Graph Signal Processing for Machine Learning

A Review and New Perspectives

Xiaowen Dong, Dorina Thanou, Laura Toni,
Michael Bronstein, Pascal Frossard

ICASSP Tutorial, June 2021

Part III

Applications, Open Challenges and New Perspectives

Outline

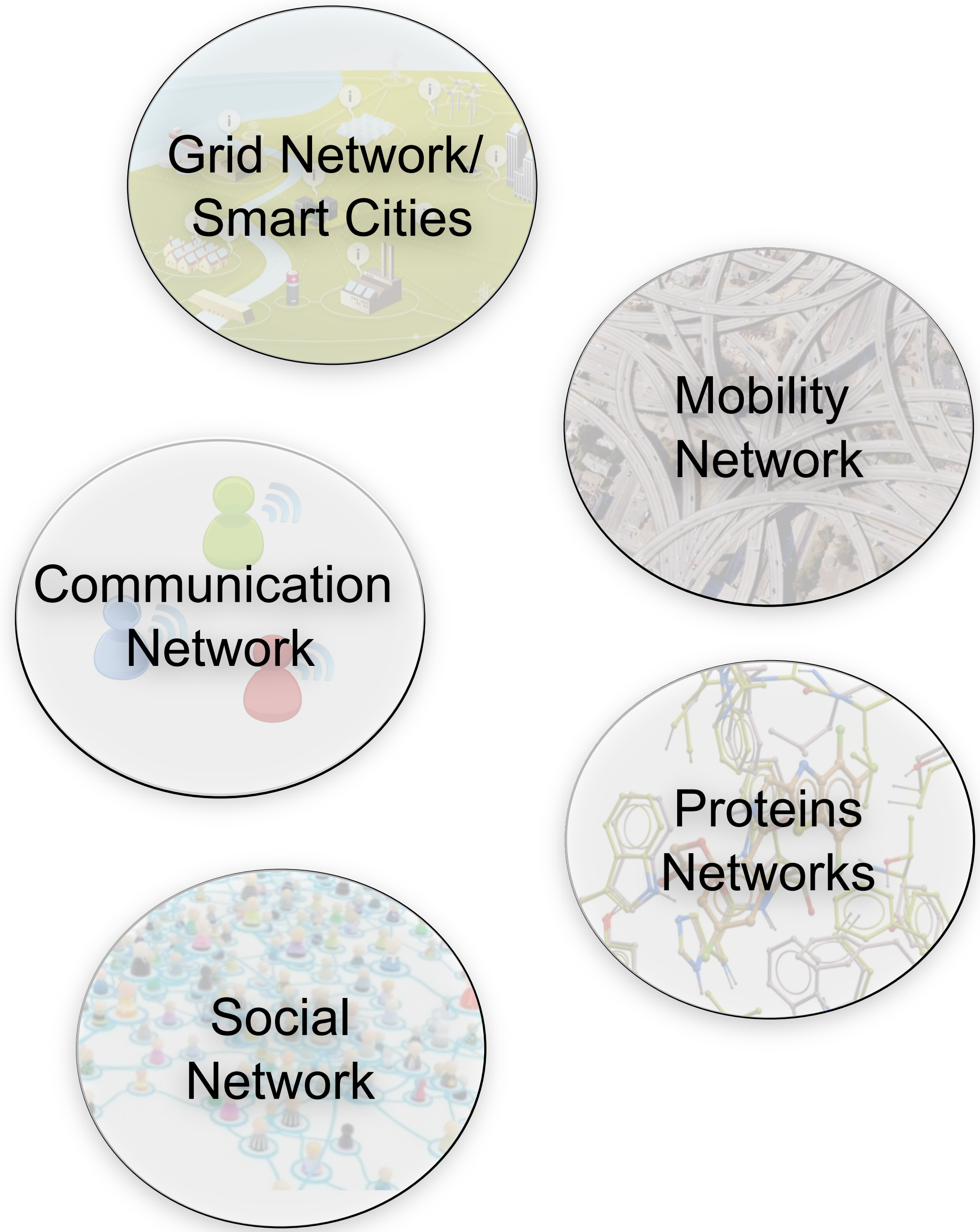
- Brief introduction to graph signal processing (GSP)
- Challenge I: GSP for exploiting data structure
- Challenge II: GSP for improving efficiency and robustness
- Challenge III: GSP for enhancing model interpretability
- Applications
- Summary, open challenges, and new perspectives

Networks are pervasive

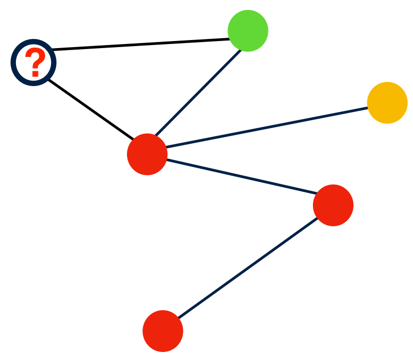
Network Science

Healthcare

Computer Vision

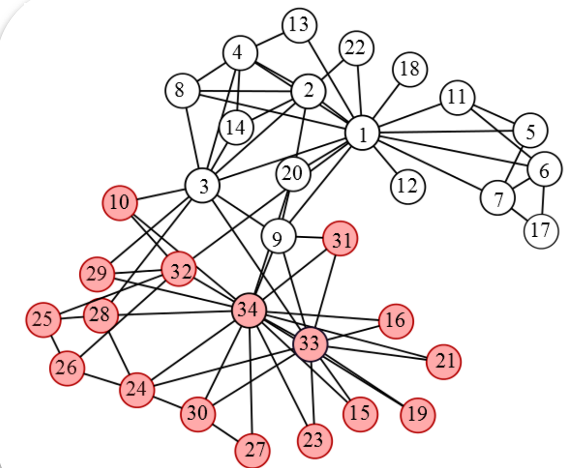


Main Problems for GSP-Based ML



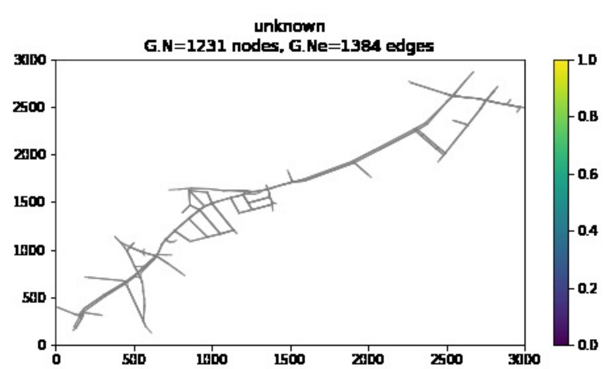
Node-Graph-

Classification



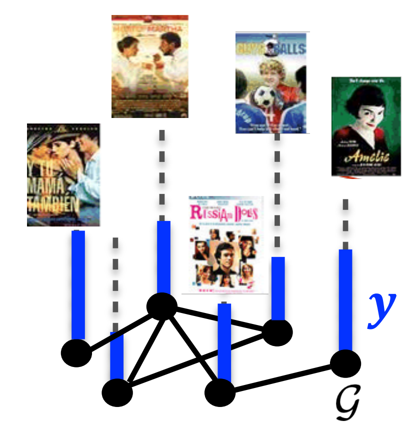
Community detection

(multi scale analysis)



Time series prediction - Network control

(dynamics of graphs)



Online Learning

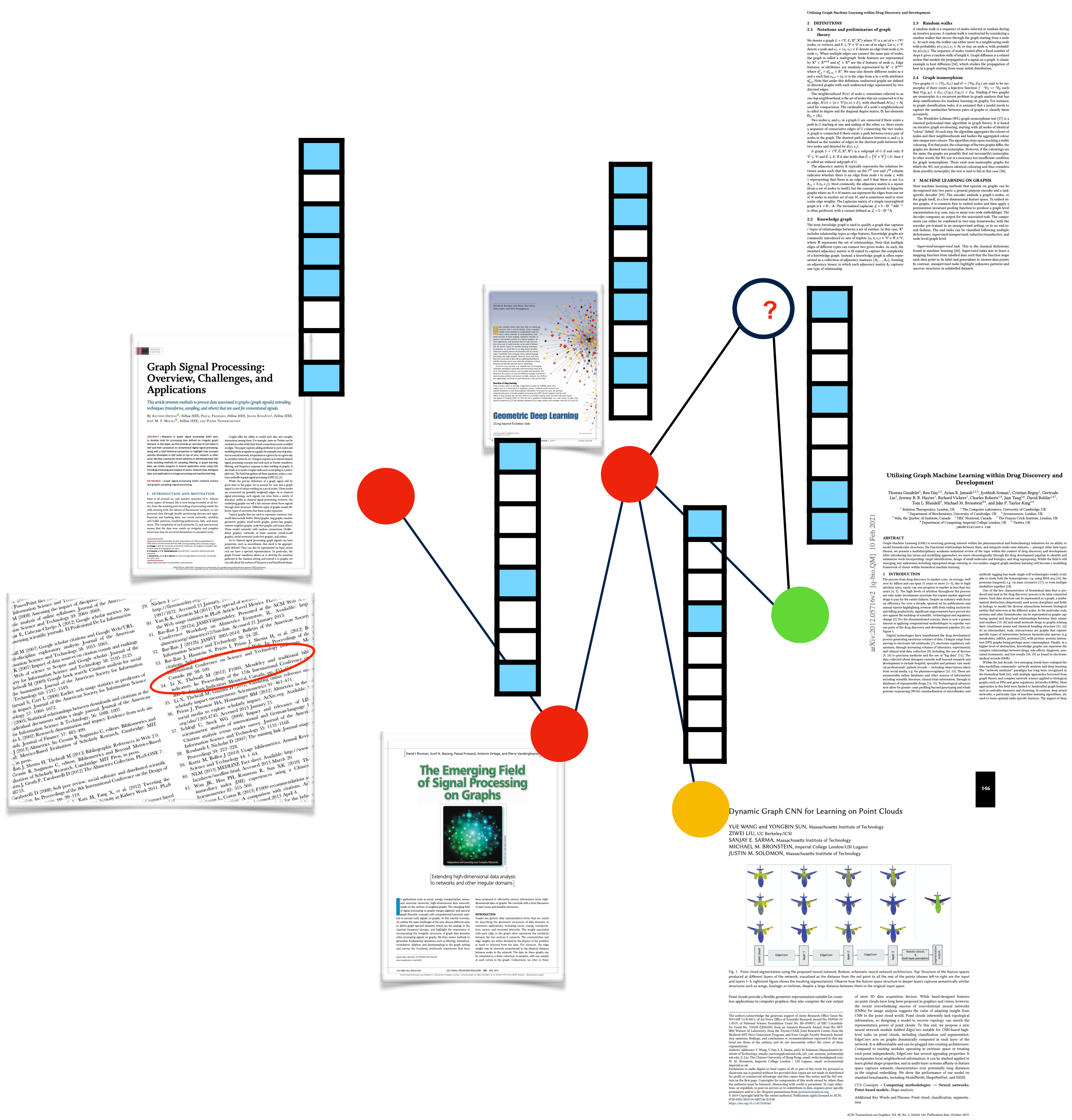
(problem structure inference)

Graph

- Implicit / Explicit

- Given / Constructed
- Inferred

Document Analysis: Node Classification



Graph

Nodes → papers

Edge → paper citation

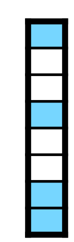
Nodes Label

Signal Processing

Drug Discovery

Computer Vision

Node Signal



content-based features

0: word not present in the paper

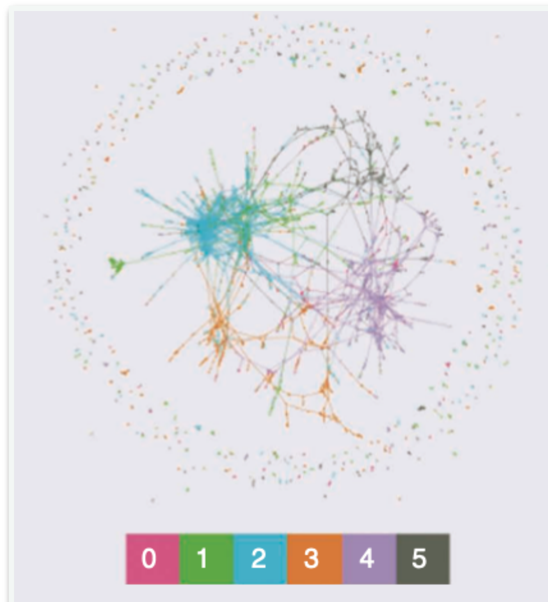
1: word present in the paper

Dataset	#Nodes	#Edges	Train/Dev/Test
Cora	2,708	5,429	140/500/1,000
CiteSeer	3,327	4,723	120/500/1,000
Pubmed	19,717	44,338	60/500/1,000

http://networkrepository.com/cora.php

Document Analysis: Node Classification

CORA-ML

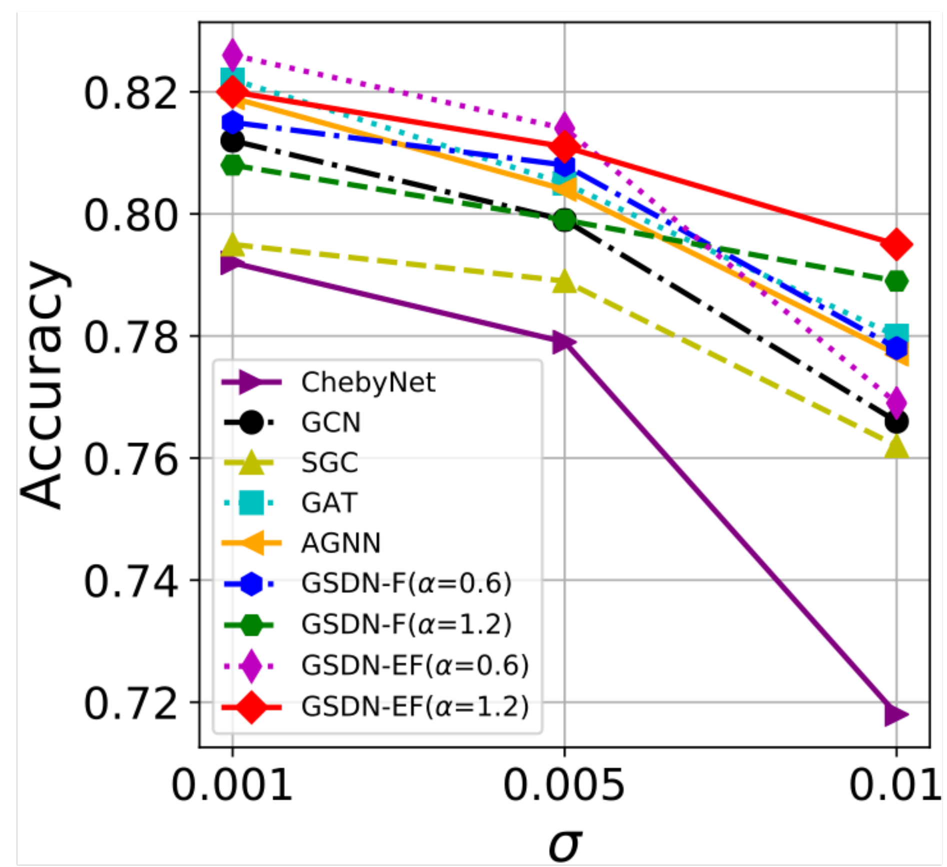


CiteSeer

Dataset	#Nodes	#Edges	Train/Dev/Test
Cora	2,708	5,429	140/500/1,000
CiteSeer	3,327	4,723	120/500/1,000
Pubmed	19,717	44,338	60/500/1,000

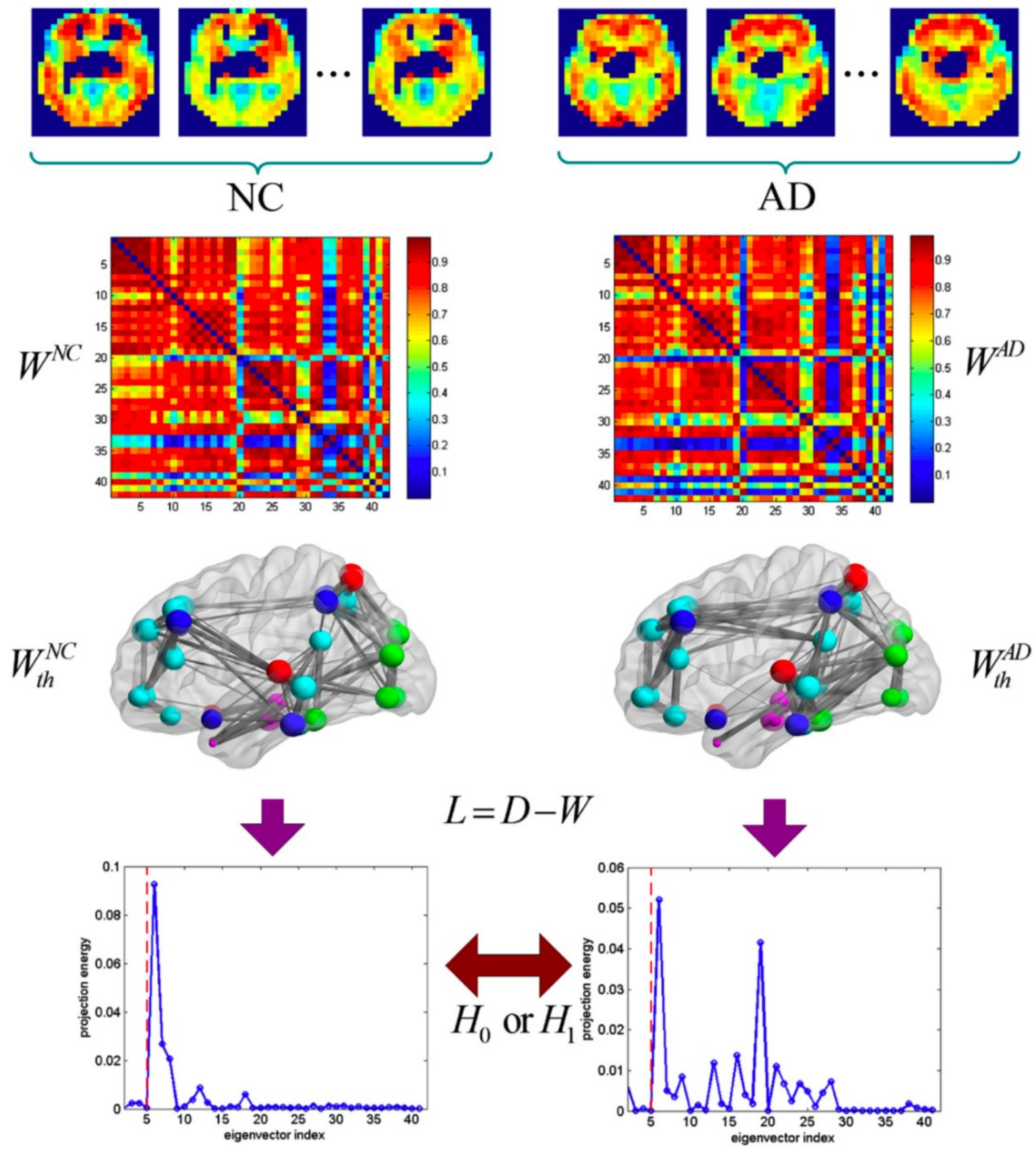
GNNs interpreted as implementing denoising and/or smoothing of graph signals.

Validation with semi-supervised node classification on noisy citation networks



Guoji Fu et al. “Understanding Graph Neural Networks from Graph Signal Denoising Perspectives”, arXiv, 2020

Neuroscience: Graph Classification



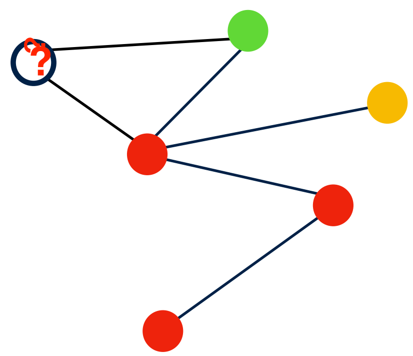
- Two graphs are build based on AD (Alzheimer’s disease) and NC (normal control)
- PET/fMRI data as graph-signals
- Edge weights describing the affinity between each pair of brain regions
- Graph classification as hypothesis testing

H_0: signal smooth on graph G_0

H_1: signal smooth on graph G_1

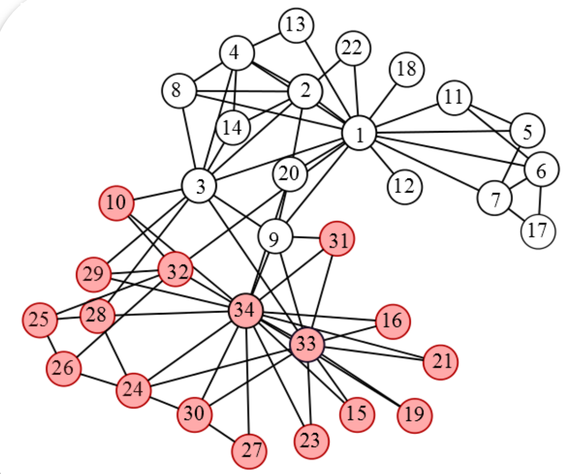
C. Hu et al., “Matched signal detection on graphs: Theory and application to brain imaging data classification”, NeuroImage, 2016.

