Sparse Partial Least Squares for Coarse Noisy Graph Alignment

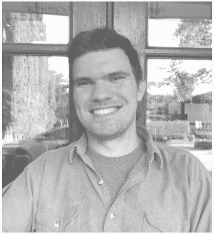
Michael Weylandt

IEEE Statistical Signal Processing Workshop 2021 (Virtual)

IC Postdoctoral Research Fellow
Univeristy of Florida Informatics Institute

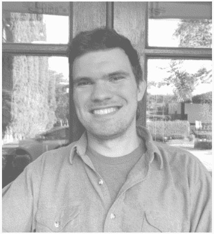
Acknowledgements

Co-Authors: George Michailidis (UF) and Mitch Roddenberry (Rice ECE)

Acknowledgements

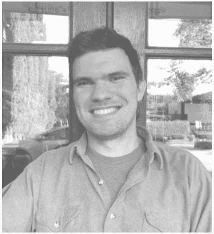
Co-Authors: George Michailidis (UF) and Mitch Roddenberry (Rice ECE)



US Intelligence Community Postdoctoral Research Fellowship

Acknowledgements

Co-Authors: George Michailidis (UF) and Mitch Roddenberry (Rice ECE)



US Intelligence Community Postdoctoral Research Fellowship

Discussions: Genevera Allen (Rice U)

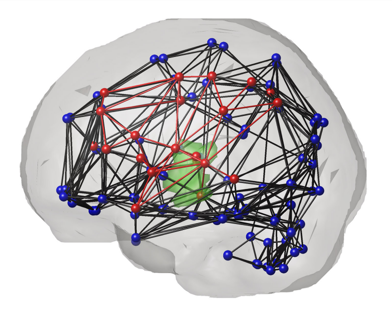
Ubiquitous Network Data: telecommunications, social media, neuroscience, sensor networks, transportation, *etc.*

Network Data

Ubiquitous Network Data: telecommunications, social media, neuroscience, sensor networks, transportation, *etc.*

Three types of network data:

- Networks as models of complex phenomena ("graphical models")



Network Data

Ubiquitous Network Data: telecommunications, social media, neuroscience, sensor networks, transportation, *etc.*

Three types of network data:

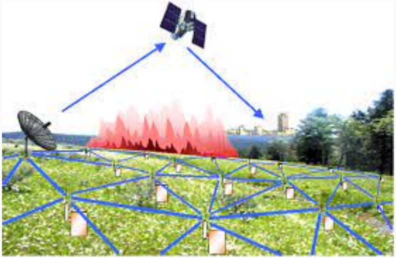
- Networks as models of complex phenomena ("graphical models")
- Observed network(s) ("network science")

Network Data

Ubiquitous Network Data: telecommunications, social media, neuroscience, sensor networks, transportation, *etc.*

Three types of network data:

- Networks as models of complex phenomena ("graphical models")
- Observed network(s) ("network science")
- Data observed on a network ("graph signal processing")



Ubiquitous Network Data: telecommunications, social media, neuroscience, sensor networks, transportation, *etc.*

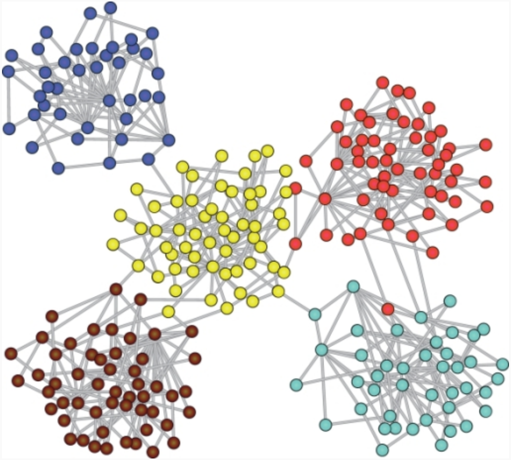
Three types of network data:

- Networks as models of complex phenomena ("graphical models")
- Observed network(s) ("network science")
- Data observed on a network ("graph signal processing")

Today: contributions to **network science** and **graph signal processing**

Community Detection

Community detection - dividing an observed network into meaningful groups ("communities")



Very well studied problem (Levina, Zhu, Amini, Vershynin, Bickel ...)

