

Hodge-Aware Subspace Detector

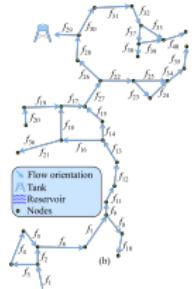
Chengen Liu Elvin Isufi

Delft University of Technology, The Netherlands

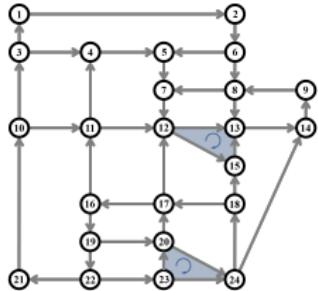
e.isufi-1@tudelft.nl

November 3, 2025

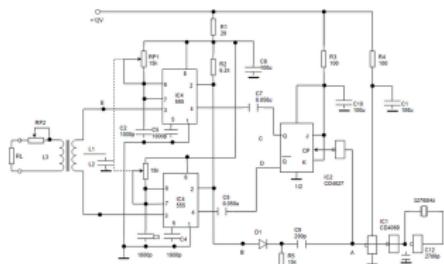
Motivation



Water flow



Traffic flow

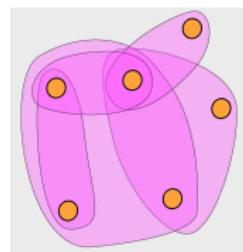
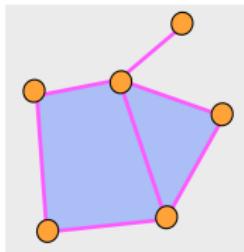
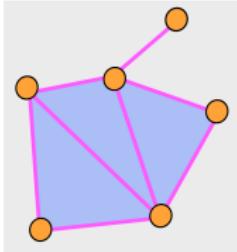


Electric current

- Water networks: detecting water leakage
- Traffic networks: detecting traffic jams
- Power networks: detecting outages

Motivation

- Detecting anomalies in signals with multi-way dependencies

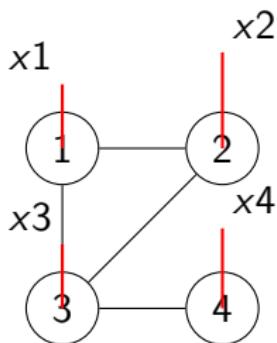


Challenge

- Data contain an **irregular** structure
- A detector with **theoretical** guarantees

Graph Structure of Node Signals

- Graph \mathcal{G} with Laplacian matrix \mathbf{L}
- Graph signals are mappings $x : \mathcal{V} \rightarrow \mathbb{R}$

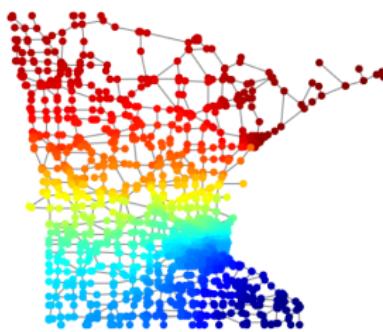


$$\mathbf{L} = \begin{pmatrix} 2 & -1 & -1 & 0 \\ -1 & 2 & -1 & 0 \\ -1 & -1 & 3 & -1 \\ 0 & 0 & -1 & 1 \end{pmatrix} \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

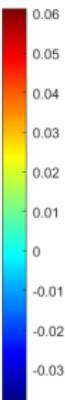
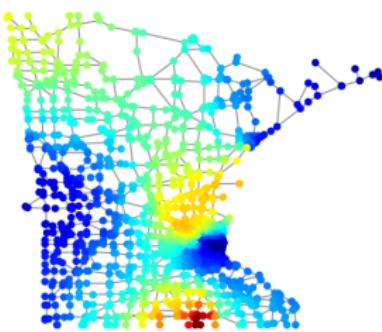
Detection with Graph Signals

- **Normal** behavior – smooth signal $\mathbf{x}^\top \mathbf{L} \mathbf{x}$ is **small**
- **Anomaly** – nonsmooth $\mathbf{x}^\top \mathbf{L} \mathbf{x}$ is **high**

