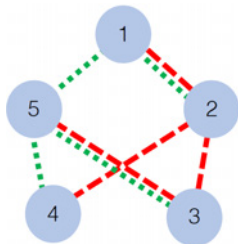# Multilayer Multimodal Network Inference

Alfred Hero

Michigan Institute for Data Science (MIDAS)
Dept of Electrical Engineering and Computer Science (EECS), Dept
of Biomedical Engineering (BME), Dept of Statistics
Program in Computational Medicine and Bioinformatics
Program in Applied and Interdisciplinary Mathematics
Program in Applied Physics
University of Michigan - Ann Arbor

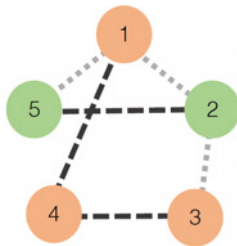July 4, 2018

Multilayer multimodal networks

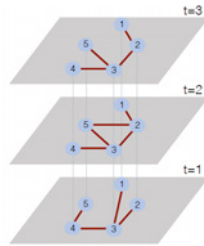Many edge types   Many vertex types   Dynamic edges
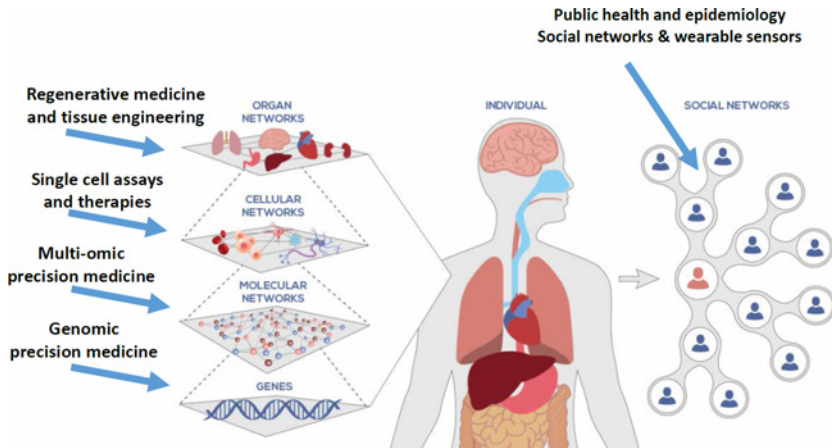

(a)   (b)   (c)

Multilayer multimodal networks

- Engineered multilayered communications networks (Salus and Vinton [1995], Addison-Wesley)
- Multiplex social networks (Verbrugge [1979], J. Social Forces)
- Multirelational terrorist networks (Carley [2005], Unpublished)
- Infection and prevention layers in epidemic networks (Jo et al. [2006], Physica A)

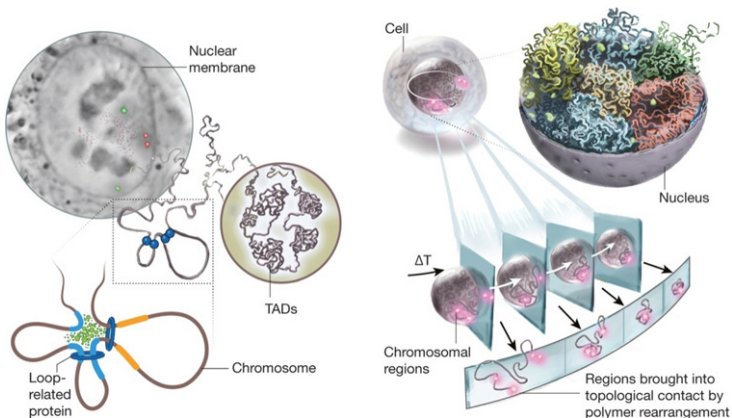## Integrative multiomic biology

Source: Institute for Systems Biology, Seattle

5D Nucleome

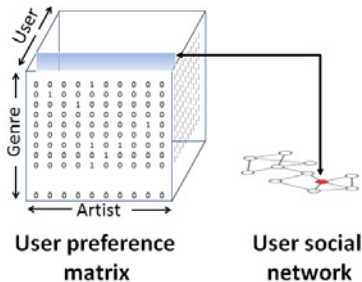The genome is contained in the cell nucleus and has five dimensions:

- form (spatial organization) - chromatin layer
- function (transcription) - gene expression layer
- time (growth/cell-cycle) - temporal layers

Source: 4D Nucleome (Dekker et al. [2017], Nature)

Network observation models

(Hsiao et al. [2014], ACM Data Mining)   (Chen et al. [2016a], Bioinformatics)
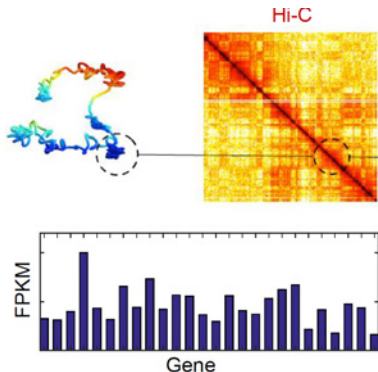
Two basic types of network data

- Edges are observed *relations* between vertex pairs
    - Friendships, kinships, email exchanges btwn agents in a social network
    - Contacts between genes in chromosomal regions from HiC
- Edges are observed *similarities* between attributes of vertex pairs
    - Inner product of pairs of preference vectors in a recommendation system
    - Correlation between mRNA expression levels of gene pairs from RNAseq

Network analysis ojectives
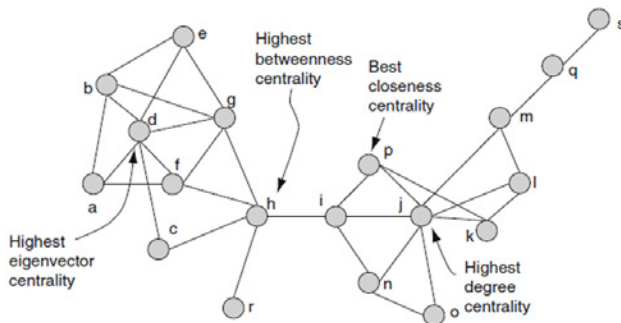
Objectives of analyst
- Find vertices of high centrality
  - Determine degree centrality, betweenness centrality, eigencentrality
- Detect clusters of tightly connected vertices
  - Extract connected componnts, perfrom deep community detection
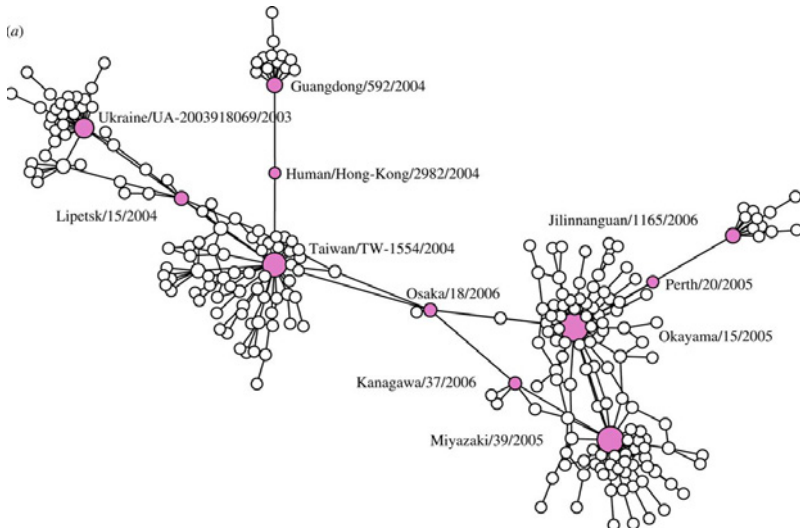- Perform statistical inference
  - Hypothesis testing, change detection, prediction

Source: Ortiz-Arroyo, 2010

Example: betweeness centrality for virus strain genotyping

Finding: High centrality nodes are virus strains that bridge between genotypes.

(a)

Guangdong/592/2004

Ukraine/UA-2003918069/2003

Human/Hong-Kong/2982/2004

Lipetsk/15/2004

Jilinnanguan/1165/2006

Taiwan/TW-1554/2004

Osaka/18/2006

Perth/20/2005

Okayama/15/2005

Kanagawa/37/2006

Miyazaki/39/2005

A. Wagner, "A genotype network reveals homoplastic cycles of convergent evolution in influenza A (H3N2) haemagglutinin," Proc. Royal Soc. B. May 2014.

Example: central vertex nomination for network disruption

Finding: Degree entropy of cancer PPI network predicts treatment resistance and 5 year cancer survivability.

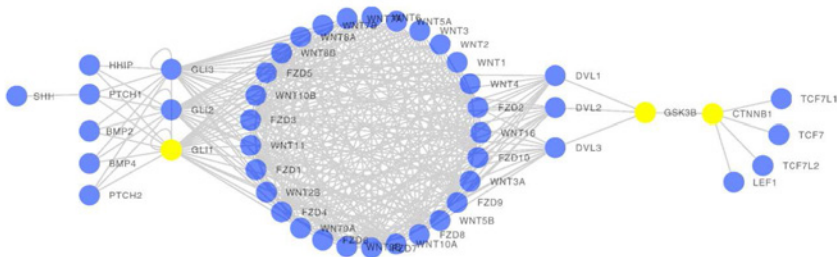Figure: Protein-Protein interaction network for basal cell carcinoma pathway as constructed from KEGG. Yellow nodes have highest betweenness centrality and can be targeted for disruption.

H. Breitkreutz, E. Rietman, J. Tuszynski (2012). "Molecular signaling network complexity is correlated with cancer patient survivability," *Proceedings of the National Academy of Sciences*, 109(23), 9209-9212.

