

Topology of Conditional Causality

A Geometric Post-Analysis of Fujitsu's Causal AI

G-RIPS FUJITSU TEAM

Farris · Misaki · Wu · Jonah

ACADEMIC MENTORS

Dr. Ogawa · Dr. Shinkawa

INDUSTRY MENTORS

Dr. Gutierrez · Dr. Higuchi

Supported By



MathCCS



TOHOKU FORUM
for CREATIVITY



FUJITSU



TOHOKU
UNIVERSITY

ipam



iTHEM.S

Fujitsu · Tohoku University · MathCCS · Tohoku Forum for Creativity · AIMR · IPAM · National Science Foundation · iTHEMS

What is a Causal Graph?



Chocolate eaten

GDP



Nobel prizes



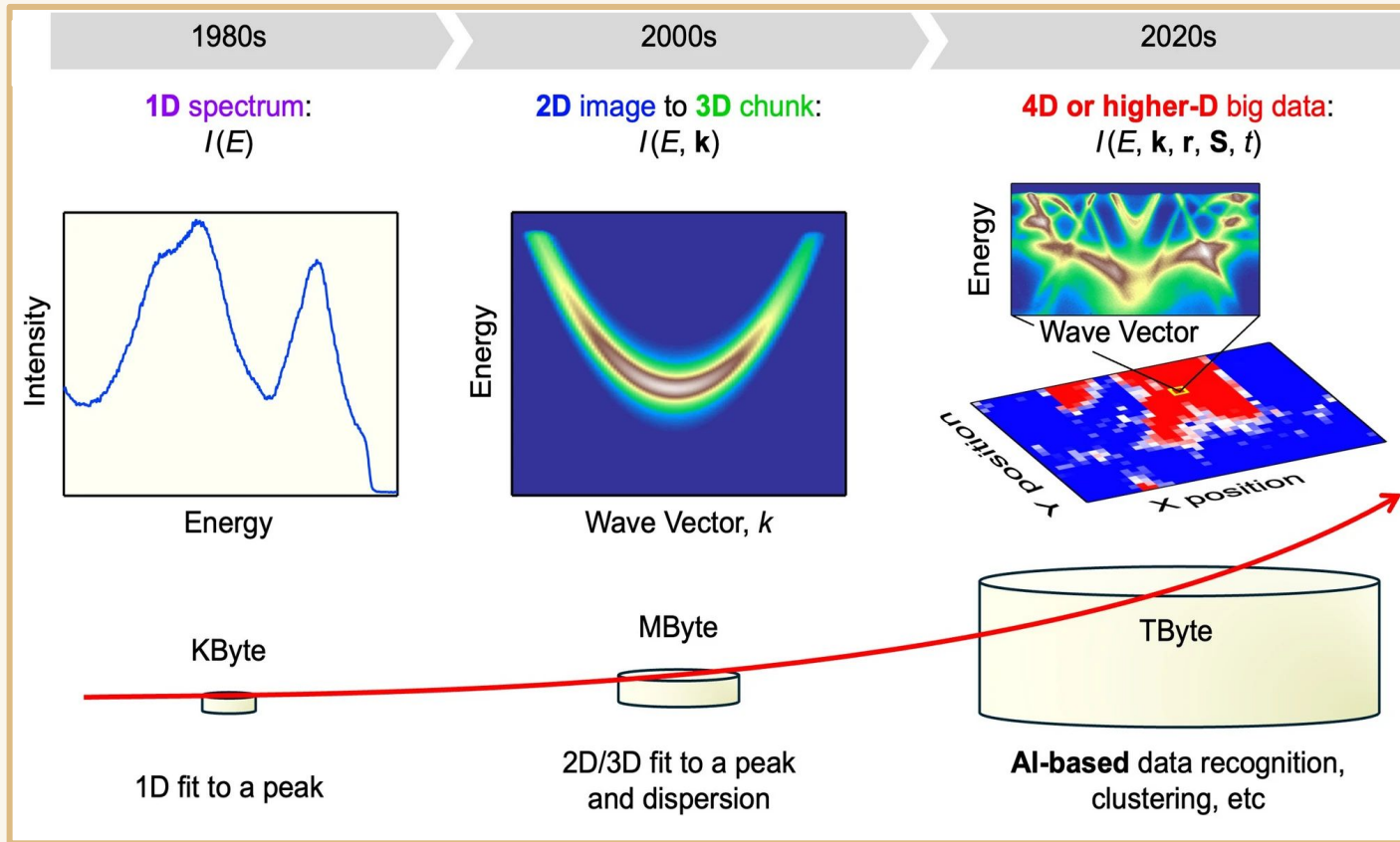
spurious correlation



causality

Background

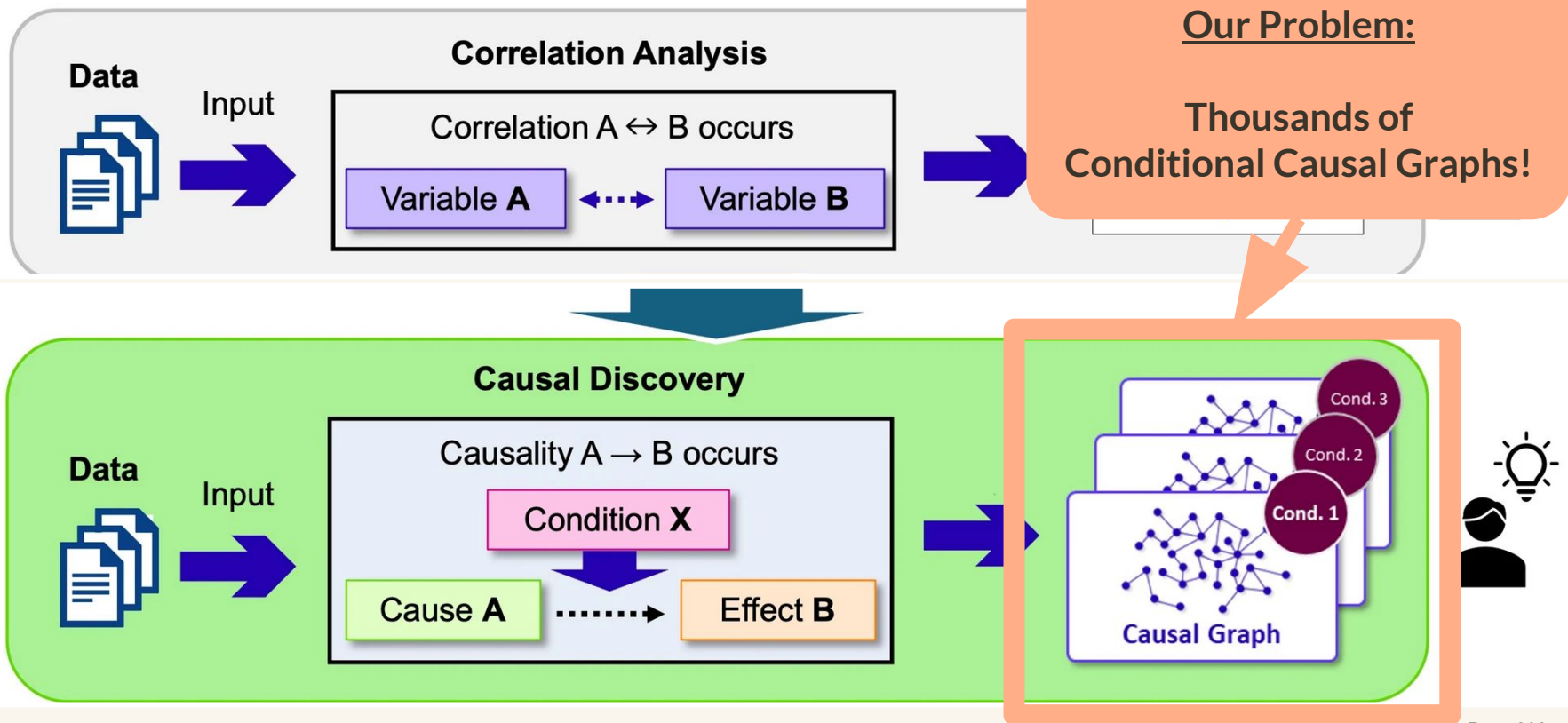
Problem : Data Explosion Massive & High-Dimensional Data



Fujitsu Causal AI

Solution : Causal AI

automatically extract causal structures



Motivation & Core Objective

Transforming Causal Graphs into Reliable Scientific Insights

Three Pillars for AI-Driven Discovery

1.Convincingness

Logical & Simple

2.Comprehensiveness

Unbiased & Exhaustive

3.Discoverability

Beyond Human Intuition

Goal: Enhance Causal AI with a mathematical framework to drive scientific discovery.

Inside Fujitsu's Causal AI

Dataset

input data



Wide Learning

Finds collection
of conditions



Direct LiNGAM

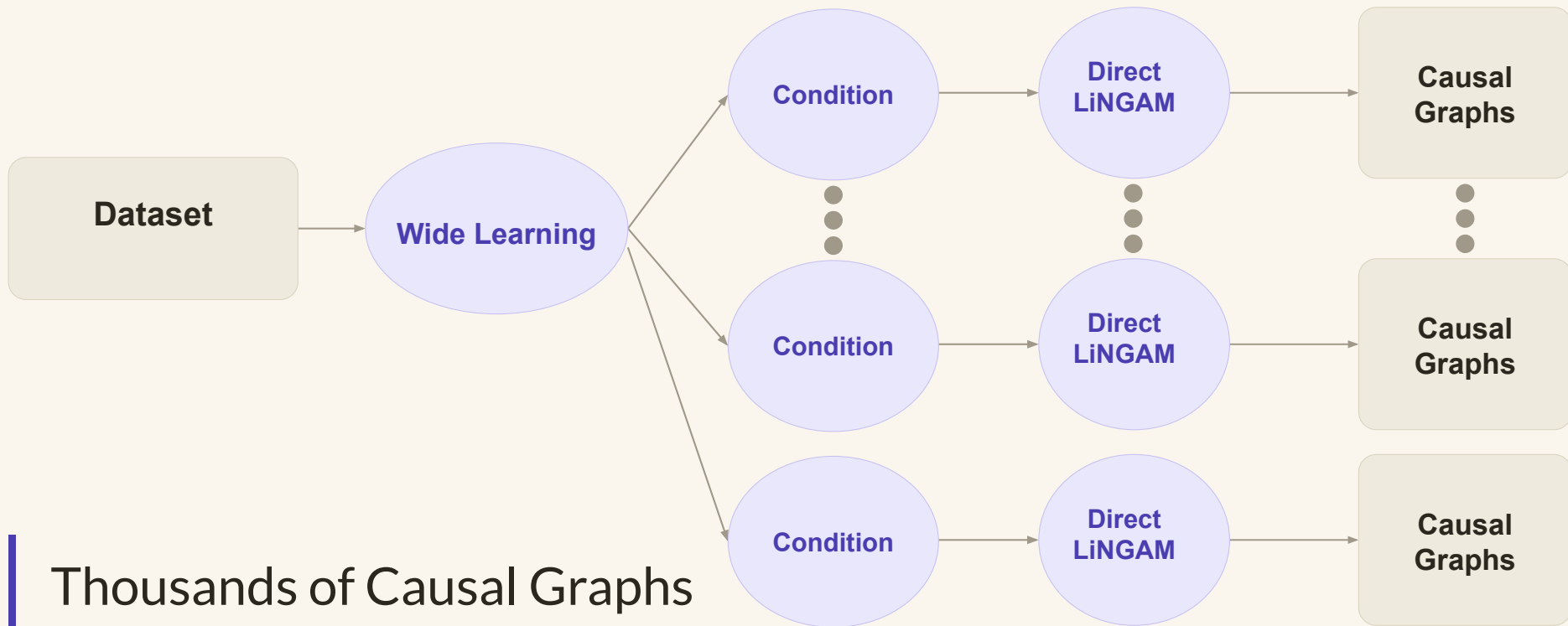
Estimate the direction of
cause and effect



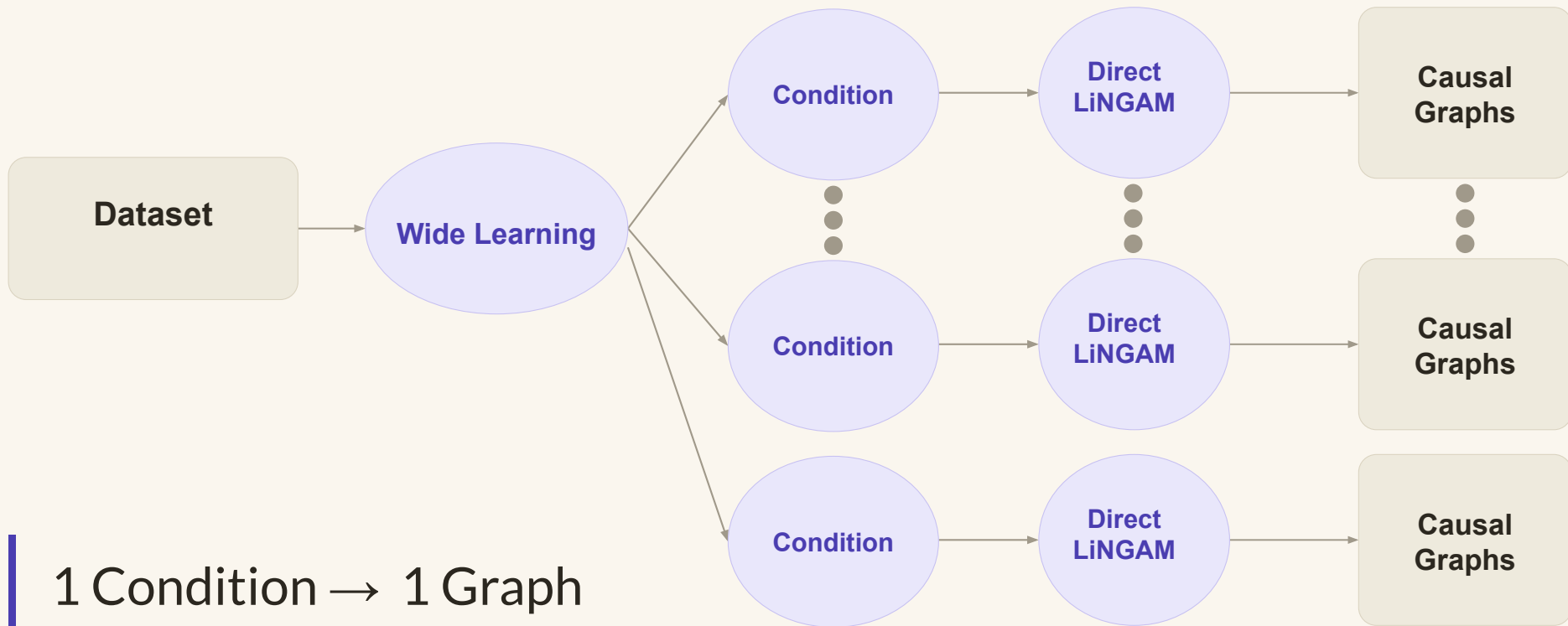
Causal Graphs

Visualizes causal
pathways

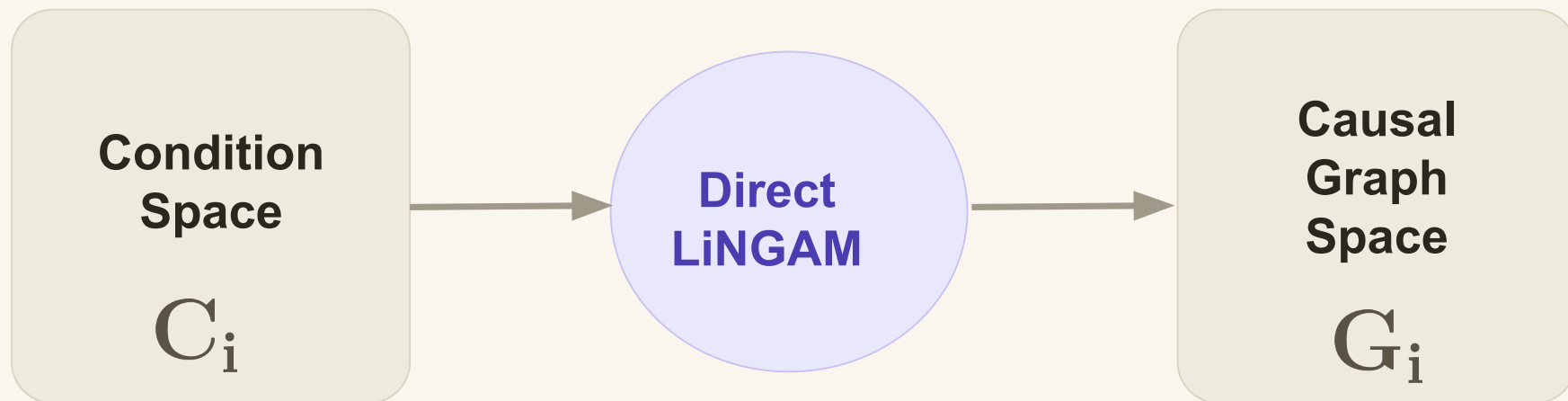
In More Detail



In More Detail



Why Use Metrics?



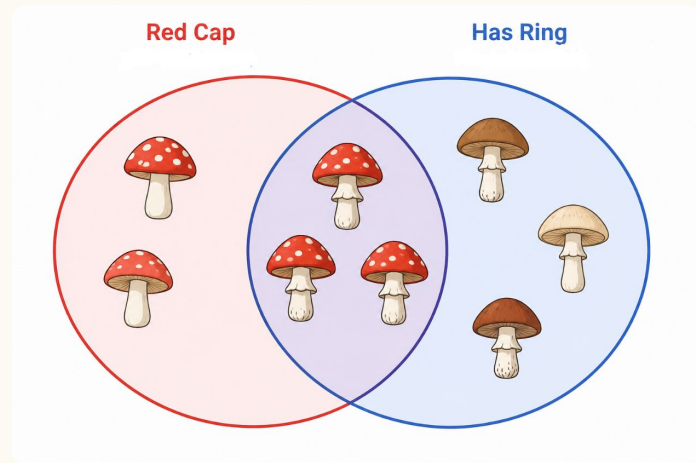
- Goal : analyze (C_i, G_i)
- Need : Metrics

A Metric on the Space of Conditions

The Jaccard Metric

How similar two sets are.

$$d_J(C_1, C_2) = 1 - \frac{|C_1 \cap C_2|}{|C_1 \cup C_2|}$$

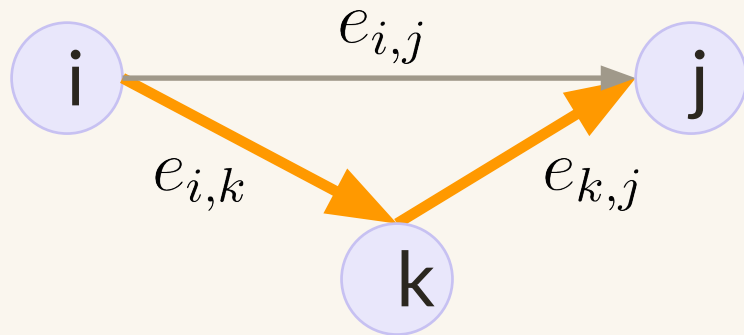
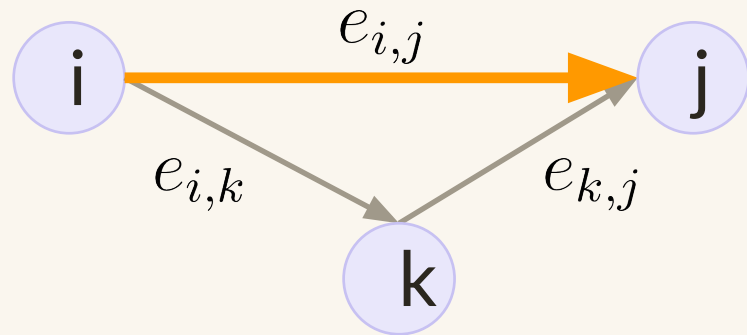


$$d_J(\text{Red}, \text{Ring}) = 1 - \frac{3}{8} = \frac{5}{8}$$

The Total Effect Metric on Causal Graphs

$$\mathbf{R}_{i,j} = \mathbf{e}_{i,j} + \mathbf{e}_{i,k}\mathbf{e}_{k,j}$$

All the ways i cause j

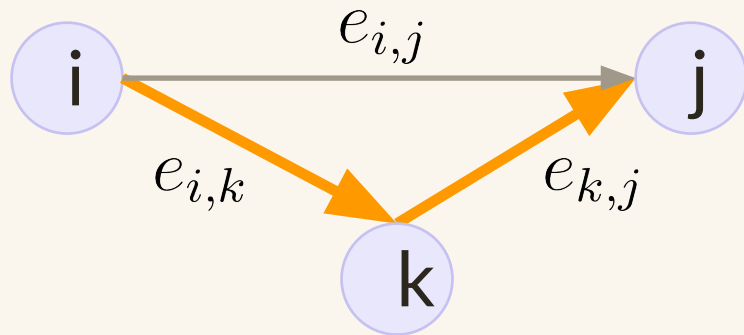
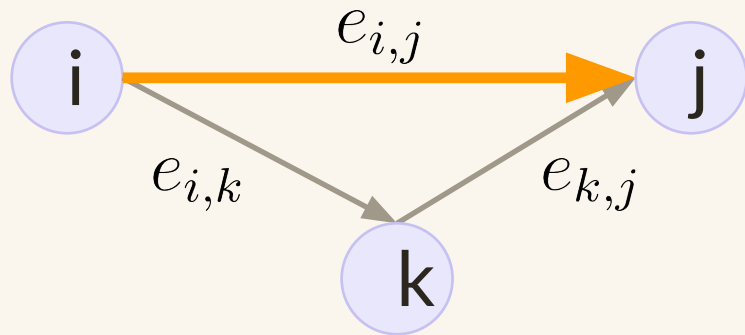


Why Total Effect Metric on Causal Graphs

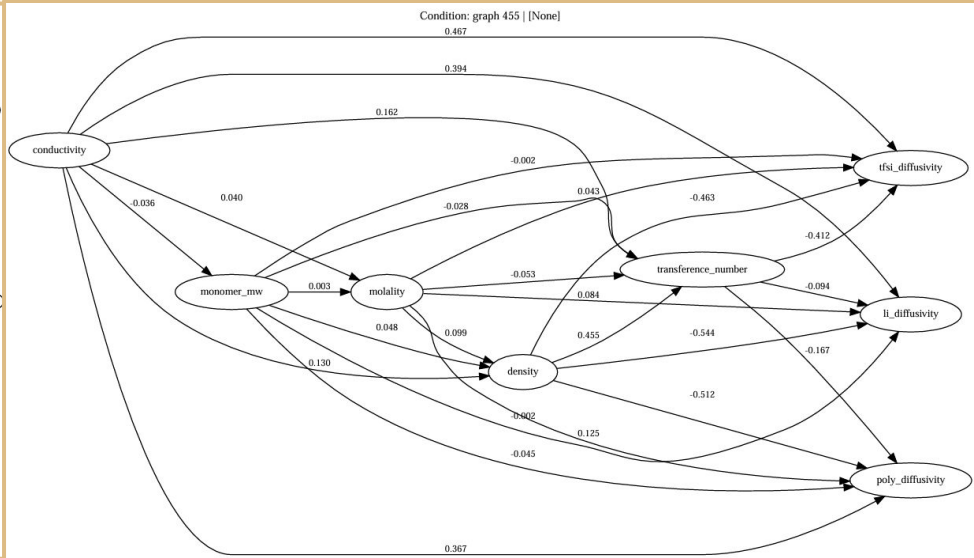
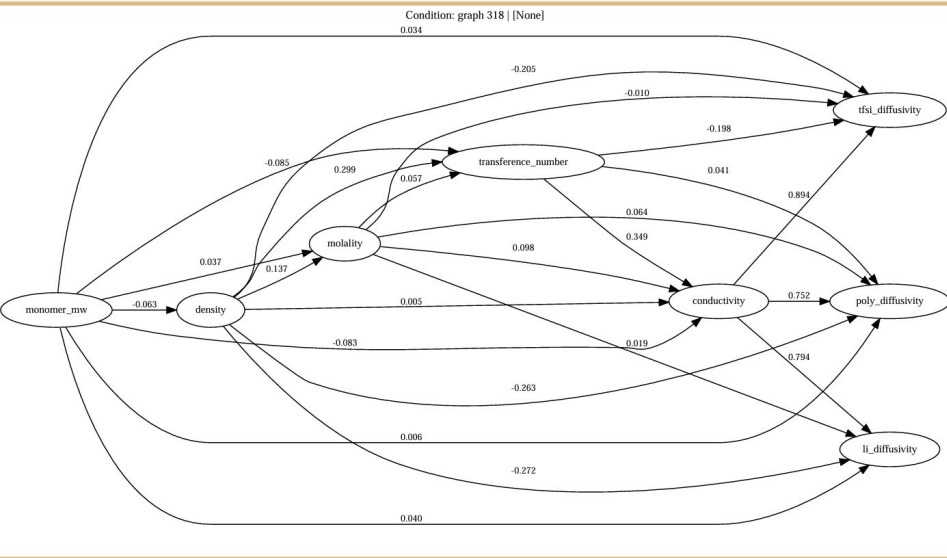
Measures how much influence

i has over **j**

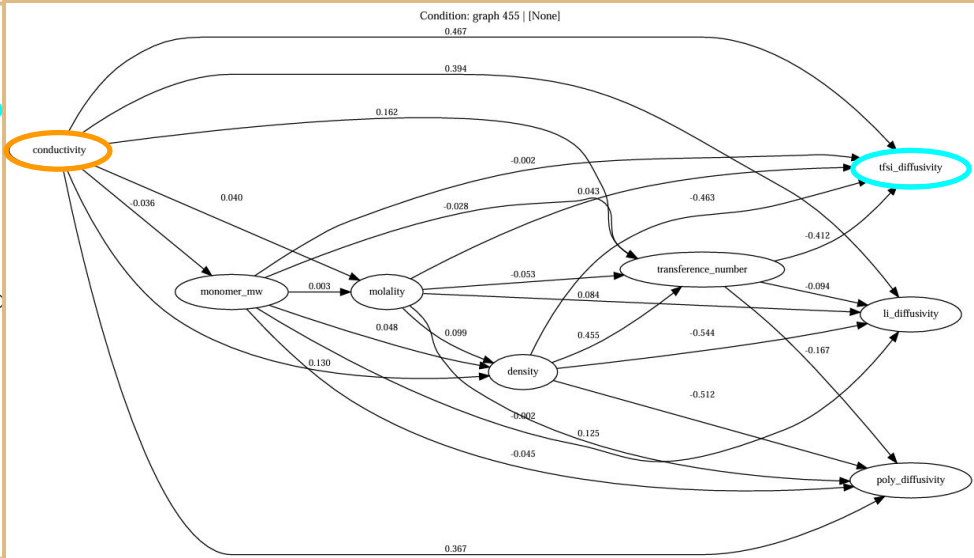
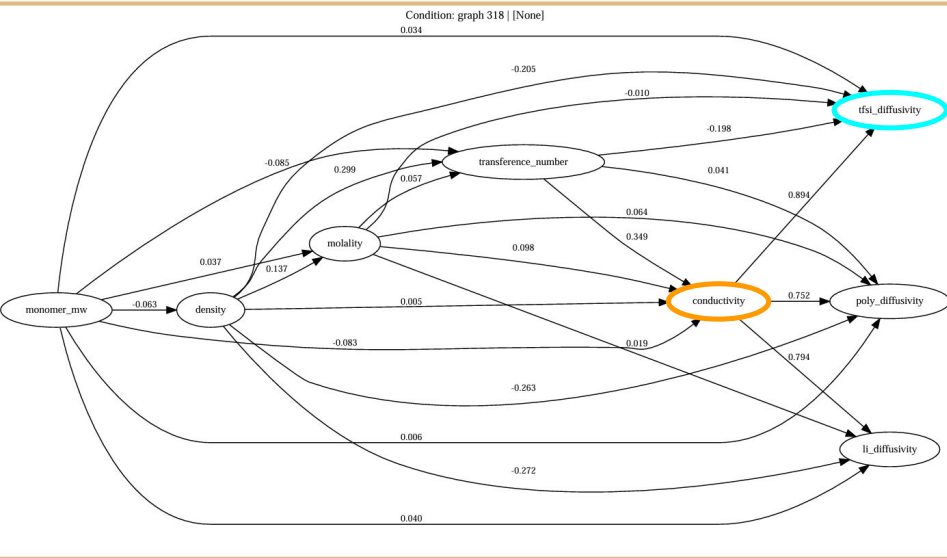
Total Effect matrices provide
comprehensiveness of causality



