# Community Structure in Multilayer Networks

Mason A. Porter
Mathematical Institute, University of Oxford

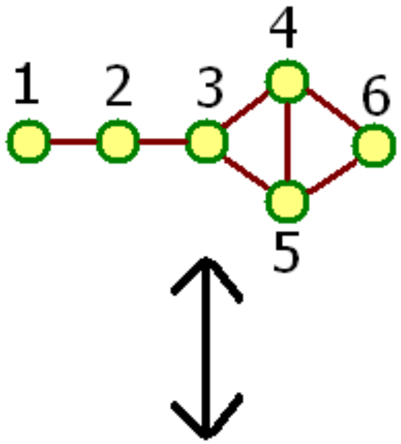
# Roadmap of Papers (Abridged)

- Review Article on Multilayer Networks
  - [1] M. Kivelä, A. Arenas, M. Barthelemy, J. P. Gleeson, Y. Moreno, & MAP, *Journal of Complex Networks*, in press (arXiv: 1309.7233)
- Expository Article on Community Structure
  - [2] M. A. Porter, J.-P. Onnela, & P. J. Mucha, *Notices of the American Mathematical Society*, Vol. 56, No. 9: 1082–1097, 1164–1166
    - Gives the state of play for community structure as of Oct. 2009. Very friendly introduction.
- Methods
  - [3] P. J. Mucha, T. Richardson, Kevin Macon, M. A. Porter, & J.-P. Onnela, Science, Vol. 328, No. 5980, 876–878 (2010)
    - Introduces our method for multilayer community detection (for "multislice" type of multilayer networks)
  - Other methods and improved understanding of the above methodology in more recent and in-progress papers.
- Applications to Neuroscience
  - [4] D. S. Bassett, N. F. Wymbs, M. A. Porter, P. J. Mucha, J. M. Carlson, & S. T. Grafton, PNAS, Vol. 118, No. 18, 7641–7646 (2011)
  - Numerous additional papers since the first one
- Other Applications
  - Political networks, coupled nonlinear-oscillator models, financial-assets networks, Lagrangian coherent structures, etc.

# Outline

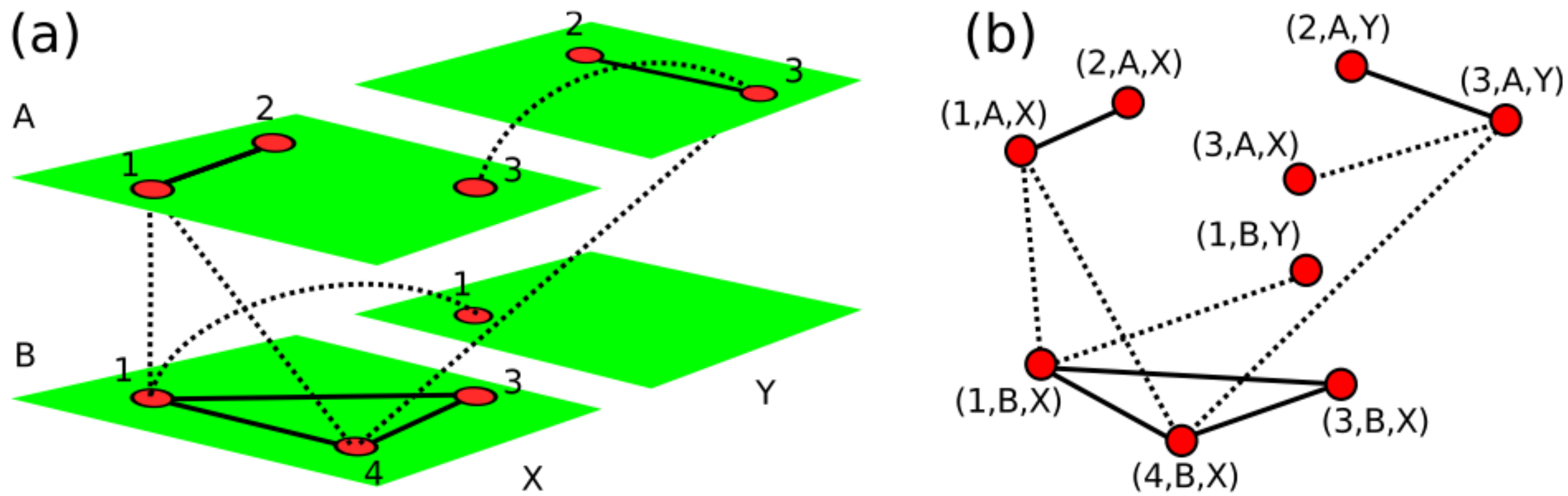
- What are *multilayer networks*?
- What is *community structure*?
- Community structure in multilayer networks
- Application to functional brain networks
- Conclusions and outlook

