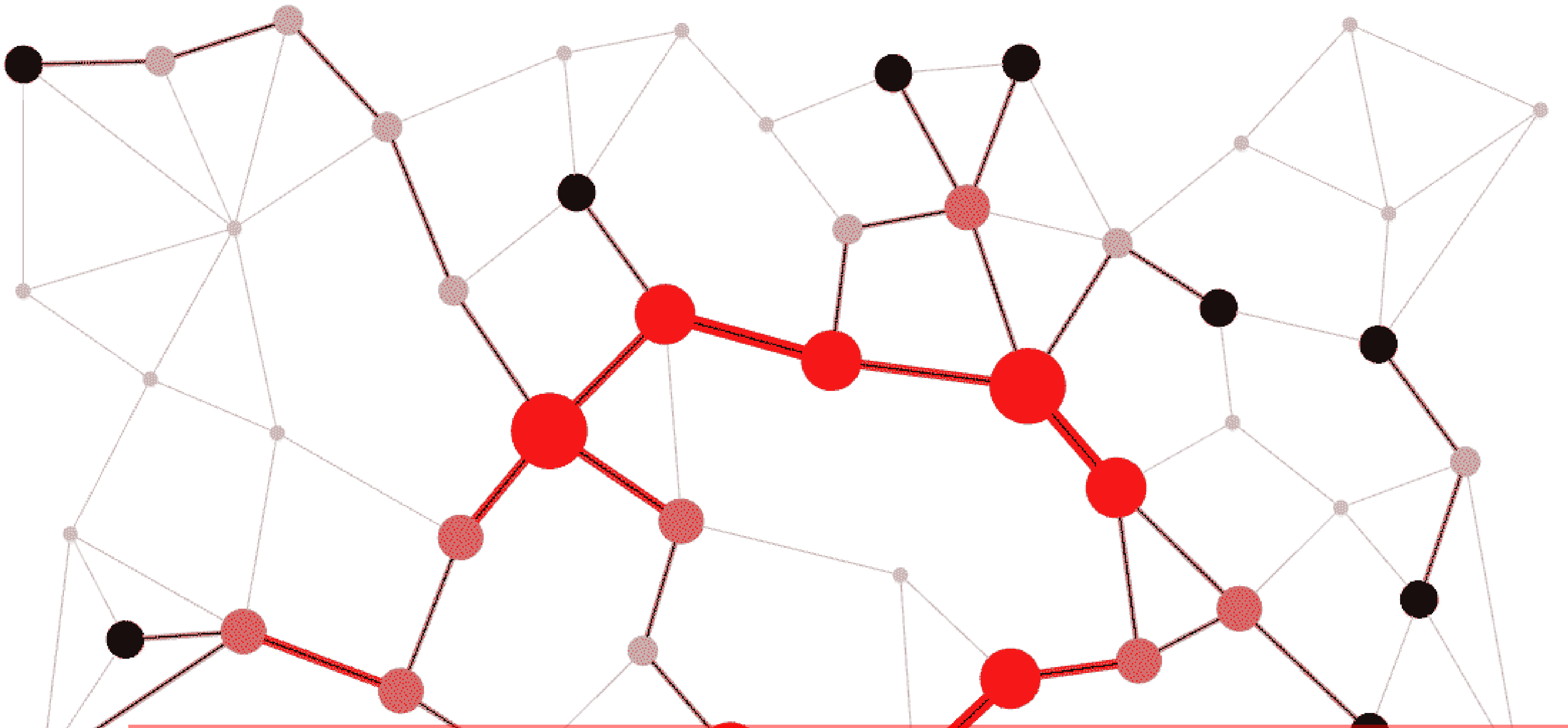


Joint Simplicial Complex Learning via Binary Linear Programming

Varun Sarathchandran, Geert Leus

GSP Workshop 2026

TU Delft

Contents

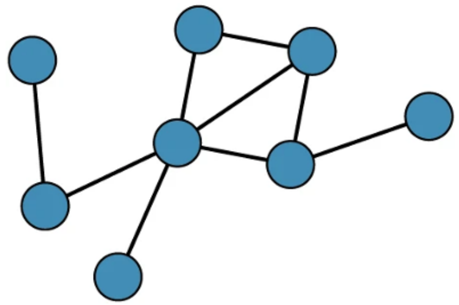
TU Delft

Hypergraphs: Moving beyond Graphs

- 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 $\mathcal{E} = \{e_1, e_2, \ldots, e_K\}$