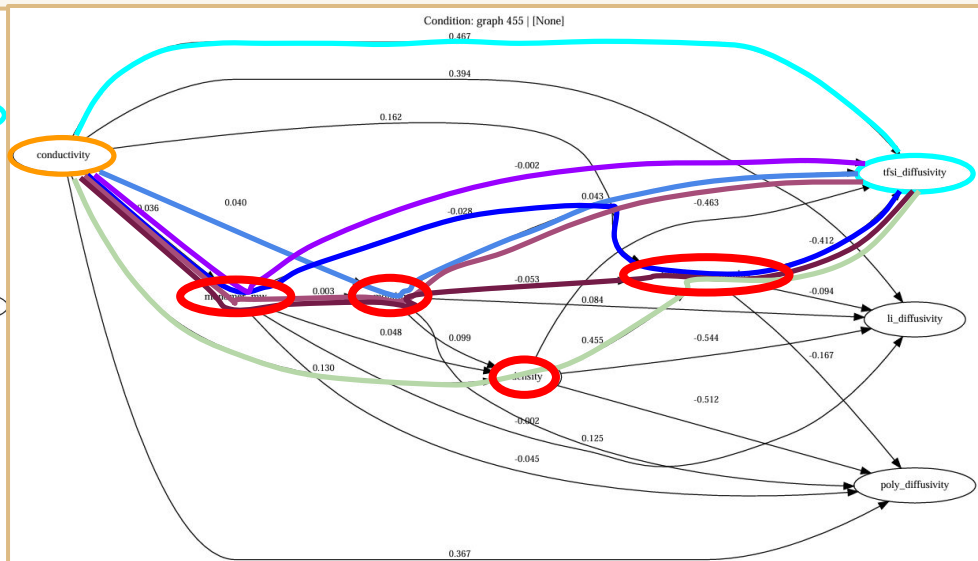
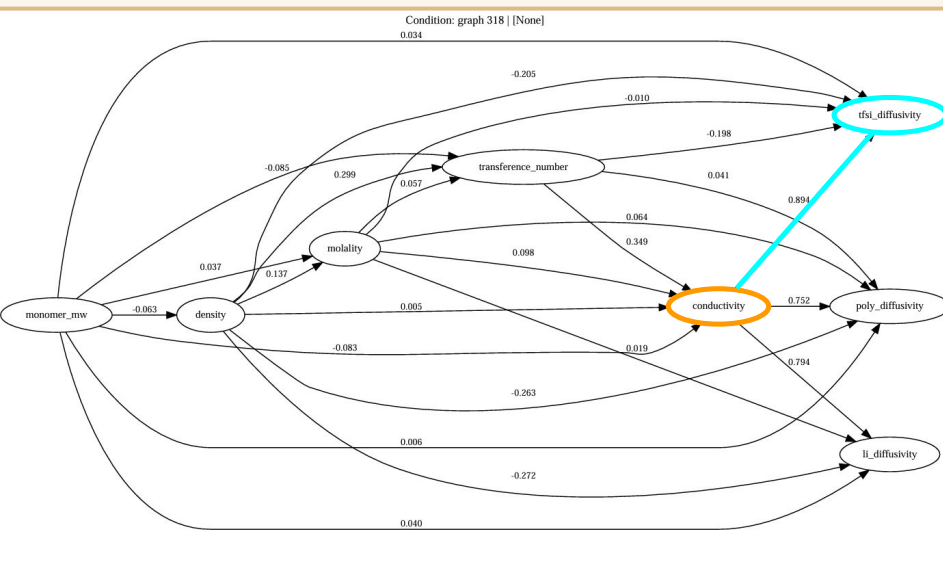
TE and Frobenius metrics



TE and Frobenius metrics



TE and Frobenius metrics

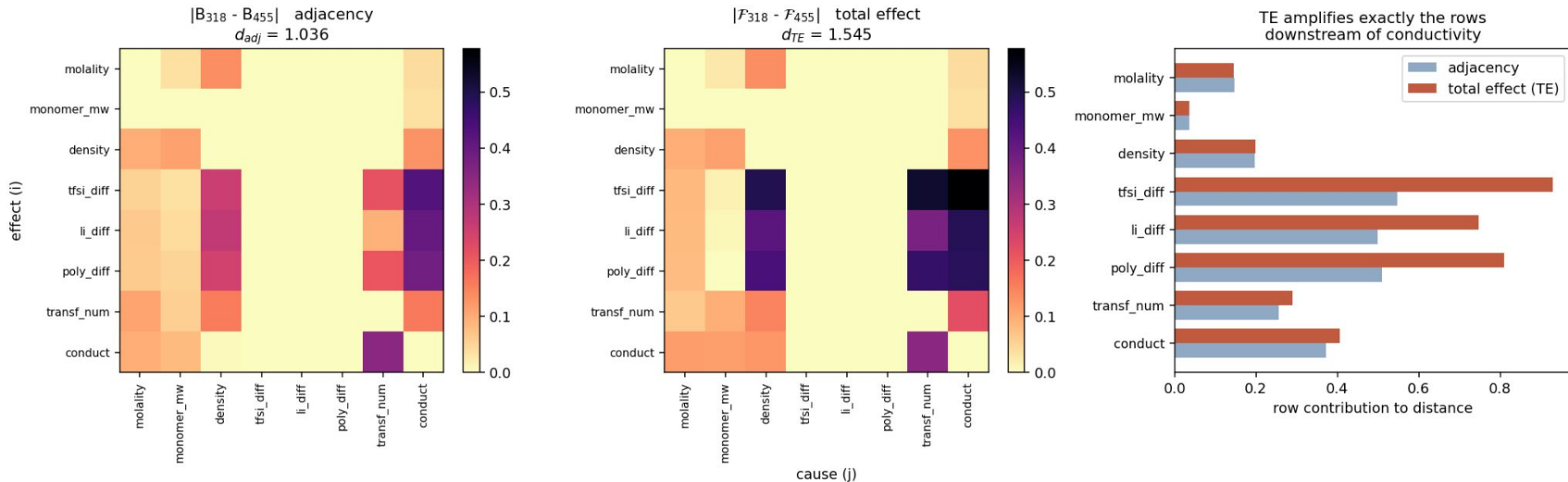


Frobenius : only catch the edges directly connects

Total Effects : collects all the paths

TE and Frobenius metrics

Same graph pair, two metrics: TE = 1.545 vs adjacency = 1.036 (x1.49)



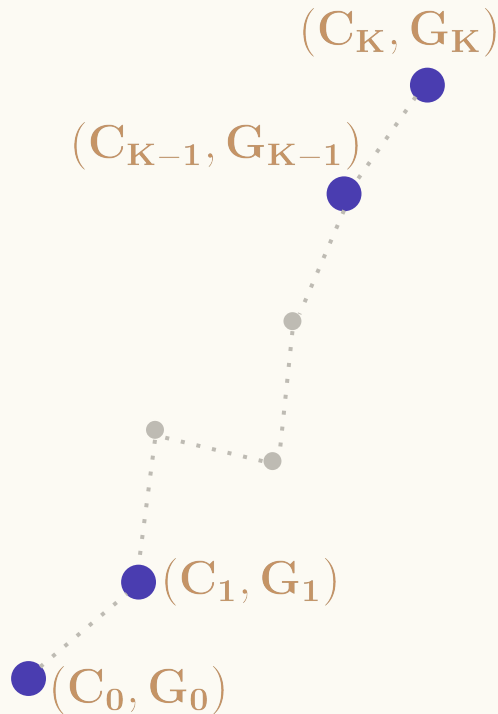
The total effect metric shows the accumulation of all paths

Path-connecting causal graphs

Path exists:
Type of Conditional homogeneity

Can we find a **path**?

No path exists:
Distinct causal structures



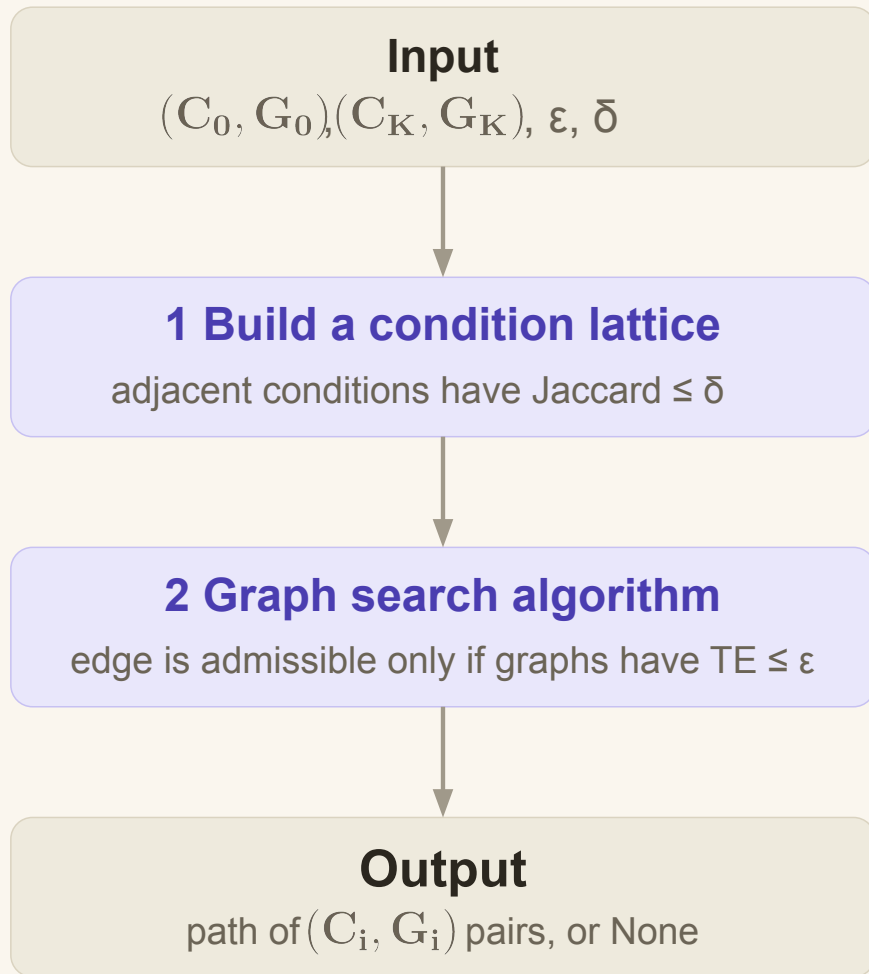
DEFINITION

(C_0, G_0) and (C_K, G_K) are **(ϵ, δ) -path connected** if there exists some $\{C_i\}$ so that at every step,

$$d_{TE}(G_i, G_{i+1}) \leq \epsilon$$

$$d_J(C_i, C_{i+1}) \leq \delta$$

The pathing algorithm



The pathing algorithm

Tradeoff!

Comprehensiveness



Computational Cost



Input

$(C_0, G_0), (C_K, G_K), \epsilon, \delta$

1 Build a condition lattice

adjacent conditions have Jaccard $\leq \delta$

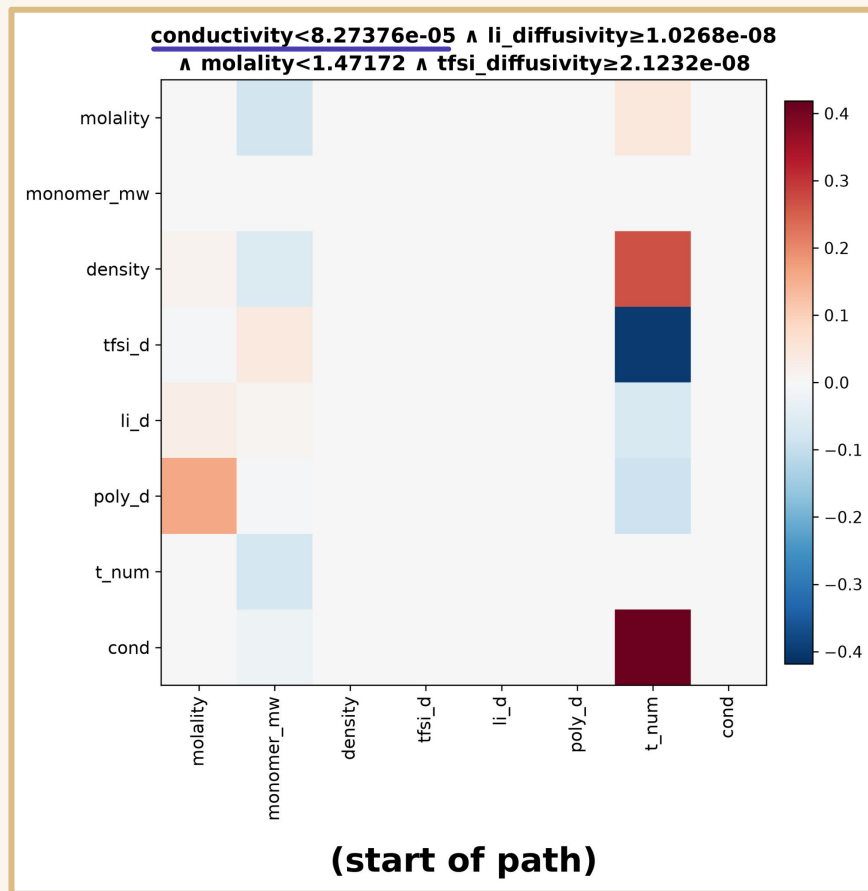
2 Graph search algorithm

edge is admissible only if graphs have TE $\leq \epsilon$

Output

path of (C_i, G_i) pairs, or None

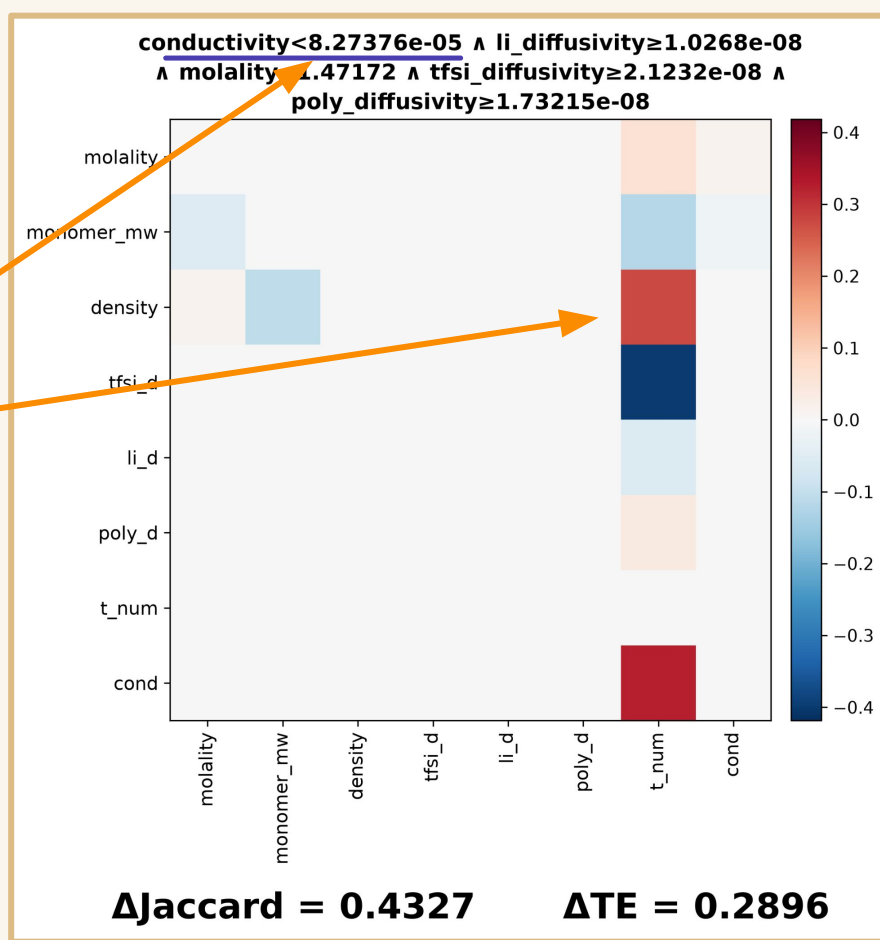
The pathing algorithm



0->1->2->3->4->5->6

The pathing algorithm

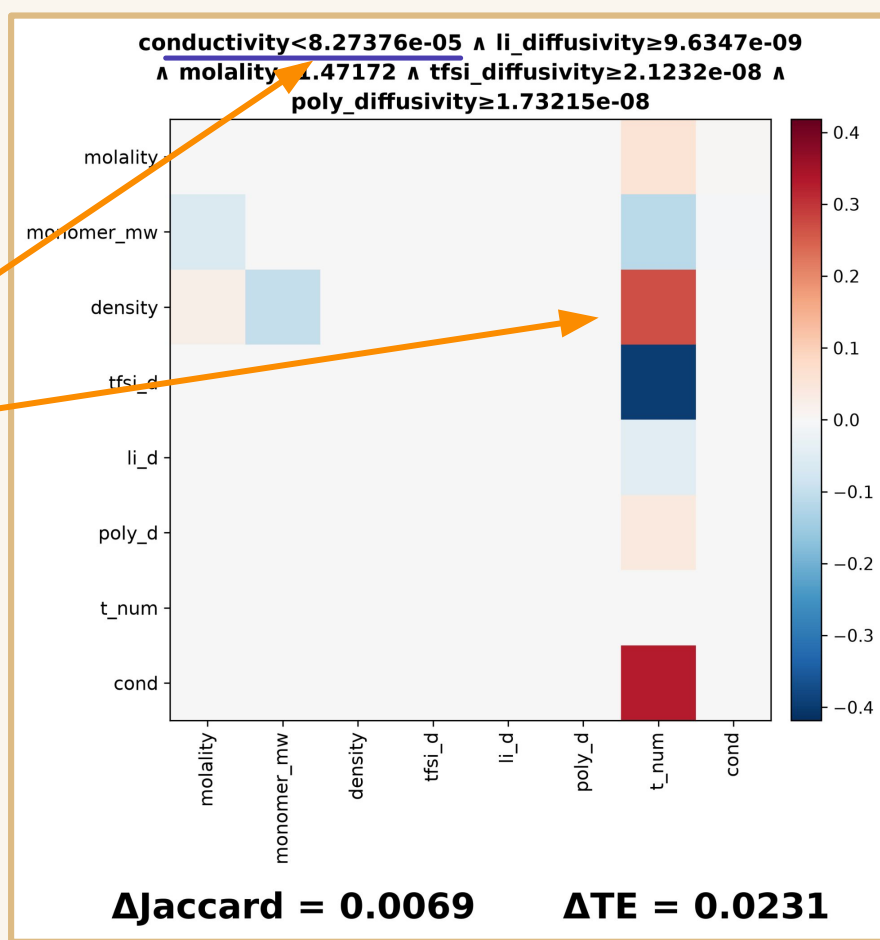
Doesn't change!



0->1->2->3->4->5->6

The pathing algorithm

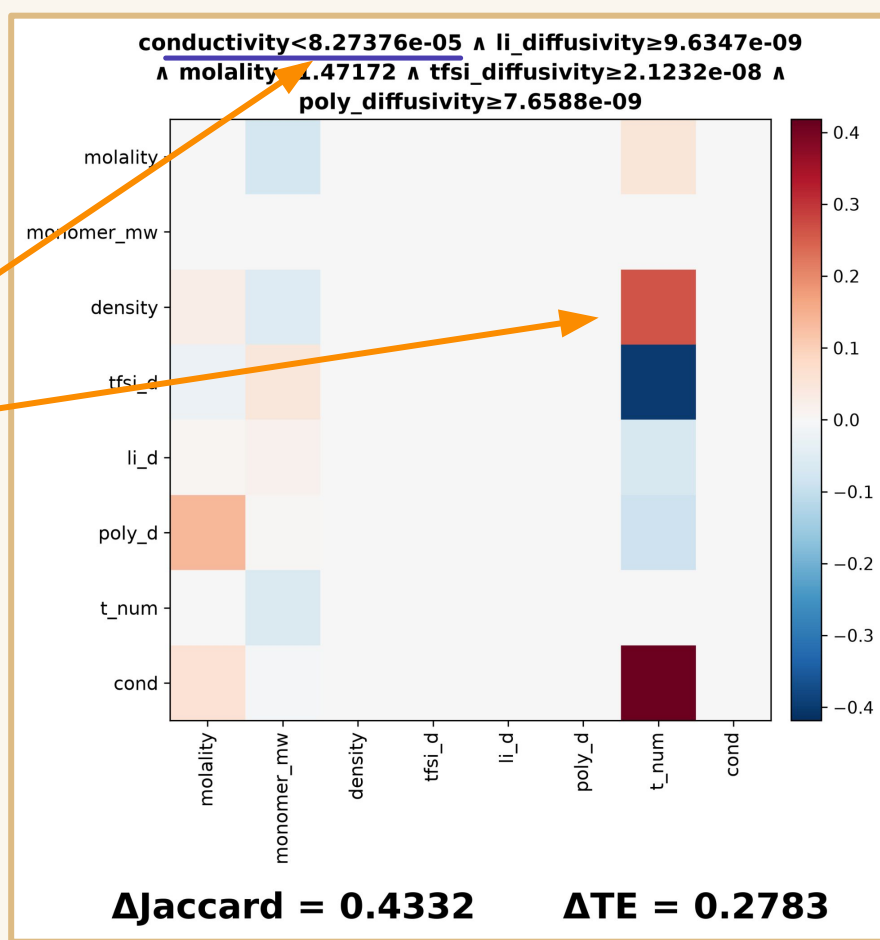
Doesn't change!



0->1->2->3->4->5->6

The pathing algorithm

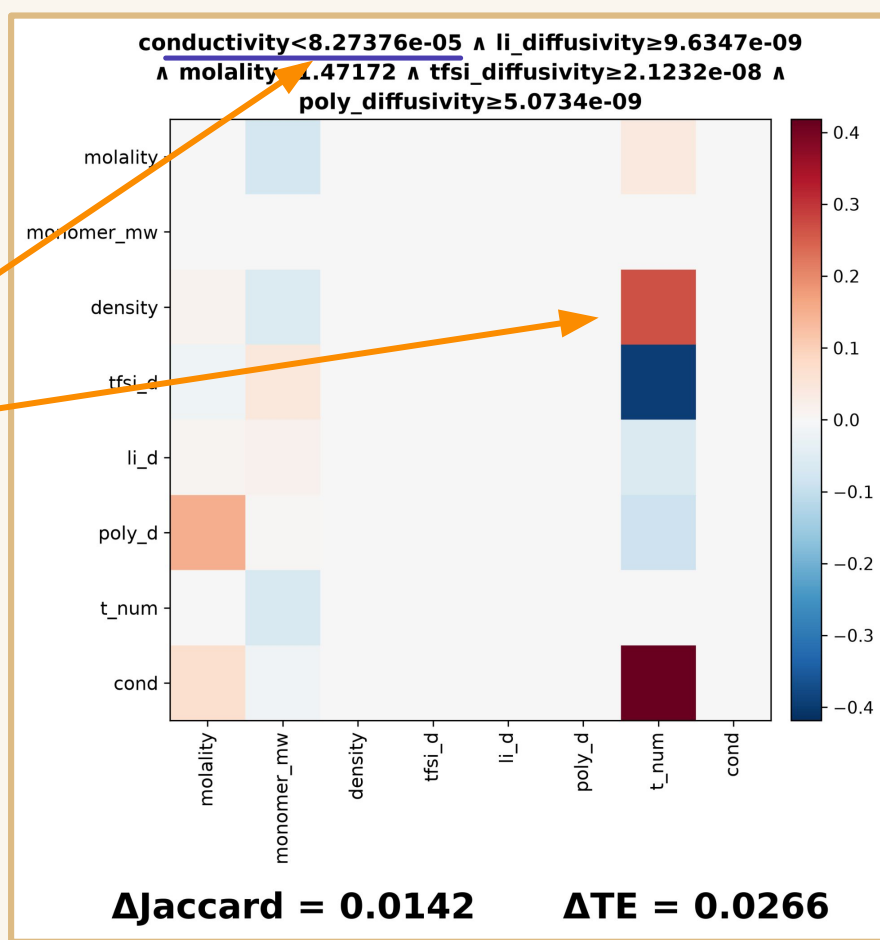
Doesn't change!



0->1->2->3->4->5->6

The pathing algorithm

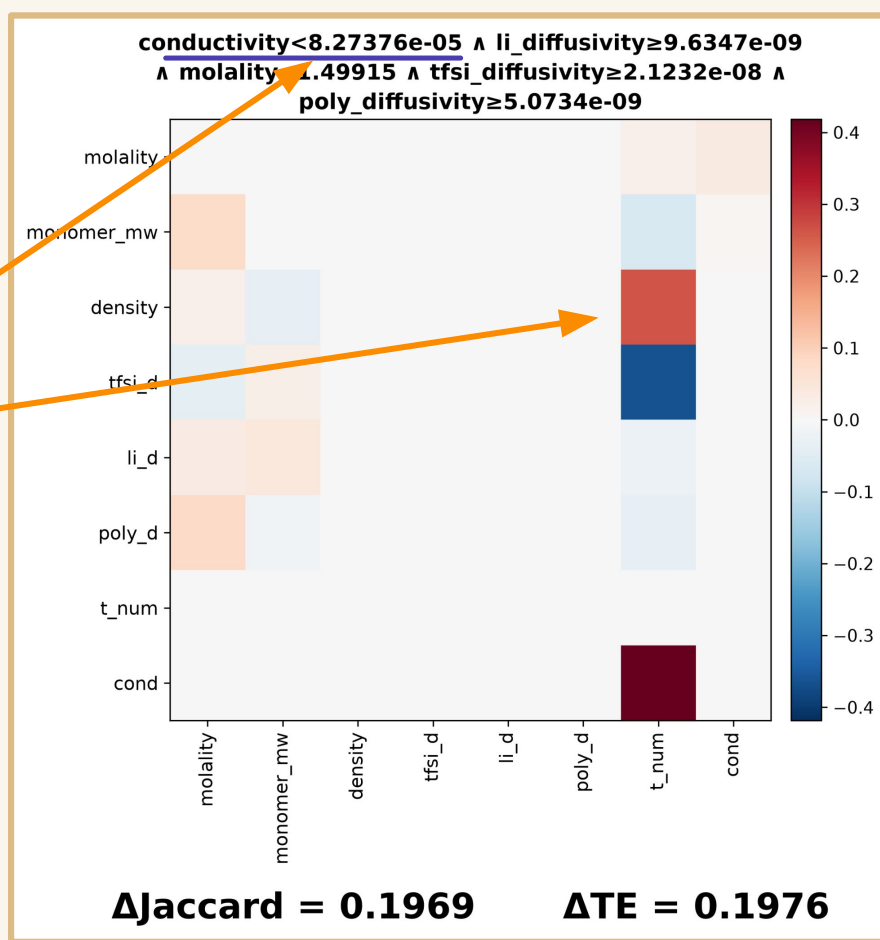
Doesn't change!



0->1->2->3->4->5->6

The pathing algorithm

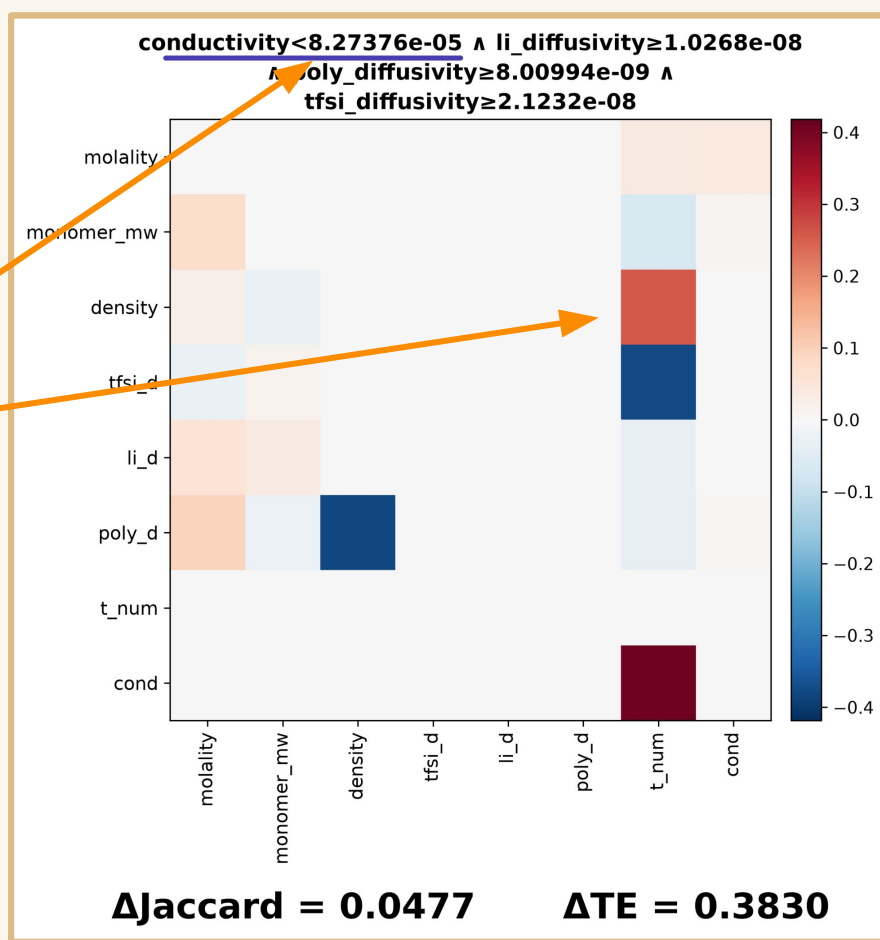
Doesn't change!



0->1->2->3->4->5->6

The pathing algorithm

Doesn't change!

