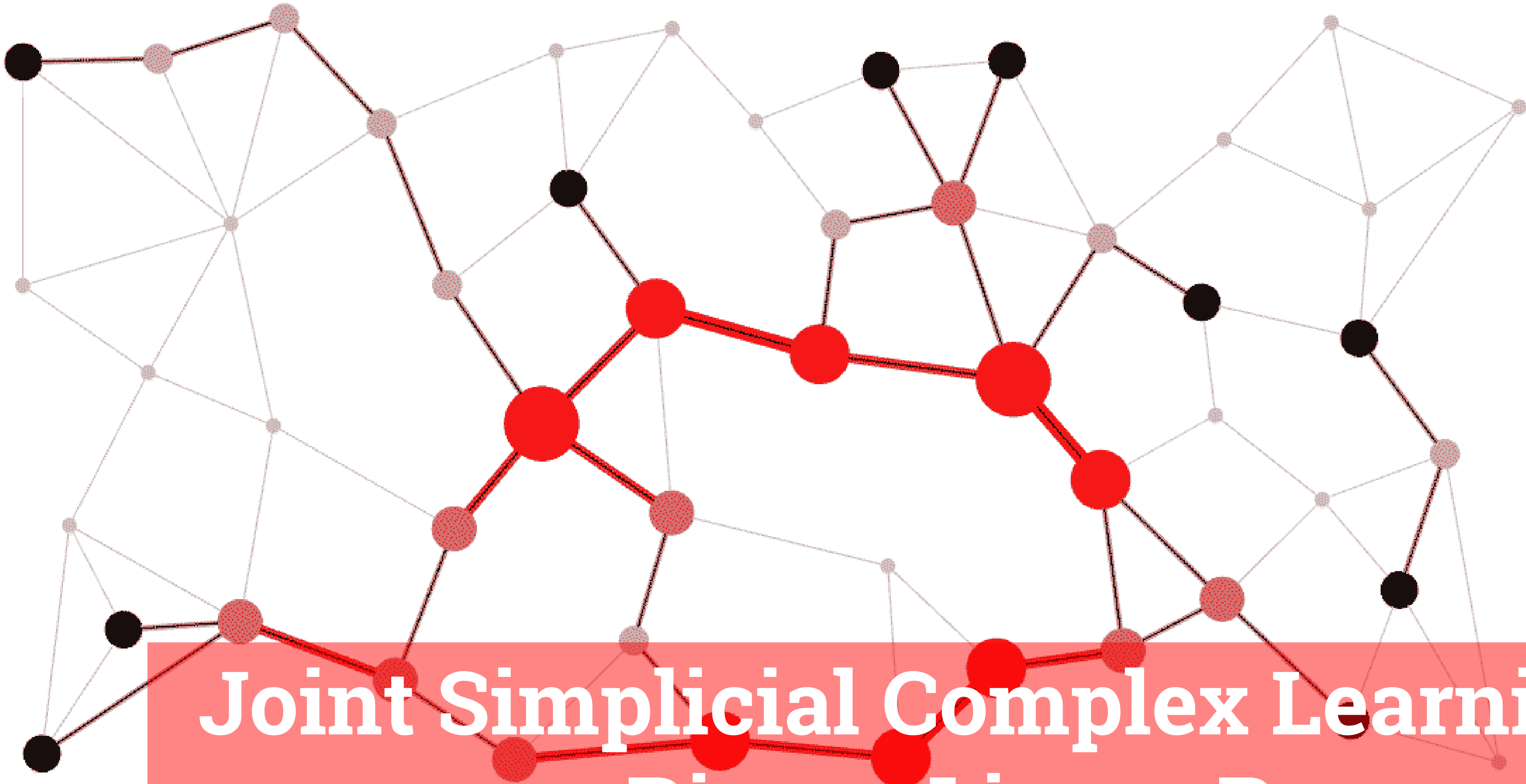


Joint Simplicial Complex Learning via Binary Linear Programming

Varun Sarathchandran
Graphs and Data, March 2026

TU Delft

Contents

TUDelft

Simplicial Complexes: An Introduction

Graphs are a set of nodes and edges:

$\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{v_1, v_2, \ldots, v_N\}$and $\mathcal{E} = \{e_1, e_2, \ldots, e_K\}$

Each edge e_k connects a pair of nodes.

