

# Learning Multiplex Graph with Inter-layer Coupling

Chenyue Zhang, Hoi-To Wai

Department of Systems Engineering & Engineering Management,  
The Chinese University of Hong Kong



LOG 2023 Meetup at Shanghai

Acknowledgement: HKRGC project #24203520

# Motivation

---



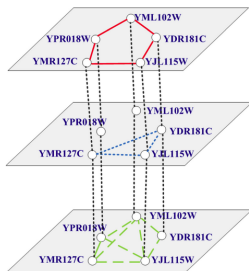
[Source: NETWORK MANAGEMENT]



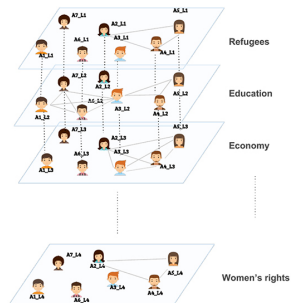
[Source: Sapien Labs]

- ▶ Graphs are natural ways to represent *social, biology, transportation, power networks*, and others.
- ▶ **Graph Signal Processing (GSP)** — extends *signal processing* to graph data and enables ‘interpretable’ **inference** of data.

# Motivation



Multilayer protein networks [Source: [Zhao et al., 2016]]



Multilayer social network [Source: [Hanteer and Rossi, 2019]]

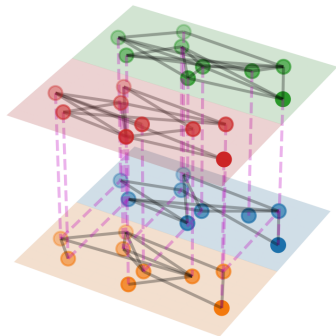
- ▶ Prior works consider **closed systems** with single layer of networks, but *networks do not live in isolation* [Kivelä et al., 2014].
- ▶ A general model is **multiplex graph** – a node is present on  $\geq 2$  **layers of graphs**, each with different topology – e.g., opinion dynamics on  $\geq 2$  topics, weather measurement stations, brain signals, etc.
- ▶ Note: We focus on *multi-attribute graph signals*.

# Goals and Contributions

To extend GSP methods to multiplex graphs: *signal analysis, topology learning*, etc. — we focus on **graph topology learning**.

## Contributions:

- ▶ **Multiplex Graph Filter** for *multi-attribute* graph signals with nonlinear intra-/inter-layer couplings.
- ▶ Interpret TV/smoothness criterion as a **matched filter** criterion – extend to handle inter-layer couplings.
- ▶ **Alternating Optimization Procedure** for efficient multiplex graph learning.

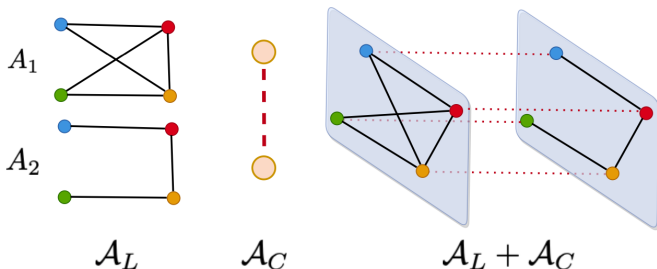


# Related Works

- ▶ **Models for multi-way graph signals** (on multiplex graphs)
  - ▶ [Zhang et al., 2023b] – **tensor GSP model**, but it lacks inter-layer coupling dynamics.
  - ▶ [Natali et al., 2020, Grassi et al., 2017] considered **product graph signal models** which is a special case of multiplex graphs.
- ▶ **Multiplex graph topology learning**
  - ▶ [Kalaitzis et al., 2013] [Kadambari and Chepuri, 2021], [Einizade and Sardouie, 2023] learn **product graphs** using graph signals via smoothness, spectral template, etc.
  - ▶ Our prior work [Zhang et al., 2023a] consider a fine-grained model for product graph learning.
- ▶ **Graph Machine Learning**
  - ▶ [Cen et al., 2019, Zhang et al., 2019] seek embeddings for graph representation on heterogeneous graph with HetGNN.
  - ▶ [Butler et al., 2023] proposed a model for convolutional learning on multigraph.
- ▶ and many others ...