Example: HiC network centrality analysis

Finding: Centrality of HiC gene-gene network related to RNAseq expression
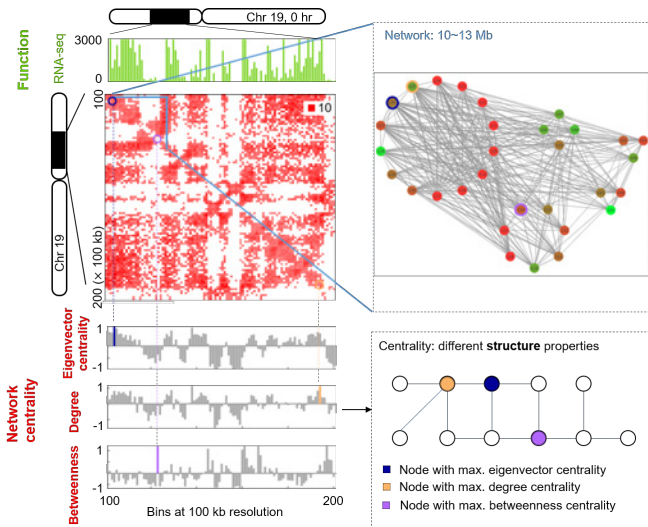
Liu et al. [2017], Genome architecture leads a bifurcation in cell identity, BioRxv

Example: HiC multiplex degree centrality over cell cycle
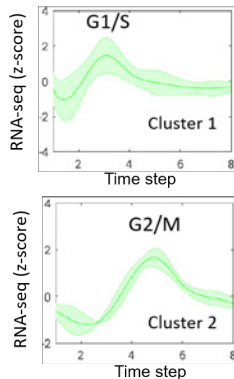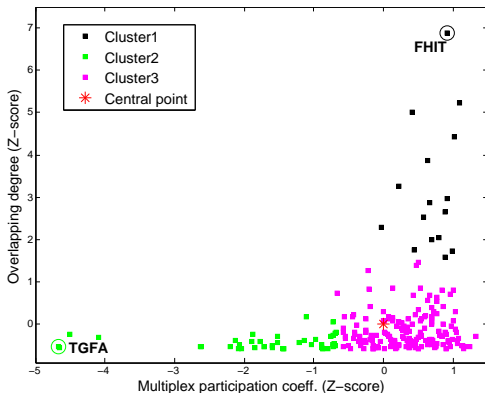
Three clusters of multiplex degree centrality (GO)

- Cluster 1 (FHIT): high centrality nodes governing pre-mitosis ($G1/S$)
- Cluster 2 (TGFA): lower centrality nodes governing mitosis ($G2/M$)

Oselio et al. [2018], Multilayer relevance networks, IEEE SPAWC

HiC degree-centrality trajectories
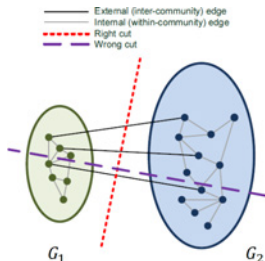
Oselio et al. [2018], Multilayer relevance networks, IEEE SPAWC

Community detection methods

Adjacency matrix partitioning

$$\mathbf{A} = \begin{array}{c} n_1 \\ n_2 \end{array} \begin{pmatrix} \mathbf{A}_1 & \mathbf{C} \\ \mathbf{C}^T & \mathbf{A}_2 \end{pmatrix}$$
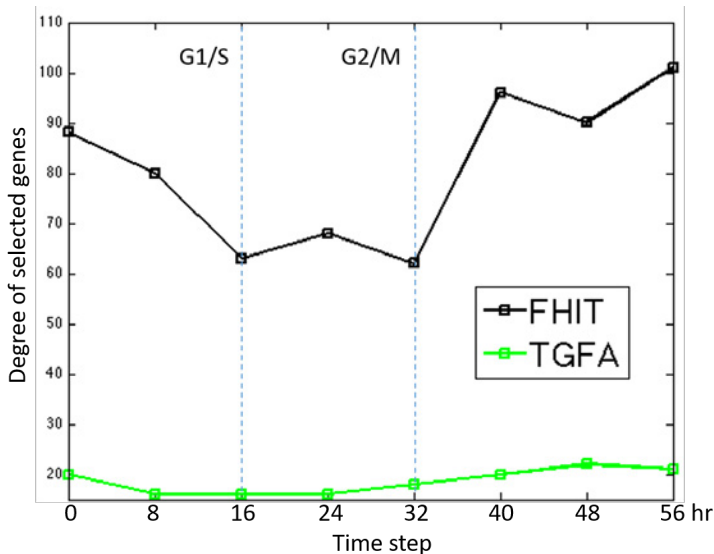
Main approaches

- Combinatorial: MINcut, MAXflow, normalized cut, cut ratio (Cheng& Wei [1991], Shi&Malik [2001])

- Parametric: Stochastic Block Models, Back Propagation (Hastings [2006], Airoldi [2008], Decelle [2011], Bickel [2012], Zhang [2014])

- Data driven: spectral methods, dendogram splitting, Louvain merging (Fiedler [1973], Ding [2001], Girvan [2002], Blondel [2008])

Community detection methods
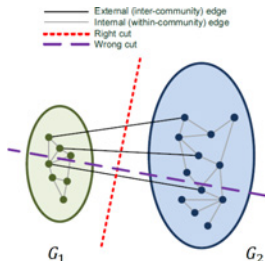
Adjacency matrix partitioning

$$\mathbf{A} = \begin{array}{c} n_1 \\ n_2 \end{array} \begin{pmatrix} \mathbf{A}_1 & \mathbf{C} \\ \mathbf{C}^T & \mathbf{A}_2 \end{pmatrix}$$

Main approaches

- Combinatorial: MINcut, MAXflow, normalized cut, cut ratio (Cheng& Wei [1991], Shi&Malik [2001])
- Parametric: Stochastic Block Models, Back Propagation (Hastings [2006], Airoldi [2008], Decelle [2011], Bickel [2012], Zhang [2014])
- Data driven: spectral methods, dendogram splitting, Louvain merging (Fiedler [1973], Ding [2001], Girvan [2002], Blondel [2008])

$\Rightarrow$ Principal questions: robustness to noise, finding #communities, prediction of detection accuracy

Chen and Hero [2018], Phase transitions and a model order selection ..., IEEE TSP

Detecting communities with noisy edges
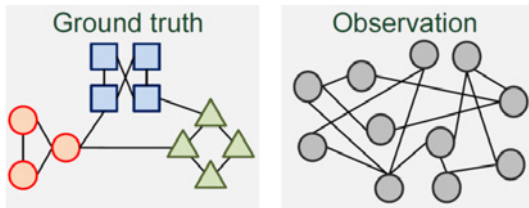
Random interconnection model (RIM): $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, $|V| = p$, $E = \{e_{ij}\}_{i>j}$

Ground truth   Observation

Detecting communities with noisy edges

Random interconnection model (RIM): $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, $|V| = p$, $E = \{e_{ij}\}_{i>j}$

Ground truth   Observation

- Binary $p \times p$ adjacency matrix $\mathbf{A}$ is corrupted by spurious edges $\mathbf{C}_{ij}$

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{C}_{12} & \mathbf{C}_{13} \\ \mathbf{C}_{21} & \mathbf{A}_2 & \mathbf{C}_{23} \\ \mathbf{C}_{31} & \mathbf{C}_{32} & \mathbf{A}_3 \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{A}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{A}_3 \end{bmatrix}}_{\mathbf{A}_0} + \underbrace{\begin{bmatrix} \mathbf{0} & \mathbf{C}_{12} & \mathbf{C}_{13} \\ \mathbf{C}_{21} & \mathbf{0} & \mathbf{C}_{23} \\ \mathbf{C}_{31} & \mathbf{C}_{32} & \mathbf{0} \end{bmatrix}}_{\mathbf{\Delta}}$$
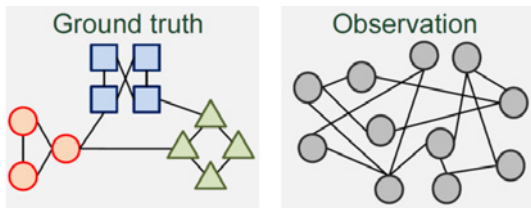
Detecting communities with noisy edges

Random interconnection model (RIM): $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, $|V| = p$, $E = \{e_{ij}\}_{i>j}$
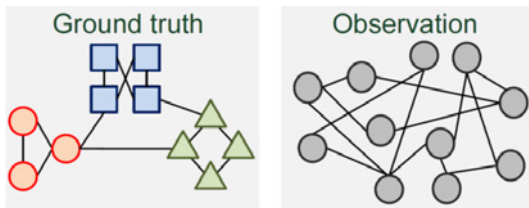

Ground truth   Observation

- Binary $p \times p$ adjacency matrix **A** is corrupted by spurious edges $\mathbf{C}_{ij}$

$$\mathbf{A} = \left[\begin{array}{ccc} \mathbf{A}_1 & \mathbf{C}_{12} & \mathbf{C}_{13} \\ \mathbf{C}_{21} & \mathbf{A}_2 & \mathbf{C}_{23} \\ \mathbf{C}_{31} & \mathbf{C}_{32} & \mathbf{A}_3 \end{array}\right] = \underbrace{\left[\begin{array}{ccc} \mathbf{A}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{A}_3 \end{array}\right]}_{\mathbf{A}_0} + \underbrace{\left[\begin{array}{ccc} \mathbf{0} & \mathbf{C}_{12} & \mathbf{C}_{13} \\ \mathbf{C}_{21} & \mathbf{0} & \mathbf{C}_{23} \\ \mathbf{C}_{31} & \mathbf{C}_{32} & \mathbf{0} \end{array}\right]}_{\mathbf{\Delta}}$$

- Ground truth graph has very sparse $\mathbf{C}_{ij}$'s: $\mathbf{A}_0$ (nearly) block diagonal
  $\Rightarrow$ Can perfectly recover communities w.h.p.
- Observed graph has noisy edges: entries of $\mathbf{C}_{ij}$'s are random
  $\Rightarrow$ Cannot reliably recover communities

Chen and Hero [2018], Phase transitions and a model order selection . . . , IEEE TSP

Spectral graph clustering: algebraic connectivity methods

- Fiedler's spectral partitioning method (Fiedler [1973], Czechoslovak Mathematical Journal)
    - Compute graph Laplacian $\mathbf{L} = \mathbf{D} - \mathbf{A}$
    - Find second smallest eigenvalue λ_2 and eigenvector $\mathbf{y} = \mathbf{y}_{2nd}$ of $\mathbf{L}$
    - Perform K-means clustering on Fiedler vector $\mathbf{y}$ with $K = 2$ groups
    - Re-partition sub-clusters until suffiicent #communities found