Hypergraphs are a set of nodes and hyperedges:

$\mathcal{H} = (\mathcal{V}, \mathcal{F})$, where $\mathcal{V} = \{v_1, v_2, \ldots, v_N\}$ and $\mathcal{F} = \{f_1, f_2, \ldots, f_K\}$

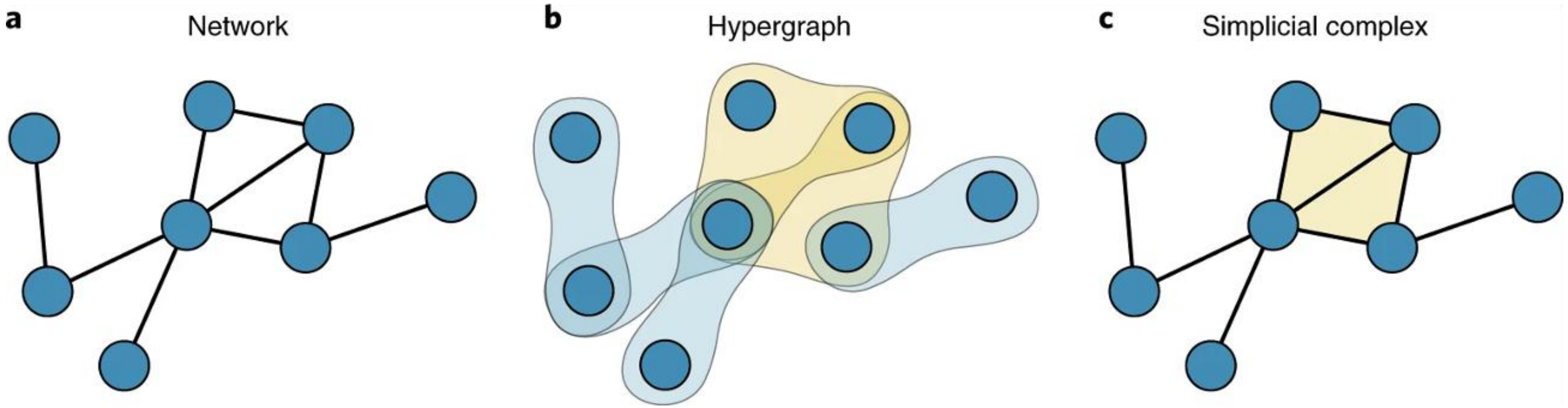
Each hyperedge f_k connects an arbitrary number of nodes.

TU Delft

Simplicial Complexes: An Introduction

Simplicial Complexes are hypergraphs with structure.

For every hyperedge connecting a set of nodes, all subsets of that set must also exist in the complex.



Each hyperedge: k-simplex. 0-simplices: nodes, 1-simplices: edges, ...

Simplicial Complexes: An Introduction

A k-simplicial signal: $\mathbf{X}_k \in \mathbb{R}^{N_k \times F_k}$,

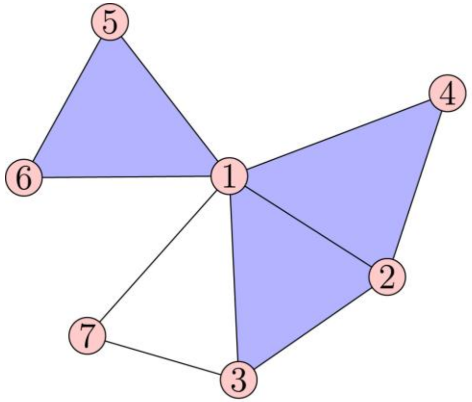
where N_k is the number of k-simplices,

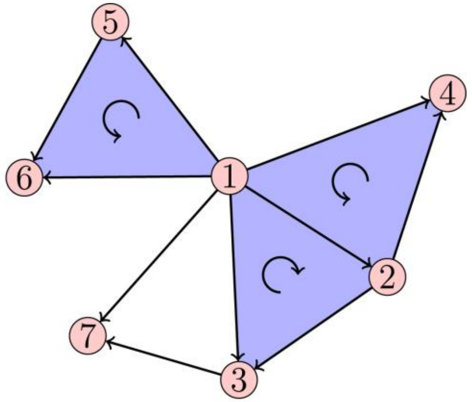
F_k is the number of features per simplex.

