Unconditional Equivalence in Causal Discovery

Danai Deligeorgaki, Department of Mathematics at KTH,
with Liam Solus (supervisor), Alex Markham and Pratik Misra

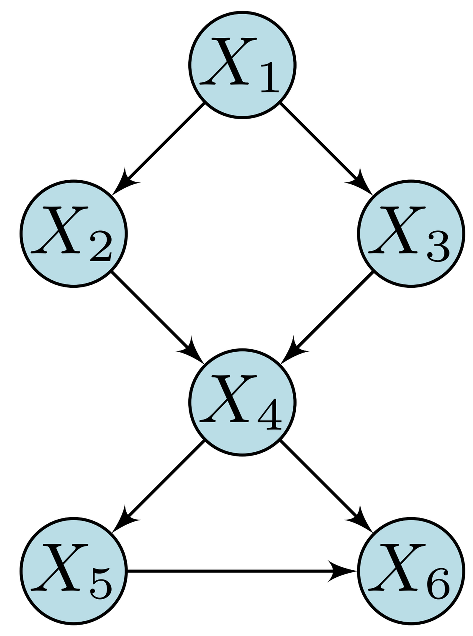
Setup

Directed acyclic graphs (DAGs) or bayesian networks are directed simple graphs without cycles.

- A DAG is used to model a **causal system**, i.e., a distribution $\mathbb{P}$ of $X_1, \ldots, X_n$ that encodes conditional independence statements.
- **Conditional independence (CI)** statements $X_A \perp\!\!\!\perp X_B \mid X_C$ in $\mathbb{P}$ can be expressed in the DAG $\mathcal{G}$ through connectivity rules.

Assumption: The distribution $\mathbb{P}$ is **Markov** to its underlying causal system: a DAG $\mathcal{G}$ that encodes exactly the CI statements that hold in the distribution.

Challenge: Given observed data from $\mathbb{P}$, can we recover $\mathcal{G}$?



MECs and UECs

Markov Equivalence:

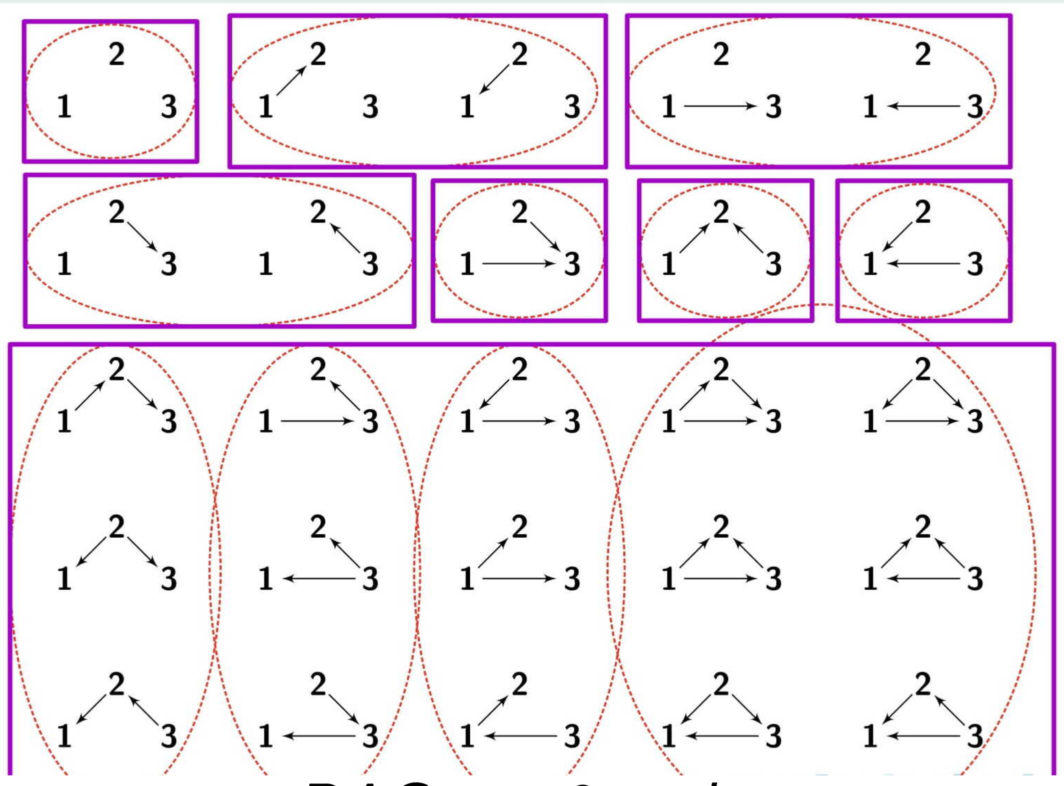
- We partition the space of DAGs into **Markov Equivalence Classes (MECs)** based on their CI relations.

Problem: Given data from $\mathbb{P}$, find the MEC of $\mathcal{G}$.

