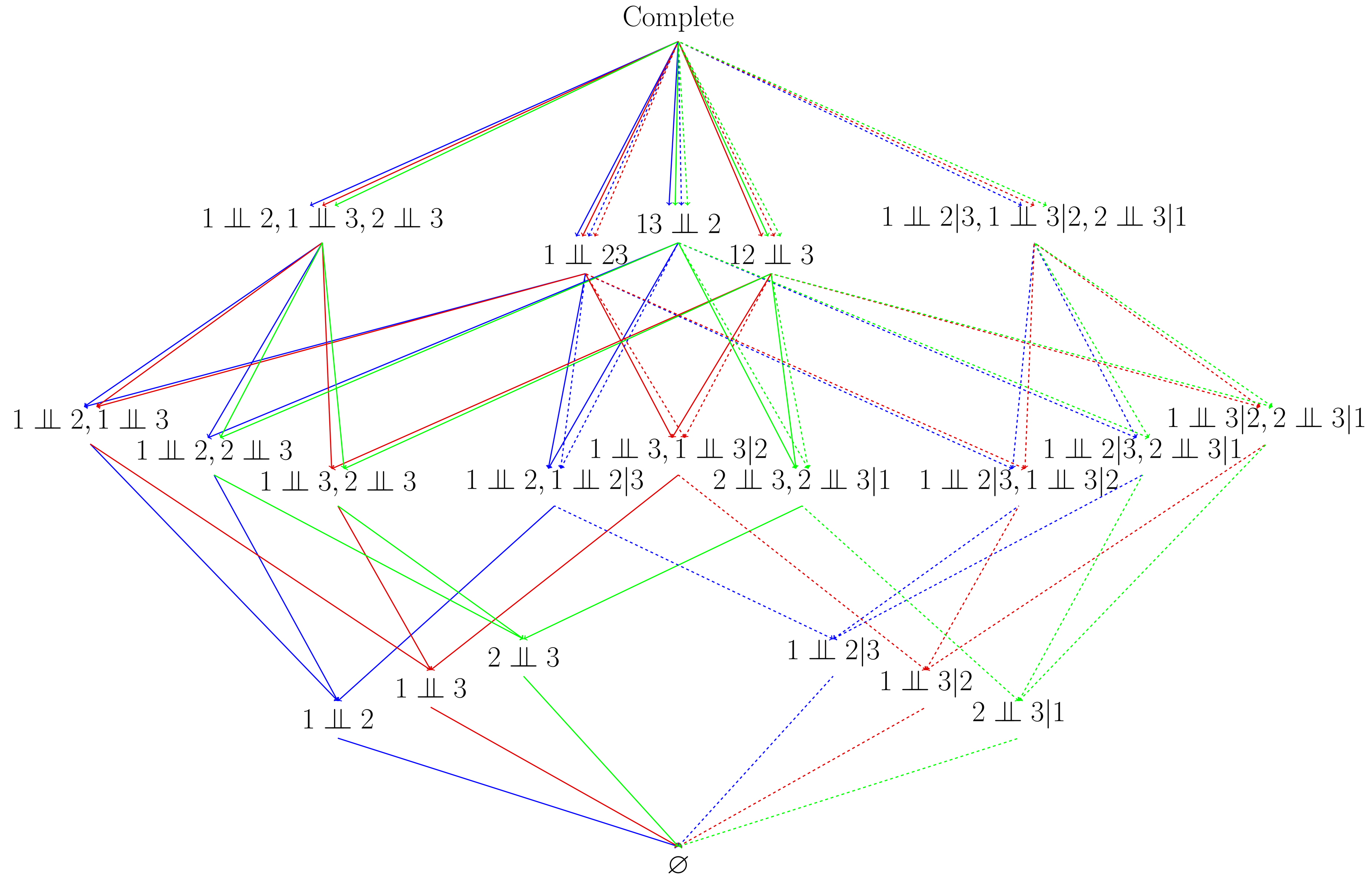
Purpose

Semigraphoids are combinatorial structures arising from models of probabilistic conditional independence. By studying the characterizations of maximal semigraphoids, we highlight the construction of two very different classes.

What is a Semigraphoid?