Main Problems for GSP-Based ML



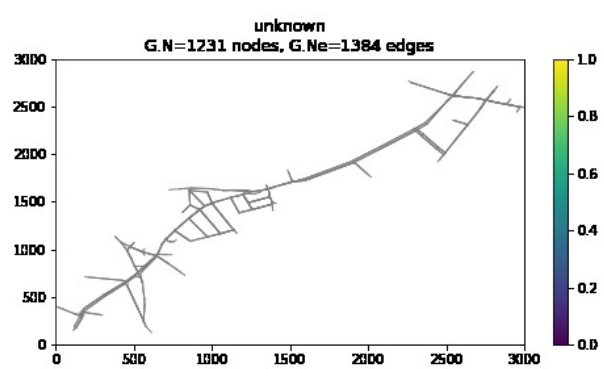
Node-
Graph-

Classification



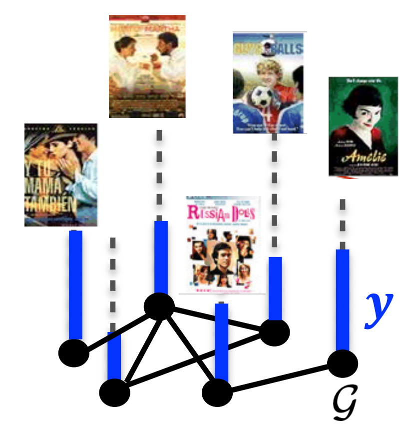
Community detection

(multi scale analysis)



Time series prediction - Network control

(dynamics of graphs)



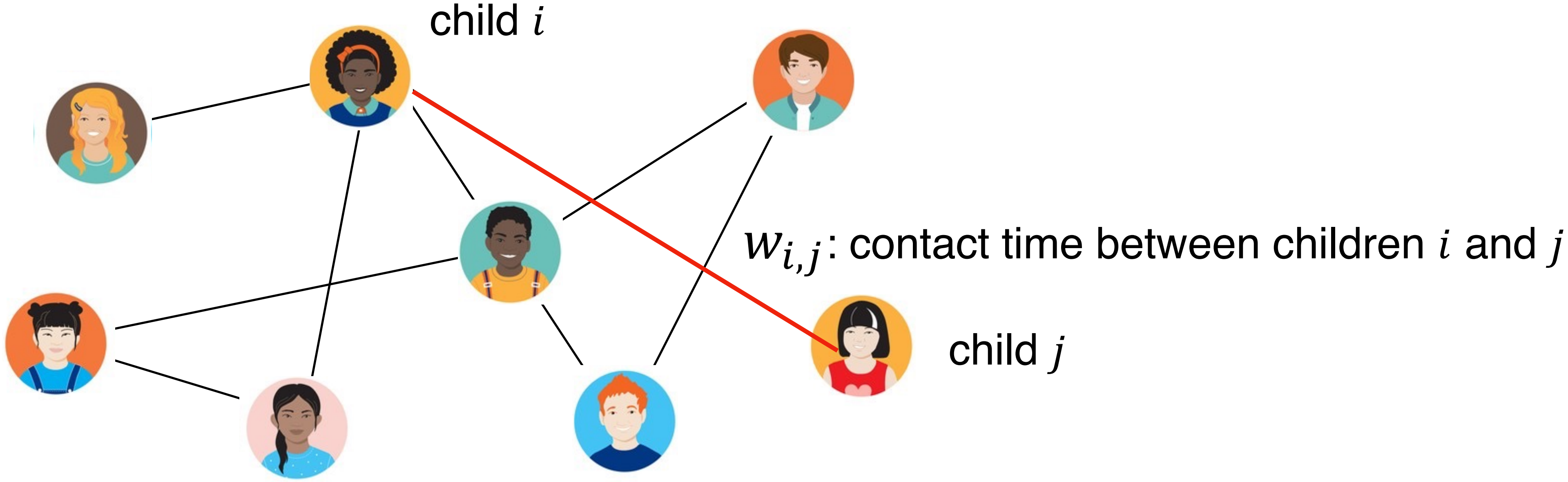
Online Learning

(problem structure inference)

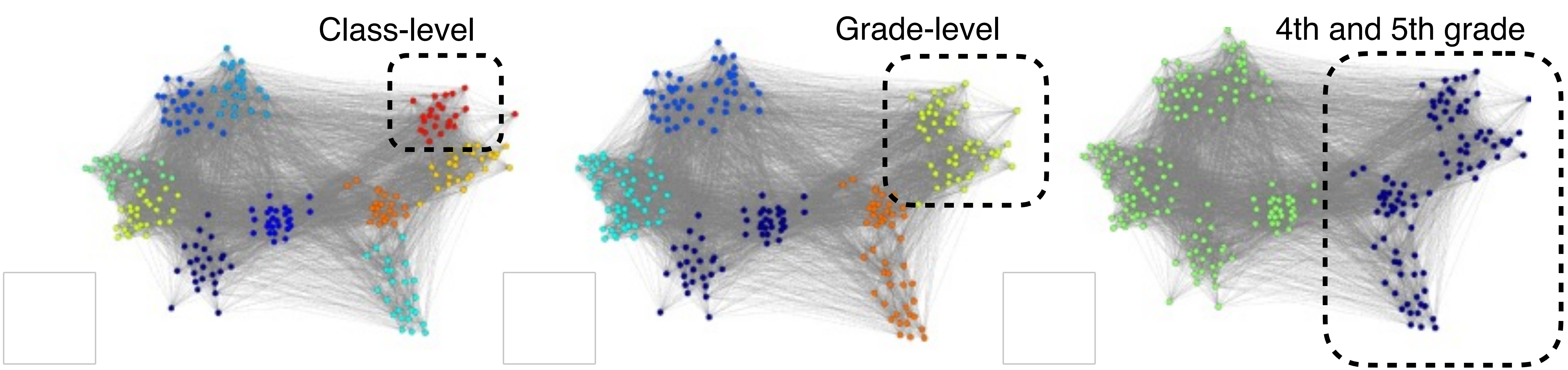
Graph

- Implicit / Explicit
- Given / Constructed
- Inferred

Community detection

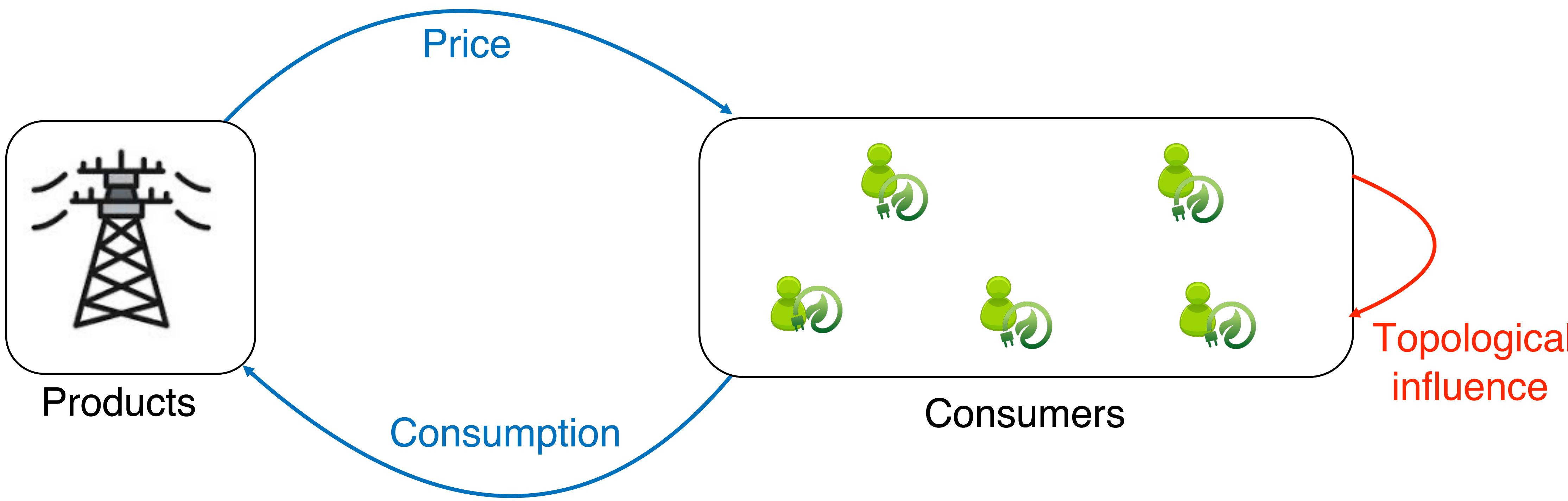


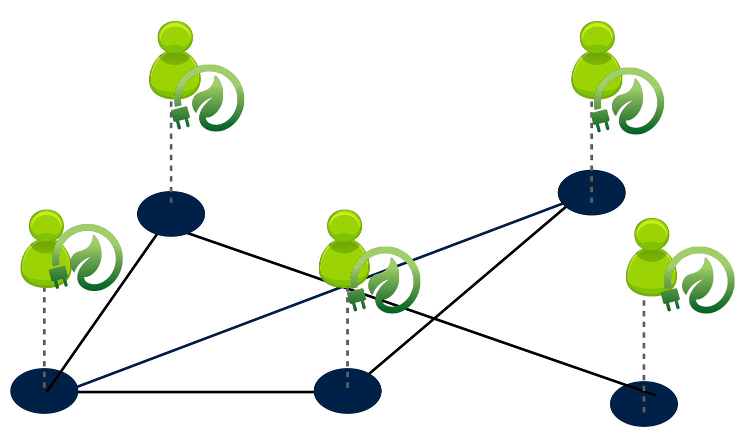
Stratification based on social interactions:
multi-scale community detection based on spectral graph wavelets at different scales



N. Tremblay and Pierre Borgnat, "Graph wavelets for multiscale community mining", IEEE TSP 2014.

Price Experiments in Consumers' Game

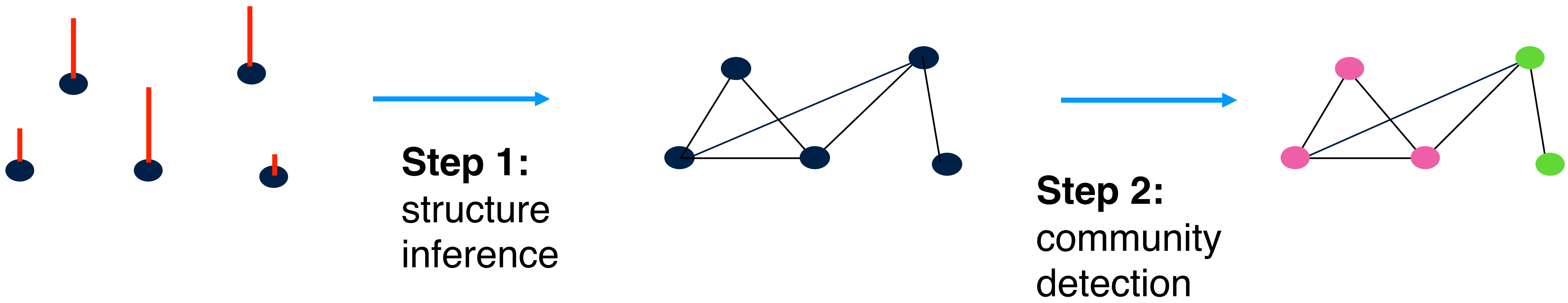




- Consumers as vertices on graph
- (Unknown) topological influence as graph weigh

- Topology inference / users consumption prediction / community detection …

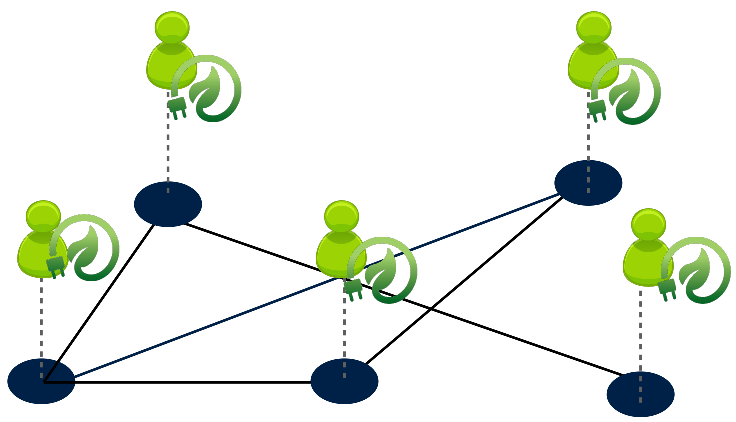
Blind Community Detection



Observations are graph signals modeled as the outputs of an unknown network process represented by a low-rank graph filter

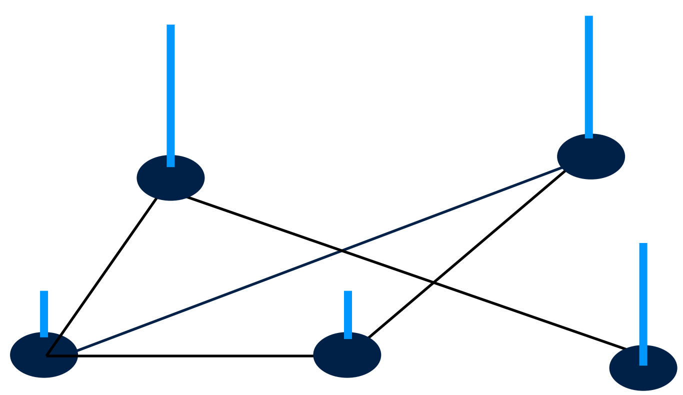
H.-T. Wai, et al., “Blind community detection from low-rank excitations of a graph filter”, IEEE TSP, 2020.
R. Ramakrishna et al., "A User Guide to Low-Pass Graph Signal Processing and Its Applications”, IEEE SPM, 2020.

Blind Community Detection: low-rank filtering

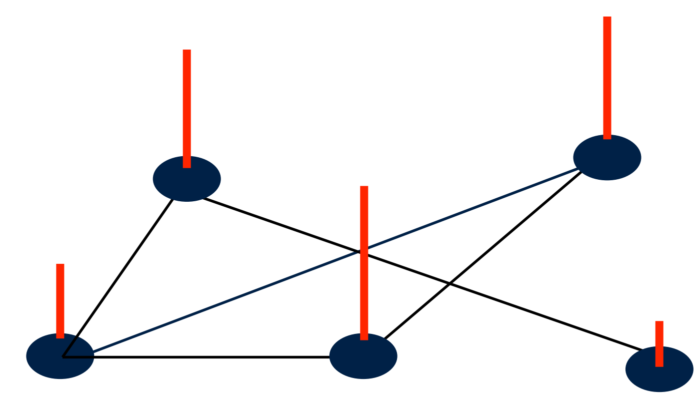


- Consumers as vertices on graph
- (Unknown) topological influence as graph weigh $\boldsymbol{A}$

- Topology inference / users consumption prediction / community detection



$$\tilde{\boldsymbol{y}}^{\ell} = (b\boldsymbol{I} - \boldsymbol{A})^{-1}\tilde{\boldsymbol{p}}^{\ell}$$



excitation signal $\tilde{\boldsymbol{p}}^{\ell}$

observed graph signal $\tilde{\boldsymbol{y}}^{\ell} := \frac{1}{L}\sum_{\tau=1}^{L} \boldsymbol{y}^{\tau} - \boldsymbol{y}^{\ell}$

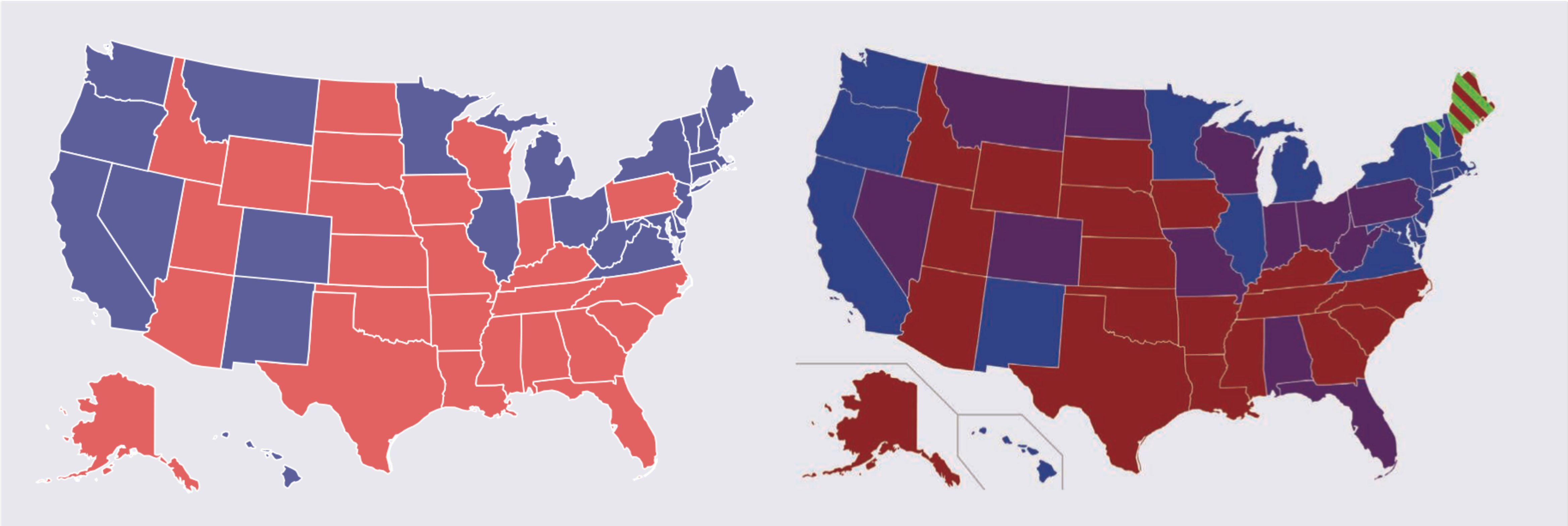
Observations are graph signals modeled as the outputs of an unknown network process represented by a graph filter

H.-T. Wai, et al., "Blind community detection from low-rank excitations of a graph filter", IEEE TSP, 2020.
R. Ramakrishna et al., "A User Guide to Low-Pass Graph Signal Processing and Its Applications", IEEE SPM, 2020.

Blind Community Detection

- The covariance matrix of **observed graph signals** is a **sketch of the Laplacian matrix** that retains coarse topological features of the graph, like communities

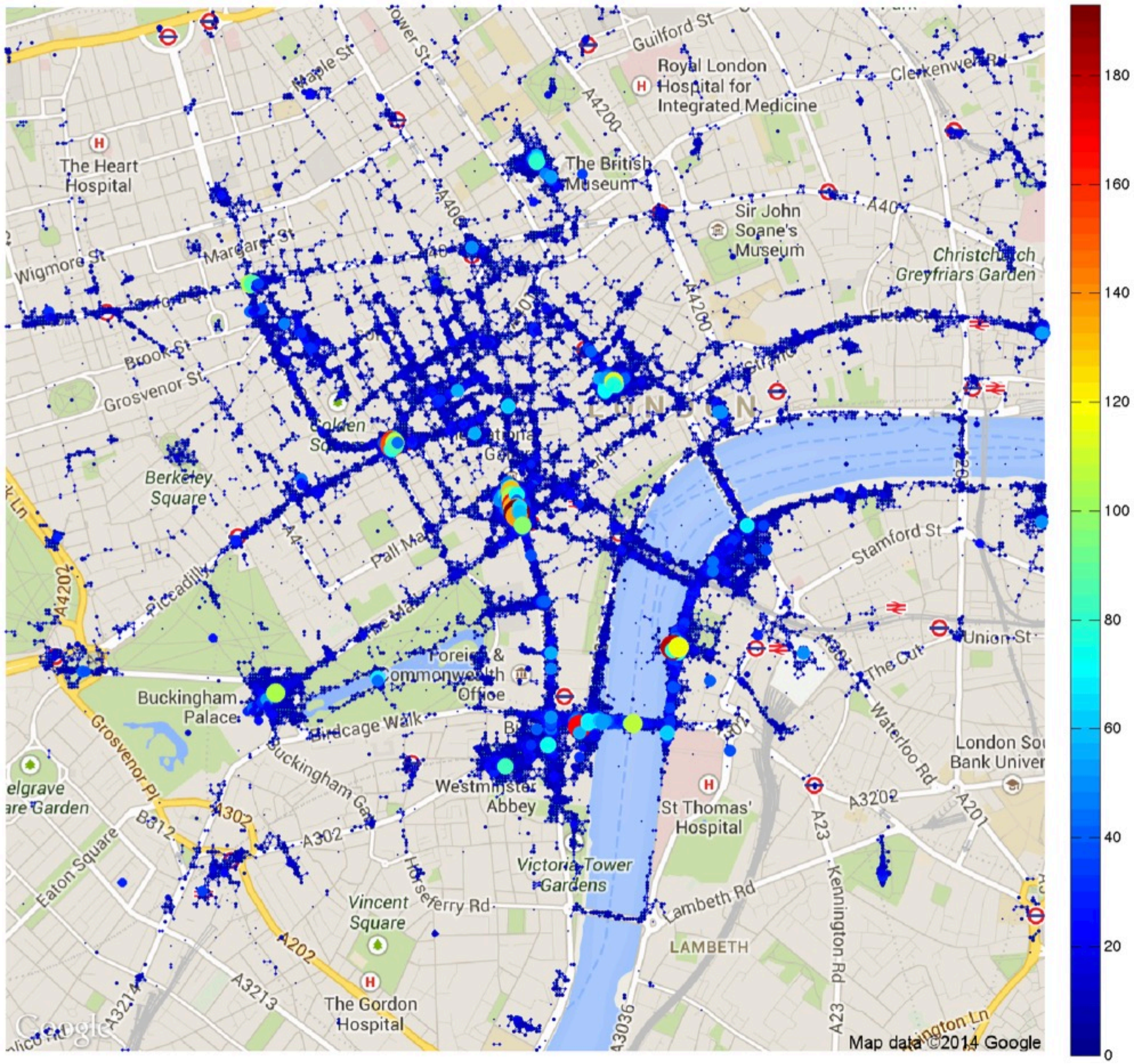
- Blind CD a**pproaches the performance of spectral clustering** (under given conditions and assuming graph filter being low-pass)



Rollcall data may be modeled as the equilibrium of an opinion dynamics process with stubborn agents

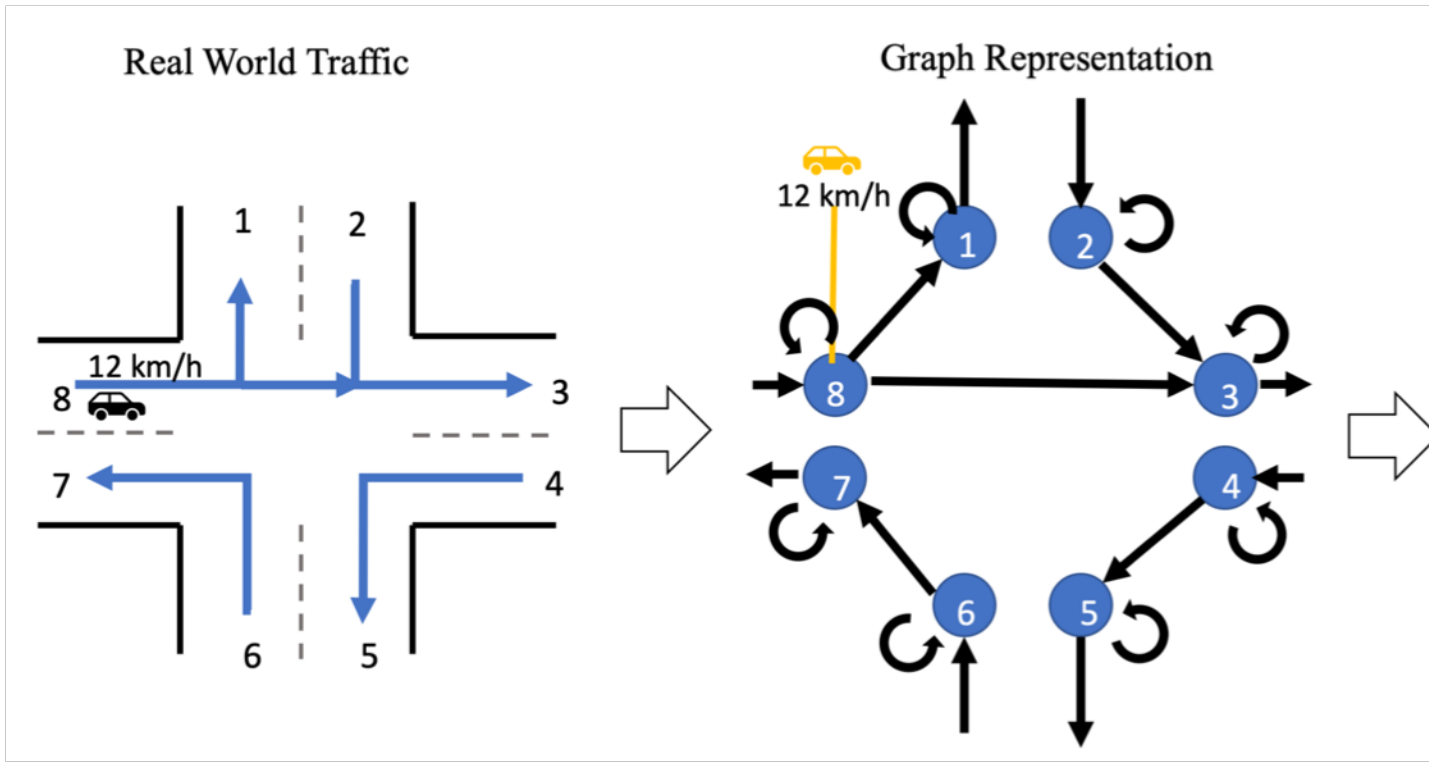
H.-T. Wai, et al., “Blind community detection from low-rank excitations of a graph filter”, IEEE TSP, 2020.
R. Ramakrishna et al., "A User Guide to Low-Pass Graph Signal Processing and Its Applications”, IEEE SPM, 2020.

Mobility inference



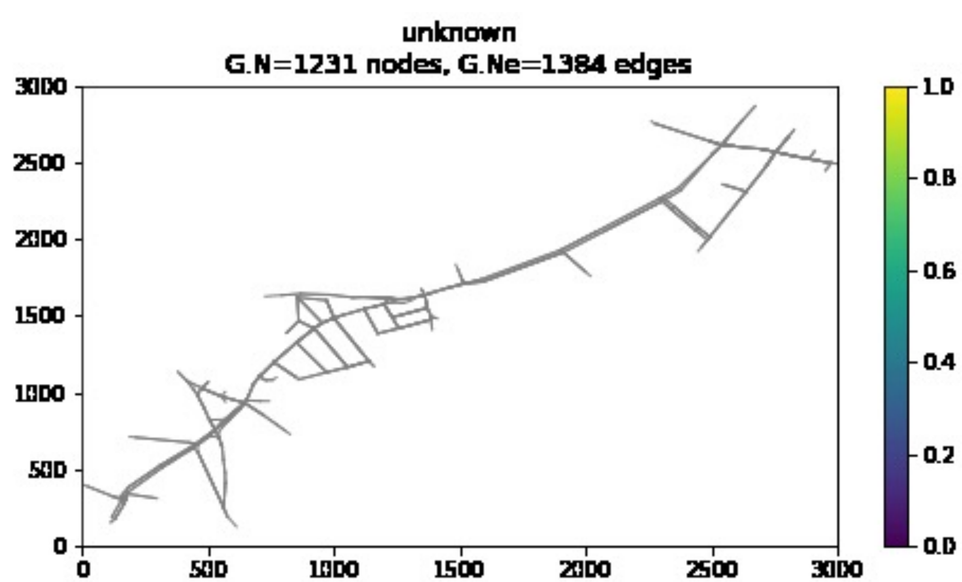
X. Dong et al. "Inference of mobility patterns via spectral graph wavelets", ICASSP, 2013.