# Multiplex Graph Model

- ▶  $G = \langle V, \mathcal{E}, \mathcal{G}^C \rangle$  with nodes  $V$ , layer edges  $\mathcal{E}$ , coupling graph  $\mathcal{G}^C$ .
- ▶ There are  $|V| = N$  nodes and  $L$  layers.
- ▶ **Layer Graphs:** For  $\ell = 1, \dots, L$ ,  $\mathcal{G}_\ell$  with supernodes  $V_\ell$ , edges  $E_\ell$ , representing intra-layer links with adjacency  $\mathbf{A}_\ell$ .
- ▶ **Coupling Graph:**  $\mathcal{G}^C$  for inter-layer links with adjacency  $\mathbf{C}$ .
- ▶ **Adjacency Matrix:** Layer-wise  $\mathbf{A}_L = \text{blkdiag}(\mathbf{A}_1, \dots, \mathbf{A}_L)$  and coupling  $\mathbf{A}_C = \mathbf{C} \otimes I_N$ . e.g: supra-adjacency:  $\mathbf{A} = \mathbf{A}_L + \mathbf{A}_C$ .



# Multi-attribute Graph Filter and Signals.

- ▶ We model **multi-attribute** graph signal as

$$\mathbf{y}^{(m)} = \mathcal{H}(\mathcal{A}_L, \mathcal{A}_C) \mathbf{x}^{(m)} + \mathbf{w}^{(m)} \in \mathbb{R}^{NL}, \quad (1)$$

- ▶  $\mathcal{H}(\mathcal{A}_L, \mathcal{A}_C)$  shall model the multiplex with distinct **intra-layer and coupling dynamics**  $\rightarrow$  a **general** multi-attribute graph filter model<sup>1</sup>:

$$\mathcal{H}(\mathcal{A}_L, \mathcal{A}_C) = \sum_{t=0}^{T-1} \sum_{j=0}^{2^t-1} h_{t,j} \prod_{i=1}^t \mathcal{A}_L^{b_{t,j}^{(i)}} \mathcal{A}_C^{1-b_{t,j}^{(i)}}, \quad b_{t,j}^{(i)} \in \{0, 1\} \quad (\text{GF})$$

- ▶ **Remark:** the tensor GSP model [Zhang et al., 2023b] essentially takes the **polynomial filter** of supra-adjacency matrix

$$\mathcal{H}(\mathcal{A}_L, \mathcal{A}_C) = \sum_{t=0}^{T-1} h_t (\mathcal{A}_L + \mathcal{A}_C)^t. \quad (\text{GF-t})$$

---

<sup>1</sup>Note: this is a multinomial with exponential number of coefficients. Similar observations are made in [Butler et al., 2023].

# Expressibility of (GF): Multiplex Graph Dynamics

It is necessary to use the model (GF) with

$$\mathcal{H}(\mathcal{A}_L, \mathcal{A}_C) = \sum_{t=0}^{T-1} \sum_{j=0}^{2^t-1} h_{t,j} \prod_{i=1}^t \mathcal{A}_L^{b_{t,j}^{(i)}} \mathcal{A}_C^{1-b_{t,j}^{(i)}}, \quad b_{t,j}^{(i)} \in \{0, 1\}$$

---

## Supra-diffusion Process:

- ▶ E.g., dynamics of epidemics [Kivelä et al., 2014]:

$$\frac{d\mathbf{y}_\ell(t)}{dt} = -\mathbf{y}_\ell(t) + \underbrace{\mathbf{A}_\ell \mathbf{y}_\ell(t)}_{\text{intra-layer}} + \underbrace{\sum_{\ell'=1}^L \mathbf{C}_{\ell,\ell'} \mathbf{y}_{\ell'}(t)}_{\text{inter-layer}} + \mathbf{x}_\ell^{(m)}.$$

- ▶ Steady-state of the diffusion process:

$$\mathbf{y}^{(m)} = \lim_{t \rightarrow \infty} \mathbf{y}(t) = (\mathbf{I}_{NL} - (\mathbf{A}_L + \mathbf{A}_C))^{-1} \mathbf{x}^{(m)},$$

- ▶ Fine with (GF) and (GF-t).



