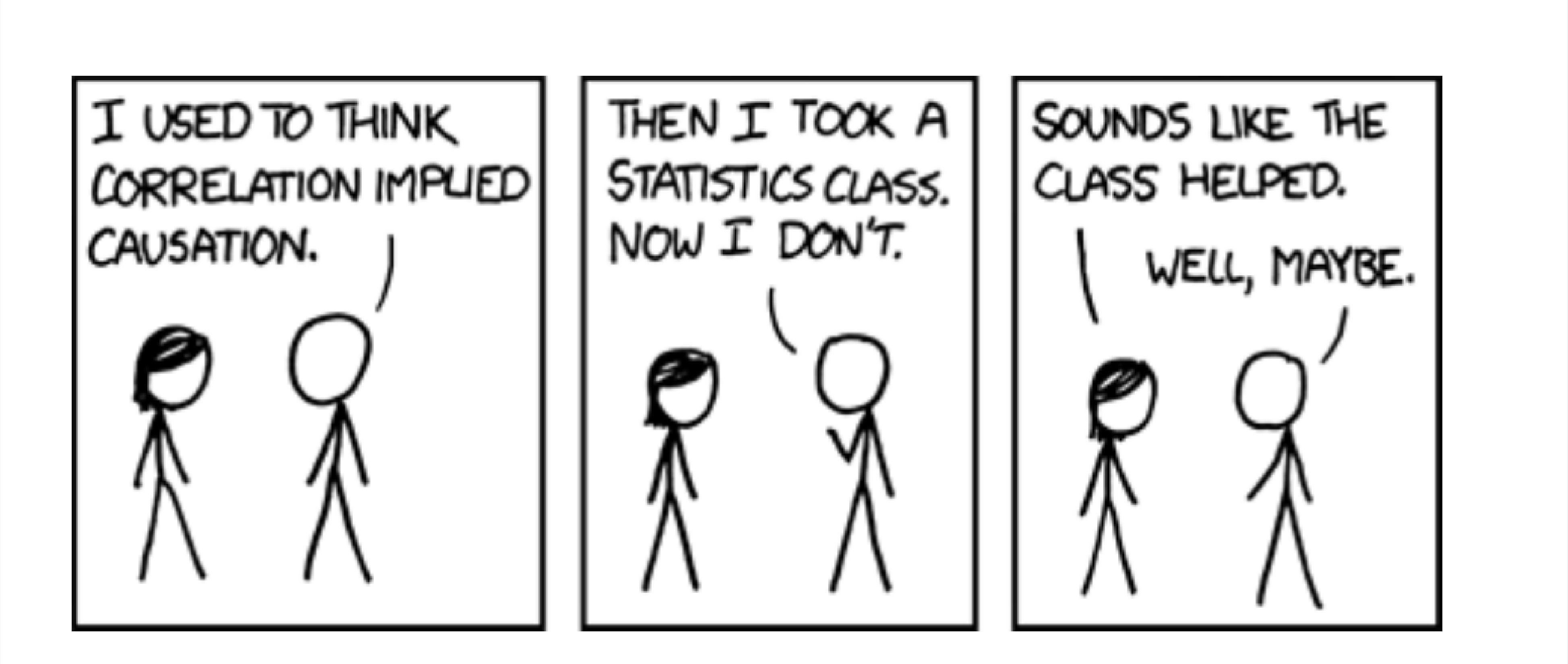
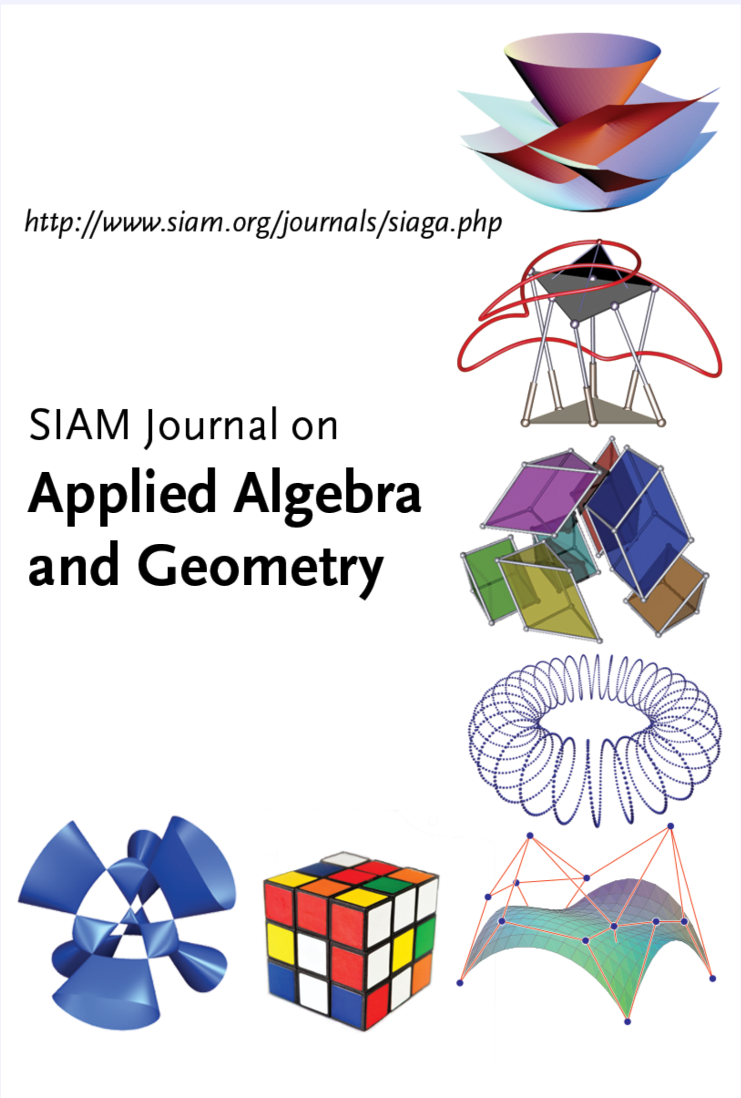
# Hypersurfaces in PC testing




*(image and text copyright Randall Munroe)*

Jan Draisma Universität Bern and TU Eindhoven

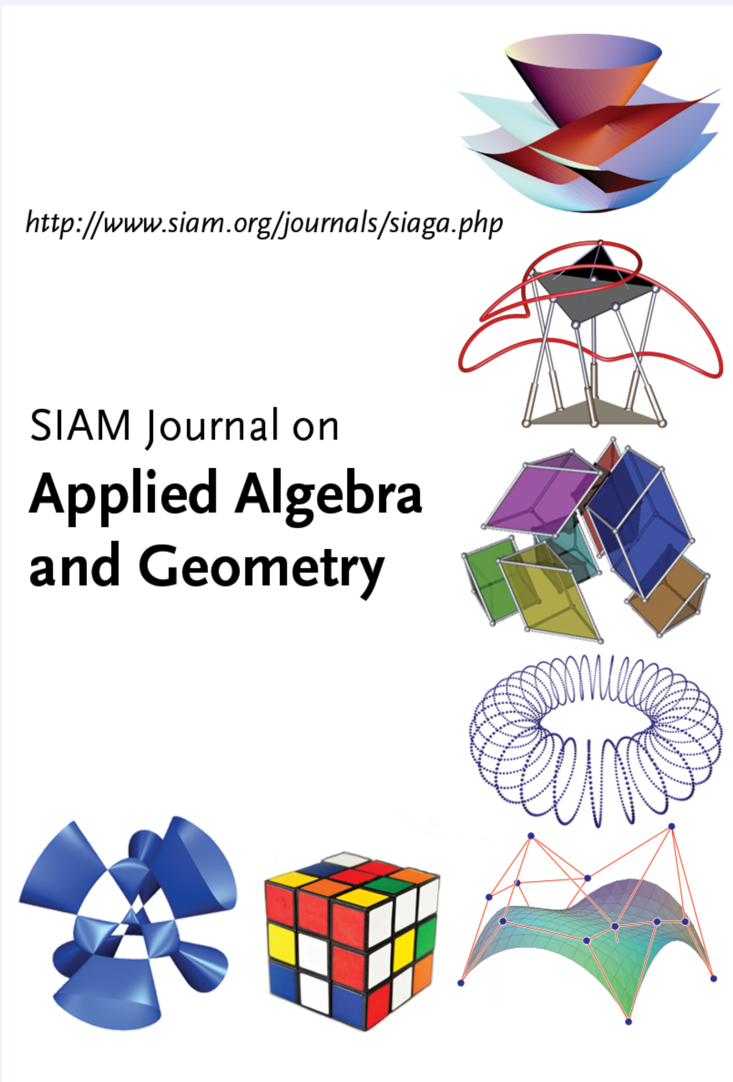
**SIAGA:**


http://www.siam.org/journals/siaga.php
SIAM Journal on
Applied Algebra
and Geometry