Transportation Network: mobility inference



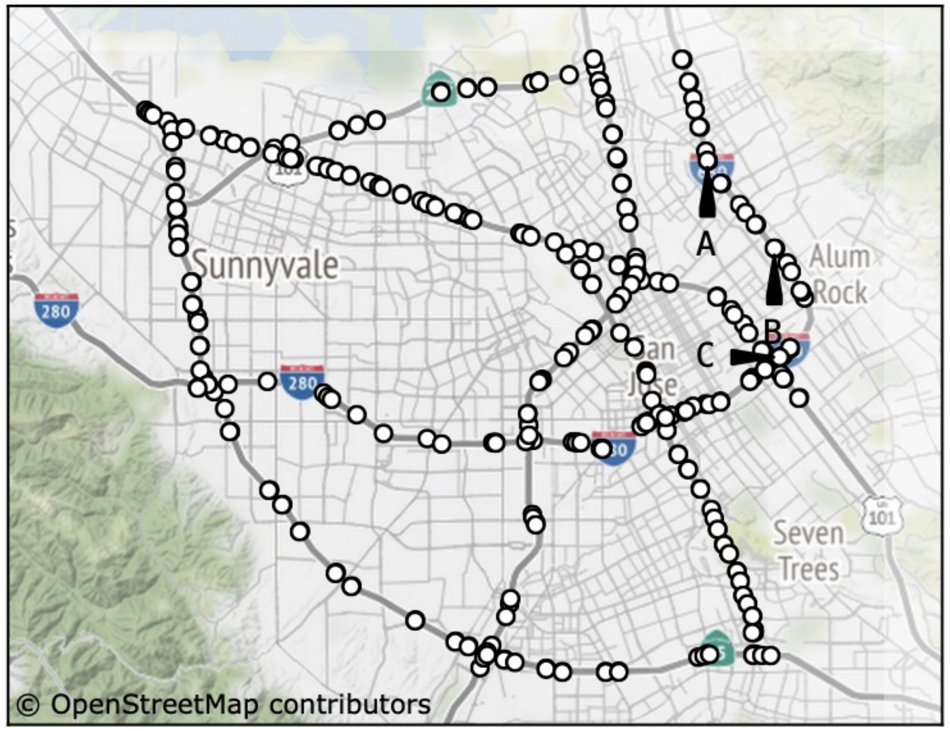
Undirected Adjacency Matrix

	1	2	3	4	5	6	7	8
1	1	0	0	0	0	0	0	1
2	0	1	1	0	0	0	0	0
3	0	1	1	0	0	0	0	1
4	0	0	0	1	1	0	0	0
5	0	0	0	1	1	0	0	0
6	0	0	0	0	0	1	1	0
7	0	0	0	0	0	1	1	0
8	1	0	0	0	0	0	0	1



Traffic propagation modelled as heat diffusion on graph signal

Transportation Network: mobility inference



Algorithm 2 Prediction of traffic features (h-steps ahead)

function PREDICTION($\mathbf{x}_t^d, h, \hat{\mathbf{H}}_t, \cdots, \hat{\mathbf{H}}_{t+h-1}$)
  Set $\mathbf{p} = \mathbf{x}_t^d$
  for $i \in [0, h-1]$ **do**
    Set $\mathbf{p} = \hat{\mathbf{H}}_{t+i}\mathbf{p}$
  end for
  $\mathbf{x}_{t+h|t} = \mathbf{p}$
return $\mathbf{x}_{t+h|t}$
end function

Heat diffusion kernels embedded into dynamic linear model to exploit topological information of the transportation network

Dynamic linear model

$$\mathbf{x}_{t+1}^d = \mathbf{H}_t\mathbf{x}_t^d + \mathbf{n}_t^d, \forall t \in [0, T-1]$$

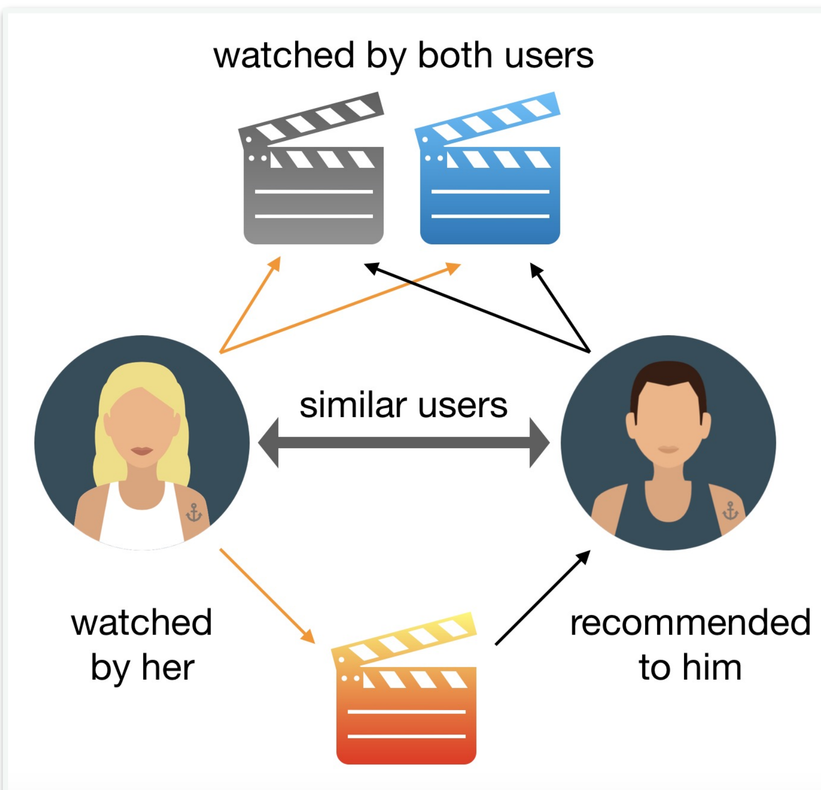
$$\mathbf{H}_t = \boxed{\mathbf{H}_t^{\mathcal{G}}(\mathcal{T})} + \boxed{\check{\mathbf{H}}_t}$$

Demand matrix (exogenous)

Internal diffusion matrix (endogenous) $\mathbf{H}^{\mathcal{G}}(\tau) = e^{-\tau\mathbf{L}(\mathcal{G})}$

S. Kwak, “Predicting traffic signals on transportation networks using spatio-temporal correlations on graphs”, arXiv 2021.

Recommender systems: Online Learning



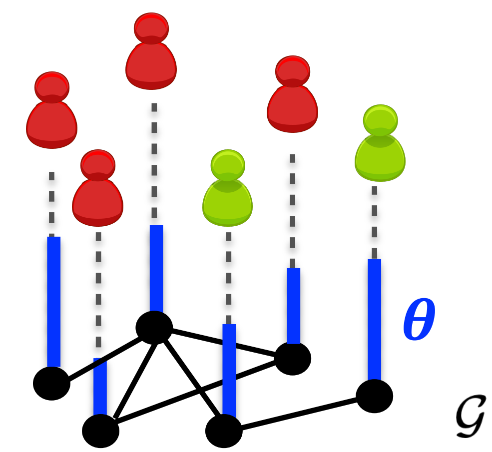
Graphs model items and users similarities

https://www.xpertup.com/downloads/movie-recommender-system/

Recommender systems: User graph

- User preferences mapped into a graph of similarities

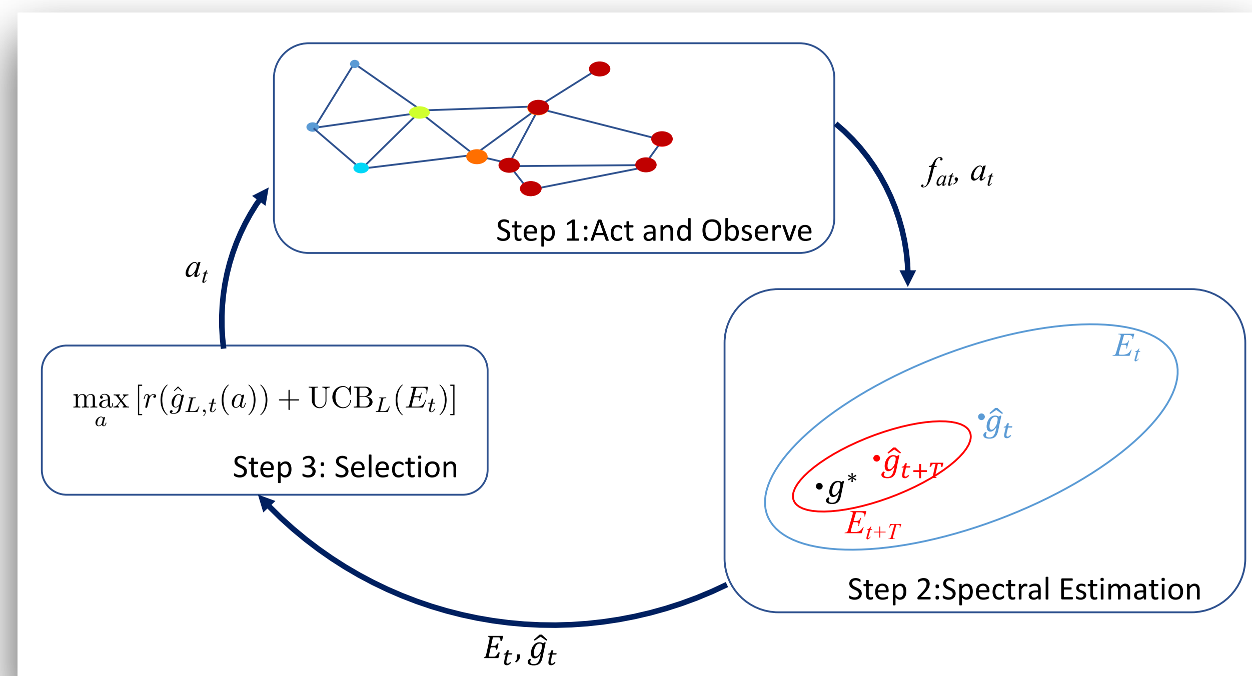
$$\boldsymbol{\Theta} = [\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_N]^T \in \mathbb{R}^{N\times d}\text{: signal on graph}$$

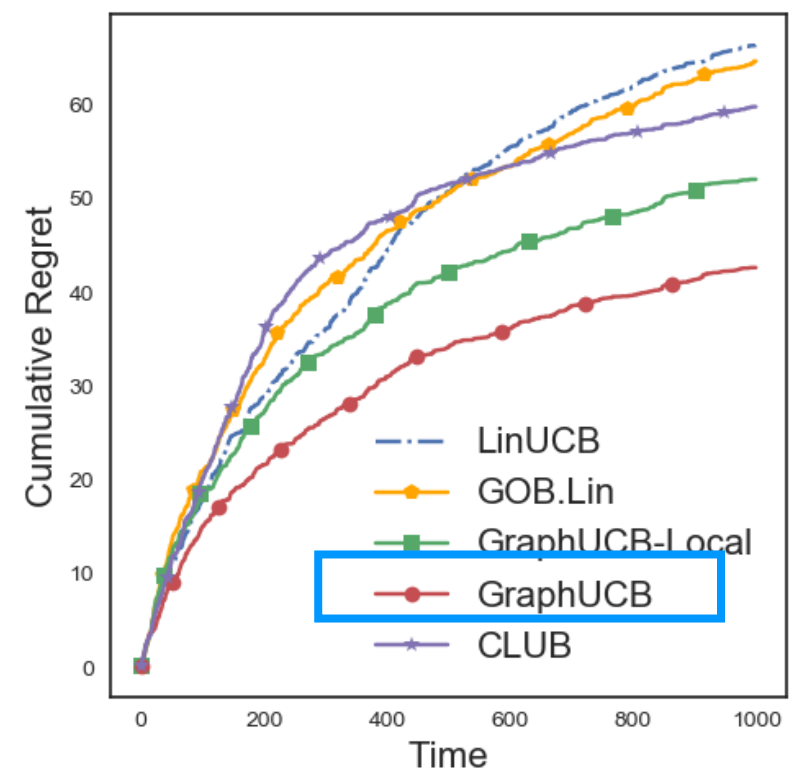


- Exploitation of smoothness prior

$$\hat{\boldsymbol{\Theta}}_t = \arg\min_{\boldsymbol{\Theta}\in\mathbb{R}^{n\times d}} \sum_{i=1}^{n} \sum_{\tau\in t_i} \underbrace{(\mathbf{x}_\tau^T\boldsymbol{\theta}_i - y_{i,\tau})^2}_{\text{fidelity term}} + \alpha\, \underbrace{tr(\boldsymbol{\Theta}^T \mathcal{L} \boldsymbol{\Theta})}_{\text{smoothness regularizer}}$$

Laplacian-regularised estimator within online learning framework

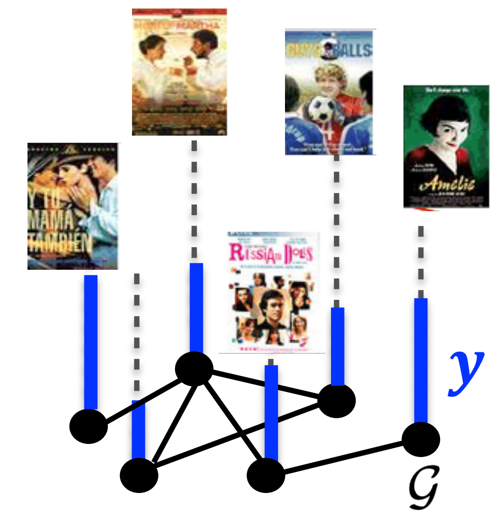




K. Yang, “Laplacian-regularized graph bandits: Algorithms and theoretical analysis”, AISTATS 2020

Recommender systems: Item graph

- Items as nodes on graph
  $\boldsymbol{y} = [\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_N]^T \in \mathbb{R}^{N \times d}$: reward (unknown)
- Signal (reward) is unknown and needs to be inferred



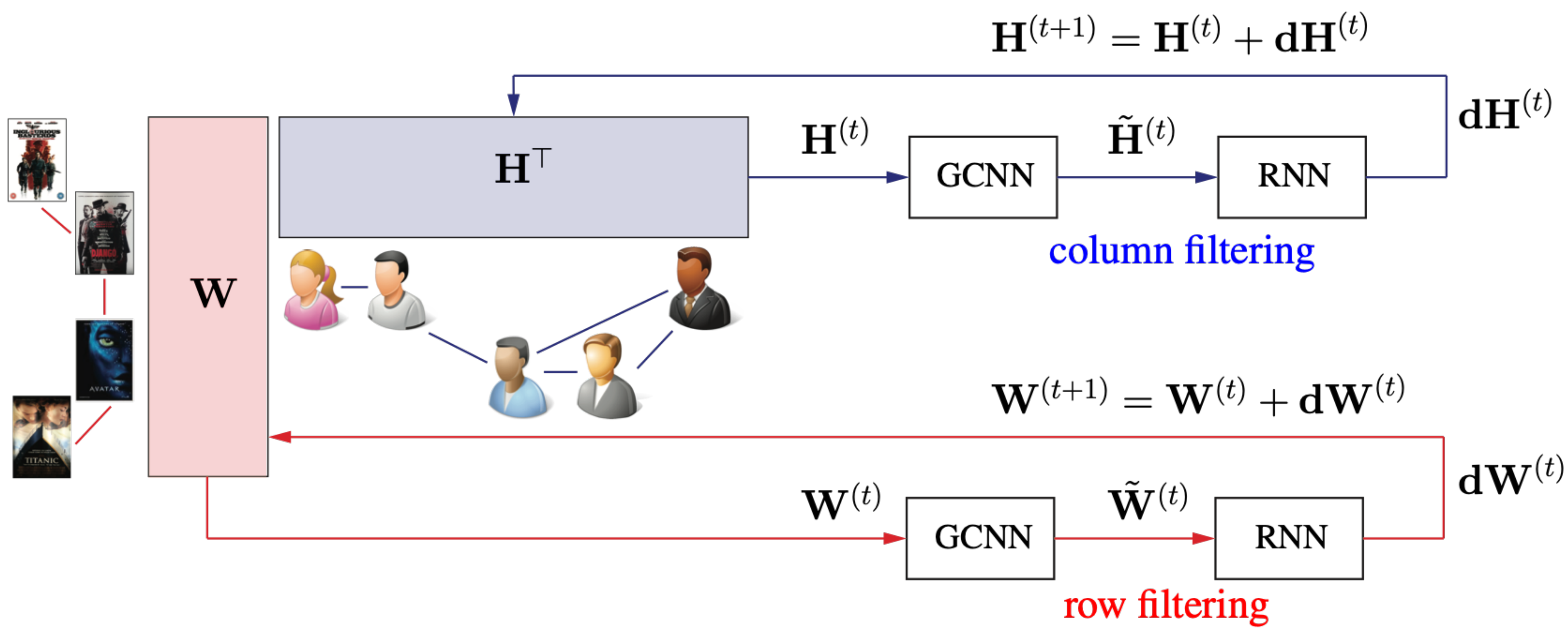
- Mean reward as smooth signal on graph $\mathbb{E}\{y_n\} = \langle \boldsymbol{\chi}_n, \boldsymbol{\alpha} \rangle$
- Recommendation problem as selection of best (with largest reward) graph node

$$n_t = \arg\max_{n} \left[\langle \boldsymbol{q}_n, \hat{\alpha} \rangle \right] + c_t \| \boldsymbol{q}_n \|_{V_t^{-1}} \quad \text{with} \quad V_t^{-1} = \boldsymbol{Q}_t \boldsymbol{Q}_t^T + (\boldsymbol{\Lambda} + \gamma \boldsymbol{I})$$

$$L = \underbrace{\begin{bmatrix} | & & | \\ \boxed{\quad \boldsymbol{q}_n \quad} & & \\ \chi_0 & \cdots & \chi_{N\text{-}1} \\ | & & | \end{bmatrix}}_{\chi} \underbrace{\begin{bmatrix} \lambda_0 & & 0 \\ & \ddots & \\ 0 & & \lambda_{N-1} \end{bmatrix}}_{\boldsymbol{\Lambda}} \underbrace{\begin{bmatrix} \text{---} & \chi_0 & \text{---} \\ & \cdots & \\ \text{---} & \chi_{N\text{-}1} & \text{---} \end{bmatrix}}_{\boldsymbol{\chi}^T}$$

M Valko et al. "Spectral bandits for smooth graph functions", ICML 2014.

Recommender systems: Matrix Completion



- Matrix completion: diffusion process as RNN casted on top of multi-graph convolutional layers
- Multi-graph convolution (spatial features), followed by LSTM (diffusion process)

F. Monti et al., "Deep geometric matrix completion: A new way for recommender systems", ICASSP 2018

Take Home Message

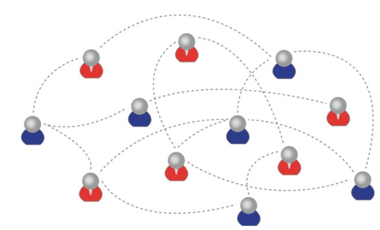
GSP Tools …

Graph Knowledge exploitation
Implicit / Explicit
Given / Constructed

- smooth- multi-resolution graph signal representation
- graph denoising
- graph sampling
- graph filter and kernel

… for ML

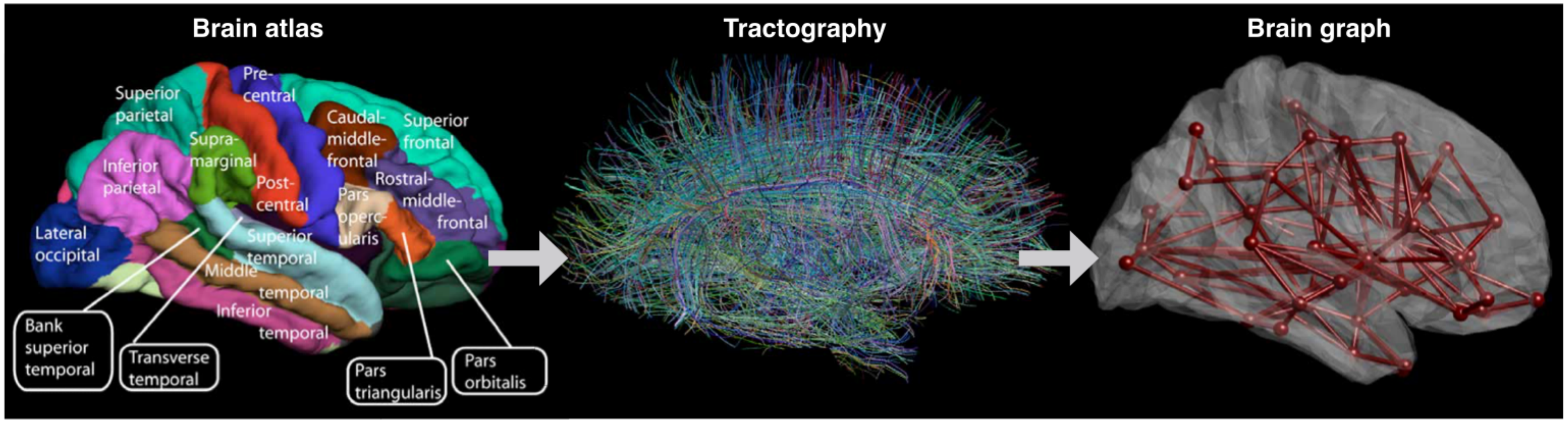
- Node/graph classification
- Community detection
- Time series (system dynamics) inference
- Online learning