Unconditional Equivalence:

- The MEC partition is a refinement of the UEC partition:

Two DAGs are in the same **Unconditional Equivalence Class (UEC)** if they encode the same unconditional independence statements $(X_A \perp\!\!\!\perp X_B \mid \emptyset)$.

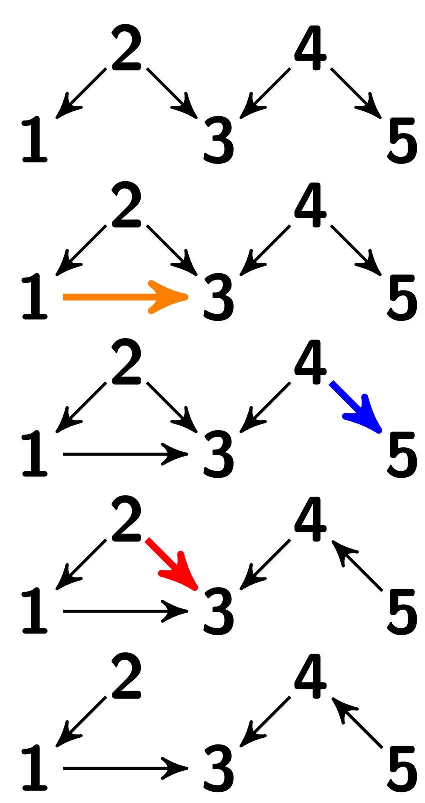


DAGs on 3 nodes

The structure of a UEC

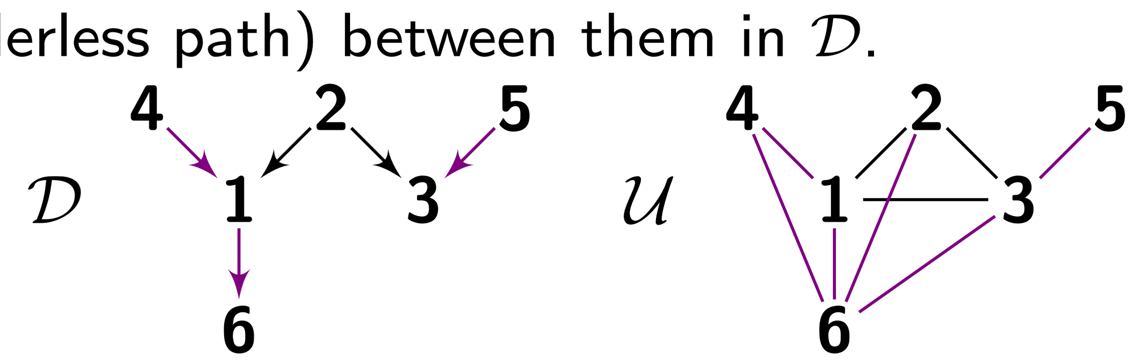
Transformational Characterization [1]:
Let $\mathcal{D}, \mathcal{D}'$ be DAGs in the same UEC. There is a sequence of edge insertions, reversals and deletions that transforms $\mathcal{D}$ into $\mathcal{D}'$ such that:

- Each edge inserted or deleted in is partially weakly covered or implied by transitivity.
- Each edge reversed in $\mathcal{D}$ is weakly covered.
- After each operation, the resulting DAG is in the same UEC as $\mathcal{D}$.



Structural Characterization:
The **unconditional dependence graph** $\mathcal{U}$ of a DAG $\mathcal{D}$ has the same nodes as $\mathcal{D}$ and an edge between v, w if there is a trek (colliderless path) between them in $\mathcal{D}$.

Example:



References

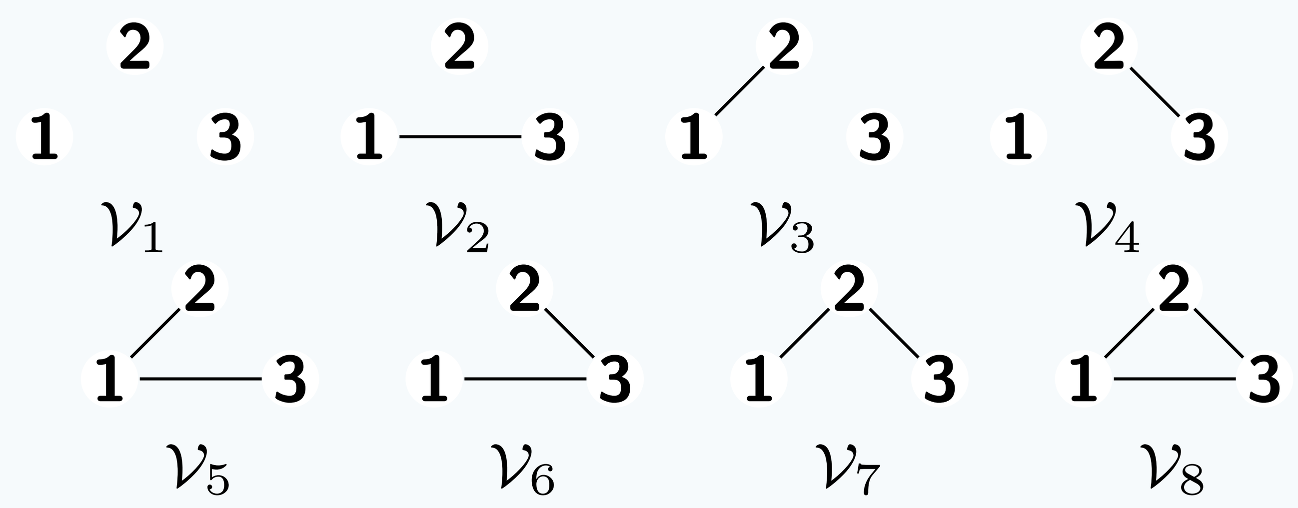
[1] *Transformational Characterization of Unconditionally Equivalent Bayesian Networks*, A. Markham, D. Deligeorgaki, P. Misra and L. Solus, 11th International Conference on Probabilistic Graphical Models (PGM), 2022.