Smooth Graph Signal on Minnesota Roadmap



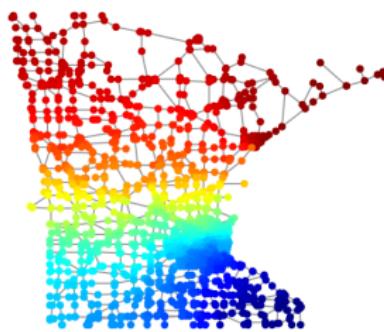
Non-Smooth Graph Signal on Minnesota Roadmap



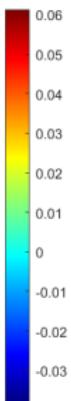
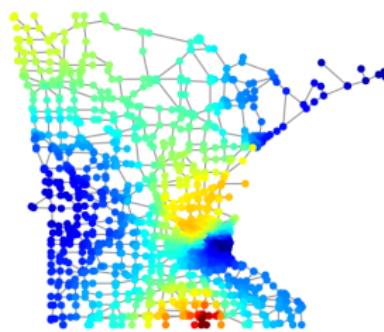
Detection with Graph Signals

- **Normal** behavior – smooth signal $\mathbf{x}^\top \mathbf{L} \mathbf{x}$ is **small**
- **Anomaly** – nonsmooth $\mathbf{x}^\top \mathbf{L} \mathbf{x}$ is **high**

Smooth Graph Signal on Minnesota Roadmap



Non-Smooth Graph Signal on Minnesota Roadmap



Goal: Develop a detection theory for signals on edges and beyond

Topological Data Structure

Simplicial complexes

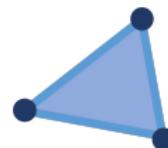
- 0-simplices: nodes – single element
- 1-simplices: edges – pair-wise relationships
- 2-simplices: triangles – relationships of 3 elements
- 3-simplices: tetrahedrons relationships of 4 elements
- ...



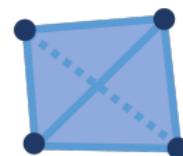
0-simplex



1-simplex



2-simplex



3-simplex

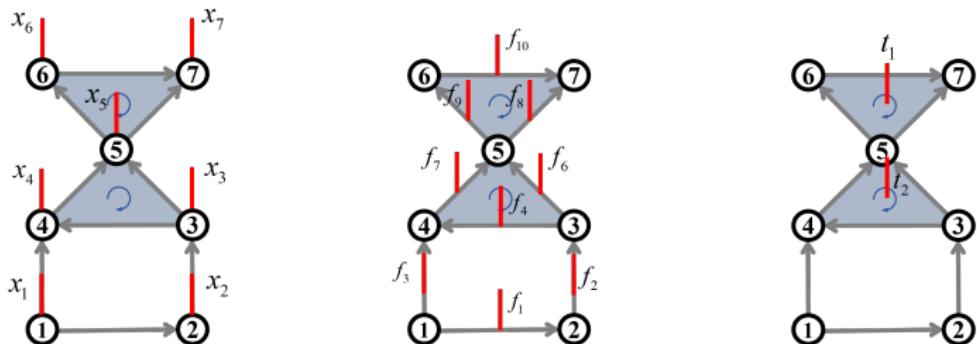
Simplicial Signals

- Mapping¹ from simplices to \mathbb{R}
- We collect k -signal into the vector $\mathbf{s}^k = [s_1^k, \dots, s_{N_k}^k]^T$

¹S. Barbarossa and S. Sardellitti (2020). "Topological signal processing over simplicial complexes". In: *IEEE Transactions on Signal Processing* 68, pp. 2992–3007

Simplicial Signals

- Mapping¹ from simplices to \mathbb{R}
- We collect k -signal into the vector $\mathbf{s}^k = [s_1^k, \dots, s_{N_k}^k]^T$
- Node signals: Measurements of temperature, water pressure
- Edge signals: Measurement of water flow, traffic flow
- Triangle signals: EEG measurements of three voxel interaction



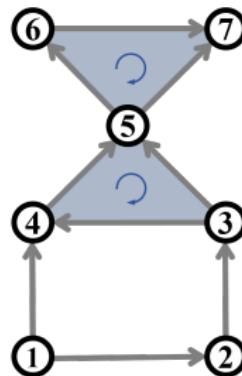
¹S. Barbarossa and S. Sardellitti (2020). "Topological signal processing over simplicial complexes". In: *IEEE Transactions on Signal Processing* 68, pp. 2992–3007