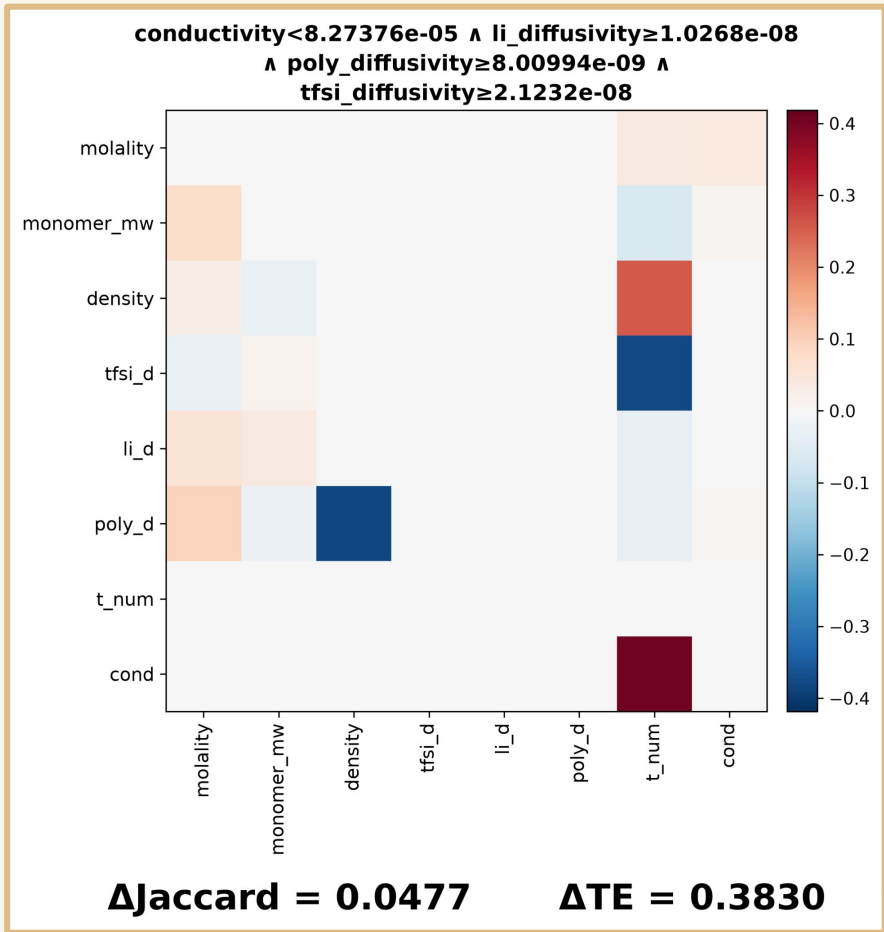


0->1->2->3->4->5->6

The pathing algorithm

Some conditions are *load bearing candidates* (**conductivity**, **tfsi_diffusivity**)

Potential **drivers** of a fixed mechanism



0->1->2->3->4->5->6

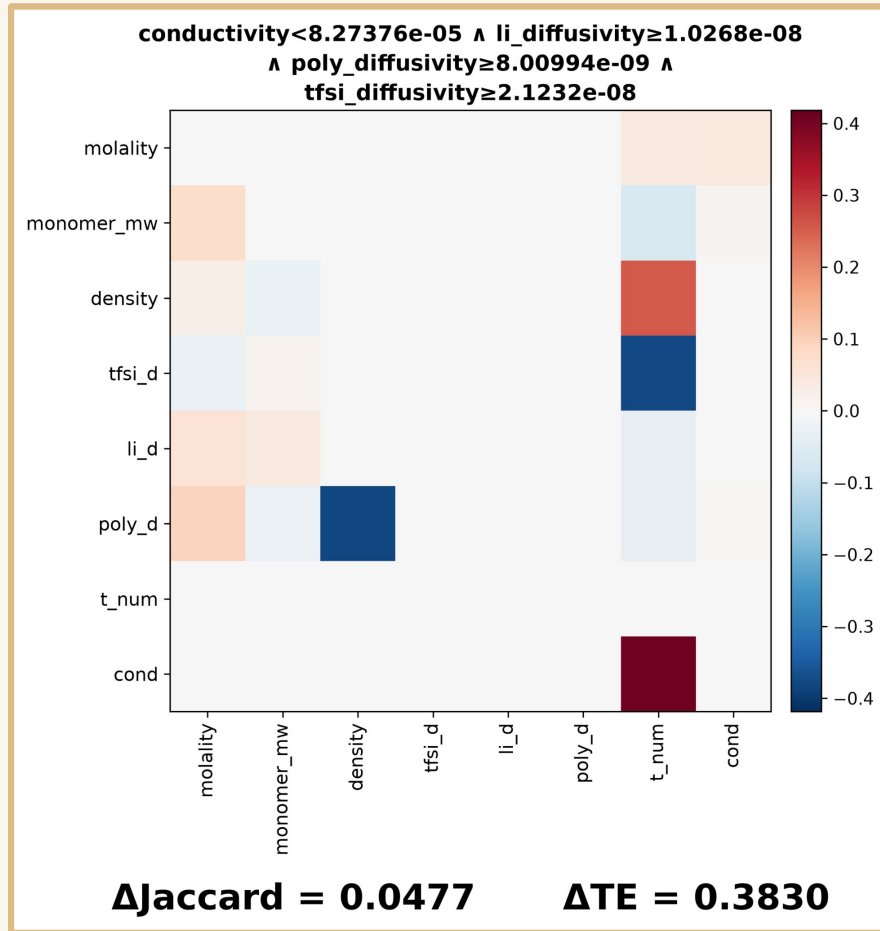
The pathing algorithm

Some conditions are *load bearing candidates* (**conductivity**, **tfsi_diffusivity**)

Potential **drivers** of a fixed mechanism

Structural finding

Not physics

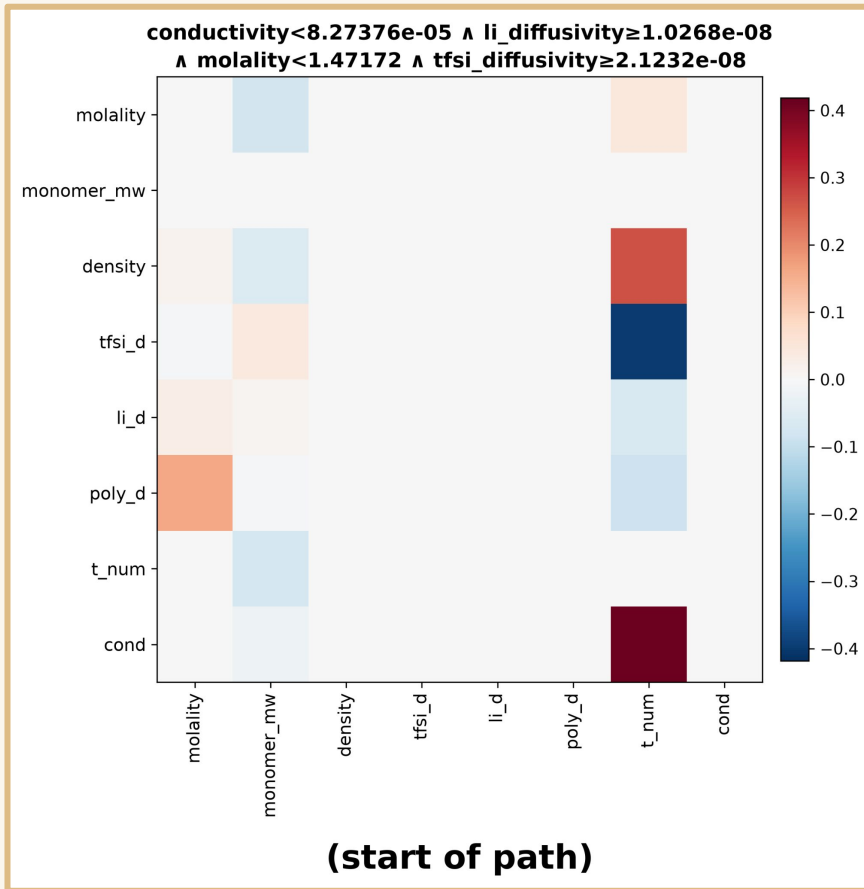


0->1->2->3->4->5->6

The pathing algorithm

Some conditions are *load bearing candidates* (**conductivity**, **tfsi_diffusivity**)

Potential **drivers** of a fixed mechanism



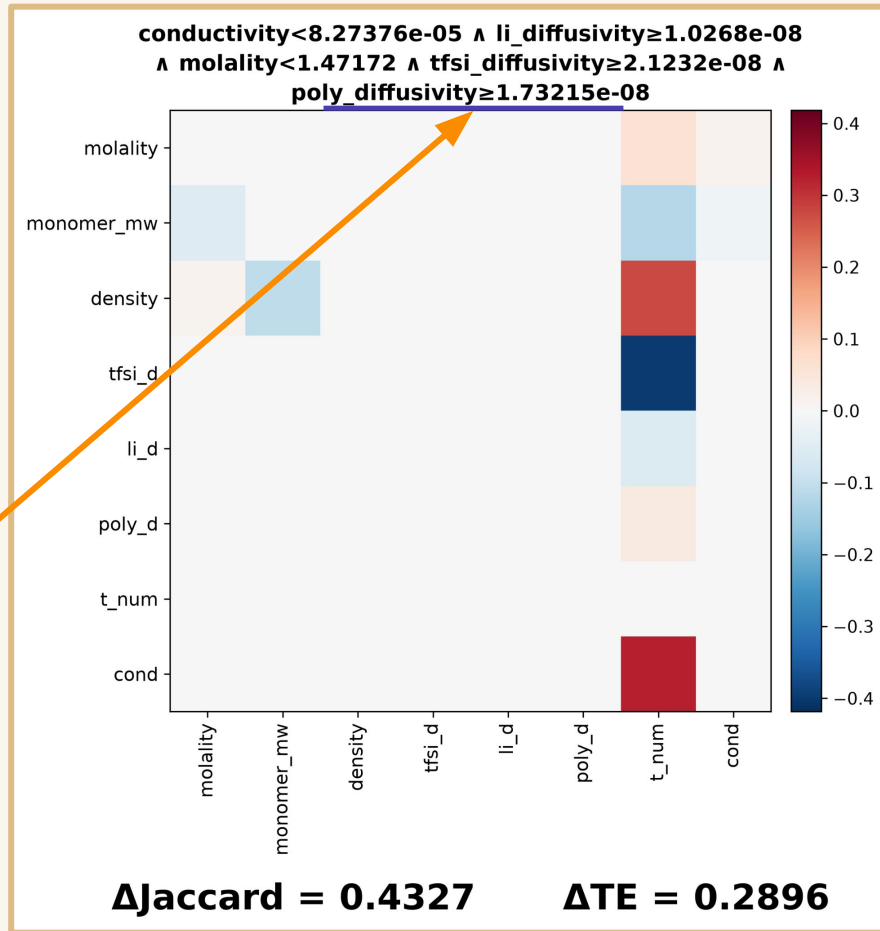
0->1->2->3->4->5->6

The pathing algorithm

Some conditions are *load bearing candidates* (**conductivity**, **tfsi_diffusivity**)

Potential **drivers** of a fixed mechanism

Changes “freely”!



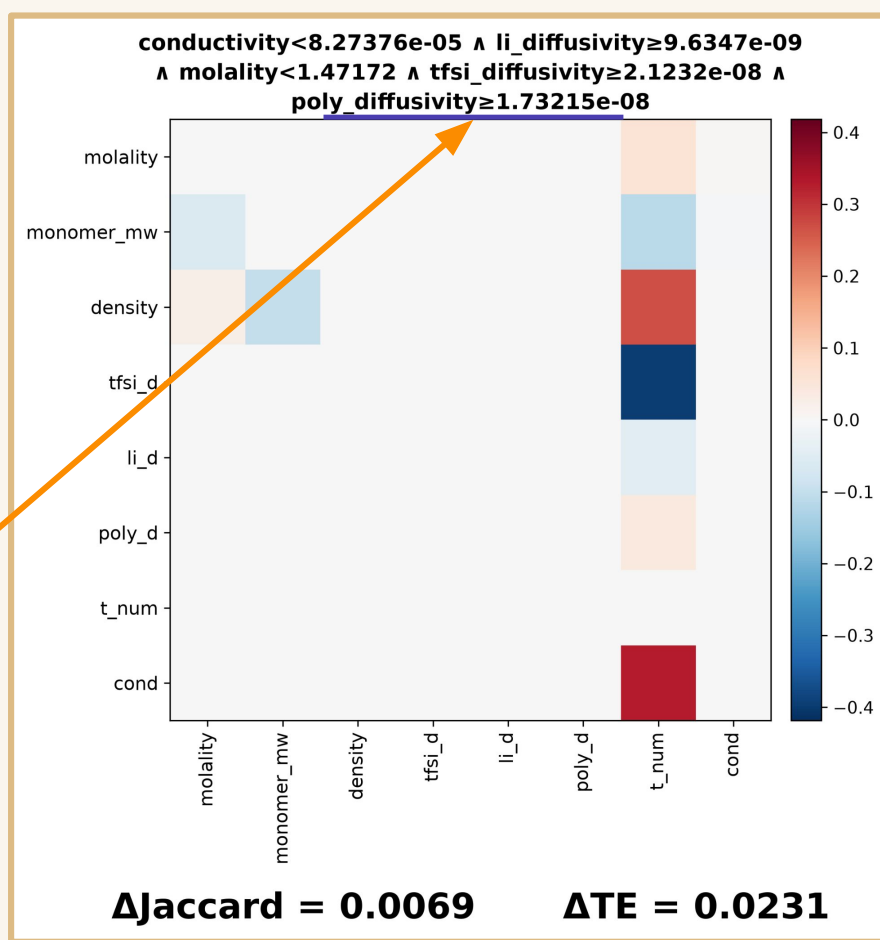
0->1->2->3->4->5->6

The pathing algorithm

Some conditions are *load bearing candidates* (**conductivity**, **tfsi_diffusivity**)

Potential **drivers** of a fixed mechanism

Changes “freely”!



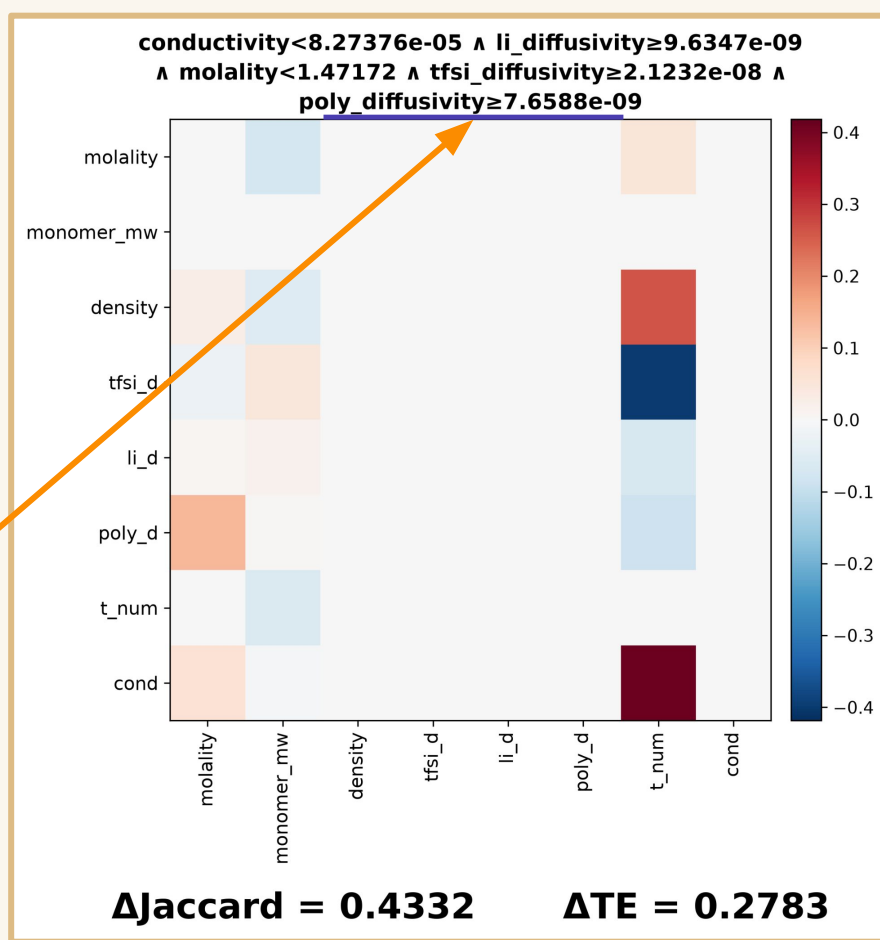
0->1->2->3->4->5->6

The pathing algorithm

Some conditions are *load bearing candidates* (**conductivity**, **tfsi_diffusivity**)

Potential **drivers** of a fixed mechanism

Changes “freely”!



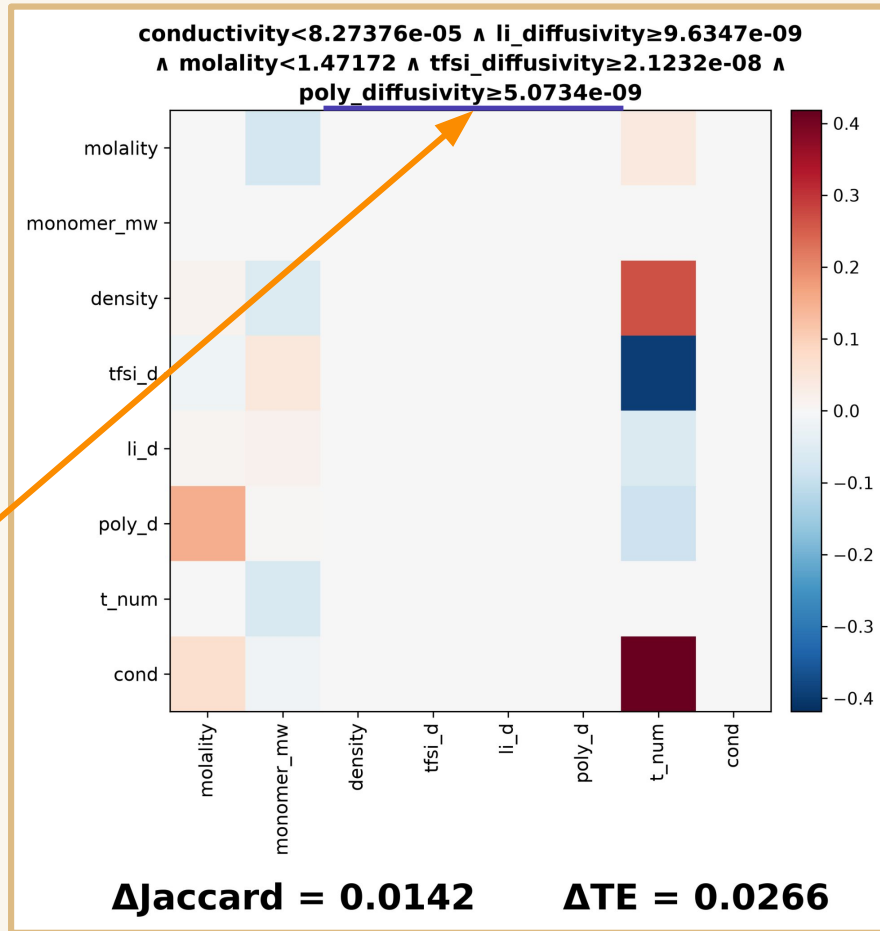
0->1->2->3->4->5->6

The pathing algorithm

Some conditions are *load bearing candidates* (**conductivity**, **tfsi_diffusivity**)

Potential **drivers** of a fixed mechanism

Changes “freely”!



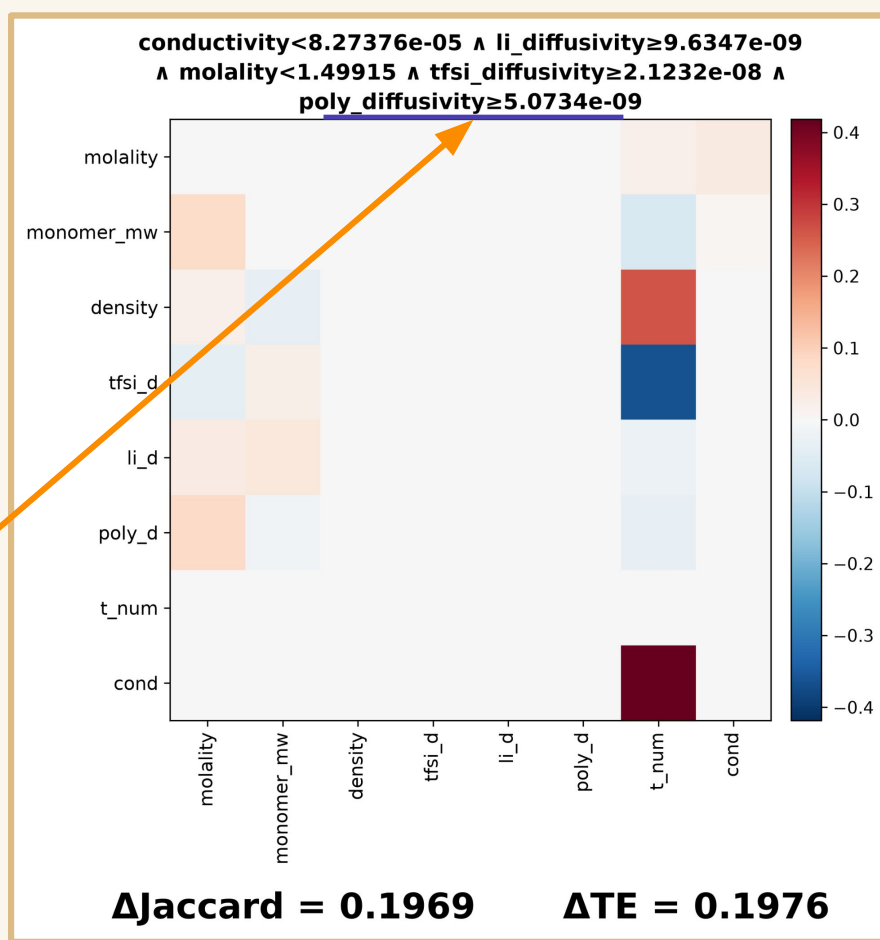
0->1->2->3->4->5->6

The pathing algorithm

Some conditions are *load bearing candidates* (**conductivity**, **tfsi_diffusivity**)

Potential **drivers** of a fixed mechanism

Changes “freely”!



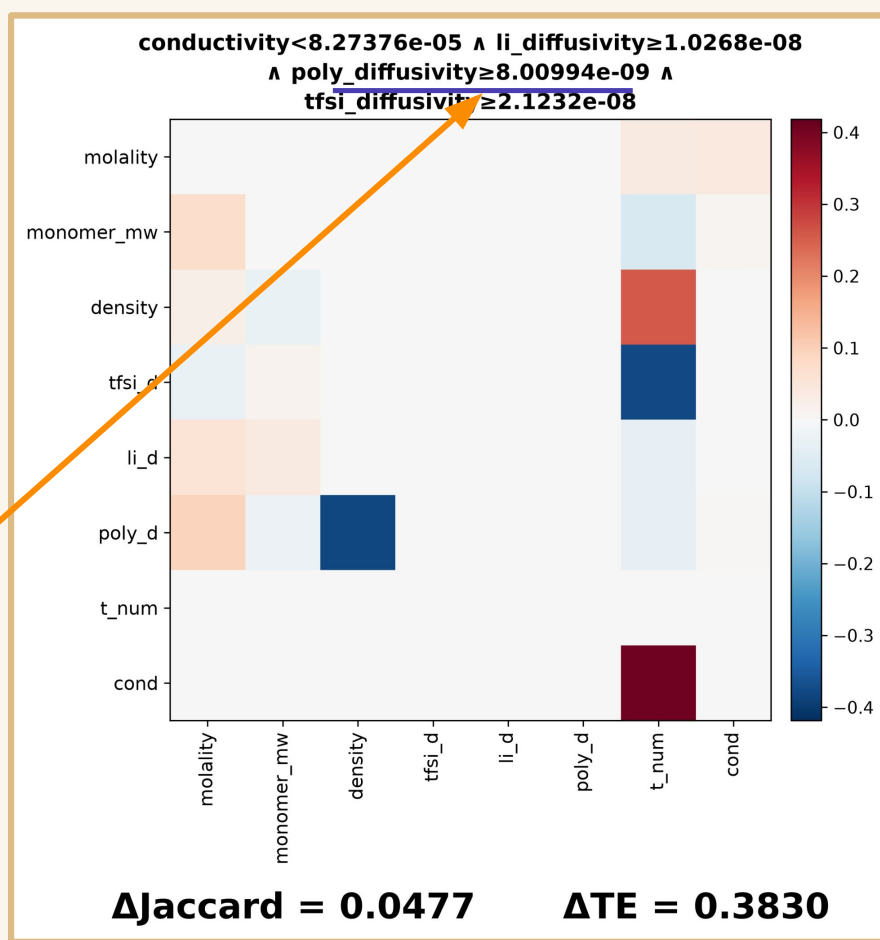
0->1->2->3->4->5->6

The pathing algorithm

Some conditions are *load bearing candidates* (**conductivity**, **tfsi_diffusivity**)

Potential **drivers** of a fixed mechanism

Changes “freely”!



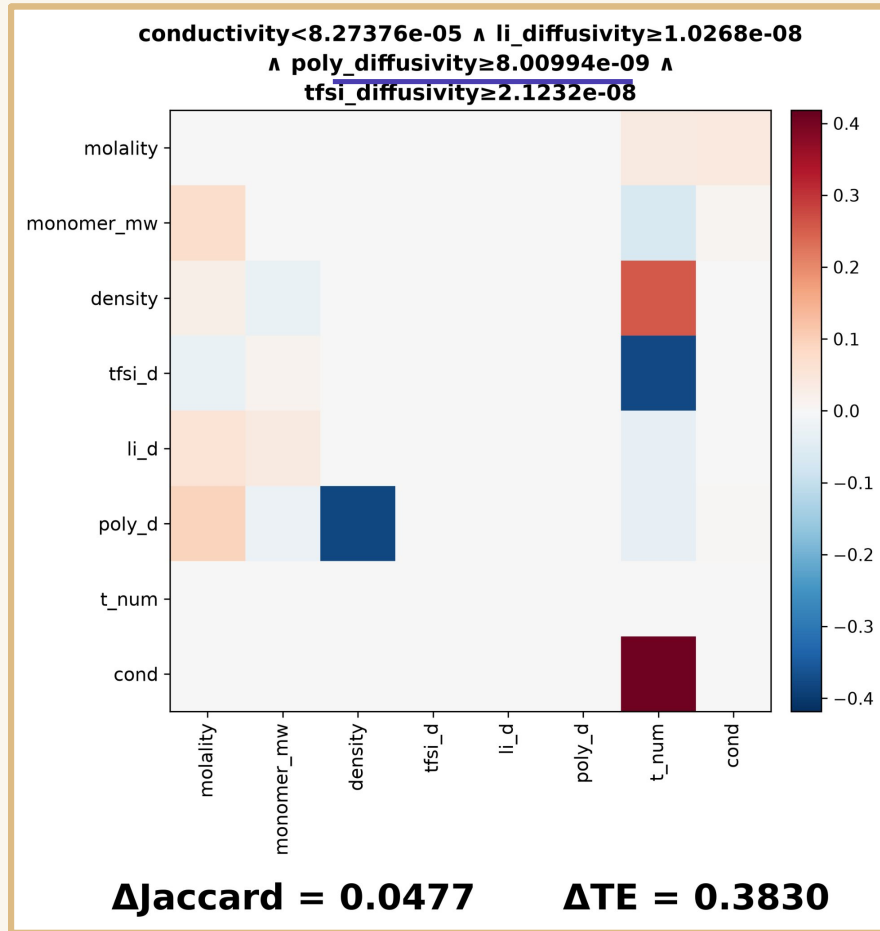
0->1->2->3->4->5->6

The path algorithm

Some conditions are *load bearing candidates* (**conductivity**, **tfsi_diffusivity**)

Potential **drivers** of a fixed mechanism

Other conditions are *mechanism irrelevant candidates* (**poly_diffusivity**)



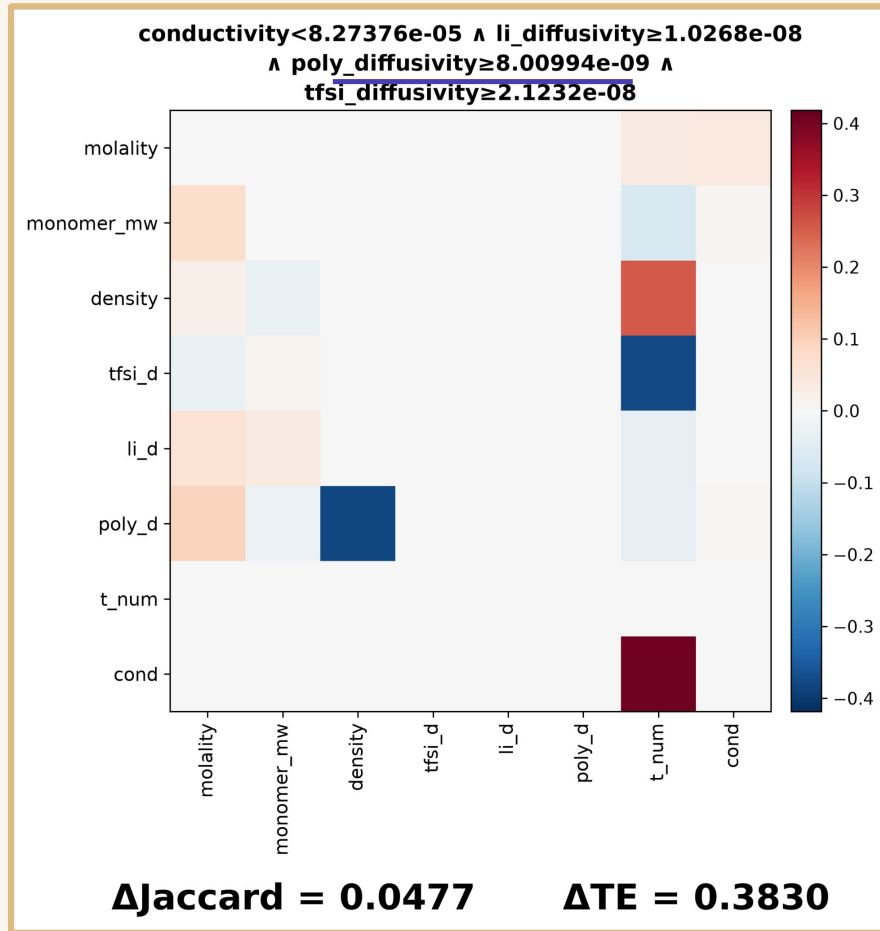
0->1->2->3->4->5->6

The path algorithm

Some conditions are *load bearing candidates* (**conductivity**, **tfsi_diffusivity**)

Discoverability!