# Two advertisements



## SIAGA:





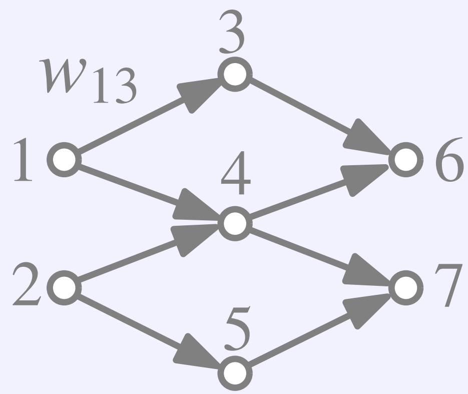
**SIAM AG 19**: 9–13 July 2019, Bern

# Warm-up: the Gessel-Viennot-Lindström Lemma

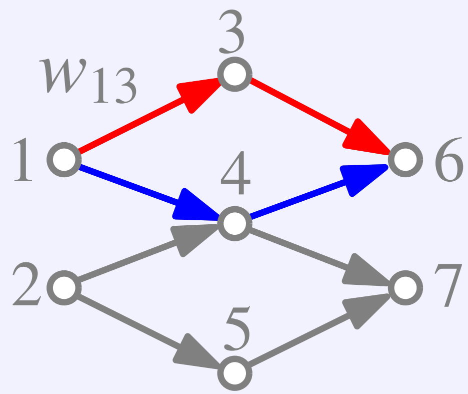
**Setting**

$\Gamma$ a directed acyclic graph (DAG) on $[m]$

for $i \to j$ a weight $w_{ij}$ $\rightsquigarrow$ *path matrix* $M$



# Warm-up: the Gessel-Viennot-Lindström Lemma




**Setting**

$\Gamma$ a directed acyclic graph (DAG) on $[m]$

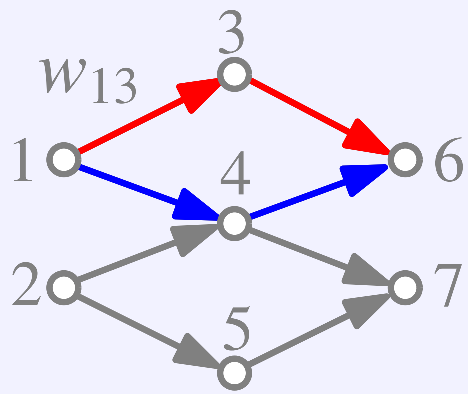
for $i \to j$ a weight $w_{ij}$ $\rightsquigarrow$ *path matrix* $M$

**Example:** $m_{16} = w_{13}w_{36} + w_{14}w_{46} = \sum_{\text{path } p: i \to j} w(p)$

**Setting**

$\Gamma$ a directed acyclic graph (DAG) on $[m]$

for $i \to j$ a weight $w_{ij}$ $\rightsquigarrow$ *path matrix* $M$




**Example:** $m_{16} = w_{13}w_{36} + w_{14}w_{46} = \sum_{\text{path } p: i \to j} w(p)$

**Lemma:** $S, S' \subseteq [m]$, $|S| = |S'| \rightsquigarrow$

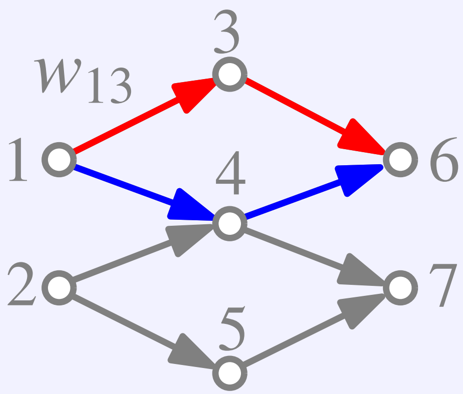
$\det M[S, S'] = \sum_{\pi: S \to S'} \operatorname{sgn}(\pi) \prod_{i \in S} m_{i,\pi(i)}$

$= \sum_{\text{path system } P: S \to S'} \operatorname{sgn}(P) w(P)$

**Setting**

$\Gamma$ a directed acyclic graph (DAG) on $[m]$

for $i \to j$ a weight $w_{ij} \rightsquigarrow$ *path matrix* $M$




**Example:** $m_{16} = w_{13}w_{36} + w_{14}w_{46} = \sum_{\text{path } p:i\to j} w(p)$

**Lemma:** $S, S' \subseteq [m]$, $|S| = |S'| \rightsquigarrow$

$\det M[S,S'] = \sum_{\pi:S\to S'} \operatorname{sgn}(\pi) \prod_{i\in S} m_{i,\pi(i)}$

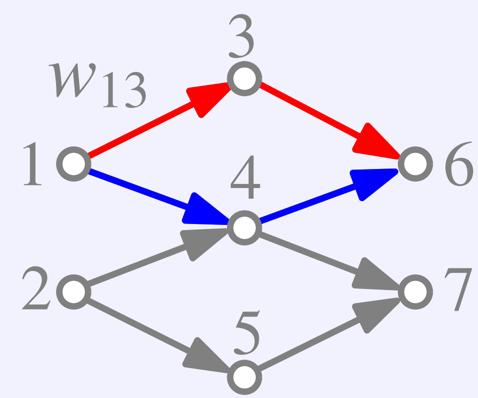
$= \sum_{\text{path system } P:S\to S'} \operatorname{sgn}(P)w(P)$ $= \sum_{\text{vertex disjoint } P:I\to J} \operatorname{sgn}(P)w(P)$!

# Warm-up: the Gessel-Viennot-Lindström Lemma

**Setting**

$\Gamma$ a directed acyclic graph (DAG) on $[m]$

for $i \to j$ a weight $w_{ij}$ $\rightsquigarrow$ *path matrix* $M$




**Example:** $m_{16} = w_{13}w_{36} + w_{14}w_{46} = \sum_{\text{path } p:i\to j} w(p)$

**Lemma:** $S, S' \subseteq [m]$, $|S| = |S'|$ $\rightsquigarrow$

$\det M[S,S'] = \sum_{\pi:S\to S'} \operatorname{sgn}(\pi) \prod_{i\in S} m_{i,\pi(i)}$

$= \sum_{\text{path system } P:S\to S'} \operatorname{sgn}(P) w(P)$ $= \sum_{\text{vertex disjoint } P:I\to J} \operatorname{sgn}(P) w(P)$!