Algebraic Representation of Simplicial Complexes

Topology can be represented by incidence matrices

- \mathbf{B}_1 : node-to-edge incidence matrix
- \mathbf{B}_2 : edge-to-triangle incidence matrix
- ...

$$\mathbf{B}_1 = \left[\begin{array}{ccccccccc} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & -1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 \end{array} \right] \quad \text{edges} \quad \text{nodes}$$
$$\mathbf{B}_2 = \left[\begin{array}{cc} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 1 & 0 \\ -1 & 0 \\ 0 & 1 \\ 0 & -1 \\ 0 & 1 \end{array} \right] \quad \text{triangles} \quad \text{edges}$$



Hodge Laplacians

The whole structure is represented by the Hodge Laplacians:

$$\mathbf{L}_0 = \mathbf{B}_1 \mathbf{B}_1^\top$$

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

$$\mathbf{L}_K = \mathbf{B}_K^\top \mathbf{B}_K$$

Hodge Laplacians

The whole structure is represented by the Hodge Laplacians:

$$\mathbf{L}_0 = \mathbf{B}_1 \mathbf{B}_1^T$$

$$\mathbf{L}_k = \mathbf{B}_k^T \mathbf{B}_k + \mathbf{B}_{k+1} \mathbf{B}_{k+1}^T \quad k = 1, \dots, K-1$$

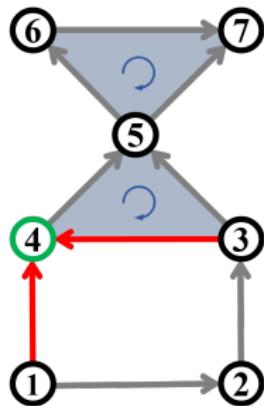
$$\mathbf{L}_K = \mathbf{B}_K^T \mathbf{B}_K$$

Two terms in Hodge Laplacians \mathbf{L}_k :

- Lower Laplacian: $\mathbf{L}_{k,\ell} = \mathbf{B}_k^T \mathbf{B}_k$ represents lower adjacencies
- Upper Laplacian: $\mathbf{L}_{k,u} = \mathbf{B}_{k+1} \mathbf{B}_{k+1}^T$ represents upper adjacencies

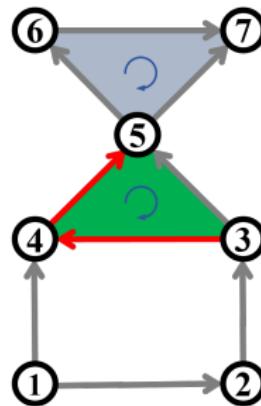
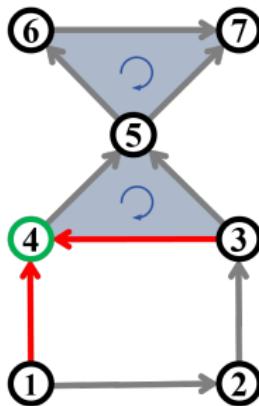
Hodge Laplacians

- Lower Laplacian: $\mathbf{L}_{k,\ell} = \mathbf{B}_k^\top \mathbf{B}_k$ represents lower adjacencies
- Edges (3, 4) and (1, 4) are lower adjacent



Hodge Laplacians

- Lower Laplacian: $\mathbf{L}_{k,\ell} = \mathbf{B}_k^T \mathbf{B}_k$ represents lower adjacencies
- Edges (3, 4) and (1, 4) are lower adjacent



- Upper Laplacian: $\mathbf{L}_{k,u} = \mathbf{B}_{k+1} \mathbf{B}_{k+1}^T$ represents upper adjacencies
- Edge (3, 4) and (4, 5) are upper adjacent

Hodge Decomposition

- Hodge Laplacian structure

$$\mathbf{L}_k = \mathbf{B}_k^T \mathbf{B}_k + \mathbf{B}_{k+1} \mathbf{B}_{k+1}^T$$

Hodge Decomposition

- Hodge Laplacian structure

$$\mathbf{L}_k = \mathbf{B}_k^T \mathbf{B}_k + \mathbf{B}_{k+1} \mathbf{B}_{k+1}^T$$

- The space of k -signal can be decomposed into three subspaces

$$\mathbb{R}^{N_k} \equiv \text{im}(\mathbf{B}_k^T) \oplus \ker(\mathbf{L}_k) \oplus \text{im}(\mathbf{B}_{k+1})$$