Simplicial Complexes: Mathematical Represenation

Incidence matrices: $\mathbf{B}_k \in \mathbb{R}^{N_{k-1} \times N_k}$

Map from $k-1$-simplices to k-simplices





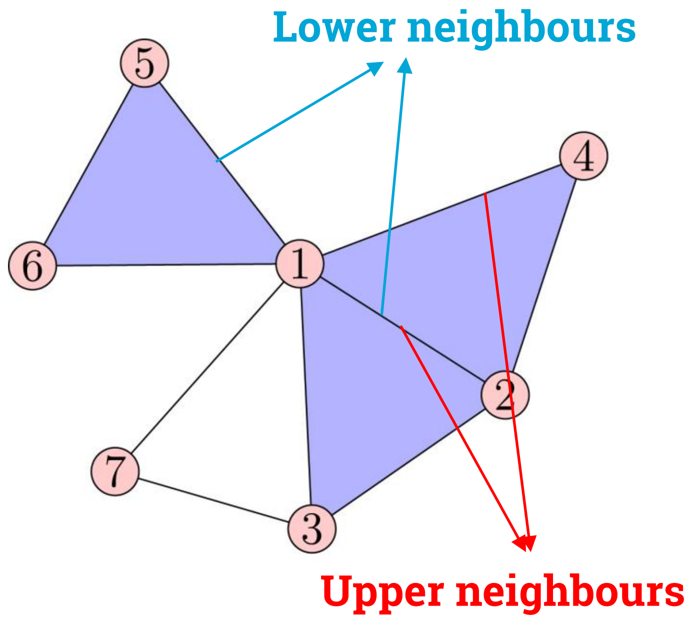
$$\mathbf{B}_1 = \overbrace{\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & -1 & 0 \end{bmatrix}}^{edges} \Bigg\} \, nodes \qquad \mathbf{B}_2 = \overbrace{\begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}}^{triangles} \Bigg\} \, edges$$

TU Delft

Simplicial Complexes: Mathematical Represenation

A simplicial complex of order K can be defined via the Hodge Laplacians:

$$\mathbf{L}_k = \underbrace{\mathbf{B}_k^\top \mathbf{B}_k}_{\mathbf{L}_k^{\text{low}}} + \underbrace{\mathbf{B}_{k+1} \mathbf{B}_{k+1}^\top}_{\mathbf{L}_k^{\text{up}}}, \quad k = 1, \dots, K-1.$$



Simplicial complex structural constraint:

$$\mathbf{B}_k \mathbf{B}_{k+1} = \mathbf{0} \ \forall \ k = 1, \dots, K-1.$$

TU Delft

Goal: identify the topology of the simplicial complex from data.

Problem Formulation

Links b/w topology and signals

Identify topology

$$\min_{\mathbf{B}_1,\mathbf{B}_2} f_1(\mathbf{B}_1, \mathbf{X}_0) + f_2(\mathbf{B}_2, \bar{\mathbf{X}}_1)$$
$$\text{s.t.} \quad \mathbf{B}_1 \in \mathcal{B}_1, \quad \mathbf{B}_2 \in \mathcal{B}_2, \quad \mathbf{B}_1\mathbf{B}_2 = \mathbf{0}.$$

Subject to feasible incidence matrices

Subject to the inclusion constraint

TU Delft

Problem Formulation

$$
\begin{aligned}
\min_{\mathbf{B}_1, \mathbf{B}_2} \quad & f_1(\mathbf{B}_1, \mathbf{X}_0) + f_2(\mathbf{B}_2, \bar{\mathbf{X}}_1) \\
\text{s.t.} \quad & \boxed{\mathbf{B}_1 \in \mathcal{B}_1, \quad \mathbf{B}_2 \in \mathcal{B}_2,}, \quad \mathbf{B}_1\mathbf{B}_2 = \mathbf{0}.
\end{aligned}
$$

Maintaining feasibility

Maintaining Feasibility

Avoid directly learning incidence matrices.