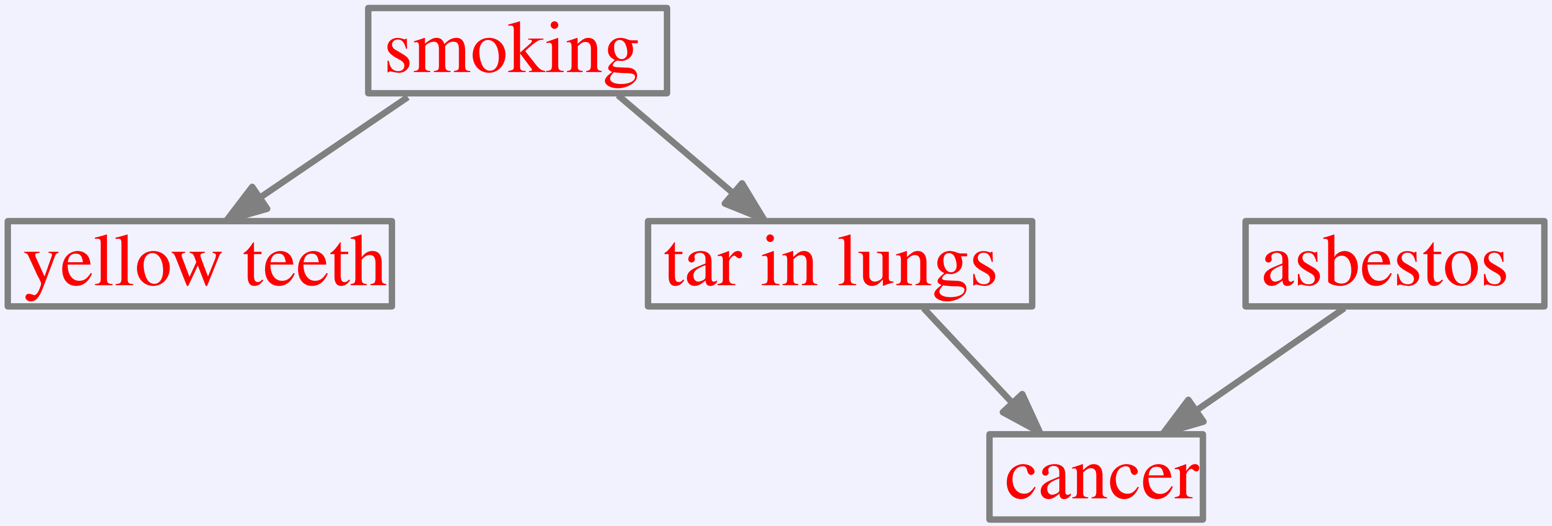
**Example:** $S = \{1,2\}, S' = \{6,7\} \rightsquigarrow \det(M[S,S']) =$

$(w_{13}w_{36})(w_{24}w_{47}) + (w_{13}w_{36})(w_{25}w_{57}) + (w_{14}w_{46})(w_{25}w_{57})$;

*the path systems intersecting at* 4 *cancel out!*

# A hypothetical DAG model

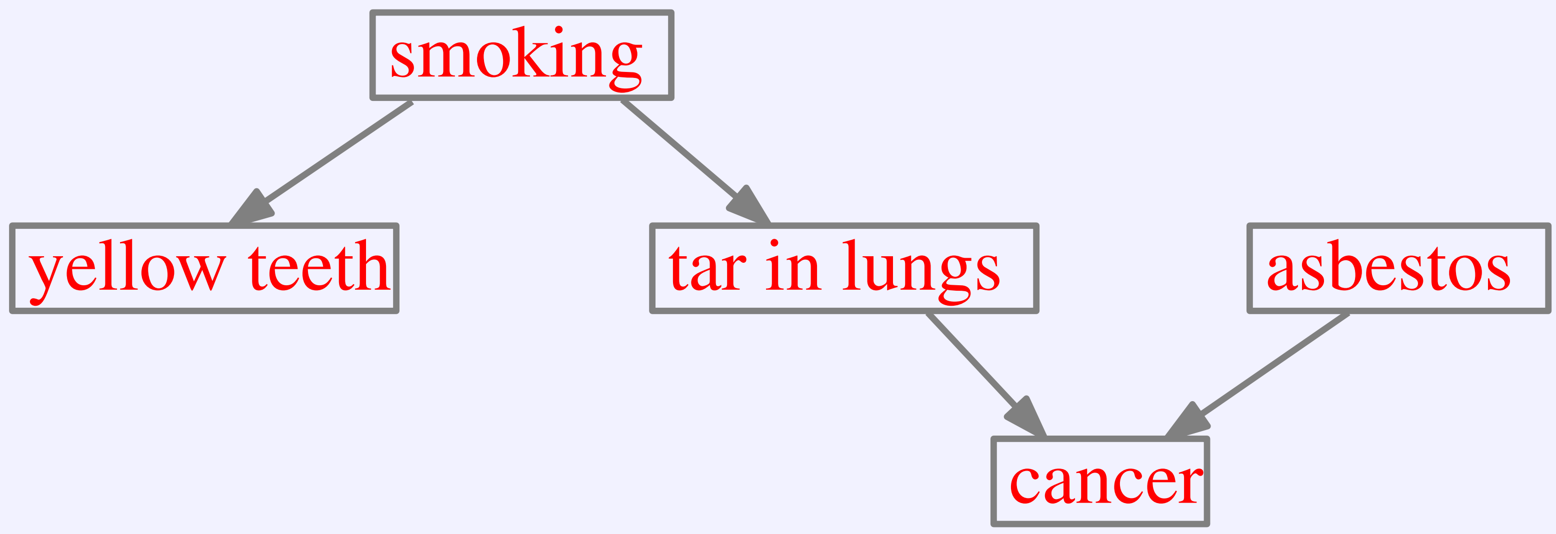





(Shalizi, 2018)

# A hypothetical DAG model






(Shalizi, 2018)

**Interpretation:**

- arrows indicate causal relationships
- graph is directed, acyclic (DAG)
- graph implies CI statements such as:

*yellow teeth and tar in lungs are independent given smoking*

# Directed Gaussian graphical models



$G = ([n], D)$ a DAG, write $i \to j$ for $(i, j) \in D$
$X_1, \ldots, X_n$: jointly Gaussian random variables

**Relations**
$X_j = \sum_{i \to j} a_{ij} X_i + \epsilon_j, \quad \epsilon \sim \mathcal{N}(0, \mathrm{diag}(\omega_1, \ldots, \omega_n) = \Omega)$

# Directed Gaussian graphical models



$G = ([n], D)$ a DAG, write $i \to j$ for $(i, j) \in D$
$X_1, \ldots, X_n$: jointly Gaussian random variables

**Relations**

$X_j = \sum_{i \to j} a_{ij} X_i + \epsilon_j, \quad \epsilon \sim \mathcal{N}(0, \mathrm{diag}(\omega_1, \ldots, \omega_n) = \Omega)$

**Parameter space**

$(a, \omega) \in \mathbb{R}^D \times \mathbb{R}^n_{>0}$

# Directed Gaussian graphical models



$G = ([n], D)$ a DAG, write $i \to j$ for $(i, j) \in D$
$X_1, \ldots, X_n$: jointly Gaussian random variables

**Relations**
$X_j = \sum_{i \to j} a_{ij} X_i + \epsilon_j, \quad \epsilon \sim \mathcal{N}(0, \mathrm{diag}(\omega_1, \ldots, \omega_n) = \Omega)$

**Parameter space**
$(a, \omega) \in \mathbb{R}^D \times \mathbb{R}^n_{>0}$

**Covariance matrix of $X$**
$(I - A)^T X = \epsilon$, so $X = (I - A)^{-T} \epsilon$, so $\Sigma = (I - A)^{-T} \Omega (I - A)^{-1}$

Here $a_{ij} = 0$ if $i \not\to j$, so $A$ is nilpotent and
$(I - A)^{-1} = I + A + A^2 + \cdots + A^{n-1}$.

# Trek expansion



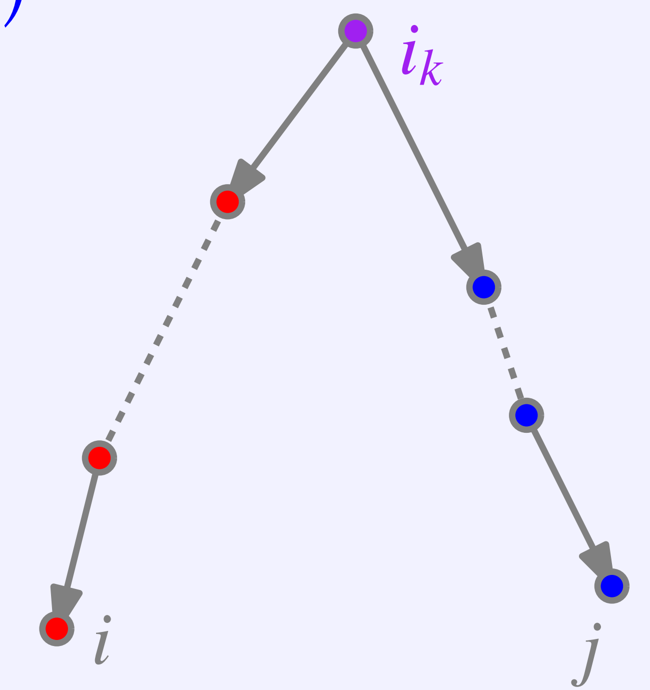
$$\Sigma = (I + A^T + (A^T)^2 + \cdots + (A^T)^{n-1}) \cdot \Omega \cdot (I + A + A^2 + \cdots + A^{n-1})$$

$\rightsquigarrow \Sigma_{ij} = \sum_{\text{treks } t: i \to j} w(t)$, where:

**Definition**

A *trek* $t : i \to j$ is a pair $[\underbrace{(i = i_0, i_1, \ldots, i_k)}_{\text{up}}, \underbrace{(i_k, i_{k+1}, \ldots, i_l = j)}_{\text{down}}]$,

$w(t) := \left(\prod_{j=1}^{k} a_{i_j, i_{j-1}}\right) \cdot \omega_{i_k} \cdot \left(\prod_{j=k}^{l-1} a_{i_j, i_{j+1}}\right)$.