Coarse Alignment via Graph Signal Processing

Graph alignment:

- Input: two graphs $\mathcal{G}_1, \mathcal{G}_2$
- Output: Mapping between $\mathcal{V}_1, \mathcal{V}_2$ so that the graphs are (nearly) "the same" under some metric

Can require exact alignment or allow inexact matching (possibly on different size graphs)

Coarse Alignment via Graph Signal Processing

Graph alignment:

- Input: two graphs $\mathcal{G}_1, \mathcal{G}_2$
- Output: Mapping between $\mathcal{V}_1, \mathcal{V}_2$ so that the graphs are (nearly) "the same" under some metric

Can require exact alignment or allow inexact matching (possibly on different size graphs)

Coarse graph alignment:

- Input: two graphs $\mathcal{G}_1, \mathcal{G}_2$
- Output: Communities in $\mathcal{G}_1, \mathcal{G}_2$ and **mapping between them**

Coarse Alignment via Graph Signal Processing

Graph alignment:

- Input: two graphs $\mathcal{G}_1, \mathcal{G}_2$
- Output: Mapping between $\mathcal{V}_1, \mathcal{V}_2$ so that the graphs are (nearly) "the same" under some metric

Can require exact alignment or allow inexact matching (possibly on different size graphs)

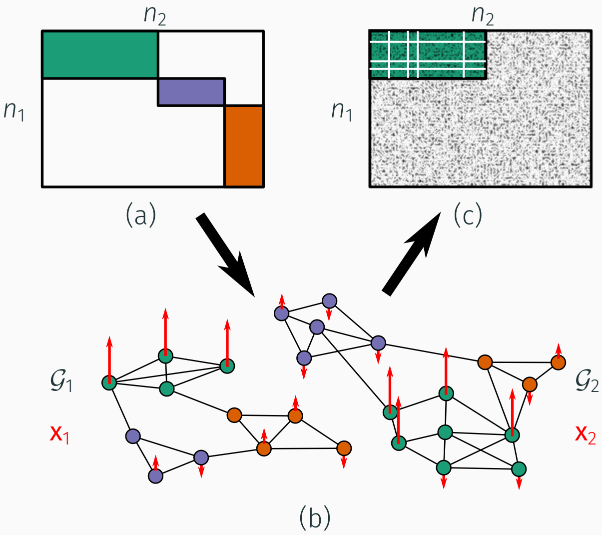
Coarse graph alignment:

- Input: two graphs $\mathcal{G}_1, \mathcal{G}_2$
- Output: Communities in $\mathcal{G}_1, \mathcal{G}_2$ and **mapping between them**

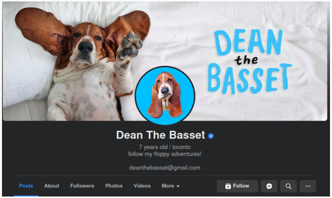
Goal: Coarse graph alignment using noisy graph signals on $\mathcal{G}_1, \mathcal{G}_2$

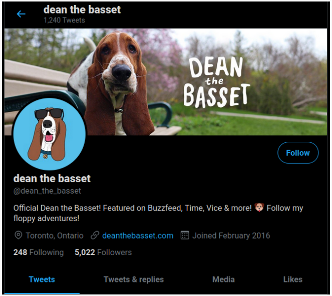
Coarse Graph Alignment via Graph Signal Processing

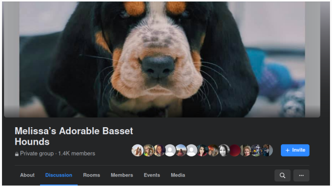
Goal: Coarse graph alignment using noisy graph signals on $\mathcal{G}_1, \mathcal{G}_2$



Example: Coarse Alignment via Graph Signal Processing







CONGA via Sparse Graph PLS

Model: two graphs $\mathcal{G}_1$, $\mathcal{G}_2$ with corresponding signals

- Graphs have "paired" community structure
- If community of $\mathcal{G}_1$ reflects signal, so does corresponding in $\mathcal{G}_2$
- No other structural assumptions on $\mathcal{G}_1, \mathcal{G}_2$