Select simplices from full complexes [1].

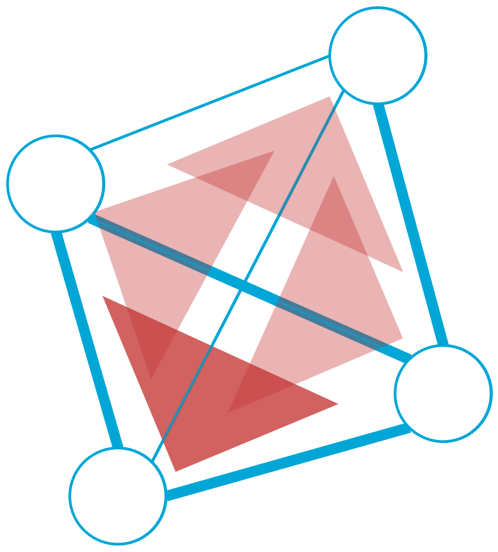
Full incidence matrices: $\bar{\mathbf{B}}_1 \in \mathbb{R}^{N_0 \times \bar{N}_1}$ and $\bar{\mathbf{B}}_2 \in \mathbb{R}^{\bar{N}_1 \times \bar{N}_2}$

Select existing simplices using $\mathbf{s}_1 \in \{0,1\}^{\overline{N_1}}$, $\mathbf{s}_2 \in \{0,1\}^{\overline{N_2}}$

Work only in the space of $\mathbf{s}_1$, $\mathbf{s}_2$

Enforce minimum number of simplices by imposing minimum cardinality

Edge selection with $\mathbf{s}_1$
Triangle selection with $\mathbf{s}_2$



TU Delft

[1] Buciulea, A., Isufi, E., Leus, G., & Marques, A. G. (2024, July). Learning the topology of a simplicial complex using simplicial signals: A greedy approach.

Problem Formulation

Links b/w topology and signals

Assume signals available on all possible edges

$$\min_{\mathbf{B}_1,\mathbf{B}_2} \quad f_1(\mathbf{B}_1, \mathbf{X}_0) + f_2(\mathbf{B}_2, \bar{\mathbf{X}}_1)$$

$$\text{s.t.} \quad \mathbf{B}_1 \in \mathcal{B}_1, \quad \mathbf{B}_2 \in \mathcal{B}_2, , \quad \mathbf{B}_1\mathbf{B}_2 = \mathbf{0}.$$

Smoothness: Linking topology to signals

Linking node signals to edges:

Node signal smoothness on edges: $\mathrm{tr}(\mathbf{X}_0^\top \bar{\mathbf{L}}_0 \mathbf{X}_0),\ \bar{\mathbf{L}}_0 = \bar{\mathbf{B}}_1 \mathrm{diag}(\mathbf{s}_1) \bar{\mathbf{B}}_1^\top$

Linking edge signals to triangles:

Edge signal curl on triangles: $\mathrm{tr}(\bar{\mathbf{X}}_1^\top \bar{\mathbf{L}}_1^{\mathrm{up}} \bar{\mathbf{X}}_1),\ \bar{\mathbf{L}}_1^{\mathrm{up}} = \bar{\mathbf{B}}_2 \mathrm{diag}(\mathbf{s}_2) \bar{\mathbf{B}}_2^\top$

Proposed new similarity measure: $\sum_{i=1}^{\bar{N}_2} [\mathbf{s}_2]_i \sum_{\substack{f,g \in \mathcal{F}(\sigma_2^{(i)}) \\ f<g}} \|[\bar{\mathbf{X}}_1]_{f,:} - [\bar{\mathbf{X}}_1]_{g,:}\|_2^2$

All measures linear in the selection variables!