Spectral graph clustering: algebraic connectivity methods
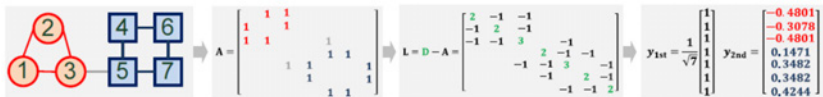
- Fiedler's spectral partitioning method (Fiedler [1973], Czechoslovak Mathematical Journal)
  - Compute graph Laplacian $\mathbf{L} = \mathbf{D} - \mathbf{A}$
  - Find second smallest eigenvalue λ_2 and eigenvector $\mathbf{y} = \mathbf{y}_{2nd}$ of $\mathbf{L}$
  - Perform K-means clustering on Fiedler vector $\mathbf{y}$ with $K = 2$ groups
  - Re-partition sub-clusters until suffiicent #communities found
- Fiedler partitioning gives balanced mincut when communities well separated (Ding et al. [2005], SIAM Data Mining)

Spectral graph clustering: algebraic connectivity methods

- Fiedler's spectral partitioning method (Fiedler [1973], Czechoslovak Mathematical Journal)
  - Compute graph Laplacian $\mathbf{L} = \mathbf{D} - \mathbf{A}$
  - Find second smallest eigenvalue λ_2 and eigenvector $\mathbf{y} = \mathbf{y}_{2nd}$ of $\mathbf{L}$
  - Perform K-means clustering on Fiedler vector $\mathbf{y}$ with $K = 2$ groups
  - Re-partition sub-clusters until suffiicent #communities found
- Fiedler partitioning gives balanced mincut when communities well separated (Ding et al. [2005], SIAM Data Mining)
- Other related approaches
  - Modularity method (Newman [2006], PNAS)
  - Non-backtracking method (Krzakala et al. [2013], PNAS)
  - Bethe free energy method (Saade et al. [2014], NIPS)

## Spectral graph clustering: algebraic connectivity methods

- Fiedler's spectral partitioning method (Fiedler [1973], Czechoslovak Mathematical Journal)
    - Compute graph Laplacian $\mathbf{L} = \mathbf{D} - \mathbf{A}$
    - Find second smallest eigenvalue λ_2 and eigenvector $\mathbf{y} = \mathbf{y}_{2nd}$ of $\mathbf{L}$
    - Perform K-means clustering on Fiedler vector $\mathbf{y}$ with $K = 2$ groups
    - Re-partition sub-clusters until suffiicent #communities found
- Fiedler partitioning gives balanced mincut when communities well separated (Ding et al. [2005], SIAM Data Mining)
- Other related approaches
    - Modularity method (Newman [2006], PNAS)
    - Non-backtracking method (Krzakala et al. [2013], PNAS)
    - Bethe free energy method (Saade et al. [2014], NIPS)

All of these methods are sensitive to spurious edges

Phase transitions: graph extraction from thresholded sample correlations

Graph $\mathcal{G}$ obtained by thresholding $p \times p$ sample correlation matrix **R** at level $\rho \in [0, 1]$
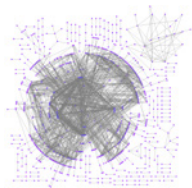
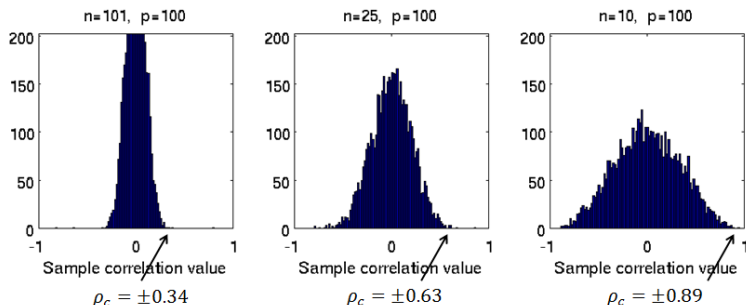Number of degree d (hub) nodes in G: $N_{d,\rho} = \sum_{i=1}^{p} I(d_i \geq d)$

$$I(d_i \geq d) = \left\{ \begin{array}{ll} 1, & \mathrm{card}\{j : j \neq i, |\mathbf{R}_{ij}| \geq \rho\} \geq d \\ 0, & o.w. \end{array} \right.$$

where

$$\mathbf{R} = \mathrm{diag}(\hat{\boldsymbol{\Sigma}})^{-1/2} \hat{\boldsymbol{\Sigma}} \mathrm{diag}(\hat{\boldsymbol{\Sigma}})^{-1/2}$$

is sample correlation matrix based on n i.i.d. sample replicates.

16

Phase transition in number of false edges and hubs

- p nodes, n samples, ρ threshold on sample correlation graph

- Critical threshold ρ_c on false positive detections of degree centrality $\geq \delta$

$$\rho_c(\delta) = \sqrt{1 - c_{\delta,n}(p-1)^{-2\delta/\delta(n-2)-2}}$$

- $c_{n,\delta} = O(n^{-3/2})$ is only weakly dependent on correlation if ~~block~~ sparse

- Hero and Rajaratnam, "Large scale correlation screening," J. Am Stat, Assoc., 2011

- Hero and Rajaratnam, "Hub screening in partial correlation graphs," IEEE Trans. on Info Theory, 2012

- Wei, Rajaratnam and Hero, "Correlation screening with relaxed sparsity," in preparation, 2018
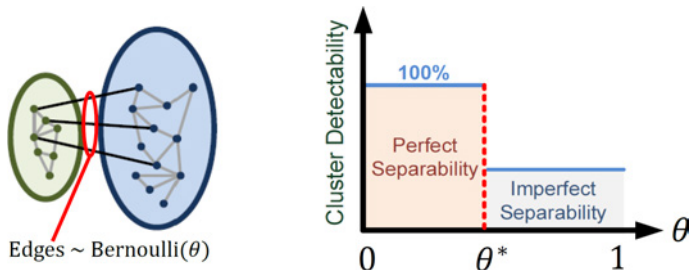
Extraction of false hubs: critical threshold ρ_c as function of n
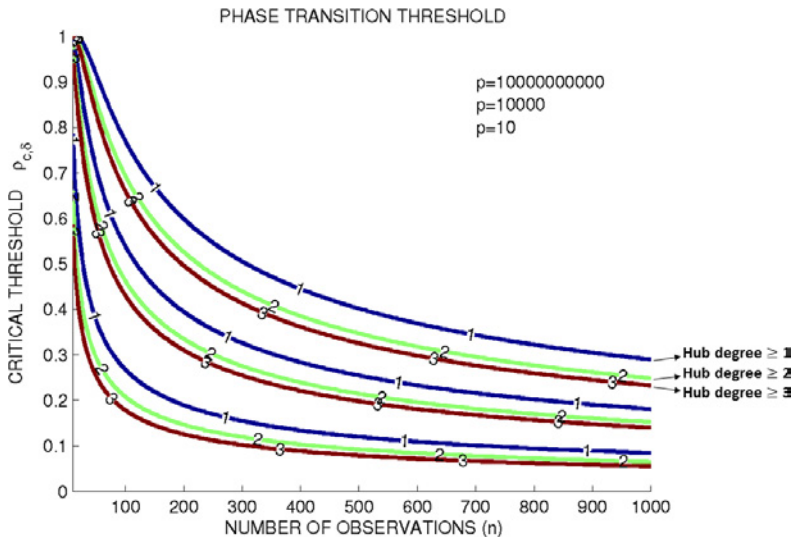
Extraction of false hubs: critical threshold ρ_c as function of n
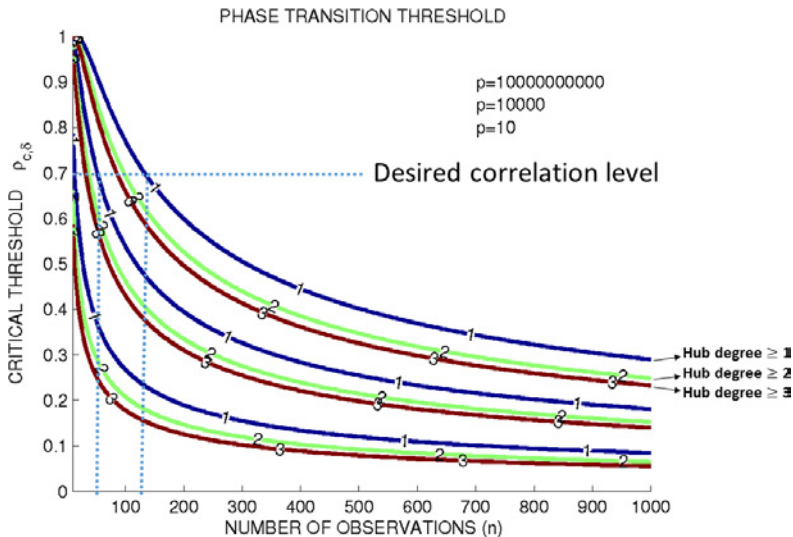
Phase transitions in community detection from noisy graph

Edges ~ Bernoulli(θ)

Theorem (Existence of phase transitions)

Let $\mathcal{G}_0$ be a graph with $p \times p$ block diagonal adjacency matrix $\mathbf{A}_0$ and let $\mathcal{G}$ be a graph with random adjacency matrix $\mathbf{A} = \mathbf{A}_0 \oplus \mathbf{\Delta}$ where $\mathbf{\Delta}$ is a random binary matrix, independent of $\mathbf{A}_0$, with Bernoulli(θ) entries.

Then, as $p \to \infty$, community detection algorithms exhibit a phase transition threshold with critical phase transition parameter θ^*: the block diagonal structure of $\mathbf{A}_0$ is recovered if $\theta < \theta^*$ while it is not recovered if $\theta > \theta^*$.