# Network

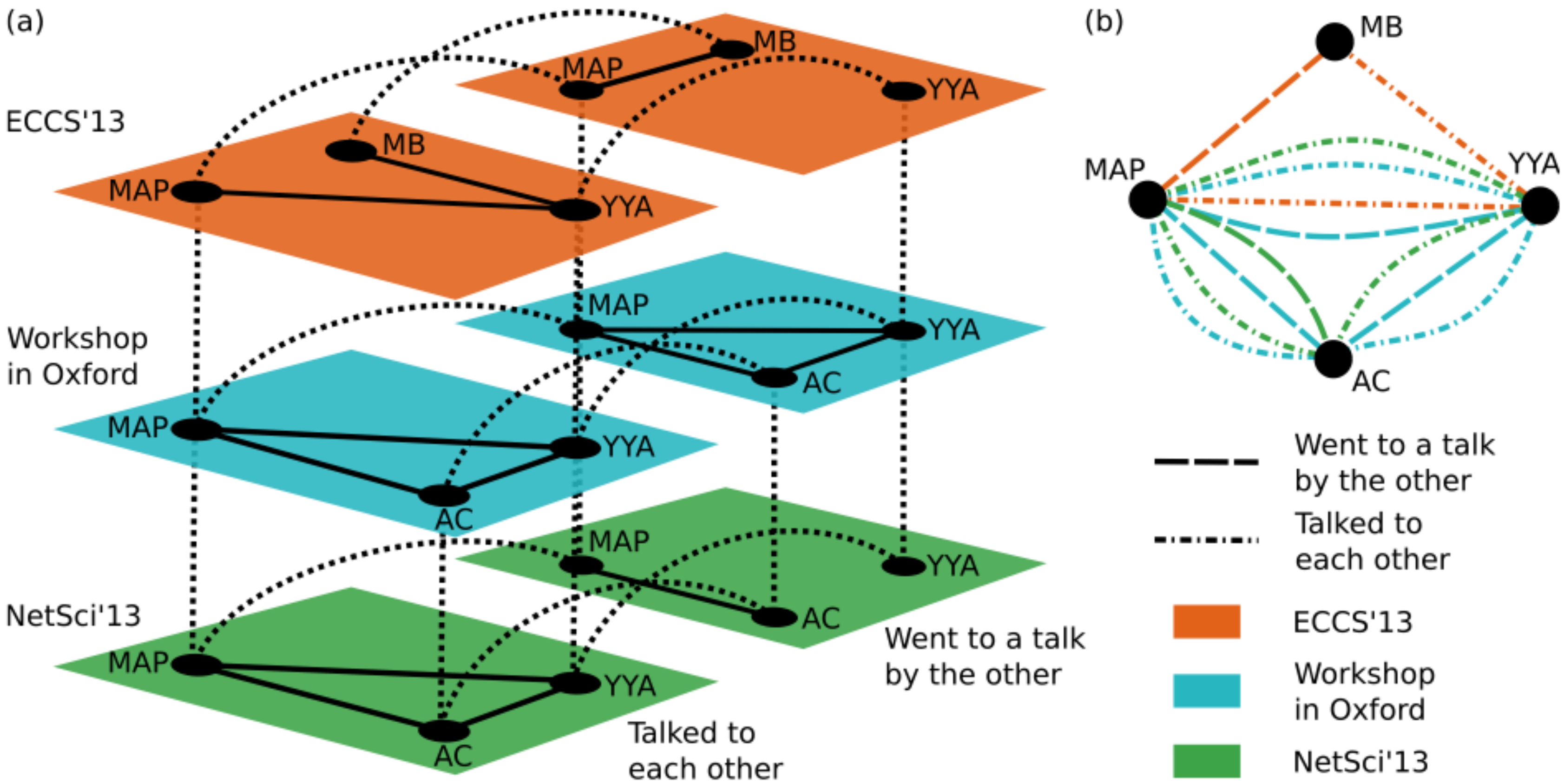



$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}$$

# Multilayer Network



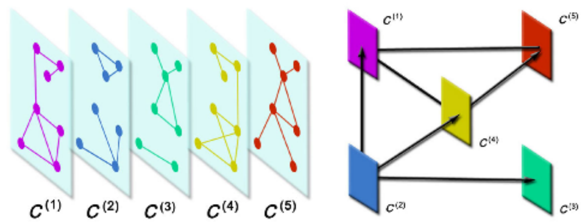
# “(Zachary) Karate Club Club” Network




Pictures courtesy of Aaron Clauset

# Multilayer Network

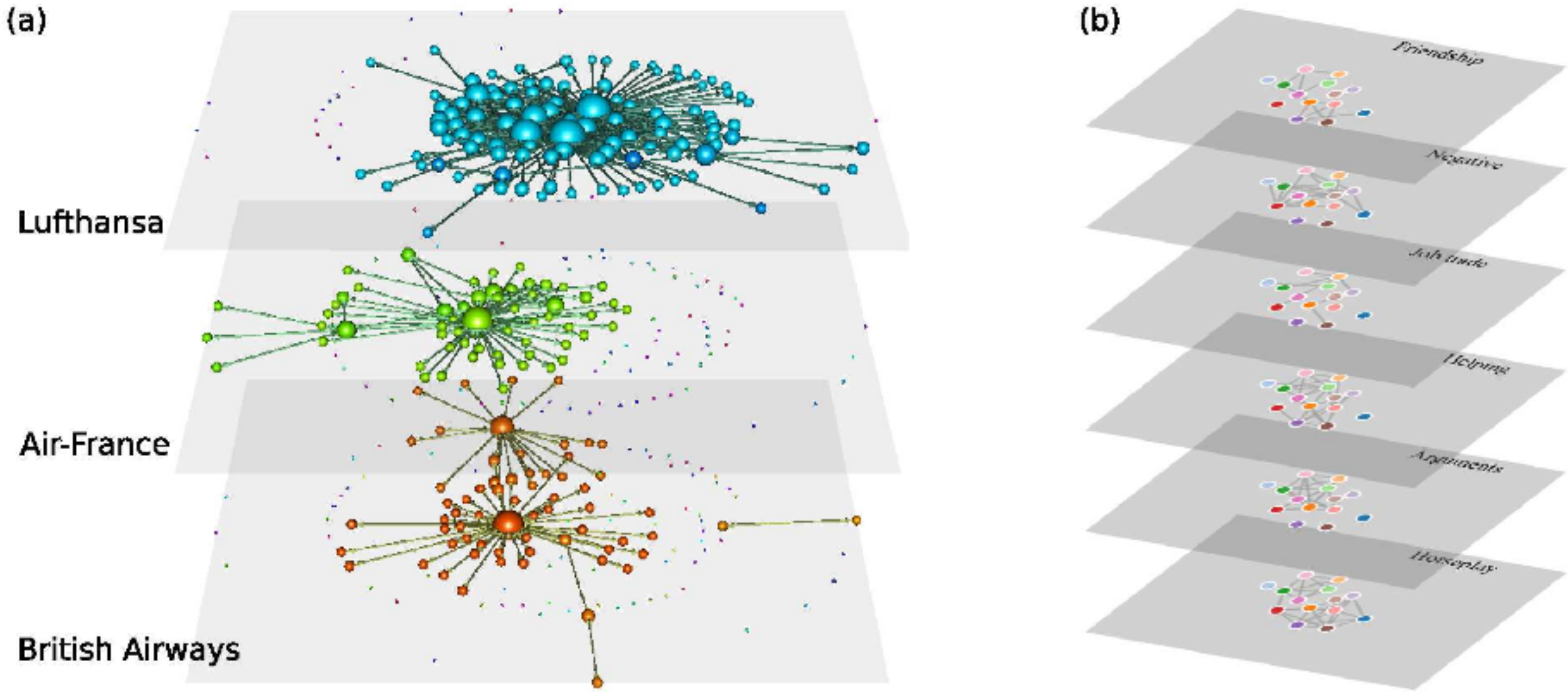
- Definition of a *multilayer network* M
  - $M = (V_M, E_M, V, L)$
    - V: set of nodes
      - As in ordinary graphs
    - L: sequence of sets of possible layers
      - One set for each additional "aspect" $d \geq 0$ beyond an ordinary network (examples: $d = 1$ in schematic on this page; $d = 2$ on last page)
    - $V_M$: set of tuples that represent *node-layers*
    - $E_M$: multilayer edge set that connects these tuples