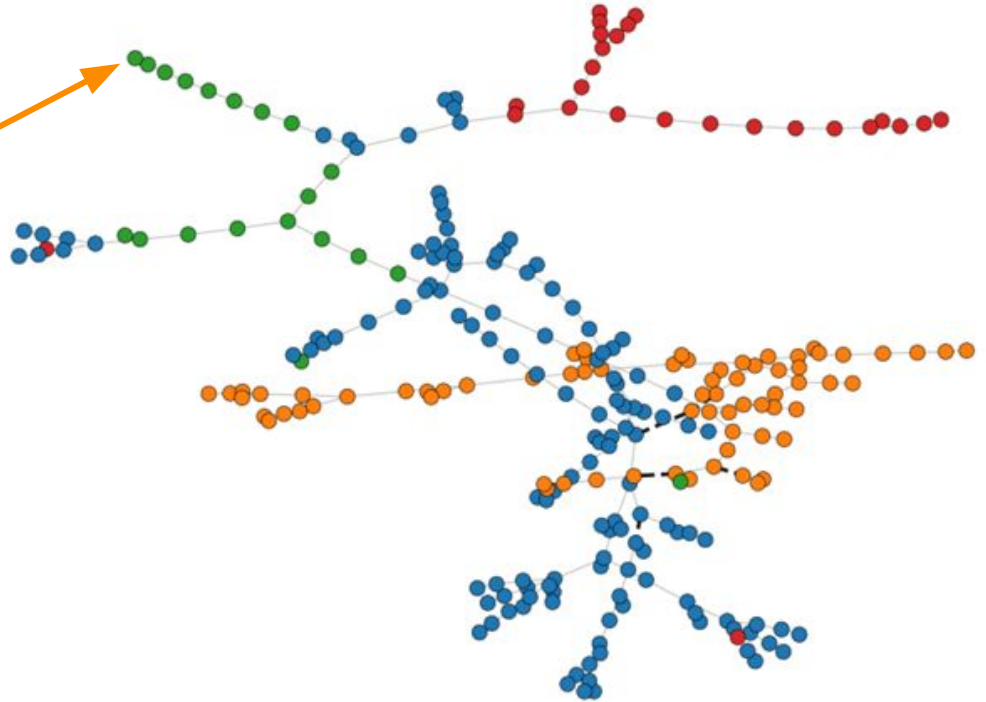
Other conditions are *mechanism irrelevant candidates* (**poly_diffusivity**)



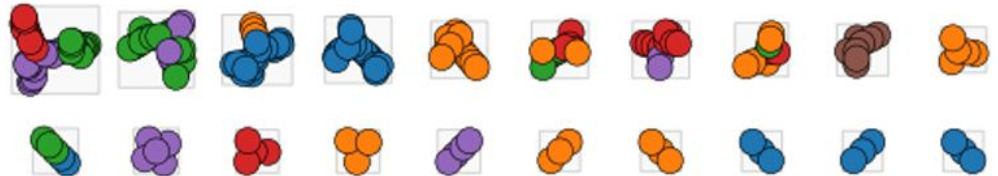
0->1->2->3->4->5->6

Path Components

Each dot is a single
condition-graph
pair

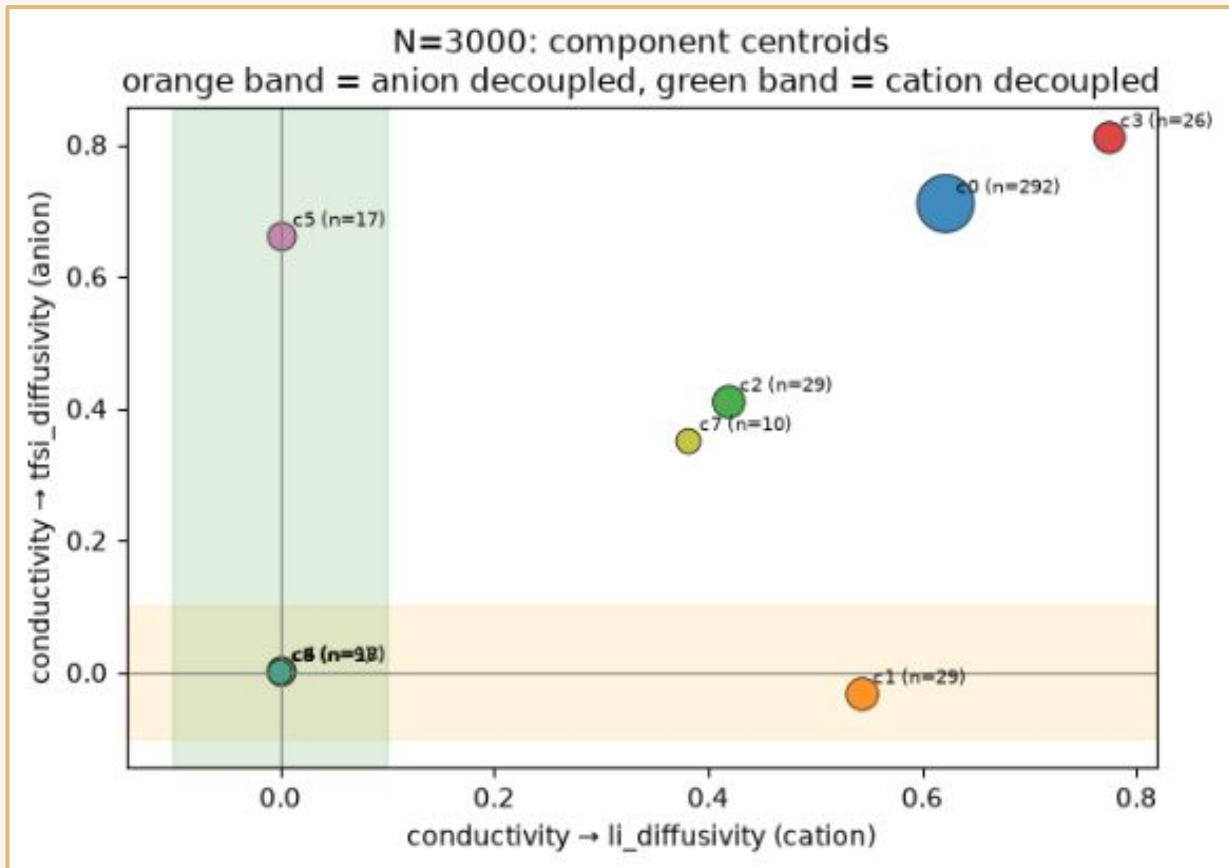


Path Components:
clustering that respect the
underlying **topologies**.



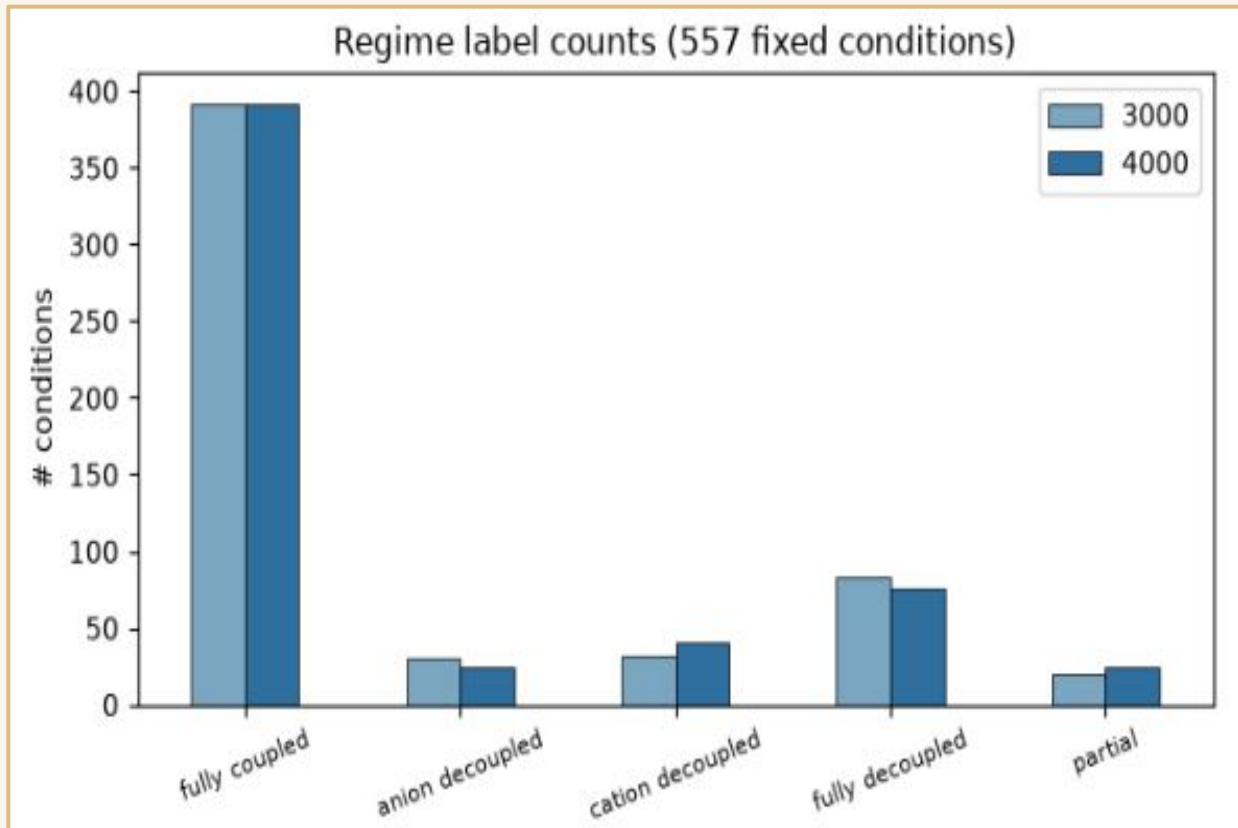
Path Components

total Effect Evaluation
across Clusters →
Clear Separation Identified



Path Components

Robust Regime Distribution
Across Sample Sizes N

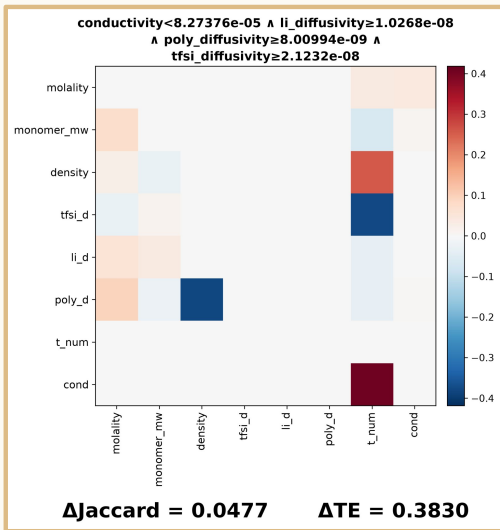


Future Work

Explore more patterns in the paths

Study more paths

apply to **more data**.



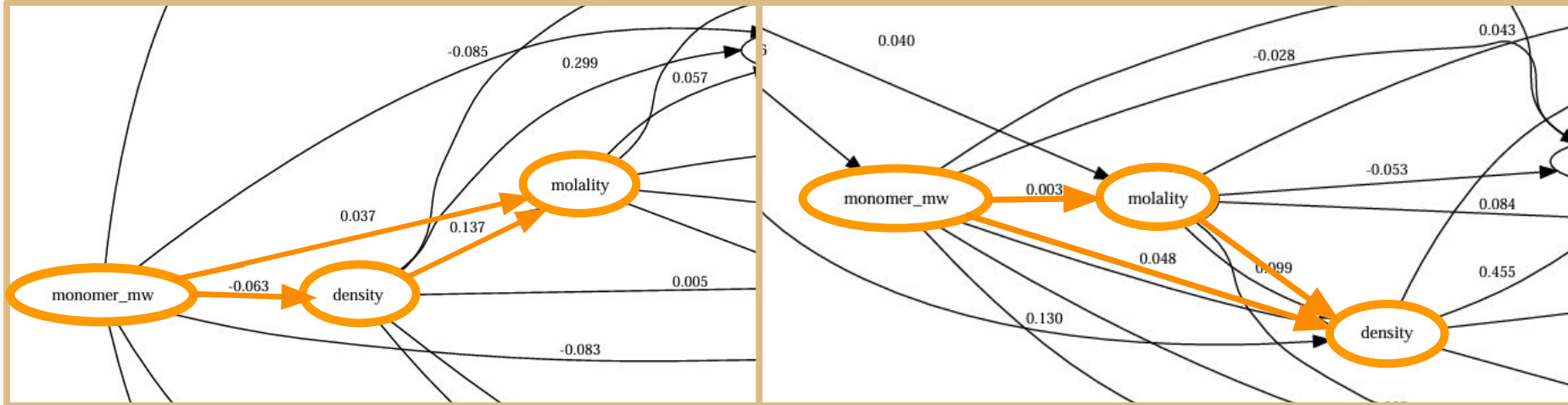
Theoretical Guarantees

Prove our conjectures regarding the robustness of pathing.

Automate Discovery

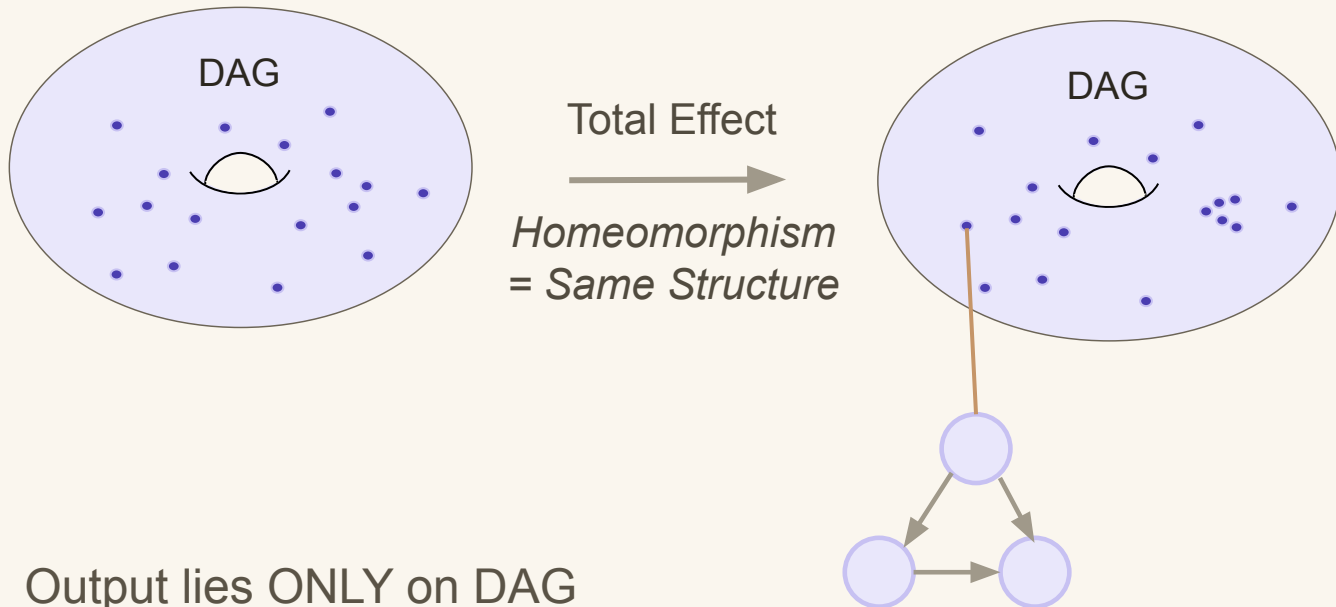
Automate detection of *mechanism irrelevancies* and *load bearing candidates* in relevant paths.

Local Pathing

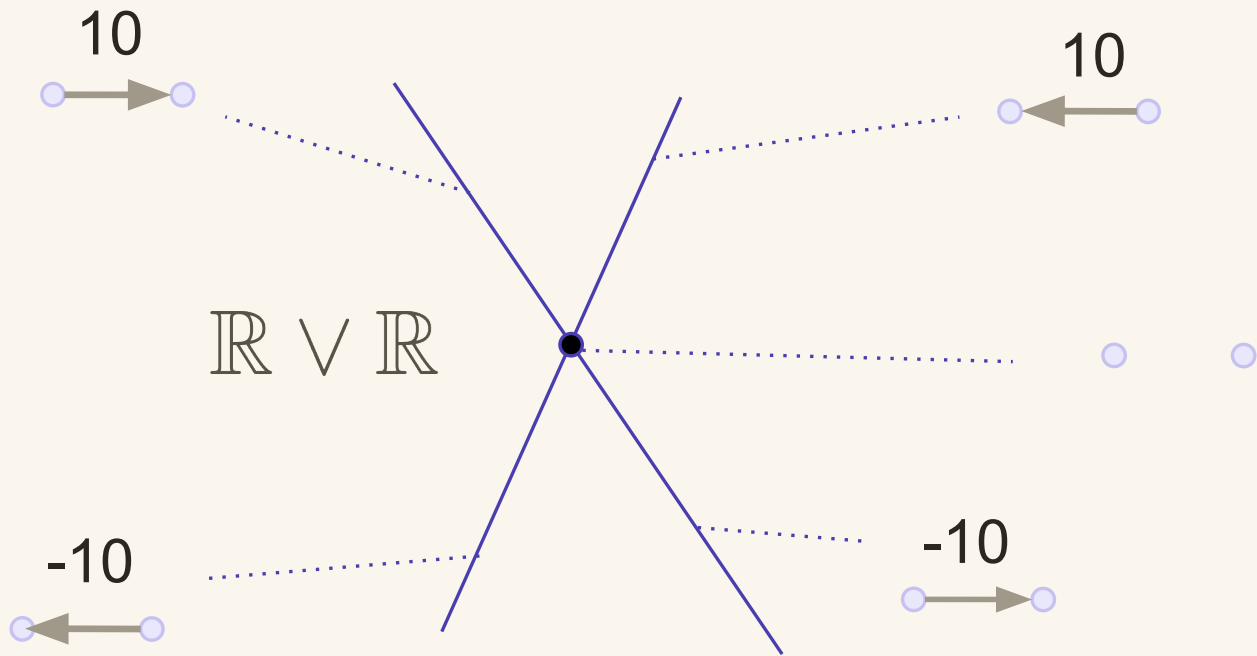


Study **substructures** as conditions change

Underlying Topology

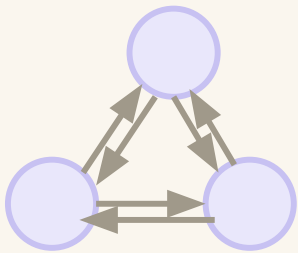


Edge Space



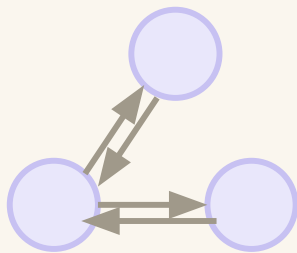
| All possibilities between fixed nodes

Nerve Structure of Topology



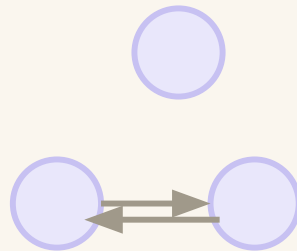
“point”

$$(\mathbb{R} \vee \mathbb{R})^3$$



“edge”

$$(\mathbb{R} \vee \mathbb{R})^2$$

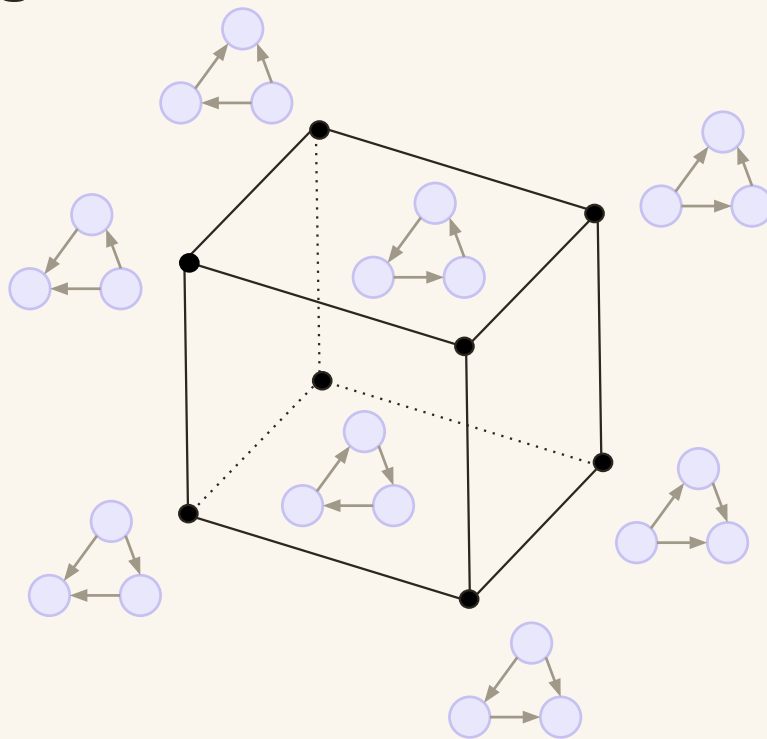


“face”

$$\mathbb{R} \vee \mathbb{R}$$

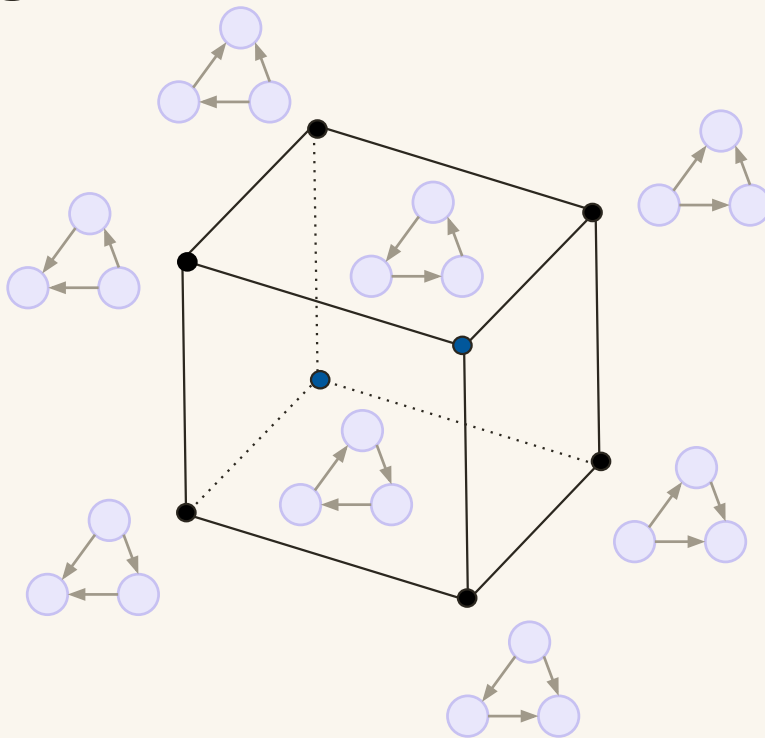
| All edge + weight possibilities

3-node Pathing



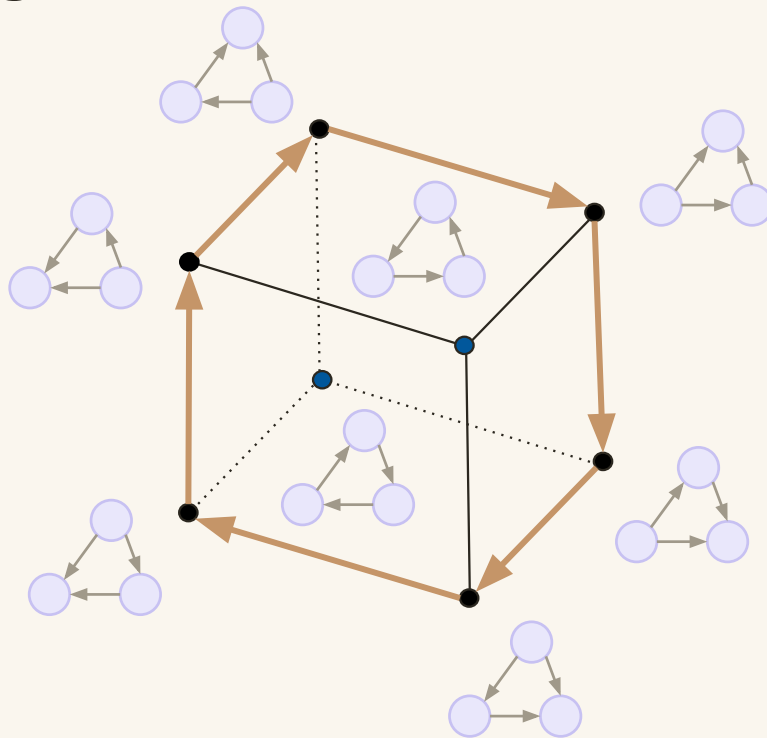
| Small example

3-node Pathing



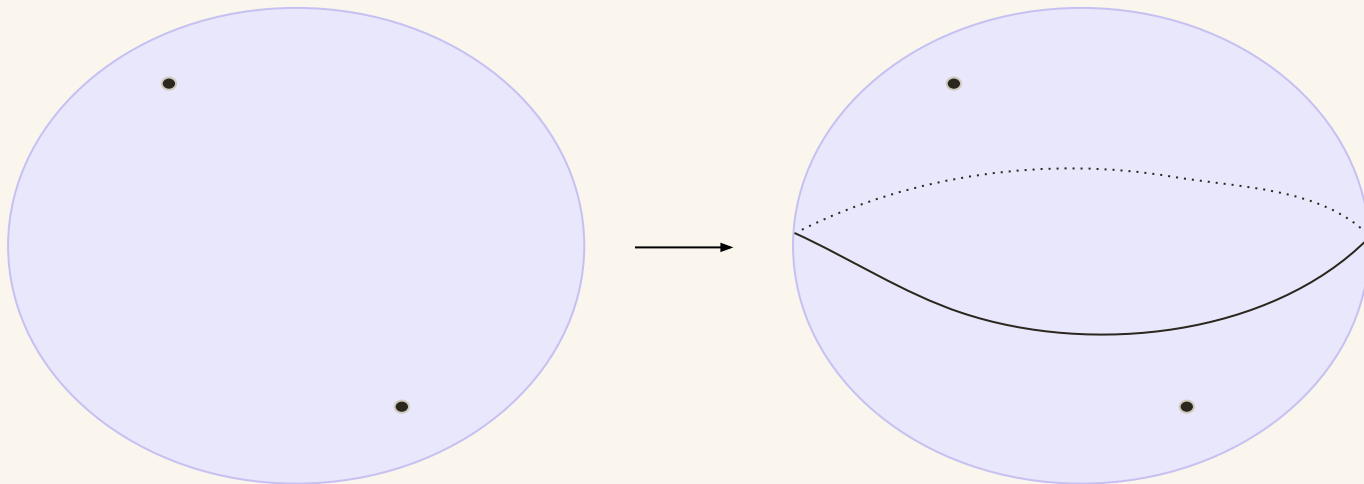
| Remove cycles

3-node Pathing



Consider loop in DAG space

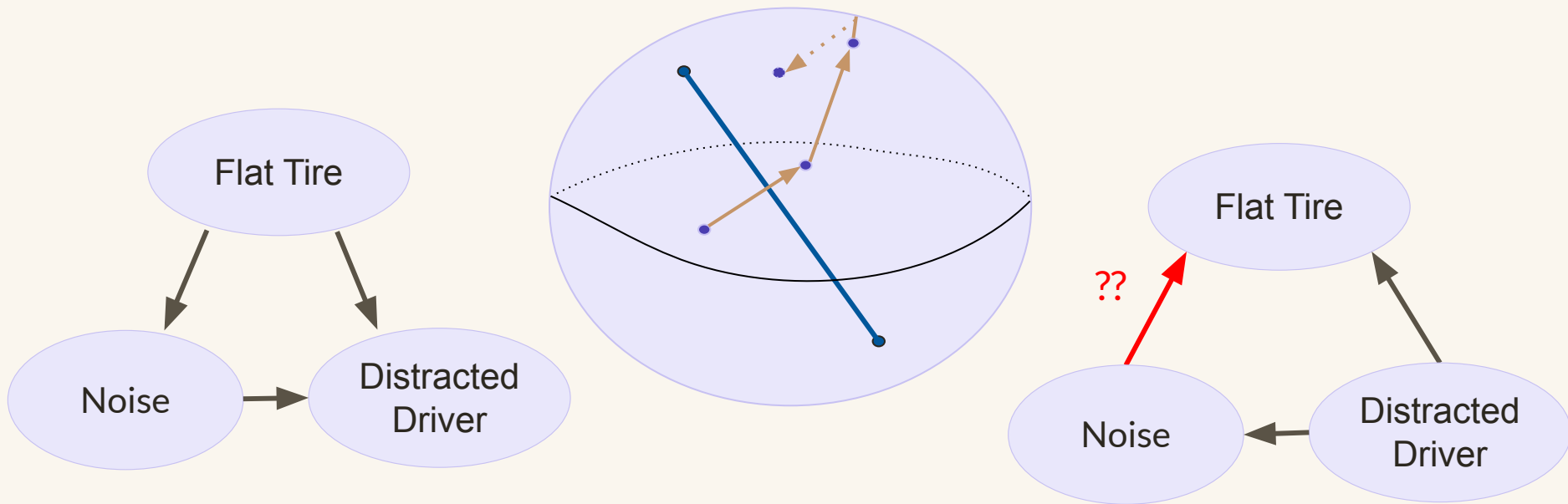
Choice of Object to Study



$$L_n = S^{n(n-1)-1} \cap \text{DAG}_n$$

Remove 0!