# Trek expansion

$$\Sigma = (I + A^T + (A^T)^2 + \cdots + (A^T)^{n-1}) \cdot \Omega \cdot (I + A + A^2 + \cdots + A^{n-1})$$

$\rightsquigarrow \Sigma_{ij} = \sum_{\text{treks } t: i \to j} w(t)$, where:

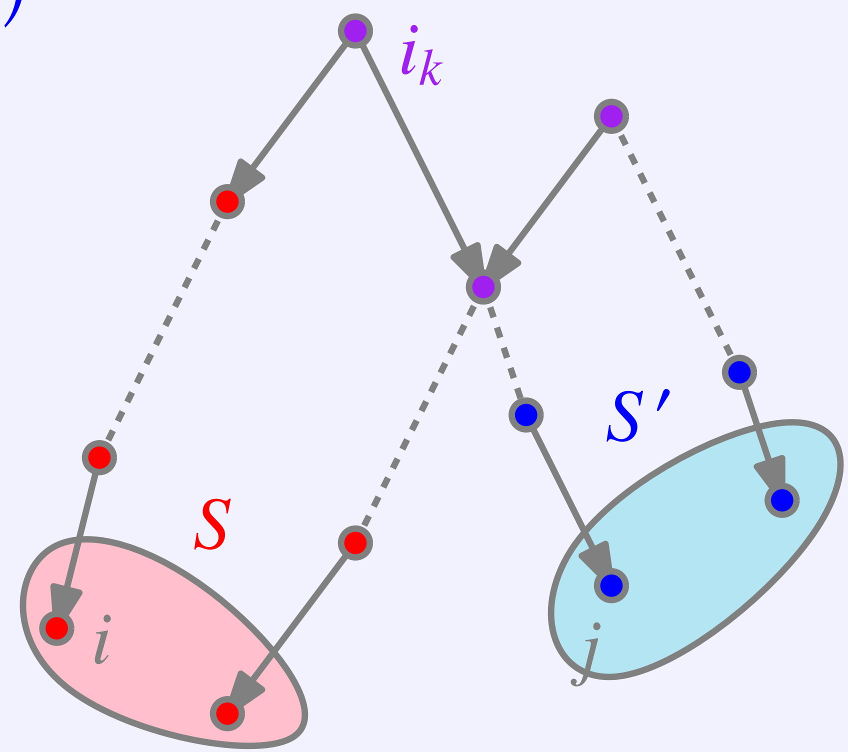
**Definition**

A *trek* $t : i \to j$ is a pair $[\underbrace{(i = i_0, i_1, \ldots, i_k)}_{\text{up}}, \underbrace{(i_k, i_{k+1}, \ldots, i_l = j)}_{\text{down}}]$,

$w(t) := \left(\prod_{j=1}^{k} a_{i_j, i_{j-1}}\right) \cdot \omega_{i_k} \cdot \left(\prod_{j=k}^{l-1} a_{i_j, i_{j+1}}\right)$.

**Theorem (Sullivant-Talaska-D)**

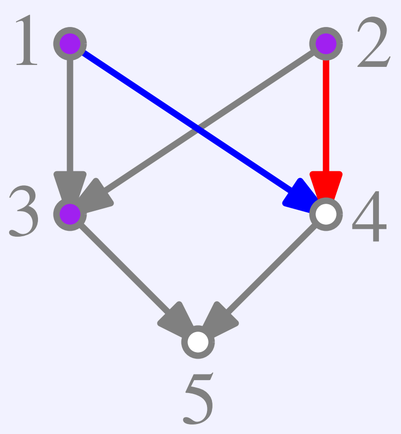
Let $S, S' \subseteq [n]$ with $|S| = |S'|$. Then $\det \Sigma[S, S'] = \sum_{T: S \to S'} \operatorname{sgn}(T) w(T)$, where the sum is over trek systems $S \to S'$ *without sided (red or blue) intersections*. Formula is cancellation-free with coefficients $\pm 2^x$.




*(Apply GVL to $G^{\mathrm{op}} \to G$.)*

# Example (Lin-Uhler-Sturmfels-Bühlmann)

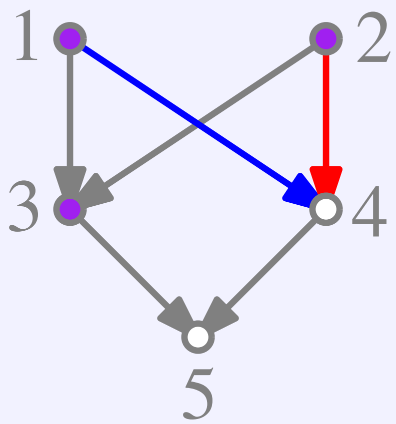





$$\det \Sigma[134, 234] = -(\omega_1 a_{14})(\omega_3)(a_{24}\omega_2) - (\omega_1 a_{13})(a_{23}\omega_2)(\omega_4)$$

# Example (Lin-Uhler-Sturmfels-Bühlmann)



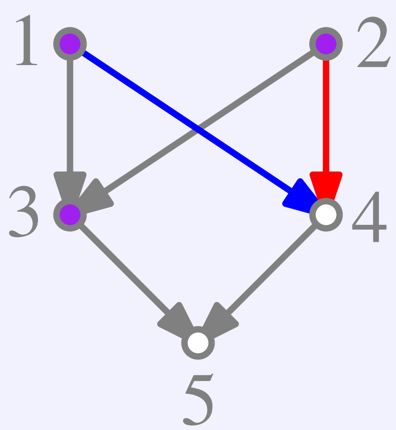



$\det \Sigma[134, 234] = -(\omega_1 a_{14})(\omega_3)(a_{24}\omega_2) - (\omega_1 a_{13})(a_{23}\omega_2)(\omega_4)$

$\det \Sigma[15, 25] = -(\omega_1 a_{13} a_{35})(a_{35} a_{23} \omega_2) - (\omega_1 a_{14} a_{45})(a_{35} a_{23} \omega_2) - (\omega_1 a_{13} a_{35})(a_{45} a_{24} \omega_2) - (\omega_1 a_{14} a_{45})(a_{45} a_{24} \omega_2)$

# Example (Lin-Uhler-Sturmfels-Bühlmann)



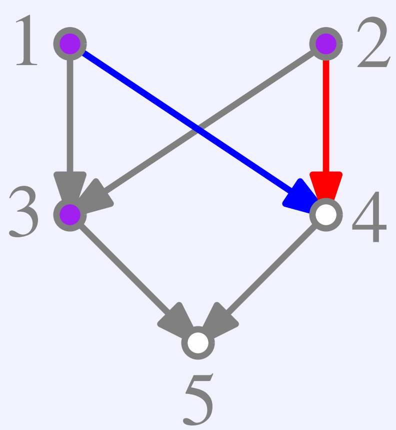



$\det \Sigma[134, 234] = -(\omega_1 a_{14})(\omega_3)(a_{24}\omega_2) - (\omega_1 a_{13})(a_{23}\omega_2)(\omega_4)$

$\det \Sigma[15, 25] = -(\omega_1 a_{13} a_{35})(a_{35} a_{23} \omega_2) - (\omega_1 a_{14} a_{45})(a_{35} a_{23} \omega_2) - (\omega_1 a_{13} a_{35})(a_{45} a_{24} \omega_2) - (\omega_1 a_{14} a_{45})(a_{45} a_{24} \omega_2)$

$\det \Sigma[134, 345] \equiv 0$

# Example (Lin-Uhler-Sturmfels-Bühlmann)




$\det \Sigma[134, 234] = -(\omega_1 a_{14})(\omega_3)(a_{24}\omega_2) - (\omega_1 a_{13})(a_{23}\omega_2)(\omega_4)$

$\det \Sigma[15, 25] = -(\omega_1 a_{13} a_{35})(a_{35} a_{23} \omega_2) - (\omega_1 a_{14} a_{45})(a_{35} a_{23} \omega_2) - (\omega_1 a_{13} a_{35})(a_{45} a_{24} \omega_2) - (\omega_1 a_{14} a_{45})(a_{45} a_{24} \omega_2)$

$\det \Sigma[134, 345] \equiv 0$

**Corollary of S-T-D**
$\det \Sigma[S, S']$ is not identically zero on $\mathbb{R}^D \times \mathbb{R}^n_{>0}$ if and only if there exists a trek system $T : S \to S'$ without sided intersection.