CONGA via Sparse Graph PLS

Model: two graphs $\mathcal{G}_1$, $\mathcal{G}_2$ with corresponding signals

- Graphs have "paired" community structure
- If community of $\mathcal{G}_1$ reflects signal, so does corresponding in $\mathcal{G}_2$
- **No other structural assumptions on $\mathcal{G}_1, \mathcal{G}_2$**

Solution: Graph-Regularized Sparse Multi-Rank PLS

$$\underset{\mathbf{U},\mathbf{V}\in\mathcal{V}_{n_1\times K}^{\mathbf{I}_{n_1}+\alpha_1\mathbf{L}_1}\times\mathcal{V}_{n_2\times K}^{\mathbf{I}_{n_2}+\alpha_2\mathbf{L}_2}}{\arg\max}\ \mathrm{Tr}(\mathbf{U}^\top\mathbf{X}_1^\top\mathbf{X}_2\mathbf{V}) - \lambda_1 P_1(\mathbf{U}) - \lambda_2 P_2(\mathbf{V})$$

CONGA via Sparse Graph PLS

Model: two graphs $\mathcal{G}_1$, $\mathcal{G}_2$ with corresponding signals

- Graphs have "paired" community structure
- If community of $\mathcal{G}_1$ reflects signal, so does corresponding in $\mathcal{G}_2$
- **No other structural assumptions on $\mathcal{G}_1, \mathcal{G}_2$**

Solution: Graph-Regularized Sparse Multi-Rank PLS

$$\underset{\mathbf{U},\mathbf{V} \in \mathcal{V}_{n_1 \times K}^{\mathbf{I}_{n_1}+\alpha_1 \mathbf{L}_1} \times \mathcal{V}_{n_2 \times K}^{\mathbf{I}_{n_2}+\alpha_2 \mathbf{L}_2}}{\arg\max} \operatorname{Tr}(\mathbf{U}^\top \mathbf{X}_1^\top \mathbf{X}_2 \mathbf{V}) - \lambda_1 P_1(\mathbf{U}) - \lambda_2 P_2(\mathbf{V})$$

Decompose $X_1^T X_2$ into two parts U, V such that:

Model: two graphs $\mathcal{G}_1$, $\mathcal{G}_2$ with corresponding signals

- Graphs have “paired” community structure
- If community of $\mathcal{G}_1$ reflects signal, so does corresponding in $\mathcal{G}_2$
- **No other structural assumptions on $\mathcal{G}_1, \mathcal{G}_2$**

Solution: Graph-Regularized Sparse Multi-Rank PLS

$$\mathop{\arg\max}_{\mathbf{U},\mathbf{V} \in \mathcal{V}_{n_1 \times K}^{\mathbf{I}_{n_1}+\alpha_1 \mathbf{L}_1} \times \mathcal{V}_{n_2 \times K}^{\mathbf{I}_{n_2}+\alpha_2 \mathbf{L}_2}} \mathrm{Tr}(\mathbf{U}^\top \mathbf{X}_1^\top \mathbf{X}_2 \mathbf{V}) - \lambda_1 P_1(\mathbf{U}) - \lambda_2 P_2(\mathbf{V})$$

Decompose $X_1^T X_2$ into two parts U, V such that:

- correlation between signals

Model: two graphs $\mathcal{G}_1$, $\mathcal{G}_2$ with corresponding signals

- Graphs have "paired" community structure
- If community of $\mathcal{G}_1$ reflects signal, so does corresponding in $\mathcal{G}_2$
- **No other structural assumptions on $\mathcal{G}_1, \mathcal{G}_2$**

Solution: Graph-Regularized Sparse Multi-Rank PLS

$$\underset{\mathbf{U},\mathbf{V} \in \mathcal{V}_{n_1 \times K}^{\mathbf{I}_{n_1} + \alpha_1 \mathbf{L}_1} \times \mathcal{V}_{n_2 \times K}^{\mathbf{I}_{n_2} + \alpha_2 \mathbf{L}_2}}{\arg\max} \mathrm{Tr}(\mathbf{U}^\top \mathbf{X}_1^\top \mathbf{X}_2 \mathbf{V}) - \lambda_1 P_1(\mathbf{U}) - \lambda_2 P_2(\mathbf{V})$$