3-node Toy Example



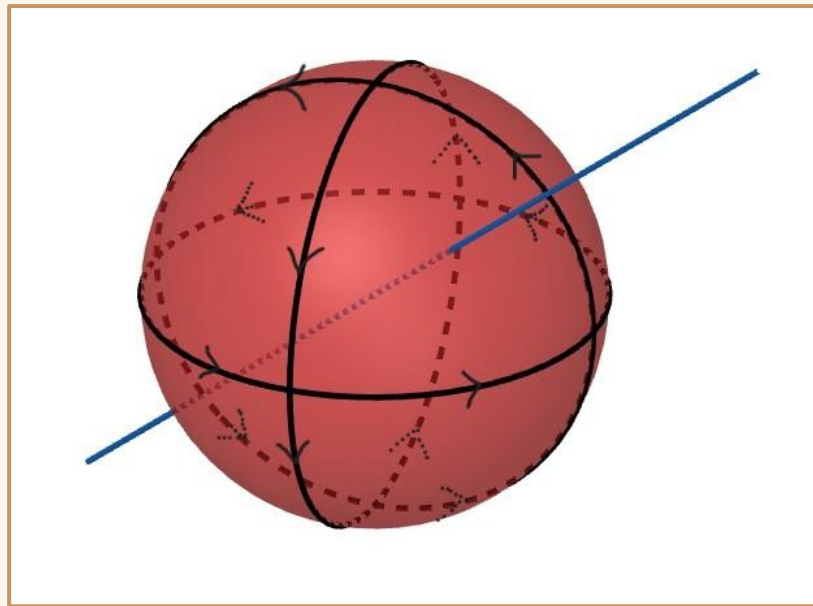
Causal graphs “wind” around this line
Half loops may represent incongruence

Projecting Along Cycles

$$M \mapsto M|_{P_{CW}}, M|_{P_{CCW}}$$

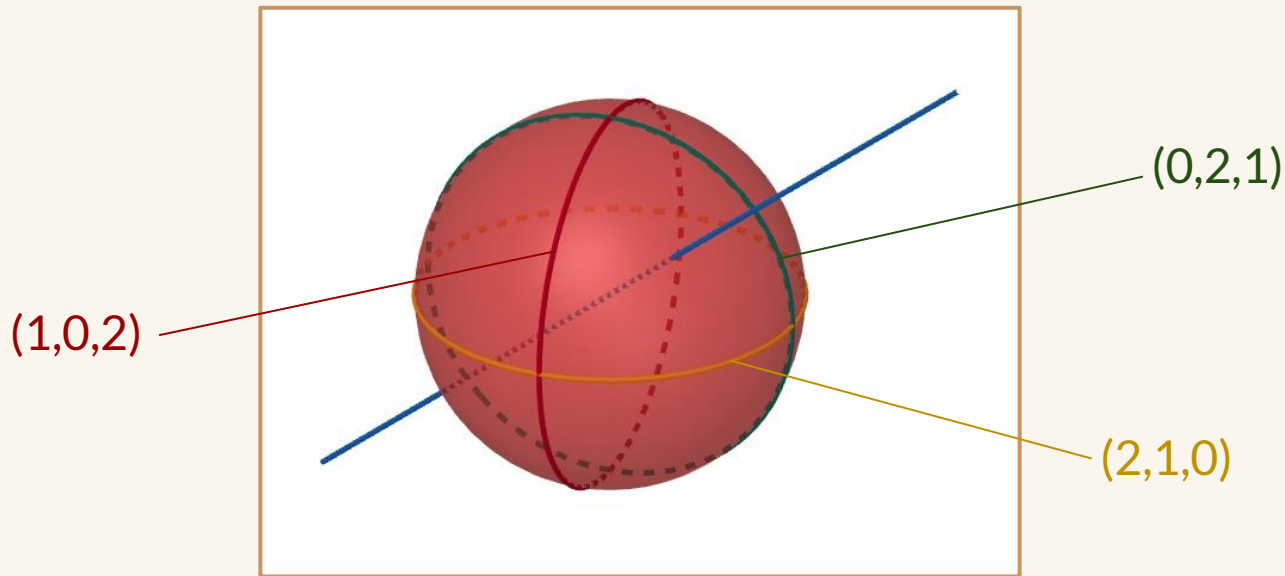
$$P_{CW} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, P_{CCW} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

$$M|_{P_{CW}}, M|_{P_{CCW}} \mapsto \text{Norm}(M|_{P_{CW}}), \text{Norm}(M|_{P_{CCW}})$$



Consider several projections on choice of 3 columns

Projecting Along Cycles (Causal Order)



Graphs with cyclic sub-causal order lie on
3 great circles

3-Node Pathing Algorithm

$$\{p_0, p_1, \dots\} \quad \max_{p_{i+1} \in C_{p_i}} (\theta(p_i, p_{i+1})) - d_J(p_i, p_{i+1}) : d_J < \epsilon = c \cdot \text{mean}(d_J)$$

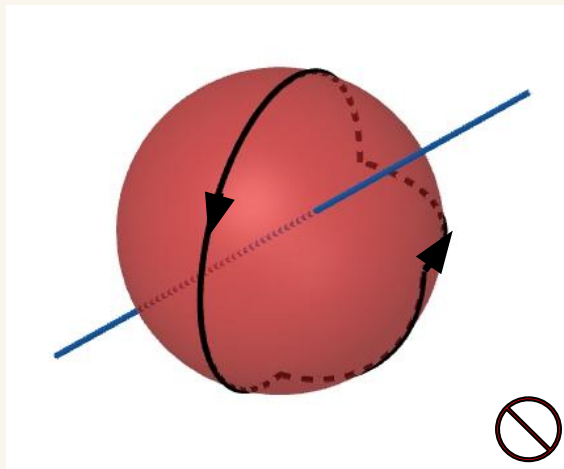
Incremental
Winding

Incremental
Jaccard

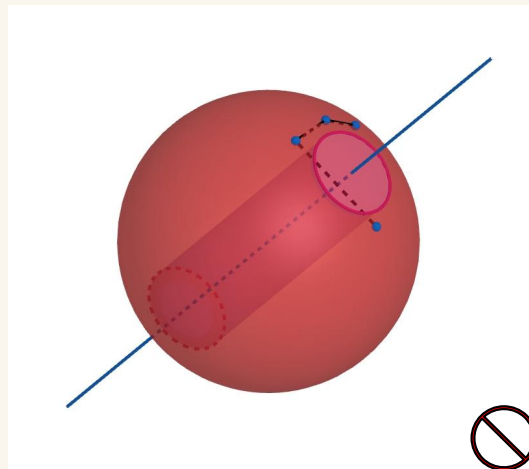
Winding algorithm conditions

3-Node Pathing Algorithm

1) No returning to great circles



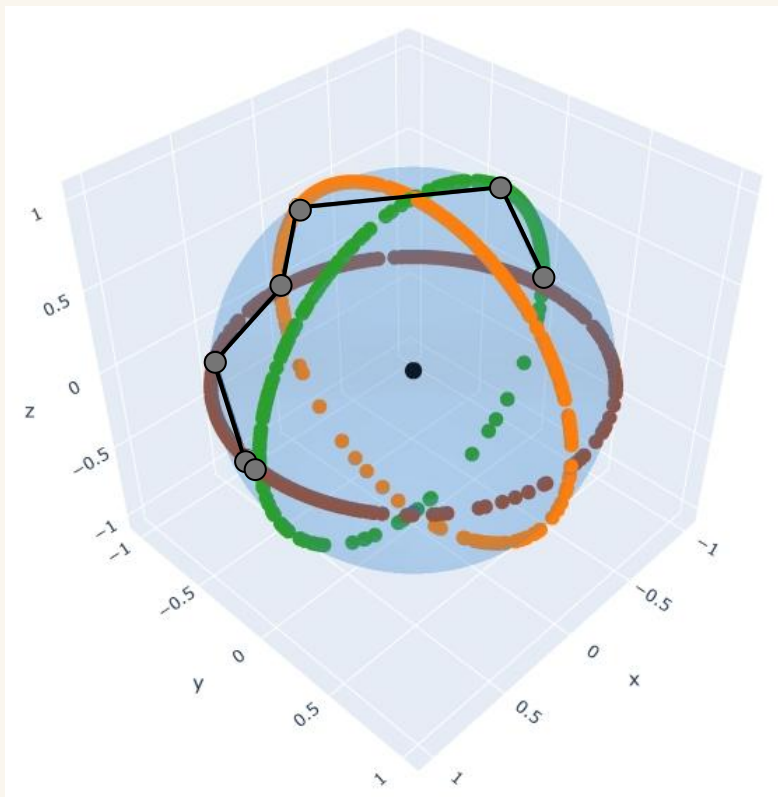
2) Avoid cylinder about line



Winding algorithm conditions

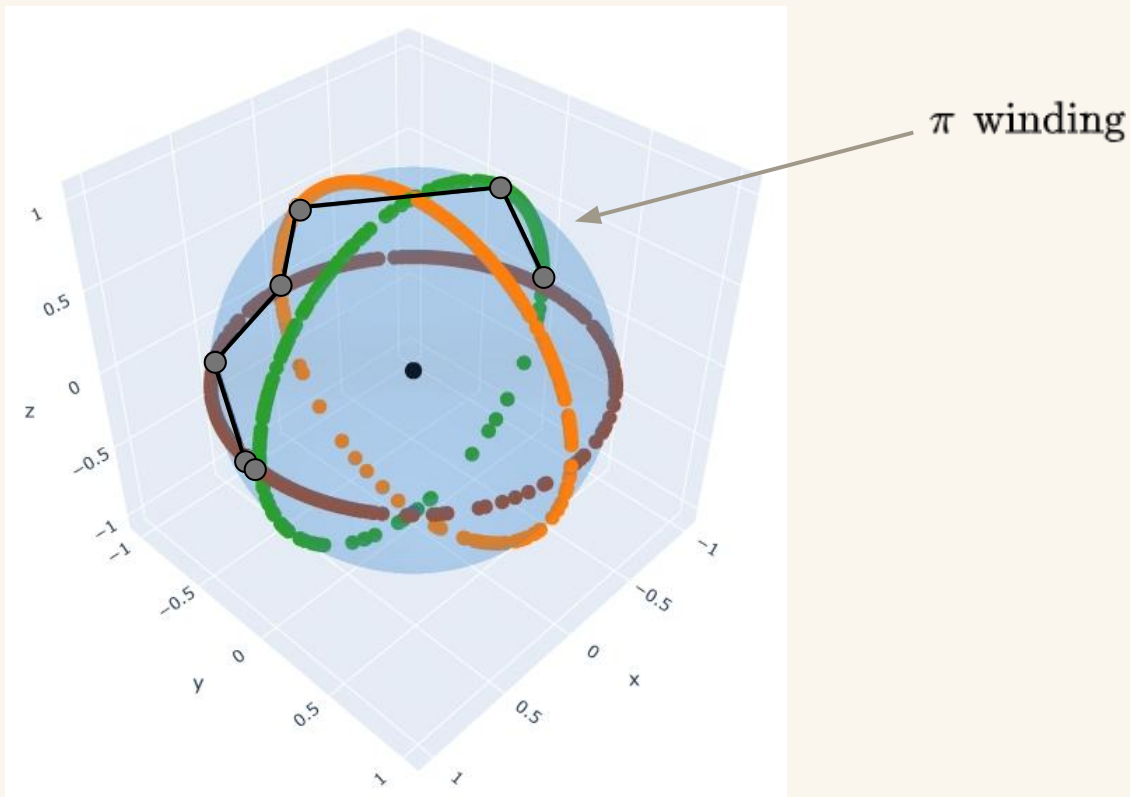
3-Node Pathing on Real Data

Winding on
real data;
Jaccard
threshold
0.3 of mean

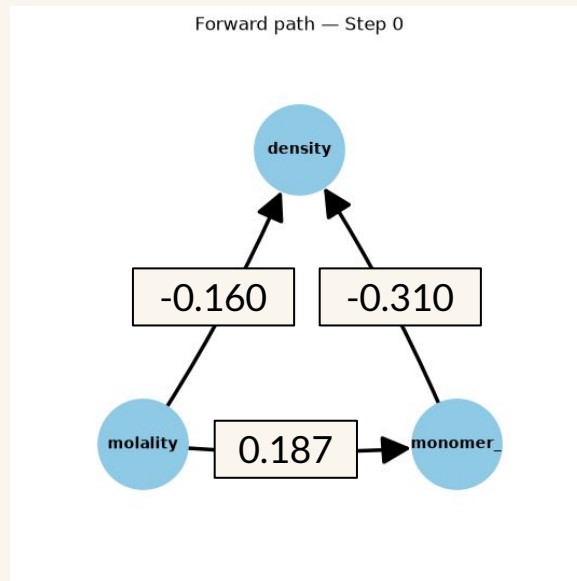
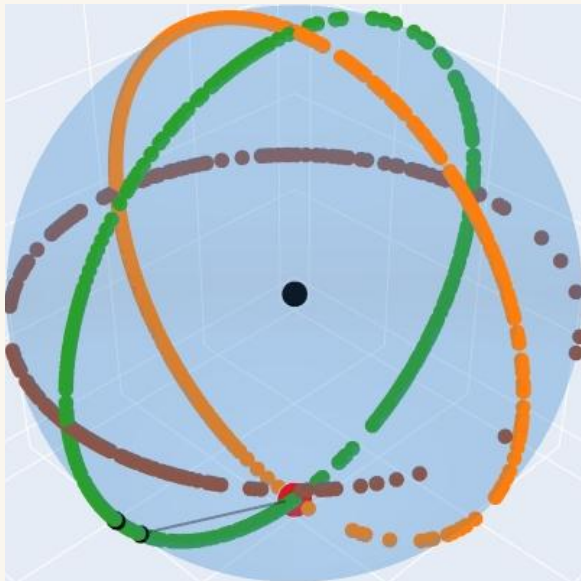


3-Node Pathing on Real Data

Winding on
real data;
Jaccard
threshold
0.3 of mean

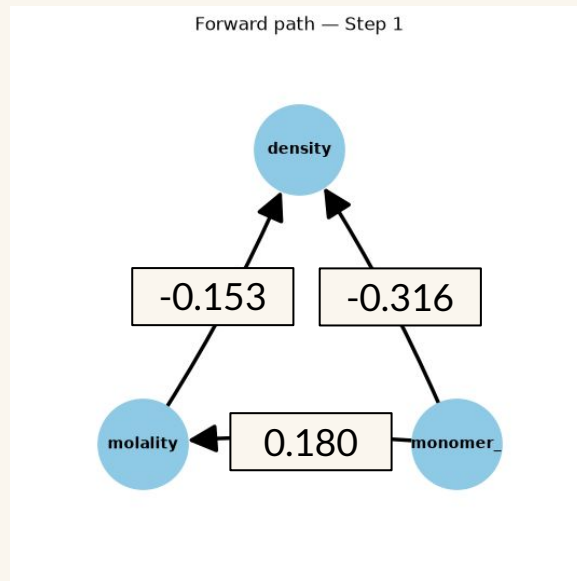
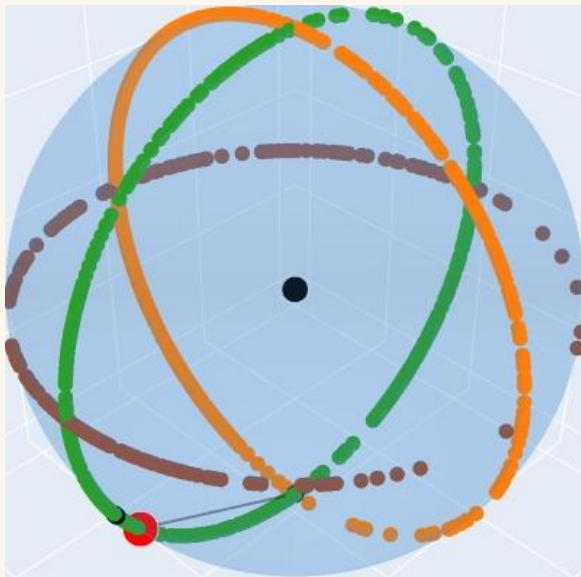


3-Node Pathing (Small Jaccard)



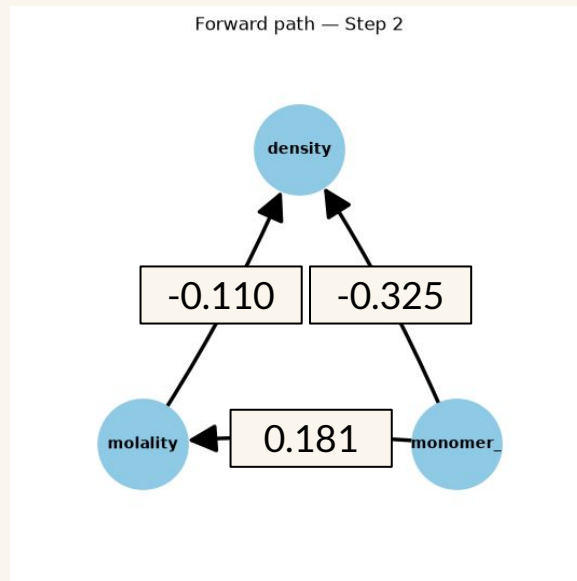
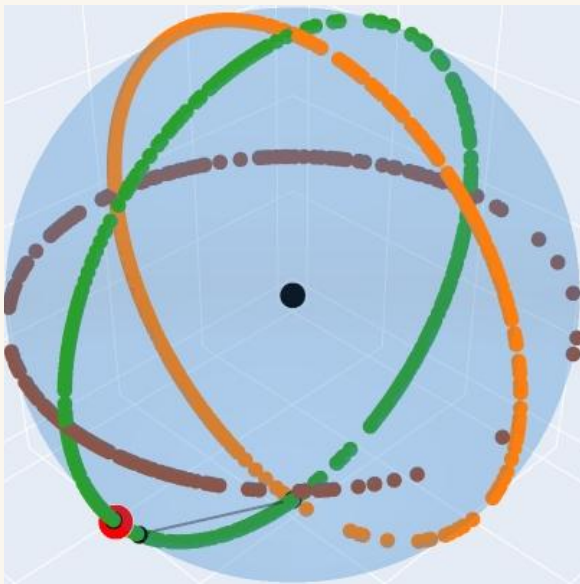
Forward - Step 0; Jaccard threshold 0.1 of mean

3-Node Pathing (Small Jaccard)



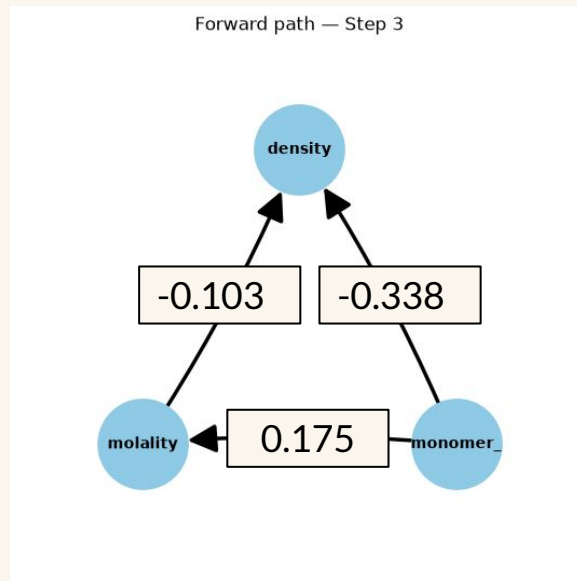
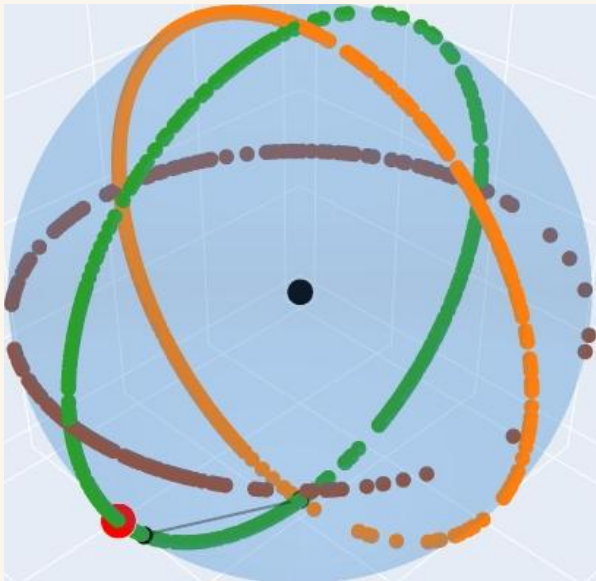
Forward - Step 1; Jaccard threshold 0.1 of mean

3-Node Pathing (Small Jaccard)



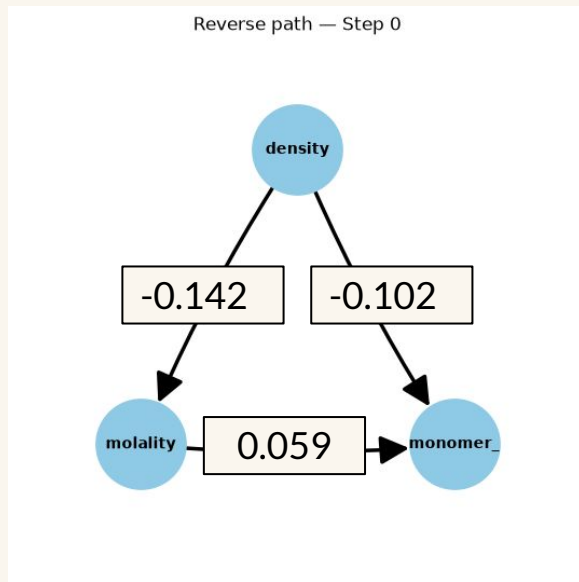
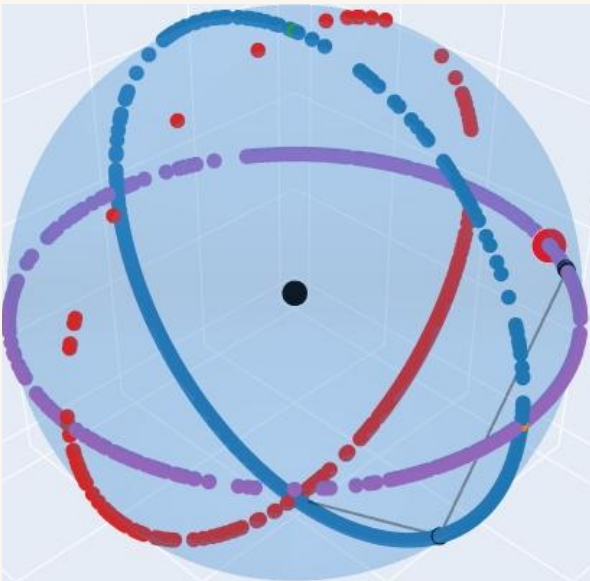
Forward - Step 2; Jaccard threshold 0.1 of mean

3-Node Pathing (Small Jaccard)



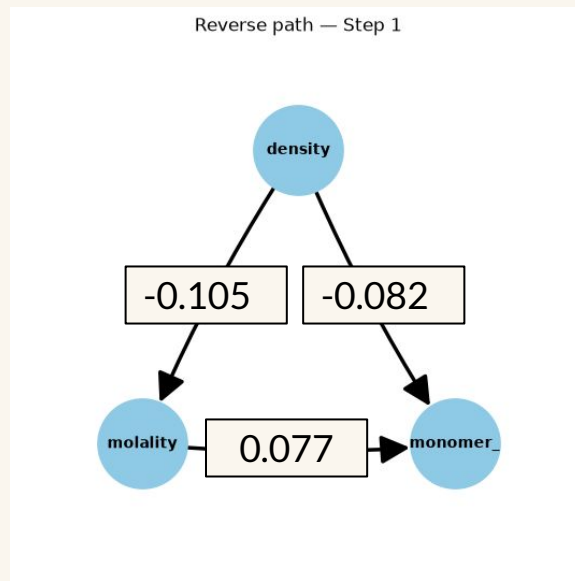
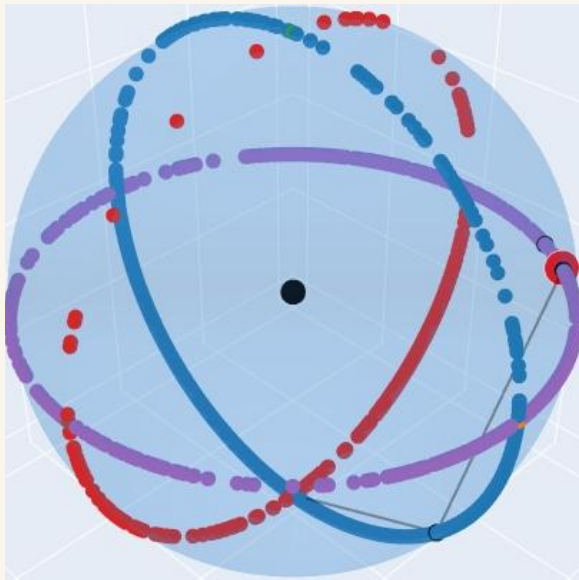
Forward - Step 3; Jaccard threshold 0.1 of mean

3-Node Pathing (Small Jaccard)



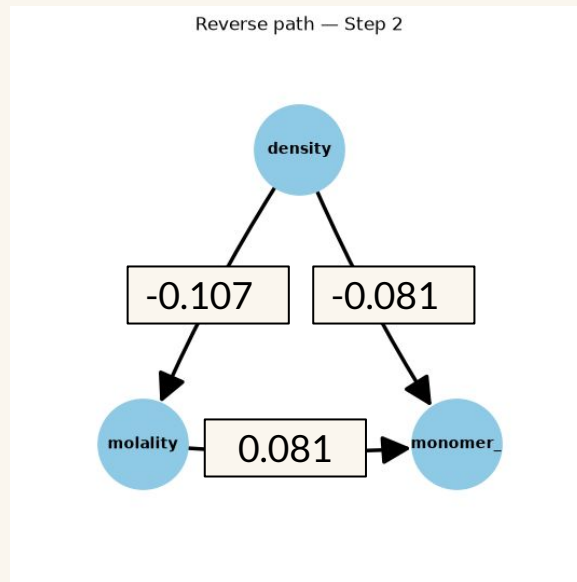
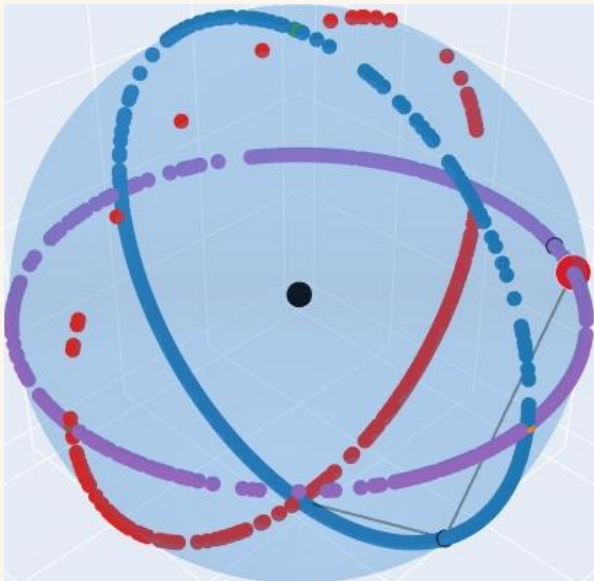
Backward- Step 0; Jaccard threshold 0.1 of mean

3-Node Pathing (Small Jaccard)



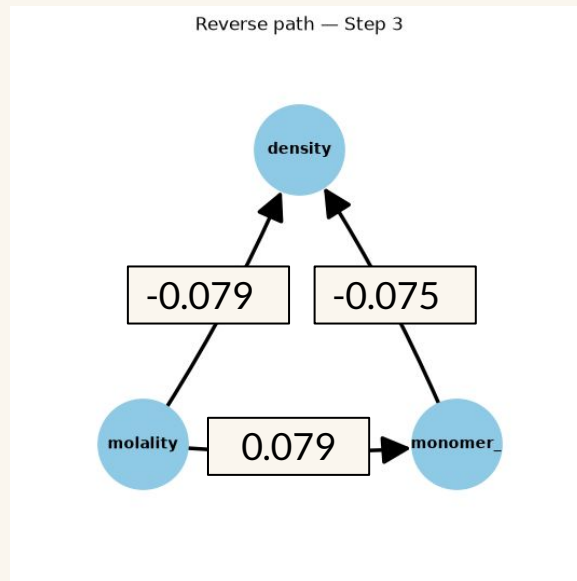
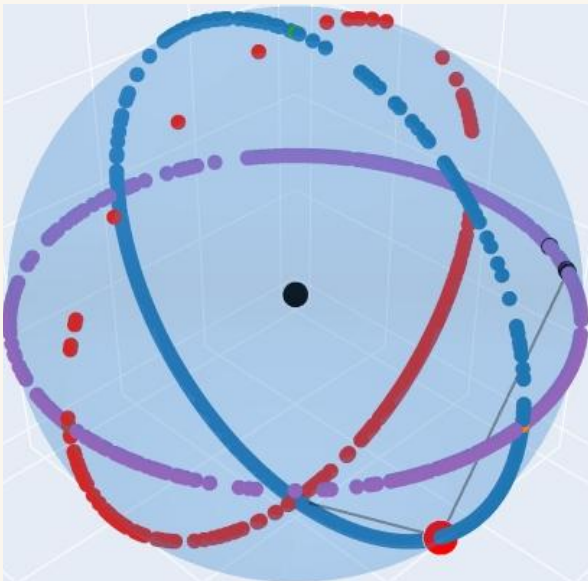
Backward- Step 1; Jaccard threshold 0.1 of mean

3-Node Pathing (Small Jaccard)



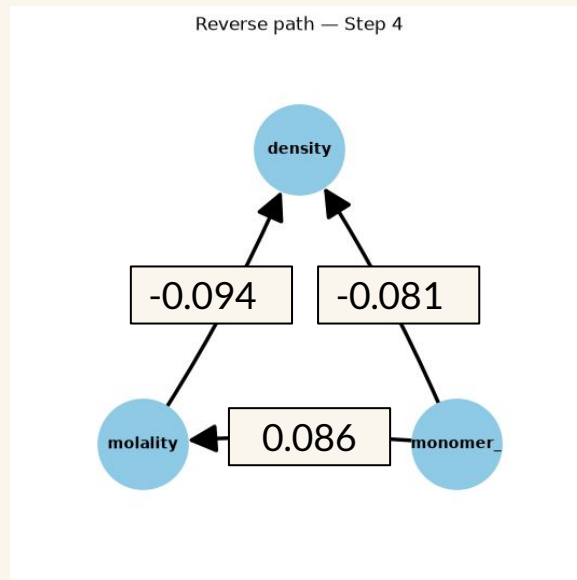
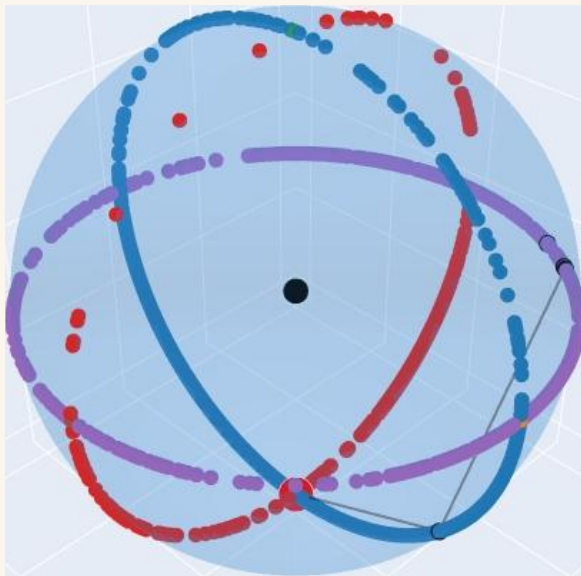
Backward- Step 2; Jaccard threshold 0.1 of mean

3-Node Pathing (Small Jaccard)



Backward- Step 3; Jaccard threshold 0.1 of mean

3-Node Pathing (Small Jaccard)