Problem Formulation

Identify topology

$$\min_{\mathbf{B}_1,\mathbf{B}_2} \quad f_1(\mathbf{B}_1,\mathbf{X}_0)+f_2(\mathbf{B}_2,\bar{\mathbf{X}}_1)$$
$$\text{s.t.} \quad \mathbf{B}_1\in\mathcal{B}_1, \quad \mathbf{B}_2\in\mathcal{B}_2,, \quad \boxed{\mathbf{B}_1\mathbf{B}_2=\mathbf{0}.}$$

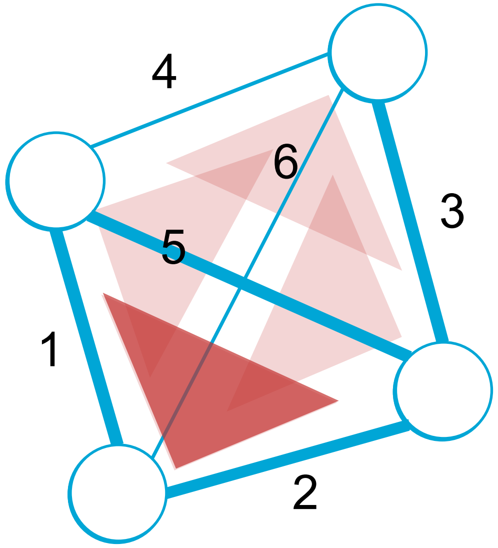
Enforcing inclusion

Enforcing Inclusion

Two main approaches currently to enforce inclusion

Hierarchical [1]: solve for edges, and then triangles.
No need to mathematically couple simplices.

Bilinear constraint [2]: $(\mathbf{1} - \mathbf{s}_1)^\top \bar{\mathbf{B}}_2^+ \mathbf{s}_2 = 0$
Requires alternating approaches.



0	0	0	1	0	1

1
1
0
0
1
0

TU Delft

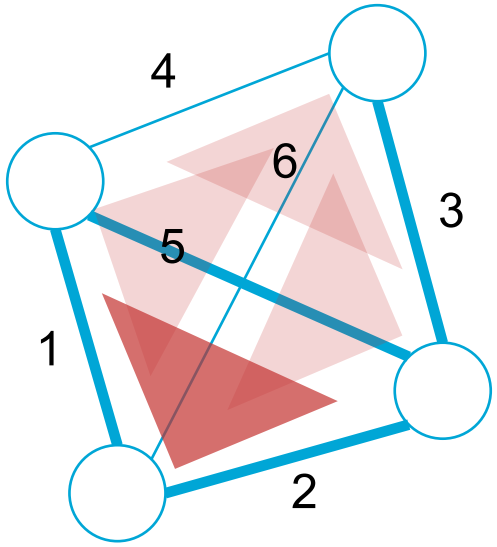
[1] Barbarossa, S., & Sardellitti, S. (2020). Topological signal processing over simplicial complexes.
[2] Buciulea, A., Isufi, E., Leus, G., & Marques, A. G. (2024, July). Learning the topology of a simplicial complex using simplicial signals: A greedy approach.

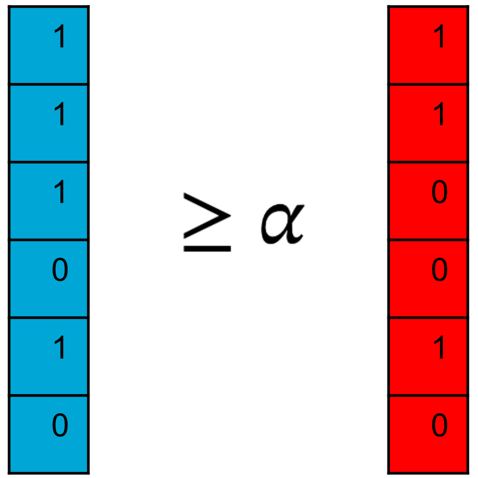
Enforcing Inclusion

We introduce a linear constraint to enforce inclusion

$$\mathbf{s}_1 \geq \alpha \bar{\mathbf{B}}_2^{+} \mathbf{s}_2$$

Allows for a joint approach without requiring alternating!