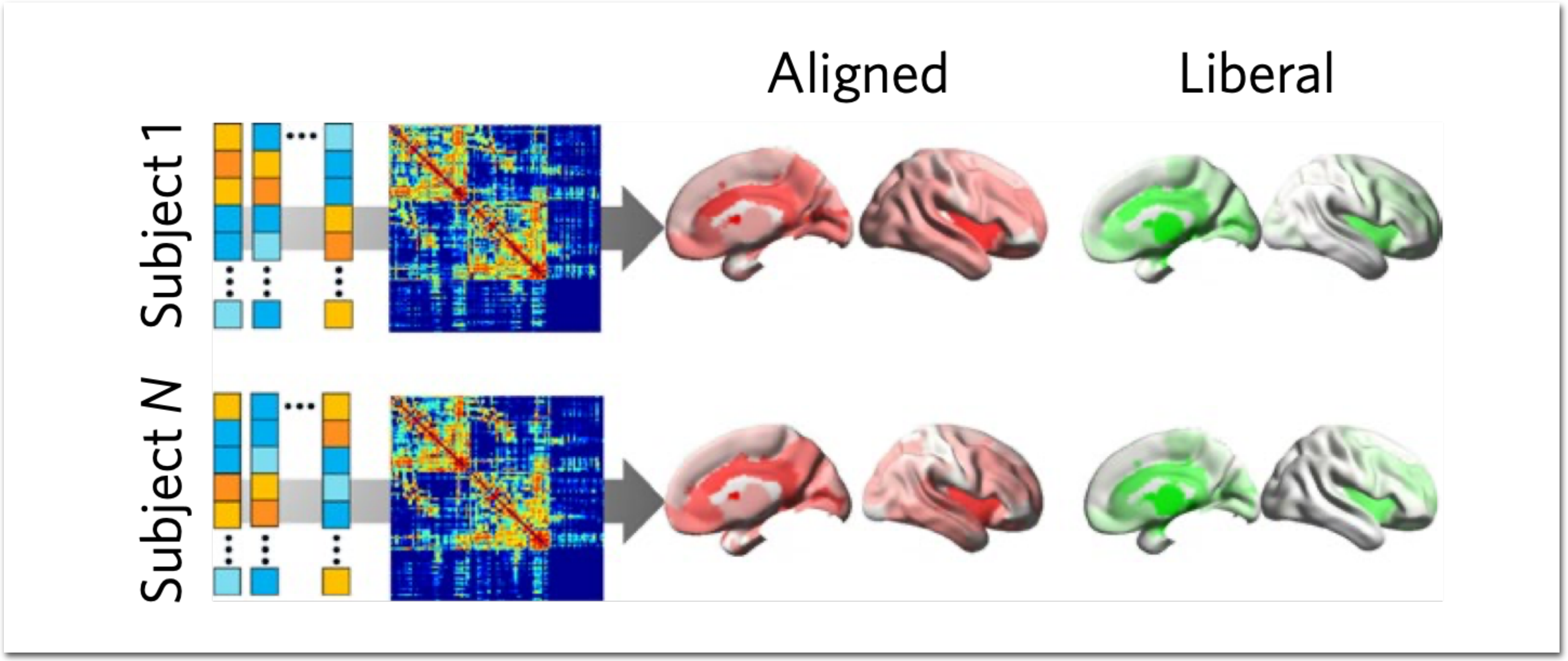
GSP/ GNN

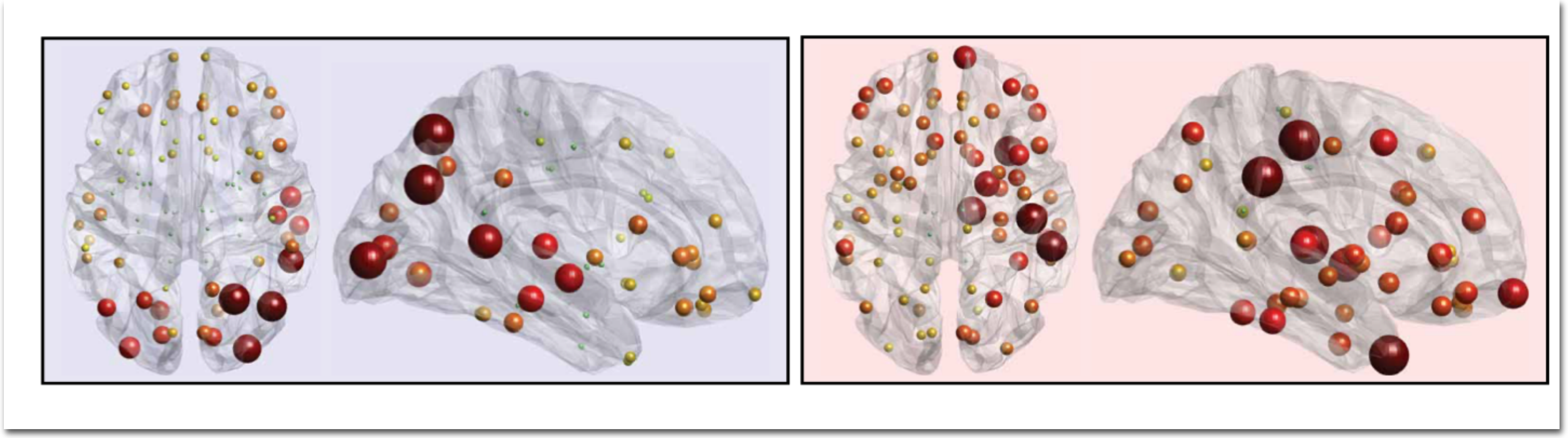
In

Healthcare and Neuroscience



Brain Analysis: GSP for Interpretability

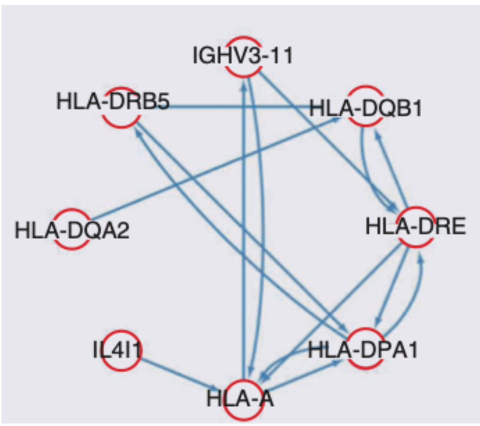




J.D. Medaglia et al., “Functional alignment with anatomical networks is associated with cognitive flexibility”, *Nat Hum Behav, 2018*
W. Huang, et al., "A Graph Signal Processing Perspective on Functional Brain Imaging”, Proceedings of the IEEE, 2018.

Gene Expression: Topology inference

To identifying gene-regulatory topologies, where nodes represent individual genes and directed edges encode causal regulatory relationships between gene pairs



Structural equation models (SEMs)

$$\boldsymbol{x}_t = \boldsymbol{A}\boxed{\boldsymbol{x}_t} + \boldsymbol{\Omega}\boxed{\boldsymbol{u}_t} + \epsilon_t$$

$$\boldsymbol{H} = (\boldsymbol{I} - \boldsymbol{A})^{-1}$$

$$\boldsymbol{\Omega} = \boldsymbol{I}$$

exogenous inputs: genotypes of the expression quantitative trait loci (eQTLs)

endogenous variables: gene-expression levels

$\boldsymbol{x}_t = \boldsymbol{H}\boldsymbol{u}_t$ Graph filtering modelling for auto-regressing SEM

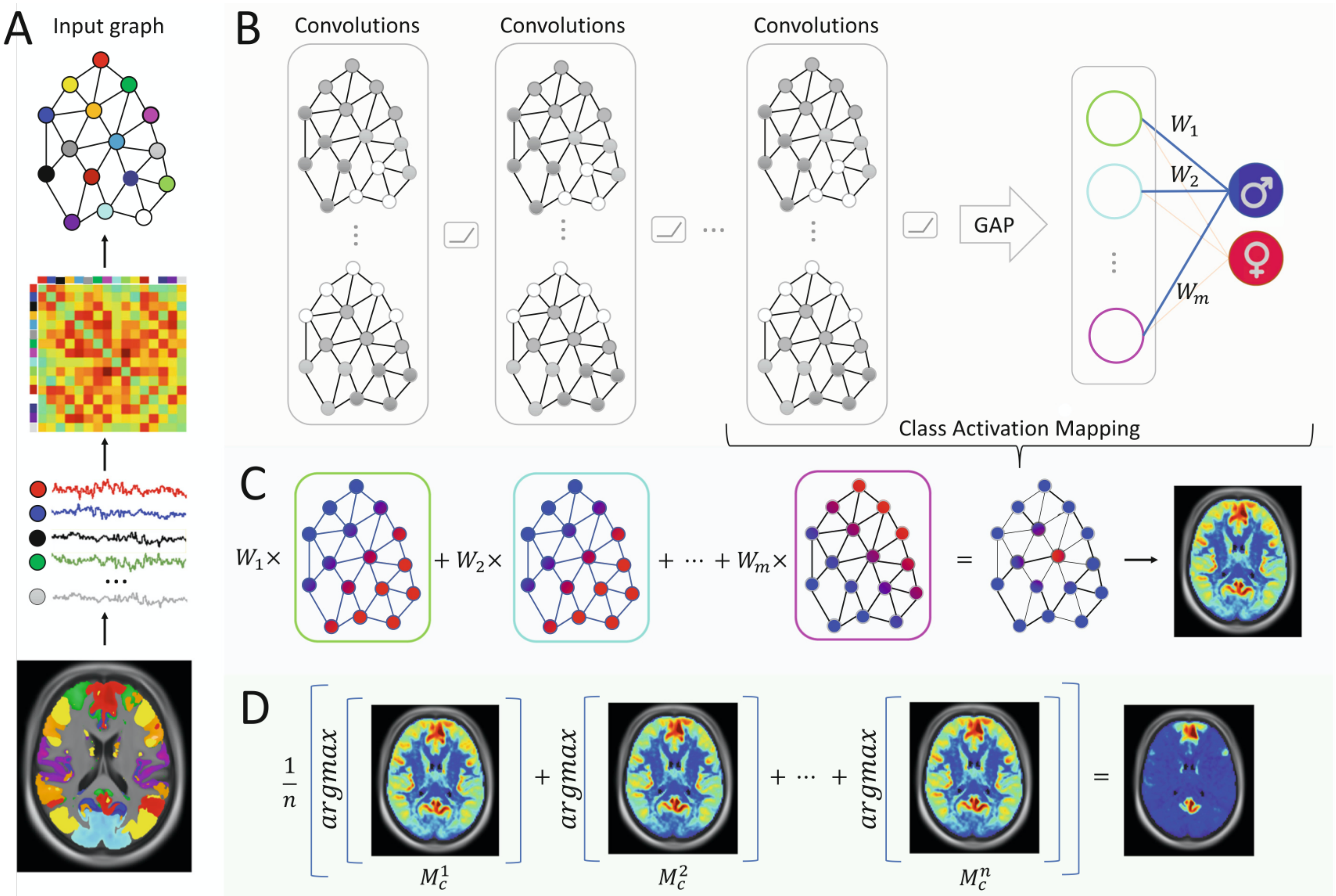
Kernel-based topology inference

$$\hat{\mathbf{A}} = \underset{\mathbf{S},\,\omega}{\operatorname{argmin}} \sum_{t=1}^{T} \left\| \mathbf{x}_t - \mathbf{A}\mathbf{x}_t + \mathbf{\Omega}\mathbf{u}_t \right\|_2^2 + \alpha \left\| \mathbf{A} \right\|_1,$$

$$\text{s.t. } \mathbf{\Omega} = \operatorname{diag}(\omega), \quad S_{ii} = 0, i = 1, \ldots, N,$$

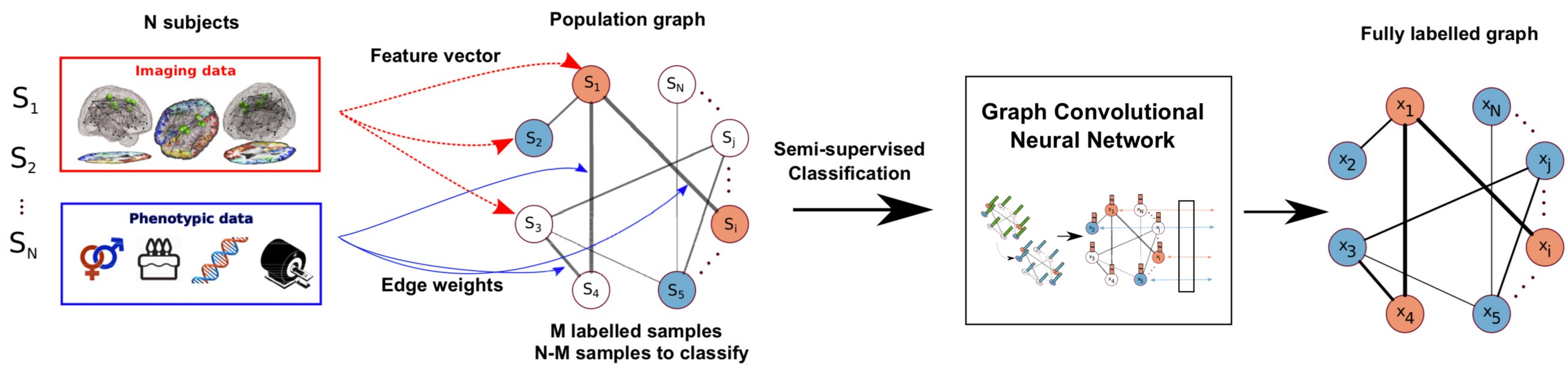
A. G. Marques, et al. "Signal processing on directed graphs", IEEE SPM, 2020.

Brain imaging: Class activation mapping



S. Arslan et al., "Graph saliency maps through spectral convolutional networks: Application to sex classification with brain connectivity." *Graphs in Biomedical Image Analysis and Integrating Medical Imaging and Non-Imaging Modalities, 2018.*

Disease prediction

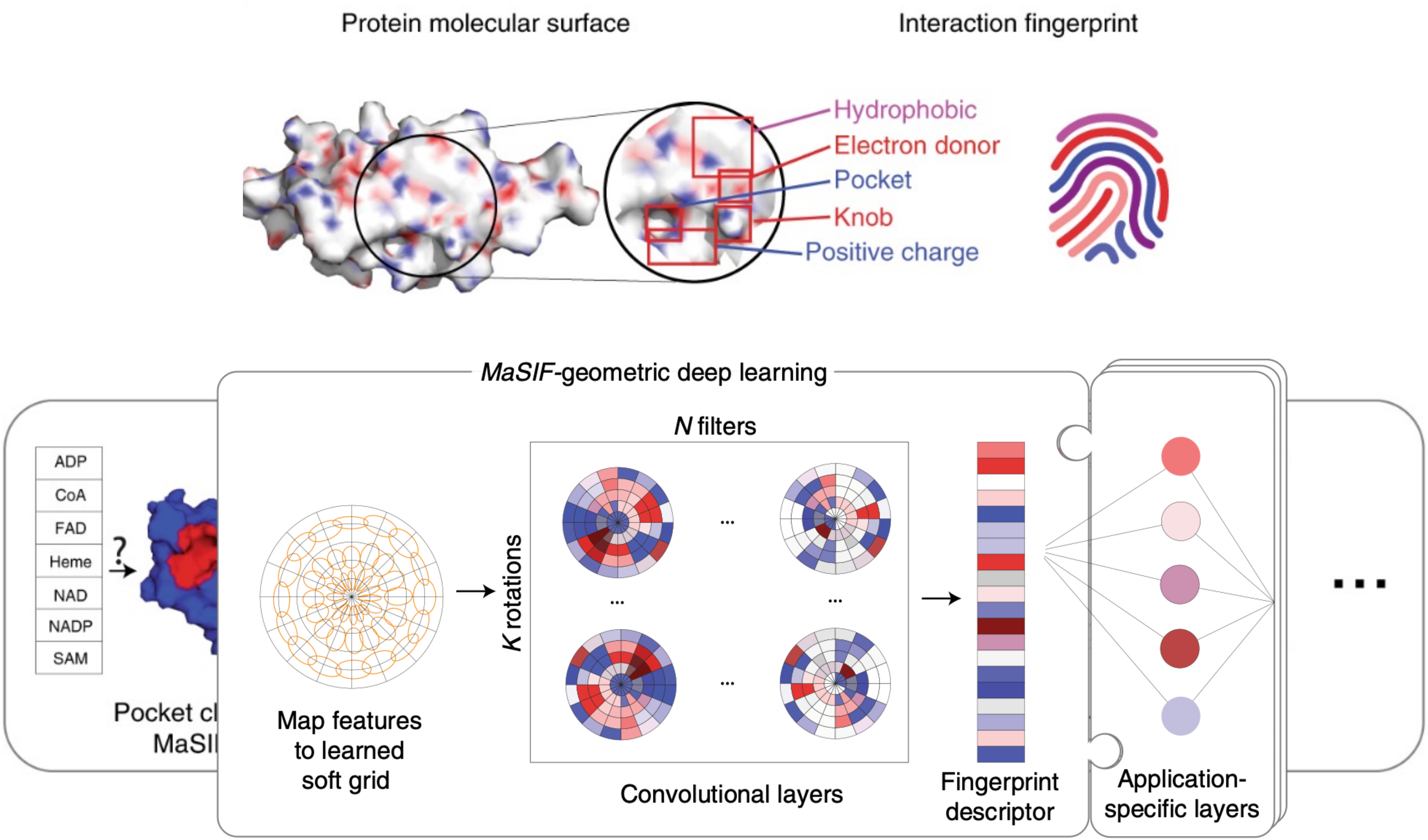


- populations modelled as a sparse graph
- phenotypic information integrated as edge weights
- imaging-based feature vectors as node signal

Subject classification as a graph labelling problem, integrating imaging and non imaging data.

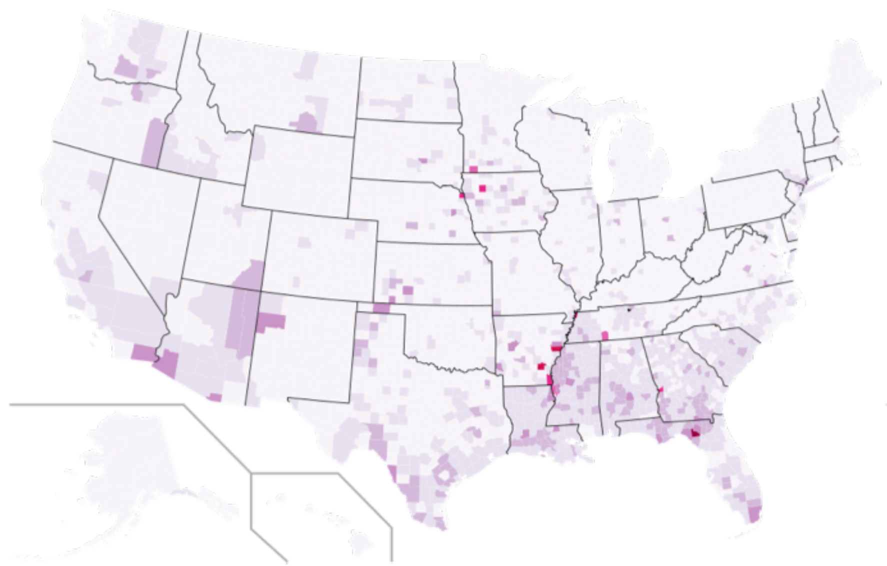
S. Parisot, “Disease prediction using graph convolutional networks”, Medical image analysis, 2018.

Protein-protein interactions

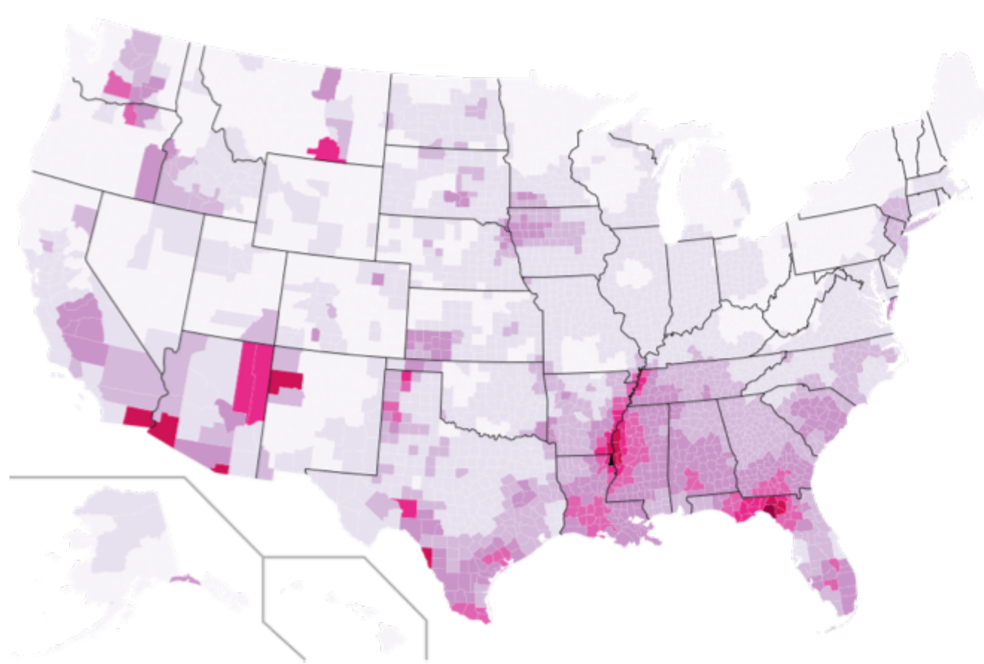


P. Gainza et al. “Deciphering interaction fingerprints from protein molecular surfaces using geometric deep learning”, Nature Methods, 2020.

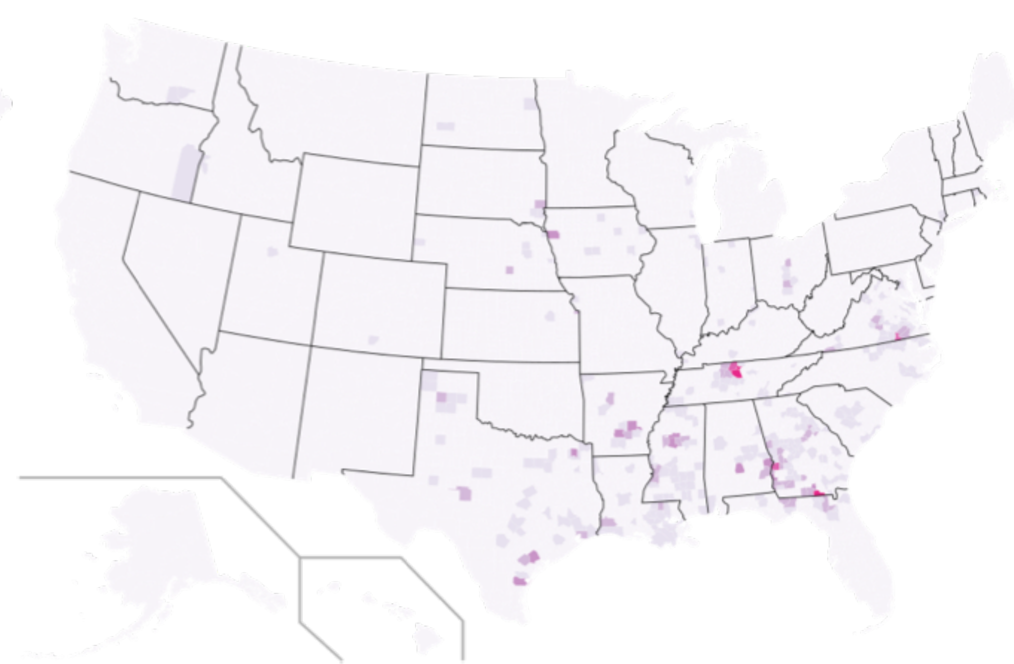
GSP for COVID-19



Cumulative number of confirmed COVID-19 cases per 100k residents for each county by Aug 31



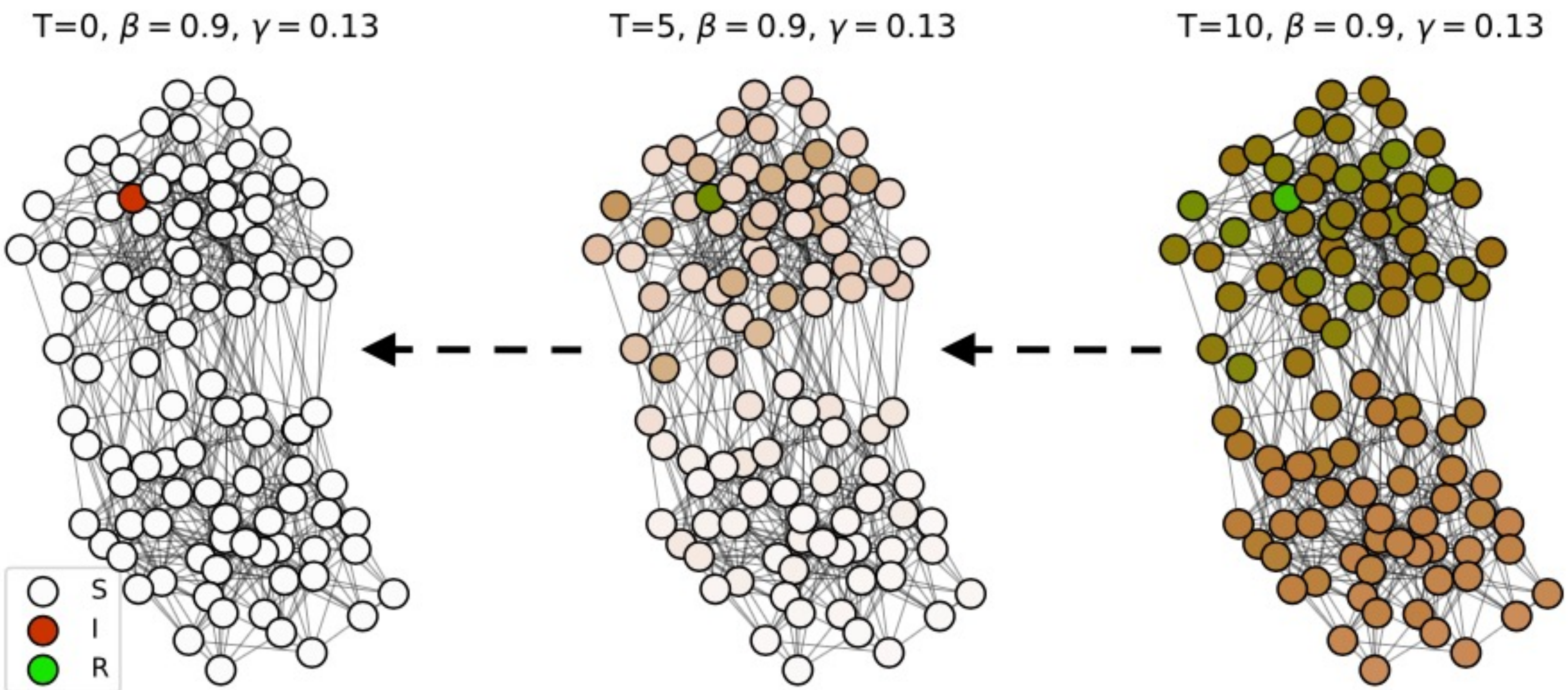
Low-pass signals of each county



High-pass signals of each county

Y. Li and G. Mateos, "Graph Frequency Analysis of COVID-19 Incidence to Identify County-Level Contagion Patterns in the United States", ICASSP 2021

GNN for COVID-19



Exploit GNNs to locate the source of the epidemics.