Hodge Decomposition

- Hodge Laplacian structure

$$\mathbf{L}_k = \mathbf{B}_k^\top \mathbf{B}_k + \mathbf{B}_{k+1} \mathbf{B}_{k+1}^\top$$

- The space of k -signal can be decomposed into three subspaces

$$\mathbb{R}^{N_k} \equiv \text{im}(\mathbf{B}_k^\top) \oplus \ker(\mathbf{L}_k) \oplus \text{im}(\mathbf{B}_{k+1})$$

- Any edge flow signal can be decomposed as

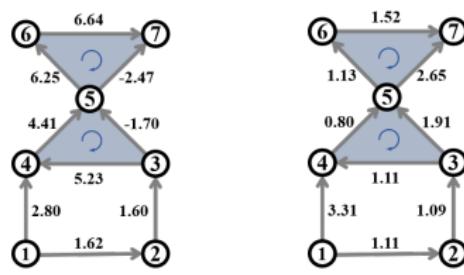
$$\mathbf{f} = \mathbf{f}_G + \mathbf{f}_C + \mathbf{f}_H$$

Hodge Components

Edge flow signal can be decomposed as

$$\mathbf{f} = \mathbf{f}_G + \mathbf{f}_C + \mathbf{f}_H$$

- Gradient flow \mathbf{f}_G : Difference of node signals



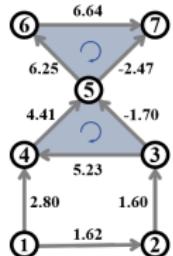
Edge flow \mathbf{f} \mathbf{f}_G

Hodge Components

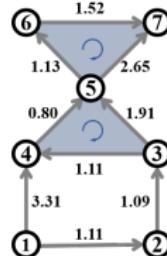
Edge flow signal can be decomposed as

$$\mathbf{f} = \mathbf{f}_G + \mathbf{f}_C + \mathbf{f}_H$$

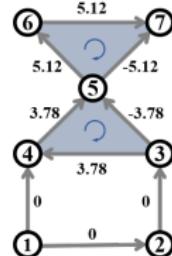
- Gradient flow \mathbf{f}_G : Difference of node signals
- Curl flow \mathbf{f}_C : Local flows circulating around triangles



Edge flow \mathbf{f}



\mathbf{f}_G



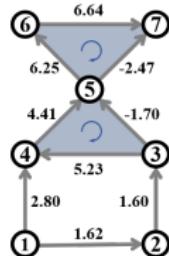
\mathbf{f}_C

Hodge Components

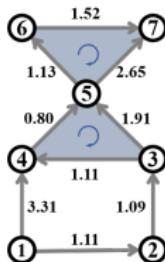
Edge flow signal can be decomposed as

$$\mathbf{f} = \mathbf{f}_G + \mathbf{f}_C + \mathbf{f}_H$$

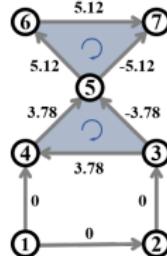
- **Gradient flow \mathbf{f}_G :** Difference of node signals
- **Curl flow \mathbf{f}_C :** Local flows circulating around triangles
- **Harmonic flow \mathbf{f}_H :** Global circulating flow



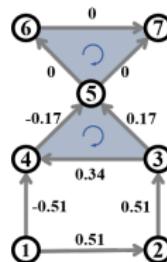
Edge flow \mathbf{f}



\mathbf{f}_C



\mathbf{f}_G



\mathbf{f}_H

Hodge Decomposition

- Eigendecomposition of the first Hodge Laplacian $\mathbf{L}_1 = \mathbf{U} \Lambda \mathbf{U}^T$
- The eigenvectors \mathbf{U} can be rearranged into $[\mathbf{U}_G \; \mathbf{U}_C \; \mathbf{U}_H]$

Hodge Decomposition

- Eigendecomposition of the first Hodge Laplacian $\mathbf{L}_1 = \mathbf{U} \Lambda \mathbf{U}^T$
- The eigenvectors \mathbf{U} can be rearranged into $[\mathbf{U}_G \ \mathbf{U}_C \ \mathbf{U}_H]$

Edge flow signal can be decomposed as

$$\mathbf{f} = \mathbf{f}_G + \mathbf{f}_C + \mathbf{f}_H$$

Hodge eigenvectors

- \mathbf{U}_G spans the **gradient** subspace, $\mathbf{f}_G = \mathbf{U}_G \hat{\mathbf{f}}_G$