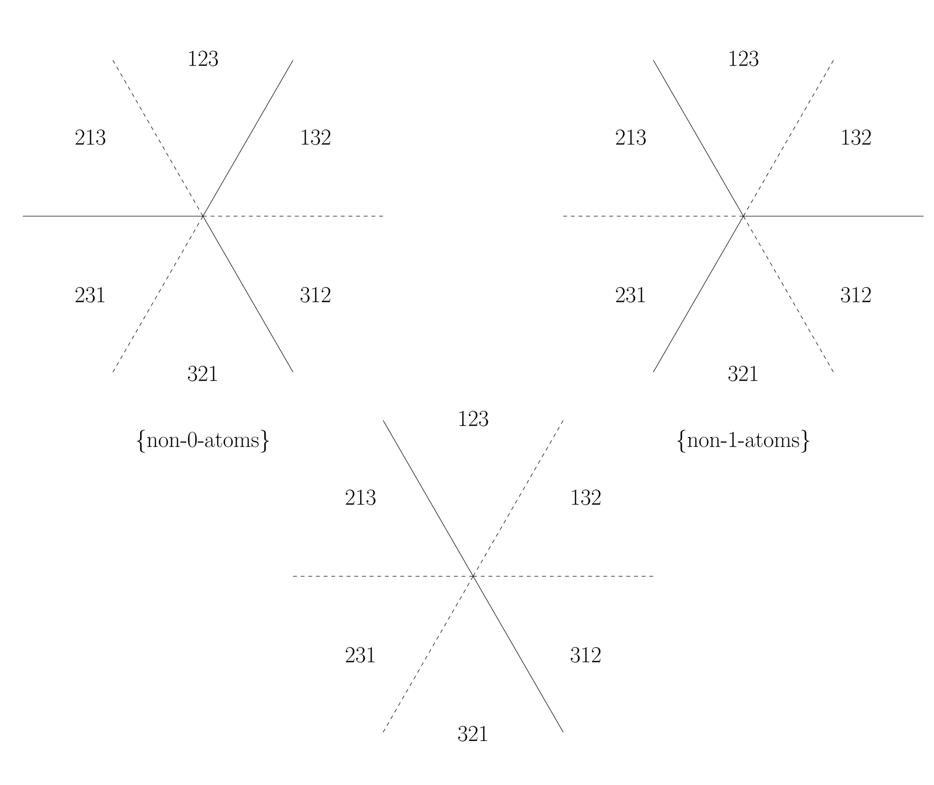
Let $a, b, c \in [n] = \{1, \dots, n\}$. A *semigraphoid* $\mathcal{S}$ is a set of triplets of the form $a \perp\!\!\!\perp b | K$, for $K \subseteq [n] \setminus ab$, that satisfy the following axiom:

$$\{a \perp\!\!\!\perp b|cK, a \perp\!\!\!\perp c|K\} \subset \mathcal{S} \Leftrightarrow \{a \perp\!\!\!\perp c|bK, a \perp\!\!\!\perp b\} \subset \mathcal{S}.$$