- Note 1: allow weighted multilayer networks by mapping edges to real numbers with $w: E_M \rightarrow \mathbb{R}$
- Note 2: $d = 0$ yields the usual single-layer ("monoplex") networks

# Visualization and Analysis Software

- Plexmath software page: http://www.plexmath.eu/?page_id=327
- Muxviz: http://muxviz.net
- M. De Domenico, M. A. Porter, & A. Arenas, arXiv:1405.0843



# Tensorial Representation

- *Adjacency tensor* for unweighted case:

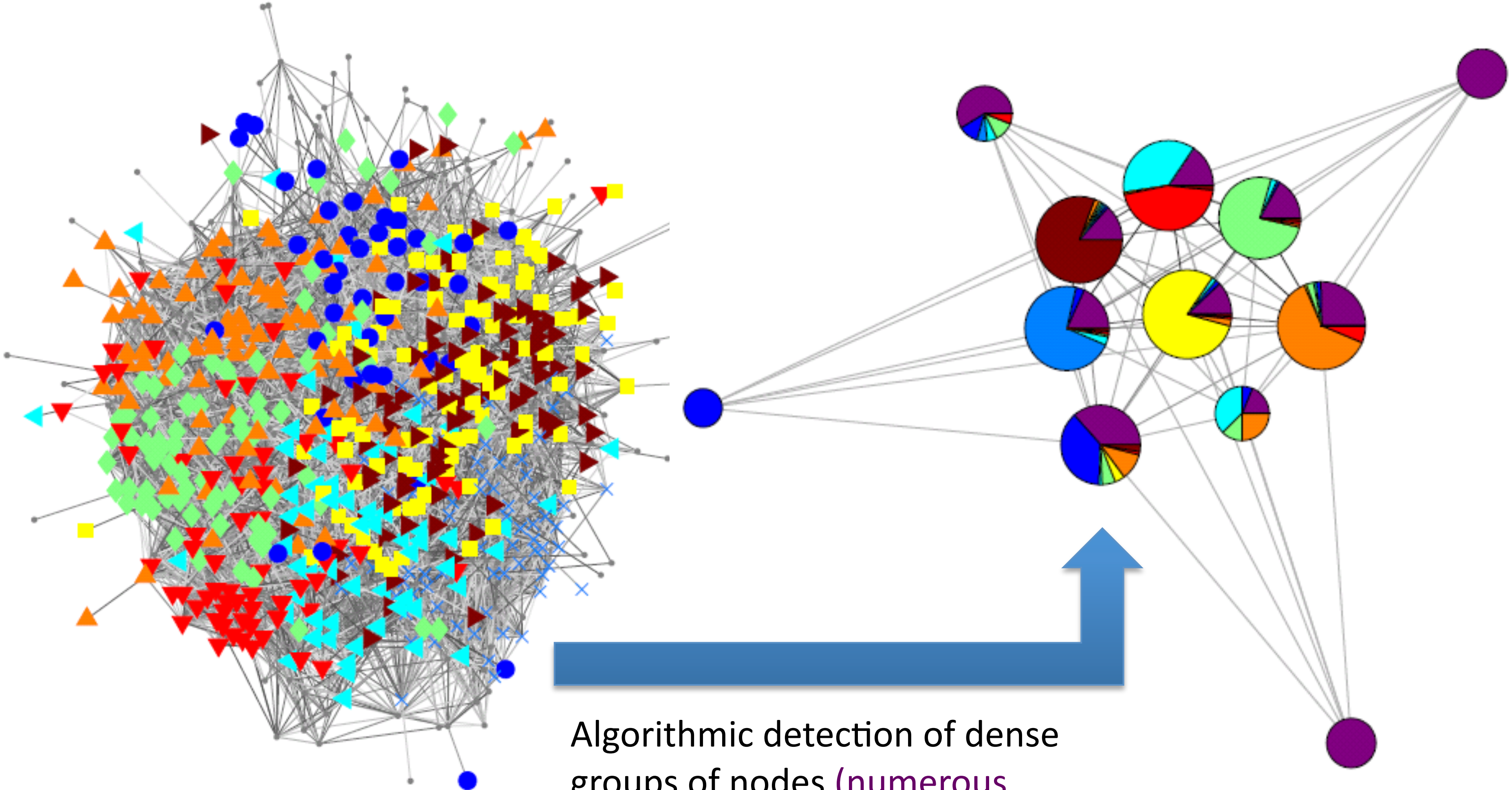
$$\mathcal{A} \in \{0,1\}^{|V|\times|V|\times|\mathbf{L}_1|\times|\mathbf{L}_1|\times\cdots\times|\mathbf{L}_d|\times|\mathbf{L}_d|}$$