# Expressibility of (GF): Multiplex Graph Dynamics

It is necessary to use the model (GF) with

$$\mathcal{H}(\mathcal{A}_L, \mathcal{A}_C) = \sum_{t=0}^{T-1} \sum_{j=0}^{2^t-1} h_{t,j} \prod_{i=1}^t \mathcal{A}_L^{b_{t,j}^{(i)}} \mathcal{A}_C^{1-b_{t,j}^{(i)}}, \quad b_{t,j}^{(i)} \in \{0, 1\}$$

---

## Opinion Dynamics:

- ▶ Evolution with mutual trust  $\mathbf{C}$  and logical matrix  $\mathbf{A}_\ell$ :

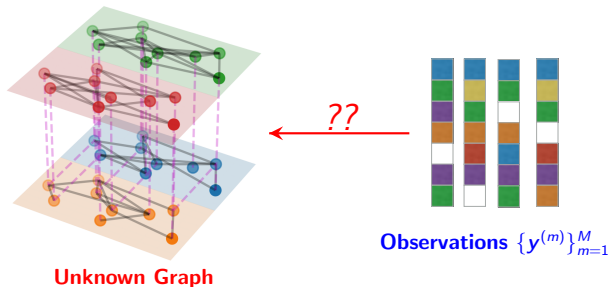
$$\mathbf{y}_\ell(t+1) = \underbrace{\mathbf{A}_\ell \sum_{\ell'=1}^L \mathbf{C}_{\ell,\ell'} \mathbf{y}_{\ell'}(t)}_{\text{coupled inter- and intra-layer}} + \mathbf{x}_\ell^{(m)},$$

- ▶ Steady-state opinions:

$$\mathbf{y}^{(m)} = \lim_{t \rightarrow \infty} \mathbf{y}(t) = (\mathbf{I}_{NL} - \mathbf{A}_L \mathbf{A}_C)^{-1} \mathbf{x}^{(m)}.$$

- ▶ **Fine with** by (GF) **but not** (GF-t).

# Multiplex Graph Learning



**Task:** Given graph signals  $\{\mathbf{y}^{(m)}\}_{m=1}^M$ , estimate multiplex graph  $\mathcal{A}_L, \mathcal{A}_C$ .

- **General idea:** Following [Dong et al., 2016], exploit **smoothness** of multi-attribute graph signals  $\rightarrow$  how to leverage (GF)?

# TV Objective and Matched Graph Filter

(Let's take a slight detour ...)

- ▶ Let  $\mathbf{S} \in \mathbb{R}^{N \times N}$  be the pairwise distance matrix of graph signals and we aim at learning the (simple) graph adjacency  $\mathbf{A}$ .
- ▶ Consider the **Dirichlet energy criterion** in [Berger et al., 2020]<sup>2</sup>:

$$\text{TV}(\hat{\mathbf{A}}) := \sum_{i,j=1}^N \hat{A}_{ij} \frac{1}{M} \|\mathbf{y}_i - \mathbf{y}_j\|^2 = \sum_{i,j=1}^N \hat{A}_{ij} S_{ij} = \langle \hat{\mathbf{A}} | \mathbf{S} \rangle. \quad (2)$$

With  $\mathbf{y}^{(m)} \approx \mathcal{H}(\mathbf{A})\mathbf{x}^{(m)}$  and under mild condition

$$\min_{\hat{\mathbf{A}}} \text{TV}(\hat{\mathbf{A}}) \stackrel{\text{approx.}}{\iff} \max_{\hat{\mathbf{A}}} \langle \hat{\mathbf{A}} | \mathcal{H}^2(\mathbf{A}) \rangle.$$

- ▶ If  $\mathcal{H}^2(\mathbf{A})$  is a low-pass graph filter [Ramakrishna et al., 2020], then its **first order approximation**<sup>3</sup> is given by  $\mathcal{H}^2(\mathbf{A}) \approx \mathbf{A}$ .
- ▶ Criterion (2) can be interpreted as a **matched filter** criterion.