Hodge Decomposition

- Eigendecomposition of the first Hodge Laplacian $\mathbf{L}_1 = \mathbf{U} \Lambda \mathbf{U}^T$
- The eigenvectors \mathbf{U} can be rearranged into $[\mathbf{U}_G \ \mathbf{U}_C \ \mathbf{U}_H]$

Edge flow signal can be decomposed as

$$\mathbf{f} = \mathbf{f}_G + \mathbf{f}_C + \mathbf{f}_H$$

Hodge eigenvectors

- \mathbf{U}_G spans the **gradient** subspace, $\mathbf{f}_G = \mathbf{U}_G \hat{\mathbf{f}}_G$
- \mathbf{U}_C spans the **curl** subspace, $\mathbf{f}_C = \mathbf{U}_C \hat{\mathbf{f}}_C$

Hodge Decomposition

- Eigendecomposition of the first Hodge Laplacian $\mathbf{L}_1 = \mathbf{U} \Lambda \mathbf{U}^T$
- The eigenvectors \mathbf{U} can be rearranged into $[\mathbf{U}_G \ \mathbf{U}_C \ \mathbf{U}_H]$

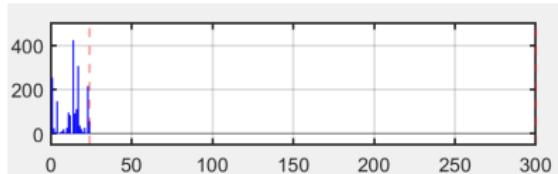
Edge flow signal can be decomposed as

$$\mathbf{f} = \mathbf{f}_G + \mathbf{f}_C + \mathbf{f}_H$$

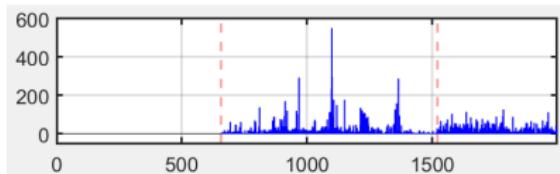
Hodge eigenvectors

- \mathbf{U}_G spans the **gradient** subspace, $\mathbf{f}_G = \mathbf{U}_G \hat{\mathbf{f}}_G$
- \mathbf{U}_C spans the **curl** subspace, $\mathbf{f}_C = \mathbf{U}_C \hat{\mathbf{f}}_C$
- \mathbf{U}_H spans the **harmonic** subspace, $\mathbf{f}_H = \mathbf{U}_H \hat{\mathbf{f}}_H$

Examples of Real-world Flows



Forex

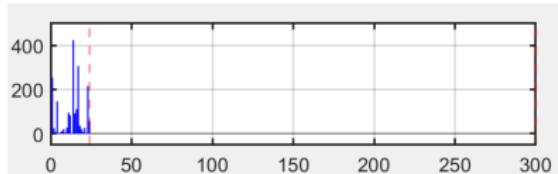


Lastfm

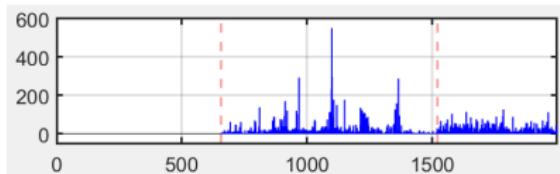
Observation-1: Normally, signal is localized in specific subspace

- Forex flow only has non-curl component
- Lastfm flow only has non-gradient component

Examples of Real-world Flows



Forex



Lastfm

Observation-1: Normally, signal is localized in specific subspace

- Forex flow only has non-curl component
- Lastfm flow only has non-gradient component

Observation-2: Under anomalies, we will have signal subspace spillage

Hypothesis Testing

- \mathbf{x} – true edge flow
- $\mathbf{f} = \mathbf{x} + \mathbf{n}$ – observed edge flow with $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I})$.