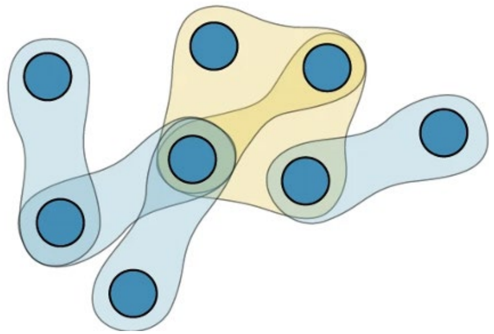
  Each edge $\mathbf{e}_\mathrm{k}$ connects a pair of nodes.



- Hypergraphs are a set of nodes and hyperedges:

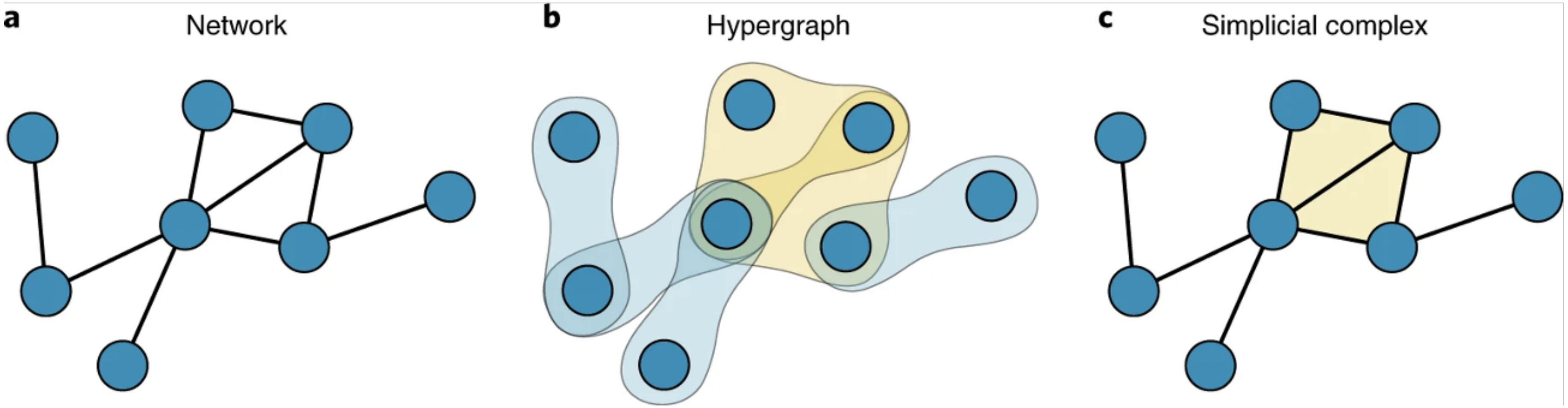
  $\mathcal{H} = (\mathcal{V}, \mathcal{F})$, where $\mathcal{V} = \{\mathrm{v}_1, \mathrm{v}_2, \ldots, \mathrm{v}_\mathrm{N}\}$ and $\mathcal{F} = \{f_1, f_2, \ldots, f_K\}$

  Each hyperedge f_k connects an arbitrary number of nodes.



Simplicial Complexes

- 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.