Problem Formulation

$$\min_{\mathbf{B}_1,\mathbf{B}_2} \quad f_1(\mathbf{B}_1, \mathbf{X}_0) + f_2(\mathbf{B}_2, \bar{\mathbf{X}}_1)$$
$$\text{s.t.} \quad \mathbf{B}_1 \in \mathcal{B}_1, \quad \mathbf{B}_2 \in \mathcal{B}_2,, \quad \mathbf{B}_1\mathbf{B}_2 = \mathbf{0}.$$

$$\begin{aligned} \min_{\mathbf{s}_1,\mathbf{s}_2} \quad & \mathbf{h}_1^\top \mathbf{s}_1 + \mathbf{h}_2^\top \mathbf{s}_2 \\ \text{s.t.} \quad & \mathbf{s}_1 \geq \alpha \bar{\mathbf{B}}_2^+ \mathbf{s}_2 \\ & \mathbf{s}_1 \in \{0,1\}^{\bar{N}_1}, \quad \mathbf{s}_2 \in \{0,1\}^{\bar{N}_2} \\ & \mathbf{1}^\top \mathbf{s}_1 \geq C_1, \quad \mathbf{1}^\top \mathbf{s}_2 \geq C_2. \end{aligned}$$

TU Delft

Baselines

Hierarchical Approach [1]:

First solve

$$\begin{aligned} \min_{\mathbf{s}_1} \quad & \mathbf{h}_1^\top \mathbf{s}_1 \\ \text{s.t.} \quad & \mathbf{1}^\top \mathbf{s}_1 \geq C_1, \quad \mathbf{s}_1 \in \{0,1\}^{\bar{N}_1} \end{aligned}$$

Then, restrict to feasible triangle set $\mathcal{T}(\hat{\mathbf{s}}_1)$ and solve

$$\begin{aligned} \min_{\mathbf{s}_2} \quad & \mathbf{h}_2^\top \mathbf{s}_2 \\ \text{s.t.} \quad & [\mathbf{s}_2]_t = 0, \quad \forall t \notin \mathcal{T}(\hat{\mathbf{s}}_1) \\ & \mathbf{1}^\top \mathbf{s}_2 \geq C_2, \quad \mathbf{s}_2 \in \{0,1\}^{\bar{N}_2}. \end{aligned}$$

TU Delft

[1] Barbarossa, S., & Sardellitti, S. (2020). Topological signal processing over simplicial complexes.

Baselines

Greedy, alternating approach [1]:

$$\begin{aligned} \min_{\mathbf{s}_1,\mathbf{s}_2} \quad & \|\mathbf{s}_1\|_0 + \|\mathbf{s}_2\|_0 + \mathbf{h}_1^\top \mathbf{s}_1 + \mathbf{h}_2^\top \mathbf{s}_2 \\ & + \gamma(\mathbf{1} - \mathbf{s}_1)^\top \bar{\mathbf{B}}_2^+ \mathbf{s}_2 \\ \text{s.t.} \quad & \mathbf{s}_1 \in \{0,1\}^{\bar{N}_1}, \quad \mathbf{s}_2 \in \{0,1\}^{\bar{N}_2} \\ & \|\mathbf{s}_1\|_0 \geq C_1, \quad \|\mathbf{s}_2\|_0 \geq C_2. \end{aligned}$$

TU Delft

[1] Buciulea, A., Isufi, E., Leus, G., & Marques, A. G. (2024, July). Learning the topology of a simplicial complex using simplicial signals: A greedy approach.