(Chen and Hero [2015b], Phys Rev E), (Chen and Hero [2015a], IEEE TSP), (Abbe and Sandon [2015], FOCS), (Nadakuditi and Newman [2012], Phys Rev)

Example: spectral graph clustering in stochastic block model

Single community in additive noise
2×2 connection probability matrix

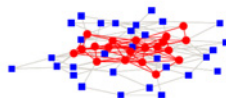$$\boldsymbol{\Theta} = \left[\begin{array}{cc} \theta_{in} & \theta_{out} \\ \theta_{out} & \theta_{out} \end{array} \right]$$
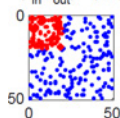
$p \times p$ graph Laplacian matrix,
$p = p_{in} + p_{out}$

$$\mathbf{L} = \left[\begin{array}{cc} \mathbf{L}_{in} + \mathbf{D}_{in} & -\mathbf{C} \\ -\mathbf{C} & \mathbf{L}_{out} \end{array} \right]$$

Fiedler vector $\mathbf{y} = \left[\begin{array}{c} \mathbf{y}_{in} \\ \mathbf{y}_{out} \end{array} \right]$ of $\mathbf{L}$

- c_{in}: avg within-community degree
- c_{out}: avg btwn-community degree

Example: spectral graph clustering in stochastic block model

Single community in additive noise

2×2 connection probability matrix

$$\mathbf{\Theta} = \left[\begin{array}{cc} \theta_{in} & \theta_{out} \\ \theta_{out} & \theta_{out} \end{array} \right]$$

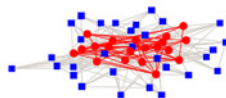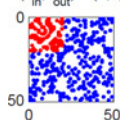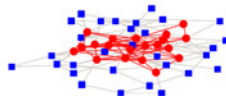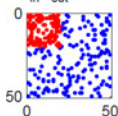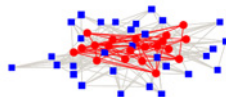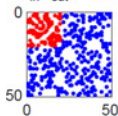$p \times p$ graph Laplacian matrix, $p = p_{in} + p_{out}$

$$\mathbf{L} = \left[\begin{array}{cc} \mathbf{L}_{in} + \mathbf{D}_{in} & -\mathbf{C} \\ -\mathbf{C} & \mathbf{L}_{out} \end{array} \right]$$

Fiedler vector $\mathbf{y} = \left[\begin{array}{c} \mathbf{y}_{in} \\ \mathbf{y}_{out} \end{array} \right]$ of $\mathbf{L}$

$(c_{in}, c_{out}) = (5,2)$

$(c_{in}, c_{out}) = (5,4)$

- c_{in}: avg within-community degree
- c_{out}: avg btwn-community degree

Theorem (Spectral clustering phase transition)

For fixed θ_{in} there exists an asymptotic threshold θ_{out}^* such that (a.s.)

$$\left\{ \begin{array}{ll} \sqrt{\frac{pp_{in}}{p_{out}}} \mathbf{y}_{in} \to \pm \mathbf{1}_{in}, \ \sqrt{\frac{pp_{in}}{p_{out}}} \mathbf{y}_{out} \to \pm \mathbf{1}_{out} & if \theta_{out} < \theta_{out}^* \\ \mathbf{1}_{in} \mathbf{y}_{in} \to 0, \ \mathbf{1}_{out} \mathbf{y}_{out} \to 0 & if \theta_{out} > \theta_{out}^* \end{array} \right.$$

as $p_{in} \to \infty$, $p_{out} \to \infty$ and $p_{in}/p_{out} \to c > 0$.

Deep community detection via iterative pruning

Algebraic connectivity (Fiedler eigenvalue) depends on $\mathcal{E}$ and $\mathcal{V}$

$$f(\mathcal{E}, \mathcal{V}) = \lambda_2(\mathcal{E}, \mathcal{V})$$

Main idea: denoise community structure by optimal pruning

q-**edge pruning** $\qquad\qquad\qquad\qquad$ q-**node pruning**

$\mathcal{E}_q = \operatorname{argmin}_{e_1,\ldots,e_q \in \mathcal{E}} f(\mathcal{E}, \mathcal{V})$ $\qquad\qquad$ $\mathcal{V}_q = \operatorname{argmin}_{v_1,\ldots,v_q \in \mathcal{V}} f(\mathcal{E}, \mathcal{V})$

Deep community detection via iterative pruning

Algebraic connectivity (Fiedler eigenvalue) depends on $\mathcal{E}$ and $\mathcal{V}$

$$f(\mathcal{E}, \mathcal{V}) = \lambda_2(\mathcal{E}, \mathcal{V})$$

Main idea: denoise community structure by optimal pruning

q-**edge pruning** q-**node pruning**
$$\mathcal{E}_q = \mathrm{argmin}_{e_1, \ldots, e_q \in \mathcal{E}} f(\mathcal{E}, \mathcal{V}) \qquad\qquad \mathcal{V}_q = \mathrm{argmin}_{v_1, \ldots, v_q \in \mathcal{V}} f(\mathcal{E}, \mathcal{V})$$

Greedy approximation: prune one at a time.
For $k = 1, \ldots, q$

Greedy edge pruning   **Greedy node pruning**
$$e^{(k)} = \mathrm{argmin}_{e \in \mathcal{E}^{(k)}} f(\mathcal{E}^{(k)}, \mathcal{V}) \qquad\qquad v^{(k)} = \mathrm{argmin}_{v \in \mathcal{V}^{(k)}} f(\mathcal{E}^{(k)}, \mathcal{V}^k))$$

Deep community detection via iterative pruning

Algebraic connectivity (Fiedler eigenvalue) depends on $\mathcal{E}$ and $\mathcal{V}$

$$f(\mathcal{E}, \mathcal{V}) = \lambda_2(\mathcal{E}, \mathcal{V})$$

Main idea: denoise community structure by optimal pruning

$$\begin{array}{cc}
\textbf{q-edge pruning} & \textbf{q-node pruning} \\
\mathcal{E}_q = \operatorname{argmin}_{e_1,\ldots,e_q \in \mathcal{E}} f(\mathcal{E}, \mathcal{V}) & \mathcal{V}_q = \operatorname{argmin}_{v_1,\ldots,v_q \in \mathcal{V}} f(\mathcal{E}, \mathcal{V})
\end{array}$$

Greedy approximation: prune one at a time.
For $k = 1, \ldots, q$

$$\begin{array}{cc}
\textbf{Greedy edge pruning} & \textbf{Greedy node pruning} \\
e^{(k)} = \operatorname{argmin}_{e \in \mathcal{E}^{(k)}} f(\mathcal{E}^{(k)}, \mathcal{V}) & v^{(k)} = \operatorname{argmin}_{v \in \mathcal{V}^{(k)}} f(\mathcal{E}^{(k)}, \mathcal{V}^k))
\end{array}$$

Fact: greedy pruning minimizes Local Fiedler Value Centrality (LFVC)

$$\begin{array}{cc}
\textbf{Edge-LFVC} & \textbf{Node-LFVC} \\
\Delta_{e_{ij}} f = (y_i - y_j)^2 & \Delta_{v_i} f = \sum_{j \in \mathcal{N}_i} (y_i - y_j)^2
\end{array}$$

Deep community detection via iterative pruning

Algebraic connectivity (Fiedler eigenvalue) depends on $\mathcal{E}$ and $\mathcal{V}$

$$f(\mathcal{E}, \mathcal{V}) = \lambda_2(\mathcal{E}, \mathcal{V})$$

Main idea: denoise community structure by optimal pruning

$$\begin{array}{cc}
\textbf{q-edge pruning} & \textbf{q-node pruning} \\
\mathcal{E}_q = \operatorname{argmin}_{e_1,\ldots,e_q \in \mathcal{E}} f(\mathcal{E}, \mathcal{V}) & \mathcal{V}_q = \operatorname{argmin}_{v_1,\ldots,v_q \in \mathcal{V}} f(\mathcal{E}, \mathcal{V})
\end{array}$$

Greedy approximation: prune one at a time.
For $k = 1, \ldots, q$

$$\begin{array}{cc}
\textbf{Greedy edge pruning} & \textbf{Greedy node pruning} \\
e^{(k)} = \operatorname{argmin}_{e \in \mathcal{E}^{(k)}} f(\mathcal{E}^{(k)}, \mathcal{V}) & v^{(k)} = \operatorname{argmin}_{v \in \mathcal{V}^{(k)}} f(\mathcal{E}^{(k)}, \mathcal{V}^k))
\end{array}$$

Fact: greedy pruning minimizes Local Fiedler Value Centrality (LFVC)

$$\begin{array}{cc}
\textbf{Edge-LFVC} & \textbf{Node-LFVC} \\
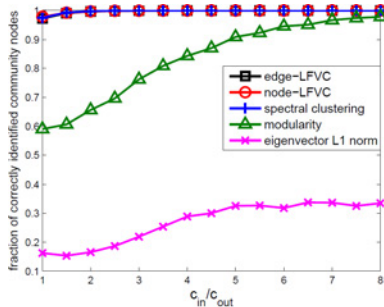\Delta_{e_{ij}} f = (y_i - y_j)^2 & \Delta_{v_i} f = \sum_{j \in \mathcal{N}_i} (y_i - y_j)^2
\end{array}$$

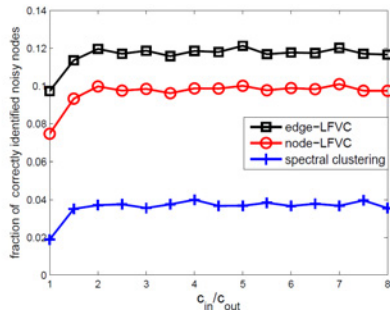Property: $f(\mathcal{E}, \mathcal{V})$ is *submodular* in $\mathcal{E}$ and $\mathcal{V}$

Implication: Greedy pruning comes within $(1 - 1/e)$ of optimal pruning

Chen and Hero [2015a], Deep community detection, IEEE TSP

## Deep Community Detection Numerical Comparisons

Sensitivity

Specificity

- Single-community stochastic block model with $p = p_{in} + p_{out} = 200$ nodes
- Community size: $p_{in} = 40$ nodes
- Stochastic block model probability parameters:
  - $\theta_{out} = 1/80$ ($c_{out} = 2$)
  - θ_{in} (c_{in}) varies over a range
- Curves represent averages of 100 trials

Chen and Hero [2015a], Deep community detection, IEEE TSP

23

Deep Community Detection application: `last.fm` social network
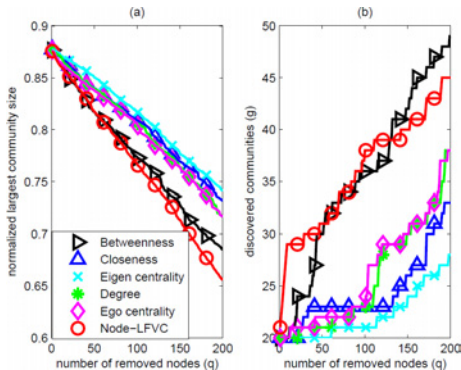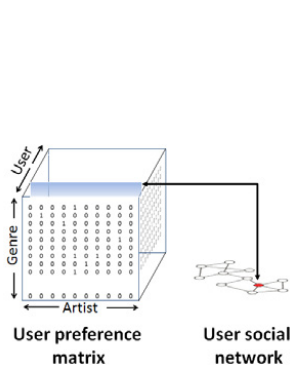
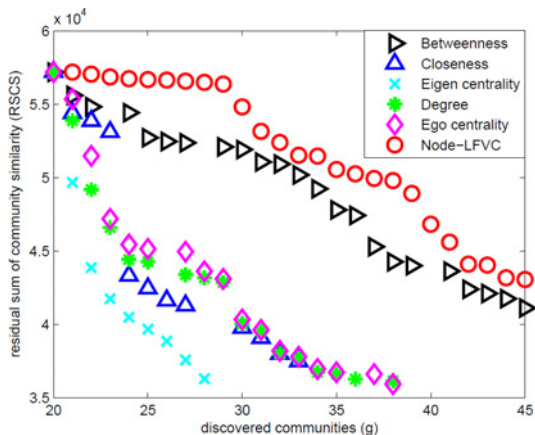Deep community detection applied to `last.fm` user social network.
Left: Decrease in algebraic connectivity as function of # nodes removed.
Right: Increase in number of discovered communities.