**Partial correlations:** $S \subseteq [n], i_0, j_0 \in [n] \setminus S$
$\rightsquigarrow \mathrm{corr}(i_0, j_0|S) := \dfrac{\det\Sigma[S+i_0, S+j_0]}{\sqrt{\det\Sigma[S+i_0, S+i_0]\,\det\Sigma[S+j_0, S+j_0]}}$ is the *partial correlation of* $i_0, j_0$ *given* $S$; it is zero iff $X_{i_0}, X_{j_0}$ indep. given $(X_s)_{s\in S}$.

**Partial correlations:** $S \subseteq [n], i_0, j_0 \in [n] \setminus S$
$\rightsquigarrow \mathrm{corr}(i_0, j_0|S) := \frac{\det \Sigma[S+i_0, S+j_0]}{\sqrt{\det \Sigma[S+i_0, S+i_0] \det \Sigma[S+j_0, S+j_0]}}$ is the *partial correlation of* $i_0, j_0$ *given* $S$; it is zero iff $X_{i_0}, X_{j_0}$ indep. given $(X_s)_{s \in S}$.

**(Consequence of the) Hammersley-Clifford Theorem**
The model $\{\Sigma \mid (a, \omega)\} \subseteq \mathbb{R}^{n \times n}$ is uniquely determined by the partial correlations that vanish identically on it.

**Remark:** Two distinct DAGs can yield the same model (*Markov equivalence*); cannot be distinguished with observational data.

**Exercise 1:** The following are equivalent:

1. $\det \Sigma[S + i_0, S + j_0] \equiv 0$;
2. there exists $C_{\text{up}}, C_{\text{down}} \subseteq [m]$ with $|C_{\text{up}}| + |C_{\text{down}}| < |S| + 1$ such that every trek from $S + i_0$ to $S + j_0$ passes through $C_{\text{up}}$ on its way up or through $C_{\text{down}}$ on its way down;
3. $S$ *d-separates* $i_0$ from $j_0$: every undirected path from $i_0$ to $j_0$ either has a *noncollider* in $S$ or else has a *collider* $\rightarrow c \leftarrow$ such that $c$ and its descendents do not belong to $S$. (Yangming Di)

**Exercise 1:** The following are equivalent:
1. $\det \Sigma[S + i_0, S + j_0] \equiv 0$;
2. there exists $C_{\text{up}}, C_{\text{down}} \subseteq [m]$ with $|C_{\text{up}}| + |C_{\text{down}}| < |S| + 1$ such that every trek from $S + i_0$ to $S + j_0$ passes through $C_{\text{up}}$ on its way up or through $C_{\text{down}}$ on its way down;
3. $S$ *d-separates* $i_0$ from $j_0$: every undirected path from $i_0$ to $j_0$ either has a *noncollider* in $S$ or else has a *collider* $\to c \leftarrow$ such that $c$ and its descendents do not belong to $S$. (Yangming Di)

**Exercise 2:** if $A$ is a square matrix, then $\det A[S, S'] = \pm \det(A^{-1})[\overline{S'}, \overline{S}] \cdot \det(A)$. (So subdeterminants of $\Sigma$ are also subdeterminants of the *concentration matrix* $\Sigma^{-1}$.)

**Exercise 1:** The following are equivalent:
1. $\det \Sigma[S + i_0, S + j_0] \equiv 0$;
2. there exists $C_{\mathrm{up}}, C_{\mathrm{down}} \subseteq [m]$ with $|C_{\mathrm{up}}| + |C_{\mathrm{down}}| < |S| + 1$ such that every trek from $S + i_0$ to $S + j_0$ passes through $C_{\mathrm{up}}$ on its way up or through $C_{\mathrm{down}}$ on its way down;
3. $S$ *d-separates* $i_0$ from $j_0$: every undirected path from $i_0$ to $j_0$ either has a *noncollider* in $S$ or else has a *collider* $\to c \leftarrow$ such that $c$ and its descendents do not belong to $S$. (Yangming Di)

**Exercise 2:** if $A$ is a square matrix, then $\det A[S, S'] = \pm \det(A^{-1})[\overline{S'}, \overline{S}] \cdot \det(A)$. (So subdeterminants of $\Sigma$ are also subdeterminants of the *concentration matrix* $\Sigma^{-1}$.)

**Exercise 3:** Fix $i_0, j_0$. If $\det \Sigma[S + i_0, S + j_0] \equiv 0$ for *some* $S \subseteq [n] \setminus \{i_0, j_0\}$, then this holds for some set $S \subseteq \mathrm{Pa}(i_0)$ or some $S \subseteq \mathrm{Pa}(j_0)$.

# Discovering the graph from $\Sigma$: PC algorithm



Start with undirected $K_n$; $E :=$ {all 2-subsets of $[n]$}.

**Phase 1: edge removal**

For $k = 0, \ldots, n-2$ do: for $\{i_0, j_0\} \in E$ and $S \subseteq [n] \setminus \{i_0, j_0\}$ a set of neighbours of $i_0$ or of $j_0$ with $|S| = k$ compute $\mathrm{corr}(i_0, j_0 | S)$. If zero, delete $\{i_0, j_0\}$ from $S$.

# Discovering the graph from $\Sigma$: PC algorithm 

Start with undirected $K_n$; $E := \{\text{all 2-subsets of } [n]\}$.

**Phase 1: edge removal**
For $k = 0, \ldots, n-2$ do: for $\{i_0, j_0\} \in E$ and $S \subseteq [n] \setminus \{i_0, j_0\}$ a set of neighbours of $i_0$ or of $j_0$ with $|S| = k$ compute $\text{corr}(i_0, j_0|S)$. If zero, delete $\{i_0, j_0\}$ from $S$.

**Phase 2: orientation**
Use partial correlations to try and orient the edges in $E$.

# Discovering the graph from $\Sigma$: PC algorithm 

Start with undirected $K_n$; $E := \{\text{all 2-subsets of } [n]\}$.

**Phase 1: edge removal**
For $k = 0, \ldots, n-2$ do: for $\{i_0, j_0\} \in E$ and $S \subseteq [n] \setminus \{i_0, j_0\}$ a set of neighbours of $i_0$ or of $j_0$ with $|S| = k$ compute $\mathrm{corr}(i_0, j_0 | S)$. If zero, delete $\{i_0, j_0\}$ from $S$.

**Phase 2: orientation**
Use partial correlations to try and orient the edges in $E$.

**Theorem (Spirtes-Glymour-Scheines)**
If $\Sigma$ is *faithful* to a *single* DAG, then it is found. If it is faithful to a single Markov equivalence class of DAGs, then that is found.

# Discovering the graph from $\Sigma$: PC algorithm 

Start with undirected $K_n$; $E := \{\text{all 2-subsets of } [n]\}$.