Backward- Step 4; Jaccard threshold 0.1 of mean

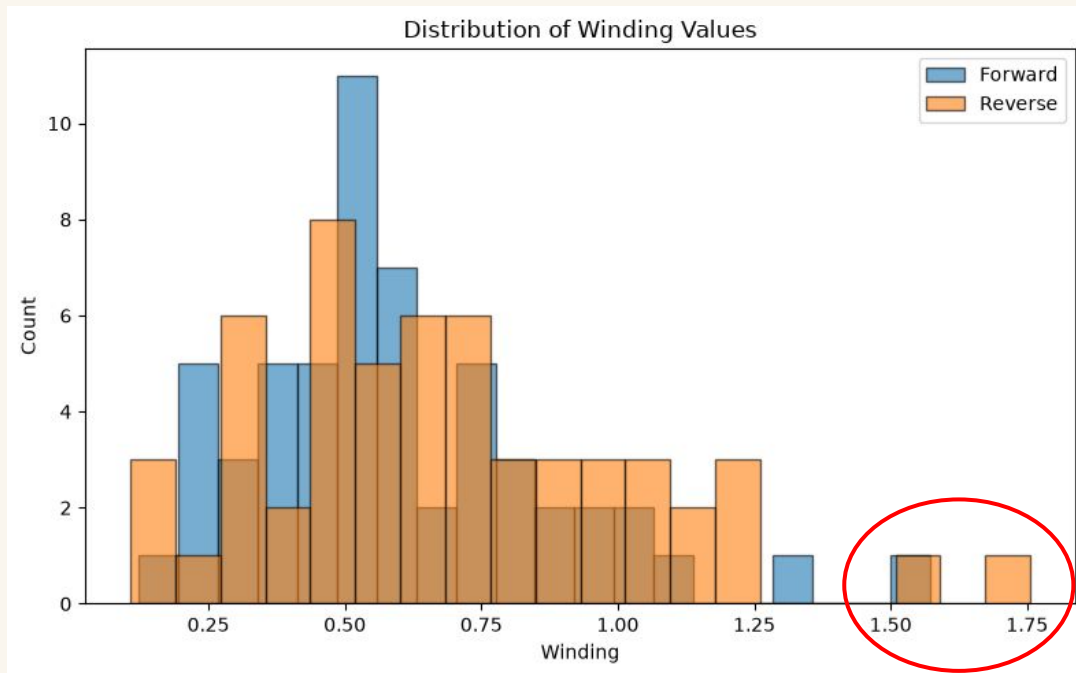
Ranked winding (Molality, Monomer Width, Density)

3-Node Pathing Full Data

Forward			Reverse		
Columns	Winding	Len	Columns	Winding	Len
(M, Li, TN)	1.653	3	(M, MW, D)	1.756	5
(M, MW, TN)	1.573	5	(M, Li, TN)	1.570	5
(MW, D, Li)	1.545	3	(M, MW, TN)	1.217	3
(D, Li, TN)	1.309	3	(M, D, TN)	1.189	3
(D, TN, Cond)	1.074	2	(Li, Poly, TN)	1.182	2
(M, D, TN)	1.051	2	(D, Li, TN)	1.149	3
(MW, TN, Cond)	1.013	4	(M, MW, Cond)	1.132	2
(M, M, TN)	0.982	3	(M, D, Li)	0.982	2
(TFSI, Li, Cond)	0.471	2	(M, Poly, TN)	0.531	2
(MW, Poly, TN)	0.882	4	(MW, D, TFSI)	0.521	3
(MW, TFSI, Poly)	0.864	4	(MW, D, Poly)	0.516	2
(M, D, Cond)	0.810	3	(M, MW, Li)	0.505	2
(M, Li, Poly)	0.729	3	(TFSI, Poly, Cond)	0.489	2
(D, TFSI, Li)	0.706	3	(M, TFSI, Cond)	0.457	2
(M, MW, Poly)	0.700	3	(Li, Poly, Cond)	0.422	2
(M, MW, D)	0.673	4	(M, TFSI, Poly)	0.410	2
(MW, Li, TN)	0.671	7	(M, Li, Poly)	0.380	2
(M, Li, Poly)	0.613	3	(M, TN, Cond)	0.374	5
(TFSI, TN, Cond)	0.609	3	(MW, Li, Poly)	0.321	3
(TFSI, Li, TN)	0.605	2	(Li, TN, Cond)	0.312	2
(M, TN, Cond)	0.600	2	(TFSI, Li, Cond)	0.311	2
(D, Li, Poly)	0.592	4	(D, Poly, Cond)	0.307	2
(M, Poly, Cond)	0.579	2	(MW, Poly, Cond)	0.295	2
(D, TFSI, Cond)	0.577	2	(M, TFSI, TN)	0.265	3
(Poly, TN, Cond)	0.565	2	(M, D, TFSI)	0.252	2
			(MW, TFSI, TN)	0.250	2
			(M, TFSI, Li)	0.211	2
			(MW, TFSI, Cond)	0.196	2
			(MW, Li, Cond)	0.169	2
			(D, TFSI, Poly)	0.554	2
			(MW, D, Cond)	0.543	2
			(TFSI, Li, Poly)	0.534	2
			(M, D, Li)	0.533	2
			(M, Poly, TN)	0.531	2
			(MW, D, TFSI)	0.521	3
			(MW, D, Poly)	0.516	2
			(M, MW, Li)	0.505	2
			(TFSI, Poly, Cond)	0.489	2
			(M, TFSI, Cond)	0.457	2
			(Li, Poly, Cond)	0.422	2
			(M, TFSI, Poly)	0.410	2
			(M, Li, Poly)	0.380	2
			(M, TN, Cond)	0.374	5
			(MW, Li, Poly)	0.321	3
			(Li, TN, Cond)	0.312	2
			(TFSI, Li, Cond)	0.311	2
			(D, Poly, Cond)	0.307	2
			(MW, Poly, Cond)	0.295	2
			(M, TFSI, TN)	0.265	3
			(M, D, TFSI)	0.252	2
			(MW, TFSI, TN)	0.250	2
			(M, TFSI, Li)	0.211	2
			(MW, TFSI, Cond)	0.196	2
			(MW, Li, Cond)	0.169	2
			(D, Poly, Cond)	0.638	4
			(M, D, Li)	0.622	2
			(TFSI, TN, Cond)	0.612	3
			(MW, TN, Cond)	0.599	2
			(D, TFSI, Li)	0.588	3
			(D, Li, Poly)	0.577	2
			(M, MW, TFSI)	0.550	2
			(Poly, TN, Cond)	0.543	4
			(MW, D, Li)	0.526	2
			(M, Li, Poly)	0.519	2
			(D, TN, Cond)	0.515	2
			(M, Poly, Cond)	0.510	2
			(D, Li, Cond)	0.498	2
			(TFSI, Poly, TN)	0.488	2
			(TFSI, Li, Cond)	0.471	2
			(D, TFSI, Poly)	0.452	2
			(M, TFSI, Poly)	0.446	2
			(MW, Poly, Cond)	0.424	2
			(M, TFSI, Li)	0.348	2
			(MW, Li, Cond)	0.330	2
			(MW, TFSI, Poly)	0.277	2
			(TFSI, Poly, Cond)	0.273	3
			(M, D, TFSI)	0.245	2
			(MW, D, Poly)	0.213	2
			(MW, D, TFSI)	0.163	2
			(Li, TN, Cond)	0.124	2
			(MW, TFSI, Li)	0.111	2
			(Li, Poly, Cond)	0.095	2

Ranked winding (TFSI, Li Diffusivity, Conductivity)

Winding Numbers as Flexibility



Winding histogram w/outliers

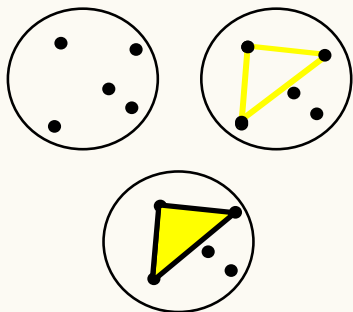
Open Questions and Potential Directions

Higher Homology

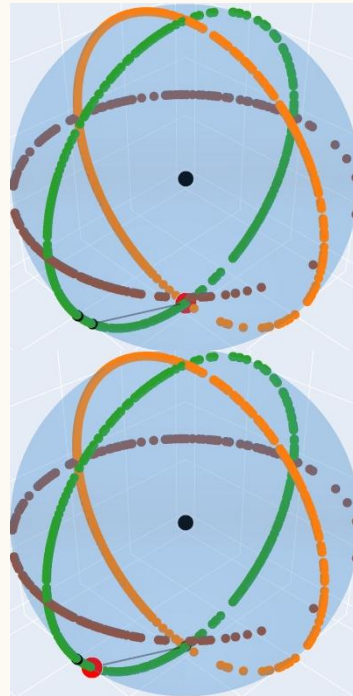
Q: What are the higher homology groups of the link complex?

Q: How does data fill that space?

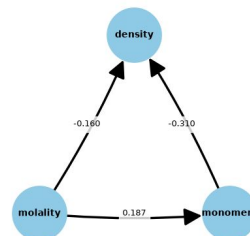
Quillen Fiber Approach
Persistent Homology on Smaller Cells



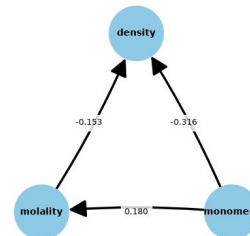
Better Winding Detection



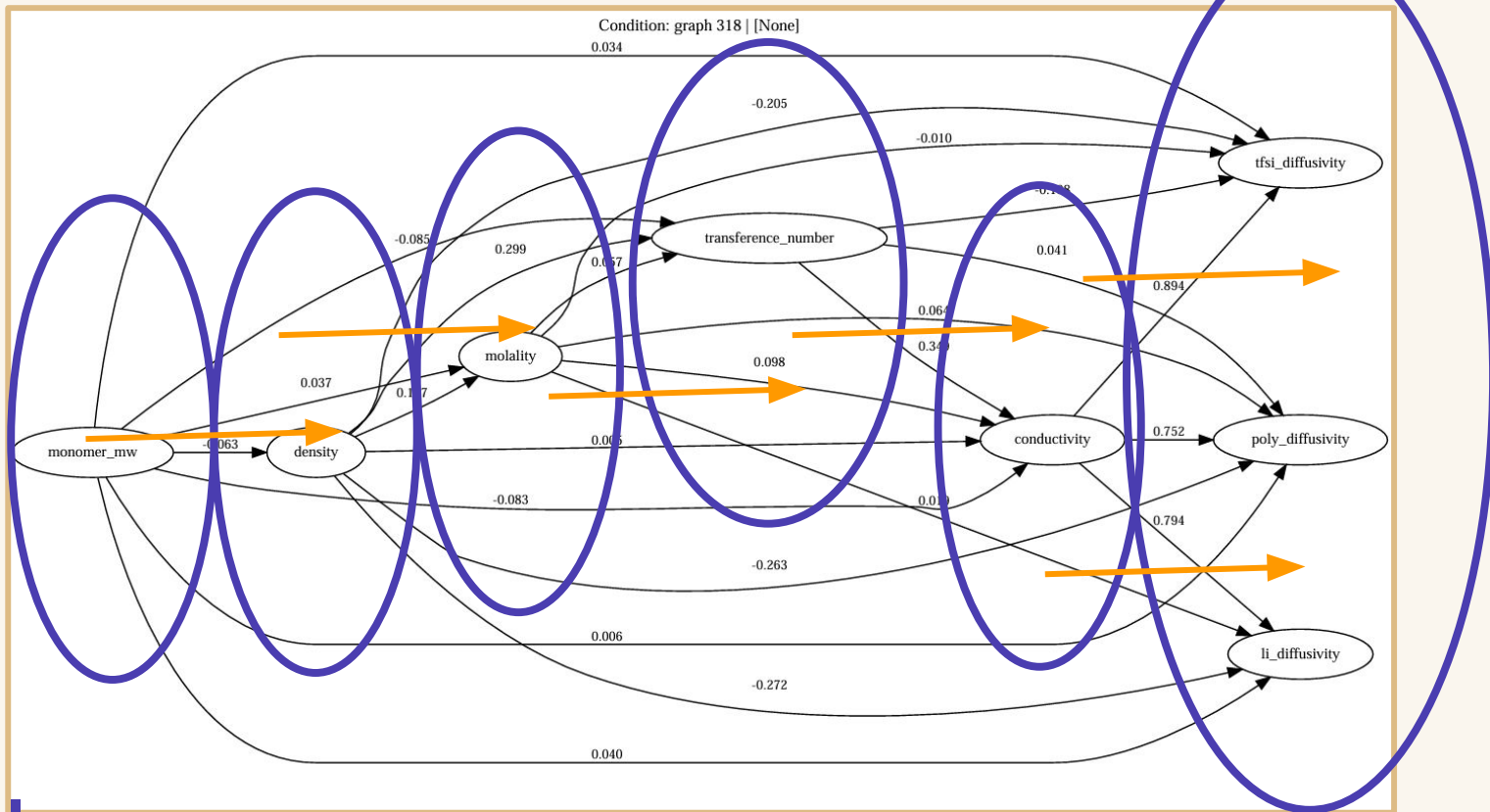
Forward path — Step 0



Forward path — Step 1



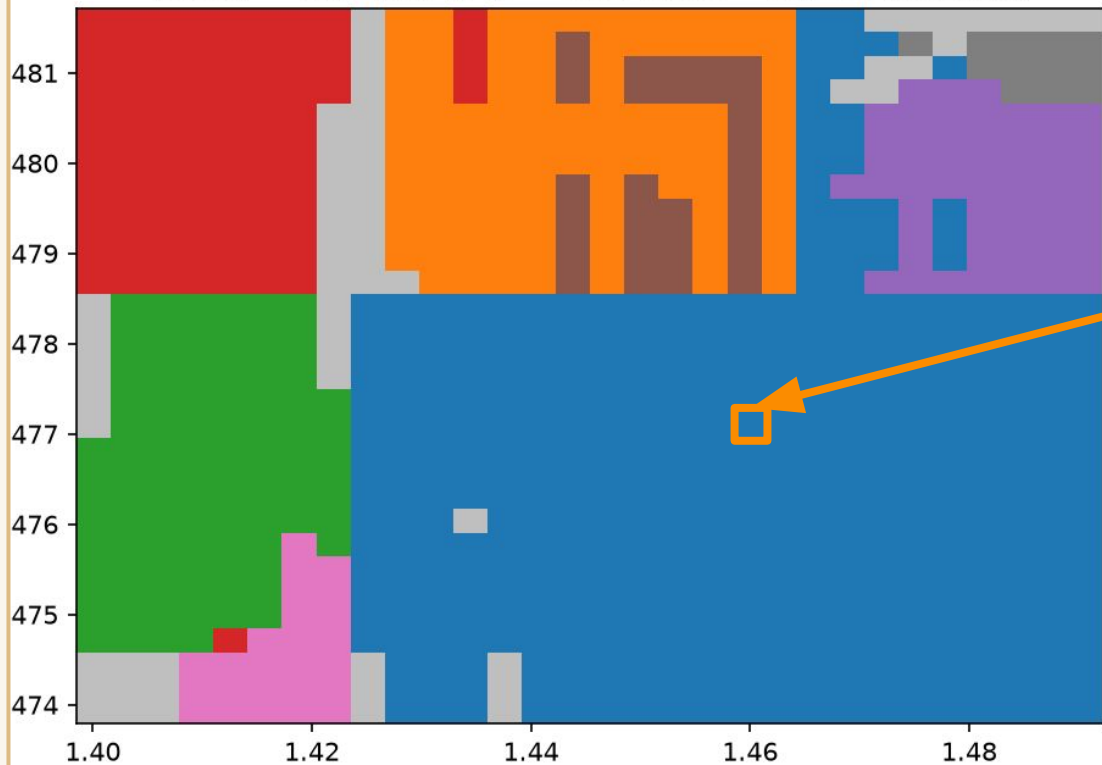
Topological Ordering



Question: When does causal ordering **change**?

Color by Topological Ordering

Causal ORDERING vs thresholds
(colour = ordering, grey = rare, white = subset too small)



X axis : Molality < X

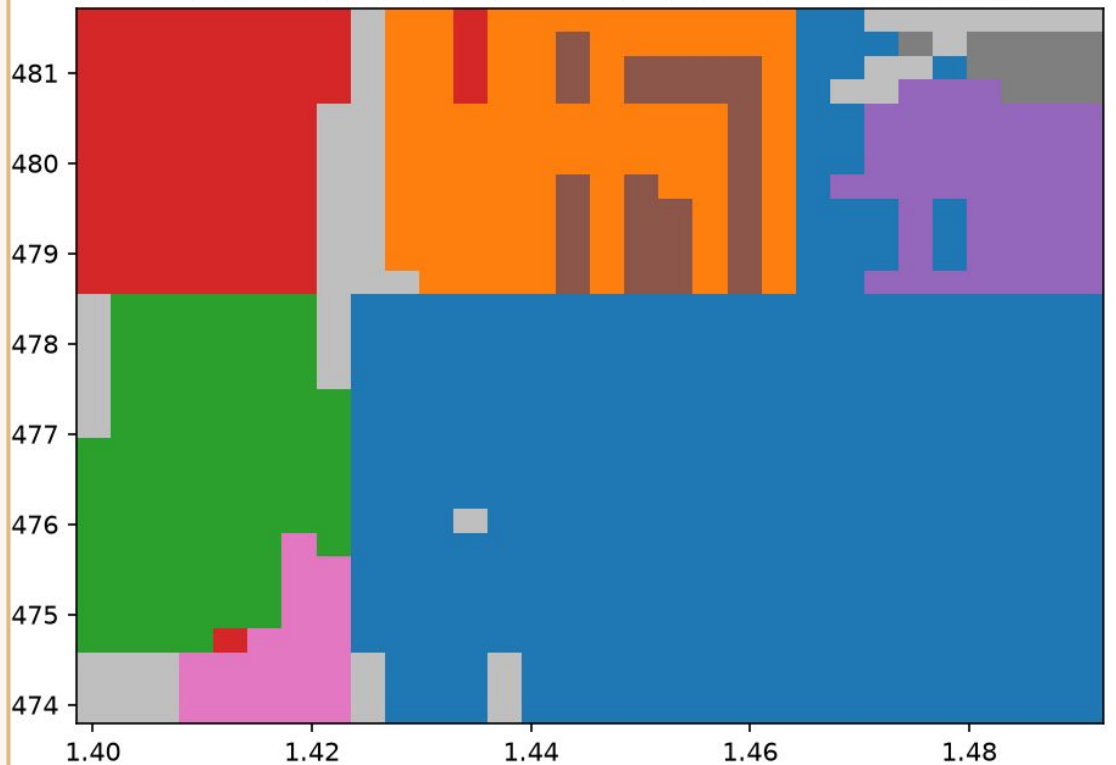
Y axis: monomer_mw < Y

Ex: Molality < 1.46 and
monomer_mw < 477

Color = Topological Ordering

Color by Topological Ordering

Causal ORDERING vs thresholds
(colour = ordering, grey = rare, white = subset too small)



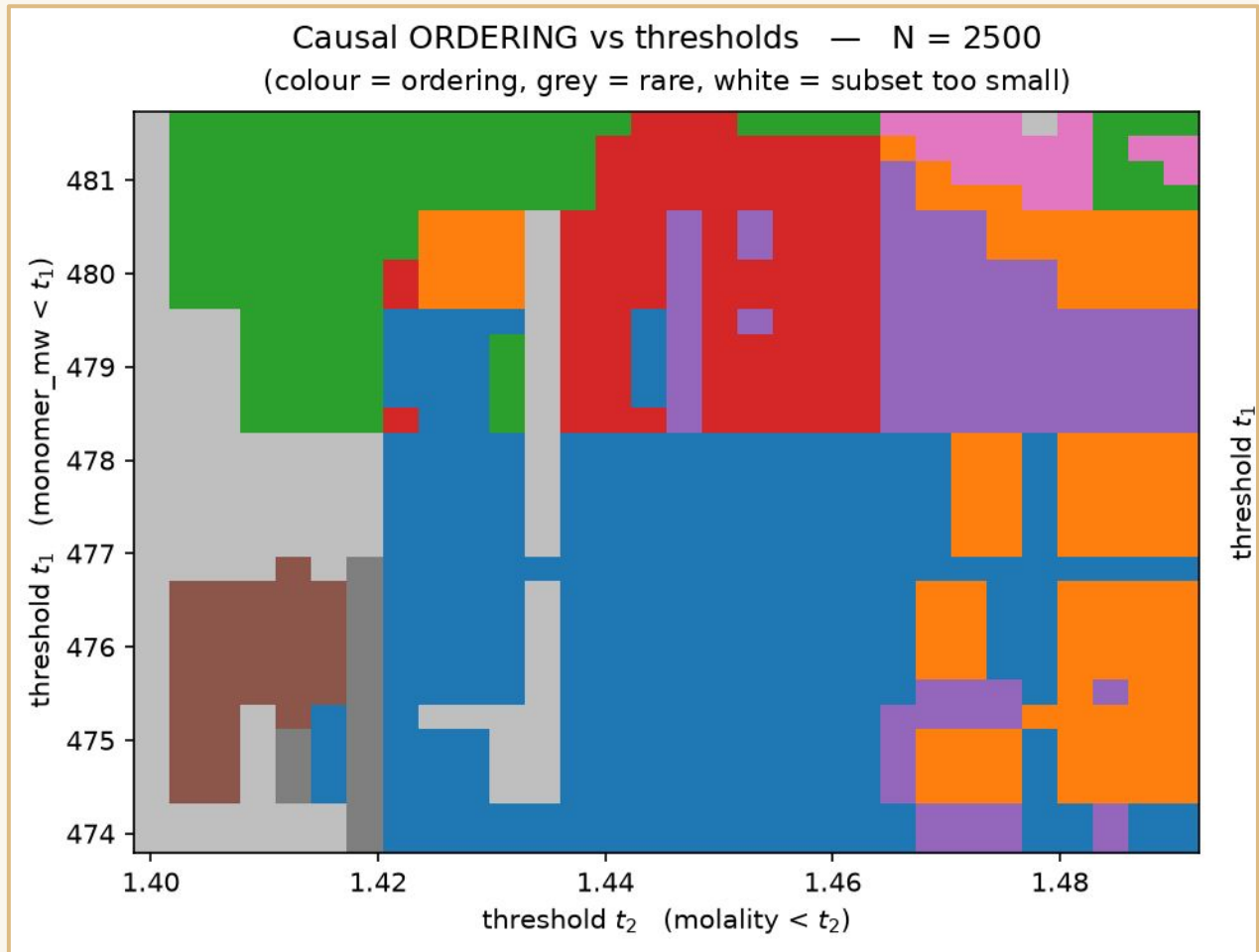
Regions are **Causal Regimes**.

Boundaries are **potential discoveries**

Topological Orders

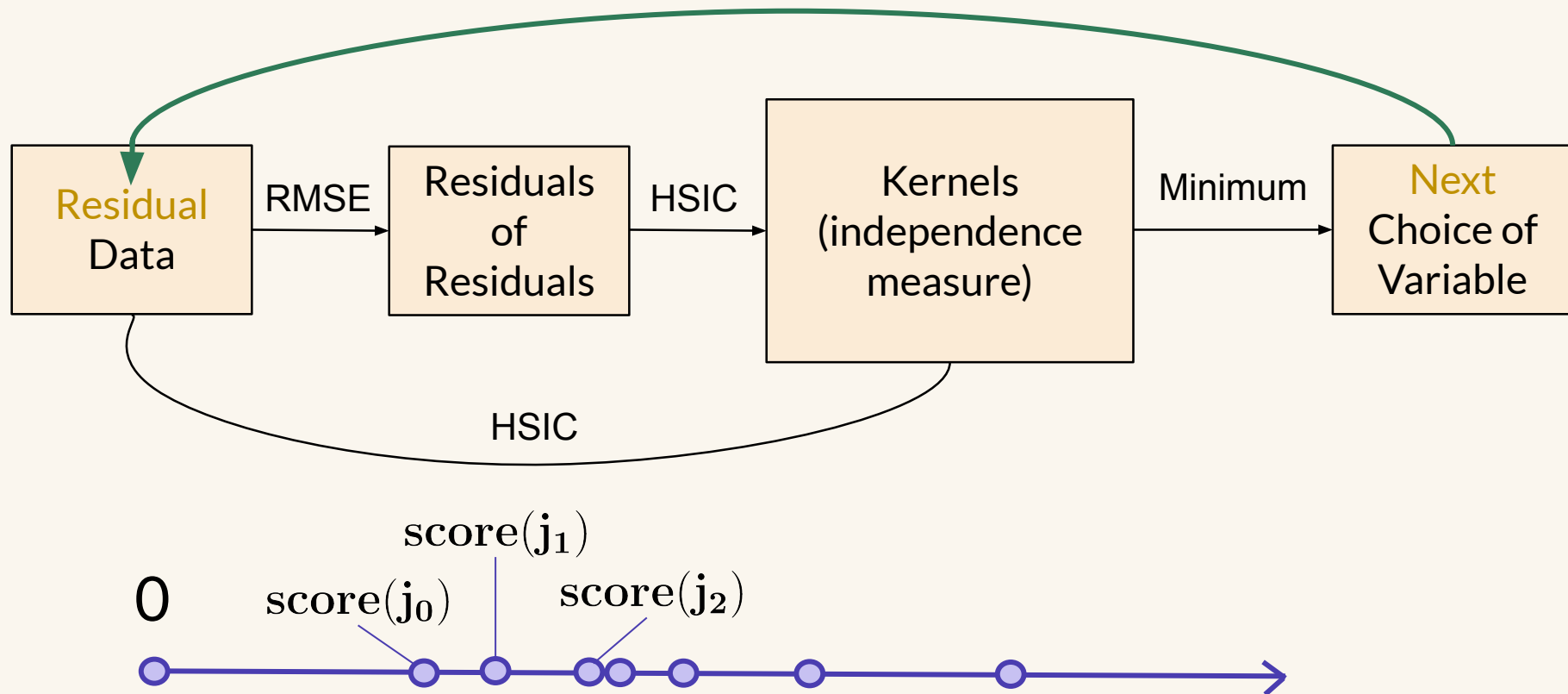
Convergence to causal regimes.
Convergence is convincing

Only **two** of **several** dimensions.
True geometry is complicated

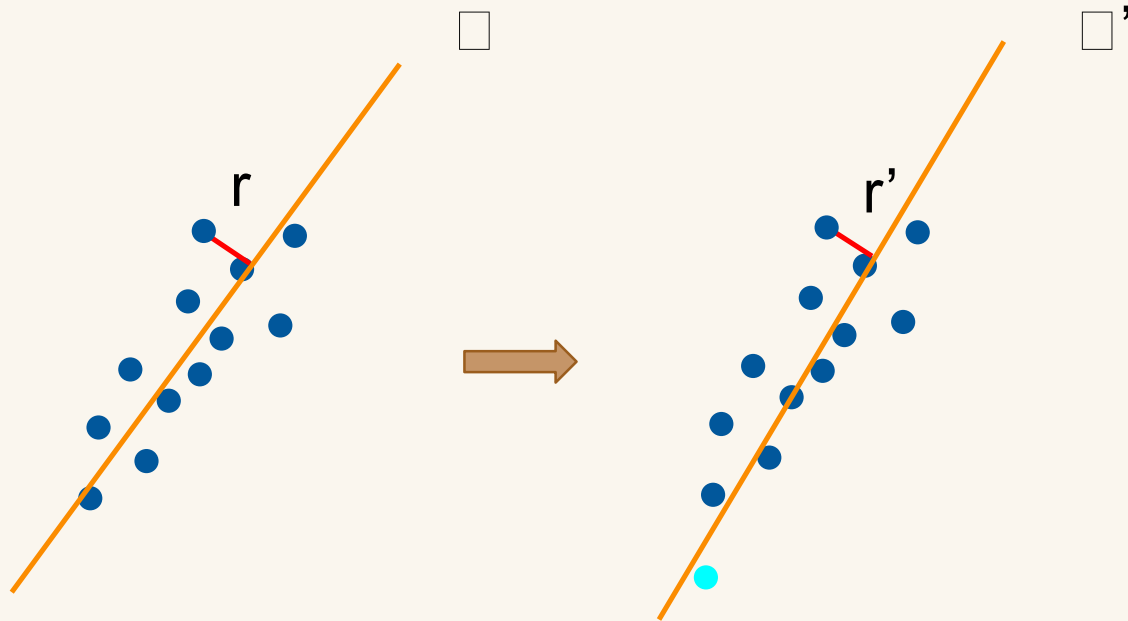


Causal Regimes as sample size increases

Direct LiNGAM Bound

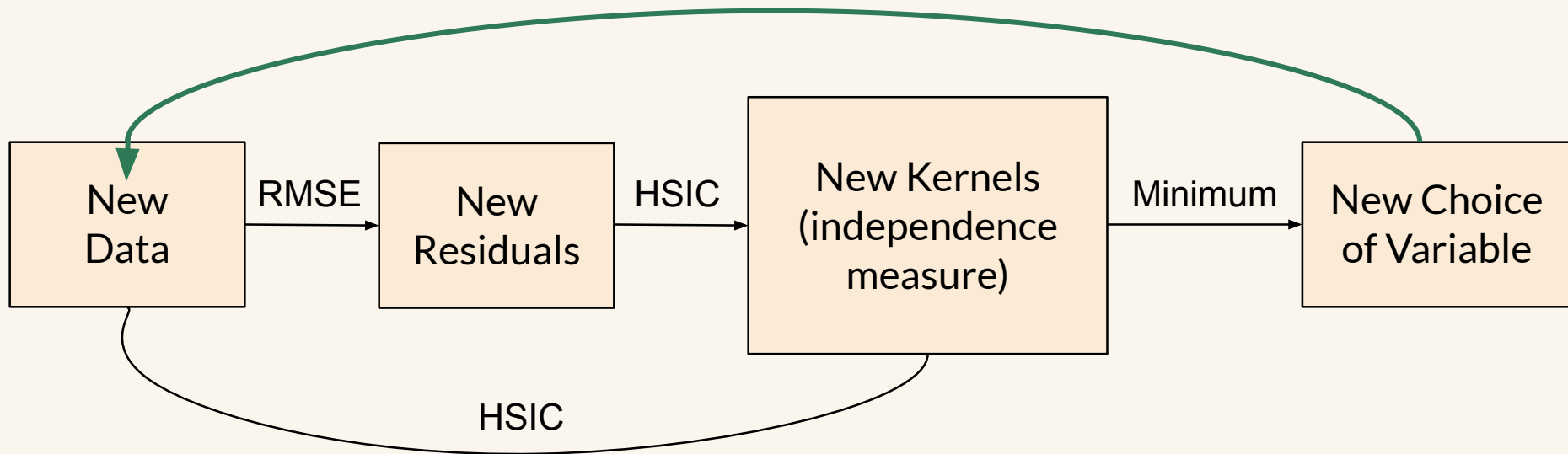


Direct LiNGAM Bound



| Change by adding a new datapoint

Direct LiNGAM Bound



Q: When old ordering = new ordering

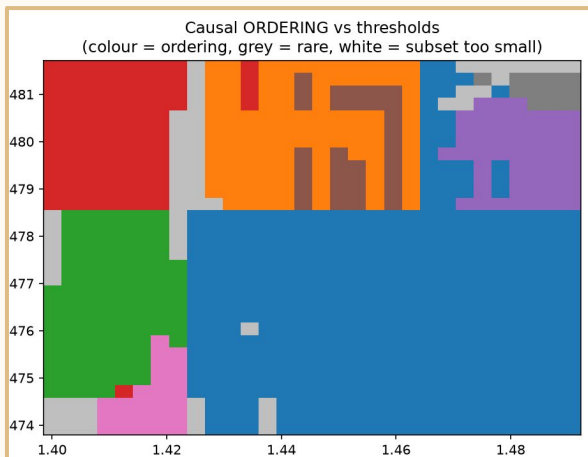
A: We have a bound

Future Work

Formalization condition space operations

For users and developers to approach discoverability.