- Node removals selected by greedy maximization of LFVC criterion
- $p = 1843$ nodes
- $|V| = 12668$ edges

Chen and Hero [2015a], Deep community detection, IEEE TSP

Deep Community Detection application: `last.fm` social network

Residual sum of community similarity: normalized sum of correlations of `last.fm` pairwise preference in each discovered community of `last.fm` user social network

Chen and Hero [2015a], Deep community detection, IEEE TSP

Community Detection with Adaptive Model Order Selection (AMOS)

- General phase transition results for K communities have been derived
- Bounds on critical phase transition threshold θ^* can often be found

$$\theta_{LB} \leq \theta^* \leq \theta_{UB}$$

- Community detection with adaptive estimation of K
    1. Perform community detection with a hypothesized value $\hat{K}$ of K
    2. Construct statistical $(1 - \alpha) \times 100\%$ confidence interval on θ
    3. If confidence interval on θ falls above θ_{UB} then reject $\hat{K}$

Community Detection with Adaptive Model Order Selection (AMOS)

- General phase transition results for K communities have been derived
- Bounds on critical phase transition threshold θ^* can often be found

$$\theta_{LB} \leq \theta^* \leq \theta_{UB}$$

- Community detection with adaptive estimation of K
  1. Perform community detection with a hypothesized value $\hat{K}$ of K
  2. Construct statistical $(1 - \alpha) \times 100\%$ confidence interval on θ
  3. If confidence interval on θ falls above θ_{UB} then reject $\hat{K}$
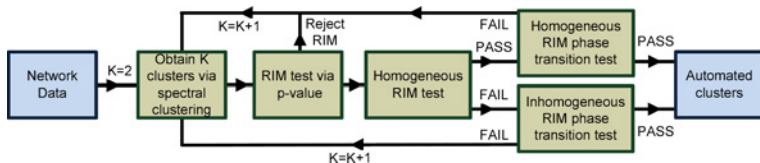
Chen and Hero [2018], Phase transions and a model selection..., IEEE TSP

Centrality and community detection in multilayer networks

Focus on L-layer multiplex graphs $\mathcal{G} = \{\mathcal{E}, \mathcal{V}\}$:

- layers that share the same nodes $\mathcal{V}$
- L sets of edges $\mathcal{E} = \mathcal{E}_1 \times \ldots \times \mathcal{E}_L$
- A tensor-valued adjacency matrix $A = A^{(1)} \times \cdots \times A^{(L)}$

Centrality and community detection in multilayer networks

Focus on L-layer multiplex graphs $\mathcal{G} = \{\mathcal{E}, \mathcal{V}\}$:

- layers that share the same nodes $\mathcal{V}$
- L sets of edges $\mathcal{E} = \mathcal{E}_1 \times \ldots \times \mathcal{E}_L$
- A tensor-valued adjacency matrix $A = A^{(1)} \times \cdots \times A^{(L)}$

Methods for integrative multiplex graph analysis (Kivelä et al. [2014])

- Scalar network aggregation
  - Multilayer majority centrality aggregation (Bródka et al. [2011], CASoN)
  - Convex layer aggregation (Chen and Hero [2017], IEEE T SIPN)

- Spectral decompoisition
  - Multicentrality spectral decomposition (Chen et al. [2016b], IEEE ICASSP)
  - Multilayer information diffusion (Mahdizadehaghdam et al. [2016], IEEE T SIPN)

- Latent variable relevance networks
  - MCMC Dynamic Stochastic Block Models (DSBM) (Yang et al. [2011], Machine Learning), (Xu and Hero [2014], IEEE JSTSP)
  - Bayesian model averaging of DSBM (Oselio et al. [2014], IEEE JSTSP)

Multilayer iterative model order selection algorithm (MIMOSA)

Assume L layers have common community structure

$$\mathbf{A}^{(l)} = \left[\begin{array}{ccc} \mathbf{A}_1^{(l)} & \mathbf{C}_{12}^{(l)} & \mathbf{C}_{13}^{(l)} \\ \mathbf{C}_{21}^{(l)} & \mathbf{A}_2^{(l)} & \mathbf{C}_{23}^{(l)} \\ \mathbf{C}_{31}^{(l)} & \mathbf{C}_{32}^{(l)} & \mathbf{A}_3^{(l)} \end{array} \right], \qquad l = 1, \ldots, L$$

MIMOSA applies AMOS to *convex combination* of layers

- $\mathbf{w} = [w_1, \ldots, w_L]$ a set of weights on the simplex $\sum_{l=1}^{L} w_l = 1$, $w_i \geq 0$
- Compute the convex combination $\overline{\mathbf{A}} = \sum_{l=1}^{L} w_l \mathbf{A}^{(l)}$
- Apply AMOS spectral clustering to the matrix $\overline{\mathbf{A}}$

Properties:

- MIMOSA is simple to implement: it flattens multilayer structure
- Phase transition analysis is straightforward
- Analysis suggests automated weight selection procedure

Chen and Hero [2017], Multilayer spectral graph clustering. . . , IEEE T SIPN

MIMOSA phase transition analysis

Theorem

Let $\mathcal{G}_0$ be a L-layer SBM whose layers have block diagonal adjacency matrices $\{\mathbf{A}_0^{(l)}\}_{l=1}^{L}$, assumed independent. Let the l-th layer of observed SBM have layer adjacency matrix $\mathbf{A}^{(l)} = \mathbf{A}_0^{(l)} \oplus \mathbf{\Delta}^{(l)}$, where $\mathbf{\Delta}^{(l)}$ is a random binary matrix with Bernoulli(θ) entries.

Then, as $p \to \infty$, the MIMOSA spectral clustering algorithm has a critical phase transition with parameter $\theta^*(\mathbf{w})$ such that any common community to all layers is recovered if $\theta < \theta^*(\mathbf{w})$ while it is not recovered if $\theta > \theta^*(\mathbf{w})$.

An optimal set of weights can be estimated that provides the most favorable value of $\theta(\mathbf{w})$
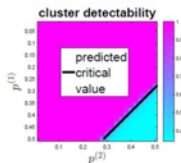
Chen and Hero [2017], Multilayer spectral graph clustering. . . , IEEE T SIPN
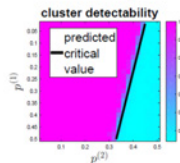
MIMOSA phase transition results

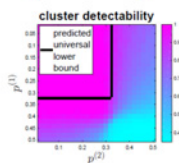- Phase transition regions on noise params ($K = 2$)
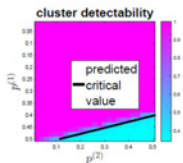
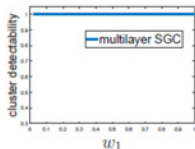$(w_1, w_2) = (0.8, 0.2)$, $(0.5, 0.5)$, $(0.2, 0.0.8)$,   Convex hull
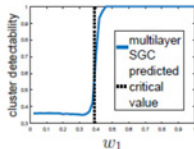
- Detectability regions of 2 layer 1000 node network: $\theta_{11} = 0.3$, $\theta_{12} = 0.2$, $\theta_{22} = 0.4$
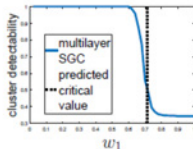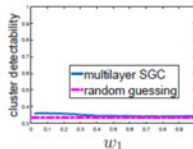
- Value of phase transition is accurately predicted

$(p^1, p^2) = (0.2, 0.2)$, $(0.2, 0.5)$, $(0.5, 0.2)$, $(0.5, 0.5)$

30

MIMOSA experimental results

Dataset	Method	K	NMI	RI	F-measure	conductance	NC
Human H1V1 genetic interaction	MIMOSA	2	-	-	-	**0.0346**	**0.0666**
	MIMOSA-uniform	2	-	-	-	0.0346	0.0666
	GenLouvain ($\gamma = 0.5$)	4	-	-	-	0.1822	0.2292
	GenLouvain ($\gamma = 1$)	4	-	-	-	0.1822	0.2292
	GenLouvain ($\gamma = 2$)	5	-	-	-	0.1458	0.3167
	SC-ML	2	-	-	-	0.1161	0.2027
	Self-Tuning	7	-	-	-	0.5627	0.8722
Pierre Auger coauthorship	MIMOSA	2	-	-	-	**0.0113**	**0.1888**
	MIMOSA-uniform	NA	-	-	-	NA	NA
	GenLouvain ($\gamma = 0.5$)	9	-	-	-	1.5207	1.8423
	GenLouvain ($\gamma = 1$)	13	-	-	-	1.2655	1.4699
	GenLouvain ($\gamma = 2$)	61	-	-	-	0.5717	0.6356
	SC-ML	2	-	-	-	1.2939	2.5181
	Self-Tuning	63	-	-	-	0.8400	0.9321

- Human HIV-1: GGI net separated into 5 layers of association - PhAssoc, DirInt, Coloc, Assoc, SupprGen (De Domenico [2014], J. Complex Networks)
- Paul Auger: cosmologist co-author network separated into 9 research topic layers (De Domenico [2015], Phys Rev X)

Comparative algorithms

- MIMOSA uniform - MIMOSA with uniform weights
- Louvain(γ) - maximizes multilayer modularity (Mucha et al. [2010], Science)
- SC-ML - multilayer spectral clustering (Dong et al. [2012], IEEE TSP)
- Self-tuning - automated model selection SGC algorithm (Zelnik-Manor and Perona [2005], NIPS)