[2] *Combinatorial and algebraic perspectives on the marginal independence structure of Bayesian networks*, D. Deligeorgaki, A. Markham, P. Misra and L. Solus, arXiv:2210.00822v1, 2022.

Traversing the space of UECs

In [2] we give a **Transformational and Structural Characterization** of the graphs representing a UEC of DAGs, using algebraic geometry (e.g., Gröbner bases) and combinatorics.

List of graphs representing a UEC on 3 nodes:

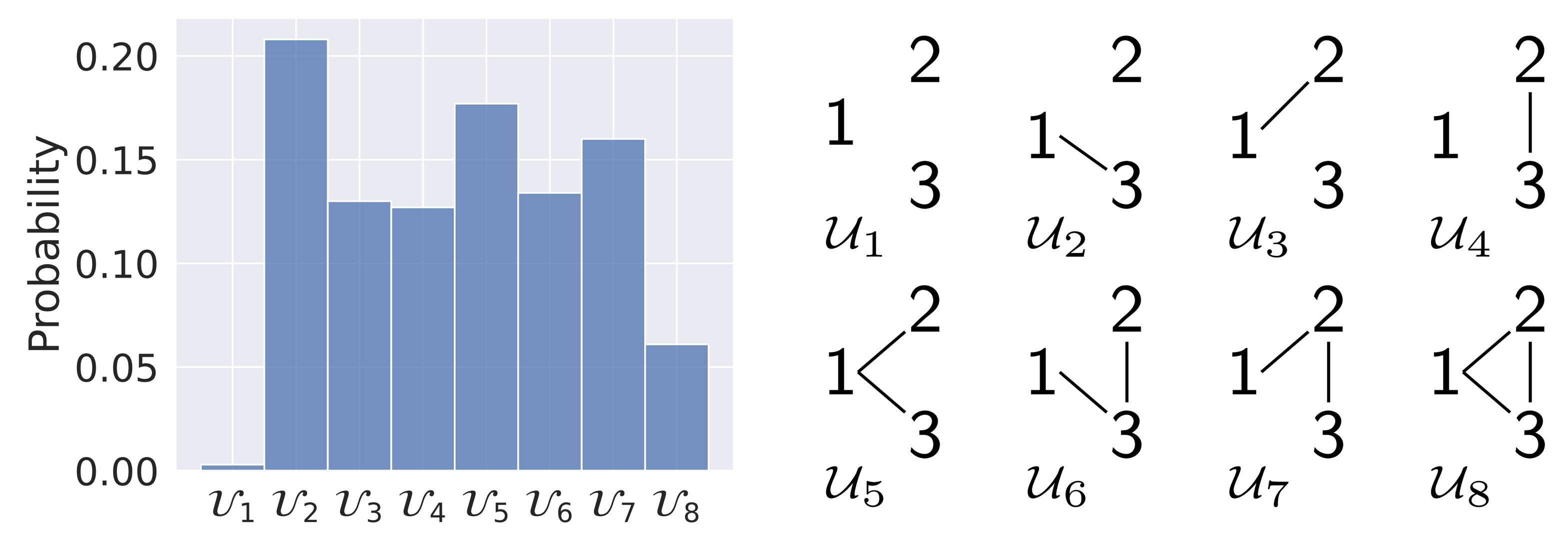


Number of UECs on 1,2,3,... nodes: 1,2,8,49,462,6424,...

Applications

Using the transformational characterization, we implemented an MCMC method for estimating the UEC of $\mathcal{G}$, **GrUES** (Gröbner-based Unconditional Equivalence Search).

A) We simulated a data set (1000 samples) over 3 nodes using the Gaussian linear model (with randomly generated weights) $X_1 = \epsilon_1, \quad X_2 = \epsilon_2, \quad X_3 = -0.9247 X_1 + \epsilon_3$.



▶ Histogram of the Markov chain indicates a MAP estimate $\pi(\mathcal{U}_2 | X) = 0.208$ (i.e., $\mathcal{U}_2$ has the highest frequency).

B) 100 linear Gaussian DAG models on 5 nodes for edge probability $p \in \{0.1, 0.2, \ldots, 0.9\}$ and random edge weights