Decompose $X_1^T X_2$ into two parts U, V such that:

- correlation between signals
- in a way that respects graph structure

$$\mathbf{U} \in \mathcal{V}^{\mathbf{I}_{n_1} + \alpha_1 \mathbf{L}_1} \Leftrightarrow \mathbf{U}^T (\mathbf{I}_{n_1} + \alpha_1 \mathbf{L}_1) \mathbf{U} = \mathbf{I}$$

CONGA via Sparse Graph PLS

Model: two graphs $\mathcal{G}_1$, $\mathcal{G}_2$ with corresponding signals

- Graphs have "paired" community structure
- If community of $\mathcal{G}_1$ reflects signal, so does corresponding in $\mathcal{G}_2$
- **No other structural assumptions on $\mathcal{G}_1, \mathcal{G}_2$**

Solution: Graph-Regularized Sparse Multi-Rank PLS

$$\underset{\mathbf{U},\mathbf{V} \in \mathcal{V}_{n_1 \times K}^{\mathbf{I}_{n_1}+\alpha_1 \mathbf{L}_1} \times \mathcal{V}_{n_2 \times K}^{\mathbf{I}_{n_2}+\alpha_2 \mathbf{L}_2}}{\arg\max} \operatorname{Tr}(\mathbf{U}^\top \mathbf{X}_1^\top \mathbf{X}_2 \mathbf{V}) - \lambda_1 P_1(\mathbf{U}) - \lambda_2 P_2(\mathbf{V})$$

Decompose $X_1^T X_2$ into two parts U, V such that:

- correlation between signals
- in a way that respects graph structure
- and selects a sparse set of nodes

CONGA via Sparse Graph PLS

Model: two graphs $\mathcal{G}_1$, $\mathcal{G}_2$ with corresponding signals

- Graphs have "paired" community structure
- If community of $\mathcal{G}_1$ reflects signal, so does corresponding in $\mathcal{G}_2$
- **No other structural assumptions on** $\mathcal{G}_1, \mathcal{G}_2$

Solution: Graph-Regularized Sparse Multi-Rank PLS

$$\underset{\mathbf{U},\mathbf{V} \in \mathcal{V}_{n_1 \times K}^{\mathbf{I}_{n_1}+\alpha_1 \mathbf{L}_1} \times \mathcal{V}_{n_2 \times K}^{\mathbf{I}_{n_2}+\alpha_2 \mathbf{L}_2}}{\arg\max} \mathrm{Tr}(\mathbf{U}^\top \mathbf{X}_1^\top \mathbf{X}_2 \mathbf{V}) - \lambda_1 P_1(\mathbf{U}) - \lambda_2 P_2(\mathbf{V})$$

Decompose $X_1^T X_2$ into two parts U, V such that:

- correlation between signals
- in a way that respects graph structure
- and selects a sparse set of nodes

Sparsity + Orthogonality $\approx$ Non-overlapping support

Paired columns of $\mathbf{U}, \mathbf{V}$ give corresponding community membership

Algorithm: Manifold Optimization

Sparse Graph-Smooth PLS Algorithm

1. Construct Graph Laplacians $\mathsf{L}_1, \mathsf{L}_2$ and associated smoothing matrices $\mathsf{S}_1 = \mathsf{I}_{n_1} + \alpha_1 \mathsf{L}_1$, $\mathsf{S}_2 = \mathsf{I}_{n_2} + \alpha_2 \mathsf{L}_2$
2. Initialize $\hat{\mathsf{U}}, \hat{\mathsf{V}}$ to the leading K singular vectors of $\mathsf{X}_1^\top \mathsf{X}_2$
3. Repeat until convergence:

$$\hat{\mathsf{U}} = \underset{\mathsf{U} \in \text{conv}\, \mathcal{V}^{\mathsf{S}_1}_{n_1 \times K}}{\arg\min} - \text{Tr}(\mathsf{U}^\top \mathsf{X}_1^\top \mathsf{X}_2 \hat{\mathsf{V}}) + \lambda_1 P_1(\mathsf{U})$$

$$\hat{\mathsf{V}} = \underset{\mathsf{V} \in \text{conv}\, \mathcal{V}^{\mathsf{S}_2}_{n_2 \times K}}{\arg\min} - \text{Tr}(\mathsf{V}^\top \mathsf{X}_2^\top \mathsf{X}_1 \hat{\mathsf{U}}) + \lambda_2 P_2(\mathsf{V})$$