\mathcal{H}_0 : \mathbf{x} lives in some specific Hodge subspace

\mathcal{H}_1 : \mathbf{x} does not live in this Hodge subspace

Hypothesis Testing

Hypothesis \mathcal{H}_0 : the subspace is spanned by eigenvectors

$$\mathbf{U}_{\Delta} \in \{\mathbf{U}_G, \mathbf{U}_C, \mathbf{U}_H, [\mathbf{U}_G, \mathbf{U}_C], [\mathbf{U}_G, \mathbf{U}_H], [\mathbf{U}_C, \mathbf{U}_H]\}.$$

The complement eigenvectors are $\mathbf{U}_{\bar{\Delta}}$ so that $\mathbf{U}_{\bar{\Delta}}^T \mathbf{U}_{\Delta} = \mathbf{0}$.

Hypothesis Testing

Hypothesis \mathcal{H}_0 : the subspace is spanned by eigenvectors

$$\mathbf{U}_{\Delta} \in \{\mathbf{U}_G, \mathbf{U}_C, \mathbf{U}_H, [\mathbf{U}_G, \mathbf{U}_C], [\mathbf{U}_G, \mathbf{U}_H], [\mathbf{U}_C, \mathbf{U}_H]\}.$$

The complement eigenvectors are $\mathbf{U}_{\bar{\Delta}}$ so that $\mathbf{U}_{\bar{\Delta}}^T \mathbf{U}_{\Delta} = \mathbf{0}$.

Hypothesis \mathcal{H}_1 : the projection of \mathbf{f} in the complement subspace is

$$\hat{\mathbf{f}}_{\bar{\Delta}} = \mathbf{U}_{\bar{\Delta}}^T \mathbf{x} + \mathbf{U}_{\bar{\Delta}}^T \mathbf{n} = \hat{\mathbf{x}}_{\bar{\Delta}} + \hat{\mathbf{n}}_{\bar{\Delta}}.$$

Hypothesis Testing

Hypothesis \mathcal{H}_0 : the subspace is spanned by eigenvectors

$$\mathbf{U}_{\Delta} \in \{\mathbf{U}_G, \mathbf{U}_C, \mathbf{U}_H, [\mathbf{U}_G, \mathbf{U}_C], [\mathbf{U}_G, \mathbf{U}_H], [\mathbf{U}_C, \mathbf{U}_H]\}.$$

The complement eigenvectors are $\mathbf{U}_{\bar{\Delta}}$ so that $\mathbf{U}_{\bar{\Delta}}^T \mathbf{U}_{\Delta} = \mathbf{0}$.

Hypothesis \mathcal{H}_1 : the projection of \mathbf{f} in the complement subspace is

$$\hat{\mathbf{f}}_{\bar{\Delta}} = \mathbf{U}_{\bar{\Delta}}^T \mathbf{x} + \mathbf{U}_{\bar{\Delta}}^T \mathbf{n} = \hat{\mathbf{x}}_{\bar{\Delta}} + \hat{\mathbf{n}}_{\bar{\Delta}}.$$

$$\begin{aligned}\mathcal{H}_0 : \quad & \hat{\mathbf{f}}_{\bar{\Delta}} = \hat{\mathbf{n}}_{\bar{\Delta}} \\ \mathcal{H}_1 : \quad & \hat{\mathbf{f}}_{\bar{\Delta}} = \mathbf{U}_{\bar{\Delta}}^T \mathbf{x} + \hat{\mathbf{n}}_{\bar{\Delta}}\end{aligned}.$$

Simplicial Hodge Detector

Generalized likelihood ratio test (GLRT)

$$T(\hat{\mathbf{f}}_{\bar{\Delta}}) = \frac{p\left(\hat{\mathbf{f}}_{\bar{\Delta}}; \hat{\mathbf{x}}_{\bar{\Delta}1}^*, \mathcal{H}_1\right)}{p\left(\hat{\mathbf{f}}_{\bar{\Delta}}; \hat{\mathbf{x}}_{\bar{\Delta}0}^*, \mathcal{H}_0\right)} \stackrel{\mathcal{H}_1}{\gtrless} \gamma$$

Simplicial Hodge Detector

Generalized likelihood ratio test (GLRT)

$$T(\hat{\mathbf{f}}_{\bar{\Delta}}) = \frac{p\left(\hat{\mathbf{f}}_{\bar{\Delta}}; \hat{\mathbf{x}}_{\bar{\Delta}1}^*, \mathcal{H}_1\right)}{p\left(\hat{\mathbf{f}}_{\bar{\Delta}}; \hat{\mathbf{x}}_{\bar{\Delta}0}^*, \mathcal{H}_0\right)} \stackrel{\mathcal{H}_1}{\gtrless} \gamma$$