- GNNs are model-agnostic
- GNNs identify P0 close to theoretical bounds accuracy

Shah, Chintan, et al. "Finding Patient Zero: Learning Contagion Source with Graph Neural Networks." *arXiv, 2020*

Take Home Message

GSP Tools …

Graph Knowledge exploitation
Implicit / Explicit
Given / Constructed

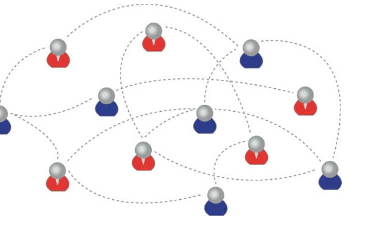
- smooth- multi-resolution graph signal representation
- graph denoising
- graph sampling
- graph filter and kernel
- graph convolution / graph clustering

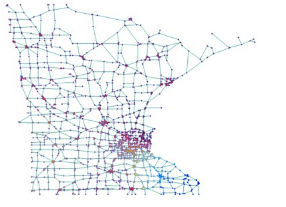
Unknown graph knowledge
Implicit

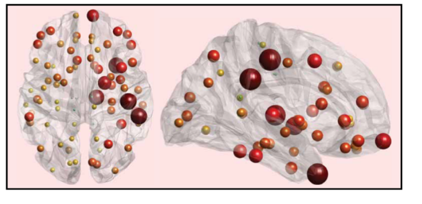
- topological inference

… for ML

- Node/graph classification
- Community detection
- Time series (system dynamics) inference
- Online learning





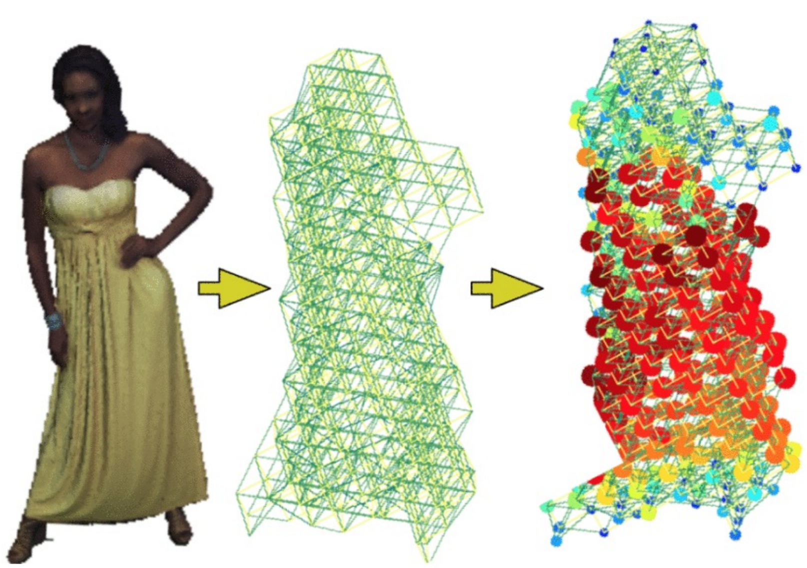


- Interpretability
- Complex dependencies
- Local (high-frequency) activation mapping
- Model-agnostic

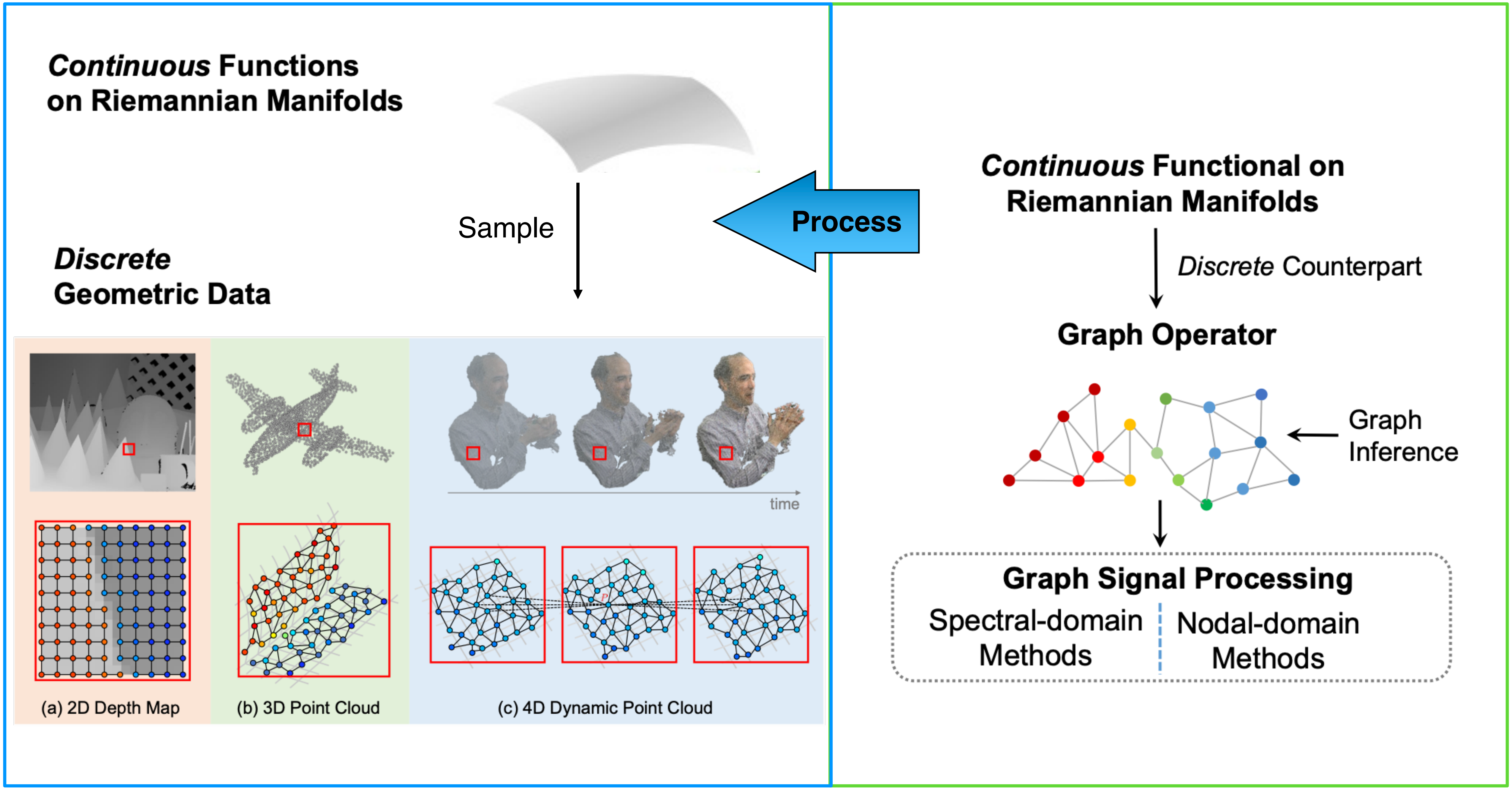
GSP-based ML

in

Computer Vision

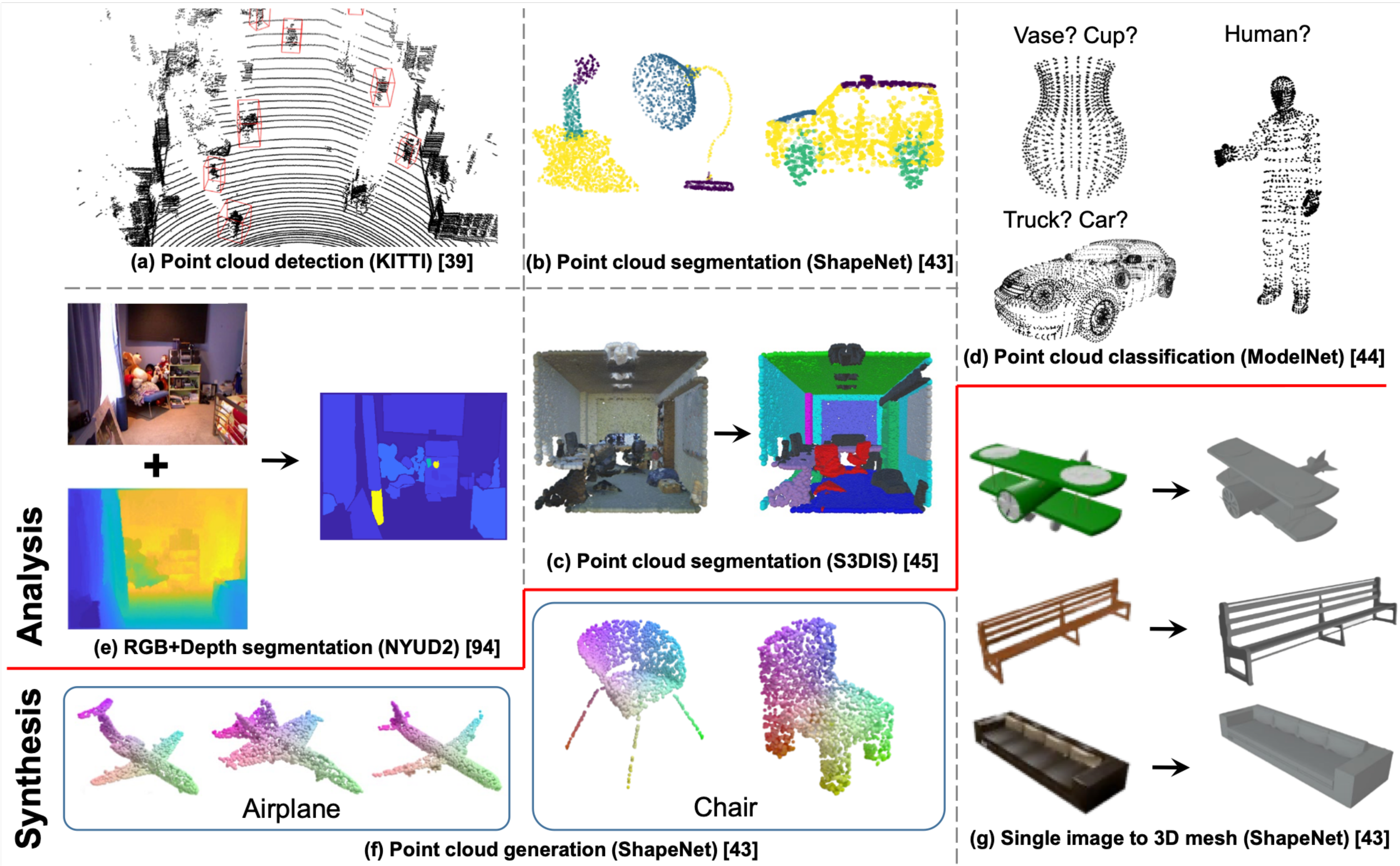


GSP for Geometric Data



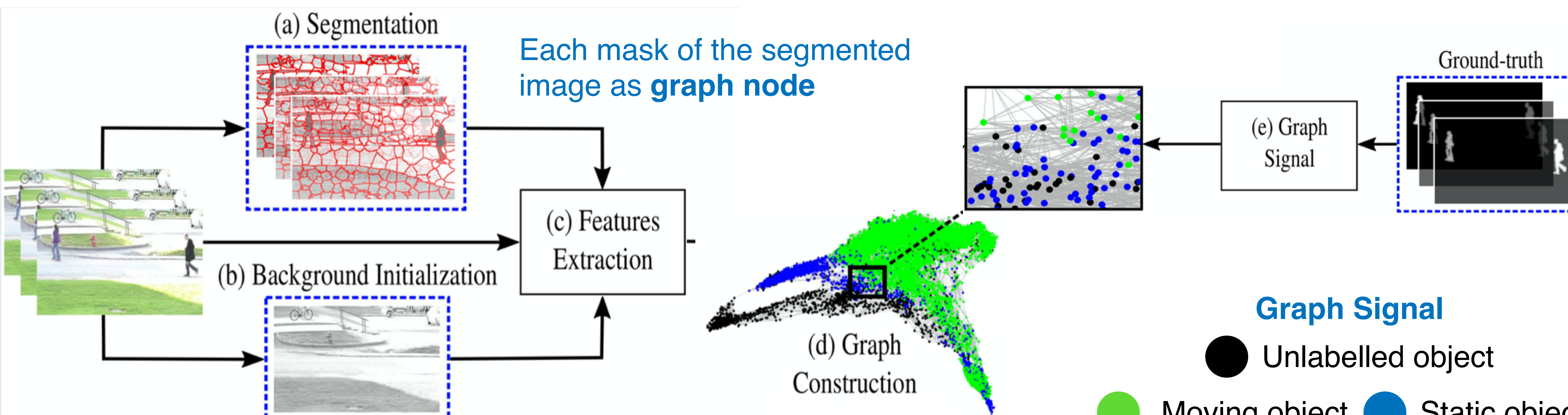
Hu, Wei, et al. "Graph Signal Processing for Geometric Data and Beyond: Theory and Applications", arXiv 2020

Analysis and Synthesis Tasks via GNNs



Hu, Wei, et al. "Graph Signal Processing for Geometric Data and Beyond: Theory and Applications", arXiv 2020
Y. Guo, et al."Deep learning for 3d point clouds: A survey", IEEE TPAMI, 2020.

Moving Object Segmentation



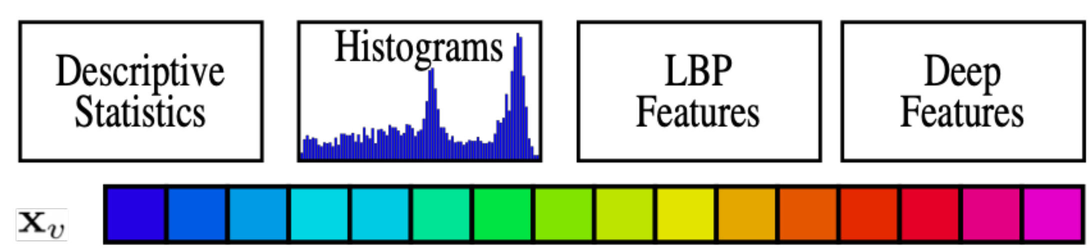
Each mask of the segmented image as **graph node**

Graph Signal

Unlabelled object

Moving object

Static object

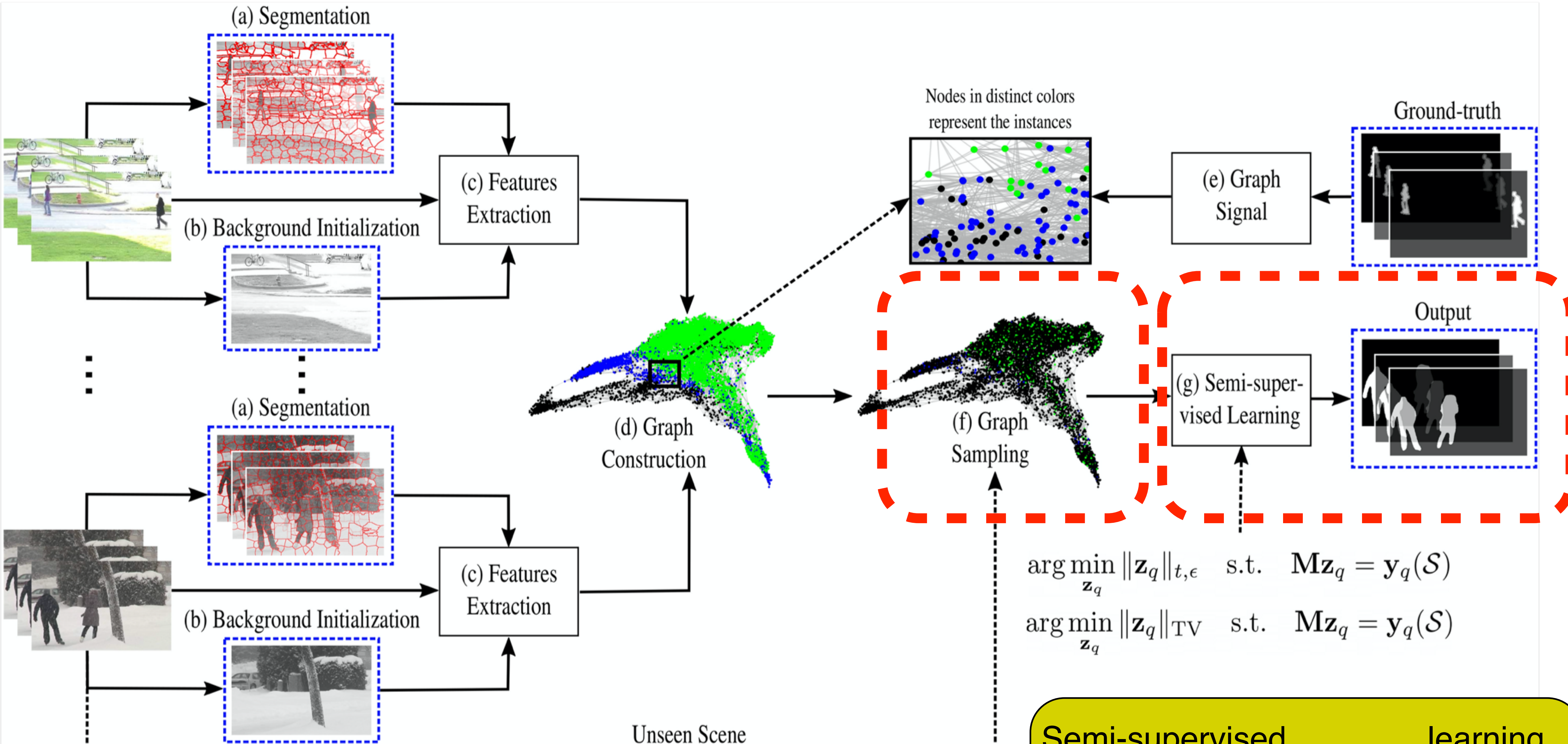


Graph weights $w_{i,j} = \exp||\boldsymbol{x}_i - \boldsymbol{x}_j||_2^2/\sigma$

Mask feature $\boldsymbol{x}_i$: concatenation of optical, flow, intensity, texture, and deep features

J. Giraldo et al. “Graph Moving Object Segmentation”, IEEE TPAMI, 2020

Moving Object Segmentation



$$\arg\min_{\mathbf{z}_q} \|\mathbf{z}_q\|_{t,\epsilon} \quad \text{s.t.} \quad \mathbf{M}\mathbf{z}_q = \mathbf{y}_q(\mathcal{S})$$

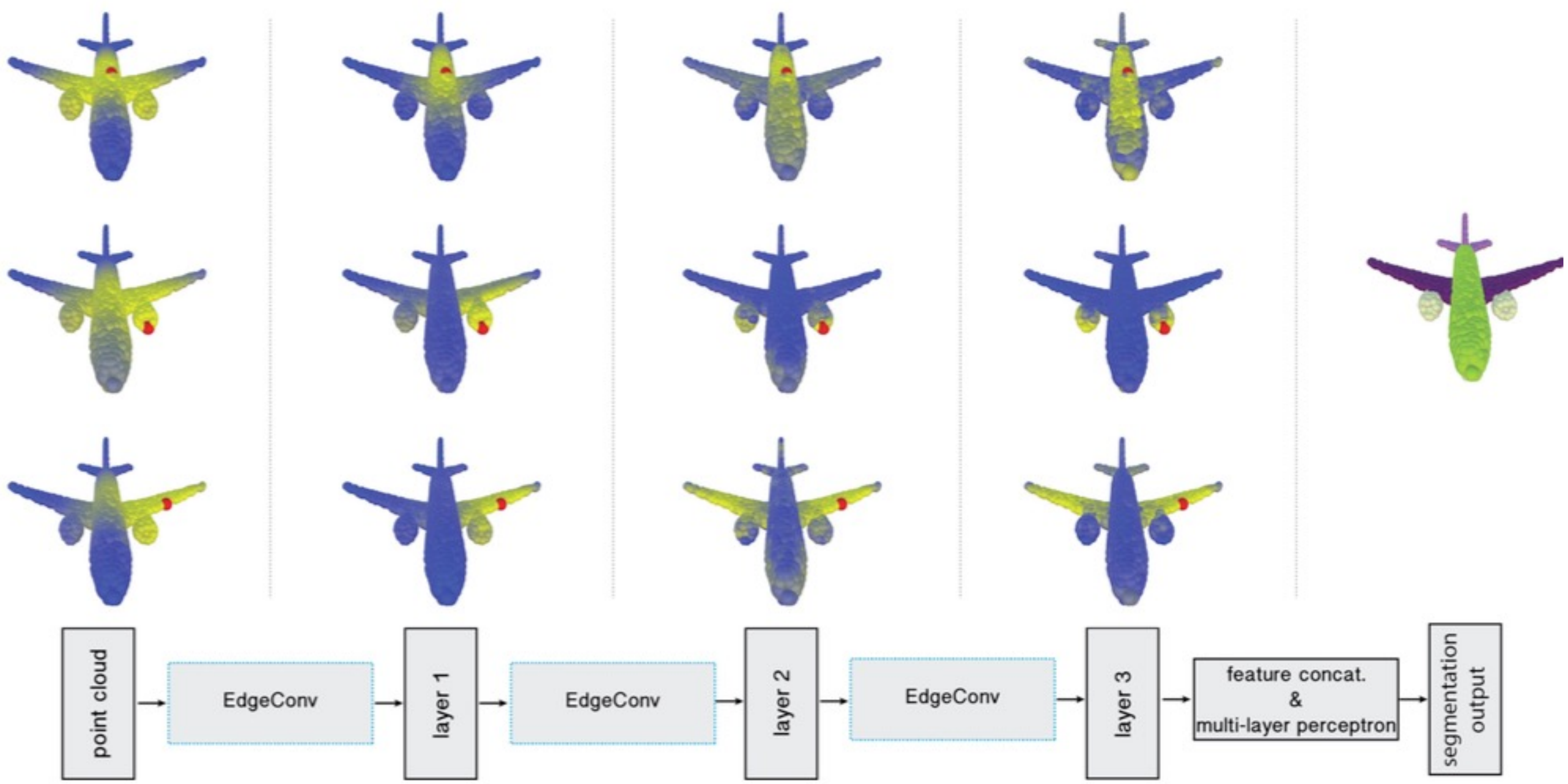
$$\arg\min_{\mathbf{z}_q} \|\mathbf{z}_q\|_{\mathrm{TV}} \quad \text{s.t.} \quad \mathbf{M}\mathbf{z}_q = \mathbf{y}_q(\mathcal{S})$$

Semi-supervised learning (classification) as graph signal learning with smooth regularizer

Graph sampling: for band limited signals, at least ρ (bandwidth) labeled nodes are needed to achieve perfect classification