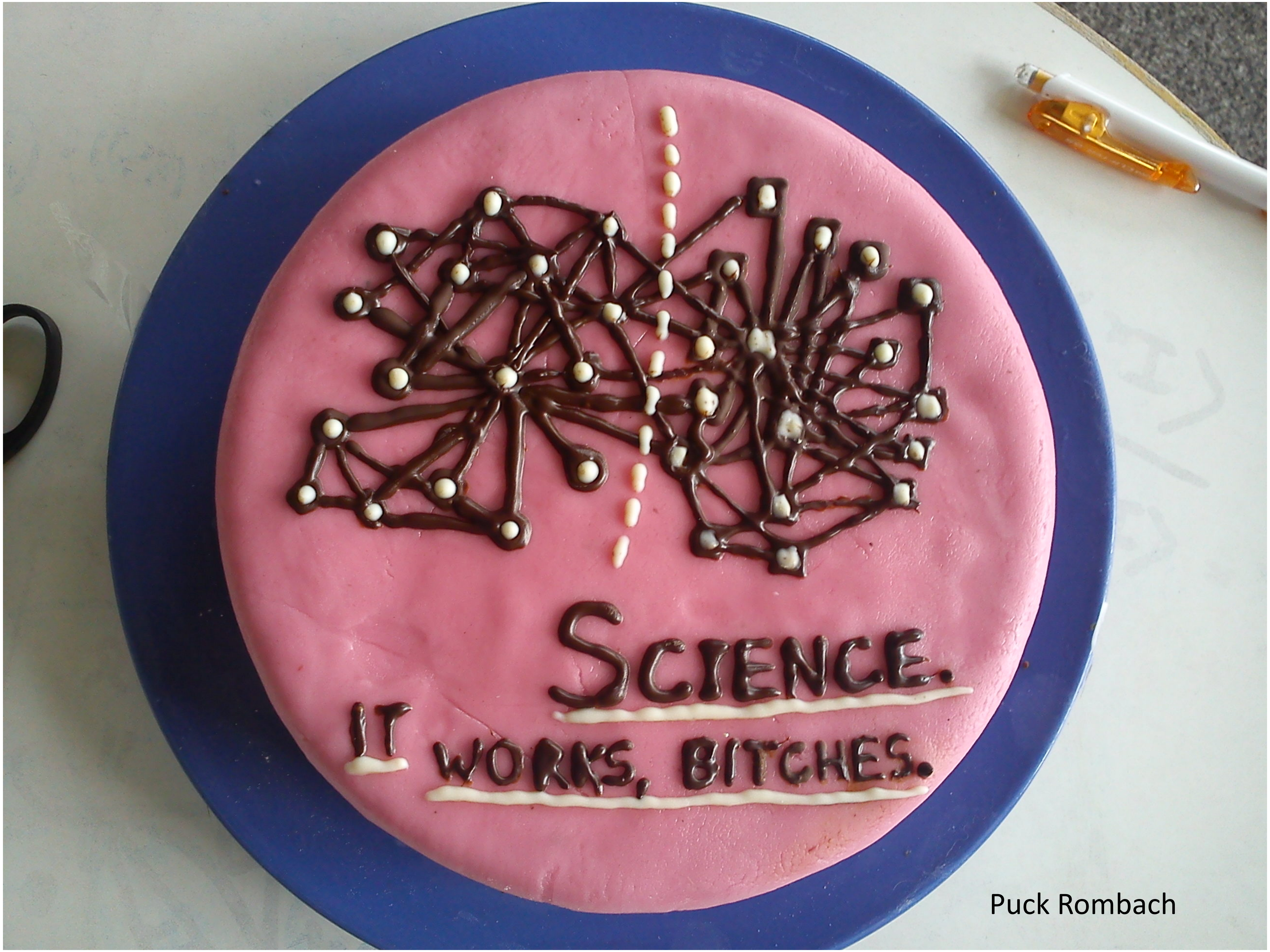
- Elements of adjacency tensor:
  - $A_{uv\boldsymbol{\alpha}\boldsymbol{\beta}} = A_{uv\alpha 1\beta 1 \ldots \alpha d\beta d} = 1$ iff $((u,\boldsymbol{\alpha}), (v,\boldsymbol{\beta}))$ is an element of $E_M$ (else $A_{uv\boldsymbol{\alpha}\boldsymbol{\beta}} = 0$)