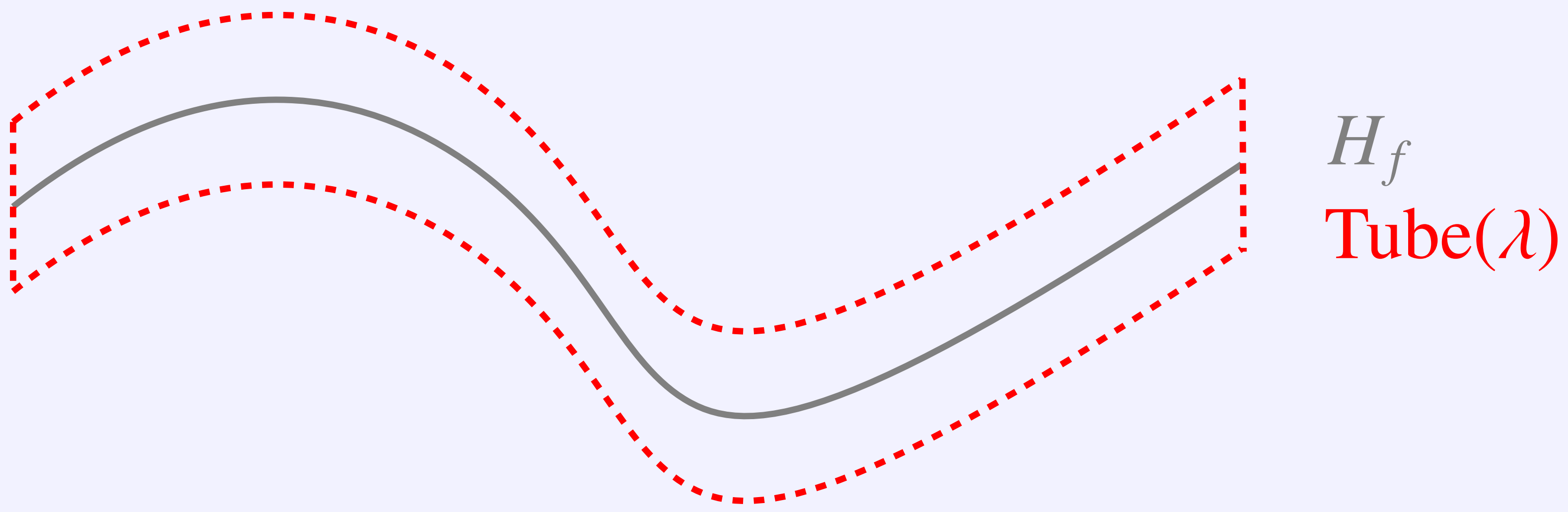
**Phase 1: edge removal**
For $k = 0, \ldots, n-2$ do: for $\{i_0, j_0\} \in E$ and $S \subseteq [n] \setminus \{i_0, j_0\}$ a set of neighbours of $i_0$ or of $j_0$ with $|S| = k$ compute $\text{corr}(i_0, j_0|S)$. If zero, delete $\{i_0, j_0\}$ from $S$.

**Phase 2: orientation**
Use partial correlations to try and orient the edges in $E$.

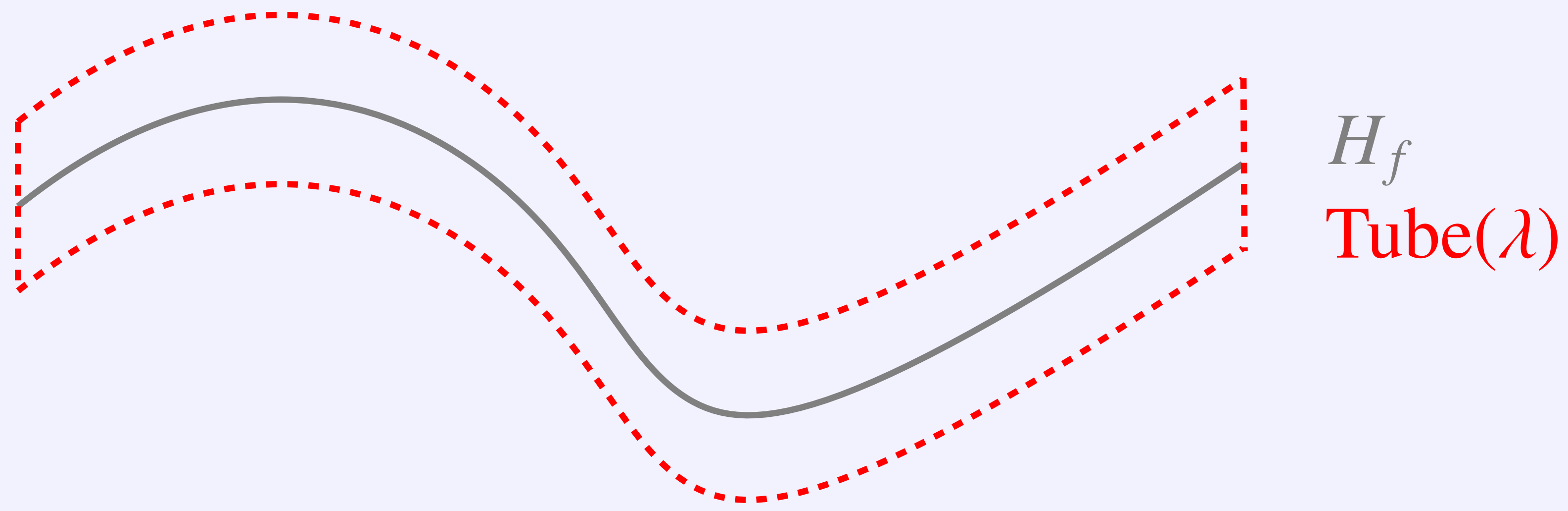
**Theorem (Spirtes-Glymour-Scheines)**
If $\Sigma$ is *faithful* to a *single* DAG, then it is found. If it is faithful to a single Markov equivalence class of DAGs, then that is found.

*In practice, "if zero" is replaced by "if* $|\text{corr}| \leq \lambda$*".*

**Geometry studied by Lin-Uhler-Sturmfels-Bühlmann:**
Suppose true DAG $G$ has $i_0 \to j_0$. Set $f := \det \Sigma[S + i_0, S + j_0]$. The PC-test $|\mathrm{corr}(i_0, j_0|S)| \leq \lambda$ describes a neighbourhood Tube($\lambda$) of the hypersurface in $H_f := \{a \mid f = 0\} \subseteq \mathbb{R}^D$ (they take $\Omega = I$).



**Geometry studied by Lin-Uhler-Sturmfels-Bühlmann:**
Suppose true DAG $G$ has $i_0 \to j_0$. Set $f := \det \Sigma[S + i_0, S + j_0]$. The PC-test $|\mathrm{corr}(i_0, j_0|S)| \leq \lambda$ describes a neighbourhood $\mathrm{Tube}(\lambda)$ of the hypersurface in $H_f := \{a \mid f = 0\} \subseteq \mathbb{R}^D$ (they take $\Omega = I$).




$\mathrm{Tube}(\lambda)$ *is the region in parameter space where the PC-test would delete* $\{i_0, j_0\}$ *erroneously. If* $H_f$ *is smooth, then* $\mathrm{Vol}(\mathrm{Tube}(\lambda))$ *(relative to a fixed density on* $\mathbb{R}^D$*) is linear in* $\lambda$ *for* $\lambda \to 0$. *If not, it may be superlinear.*

**Theorem (Lin-Uhler-Sturmfels-Bühlman)**
As $\lambda \to 0$, $\mathrm{Vol}(\mathrm{Tube}(\lambda)) \approx C\lambda^{\ell}(-\ln\lambda)^{m-1}$ for some $C > 0$ and $(\ell, m)$ the *real log canonical threshold* of $f$.

**Theorem (Lin-Uhler-Sturmfels-Bühlman)**
As $\lambda \to 0$, $\mathrm{Vol}(\mathrm{Tube}(\lambda)) \approx C\lambda^{\ell}(-\ln\lambda)^{m-1}$ for some $C > 0$ and $(\ell, m)$ the *real log canonical threshold* of $f$.

They computed $(\ell, m)$ for many tuples $(G, i_0, j_0, S)$. In particular:

**Theorem (L-U-S-B)**
If $G$ is a complete DAG on $n \leq 6$ vertices, then $H_f$ is smooth for all choices of $i_0, j_0, S$.