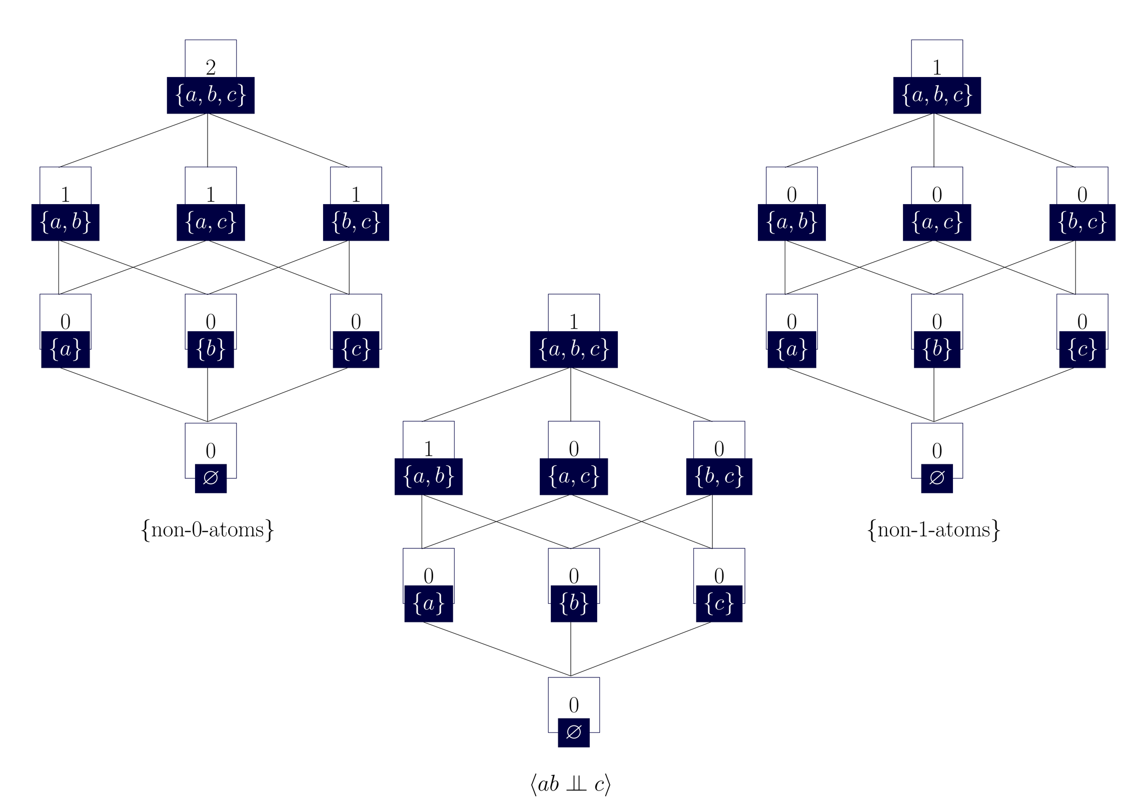
Convex Rank Tests

- The S_n-*fan* in $\mathcal{R}^n$ is defined by the planes $x_i = x_j$.
- A *convex rank test* gives a coarsening of the S_n-fan. A fan achieved by erasing rays of the S_n-fan.
- There is a bijection between convex rank tests and semigraphoids. [2]



Structural Imsets

- An *imset over* $[n]$ is an element of $\mathbb{Z}^{\mathcal{P}([n])}$.
- The imset $u_{a \perp\!\!\!\perp b|K} = \delta_{abK} + \delta_K - \delta_{aK} - \delta_{bK}$.
- For $m, u \in \mathbb{Z}^{\mathcal{P}([n])}$, the l-standardized imset of the semigraphoid u is $\langle m, u \rangle = \sum_{s \subseteq [n]} m_s u_s$. [3]
- The semigraphoid $\mathcal{M}^m$ is defined to be $\mathcal{M}^m = \{u_{a \perp\!\!\!\perp b|K} : \langle m, u_{a \perp\!\!\!\perp b|K} \rangle = 0\}$.



Classification on Four Random Variables

- *Proposition.* For any $k \in [n-2]$, the semigraphoid $\cup_{l \in \{1,\dots,n-2\} \setminus k} \{l - \text{atoms}\}$ is maximal. We call this semigraphoid the *non-k-atoms.*

1. non-0-atoms $= \cup_{k \in \{1,2\}} \{k\text{-atoms}\}$
2. non-1-atoms $= \cup_{k \in \{0,2\}} \{k\text{-atoms}\}$
3. non-2-atoms $= \cup_{k \in \{0,1\}} \{k\text{-atoms}\}$

- Let $A \subseteq [n]$ and $B = [n] \setminus A$. Given two semigraphoids $\mathcal{S} \subseteq CI(A), \mathcal{T} \subseteq CI(B)$, we define the *direct sum* of S and T to be the semigraphoid:
  $\mathcal{S} \oplus \mathcal{T} = \langle \mathcal{S} \cup \mathcal{T} \cup \{a \perp\!\!\!\perp b|K : a \in A, b \in B, K \subset AB \setminus ab\} \rangle.$
- *Proposition.* Let N be an index set and $A \subset N$. The direct sum of a maximal semigraphoid on A and the complete model on $N \setminus A$ is maximal on N.

4. $\langle ab \perp\!\!\!\perp c \rangle \oplus \langle \varnothing \rangle_{\{d\}}$
5. $\langle a \perp\!\!\!\perp b, a \perp\!\!\!\perp c, b \perp\!\!\!\perp c \rangle \oplus \langle \varnothing \rangle_{\{d\}}$
6. $\langle a \perp\!\!\!\perp b|c, a \perp\!\!\!\perp c|b, b \perp\!\!\!\perp c|a \rangle \oplus \langle \varnothing \rangle_{\{d\}}$

Complementary Maximals

- Given a semigraphoid $\mathcal{S}$, let $\mathcal{S}_k = \mathcal{S} \cap \{k\text{-atoms}\}$ for each k. Define its *complement* $\mathcal{S}^C$ by
  $$(\mathcal{S}^C)_k = \begin{cases} \{k\text{-atoms}\} & \text{if } \mathcal{S}_k = \{k\text{-atoms}\} \\ \{k\text{-atoms}\} \setminus \mathcal{S}_k & \text{else.} \end{cases}$$
- *Proposition.* The complement of the direct sum $\{\text{non-}k\text{-atoms}\} \oplus \varnothing_n$ is maximal.