---

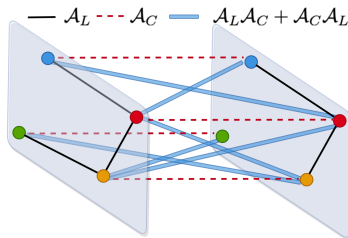
<sup>2</sup>Graph learning methods based on quadratic TV such as [Dong et al., 2016] can be interpreted similarly.

<sup>3</sup>In general  $\mathcal{H}(\mathbf{A})$  is not known a-priori, 1st order approx is the best we can do.

# Tractable Approximation to (GF)

- ▶ (GF) is **intractable** in general  $\because$  exponential no. of parameters.
- ▶ Inspired by the examples, consider the **approximation**<sup>4</sup>:

$$\mathcal{H}(\mathcal{A}_L, \mathcal{A}_C) \approx \overline{\mathcal{H}}_1(\mathcal{A}_L) + \overline{\mathcal{H}}_2(\mathcal{A}_C) + \overline{\mathcal{H}}_3(\mathcal{A}_L \mathcal{A}_C + \mathcal{A}_C \mathcal{A}_L) \quad (\text{a-GF})$$



- ▶  $\overline{\mathcal{H}}_1, \overline{\mathcal{H}}_2, \overline{\mathcal{H}}_3$  are polynomials.
- ▶  $\overline{\mathcal{H}}_1(\mathcal{A}_L), \overline{\mathcal{H}}_2(\mathcal{A}_C)$  model intra- and inter-layer graph dynamics,
- ▶  $\overline{\mathcal{H}}_3$  captures **two-hops** neighbors and cross-layer interactions.

<sup>4</sup>Also assume that  $\mathcal{H}(\mathcal{A}_L, \mathcal{A}_C)\mathcal{H}(\mathcal{A}_L, \mathcal{A}_C)^\top$  obeys a similar form.

# Matched Graph Filter for Multiplex Graph Learning

- ▶  $\mathbf{S} \in \mathbb{R}^{NL \times NL}$  = pairwise distance matrix of **multi-attribute** signals.
- ▶ Consider a **generalized Dirichlet energy criterion**

$$\text{TV}(\mathcal{A}_L, \mathcal{A}_C) := \sum_{i,j=1}^{NL} [\hat{h}(\mathcal{A}_L, \mathcal{A}_C)]_{ij} S_{ij} = \langle \hat{h}(\mathcal{A}_L, \mathcal{A}_C) | \mathbf{S} \rangle. \quad (3)$$

- ▶ With (**a-GF**), we have

$$\mathbf{S} \approx \overline{\mathcal{H}}_1(\mathcal{A}_L) + \overline{\mathcal{H}}_2(\mathcal{A}_C) + \overline{\mathcal{H}}_3(\mathcal{A}_L \mathcal{A}_C + \mathcal{A}_C \mathcal{A}_L).$$

- ▶ **Assumption H1:**  $\overline{\mathcal{H}}_1(\cdot), \overline{\mathcal{H}}_2(\cdot), \overline{\mathcal{H}}_3(\cdot)$  are **low-pass graph filters**.
- ▶ Under H1, the **matched multiplex graph filter** design:

$$\hat{h}(\mathcal{A}_L, \mathcal{A}_C) = \mathcal{A}_L + \mathcal{A}_C + \lambda(\mathcal{A}_C \mathcal{A}_L + \mathcal{A}_L \mathcal{A}_C).$$

→ A **high-order smoothness** metric!