**Proof method:** find a power of $\det(\Sigma[S + i_0 + j_0, S + i_0 + j_0])$ in the *Jacobian ideal* $\langle \{f\} \cup \{\frac{\partial f}{\partial a_{ij}} | i \to j\}\rangle$; e.g. 240 cases for $n = 6$.

**Theorem (Lin-Uhler-Sturmfels-Bühlman)**
As $\lambda \to 0$, $\mathrm{Vol}(\mathrm{Tube}(\lambda)) \approx C\lambda^{\ell}(-\ln\lambda)^{m-1}$ for some $C > 0$ and $(\ell, m)$ the *real log canonical threshold* of $f$.

They computed $(\ell, m)$ for many tuples $(G, i_0, j_0, S)$. In particular:

**Theorem (L-U-S-B)**
If $G$ is a complete DAG on $n \leq 6$ vertices, then $H_f$ is smooth for all choices of $i_0, j_0, S$.

**Proof method:** find a power of $\det(\Sigma[S + i_0 + j_0, S + i_0 + j_0])$ in the *Jacobian ideal* $\langle \{f\} \cup \{\frac{\partial f}{\partial a_{ij}} | i \to j\} \rangle$; e.g. 240 cases for $n = 6$.

**Question (L-U-S-B)**
Also for $n > 6$?

# New theorem



**Theorem (D, arXiv: 1806.00320)**
Let $G$ be a DAG on $[n]$, $i_0, j_0 \in [n]$ distinct and $S \subseteq [n] \setminus \{i_0, j_0\}$.
Assume that $i_0 \to j_0$ and that $i_0 \to s$ for each $s \in S$ *below* $j_0$.
Then $H_f := \{a \in \mathbb{R}^D \mid \det \Sigma[S + i_0, S + j_0] = 0\}$ is smooth.

# New theorem

**Theorem (D, arXiv: 1806.00320)**
Let $G$ be a DAG on $[n]$, $i_0, j_0 \in [n]$ distinct and $S \subseteq [n] \setminus \{i_0, j_0\}$.
Assume that $i_0 \to j_0$ and that $i_0 \to s$ for each $s \in S$ *below* $j_0$.
Then $H_f := \{a \in \mathbb{R}^D \mid \det \Sigma[S + i_0, S + j_0] = 0\}$ is smooth.

Treat the $a_{i,j}$ as variables. Let $J$ be the ideal in $\mathbb{R}[a_{ij} \mid i \to j]$ generated by the partial derivatives of $f$. Goal: $V_{\mathbb{R}}(J) = \emptyset$.

**Lemma 1**
The $a_{sj}$, $s \in S$, $j \in [n]$, $s \to j$ don't appear in $f$, so w.l.o.g. $s \not\to j$.

**Theorem (D, arXiv: 1806.00320)**
Let $G$ be a DAG on $[n]$, $i_0, j_0 \in [n]$ distinct and $S \subseteq [n] \setminus \{i_0, j_0\}$. Assume that $i_0 \to j_0$ and that $i_0 \to s$ for each $s \in S$ *below* $j_0$. Then $H_f := \{a \in \mathbb{R}^D \mid \det \Sigma[S + i_0, S + j_0] = 0\}$ is smooth.

Treat the $a_{i,j}$ as variables. Let $J$ be the ideal in $\mathbb{R}[a_{ij} \mid i \to j]$ generated by the partial derivatives of $f$. Goal: $V_{\mathbb{R}}(J) = \emptyset$.

**Lemma 1**
The $a_{sj}$, $s \in S$, $j \in [n]$, $s \to j$ don't appear in $f$, so w.l.o.g. $s \not\to j$.

**Lemma 2**
For $s \in S, i_0 \to s$ the variable $a_{i_0,s}$ appears at most linearly in $f$, with coefficient $\pm \det \Sigma[S + i_0, S + j_0 - s + i_0]$; so this is in $J$.

# New theorem



**Lemma 3**

The variable $a_{i_0,j_0}$ appears at most linearly in $f$, with coefficient $\pm(\det \Sigma[S+i_0, S+i_0] - g)$ where $g = \sum_{T:S+i_0 \to S+i_0} \mathrm{sgn}(T) w(T)$, the sum over $T$, no sided intersections, *passing* $j_0$ *on way down*.

# New theorem



**Lemma 3**

The variable $a_{i_0,j_0}$ appears at most linearly in $f$, with coefficient $\pm(\det \Sigma[S+i_0, S+i_0] - g)$ where $g = \sum_{T:S+i_0 \to S+i_0} \mathrm{sgn}(T) w(T)$, the sum over $T$, no sided intersections, *passing* $j_0$ *on way down*.

Let $p_{j_0,s}$ be the sum of all weights of paths from $j_0$ to $s$.

**Lemma 4**

$g = \sum_{s \in S} \pm \det \Sigma[S+i_0, S+i_0-s+j_0] \cdot p_{j_0,s}$.

**Lemma 3**

The variable $a_{i_0,j_0}$ appears at most linearly in $f$, with coefficient $\pm(\det\Sigma[S+i_0, S+i_0] - g)$ where $g = \sum_{T:S+i_0\to S+i_0} \mathrm{sgn}(T)w(T)$, the sum over $T$, no sided intersections, *passing* $j_0$ *on way down*.

Let $p_{j_0,s}$ be the sum of all weights of paths from $j_0$ to $s$.

**Lemma 4**

$g = \sum_{s\in S} \pm\det\Sigma[S+i_0, S+i_0-s+j_0]\cdot p_{j_0,s}.$

**Proof of theorem:** Modulo $J$ we have

$\det\Sigma[S+i_0, S+i_0] = g = \sum_{s\in S} \pm\Sigma[S+i_0, S+i_0-s+j_0]\cdot p_{j_0,s}.$

**Lemma 3**
The variable $a_{i_0,j_0}$ appears at most linearly in $f$, with coefficient $\pm(\det\Sigma[S+i_0, S+i_0] - g)$ where $g = \sum_{T:S+i_0\to S+i_0} \mathrm{sgn}(T)w(T)$, the sum over $T$, no sided intersections, *passing* $j_0$ *on way down*.

Let $p_{j_0,s}$ be the sum of all weights of paths from $j_0$ to $s$.

**Lemma 4**
$g = \sum_{s\in S} \pm\det\Sigma[S+i_0, S+i_0-s+j_0]\cdot p_{j_0,s}$.

**Proof of theorem:** Modulo $J$ we have
$\det\Sigma[S+i_0, S+i_0] = g = \sum_{s\in S} \pm\Sigma[S+i_0, S+i_0-s+j_0]\cdot p_{j_0,s}$.

This is in $J$ by Lemma 2. Now use that for $a\in\mathbb{R}^D$, $\Sigma[S+i_0, S+i_0]$ is positive definite, hence has nonzero determinant. $\square$

# Discussion: a conjectured volume inequality 

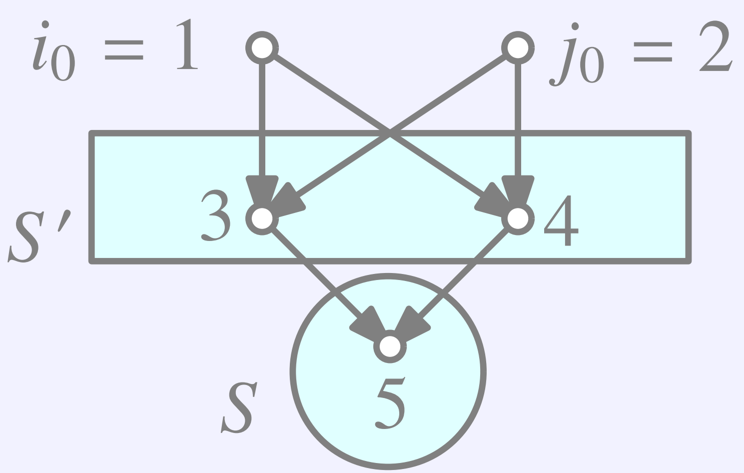
**Setting**

Let $i_0, j_0, S', S$ be disjoint and such that each element of $S'$ is a descendant of $i_0$ and of $j_0$, and each element of $S$ is a descendant of each element of $S'$. L-U-S-B conjecture, for $\lambda \in [0, 1]$:

$\mathrm{Vol}\{a : |\mathrm{corr}(i_0, j_0|S)| \leq \lambda\} \geq \mathrm{Vol}\{a : |\mathrm{corr}(i_0, j_0|S')| \leq \lambda\}$

**Motivation**

*'It is widely believed that collider-stratification bias tends to attenuate when it arises from more extended paths.'*



# Discussion: a conjectured volume inequality



**Setting**

Let $i_0, j_0, S', S$ be disjoint and such that each element of $S'$ is a descendant of $i_0$ and of $j_0$, and each element of $S$ is a descendant of each element of $S'$. L-U-S-B conjecture, for $\lambda \in [0, 1]$:

$\mathrm{Vol}\{a : |\mathrm{corr}(i_0, j_0|S)| \leq \lambda\} \geq \mathrm{Vol}\{a : |\mathrm{corr}(i_0, j_0|S')| \leq \lambda\}$

**Motivation**

*'It is widely believed that collider-stratification bias tends to attenuate when it arises from more extended paths.'*

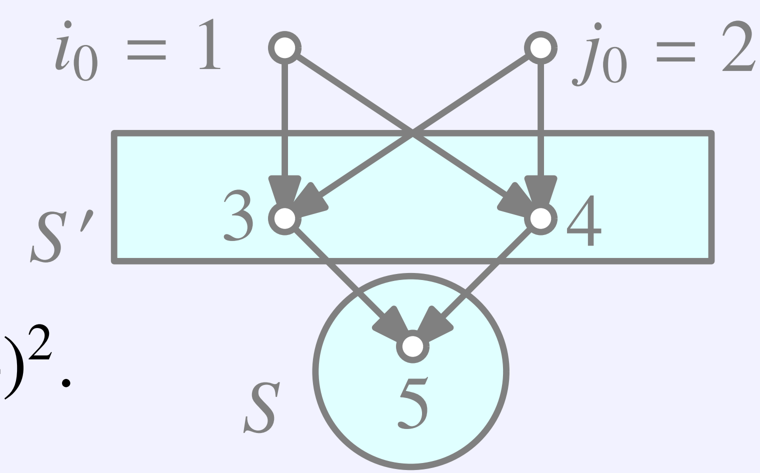
But taking

$a^*_{13} = -3, a^*_{14} = -2,$

$a^*_{23} = 8, a^*_{24} = 10,$

$a^*_{3,5} = 2, a^*_{4,5} = 0$ yields

$\mathrm{corr}(1, 2|5)^2 = \frac{1024}{1189} > \frac{88}{105} = \mathrm{corr}(1, 2|3, 4)^2.$



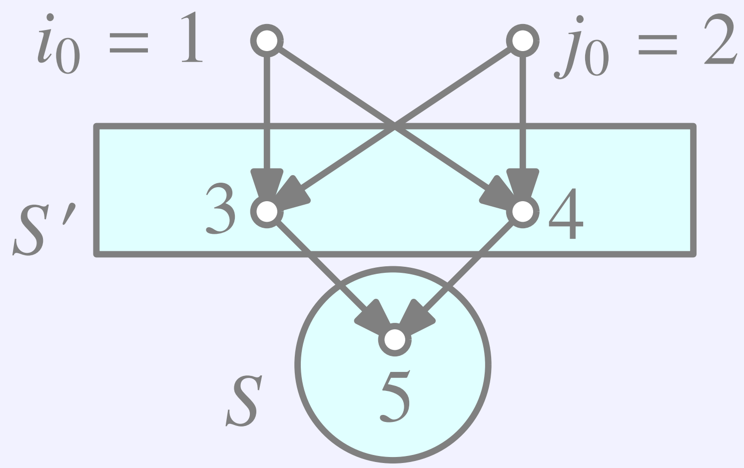
# Discussion: a conjectured volume inequality



$\mathrm{corr}(1,2|5)^2 = \frac{1024}{1189} > \frac{88}{105} = \mathrm{corr}(1,2|3,4)^2$.

So concentrating the mass near this point $a^*$ and taking a suitable $\lambda \in (\sqrt{\frac{88}{105}}, \sqrt{\frac{1024}{1189}})$, we find

$\mathrm{Vol}\{a : |\mathrm{corr}(1,2|5)| < \lambda\} < \mathrm{Vol}\{a : |\mathrm{corr}(1,2|3,4)| < \lambda\}$



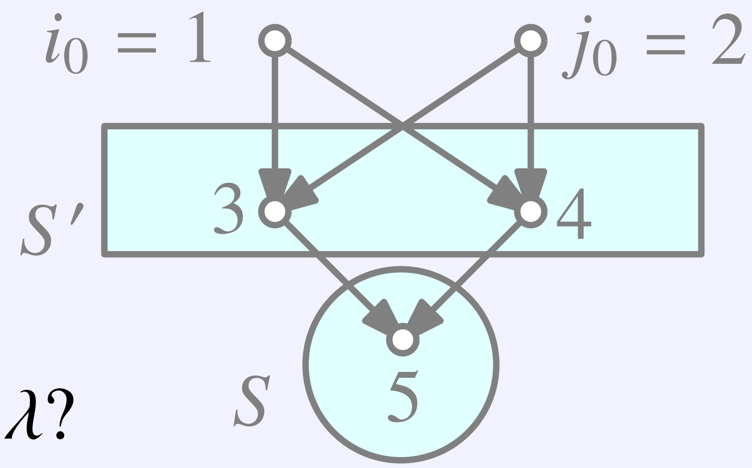
# Discussion: a conjectured volume inequality



$\text{corr}(1, 2|5)^2 = \frac{1024}{1189} > \frac{88}{105} = \text{corr}(1, 2|3, 4)^2.$

So concentrating the mass near this point $a^*$ and taking a suitable $\lambda \in (\sqrt{\frac{88}{105}}, \sqrt{\frac{1024}{1189}})$, we find

$\text{Vol}\{a : |\text{corr}(1, 2|5)| < \lambda\} < \text{Vol}\{a : |\text{corr}(1, 2|3, 4)| < \lambda\}$




**Interesting open problems:**

- after fixing the density, true for small $\lambda$?
- true at least for the real log canonical thresholds?
- more generally, can the rlct be computed directly from the graph? (L-U-S-B solve this, e.g., for trees without colliders.)

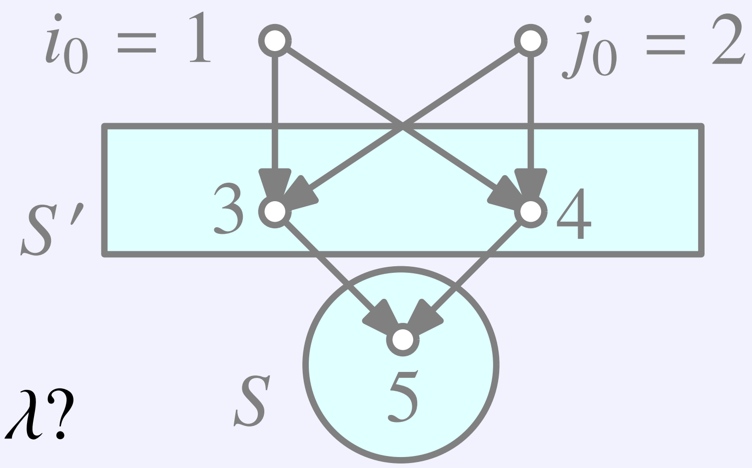
# Discussion: a conjectured volume inequality



$\text{corr}(1,2|5)^2 = \frac{1024}{1189} > \frac{88}{105} = \text{corr}(1,2|3,4)^2$.

So concentrating the mass near this point $a^*$ and taking a suitable $\lambda \in (\sqrt{\frac{88}{105}}, \sqrt{\frac{1024}{1189}})$, we find

$\text{Vol}\{a : |\text{corr}(1,2|5)| < \lambda\} < \text{Vol}\{a : |\text{corr}(1,2|3,4)| < \lambda\}$




**Interesting open problems:**

- after fixing the density, true for small $\lambda$?
- true at least for the real log canonical thresholds?
- more generally, can the rlct be computed directly from the graph? (L-U-S-B solve this, e.g., for trees without colliders.)

THANK YOU!