J. Giraldo et al. “Graph Moving Object Segmentation”, IEEE TPAMI, 2020

3D Point Cloud Segmentation

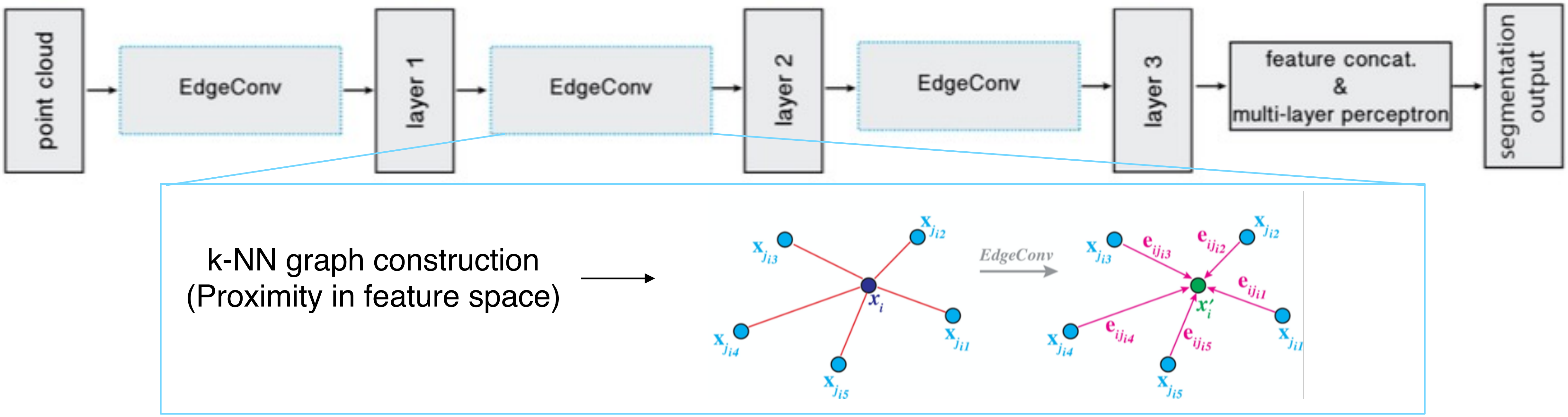


Graph convolution-like operators on the edges connecting neighboring pairs of points, in the spirit of graph neural networks

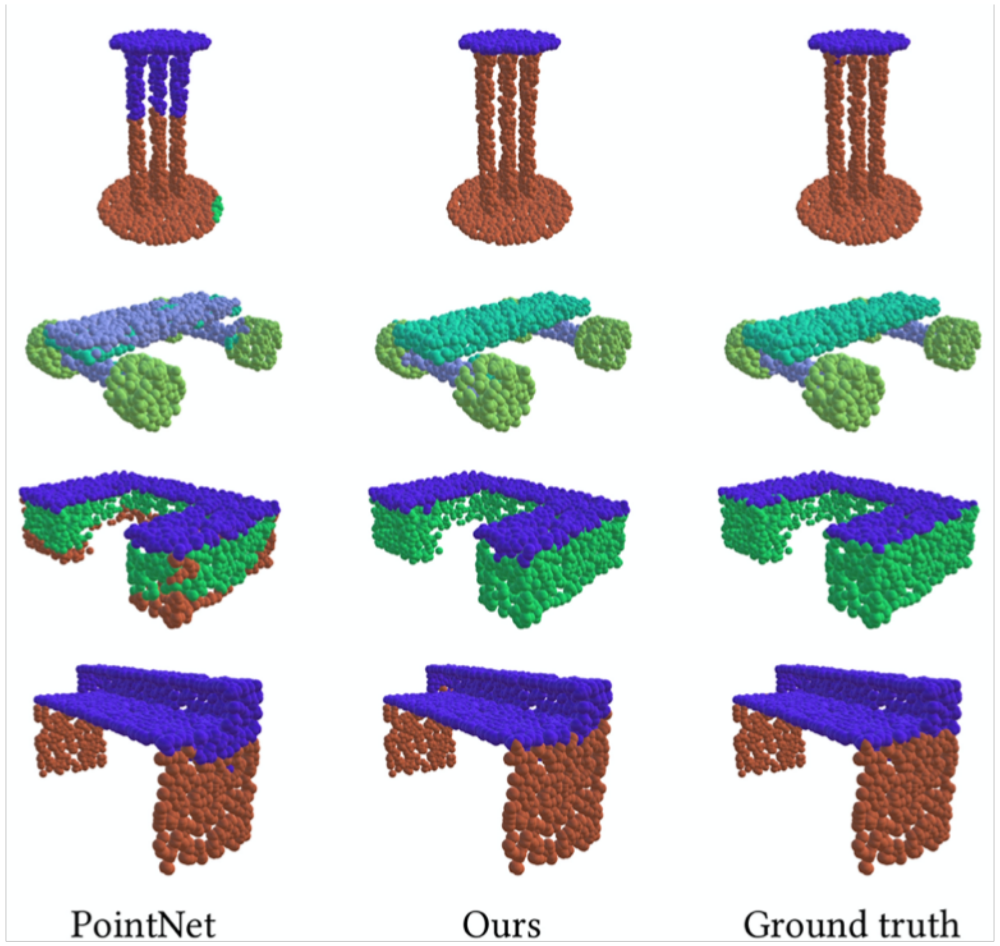
Feature space structure in deeper layers captures semantically similar structures

Y. Wang, et al. “Dynamic Graph CNN for Learning on Point Clouds”, ACM Trans. Graph, 2019.

3D Point Cloud Segmentation

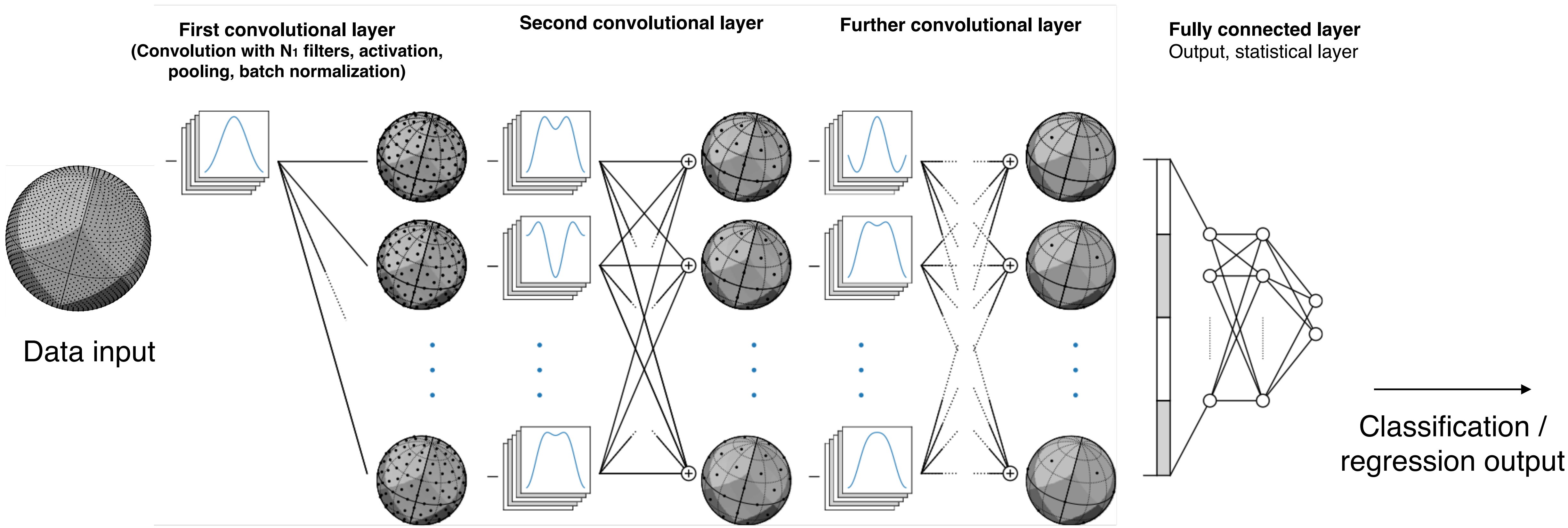


- Dynamic graph constructed at each layer based on features similarity
- Aggregation of edge features
- Translation-invariance and non-locality properties

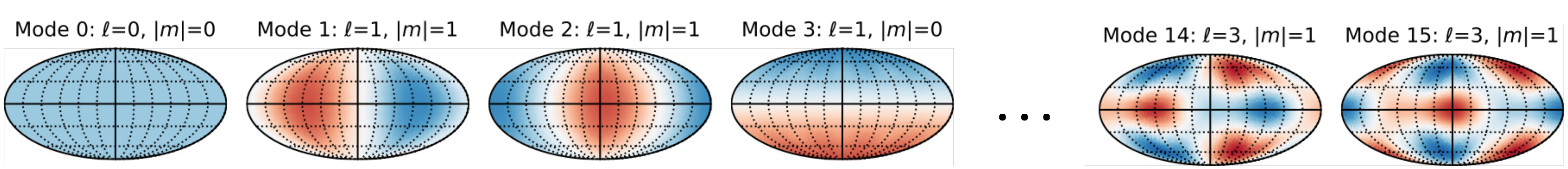


Y. Wang, et al. "Dynamic Graph CNN for Learning on Point Clouds", ACM Trans. Graph, 2019.

Cosmology: DeepSphere



- Sphere modelled with a graph and convolutions are performed on the graph
- Down-sampling operation (based on hierarchical pixelization of the sphere) to achieve multiple scales data analysis while preserving the spatial localization of features



Perraudin et al., “DeepSphere”, Astronomy and Computing, 2019

Take Home Message

GSP Tools …

Graph Knowledge exploitation
Implicit / Explicit
Given / Constructed

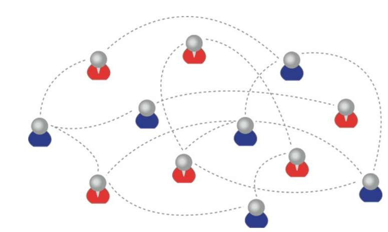
- smooth- multi-resolution graph signal representation
- graph denoising
- graph sampling
- graph filter and kernel
- graph convolution / graph clustering

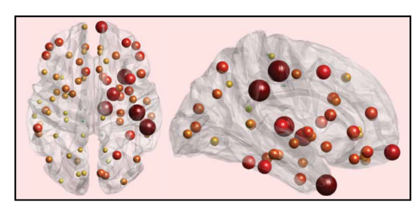
Unknown graph knowledge
Implicit

- topological inference

… for ML

- Node/graph classification
- Community detection
- Time series (system dynamics) inference
- Online learning





- Interpretability
- Complex dependencies
- Local (high-frequency) activation mapping
- Model-agnostic

- Translation invariance
- Non-locality properties
- Robustness to noise
- Sampling for computation efficiency

Outline

- Brief introduction to graph signal processing (GSP)
- Challenge I: GSP for exploiting data structure
- Challenge II: GSP for improving efficiency and robustness
- Challenge III: GSP for enhancing model interpretability
- Applications
- Summary, open challenges, and new perspectives

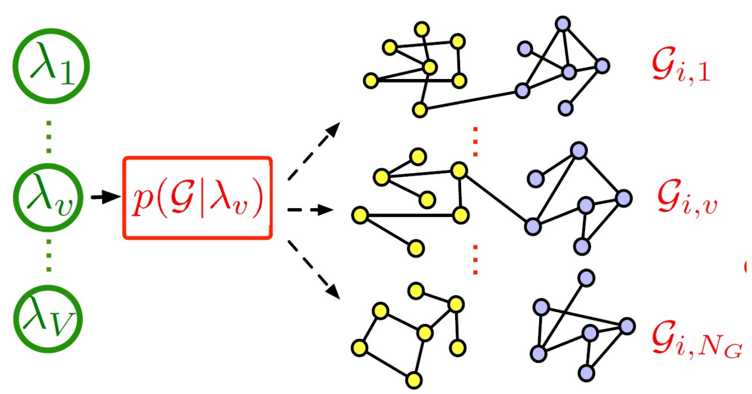
Open Challenges

GSP Tools

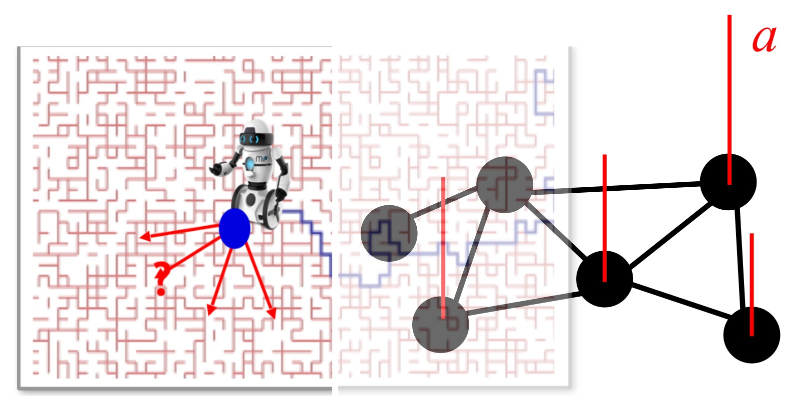
GSP-based ML

ML

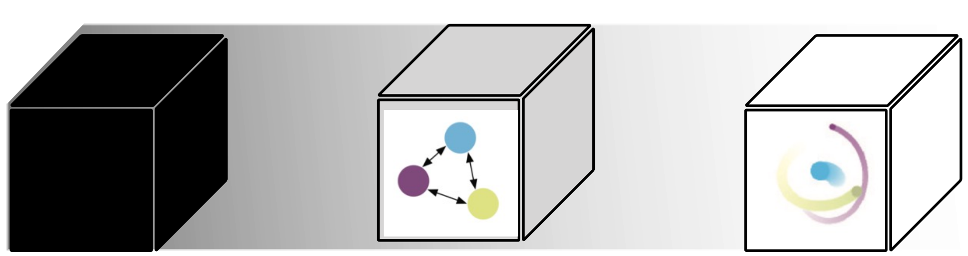
This tutorial



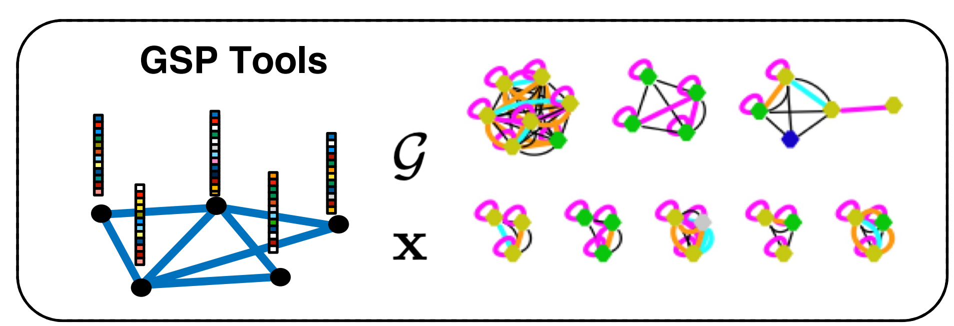
GSP and probabilistic models



GSP-based decision making strategies

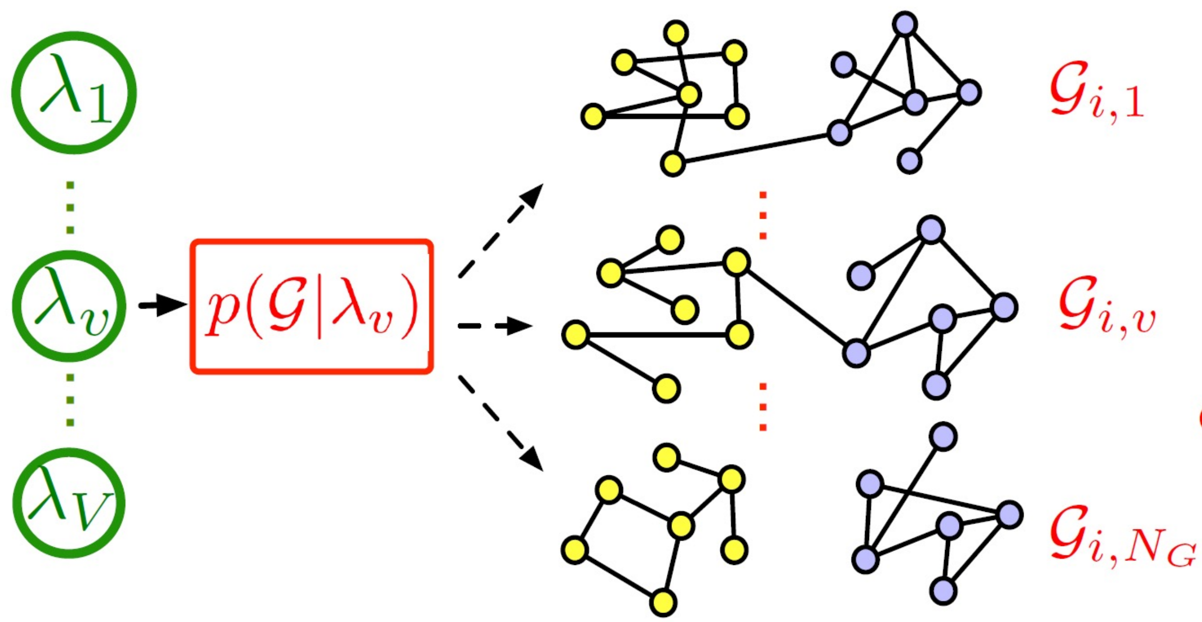


GSP and model interpretability

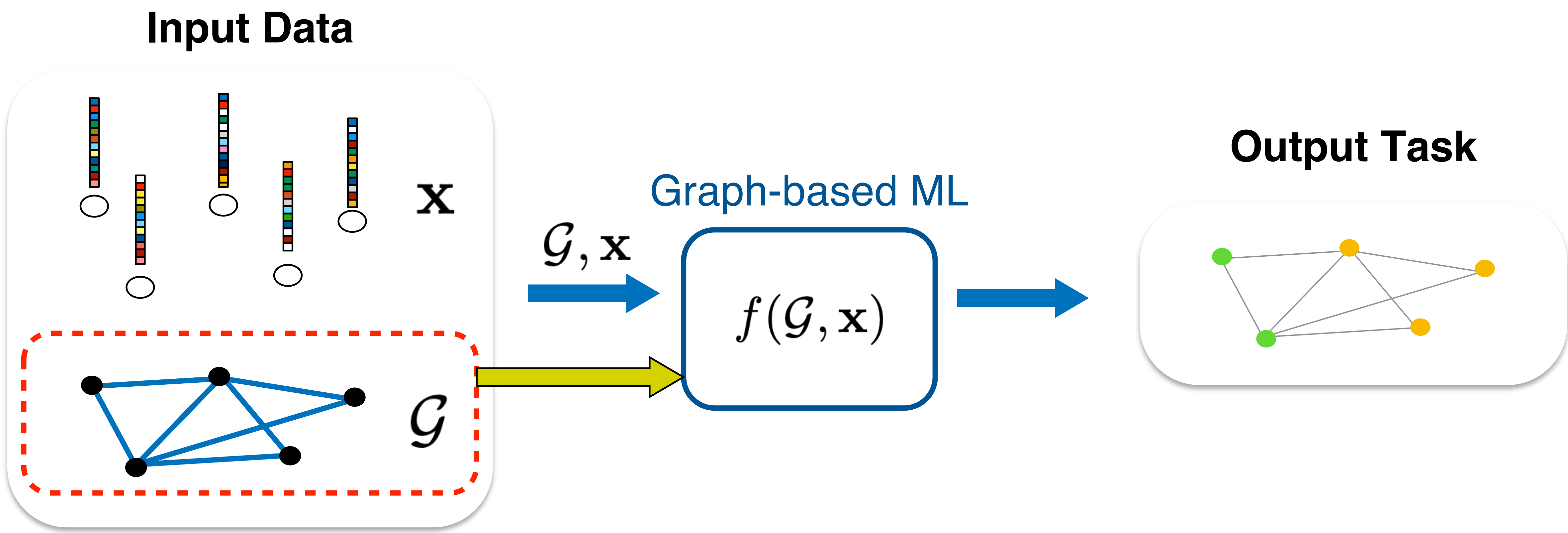


GSP and higher-order structure

GSP and Probabilistic Models



GSP and Probabilistic Models



- Most works assume graph is known a priori or fixed (deterministic setting)
- Real world networks are noisy and/or evolving over time

Challenge I: To take into account the topology uncertainty in graph-based machine learning tasks

Topological uncertainty in GSP: Modelling

Challenge I: To take into account topological uncertainty in graph-based SP and ML tasks

- How do we model topological noise?
  - Random graph model for topological noise [1]

ground truth adjacency matrix

$$y = x + n, n \sim \mathcal{N}(\mu, \sigma^2) \xrightarrow{?} \boldsymbol{W} = \boxed{\boldsymbol{A}} + \boxed{\boldsymbol{E}}$$

error matrix

  - What is the impact of the topological noise on filtering [1,2,3]?

[1] J. Miettinen, "Modelling Graph Errors: Towards Robust Graph Signal Processing", arXiv, 2020.

[2] E. Isufi, et al., "Filtering random graph processes over random time-varying graphs", IEEE TSP, 2017.

[3] E. Ceci, S. Barbarossa, "Graph Signal Processing in the Presence of Topology Uncertainties", IEEE TSP, 2020.

Topological uncertainty in GSP: Robustness

Challenge I: To take into account topological uncertainty in graph-based SP and ML tasks

- How do we build algorithms resilient to noisy topologies?
  - [3] proposes a robust signal recovery algorithm, under assumption of small perturbation, that incorporates statistical knowledge about topology uncertainty
  - [4] robustifies LMS with respect to mismatches in the presumed graph topology
  - [5] presents a robust formulation for graph-filter identification from input-output observations

[3] E. Ceci, S. Barbarossa, “Graph Signal Processing in the Presence of Topology Uncertainties”, IEEE TSP, 2020.

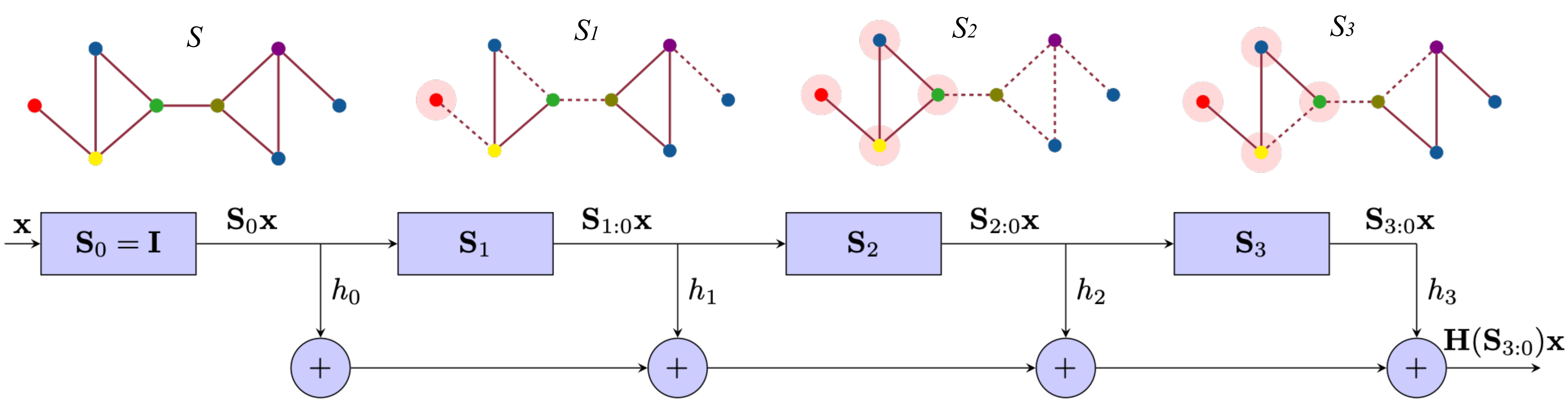
[4] J. Miettinen et al. "Robust Least Mean Squares Estimation of Graph Signals” ICASSP, 2019.

[5] S. Rey and A. G. Marques, “Robust graph-filter identification with graph denoising regularization,” ICASSP, 2021.