Simplicial Hodge Detector (SHD)

Hypothesis \mathcal{H}_0 : $\hat{\mathbf{x}}_{\bar{\Delta}0}^* = \mathbf{0}$

Hypothesis \mathcal{H}_1 : $\hat{\mathbf{x}}_{\bar{\Delta}1}^* = \hat{\mathbf{f}}_{\bar{\Delta}}$

The SHD is

$$T_{\text{SHD}}(\hat{\mathbf{f}}_{\bar{\Delta}}) = \|\hat{\mathbf{f}}_{\bar{\Delta}}\|_2^2 / \sigma^2 \stackrel{\mathcal{H}_1}{\gtrless} \gamma$$

Spectral priors (sparse, low-pass) can be used to estimate the signal

Performance Guarantees

The detector $T_{\text{SHD}}(\hat{\mathbf{f}}_{\bar{\Delta}})$ has a Chi-square distribution

$$T_{\text{SHD}}(\hat{\mathbf{f}}_{\bar{\Delta}}) \sim \begin{cases} \chi^2_{N_{\bar{\Delta}}} & \text{under } \mathcal{H}_0 \\ \chi^2_{N_{\bar{\Delta}}}(\delta) & \text{under } \mathcal{H}_1 \end{cases}$$

Performance Guarantees

The detector $T_{\text{SHD}}(\hat{\mathbf{f}}_{\bar{\Delta}})$ has a Chi-square distribution

$$T_{\text{SHD}}(\hat{\mathbf{f}}_{\bar{\Delta}}) \sim \begin{cases} \chi^2_{N_{\bar{\Delta}}} & \text{under } \mathcal{H}_0 \\ \chi^2_{N_{\bar{\Delta}}}(\delta) & \text{under } \mathcal{H}_1 \end{cases}$$

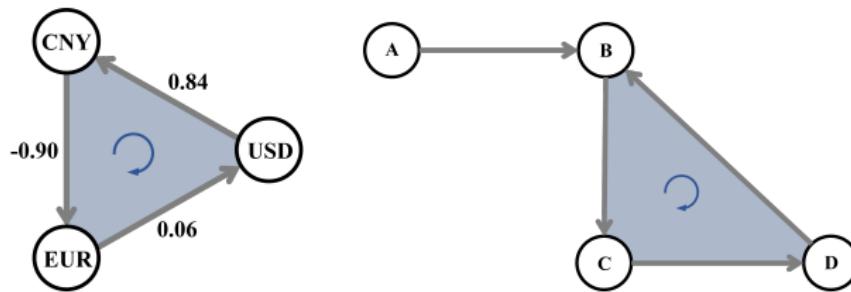
The probability of false alarm and detection are

$$P_{\text{FA}} \triangleq \Pr\{T_{\text{SHD}}(\hat{\mathbf{f}}_{\bar{\Delta}}) > \gamma; \mathcal{H}_0\} = Q_{\chi^2_{N_{\bar{\Delta}}}}(\gamma)$$

$$P_{\text{D}} \triangleq \Pr\{T_{\text{SHD}}(\hat{\mathbf{f}}_{\bar{\Delta}}) > \gamma; \mathcal{H}_1\} = Q_{\chi^2_{N_{\bar{\Delta}}}(\delta)}(\gamma)$$

Experimental Setup

Dataset	\mathcal{H}_0	\mathcal{H}_1
Forex	Gradient flow	$\mathbf{x} = \mathbf{B}_2 \mathbf{t}, \mathbf{t} \sim \mathcal{N}(\mathbf{0}, 0.3 \times 10^{-1} \mathbf{I})$
Lastfm	Non-gradient flow	$\mathbf{x} = \mathbf{B}_1^\top \mathbf{v}, \mathbf{v} \sim \mathcal{N}(\mathbf{0}, 5.0 \times 10^1 \mathbf{I})$
Chicago	$\mathbf{x} = \mathbf{B}_2 \mathbf{t}, \mathbf{t} \sim \mathcal{N}(\mathbf{0}, 2.3 \times 10^{-2} \mathbf{I})$	$\hat{\mathbf{x}}_{G,i} \sim \mathcal{N}(\exp(-i/5)^\top, 0.01)$