Simulated Experiments

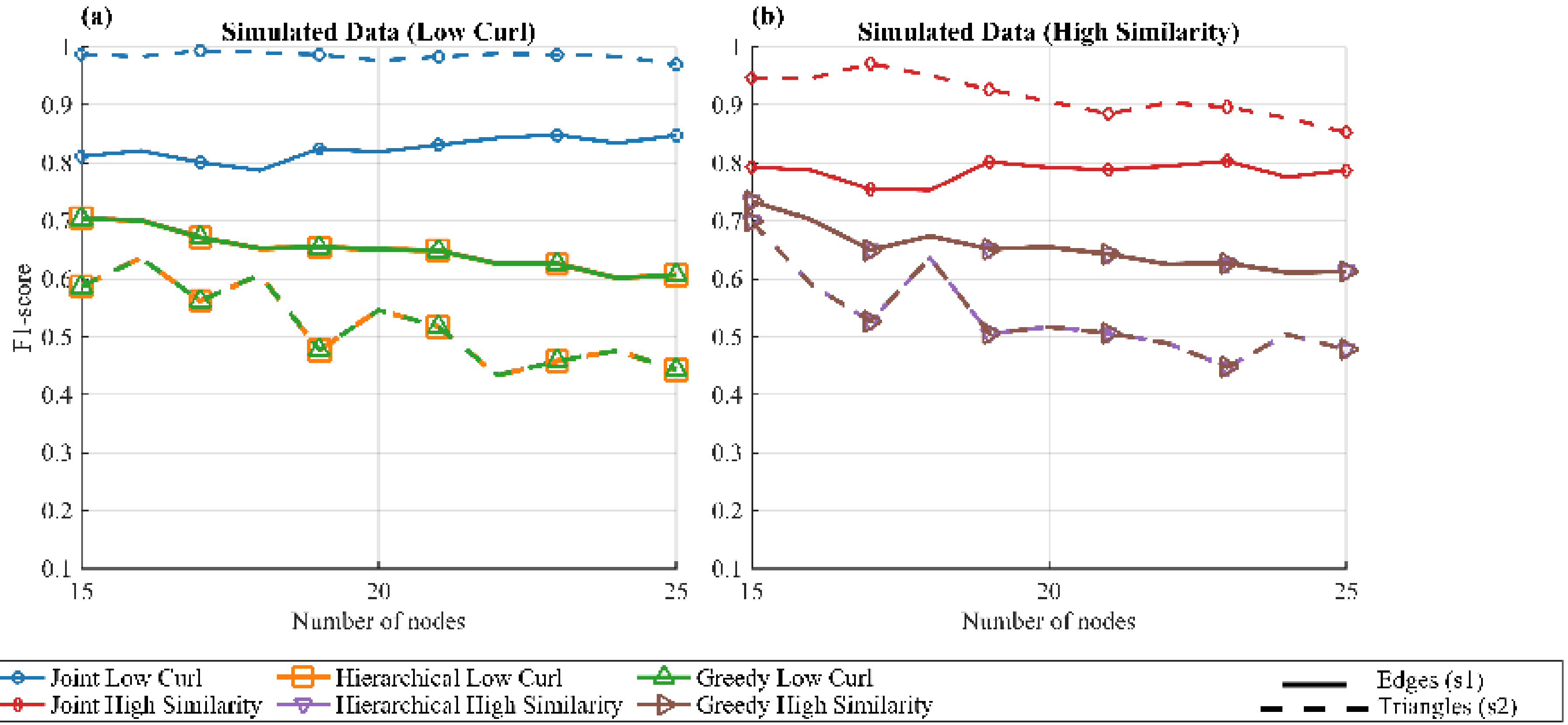
Sample an Erdős-Rényi graph with a certain edge probability.

Smooth node signals are then generated by filtering white noise through the graph Laplacian.

From the set of feasible triangles, half are randomly chosen.

Edge signals are generated similarly, using the appropriate Laplacian (associated to either curl or similarity).

Simulated Experiments



Co-Authorship Experiments

Nodes: authors, node signals: frequency of keywords used.