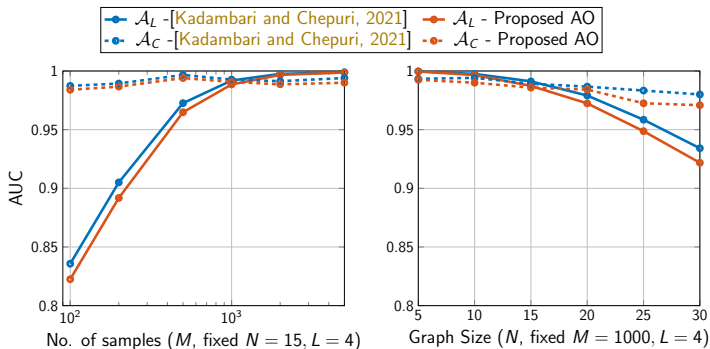
# Multiplex Graph Learning - Algorithm

Under H1, we formulate the **bi-convex** problem:

$$\min_{\hat{\mathcal{A}}_L, \hat{\mathcal{A}}_C \in \mathcal{A}} \left\langle \hat{\mathcal{A}}_L + \hat{\mathcal{A}}_C + \lambda \left( \hat{\mathcal{A}}_L \hat{\mathcal{A}}_C + \hat{\mathcal{A}}_C \hat{\mathcal{A}}_L \right) \mid \mathbf{s} \right\rangle + \alpha \left( \|\hat{\mathcal{A}}_L\|_F^2 + \|\hat{\mathcal{A}}_C\|_F^2 \right) \quad (4)$$

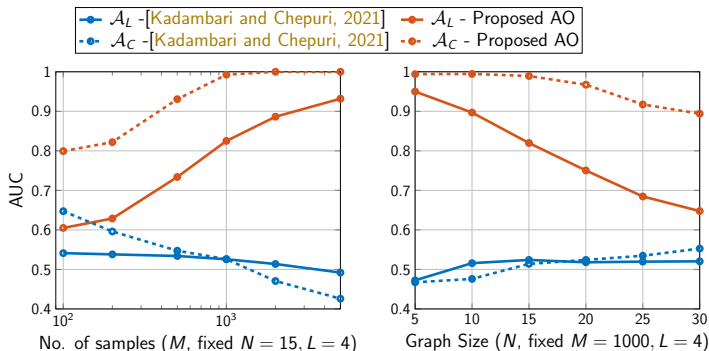
- ▶ **Algorithm:** **alternating optimization** (AO) can be applied:
  - ▶ Fix  $\hat{\mathcal{A}}_C$  and solve for  $\hat{\mathcal{A}}_L \rightarrow$  fix  $\hat{\mathcal{A}}_L$  and solve for  $\hat{\mathcal{A}}_C \rightarrow \dots$
- ▶ The AO subproblems are **separable and tractable** – each involves **convex** problems size of  $N \times N$  or  $L \times L$ .
- ▶ AO finds a stationary point of (4) as iteration number goes to  $\infty$  [Grippio and Sciandrone, 2000].
- ▶ **Remark:** when  $\lambda = 0$ , the problem reduces into that of [Kadambari and Chepuri, 2021].

# Topology Reconstruction under strong coupling



- ▶ Weak coupling:  $\mathcal{H}_{\text{wk}}(\mathcal{A}_L, \mathcal{A}_C) = (\mathbf{I} - \tau_{\text{wk}}(\mathcal{A}_L + \mathcal{A}_C))^{-1}$ .
- ▶ Observation: AUC performance generally improves as  $M$  increases/deteriorates as  $N$  increases.
- ▶ Proposed AO ( $\lambda = 0.1$ ) attains **similar** performance to the benchmark.

# Topology Reconstruction under strong coupling



- ▶ Strong coupling:  $\mathcal{H}_{\text{str}}(\mathcal{A}_L, \mathcal{A}_C) = (I - \tau_{\text{str}} \mathcal{A}_L \mathcal{A}_C)^{-1}$ .
- ▶ Benchmark fails in estimating graph topologies under strong coupling.
- ▶ Proposed AO ( $\lambda = 5$ ) recovers topology effectively regardless of layer coupling **robustly**.



# Summary

**Takeaway:** Distinct **inter-layer and intra-layer interactions** dynamics require careful modeling for multiplex graph learning.