4. Return $\hat{\mathsf{U}}$ and $\hat{\mathsf{V}}$

Sparse Graph-Smooth PLS Algorithm

1. Construct Graph Laplacians $\mathbf{L}_1, \mathbf{L}_2$ and associated smoothing matrices $\mathbf{S}_1 = \mathbf{I}_{n_1} + \alpha_1 \mathbf{L}_1$, $\mathbf{S}_2 = \mathbf{I}_{n_2} + \alpha_2 \mathbf{L}_2$
2. Initialize $\hat{\mathbf{U}}, \hat{\mathbf{V}}$ to the leading K singular vectors of $\mathbf{X}_1^\top \mathbf{X}_2$
3. Repeat until convergence:

$$\hat{\mathbf{U}} = \underset{\mathbf{U} \in \operatorname{conv} \mathcal{V}_{n_1 \times K}^{\mathbf{S}_1}}{\arg\min} - \operatorname{Tr}(\mathbf{U}^\top \mathbf{X}_1^\top \mathbf{X}_2 \hat{\mathbf{V}}) + \lambda_1 P_1(\mathbf{U})$$

$$\hat{\mathbf{V}} = \underset{\mathbf{V} \in \operatorname{conv} \mathcal{V}_{n_2 \times K}^{\mathbf{S}_2}}{\arg\min} - \operatorname{Tr}(\mathbf{V}^\top \mathbf{X}_2^\top \mathbf{X}_1 \hat{\mathbf{U}}) + \lambda_2 P_2(\mathbf{V})$$

4. Return $\hat{\mathbf{U}}$ and $\hat{\mathbf{V}}$

Application of Multi-Rank Regularized PCA (Weylandt, 2019) to PLS

Algorithm: Manifold Optimization

Sparse Graph-Smooth PLS Algorithm

1. Construct Graph Laplacians $\mathbf{L}_1, \mathbf{L}_2$ and associated smoothing matrices $\mathbf{S}_1 = \mathbf{I}_{n_1} + \alpha_1 \mathbf{L}_1$, $\mathbf{S}_2 = \mathbf{I}_{n_2} + \alpha_2 \mathbf{L}_2$
2. Initialize $\hat{\mathbf{U}}, \hat{\mathbf{V}}$ to the leading K singular vectors of $\mathbf{X}_1^\top \mathbf{X}_2$
3. Repeat until convergence:

$$\hat{\mathbf{U}} = \underset{\mathbf{U} \in \operatorname{conv} \mathcal{V}^{\mathbf{S}_1}_{n_1 \times K}}{\arg\min} -\operatorname{Tr}(\mathbf{U}^\top \mathbf{X}_1^\top \mathbf{X}_2 \hat{\mathbf{V}}) + \lambda_1 P_1(\mathbf{U})$$

$$\hat{\mathbf{V}} = \underset{\mathbf{V} \in \operatorname{conv} \mathcal{V}^{\mathbf{S}_2}_{n_2 \times K}}{\arg\min} -\operatorname{Tr}(\mathbf{V}^\top \mathbf{X}_2^\top \mathbf{X}_1 \hat{\mathbf{U}}) + \lambda_2 P_2(\mathbf{V})$$

4. Return $\hat{\mathbf{U}}$ and $\hat{\mathbf{V}}$

Application of Multi-Rank Regularized PCA (Weylandt, 2019) to PLS

Solve $\hat{\mathbf{U}}, \hat{\mathbf{V}}$-subproblems via Manifold ADMM (Kovnatsky *et al.*, 2016)