7. $\left\langle \begin{Bmatrix} a \perp\!\!\!\perp b \\ a \perp\!\!\!\perp c \\ b \perp\!\!\!\perp c \end{Bmatrix}, \begin{Bmatrix} a \perp\!\!\!\perp b|c \\ a \perp\!\!\!\perp c|b \\ b \perp\!\!\!\perp c|a \end{Bmatrix}, \{2-\text{atoms}\} \right\rangle$
8. $\left\langle \{0-\text{atoms}\}, \begin{Bmatrix} a \perp\!\!\!\perp b|c \\ a \perp\!\!\!\perp c|b \\ b \perp\!\!\!\perp c|a \end{Bmatrix}, \begin{Bmatrix} a \perp\!\!\!\perp b|cd \\ a \perp\!\!\!\perp c|bd \\ b \perp\!\!\!\perp c|ad \end{Bmatrix} \right\rangle$

- The last two are complements and can be found by strategically choosing atoms without consequences.

	0-atoms	1-atoms	2-atoms
9.	$\langle \; a \perp\!\!\!\perp b,$	$\sim\begin{Bmatrix} a \perp\!\!\!\perp x|b \\ b \perp\!\!\!\perp x|a \end{Bmatrix}_{x \in [4] \setminus ab},$	$c \perp\!\!\!\perp d|ab \; \rangle$
10.	$\langle \; \sim\{a \perp\!\!\!\perp b\},$	$\begin{Bmatrix} a \perp\!\!\!\perp x|b \\ b \perp\!\!\!\perp x|a \end{Bmatrix}_{x \in [4] \setminus ab},$	$\sim\{c \perp\!\!\!\perp d|ab\} \; \rangle$

Connections with Imsets

- *Proposition.* The imset for $\{\text{non-}k\text{-atoms}\}$ is
  $$\sum_{|s| \geq k+2} \delta_s(|s| - k - 1).$$
- *Proposition.* The direct sum of $\{\text{non-}k\text{-atoms}\}$ on $[n-1]$ and $\varnothing_n$ is:
  $$\{\text{non-}k\text{-atoms}\}_{[n-1]} \oplus \varnothing_n = CI(n) \setminus \cup_{a \perp\!\!\!\perp b|K} \{a \perp\!\!\!\perp b|K, a \perp\!\!\!\perp b|nK\}.$$
- *Proposition.* Given the semigraphoid $\{\text{non-}k\text{-atoms}\}$ on $[n]$, we have:
  $$\{\text{non-}k\text{-atoms}\}_{[n]} \oplus \varnothing_{n+1} \cdots \oplus \varnothing_{n+r} = CI(n+r) \setminus \cup_{a \perp\!\!\!\perp b|K, k\text{-atom on } [n]} \{a \perp\!\!\!\perp b|KT : T \subseteq \{n+1, \dots, n+r\}\}.$$
- *Proposition.* Suppose $\mathcal{M}^m$ is maximal. Then $\mathcal{M}^m \oplus \varnothing_n = \mathcal{M}^{\overline{m}}$ where $\overline{m} = \sum_{s \subseteq [n-1]} m_s(\delta_s + \delta_{sn})$.

Five Random Variables

- From [3], there are 1319 maximal submodular semigraphoids up to permutation. We have found:
- 10 direct sums
- 4 sets of non-k-atoms
- 3 complements of direct sums on non-k-atoms
- 4 non-submodular that include the set $\{0\text{-atoms}, 2\text{-atoms}\}$.

References

[1] F. Matus, *Towards Classification of Semigraphoids*, Discrete Mathematics, **277**, pp. 115-145, 2004.

[2] J. Morton, L.Pachter, A. Shiu, B. Sturmfels, O. Wienand, *Convex Rank Tests and Semigraphoids*, SIAM Journal on Discrete Mathematics, **23**, pp. 1117-1134, 2009.

[3] M. Studeny, *Probabilistic Conditional Independence Structures*, Springer Series in Information Science and Statistics, Springer-Verlag, London, UK, 2005.

San Diego State University