We have introduced a method for learning multiplex network structures from multi-attribute graph signals:

- ▶ **General multiplex graph filter** to model complex signal interactions.
- ▶ **Matched filter** perspective to graph learning by smoothness → a **high-order smoothness metric** aimed at inter-layer coupling.
- ▶ An **efficient AO procedure** for learning graph topologies.
- ▶ **Future work:** modeling of multiplex graph signals, adopting other GSP tools, ...

**Thank you!**

# References

- [Berger et al., 2020] Berger, P., Hannak, G., and Matz, G. (2020). Efficient graph learning from noisy and incomplete data. *IEEE Transactions on Signal and Information Processing over Networks*, 6:105–119.
- [Butler et al., 2023] Butler, L., Parada-Mayorga, A., and Ribeiro, A. (2023). Convolutional learning on multigraphs. *IEEE Transactions on Signal Processing*, 71:933–946.
- [Cen et al., 2019] Cen, Y., Zou, X., Zhang, J., Yang, H., Zhou, J., and Tang, J. (2019). Representation learning for attributed multiplex heterogeneous network. In *KDD*, pages 1358–1368.
- [Dong et al., 2016] Dong, X., Thanou, D., Frossard, P., and Vandergheynst, P. (2016). Learning laplacian matrix in smooth graph signal representations. *IEEE Transactions on Signal Processing*, 64(23):6160–6173.
- [Einizade and Sardouie, 2023] Einizade, A. and Sardouie, S. H. (2023). Learning product graphs from spectral templates. *IEEE Transactions on Signal and Information Processing over Networks*.
- [Grassi et al., 2017] Grassi, F., Loukas, A., Perraudin, N., and Ricaud, B. (2017). A time-vertex signal processing framework: Scalable processing and
- IEEE Transactions on Signal Processing*, 66(3):817–829.
- [Grippo and Sciandrone, 2000] Grippo, L. and Sciandrone, M. (2000). On the convergence of the block nonlinear gauss–seidel method under convex constraints. *Operations research letters*, 26(3):127–136.
- [Hanteer and Rossi, 2019] Hanteer, O. and Rossi, L. (2019). An innovative way to model twitter topic-driven interactions using multiplex networks. *Frontiers in big data*, 2:9.
- [Kadambari and Chepuri, 2021] Kadambari, S. K. and Chepuri, S. P. (2021). Product graph learning from multi-domain data with sparsity and rank constraints. *IEEE Transactions on Signal Processing*, 69:5665–5680.
- [Kalaitzis et al., 2013] Kalaitzis, A., Lafferty, J., Lawrence, N. D., and Zhou, S. (2013). The bigraphical lasso. In *ICML*, pages 1229–1237. PMLR.
- [Kivelä et al., 2014] Kivelä, M., Arenas, A., Barthélemy, M., Gleeson, J. P., Moreno, Y., and Porter, M. A. (2014). Multilayer networks. *Journal of complex networks*, 5(1):202–258.
- Forecasting multi-dimensional processes over graphs. In *ICASSP*, pages 5575–5579. IEEE.
- [Ramakrishna et al., 2020] Ramakrishna, R., Wai, H. T., and Scaglione, A. (2020). A user guide to low-pass graph signal processing and its applications: Tools and applications. *IEEE Signal Processing Magazine*, 37(6):74–85.
- [Zhang et al., 2023a] Zhang, C., He, Y., and Wai, H.-T. (2023a). Product graph learning from multi-attribute graph signals with inter-layer coupling. In *ICASSP*, pages 1–5. IEEE.
- [Zhang et al., 2019] Zhang, C., Song, D., Huang, C., Swami, A., and Chawla, N. V. (2019). Heterogeneous graph neural network. In *KDD*, pages 793–803.
- [Zhang et al., 2023b] Zhang, S., Deng, Q., and Ding, Z. (2023b). Signal processing over multilayer graphs: Theoretical foundations and practical applications. *IEEE Internet of Things Journal*.
- [Zhao et al., 2016] Zhao, B., Hu, S., Li, X., Zhang, F., Tian, Q., and Ni, W. (2016). An efficient method for protein function annotation based on multilayer protein