Topological uncertainty in graph ML

Challenge I: To take into account topological uncertainty in graph-based SP and ML tasks

- How do we incorporating uncertainty into learning algorithms?
  - [6] proposes a GNN architecture where the distributed graph convolution module accounts for the random network changes

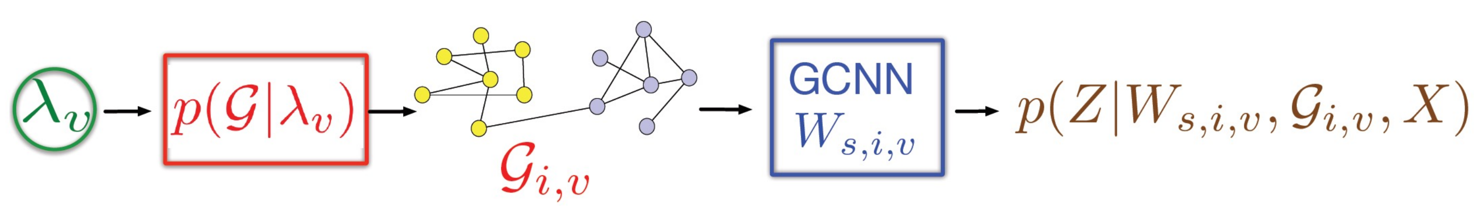


[6] Gao et al. “Stochastic Graph Neural Networks”, ICASSP and arXiv 2020.

Topological uncertainty in graph ML

Challenge I: To take into account topological uncertainty in graph-based SP and ML tasks

- How do we incorporating uncertainty into learning algorithms?
  - Bayesian approach: compute posterior associated with graph generative model so that new graph instances can be resampled [7,8]



$$p(\mathbf{Z}|\mathbf{Y}_{\mathcal{L}}, \mathbf{X}, \mathcal{G}_{obs}) = \int p(\mathbf{Z}|W, \mathcal{G}, \mathbf{X}) p(W|\mathbf{Y}_{\mathcal{L}}, \mathbf{X}, \mathcal{G}) p(\mathcal{G}|\lambda) p(\lambda|\mathcal{G}_{obs})\, dW\, d\mathcal{G}\, d\lambda\,,$$

$$\approx \frac{1}{V} \sum_{v=1}^{V} \frac{1}{N_G S} \sum_{i=1}^{N_G} \sum_{s=1}^{S} p(\mathbf{Z}|W_{s,i,v}, \mathcal{G}_{i,v}, \mathbf{X})\,.$$

figure from https://github.com/huawei-noah/BGCN

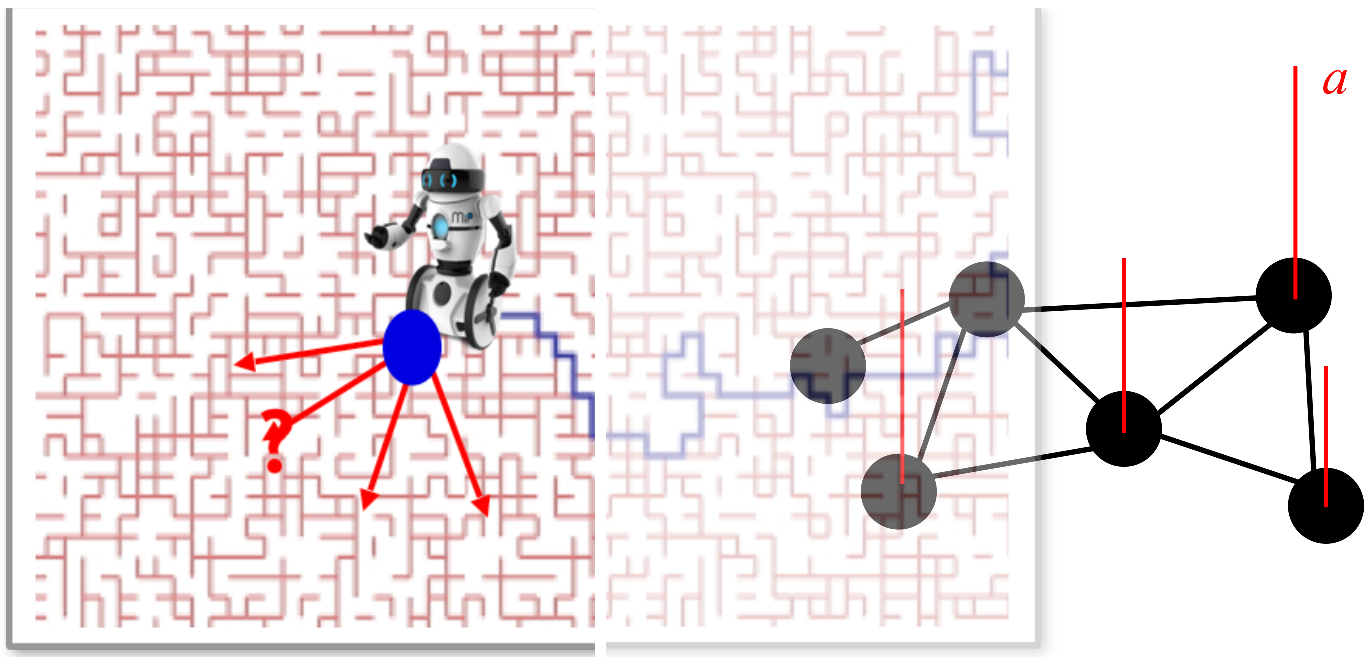
[7] Y. Zhang, et al. "Bayesian graph convolutional neural networks for semi-supervised classification", AAAI, 2019.
[8] Elinas et al. "Variational Inference for Graph Convolutional Networks in the Absence of Graph Data and Adversarial Settings", NeurIPS 2020.

Topological uncertainty in graph ML

Challenge I: To take into account topological uncertainty in graph-based SP and ML tasks

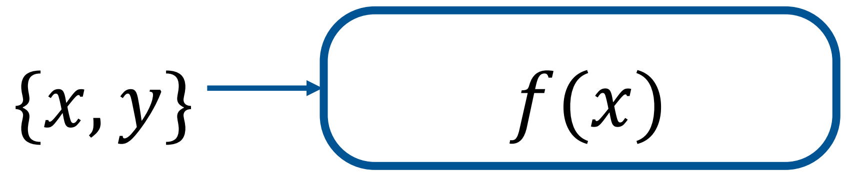
- How do we understand topological noise and its impact?
- How do we build algorithms resilient to noisy topologies?
- How do we incorporating uncertainty into learning algorithms?

GSP-Based Decision Making Strategies



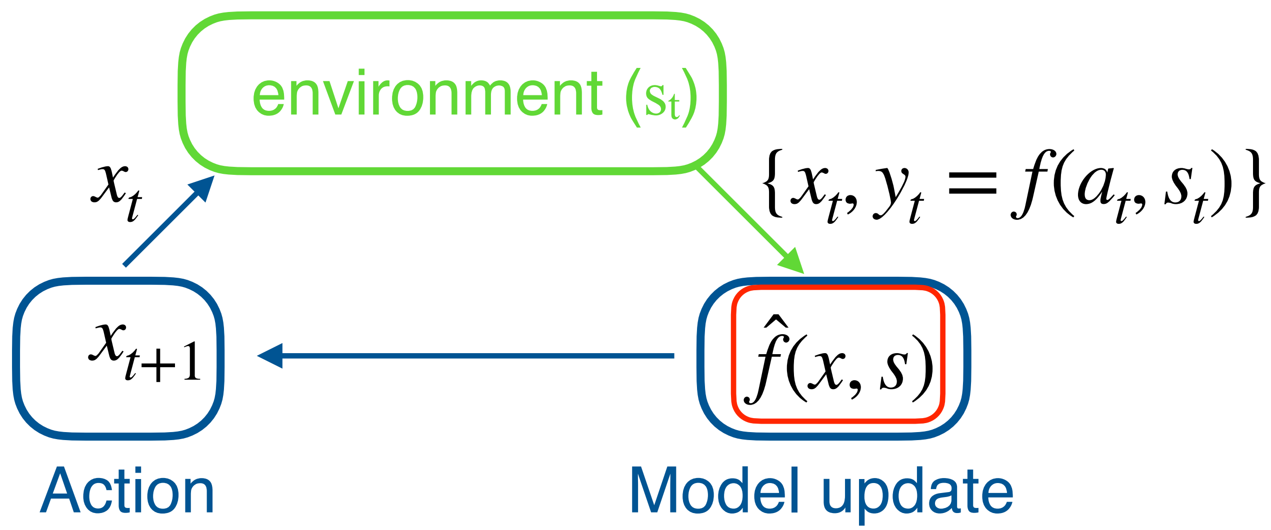
GSP-Based Decision Making Strategies

Supervised learning



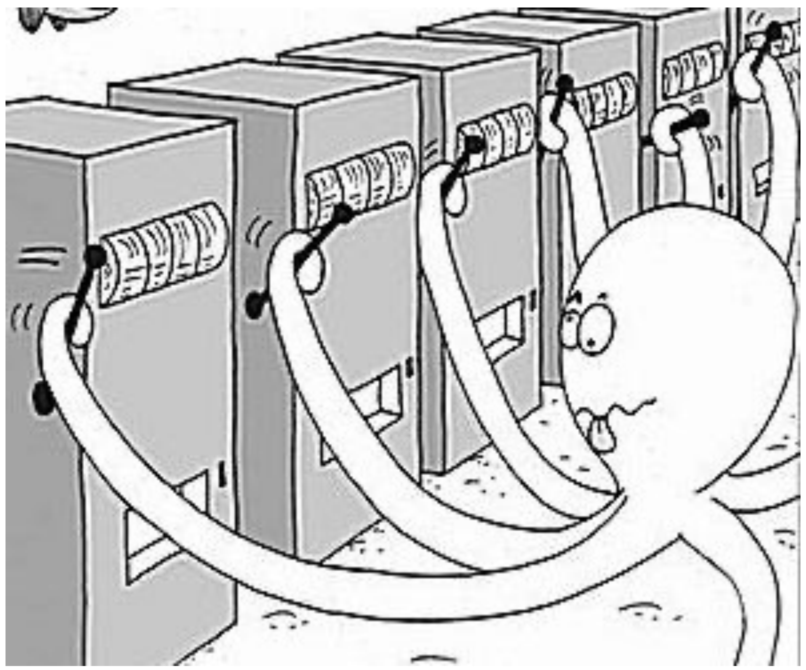
Decision Making Strategies (DMSs)

- Reinforcement learning
- Multi-arm bandit problems

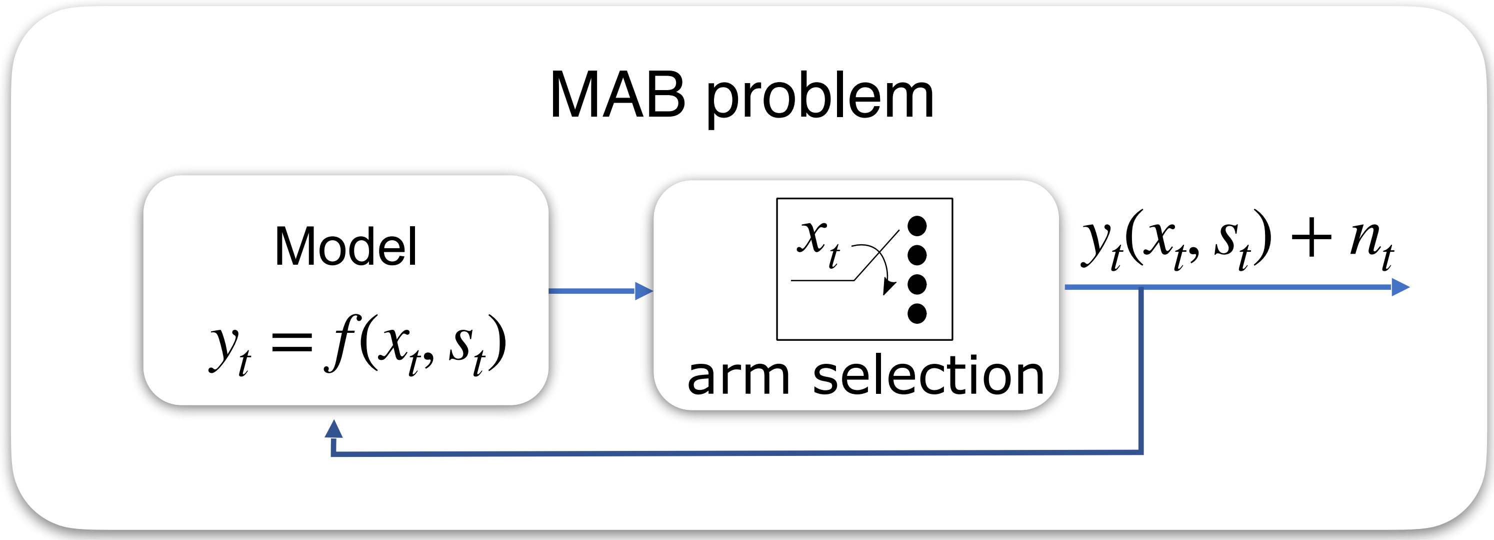


Optimize sequential actions in a way that maximizes the expected reward, when the environment's model is **uncertain a priori**

GSP for Multi-Arm Bandit



https://blogs.mathworks.com/images/loren/2016/multiarmedbandit.jpg

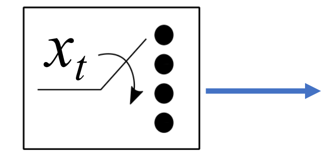


- x_t: selected action
- n_t: additive noise
- y_t: mean payoff
- s_t: context / user
- f: **unknown** model

High-dimensional search space ?

Optimize sequential actions in a way that maximizes the expected reward, when the environment's model is **uncertain a priori**

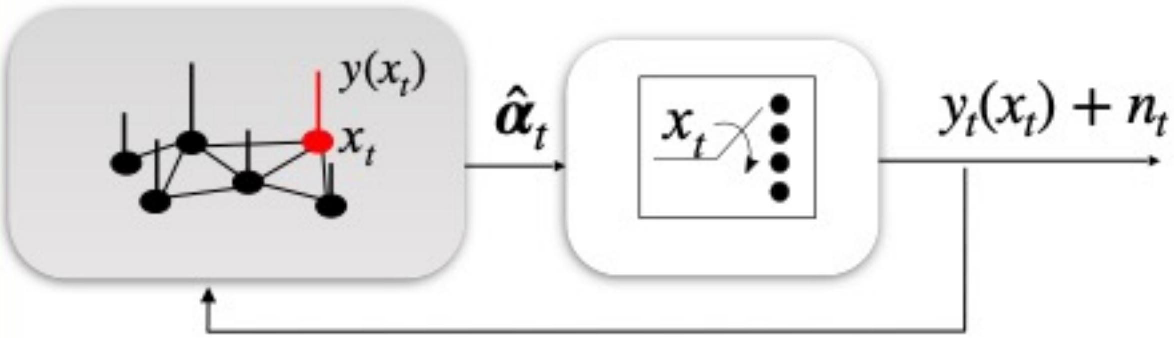
GSP for Multi-Arm Bandit



arm selection

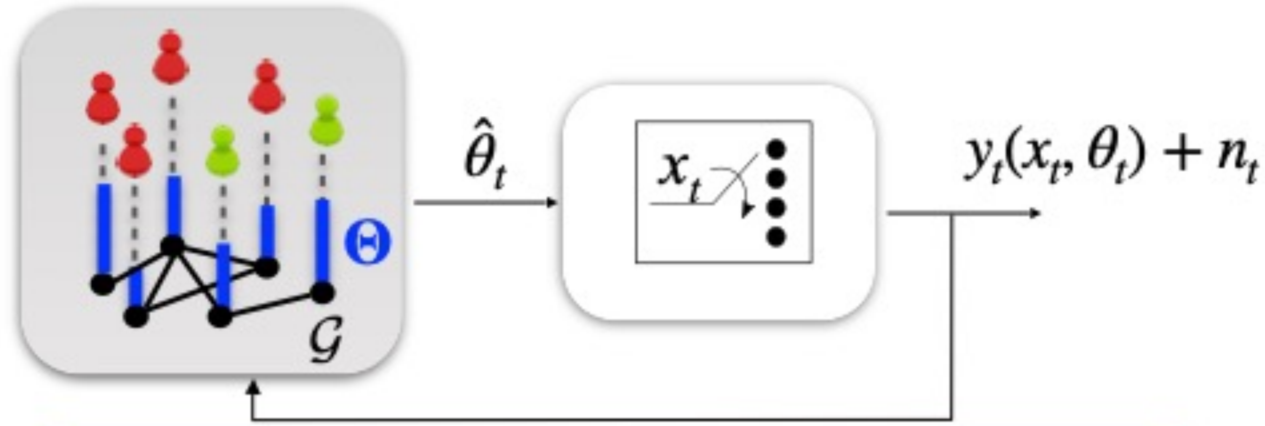
$y_t = f(x_t, s_t)$

Payoff as signal on G [1]



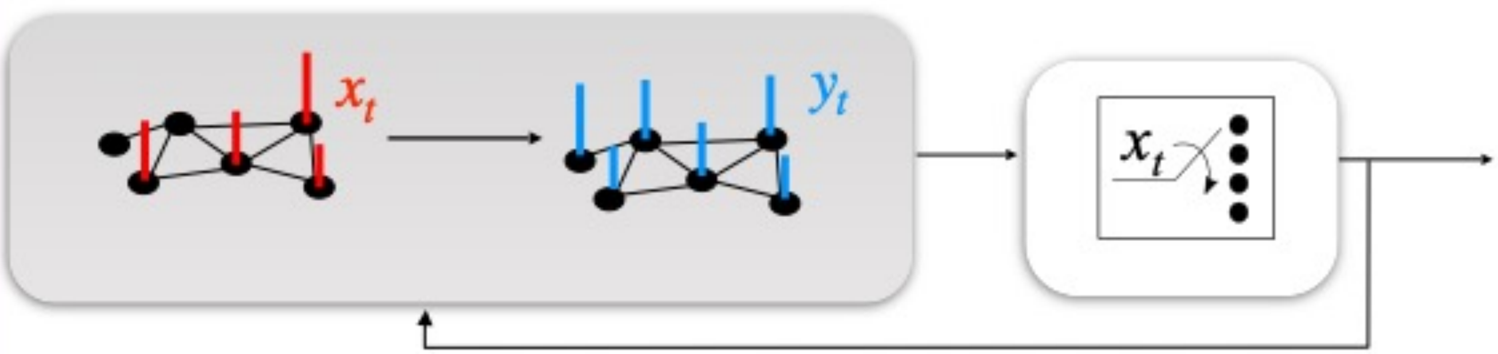
$$y = \langle \chi_G(x_t), \hat{\boldsymbol{\alpha}} \rangle$$

Context Features as signal on G [2]



$$\hat{\boldsymbol{\Theta}}_t = \arg \min_{\boldsymbol{\Theta} \in \mathbb{R}^{n \times d}} \sum_{i=1}^{n} \sum_{\tau \in t_i} (\mathbf{x}_\tau^T \boldsymbol{\theta}_i - y_{i,\tau})^2 + \alpha \; tr(\boldsymbol{\Theta}^T \mathscr{L} \boldsymbol{\Theta})$$

System Dynamics on G [3]

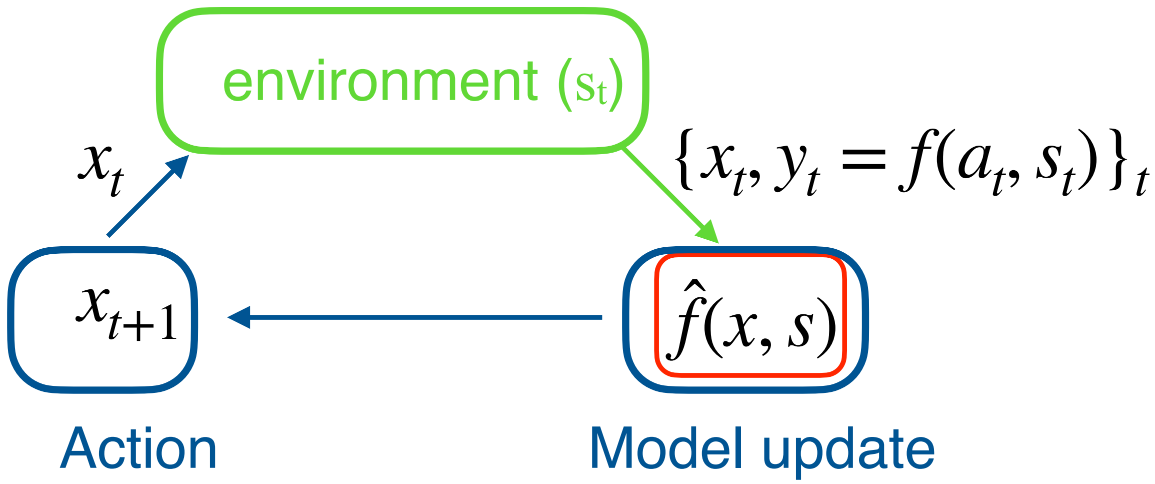


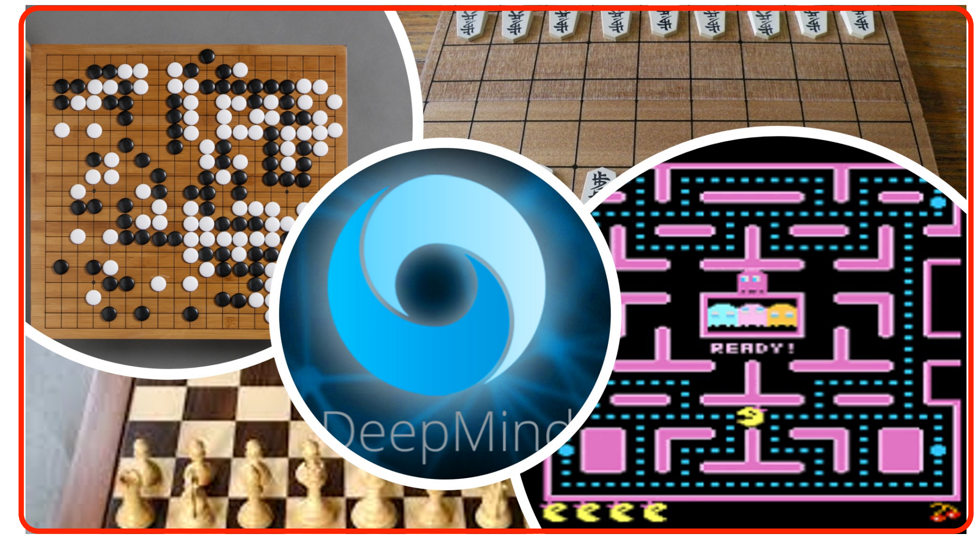
$y_t = g_L(x_t) + n_t,$

$g_L(\cdot)$ graph dictionary

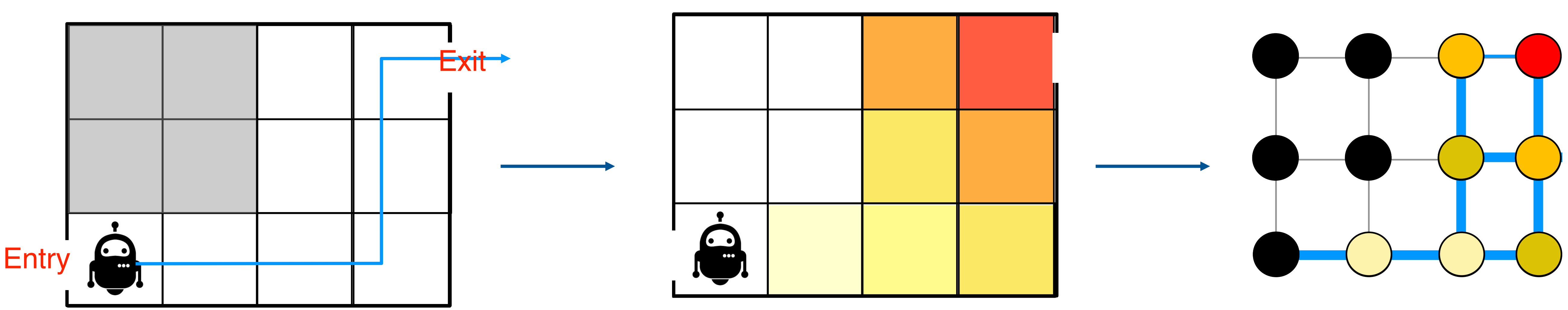
[1] M Valko et al. "Spectral bandits for smooth graph functions", ICML 2014.
[2] K. Yang, "Laplacian-regularized graph bandits: Algorithms and theoretical analysis", AISTATS 2020
[3]L. Toni, "Spectral MAB for unknown graph processes", EUSIPCO, 2018

GSP for RL





High-dimensional state-action space



Learning model / value function

Learning graph edge weights / signal on graph

C. M. Machado et al. "Eigenoption discovery through the deep successor representation", ICLR, 2018
S. Madjiheurem, "Representation learning on graphs: A reinforcement learning application", AISTATS, 2019.
S. Rozada et al., "Low-rank State-action Value-function Approximation", arXiv 2104.08805v1, 2021

GSP-Based Decision Making Strategies

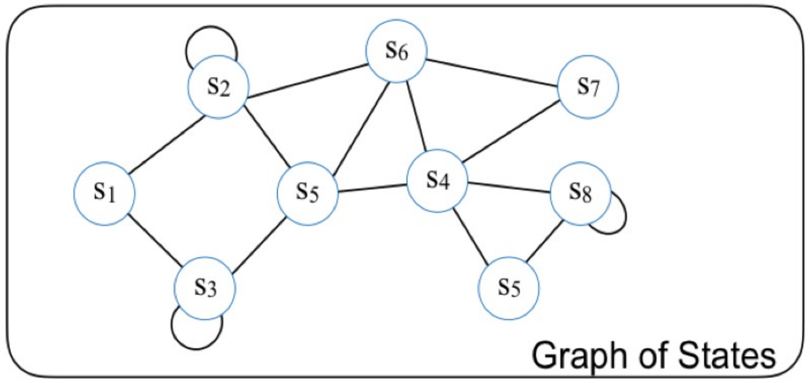
Challenge II: how can GSP tools be applied to DMSs to improve efficiency, complexity, and robustness?