Sparse Graph-Smooth PLS Algorithm (Greedy Variant for Large Graphs)

1. Construct Graph Laplacians and smoothing matrices with leading eigenvalues $\ell_1 = \lambda_{\max}(\mathsf{S}_1)$ and $\ell_2 = \lambda_{\max}(\mathsf{S}_2)$
2. Initialize $\mathbf{C}_1 := \mathsf{X}_1^\top \mathsf{X}_2$
3. For $k = 1, \dots K$:
   - 3.1 Initialize $\hat{\mathbf{u}}_k, \hat{\mathbf{v}}_k$ to the leading singular vectors of $\mathbf{C}_k$
   - 3.2 Repeat until convergence:
     - 3.2.1 iterate $\mathbf{u}$-update: $\hat{\mathbf{u}}_k \propto \text{prox}_{\frac{\lambda_1}{\ell_1} P_1(\cdot)}\left(\mathbf{u}_k + \ell_1^{-1}\left(\mathbf{C}_k\hat{\mathbf{v}}_k - \mathsf{S}_1\mathbf{u}_k\right)\right)$
     - 3.2.2 iterate $\mathbf{v}$-update: $\hat{\mathbf{v}}_k \propto \text{prox}_{\frac{\lambda_2}{\ell_2} P_2(\cdot)}\left(\mathbf{v}_k + \ell_2^{-1}\left(\mathbf{C}_k^\top\hat{\mathbf{u}}_k - \mathsf{S}_2\mathbf{v}_k\right)\right)$
   - 3.3 Set $\mathbf{C}_{k+1} := \mathbf{C}_k - \frac{\mathbf{C}_k\hat{\mathbf{v}}_k\hat{\mathbf{u}}_k^T\mathbf{C}_k}{\hat{\mathbf{u}}_k^T\mathbf{C}_k\hat{\mathbf{v}}_k}$
4. Return $\{\hat{\mathbf{u}}_k\}_{k=1}^K$ and $\{\hat{\mathbf{v}}_k\}_{k=1}^K$

Combines Rank-1 Sparse and Functional PCA (Allen and Weylandt, 2019) with Schur Deflation (Weylandt, 2019)

Regularity and Optimality Guarantees

The CONGA problem is well-posed

1. $|\mathbf{U}^*_{\cdot k}\|_{S_1}$ is either 0 or 1 for all k. Similarly for $\mathbf{V}^*$.
2. If $(\mathbf{U}^*, \mathbf{V}^*) \neq (0, 0)$, the SGPLS solution $(\mathbf{U}^*, \mathbf{V}^*)$ depends smoothly on all (non-zero) regularization parameters.

Theoretical Properties

Regularity and Optimality Guarantees

The CONGA problem is well-posed

1. $|\mathbf{U}^*_{\cdot k}\|_{S_1}$ is either 0 or 1 for all k. Similarly for $\mathbf{V}^*$.
2. If $(\mathbf{U}^*, \mathbf{V}^*) \neq (0, 0)$, the SGPLS solution $(\mathbf{U}^*, \mathbf{V}^*)$ depends smoothly on all (non-zero) regularization parameters.

Furthermore, the proposed algorithms are well-behaved:

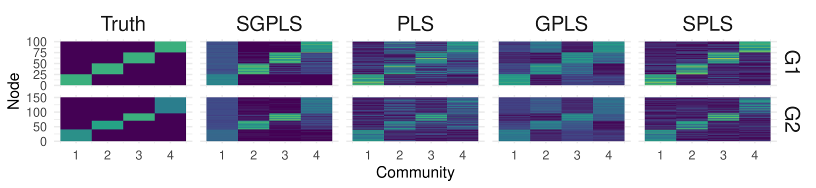
1. Step 3.2.1 of the greedy algorithm converges to a stationary point of

$$\underset{\mathbf{u}:\mathbf{u}^\top(\mathbf{I}+\alpha_1\mathbf{L}_1)\mathbf{u}\leq 1}{\arg\min} \frac{1}{2}\|\mathbf{C}_k\hat{\mathbf{v}}_k - \mathbf{u}\|_2^2 + \lambda_1 P_1(\mathbf{u}) + \frac{\alpha_1}{2}\mathbf{u}^\top\mathbf{L}_1\mathbf{u}$$