MIMOSA experimental results

Dataset	Method	K	NMI	RI	F-measure	conductance	NC
109th Congress votes - Budget	MIMOSA	2	0.7959	0.9224	0.9220	0.2713	0.4975
	MIMOSA-uniform	2	**0.8778**	**0.9604**	**0.9603**	0.2702	0.5055
	GenLouvain ($\gamma = 0.5$)	2	0.7959	0.9224	0.9220	0.2713	0.4978
	GenLouvain ($\gamma = 1$)	2	0.7959	0.9224	0.9220	0.2713	0.4978
	GenLouvain ($\gamma = 2$)	55	0.3822	0.6915	0.5539	**0.1500**	**0.1959**
	SC-ML	2	0.7610	0.9040	0.9036	0.2742	0.5089
	Self-Tuning	3	0.8488	0.9164	0.9087	1.5046	1.8011
109th Congress votes - Energy	MIMOSA	2	**0.7290**	**0.8861**	**0.8855**	**0.1151**	**0.2086**
	MIMOSA-uniform	2	0.6716	0.8513	0.8508	0.1154	0.2178
	GenLouvain ($\gamma = 0.5$)	2	0.5403	0.8182	0.8173	**0.1151**	**0.2086**
	GenLouvain ($\gamma = 1$)	2	0.5403	0.8182	0.8173	**0.1151**	**0.2086**
	GenLouvain ($\gamma = 2$)	7	0.6371	0.8521	0.8422	0.3145	0.3593
	SC-ML	2	0.6716	0.8513	0.8508	0.1154	0.2178
	Self-Tuning	4	0.6310	0.8521	0.8424	1.0204	1.0970
109th Congress votes - Security	MIMOSA	2	0.6105	0.8513	0.8506	0.0400	0.0785
	MIMOSA-uniform	2	0.6304	0.8513	0.8506	0.0400	0.0785
	GenLouvain ($\gamma = 0.5$)	2	0.5816	0.8345	0.8337	0.0400	0.0770
	GenLouvain ($\gamma = 1$)	2	**0.6598**	**0.8685**	**0.8678**	0.0400	0.0770
	GenLouvain ($\gamma = 2$)	4	0.6181	0.8515	0.8477	**0.0204**	**0.0492**
	SC-ML	2	0.6304	0.8513	0.8506	0.0400	0.0785
	Self-Tuning	2	0.6304	0.8513	0.8506	0.0400	0.0785

- Voting network of 100 senators in 109th US Congress. Each layer represents a bill.

Chen and Hero [2017], Multilayer iterative model order selection..., IEEE T SIPN

Multimodal relevance networks

A *relevance network* is a graph $\mathcal{G} = \{\mathcal{E}, \mathcal{V}\}$, $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$, whose edges $\mathcal{E}$ capture a dependency relation between vertices.

Some examples:

- Correlation relevance networks (Lee et al. [1986], LANL TR)
- Partial correlation relevance networks (Lauritzen [1996], Oxford Univ Press)
- Mutual information relevance networks (Butte and Kohane [1999], Biocomputing)
- Directed information influence networks (Rao et al. [2007], J. Bionf Comp Bio)
- Stochastic block models (Oselio et al. [2014], IEEE JSTSP)

Multimodal relevance networks

A *relevance network* is a graph $\mathcal{G} = \{\mathcal{E}, \mathcal{V}\}$, $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$, whose edges $\mathcal{E}$ capture a dependency relation between vertices.

Some examples:

- Correlation relevance networks (Lee et al. [1986], LANL TR)
- Partial correlation relevance networks (Lauritzen [1996], Oxford Univ Press)
- Mutual information relevance networks (Butte and Kohane [1999], Biocomputing)
- Directed information influence networks (Rao et al. [2007], J. Bionf Comp Bio)
- Stochastic block models (Oselio et al. [2014], IEEE JSTSP)

$\Rightarrow$ multimodal relevance network approaches model interlayer dependencies

Undirected relevance networks: graphical model of conditional dependencies

Nodes $U, V \in \mathcal{V}$ have edge if they are conditionally
dependent given all other nodes in $\mathcal{V}$

- X, Y, Z are random variables with pdf

$$f(X, Y, Z) = f(X, Y|Z)f(Z)$$

$\Rightarrow$ Generally gives a complete graph $\mathcal{G}$

- X, Y form a *community* in $\mathcal{G}$ if:

$$f(X, Y|Z) = f(X, Y) \neq f(X)f(Y)$$

$\Rightarrow$ Remove edges from (X, Z) and (Y, Z)

- X, Y are conditionally independent given Z if

$$f(X, Y|Z) = f(X|Z)f(Y|Z)$$

$\Rightarrow$ Remove edge from (X, Y)

(Lauritzen [1996], Oxford Univ Press)
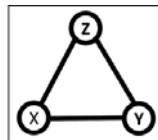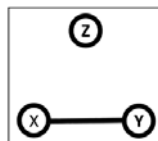
X,Y,Z dependent

X,Y community
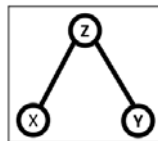
X,Y conditionally independent

Undirected relevance networks: graphical model of conditional dependencies

Natural measure of conditional dependency is MI

$$\mathrm{MI}(X, Y|Z) = \int f(x, y, z) \ln \left(f(x, y|z)/f(x|z)f(y|z) \right) dxdydz$$
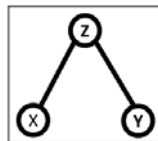
- X, Y, Z are random variables with pdf

$$f(X, Y, Z) = f(X, Y|Z)f(Z)$$

$\Rightarrow$ MI$(X, Y|Z)$, MI$(Z, X|Y)$, MI$(Z, Y|X) > 0$

- X, Y form a *community* in $\mathcal{G}$ if:

$$f(X, Y|Z) = f(X, Y) \neq f(X)f(Y)$$

$\Rightarrow$ MI$(X, Z|Y) = 0$ and MI$(Y, Z|X) = 0$

- X, Y are conditionally independent given Z if

$$f(X, Y|Z) = f(X|Z)f(Y|Z)$$

$\Rightarrow$ MI$(X, Y|Z) = 0$
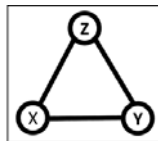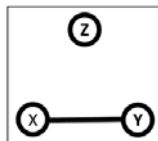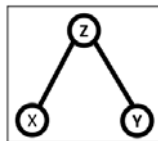
(Butte and Kohane [1999], Biocomputing)

X,Y,Z dependent

X,Y community

X,Y conditionally independent

35

Undirected relevance networks: graphical model of conditional dependencies

Gaussian case: $\mathbf{U} = [X, Y, Z] \sim \mathcal{N}(0, \boldsymbol{\Sigma})$

$$\boldsymbol{\Sigma} = \mathrm{cov}(\mathbf{U}) = \begin{bmatrix} \sigma_X^2 & \sigma_{XY} & \sigma_{XZ} \\ \sigma_{YX} & \sigma_Y^2 & \sigma_{YZ} \\ \sigma_{ZX} & \sigma_{ZY} & \sigma_Z^2 \end{bmatrix}$$

The MI takes simple form:

$$\mathrm{MI}(X, Y | Z) = -\frac{1}{2} \ln |\rho_{XY|Z}|$$

where $\rho_{XY|Z}$ is the partial correlation coefficient between X, Y

$$\rho_{XY|Z} = |(\mathbf{P})_{12}|$$

for which $\mathbf{P}$ is the partial correlation matrix

$$\mathbf{P} = [\mathrm{diag}(\boldsymbol{\Sigma}^{-1})]^{-1/2} \boldsymbol{\Sigma}^{-1} [\mathrm{diag}(\boldsymbol{\Sigma}^{-1})]^{-1/2}$$

(Lauritzen [1996], Oxford Univ Press)

X,Y,Z dependent

X,Y community

X,Y conditionally independent

36

Undirected relevance networks: graphical model of conditional dependencies

Stochastic representation of conditional dependencies
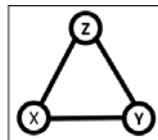
$$
\begin{aligned}
a_1 X + b_1 Y + c_1 Z &= N_1 \\
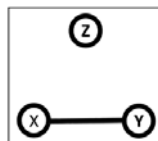b_2 Y + c_2 Z &= N_2 \\
c_3 Z &= N_3
\end{aligned}
$$

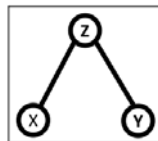where $[N_1, N_2, N_3] \sim \mathcal{N}(0, \mathbf{I})$ is isotropic Gaussian noise

Examples:

- X, Y, Z dependent: all coefficients a, b, c are non-zero
- X, Y community: $c_1 = c_2 = 0$
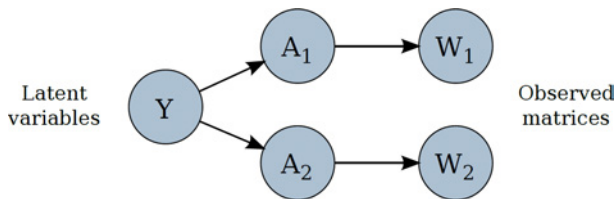- X, Y conditionally independendent: $b_1 = 0$

X,Y,Z dependent

X,Y community

X,Y conditionally independent

## Latent variable multilayer relevance network model

Markov diagram implies

$$f(W_1, W_2|A_1, A_2, Y) = f(W_1|A_1)f(W_2|A_2), \quad f(A_1, A_2|Y) = f(A_1|Y)f(A_2|Y)$$

- Observed matrices W_1, W_2: measured/estimated intralayer dependencies
- Adjacency matrices A_1, A_2: specify sparsity patterns of $\mathbb{E}[W_1]$ and $\mathbb{E}[W_2]$
- Latent variable Y: determines dependencies between A_1 and A_2

Oselio et al. [2014], Multilayer graph analysis for dynamic social networks, IEEE J STSP