Results

SNR = -12 dB

Dataset	SHD-Th.	SHD-Exp.	B-SMSD
Forex	0.71	0.71	0.45
Lastfm	0.99	0.99	0.39
Chicago	0.75	0.75	0.52

- Theoretical and empirical results match each other
- Line-graph detector B-SMSD doesn't consider high-order structures

Conclusion

- A matched subspace detection theory for topological signals
- Relied on Hodge subspace signal projections
- Principled and mathematically tractable
- Allows including topological spectral prior in the detection

Conclusion

- A matched subspace detection theory for topological signals
- Relied on Hodge subspace signal projections
- Principled and mathematically tractable
- Allows including topological spectral prior in the detection
- FW: **Jointly** consider all signals and **sampled** measurements

Conclusion

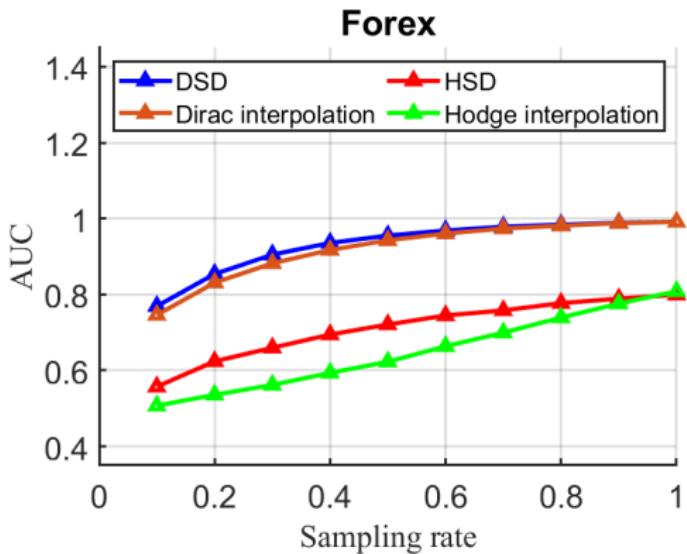
- A matched subspace detection theory for topological signals
- Relied on Hodge subspace signal projections
- Principled and mathematically tractable
- Allows including topological spectral prior in the detection
- FW: **Jointly** consider all signals and **sampled** measurements

Thank you!

Extended Results-Matched Topological Subspace Detector

Detector with missing values

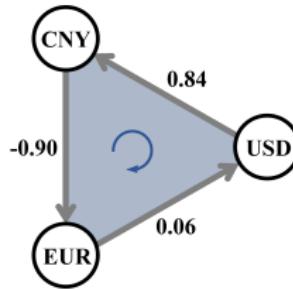
$$T(\mathbf{f}) = \frac{\|\mathbf{f} - \mathbf{U}_{\Delta\Theta}\hat{\mathbf{x}}_0^*\|_2^2 - \|\mathbf{f} - \mathbf{U}_{\Theta}\hat{\mathbf{x}}_1^*\|_2^2}{\sigma^2} \begin{cases} \geqslant \gamma & \mathcal{H}_1 \\ & \mathcal{H}_0 \end{cases}$$



Datasets

Forex(real-world)

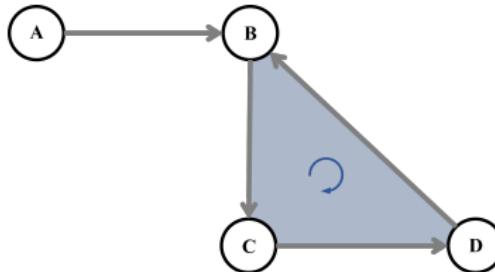
- Exchange rates between 25 different currencies
- 25 nodes, 300 edges and 2300 triangles
- Nodes: 25 different currencies
- Edges: logarithm exchange rate
- Triangles: relationships for three different currencies
- Task: Detecting whether the currency exchange flow is curl-free



Datasets

Lastfm(real-world)

- Users switching artists while playing music
- 657 nodes, 1997 edges and 1276 triangles
- Nodes: 657 different artists
- Edges: transitions between two artists
- Triangles: relationships for three different artists
- Task: Detecting whether the switch flow is divergence-free



Chicago(synthetic)

- Traffic network of Chicago
- 546 nodes, 1088 edges and 112 triangles
- Nodes: 546 different junctions
- Edges: 1088 roads between junctions
- Triangles: areas enclosed by three roads
- Task: Detecting whether the synthetic flow is divergence-free