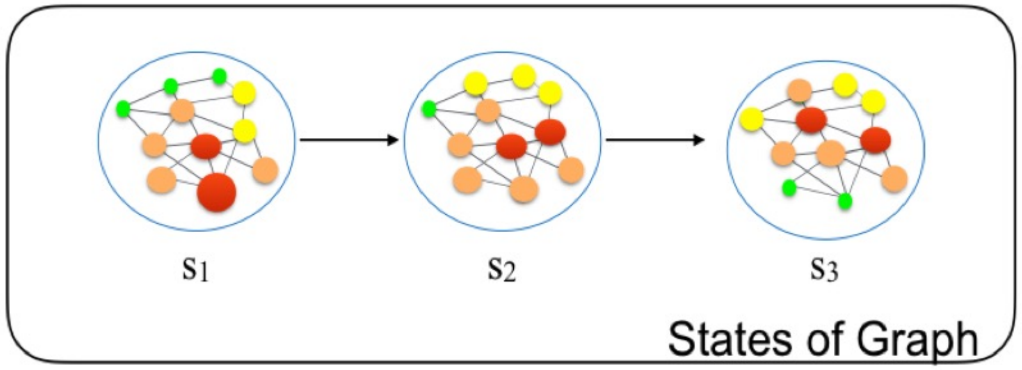
- GSP to improve data efficiency by learning in the spectral domain or by regularising on G — bandit and RL
    - Graph is not usually inferred (Topological inference)
    - Graph uncertainty is not considered (Topological uncertainty)
    - GSP-based analysis for further guarantees (Graph-based Regret bounds)

GSP-Based Decision Making Strategies

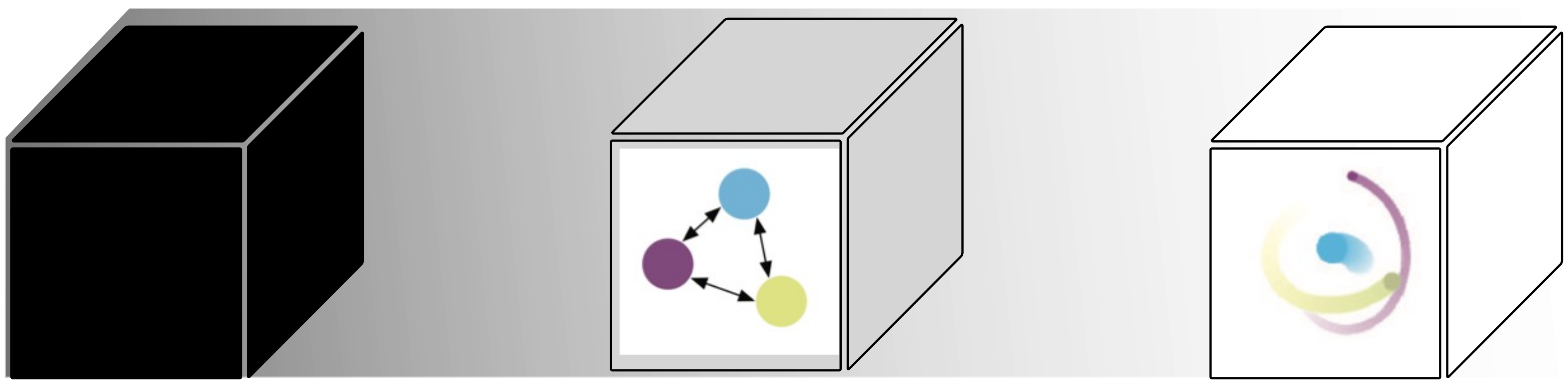
Challenge II: how can GSP tools be applied to DMSs to improve efficiency, complexity, and robustness?

- GSP to improve data efficiency by learning in the spectral domain or by regularising on G — bandit and RL
- GSP to improve accuracy/robustness
- GSP to improve computational efficiency
- GSP to model system dynamics





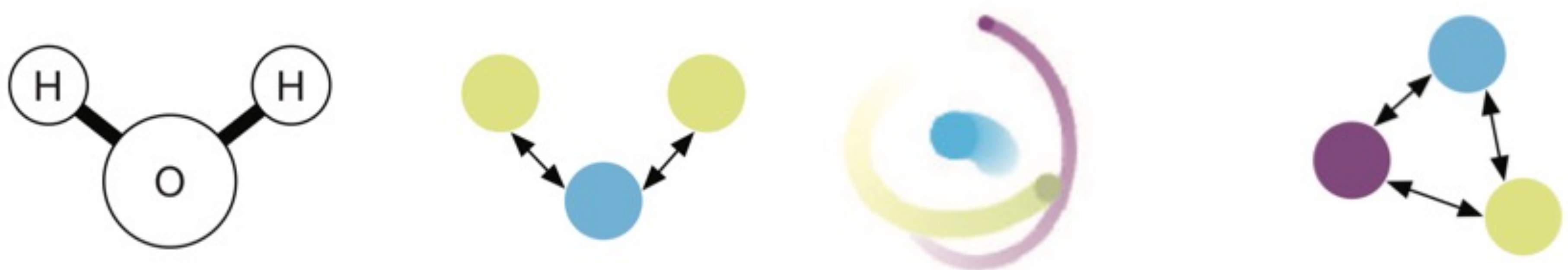
GSP and Model Interpretability



GSP and Model Interpretability

Challenge III: how can GSP tools help enhance interpretability of machine learning models?

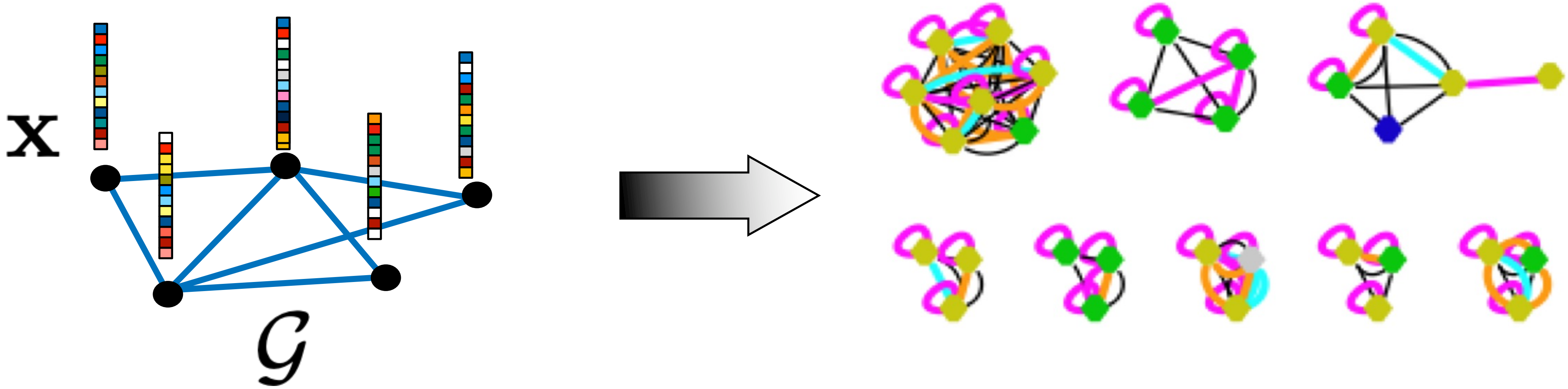
- Modelling the structure of the data with a graph could be a way of introducing domain knowledge (e.g., physical interactions)



- Graph filters may be designed (via e.g., anisotropic filters or adapting attention mechanisms) to enhance model interpretability

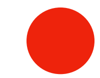
P. W. Battaglia et al., “Relational inductive biases, deep learning, and graph networks,” 2018, arXiv:1806.01261.
F. Monti et al., “Geometric deep learning on graphs and manifolds using mixture model CNNs,” CVPR, 2017.

GSP and Higher-Order Structure

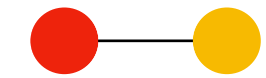


GSP and High-order Structure

- Graphs capture pairwise (lower-order) relationship between nodes



1st- level



2nd-level motif

- Higher-order structures play a key role in understanding the fundamental structures that control the behaviour of many complex systems

3rd-level motif

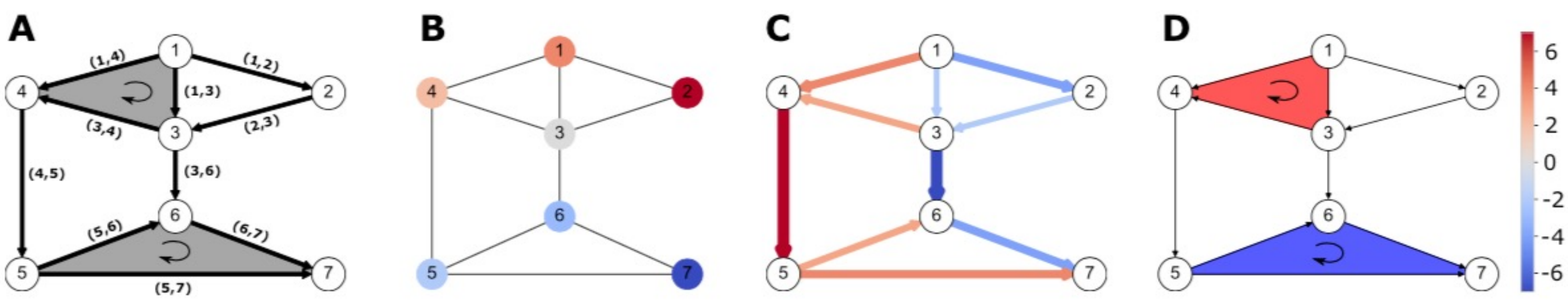
- Motifs have been used to design GNN models that are capable of handling directed graphs

Challenge IV: to extend GSP tools to higher-order structures, such as motifs, simplicial complexes, and hypergraphs

R. Benson et al., "Higher- order organization of complex networks" *Science*, 2016.
F. Monti, et al., "MotifNet: A motif-based Graph Convolutional Network for directed graphs," IEEE DSW, 2018.

GSP and Higher-order Structure

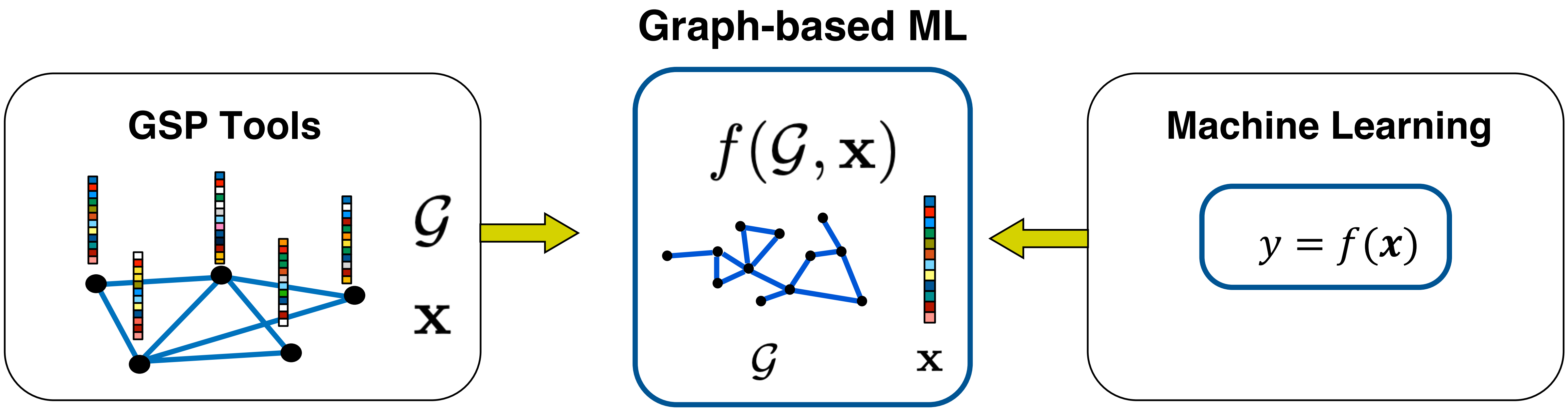
Challenge IV: to extend GSP tools to higher-order structures, such as motifs, simplicial complexes, and hypergraphs



Signals on simplicial complexes of different order

Barbarossa and Sardellitti, “Topological Signal Processing over Simplicial Complexes,” TSP, 2020.
Schaub et al., “Signal Processing on Higher-Order Networks: Livin' on the Edge ... and Beyond,” arXiv, 2021.

Conclusions



- enable convolution & hierarchical modelling on graphs
- improve efficiency & robustness of (graph-based) ML models
- interpret data structure & learning models on graphs

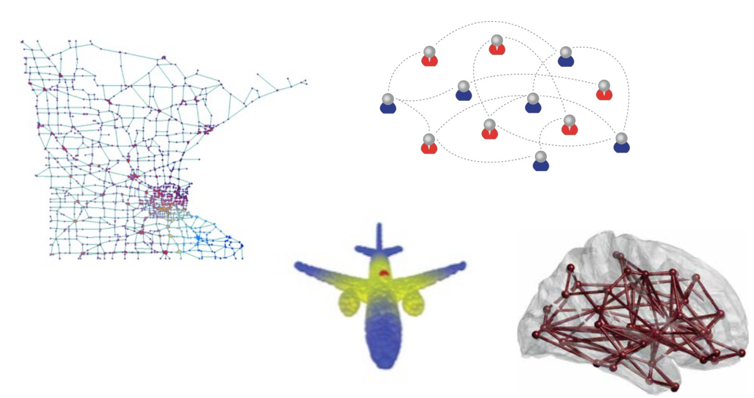
Conclusions

GSP Tools / Benefits	Graph-based regularisation	Graph filters & transforms	GSP-related learning models
Exploiting Data Structure	GP & kernels on graphs	multiscale clustering	CNNs on graphs
Improve efficiency & robustness on graphs	multi-task learning	spectral clustering	few-shot learning
Interpret data structure & learning models on graphs	interpreting DNNs	topology inference	attention models

Tasks

- Node / graph classification
- Community Detection
- Topology inference
- Dynamic Inference
- Online learning

Applications



Open Challenges and Perspectives

- Probabilistic models
- Decision Making Strategies
- Model interpretability
- Higher-order structures

Dong et al., “Graph signal processing for machine learning,” IEEE SPM, 2020.

References

- M. Vargas-Vieyra, et al, “Joint Learning of the Graph and the Data Representation for Graph-Based Semi-Supervised Learning”, TextGraphs 2020.
- Chenhui Hu et al., “Matched signal detection on graphs: Theory and application to brain imaging data classification”, NeuroImage, vol.125, pp. 587-600, 2016
- N. Tremblay and P. Borgnat, “Graph wavelets for multiscale community mining”, IEEE TSP, vol. 62, no. 20, pp. 5227-5239, 2014.
- H.-T. Wai, et al., “Blind community detection from low-rank excitations of a graph filter”, IEEE TSP, vol. 68, pp. 436–451, Dec. 2020.
- R. Ramakrishna, H.-T. Wai and A. Scaglione, "A User Guide to Low-Pass Graph Signal Processing and Its Applications: Tools and Applications" IEEE Signal Processing Magazine, vol. 37, no. 6, pp. 74-85, 2020.
- W. Huang, et al., “A graph signal processing perspective on functional brain imaging”, IEEE Proceeding, vol. 105, no. 5, pp. 868-885, 2018.
- F. Monti et al., "Deep geometric matrix completion: A new way for recommender systems”, ICASSP 2018.
- K. Yang, “Laplacian-regularized graph bandits: Algorithms and theoretical analysis”, AISTATS 2020.
- L. Toni, "Spectral MAB for unknown graph processes”, EUSIPCO, 2018
- M Valko et al. "Spectral bandits for smooth graph functions”, ICML 2014.
- S. Arslan et al., "Graph saliency maps through spectral convolutional networks: Application to sex classification with brain connectivity”, Graphs in Biomedical Image Analysis and Integrating Medical Imaging and Non-Imaging Modalities, 2018.
- L. Stanković, et al. “Data Analytics on Graphs Part I: Graphs and Spectra on Graphs”, Foundations and Trends in Machine Learning: Vol. 13, No. 1, pp 1–157, 2020.
- X. Dong et al. "Inference of mobility patterns via spectral graph wavelets”, ICASSP, 2013.

References

- S. Kwak, N. Geroliminis, and P. Frossard, “Predicting traffic signals on transportation networks using spatio-temporal correlations on graphs”, arXiv:2104.13414, 2021.
- Guoji Fu et al. “Understanding Graph Neural Networks from Graph Signal Denoising Perspectives”, arXiv:2006.04386, 2020.
- A. G. Marques, et al. "Signal processing on directed graphs: The role of edge directionality when processing and learning from network data”, IEEE SPM, vol. 37, no. 6, pp. 99-116, 2020.
- M. Cheung, et al. “Graph signal processing and deep learning: Convolution, pooling, and topology”, IEEE SPM, vol. 37, no. 6, pp. 139-149, Nov. 2020.
- J. H. Giraldo, et al., "Graph Moving Object Segmentation”, IEEE TPAMI, doi: 10.1109/TPAMI.2020.3042093, 2020.
- Yue Wang, et al. “Dynamic Graph CNN for Learning on Point Clouds”, ACM Trans. Graph. vol. 38, no. 146, Nov. 2019.
- N. Perraudin et al., “DeepSphere: Efficient spherical convolutional neural network with HEALPix sampling for cosmological applications”, Astronomy and Computing, vol. 27, pp. 130-146, 2019.
- Hu, Wei, et al. "Graph Signal Processing for Geometric Data and Beyond: Theory and Applications”, arXiv:2008.01918, 2020.
- Z. Fu and W. Hu, "Dynamic Point Cloud Inpainting via Spatial-Temporal Graph Learning," IEEE TMM, 2020.
- S. Parisot, “Disease prediction using graph convolutional networks: application to autism spectrum disorder and Alzheimer’s disease”, Medical image analysis, vol. 48, pp. 117-130, 2018.
- Shah, Chintan, et al. "Finding Patient Zero: Learning Contagion Source with Graph Neural Networks”, arXiv:2006.11913, 2020

References

- P. Gainza et al. “Deciphering interaction fingerprints from protein molecular surfaces using geometric deep learning”, Nature Methods 2020 Feb;17(2):184-192, Dec. 2019.
- Y. Zhang, et al. "Bayesian graph convolutional neural networks for semi-supervised classification”, AAAI 2019.
- E. Isufi, A. Loukas, A. Simonetto, and G. Leus, “Filtering random graph processes over random time-varying graphs,” IEEE TSP, vol. 65, no. 16, pp. 4406–4421, Aug. 2017.
- J. Miettinen, S. A. Vorobyov and E. Ollila, "Robust Least Mean Squares Estimation of Graph Signals," *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019, pp. 5416-5420, doi: 10.1109/ICASSP.2019.8683193.
- S. Madjiheurem and L. Toni, "Representation learning on graphs: A reinforcement learning application”, AISTATS, 2019.
- S. Rozada et al., “Low-rank State-action Value-function Approximation”, arXiv 2104.08805v1, 2021.
- C. M. Machado et al. "Eigenoption discovery through the deep successor representation”, ICLR, 2018.
- P. W. Battaglia et al., “Relational inductive biases, deep learning, and graph networks,” 2018, arXiv:1806.01261.
- F. Monti, D. Boscaini, J. Masci, E. Rodola, J. Svoboda, and M. M. Bronstein, “Geometric deep learning on graphs and manifolds using mixture model CNNs,” in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2017, pp. 5115–5124.
- Fu, Dongqi et al.,"Local motif clustering on time-evolving graphs”, ACM SIGKDD, 2020.
- A. R. Benson, D. F. Gleich, and J. Leskovec, “Higher- order organization of complex networks,” *Science*, vol. 353, no. 6295, pp. 163–166, 2016.
- S. Rey and A. G. Marques, “Robust graph-filter identification with graph denoising regularization,” ICASSP, 2021.

References

- J.D. Medaglia, J.D. et al., “Functional alignment with anatomical networks is associated with cognitive flexibility”, *Nat Hum Behav* 2, pp. 156–164, 2018.
- W. Huang, et al., "A Graph Signal Processing Perspective on Functional Brain Imaging" IEEE Proc., vol. 106, no. 5, pp. 868-885, May 2018.
- Y. Guo, et al.”Deep learning for 3d point clouds: A survey”, IEEE TPAMI, 2020.
- Y. Zhang and M. Rabbat, “A Graph-CNN for 3D Point Cloud Classification”, ICASSP 2018.
- Y. Li and G. Mateos, "Graph Frequency Analysis of COVID-19 Incidence to Identify County-Level Contagion Patterns in the United States”, ICASSP 2021.
- https://deepmind.com/blog/article/traffic-prediction-with-advanced-graph-neural-networks.

Image Credits

- https://upload.wikimedia.org/wikipedia/commons/7/7e/Scientific_citations.png
- https://www.xpertup.com/downloads/movie-recommender-system/
- https://image.slidesharecdn.com/copyofspotifydiscoverweekly-170701234254/95/spotify-discover-weekly-the-machine-learning-behind-your-music-recommendations-1-638.jpg?cb=1498952892
- https://images.immediate.co.uk/production/volatile/sites/3/2019/11/bbc-iplayer-862f585.jpg?quality=90&resize=620,413
- https://miro.medium.com/max/3200/1*dMR3xmufnmKiw4crlisQUA.png
- https://lh3.googleusercontent.com/hRtL9HSoj2dthO5BLoS_7GsyDwrc_eGyvL19ZnabdvcKP4Q9F8oNRyby5kKiB4A7xh4GEhIS4tqtqymZApZV3LCwzCSn5E0lT4674Q=w1440
- https://image.shutterstock.com/image-vector/set-happy-smiling-schoolkids-vector-260nw-1798071838.jpg
- https://i.pinimg.com/originals/5f/05/b0/5f05b07d049661573f223de6635e3f6e.jpg
- https://github.com/huawei-noah/BGCN
- https://miro.medium.com/max/2000/1*QjQc4nRUWYt79UUSBPof4w.png
- https://journals.plos.org/plosone/article/figure/image?download&size=large&id=info:doi/10.1371/journal.pone.0031437.g006

Thank you!