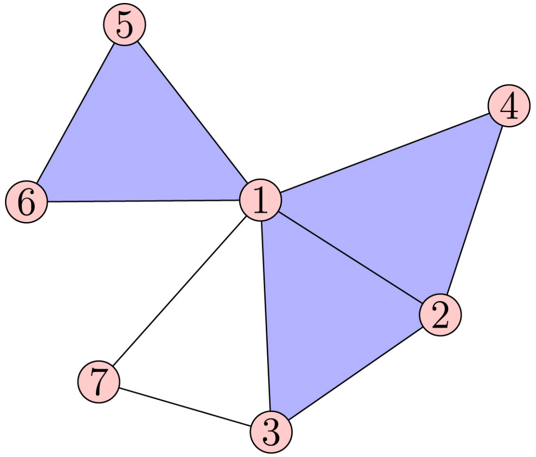
- Each hyperedge: k-simplex. 0-simplices: nodes, 1-simplices: edges, …

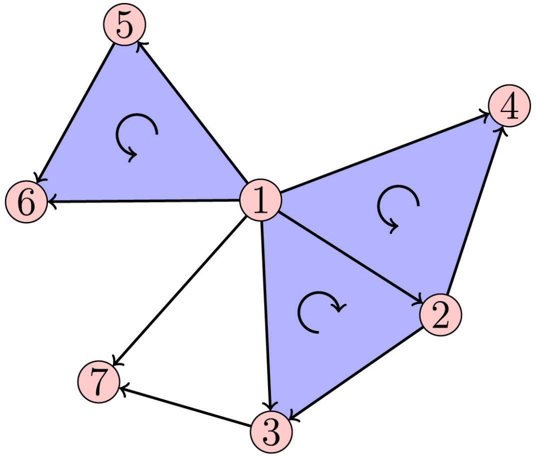
TU Delft

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

Simplicial Complexes: Defining Topology

- Given a topology, how do we mathematically represent it?
- Incidence matrices: $\mathbf{B}_k \in \mathbb{R}^{N_{k-1} \times N_k}$
- Map from $\mathbf{k-1}$-simplices to $\mathbf{k}$-simplices



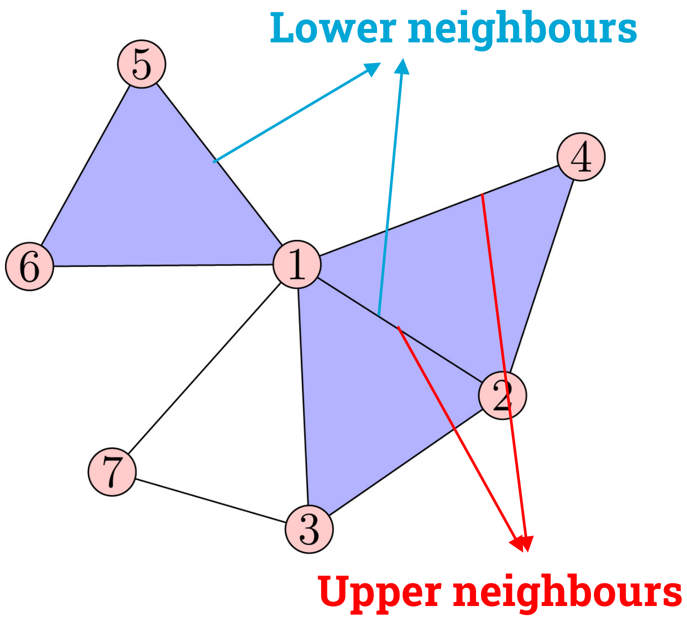


$$\mathbf{B}_1 = \overbrace{\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & -1 & 0 \end{bmatrix}}^{edges} \left.\vphantom{\begin{bmatrix}0\\0\\0\\0\\0\\0\\0\end{bmatrix}}\right\} nodes \quad \mathbf{B}_2 = \overbrace{\begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}}^{triangles} \left.\vphantom{\begin{bmatrix}0\\0\\0\\0\\0\\0\\0\end{bmatrix}}\right\} edges$$

Simplicial Complexes: Defining Topology

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

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



- Simplicial complex structural constraint:

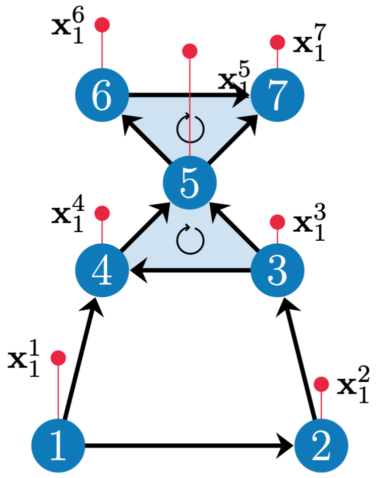
$$\mathbf{B}_k \mathbf{B}_{k+1} = \mathbf{0} \;\; \forall \; k = 1, \ldots, K-1.$$

Simplicial Complexes: Data

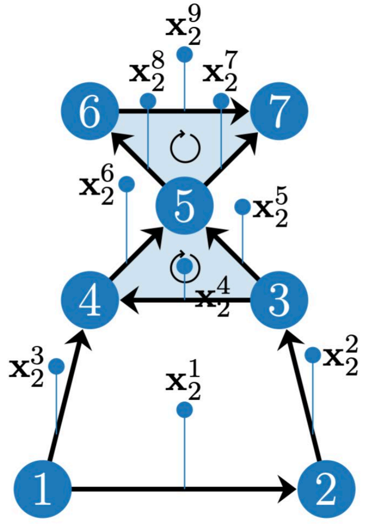
- A k-simplicial signal: $\mathbf{X}_k \in \mathbb{R}^{N_k \times F_k}$,

  where N_k is the number of k-simplices,

  F_k is the number of features per simplex.



Node signals



Edge signals

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

Problem Formulation

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

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

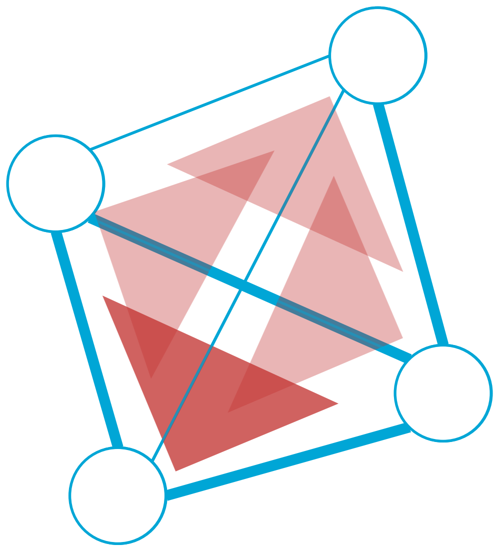
Maintaining feasibility

TU Delft

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 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

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

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

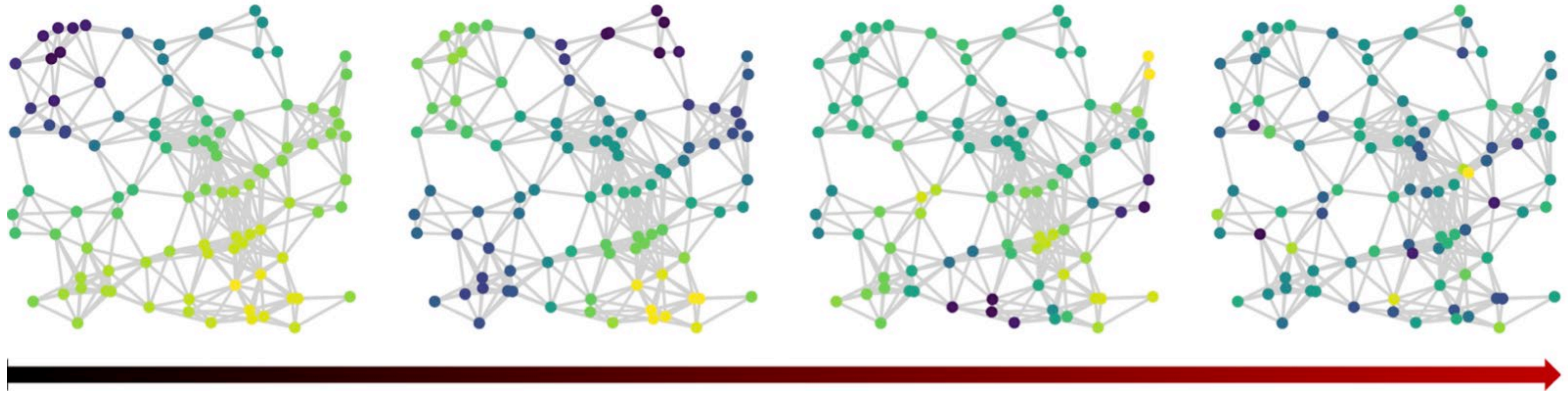
TU Delft

Smoothness: Linking topology to signals

- Linking node signals to edges:

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

  Captures total variation of node signals across edges



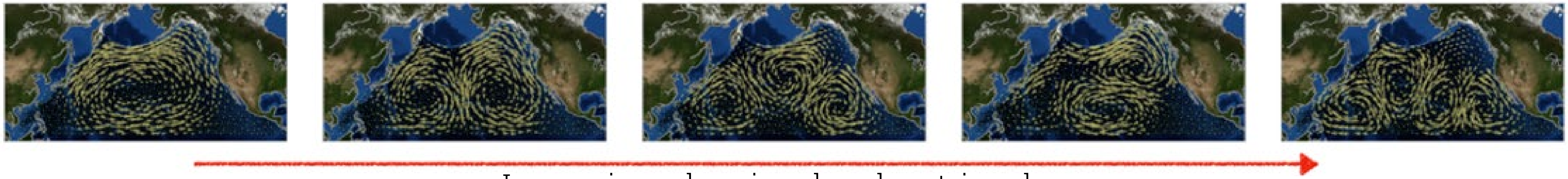
Increasing node signal variation across edges

Smoothness: Linking topology to signals

- Linking edge signals to triangles:

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

  Measures the rotational component of edge signals around triangles



Increasing edge signal curl on triangles

Smoothness: Linking topology to signals

- Linking edge signals to triangles:

  No existing measure which captures edge signal variation within a triangle
  Proposed measure: sum of pair differences of edge signals within a triangle

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

  Generalizes graph signal smoothness to higher orders.

All smoothness measures are linear in the selection variables.

TU Delft

Problem Formulation

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

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

Enforcing inclusion

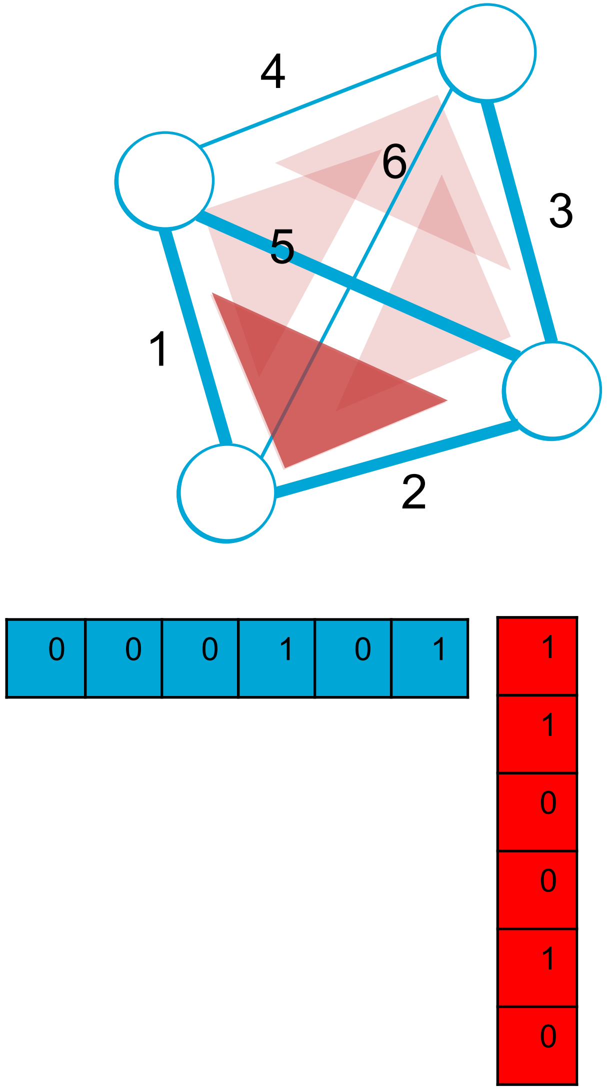
TU Delft

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.