Construct a formal way to measure dataset.



Theoretical Guarantees

Develop stability and rigid theory for convincingness.

Universal theory

Develop theory perfectly fits the assumption of DirectLiNGAM

Summary

Conditional Pathing

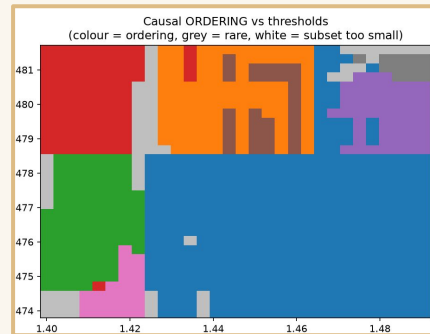
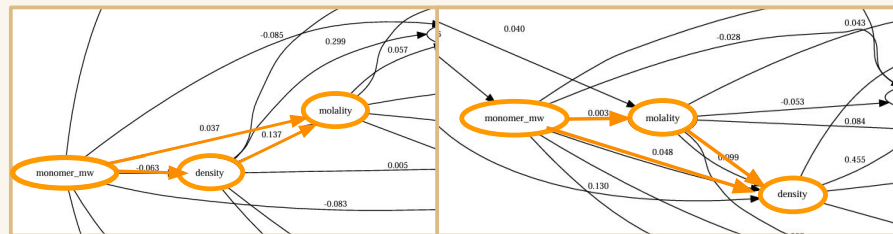
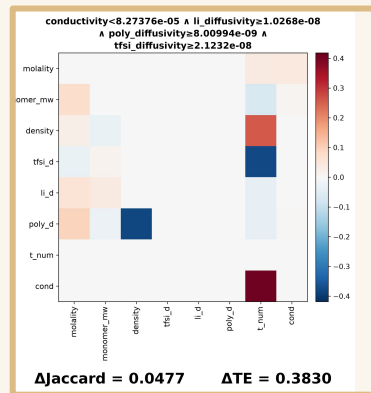
Highlight primary mechanism
Discoverability
Comprehensiveness $\leftrightarrow$ compute

Local Conditional Pathing

Substructure : discoverability
Winding number : convincingness

Topological Ordering

Change of causal order : discoverability



Questions?

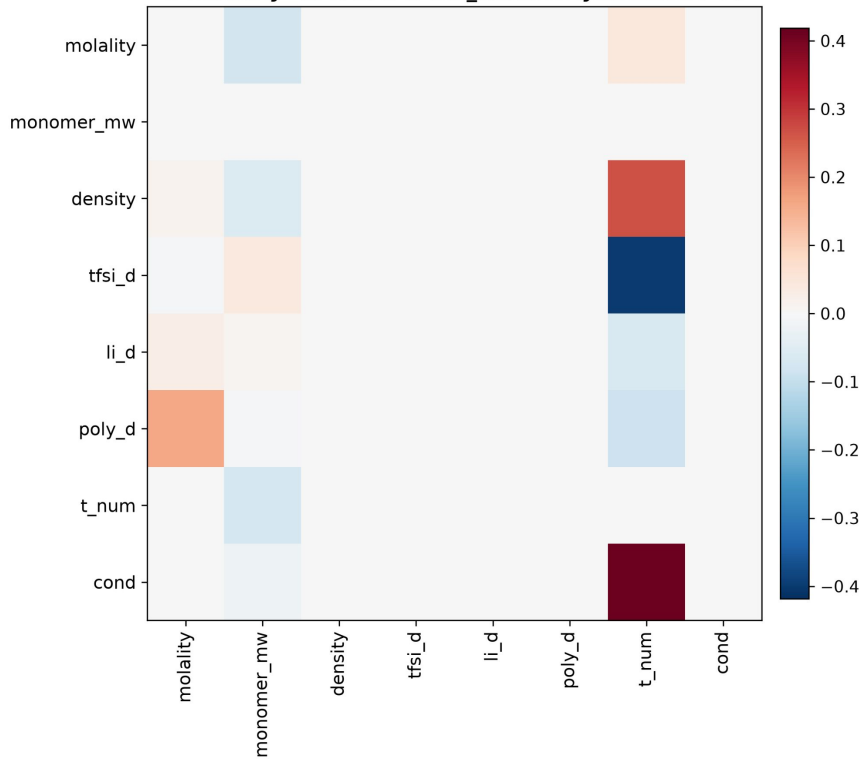


Technical Slides

Technical Slides

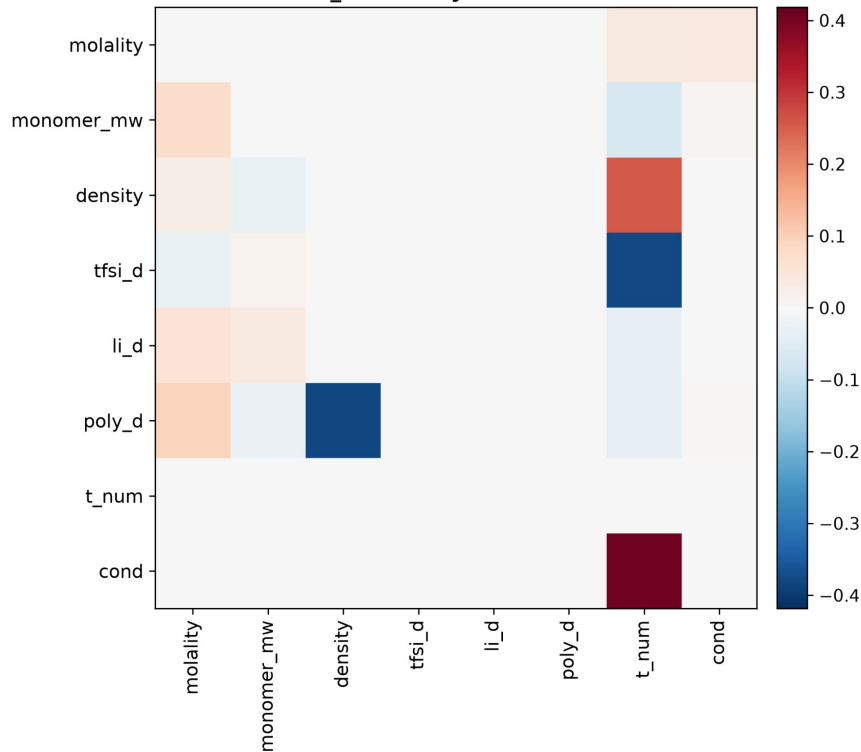
Technical Slides

**conductivity < 8.27376e-05 $\wedge$ li_diffusivity $\geq$ 1.0268e-08
 $\wedge$ molality < 1.47172 $\wedge$ tfsi_diffusivity $\geq$ 2.1232e-08**



(start of path)

**conductivity < 8.27376e-05 $\wedge$ li_diffusivity $\geq$ 1.0268e-08
 $\wedge$ poly_diffusivity $\geq$ 8.00994e-09 $\wedge$
tfsi_diffusivity $\geq$ 2.1232e-08**



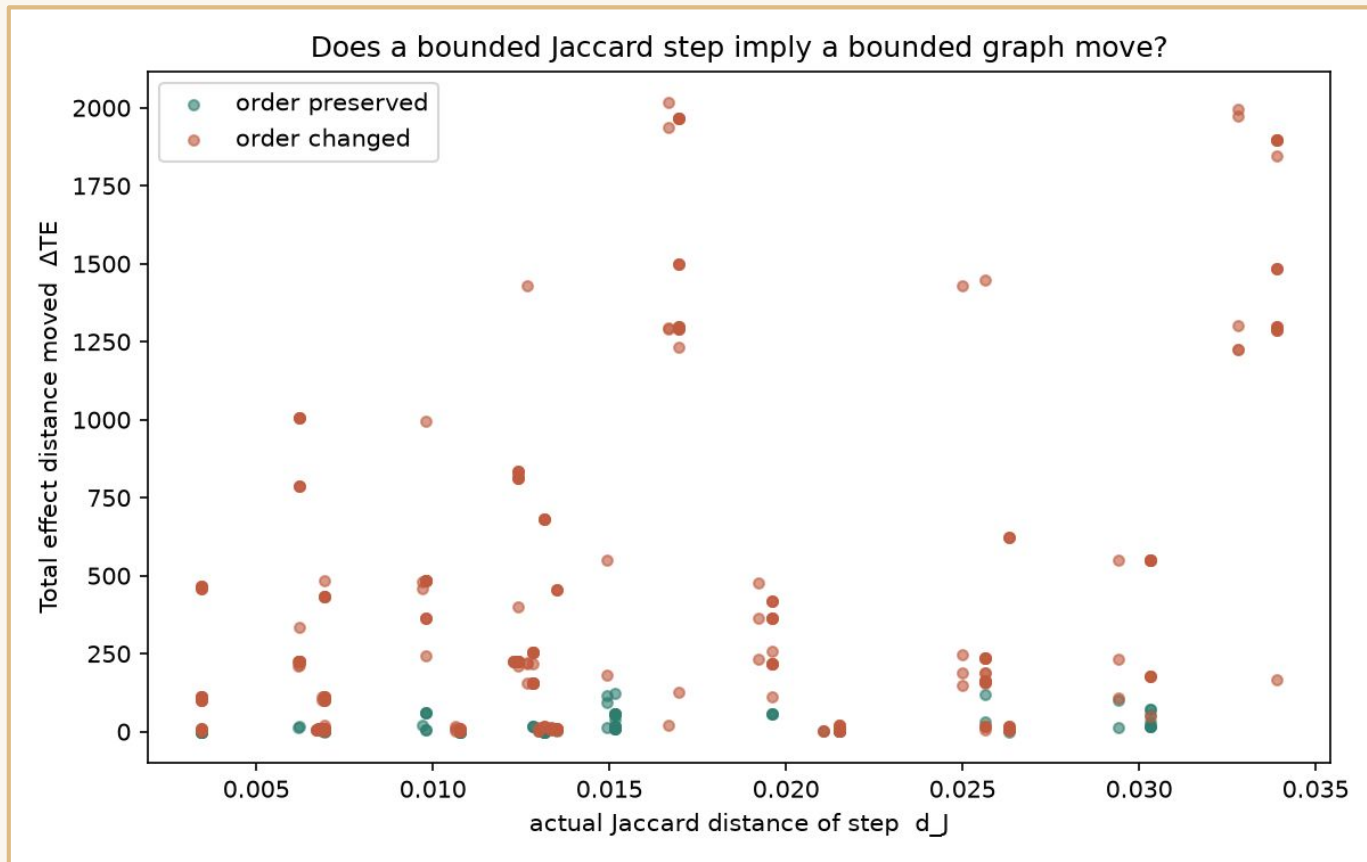
Δ Jaccard = 0.0477

Δ TE = 0.3830

Technical Slides

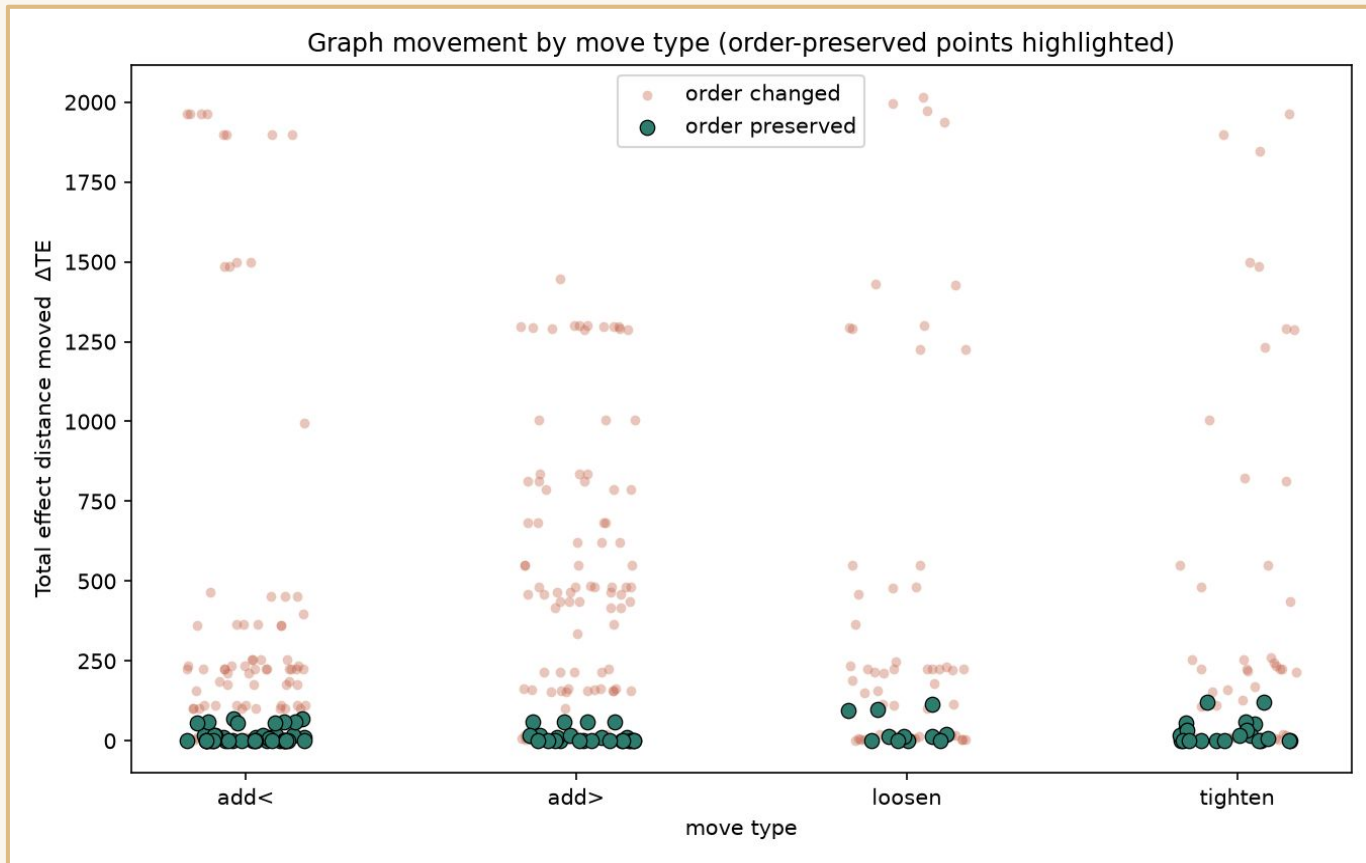
Tie Set detection

Sample size
matters here



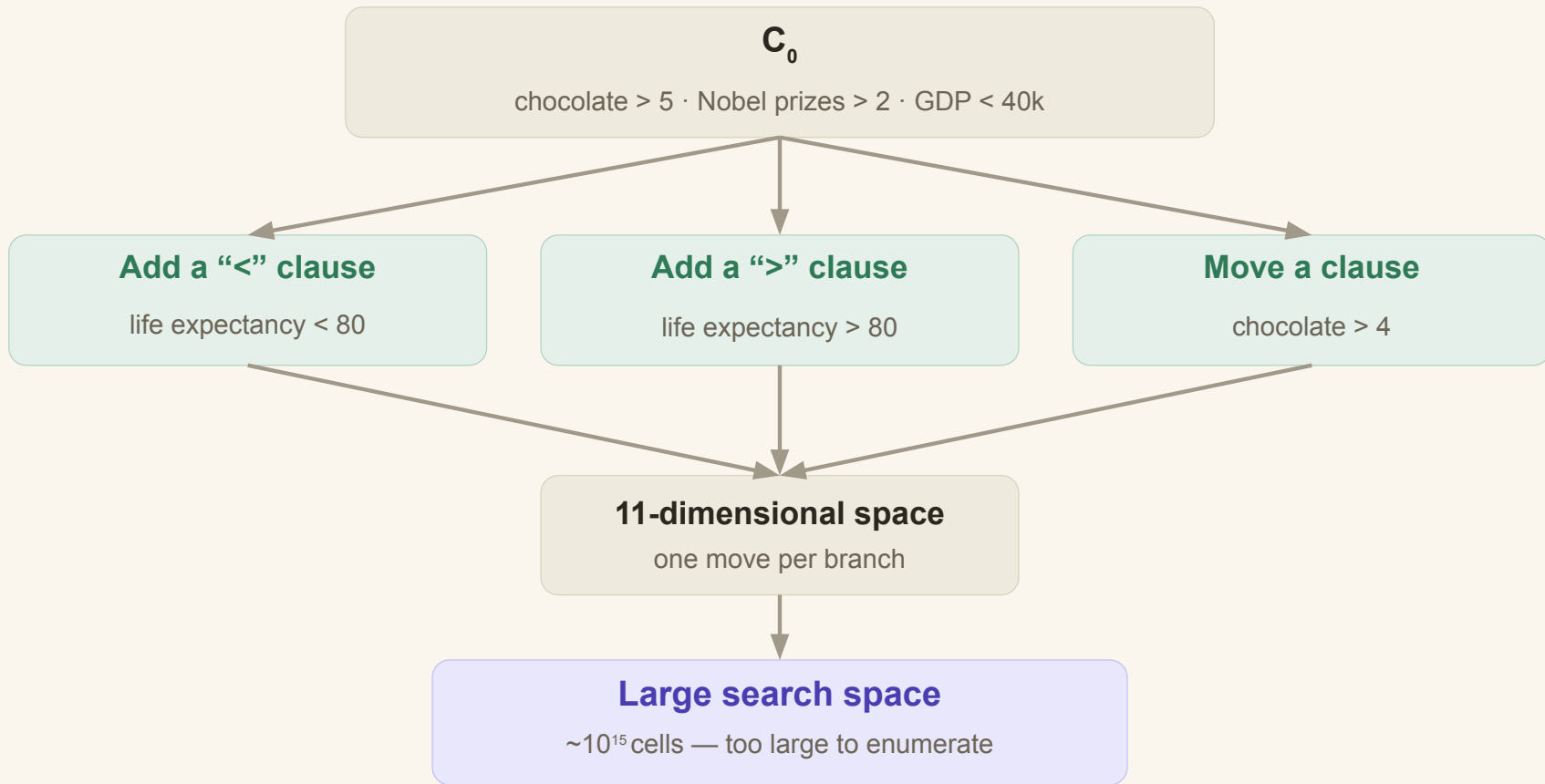
Tie Set Detection

There's stability
when **the causal
ordering is
preserved.**

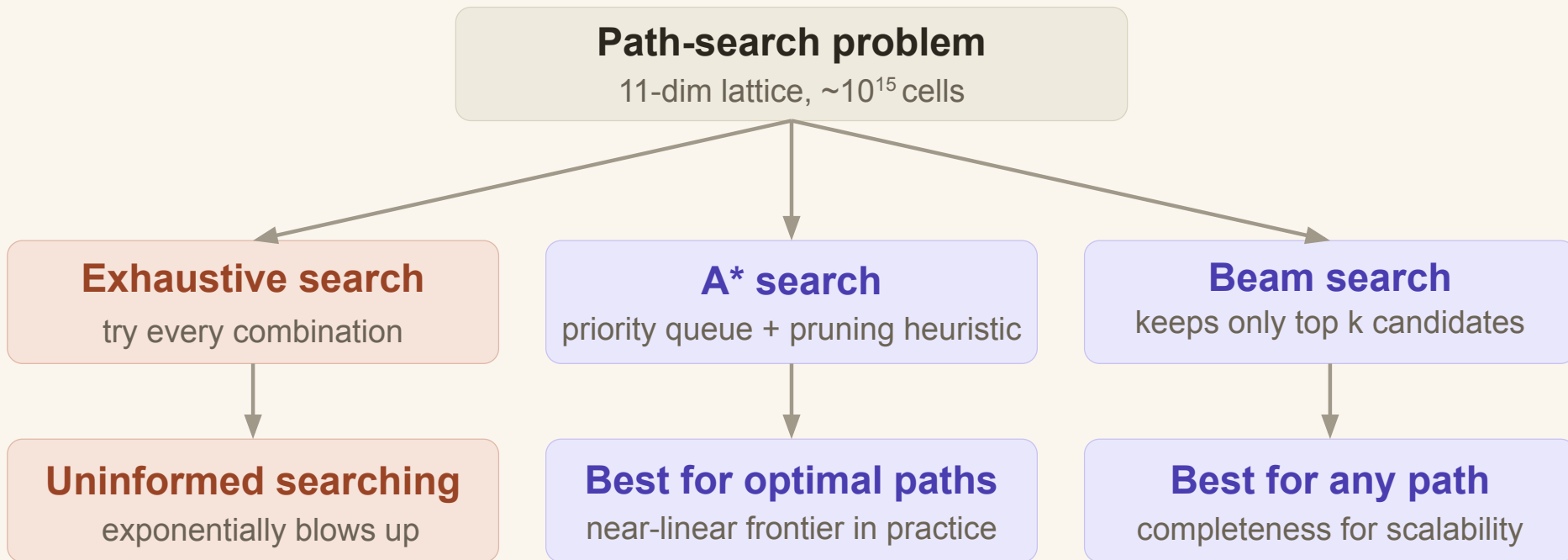


Technical Slide: The Condition Lattice

7 features — chocolate consumed, Nobel prizes, GDP, ...

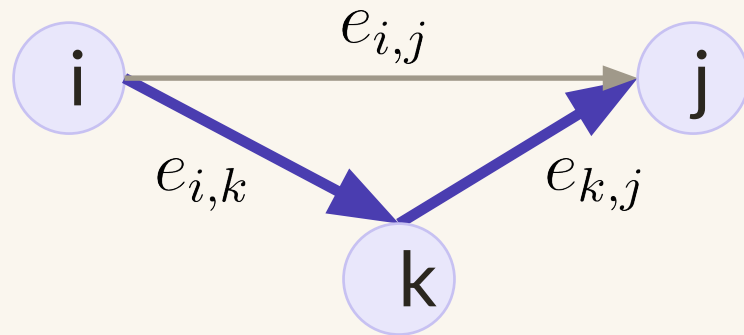
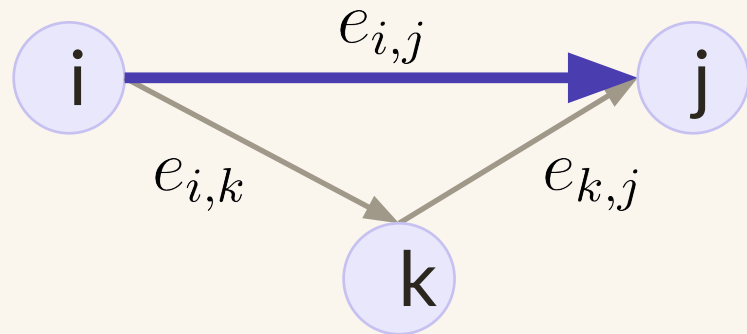


Technical Slide: Graph Searching

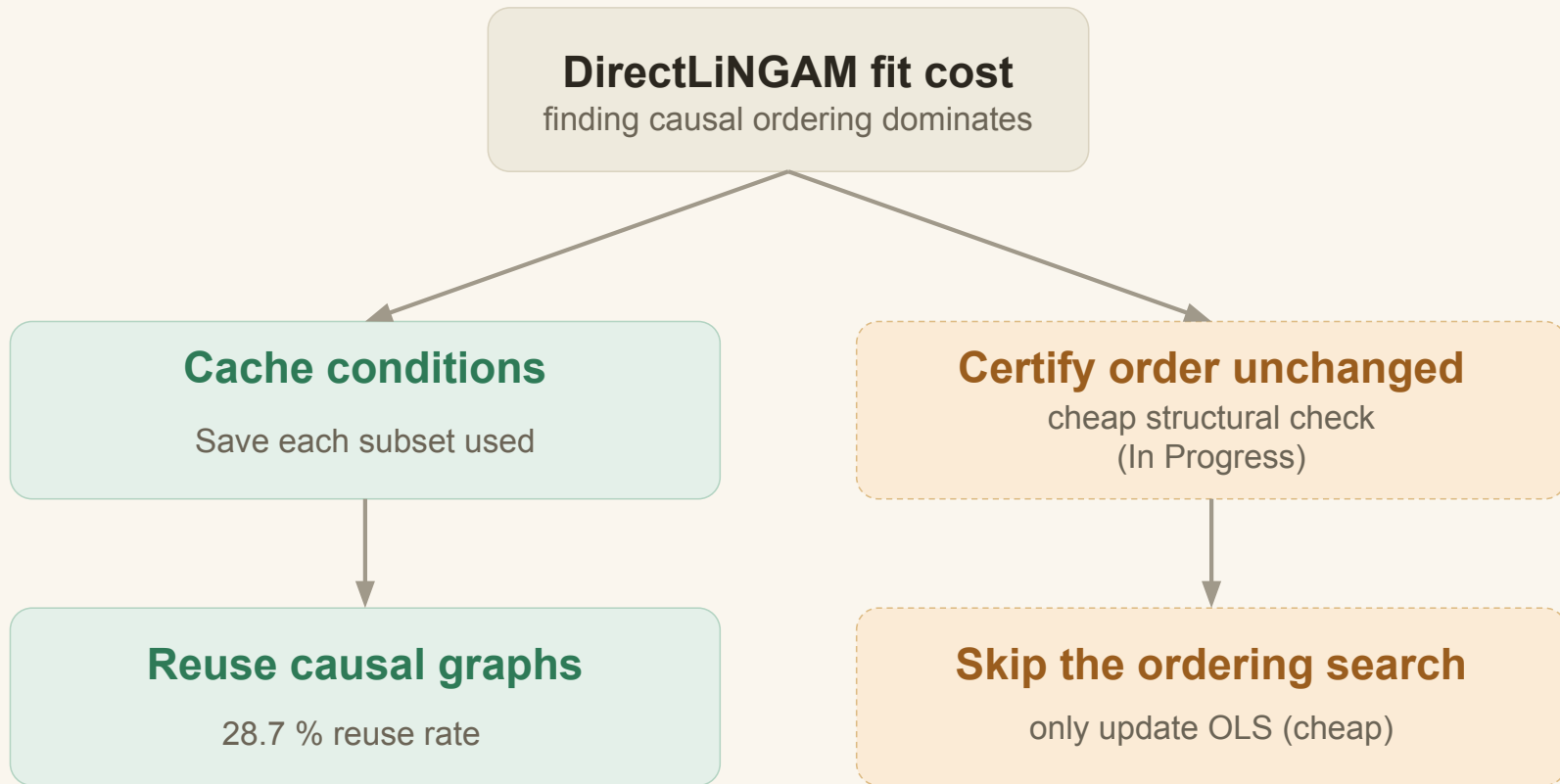


Technical Slide: TE Metric

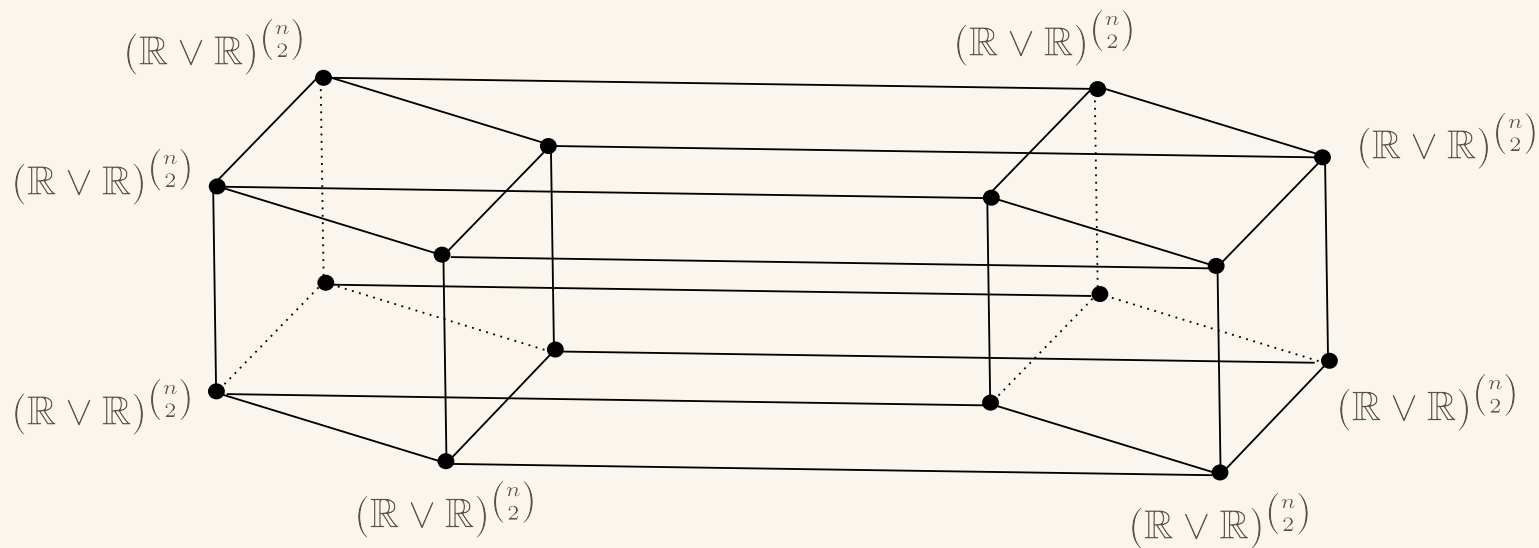
- The $e_{i,j}$ are actually the entries of adjacency matrix of DAG
- $\mathbf{P}(\mathbf{G}) = (\mathbf{I} - \mathbf{M}(\mathbf{G}))^{-1}$
- $\mathbf{P}(\mathbf{G})$ is a non-degenerate DAG space.