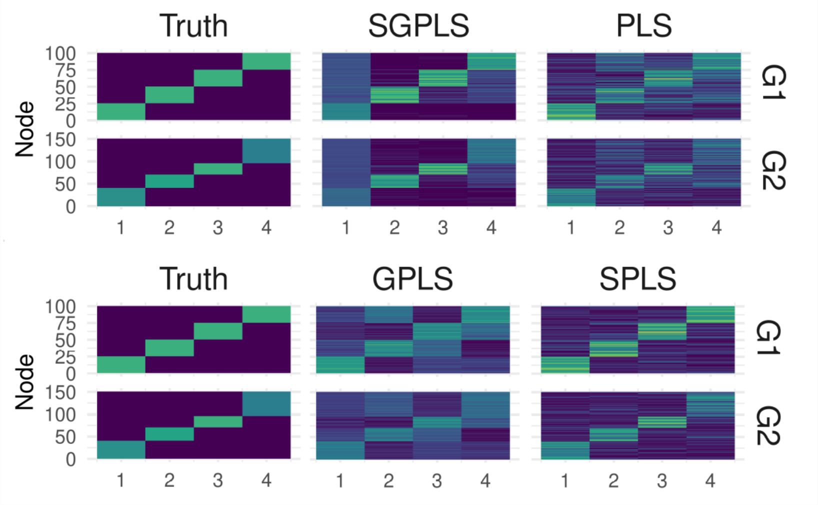
   Furthermore, if P_1 is convex, the convergence is monotone, at an $\mathcal{O}(1/K)$ rate, and to a global solution. Similar for Step 3.2.2
2. If P_1, P_2 are both convex, then $(\hat{\mathbf{u}}_k, \hat{\mathbf{v}}_k)$ returned by each step 3(b) of the greedy algorithm is both a coordinate-wise global maximum (Nash point) and a stationary point of SGPLS.

Simulation:

- $\mathcal{G}_1$: 4-block SBM (25, 25, 25, 25) with $p = 0.95, q = 0.2$
- $\mathcal{G}_2$: 4-block SBM (40, 30, 25, 55) with $p = 0.95, q = 0.2$
- m = 1000 signals: $\text{SNR}_{\mathcal{G}_1} = 0.2$, $\text{SNR}_{\mathcal{G}_2} = 0.163$
- Oracle tuning for all methods



CONGA via Sparse Graph PLS



Conclusions

Coarse Network Alignment:

Conclusions

Coarse Network Alignment:

- Community detection from multiple unaligned graphs
- Identifies matched communities based on common signals

Conclusions

Coarse Network Alignment:

- Community detection from multiple unaligned graphs
- Identifies matched communities based on common signals

Coarse Network Alignment: ArXiv 2104.02810

Thank you!

References

Allen, Genevera I. and Michael Weylandt (2019). "Sparse and Functional Principal Components Analysis". In: *DSW 2019: Proceedings of the 2nd IEEE Data Science Workshop*. Ed. by George Karypis, George Michailidis, and Rebecca Willett. Minneapolis, Minnesota: IEEE, pp. 11–16.

Kovnatsky, Artiom, Klaus Glashoff, and Michael M. Bronstein (2016). "MADMM: A Generic Algorithm for Non-Smooth Optimization on Manifolds". In: *ECCV 2016: Proceedings of the 14th European Conference on Computer Vision*. Ed. by Bastian Leibe, Jiri Matas, Nicu Sebe, and Max Welling. Vol. 9909. Lecture Notes in Computer Science. Springer, pp. 680–696.

Weylandt, Michael (2019). "Multi-Rank Sparse and Functional PCA: Manifold Optimization and Iterative Deflation Techniques". In: *CAMSAP 2019: Proceedings of the 8th IEEE Workshop on Computational Advances in Multi-Sensor Adaptive Processing*. Ed. by Geert Leus and Antonio G. Marques. Le Gosier, Guadaloupe, pp. 500–504.