## 5D Nucleome fibroblast proliferation experiment
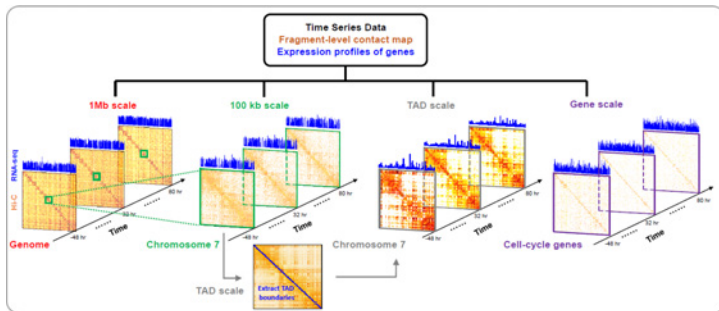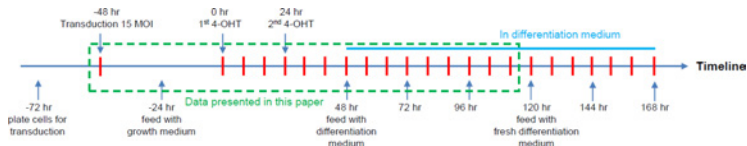
(Chen et al. [2015], PNAS), (Chen et al. [2016a], Bioinformatics), (Liu et al. [2017], BioRxv), (Chen et al. [2017], Nucleus), (Liu et al. [2018], BioRxv)

## Chromatin gene structure layer: Chromosome Conformation Capture (Hi-C)

Hi-C matrices for 23 chromosomes at 100 kilo-base pair (Kb) resolution for fibroblast proliferation
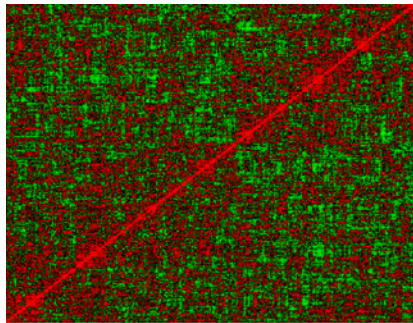
[I. Rajapakse, 4D Genome, https://rajapakse.lab.medicine.umich.edu/]
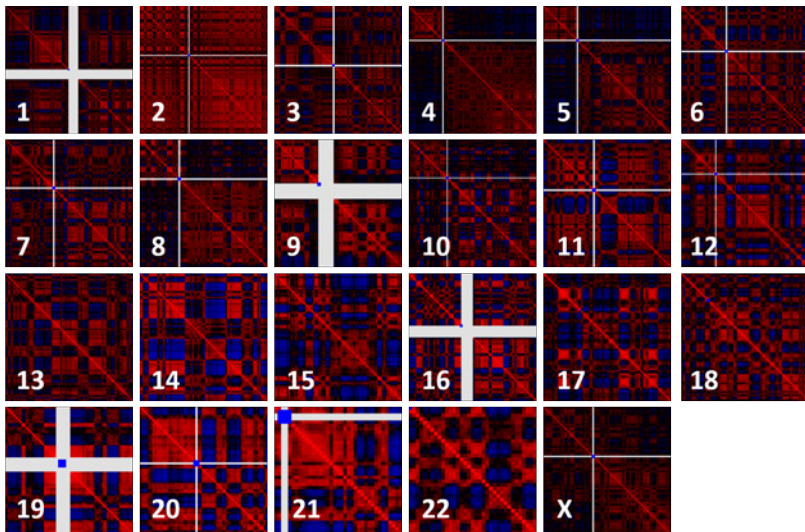
Functional gene expression layer: RNAseq

RNAseq data for Chromosome 4 [I. Rajapakse, 4D Genome, https://rajapakse.lab.medicine.umich.edu/]
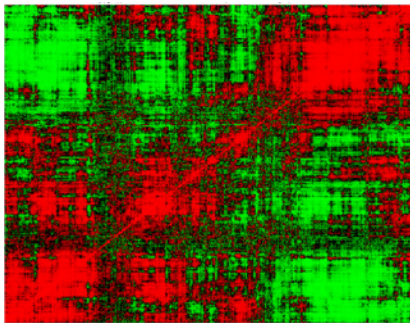
Correlation matrix

Partial correlation matrix

Correlation screening theory specifies correlation threshold value $\rho \Rightarrow \mathbf{A}^{RNA}$
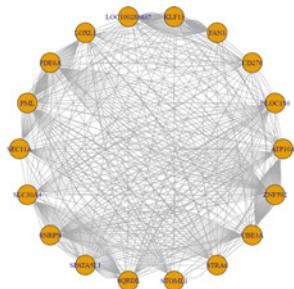
- $24 = 8 \times 3$ samples. 764 genes (zero expression genes removed)
- Critical phase transition threshold: $\rho_c = 0.72$
- 5% false positive threshold: $\rho = 0.87$

Latent multilayer community detection: motivation - multiomic networks

RNAseq parcorr and HiC contact graphs for genes in a TAD on chromosome 15
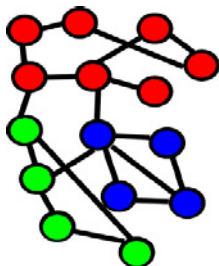

RNAseq correlation graph (Bernoulli)   HiC contact graph (Poisson)

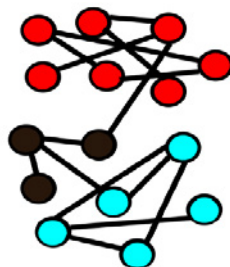Objective: partition genes into common multiomic communities and private omic communities

Tiomoko Ali et al. [2018], Latent heterogeneous multilayer community detection, arXiv

Multilayer community model
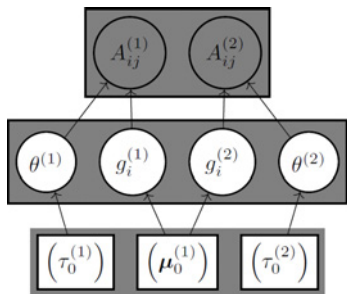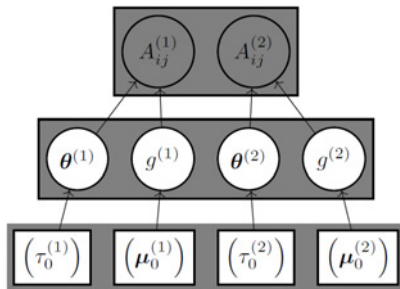

Layer 1
RNAseq correlation graph

Layer 2
HiC contact graph

- Red vertices: shared community in RNAseq and HiC networks
- Green and blue vertices: private communities in RNAseq network
- Black and turquoise vertices: private communities in HiC network

Tiomoko Ali et al. [2018], Latent heterogeneous multilayer community detection, arXiv

Latent multilayer community model

(a) $g_i^{(1)} \in [1, \ldots, K]$.

(b) $g_i^{(1)} \notin [1, \ldots, K]$.

Graphical model for shared community (left) and private community (right) - shaded circles represent deterministic parameters.

- $g_i^{(1)}, g_i^{(2)}$ community labels of vertex i in layers 1, 2
- $\theta^{(1)}, \theta^{(2)}$: SBM community parameters
- $A_{ij}^{(1)}$ and $A_{ij}^{(2)}$ are conditionally indep. weighted adjacency matrices

$$A_{ij}^{(1)} \sim f_{A_{ij}^{(1)} | \theta^{(1)}, g_i^{(1)}}, \qquad A_{ij}^{(2)} \sim f_{A_{ij}^{(2)} | \theta^{(2)}, g_i^{(2)}}$$

- $\mu^{(1)}, \mu^{(2)}$ and $\tau^{(1)}, \tau^{(2)}$ are generative distributions
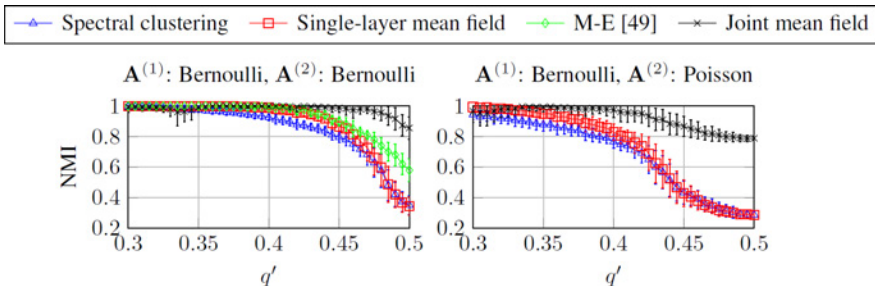
Shared community model
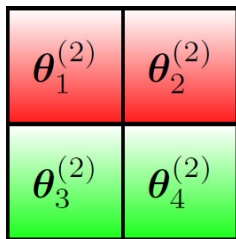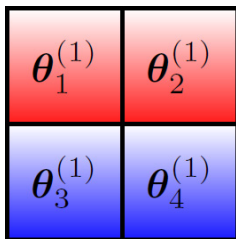
SBM community matrices $\theta^{(1)}$ and $\theta^{(2)}$ for layer 1 and layer 2, respectively.

- Proposed inference algorithm: joint mean field variational Bayes.
- Extends (Wilson et al. [2017], JMLR) to simultaneously extract shared *and* private communities
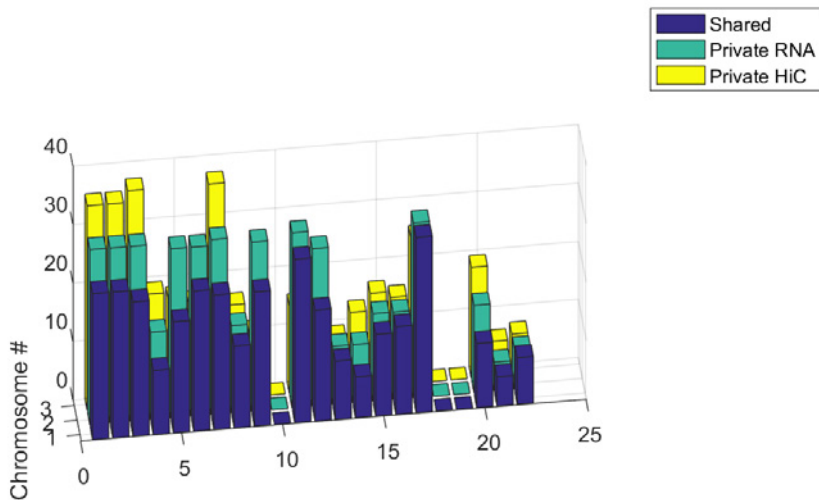
Tiomoko Ali et al. [2018], Latent heterogeneous multilayer community detection, arXiv

Simulation comparisons

- 500 vertices with Bernoulli or Poisson edges in each layer
- $K = 2$ shared communities each with 4 vertices
- Multilayer community detection algorithms implemented over 100 trials
- Comparisons for 2nd layer edge noise parameter $q' \in [0, 0.5]$
  - Spectral clustering (Chen and Hero [2017], IEEE T SPISN)
  - Single layer mean field (Aicher et al. [2014], JMLR)
  - Multilayer extraction (M-E) algorithm (Wilson et al. [2017], JMLR)
  - Proposed joint layer mean field (Tiomoko Ali et al. [2018], arXiv 2018)