Technical Slides: Optimizations



Technical Slides



Technical Slides

$$\begin{aligned} \epsilon_k &> \sum_{i \in U, i \neq k} \left(\frac{2n-1}{n^2(n-1)^2} \text{tr}(K_1 K_2) + \frac{2n}{n^2(n+1)^2(n-1)^2} (\mathbf{1}_n^T K_1 \mathbf{1}_n)(\mathbf{1}_n^T K_2 \mathbf{1}_n) + \frac{2}{(n+1)n^2} (k_1^T K_2 \mathbf{1} + K_1 k_2 \mathbf{1}_n^T) \right) \\ &\quad + \sum_{i \in U, i \neq j} \left(\frac{6n-2}{n^2(n+1)(n-1)^2} \mathbf{1}_n^T K_1 K_2 \mathbf{1}_n \right). \end{aligned}$$

$$(k_1)_k = e^{\frac{-||x_{(k,i)} - l_i||^2}{2\sigma^2}}, \quad (k_2)_k = e^{\frac{-||x_{(k,i)} - \hat{b}x_{(k,j)} - (l_i - \hat{b}l_j)||^2}{2\sigma^2}} = e^{\frac{-||x_{(k,i)} - l_i - \hat{b}(x_{(k,j)} - l_j)||^2}{2\sigma^2}}.$$

$$\left| \hat{\beta}' - \hat{\beta} \right| \leq \frac{\frac{n}{n+1} \left(|I||J| + |\hat{\beta}|J^2 \right)}{S_{xx}}.$$

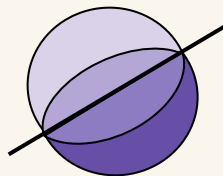
Approach 2: Topology of Causal Graph Space

$$\text{DAG}_n = \{M \in M_n(\mathbb{R}) \mid \forall k \in \mathbb{N}, \forall i, j \in [n], (M^k)_{i,i} = 0, M_{i,j} \neq 0 \implies M_{j,i} = 0\}$$

$$L_n = S^{n(n-1)-1} \cap \text{DAG}_n$$



$$L_n \simeq \bigsqcup_{m=1}^n S_+^{m-1}$$



A more precise formulation

Approach 2: Topology of Causal Graph Space

$$0 \rightarrow C_n \xrightarrow{\partial_n} C_{n-1} \xrightarrow{\partial_{n-1}} \dots \xrightarrow{\partial_2} C_1 \xrightarrow{\partial_1} C_0 \rightarrow 0$$

$$\partial_n : C_n \rightarrow C_{n-1}, G \mapsto \sum_{e \in G} (-1)^{\text{suborder}(e)} G \setminus \{e\}$$



$$\partial_m = \begin{pmatrix} 1 & 1 & 1 & \dots \\ -1 & -1 & -1 & \dots \\ 1 & 0 & 0 & \dots \\ 0 & 1 & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad (\# \text{ of } (m+1)\text{-edge DAGs}, \# \text{ of } m\text{-edge DAGs})$$

Linking space chain complex

Approach 2: Topology of Causal Graph Space

$$\partial_2 = \left(\begin{array}{cccc|c} 1 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 1 & \cdots \\ -1 & 0 & -1 & 0 & \cdots \\ 0 & -1 & 0 & -1 & \cdots \\ \hline \vdots & \vdots & \vdots & \vdots & \ddots \end{array} \right)_{12 \times 48}$$

$$\beta_0(L_3) = \dim C_0 - \text{rank}(\partial_1) = 12 - 11 = 1,$$

$$\beta_1(L_3) = \dim C_1 - \text{rank}(\partial_1) - \text{rank}(\partial_2) = 48 - 11 - 36 = 1,$$

$$\beta_2(L_3) = \dim C_2 - \text{rank}(\partial_2) - \text{rank}(\partial_3) = 48 - 36 - 0 = 12,$$

$$\beta_k(L_3) = 0, \quad k \geq 3.$$

$$H_0(L_3; \mathbb{Z}) \cong \mathbb{Z},$$

$$H_1(L_3; \mathbb{Z}) \cong \mathbb{Z},$$

$$H_2(L_3; \mathbb{Z}) \cong \mathbb{Z}^{12},$$

$$H_k(L_3; \mathbb{Z}) \cong 0, \quad k \geq 3.$$

Computing homology for $n=3$

Approach 2: Topology of Causal Graph Space

$$\Delta_n^{\text{DAG}} \simeq S^{n-2} \implies H_{n-2}(\Delta_n^{\text{DAG}}; \mathbb{Z}) \cong \mathbb{Z} \implies H_{n-2}(\tilde{\Delta}_n^{\text{DAG}}; \mathbb{Z}) \neq 0.$$

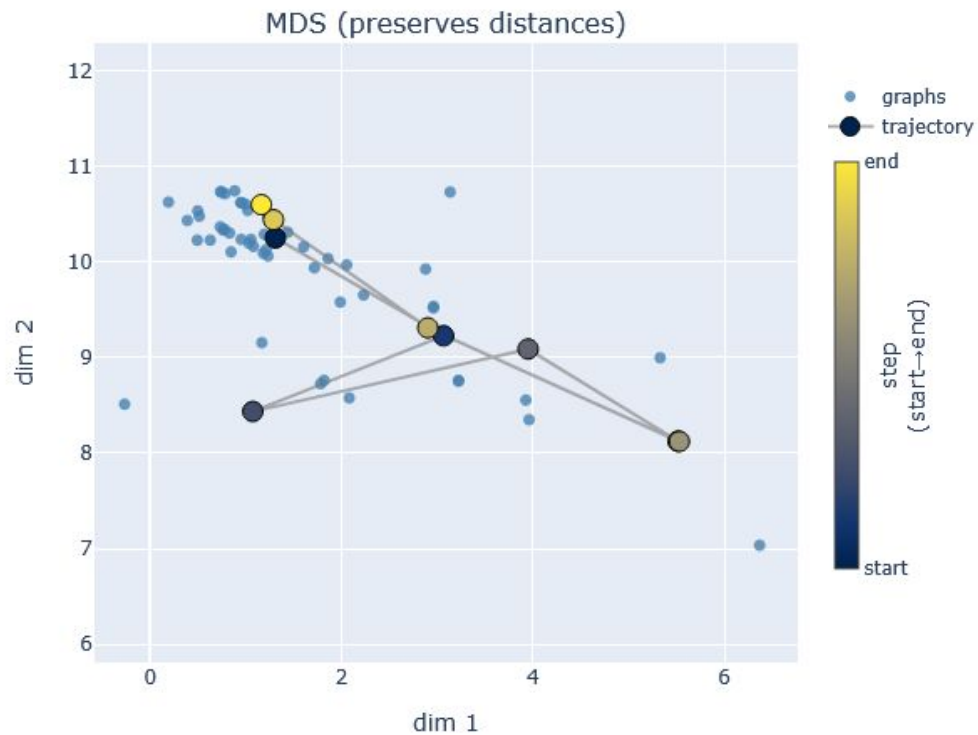
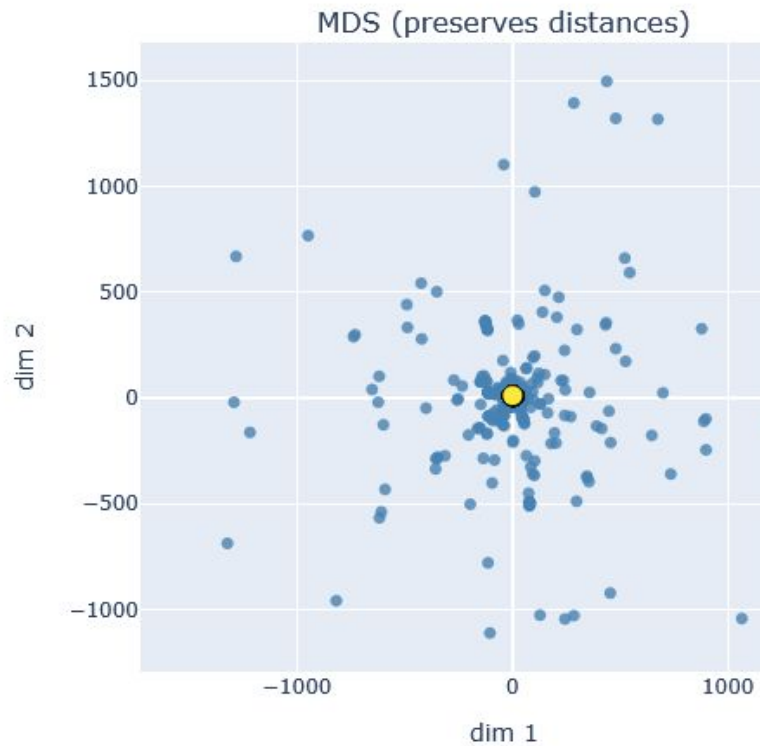
(Bjorner & Welker, 1999)

$$s_* : H_k(\Delta_n^{\text{DAG}}; \mathbb{Z}) \hookrightarrow H_k(\tilde{\Delta}_n^{\text{DAG}}; \mathbb{Z}).$$

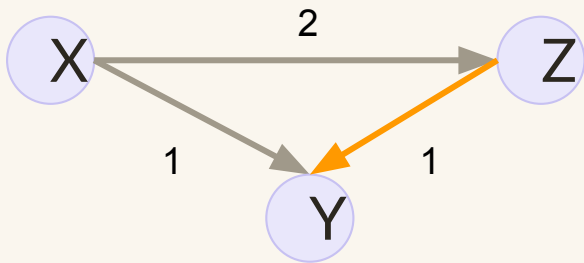
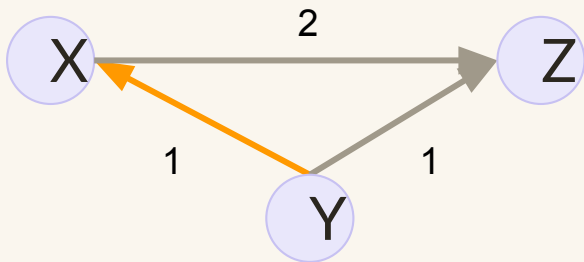
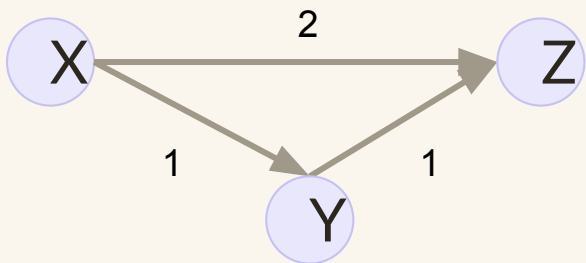
n	β_0	β_1	β_2	β_3
1	0	—	—	—
2	4	0	—	—
3	1	1	12	—
4	1	0	1	0

Higher homology

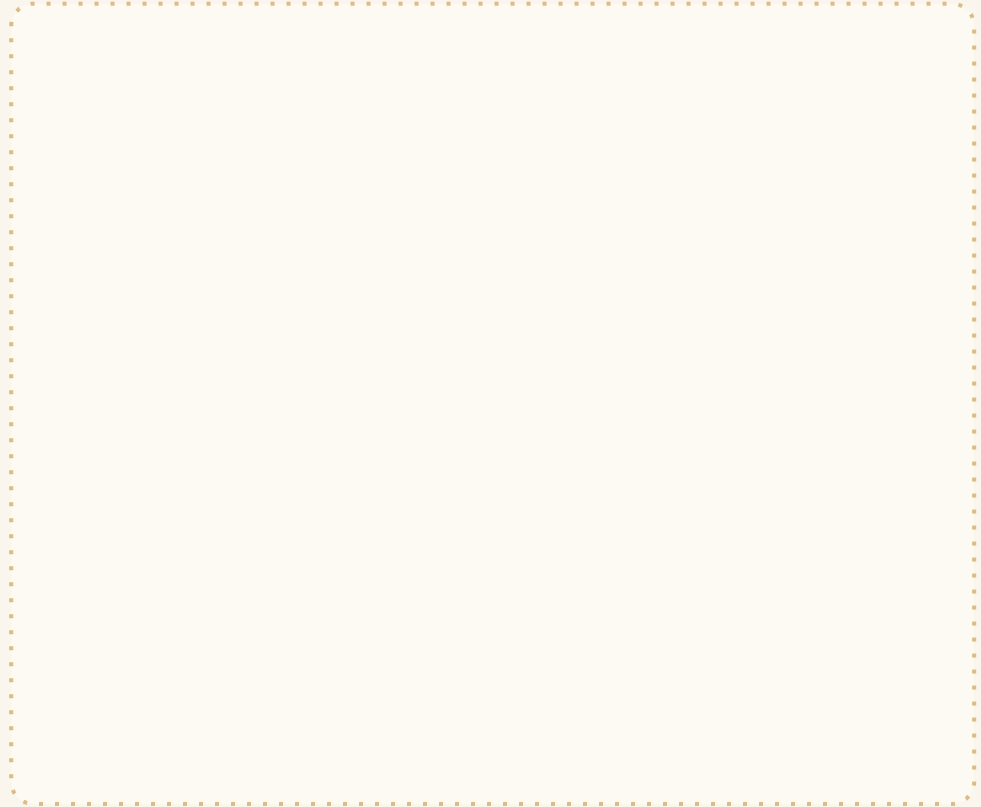
Extra Slides: Pathing in Practice



Comparison between TE and Frobenius metrics



Technical Slides



Summary

Conditional Pathing

- Highlight primary mechanism
- Discoverability
- Future work: Automation

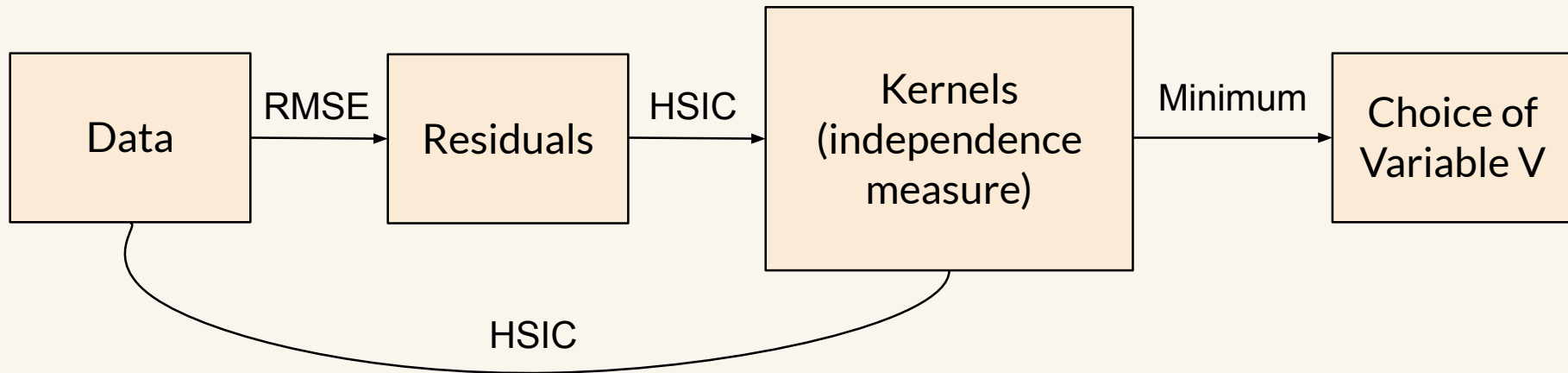
Local Conditional Pathing

- Substructure : discoverability
- Winding number : convincingness
- Future work : try more dataset

Topological Ordering

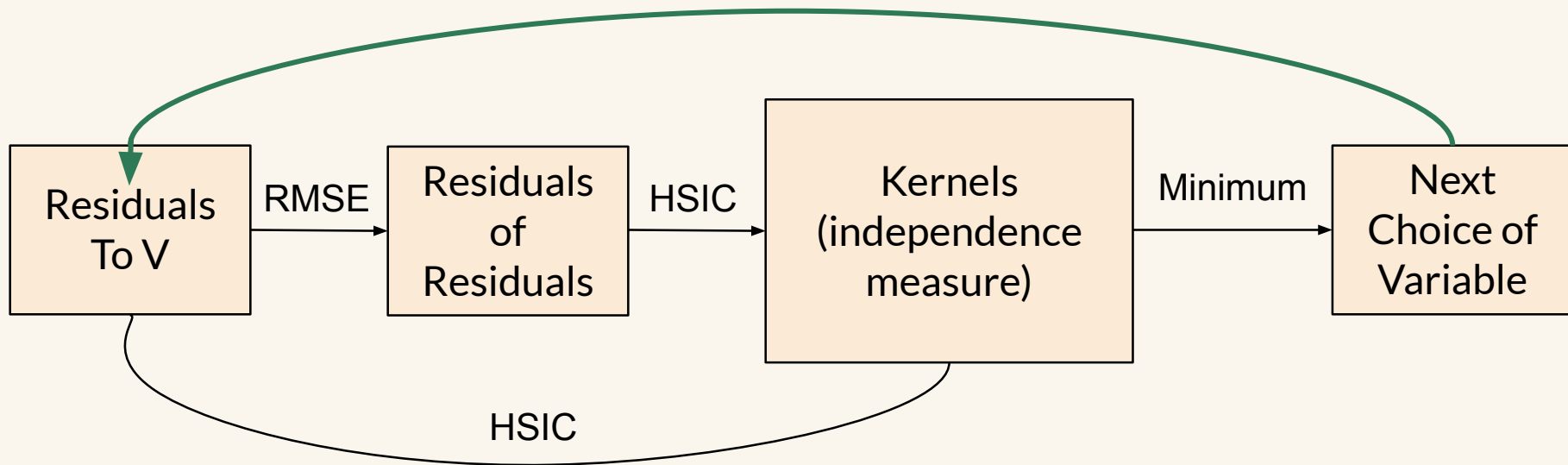
- Change of causal order : discoverability
- Future work: theory development
- Future work: try more dataset
- Future work: prove conjectures

Direct LiNGAM Bound

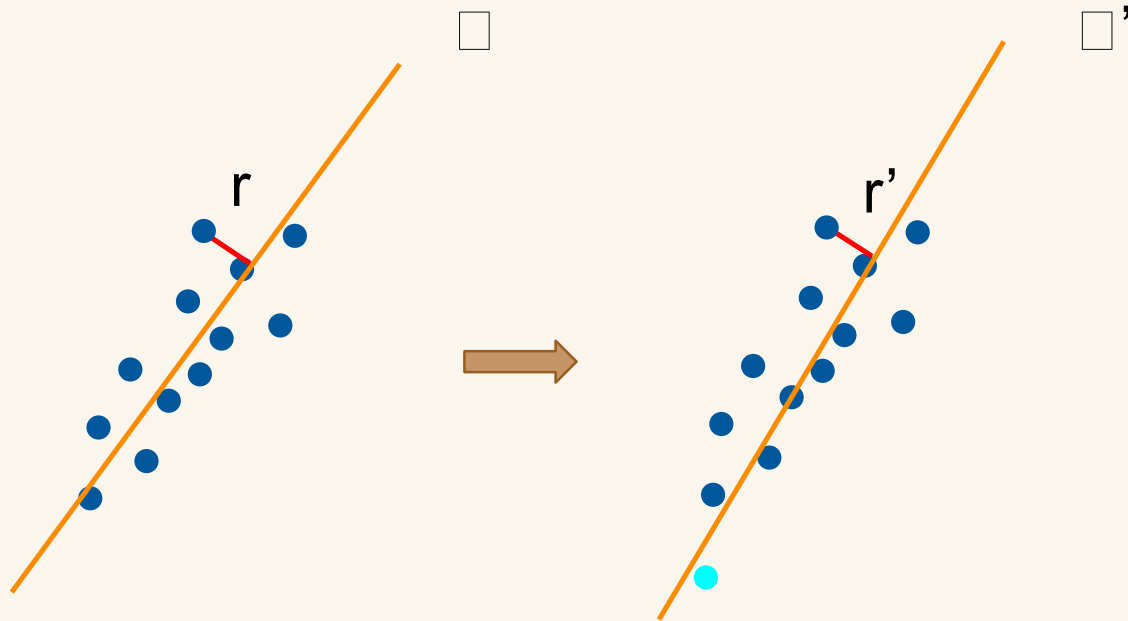


| Direct LiNGAM Causal Order Determination Step 1

Direct LiNGAM Bound

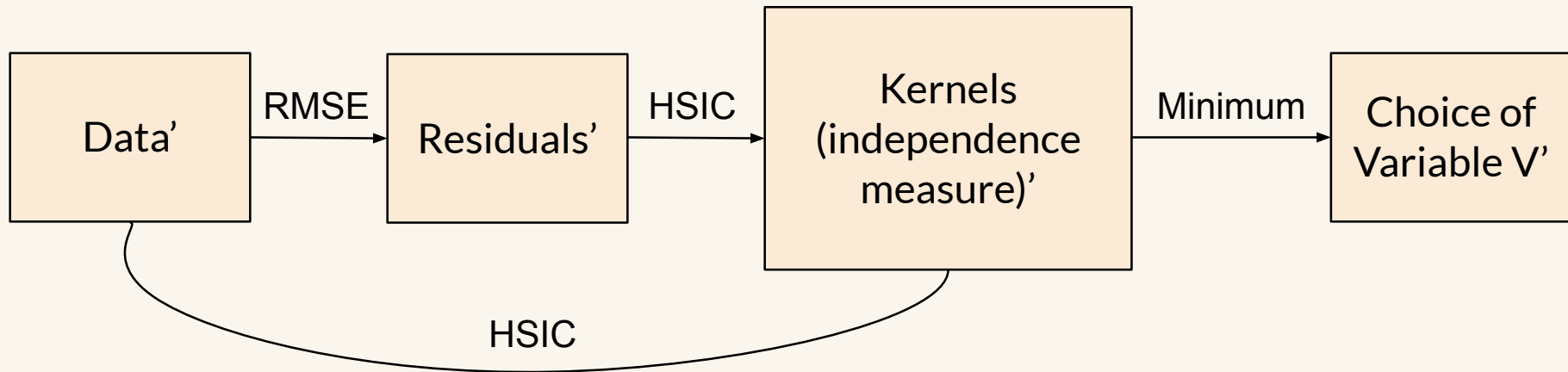


Direct LiNGAM Bound



| Change by adding a new datapoint

Direct LiNGAM Bound

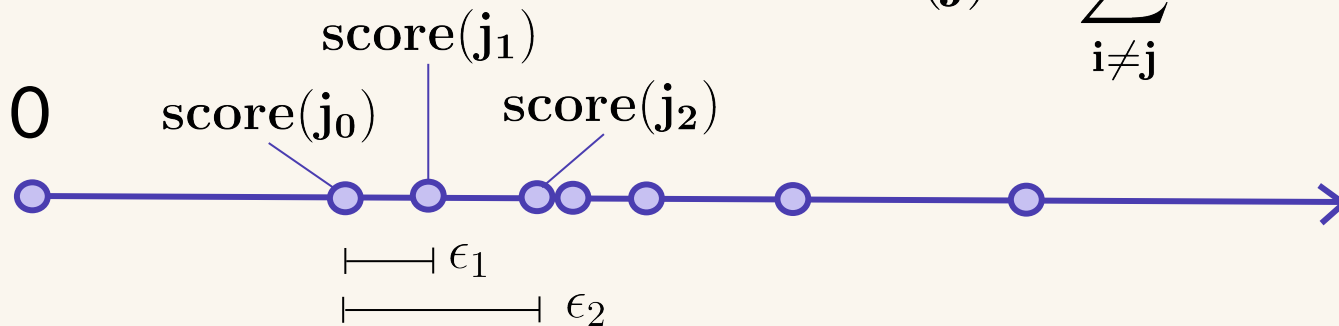


Q: What restrictions need to hold to make $V=V'$?

Direct LiNGAM Bound

$$\epsilon_m = \text{score}(\mathbf{j}_m) - \text{score}(\mathbf{j}_0)$$

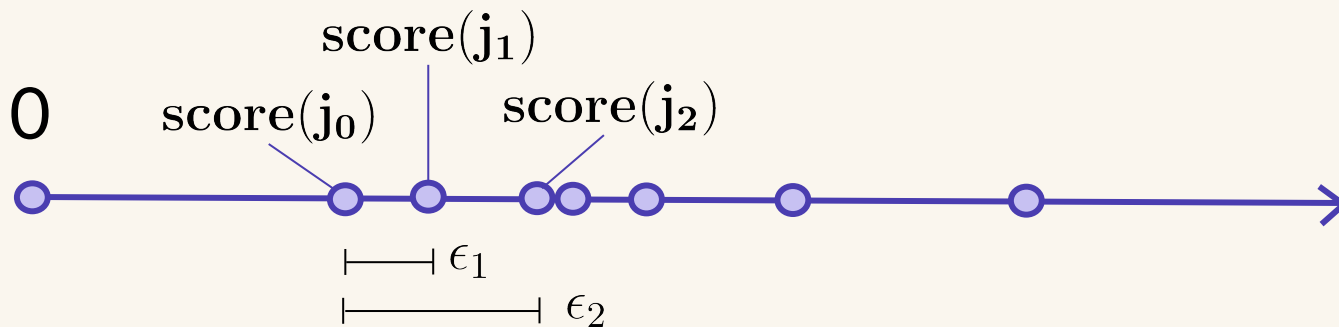
$$\text{score}(\mathbf{j}) = \sum_{i \neq j} \text{Kernel}(\mathbf{i}, \mathbf{j})$$



| Induced change in causal order

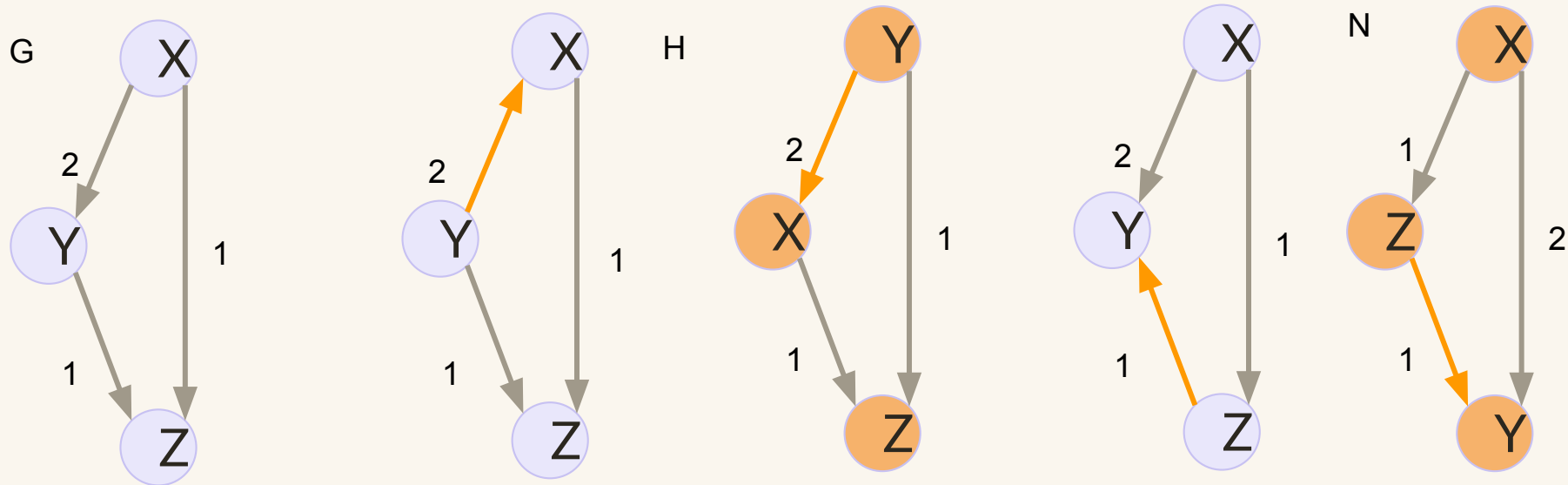
Direct LiNGAM Bound

$$|\Delta \text{score}(j_0)| + |\Delta \text{score}(j_m)| \leq \epsilon_m$$

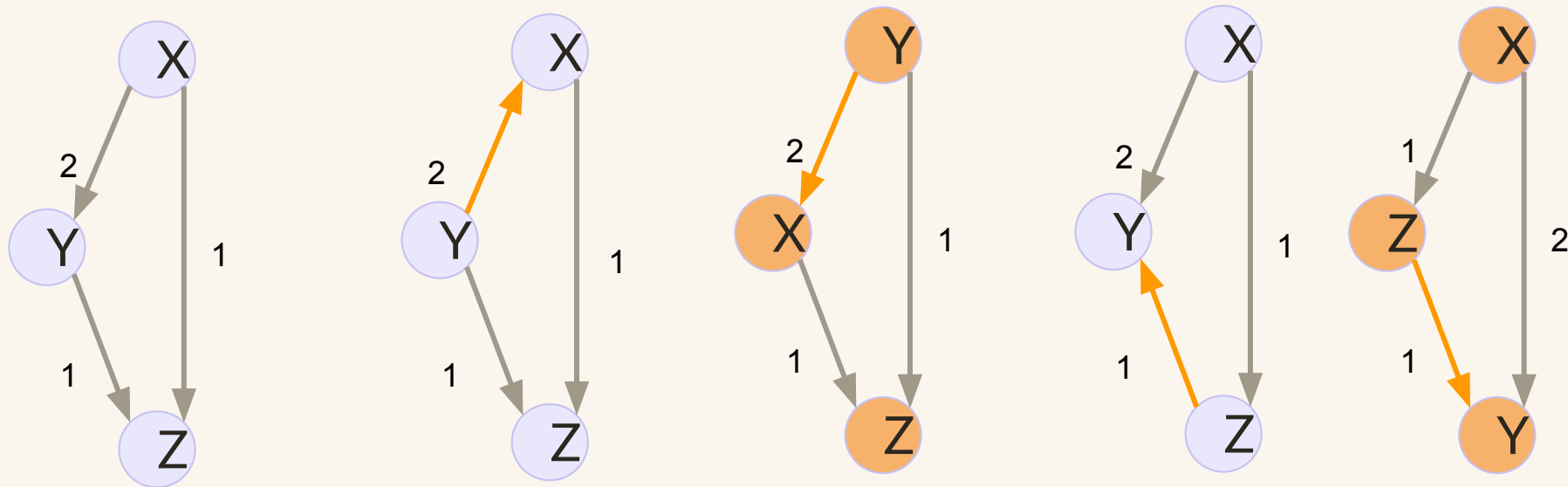


| Bound to preserve order

Comparison between TE and Frobenius metrics



Comparison between TE and Frobenius metrics



Indeed, reversing arrows corresponding to smaller lexicographic changes in the causal order gives smaller changes in the total effect metric.