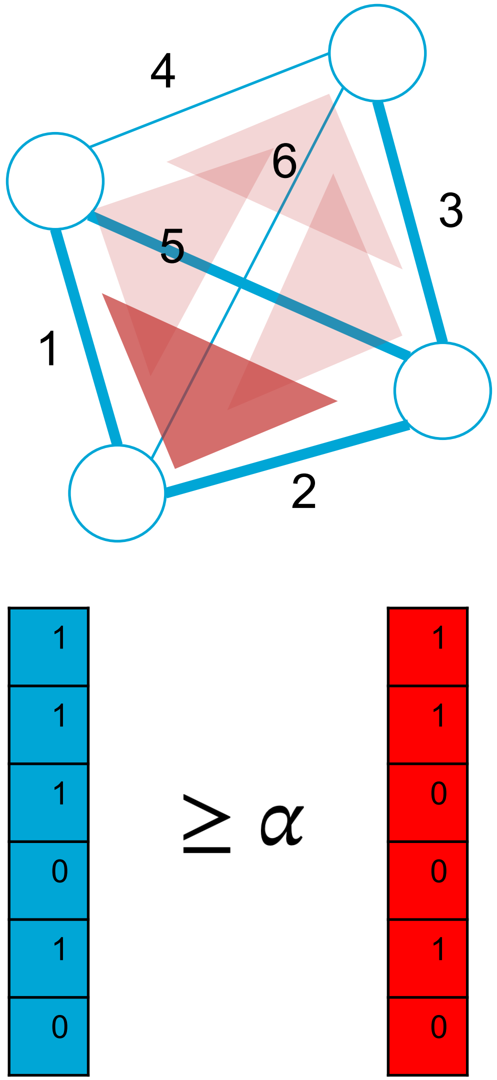
[1] Barbarossa, S., & Sardellitti, S. (2020). Topological signal processing over simplicial complexes.
[2] Buciulea, A., Isufi, E., Leus, G., & Marques, A. G. (2024, July). Learning the topology of a simplicial complex using simplicial signals: A greedy approach.

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!



TU Delft

Problem Formulation

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

$\Rightarrow$

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

Baselines

- Hierarchical Approach [1]:

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

- 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]:

$$\min_{\mathbf{s}_1,\mathbf{s}_2} \quad \|\mathbf{s}_1\|_0 + \|\mathbf{s}_2\|_0 + \mathbf{h}_1^\top \mathbf{s}_1 + \mathbf{h}_2^\top \mathbf{s}_2 + \gamma(\mathbf{1} - \mathbf{s}_1)^\top \bar{\mathbf{B}}_2^+ \mathbf{s}_2$$

$$\text{s.t.} \quad \mathbf{s}_1 \in \{0,1\}^{\bar{N}_1}, \quad \mathbf{s}_2 \in \{0,1\}^{\bar{N}_2}, \quad \|\mathbf{s}_1\|_0 \geq C_1, \quad \|\mathbf{s}_2\|_0 \geq C_2.$$

- Form a cost per edge, choose the edges with the lowest cost.
- Form a cost per triangle, choose the triangles with the lowest cost.
- Alternate.
- Inclusion penalty embedded in cost.

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).

TU Delft

Simulated Experiments

Real Data Experiments : Co-Authorship Networks

- 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.

Going Forward

- Learning weighted simplicial complexes.

  Needs special requirement: proposed constraint works only for binary support.

  $\mathbf{s}_1 \geq \alpha \bar{\mathbf{B}}_2^+ \mathbf{s}_2$ imposes the requirement that each edge weight be at least the sum of the weights of its incident triangles

- Tracking time-varying simplicial complexes.

Thanks!
Questions?

Find the paper on Arxiv

V.Sarathchandran@tudelft.nl

TUDelft