HiC/RNAseq fibroblast proliferation data: discovered communities

Tiomoko Ali et al. [2018], Latent heterogeneous multilayer community detection, arXiv

HiC/RNAseq fibroblast proliferation data: Chromosome 4 communities
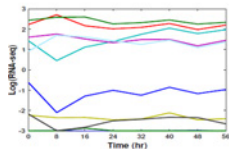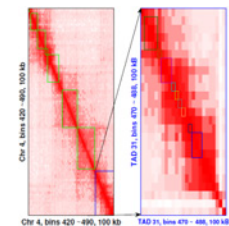
(a) Shared cluster   (b) RNA-seq private cluster   (c) HiC private cluster

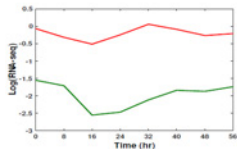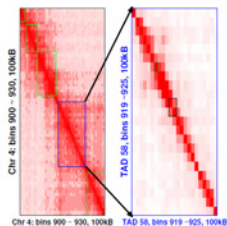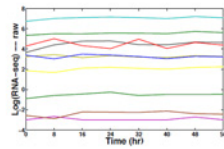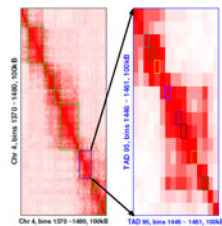Tiomoko Ali et al. [2018], Latent heterogeneous multilayer community detection, arXiv

Summary

- Community detection and centrality for multimodal multilayer networks

- Multilayer network analysis tools are less mature than single layer tools

- Two basic kinds of multilayer networks
  - Dynamic networks: layers are homogeneous (similar dependency mechanisms)
  - Multiomic networks: the layers are inhomogeneous (different dependency mechanisms)

- Latent multilayer dependency model is a probabilistic model that accounts for interlayer interactions

- Open problems:
  - Centrality measures: beyond multilayer mean and dispersion
  - Critical phase transitions: beyond independent layers
  - Community detection: beyond universal layer sharing of communiites

Acknowledgements

Emmanuel Abbe and Colin Sandon. Community detection in general stochastic block models: Fundamental limits and efficient algorithms for recovery. In *Foundations of Computer Science (FOCS), 2015 IEEE 56th Annual Symposium on*, pages 670–688. IEEE, 2015.

Christopher Aicher, Abigail Z Jacobs, and Aaron Clauset. Learning latent block structure in weighted networks. *Journal of Complex Networks*, 3(2):221–248, 2014.

Piotr Bródka, Krzysztof Skibicki, Przemysław Kazienko, and Katarzyna Musiał. A degree centrality in multi-layered social network. In *Computational Aspects of Social Networks (CASoN), 2011 International Conference on*, pages 237–242. IEEE, 2011.

Atul J Butte and Isaac S Kohane. Mutual information relevance networks: functional genomic clustering using pairwise entropy measurements. In *Biocomputing 2000*, pages 418–429. World Scientific, 1999.

Kathleen M Carley. Dynamic network analysis for counter-terrorism. *Unpublished manuscript*, 2005.

Haiming Chen, Jie Chen, Lindsey A Muir, Scott Ronquist, Walter Meixner, Mats Ljungman, Thomas Ried, Stephen Smale, and Indika Rajapakse. Functional organization of the human 4d nucleome. *Proceedings of the National Academy of Sciences*, 112(26):8002–8007, 2015.

Haiming Chen, Laura Seaman, Sijia Liu, Thomas Ried, and Indika Rajapakse. Chromosome conformation and gene expression patterns differ profoundly in human fibroblasts grown in spheroids versus monolayers. *Nucleus*, 8(4):383–391, 2017.

Jie Chen, Alfred O Hero III, and Indika Rajapakse. Spectral identification of topological domains. *Bioinformatics*, 32(14):2151–2158, 2016a.

Pin-Yu Chen and Alfred O Hero. Deep community detection. *IEEE Transactions on Signal Processing*, 63(21):5706–5719, 2015a.

Pin-Yu Chen and Alfred O Hero. Universal phase transition in community detectability under a stochastic block model. *Physical Review E*, 91(3):032804, 2015b.

Pin-Yu Chen and Alfred O Hero. Multilayer spectral graph clustering via convex layer aggregation: Theory and algorithms. *IEEE Transactions on Signal and Information Processing over Networks*, 3(3):553–567, 2017.

Pin-Yu Chen and Alfred O Hero. Phase transitions and a model order selection criterion for spectral graph clustering. *IEEE Trans. on Signal Processing*, page To appear, 2018.

Pin-Yu Chen, Sutanay Choudhury, and Alfred O Hero. Multi-centrality graph spectral decompositions and their application to cyber intrusion detection. In *Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on*, pages 4553–4557. IEEE, 2016b.

Job Dekker, Andrew S Belmont, Mitchell Guttman, Victor O Leshyk, John T Lis, Stavros Lomvardas, Leonid A Mirny, Clodagh C Oshea, Peter J Park, Bing Ren, et al. The 4d nucleome project. *Nature*, 549(7671):219, 2017.

Chris Ding, Xiaofeng He, and Horst D Simon. On the equivalence of nonnegative matrix factorization and spectral clustering. In *Proceedings of the 2005 SIAM International Conference on Data Mining*, pages 606–610. SIAM, 2005.

Xiaowen Dong, Pascal Frossard, Pierre Vandergheynst, and Nikolai Nefedov. Clustering with multi-layer graphs: A spectral perspective. *IEEE Transactions on Signal Processing*, 60(11):5820–5831, 2012.

Miroslav Fiedler. Algebraic connectivity of graphs. *Czechoslovak mathematical journal*, 23(2):298–305, 1973.

Ko-Jen Hsiao, Alex Kulesza, and Alfred Hero. Social collaborative retrieval. In *Proceedings of the 7th ACM international conference on Web search and data mining*, pages 293–302. ACM, 2014.

Hang-Hyun Jo, Seung Ki Baek, and Hie-Tae Moon. Immunization dynamics on a two-layer network model. *Physica A: Statistical Mechanics and its Applications*, 361(2):534–542, 2006.

Mikko Kivelä, Alex Arenas, Marc Barthelemy, James P Gleeson, Yamir Moreno, and Mason A Porter. Multilayer networks. *Journal of complex networks*, 2(3):203–271, 2014.

Florent Krzakala, Cristopher Moore, Elchanan Mossel, Joe Neeman, Allan Sly, Lenka Zdeborová, and Pan Zhang. Spectral redemption in clustering sparse networks. *Proceedings of the National Academy of Sciences*, 110(52):20935–20940, 2013.

S.L. Lauritzen. *Graphical models*, volume 17. Oxford University Press, USA, 1996.

YC Lee, Gary Doolen, HH Chen, GZ Sun, Tom Maxwell, and HY Lee. Machine learning using a higher order correlation network. Technical report, Los Alamos National Lab., NM (USA); Maryland Univ., College Park (USA), 1986.

Sijia Liu, Haiming Chen, Scott Ronquist, Laura Seaman, Nicholas Ceglia, Walter Meixner, Lindsey A Muir, Pin-Yu Chen, Gerald Higgins, Pierre Baldi, et al. Genome architecture leads a bifurcation in cell identity. *bioRxiv*, page 151555, 2017.

Sijia Liu, Pin-Yu Chen, Alfred Hero, and Indika Rajapakse. Dynamic network analysis of the 4d nucleome. *bioRxiv*, page 268318, 2018.

Shahin Mahdizadehaghdam, Han Wang, Hamid Krim, and Liyi Dai. Information diffusion of topic propagation in social media. *IEEE Transactions on Signal and Information Processing over Networks*, 2(4):569–581, 2016.

Peter J Mucha, Thomas Richardson, Kevin Macon, Mason A Porter, and Jukka-Pekka Onnela. Community structure in time-dependent, multiscale, and multiplex networks. *science*, 328(5980):876–878, 2010.

Raj Rao Nadakuditi and Mark EJ Newman. Graph spectra and the detectability of community structure in networks. *Physical review letters*, 108(18):188701, 2012.

M.E.J. Newman. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, 103(23):8577–8582, 2006.

Brandon Oselio, Alex Kulesza, and Alfred O Hero. Multi-layer graph analysis for dynamic social networks. *IEEE Journal of Selected Topics in Signal Processing*, 8(4):514–523, 2014.

Brandon Oselio, Sijia Liu, and Alfred Hero. Multilayer relevance networks. In *Proceedings of the IEEE Conf. on Signal Processing and Wireless Communications*. IEEE, 2018.

Arvind Rao, Alfred O Hero III, James Douglas Engel, et al. Motif discovery in tissue-specific regulatory sequences using directed information. *EURASIP Journal on Bioinformatics and Systems Biology*, 2007:3, 2007.

Alaa Saade, Florent Krzakala, and Lenka Zdeborová. Spectral clustering of graphs with the bethe hessian. In *Advances in Neural Information Processing Systems*, pages 406–414, 2014.

Peter H Salus and G Vinton. *Casting the Net: From ARPANET to Internet and Beyond...* Addison-Wesley Longman Publishing Co., Inc., 1995.

Hafiz Tiomoko Ali, Sijia Liu, Yasin Yilmaz, Indika Rajapakse, Romain Couillet, and Alfred O Hero. Latent heterogeneous multilayer community detection. *arXiv:1806.07963*, 2018.

Lois M Verbrugge. Multiplexity in adult friendships. *Social Forces*, 57(4):1286–1309, 1979.

James D Wilson, John Palowitch, Shankar Bhamidi, and Andrew B Nobel. Community extraction in multilayer networks with heterogeneous community structure. *The Journal of Machine Learning Research*, 18(1):5458–5506, 2017.

Kevin S Xu and Alfred O Hero. Dynamic stochastic blockmodels for time-evolving social networks. *IEEE Journal of Selected Topics in Signal Processing*, 8(4):552–562, 2014.

Tianbao Yang, Yun Chi, Shenghuo Zhu, Yihong Gong, and Rong Jin. Detecting communities and their evolutions in dynamic social networksa bayesian approach. *Machine learning*, 82(2):157–189, 2011.

Lihi Zelnik-Manor and Pietro Perona. Self-tuning spectral clustering. In *Advances in neural information processing systems*, pages 1601–1608, 2005.