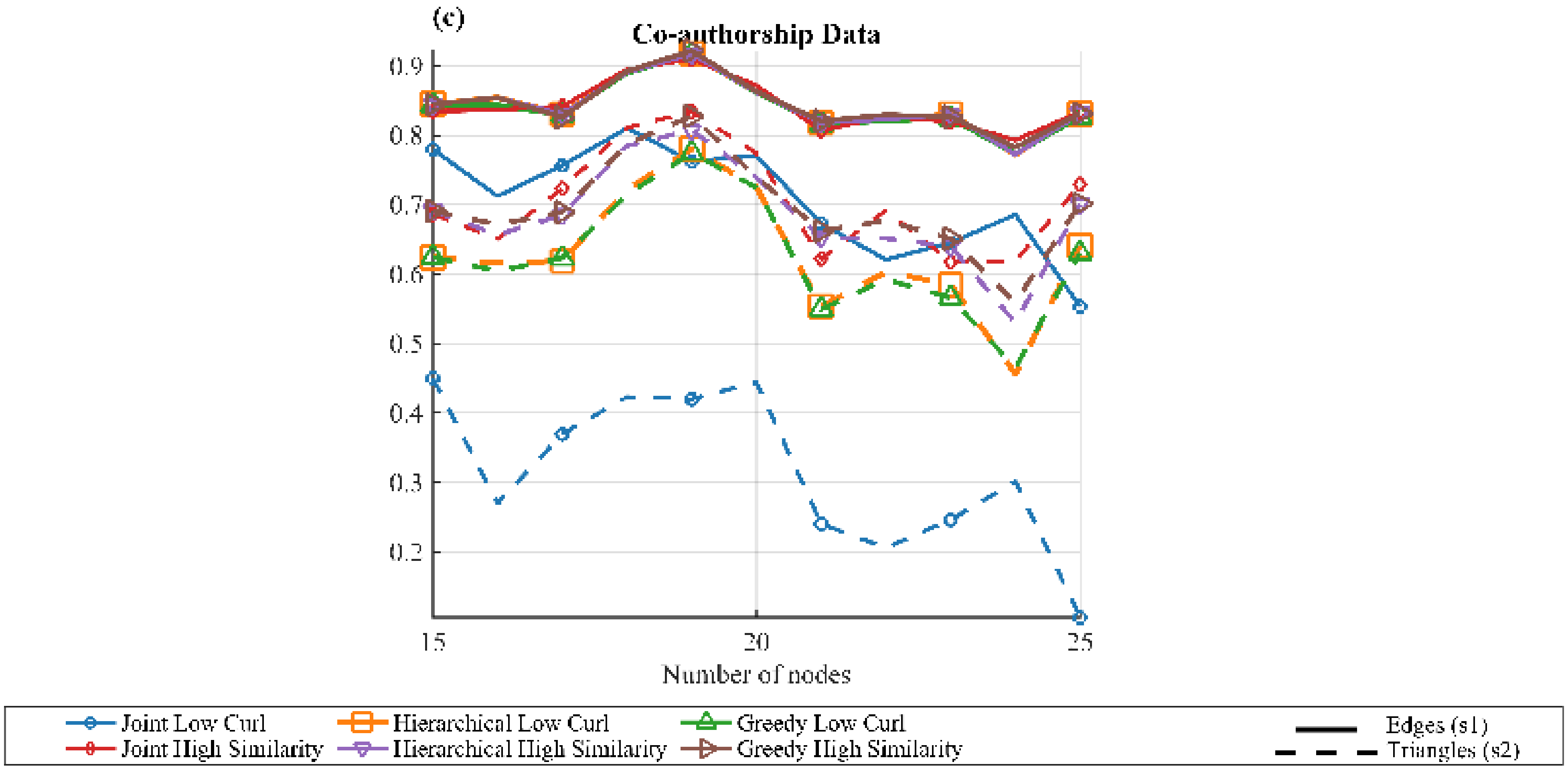
Edges and triangles: co-authors of the same paper.

Edge signals: element wise minimum of node signals.

Important property:

Edge signals are not low curl but have high similarity.

Co-Authorship Experiments



Conclusion

Solving jointly captures better structural relationships: allowed by the linear constraint.

The joint method truly reflects how well the higher simplices adhere to the prior- no reliance on feasibility.

Future work: extending to weighted simplicial complexes, tracking time-varying topology.

TU Delft

Questions