# The literature is messy.

| Name | Aligned | Disj. | Eq. Size | Diag. | Lcoup. | Cat. | $\|L\|$ | $d$ | Example refs. |
|---|---|---|---|---|---|---|---|---|---|
| Multilayer network | | | | ✓ | ✓ | ✓ | Any | 1 | [58] |
| | ✓† | | ✓† | | | | Any | 1 | [79] |
| Multiplex network | ✓† | | ✓† | ✓ | | | Any | 1 | [78, 79] |
| | ✓ | | ✓ | ✓ | ✓ | ✓ | Any | 1 | [24, 34, 49, 62, 125, 198, 287] |
| | ✓ | | ✓ | ✓ | ✓ | ✓ | 2 | 1 | [154, 180, 182] |
| | | | | ✓ | ✓ | ✓ | Any | 1 | [70] |
| | ✓ | | ✓ | ✓ | ✓ | | Any | 1 | [71, 242, 243] |
| Multivariate network | ✓ | | ✓ | ✓ | ✓ | ✓ | Any | 1 | [209] |
| Multinetwork | ✓ | | ✓ | ✓ | ✓ | ✓ | Any | 1 | [14] |
| | ✓ | | ✓ | ✓ | ✓ | ✓ | Any | 2 | [15] |
| Multirelational network | ✓ | | ✓ | ✓ | ✓ | ✓ | Any | 1 | [55, 119, 252, 278] |
| Multirelational data | ✓ | | ✓ | ✓ | ✓ | ✓ | Any | 1 | [160, 197] |
| Multilayered network | ✓ | | ✓ | ✓ | ✓ | ✓ | Any | 1 | [45–47, 242] |
| Multidimensional network | ✓ | | ✓ | ✓ | ✓ | ✓ | Any | 1 | [18, 31–33, 69, 140, 264] |
| | ✓ | | ✓ | ✓ | ✓ | ✓ | Any | 3 | [141] |
| Multislice network | ✓† | | ✓† | ✓ | | | Any | 1 | [22, 56, 187, 188] |
| Multiplex of interdep. networks | ✓ | | ✓ | ✓ | ✓ | ✓ | Any | 1 | [111] |
| Hypernetwork | ✓ | | ✓ | ✓ | ✓ | ✓ | Any | 1 | [131, 247] |
| Overlay network | ✓ | | ✓ | ✓ | ✓ | ✓ | 2 | 1 | [97, 170] |
| Composite network | ✓ | | ✓ | ✓ | ✓ | ✓ | 2 | 1 | [282] |
| Multilevel network | | | | ✓ | ✓ | ✓ | Any | 1 | [70, 74] |
| ** | | ✓ | | | | | Any | 1 | [153, 272] |
| Multiweighted graph | ✓ | | ✓ | ✓ | ✓ | ✓ | Any | 1 | [218] |
| Heterogeneous network | | ✓ | | | | | 2 | 1 | [55, 294] |
| Multitype network | | ✓ | | | | | Any | 1 | [8, 120, 269] |
| Interconnected networks | | ✓ | ✓ | | | | 2 | 1 | [81, 164] |
| | | ✓ | | | | | 2 | 1 | [225, 229] |
| Interdependent networks* | | | ✓ | | | | Any | 1 | [244] |
| | | | ✓ | | | | 2 | 1 | [173] |
| | | ✓ | | | | | 2 | 1 | [48, 110] |
| | ✓ | | ✓ | ✓ | ✓ | ✓ | Any | 1 | [25] |
| Partially interdep. networks* | | | ✓ | | | | 2 | 1 | [244] |
| Network of networks* | | | ✓ | | | | Any | 1 | [98] |
| Coupled networks | | | | ✓ | ✓ | ✓ | Any | 1 | [288] |
| Interconnecting networks | | | | ✓ | ✓ | ✓ | 2 | 1 | [286] |
| Interacting networks | | ✓ | | | | | Any | 1 | [85, 155] |
| | | ✓ | | | | | 2 | 1 | [48] |
| Heterogenous information net | | ✓ | | | | | Any | 1 | [258] |
| ** | | | | | | | Any | 2 | [77, 255–257] |
| Meta-matrix, meta-network | | | | | | | Any | 2 | [60, 61, 266] |

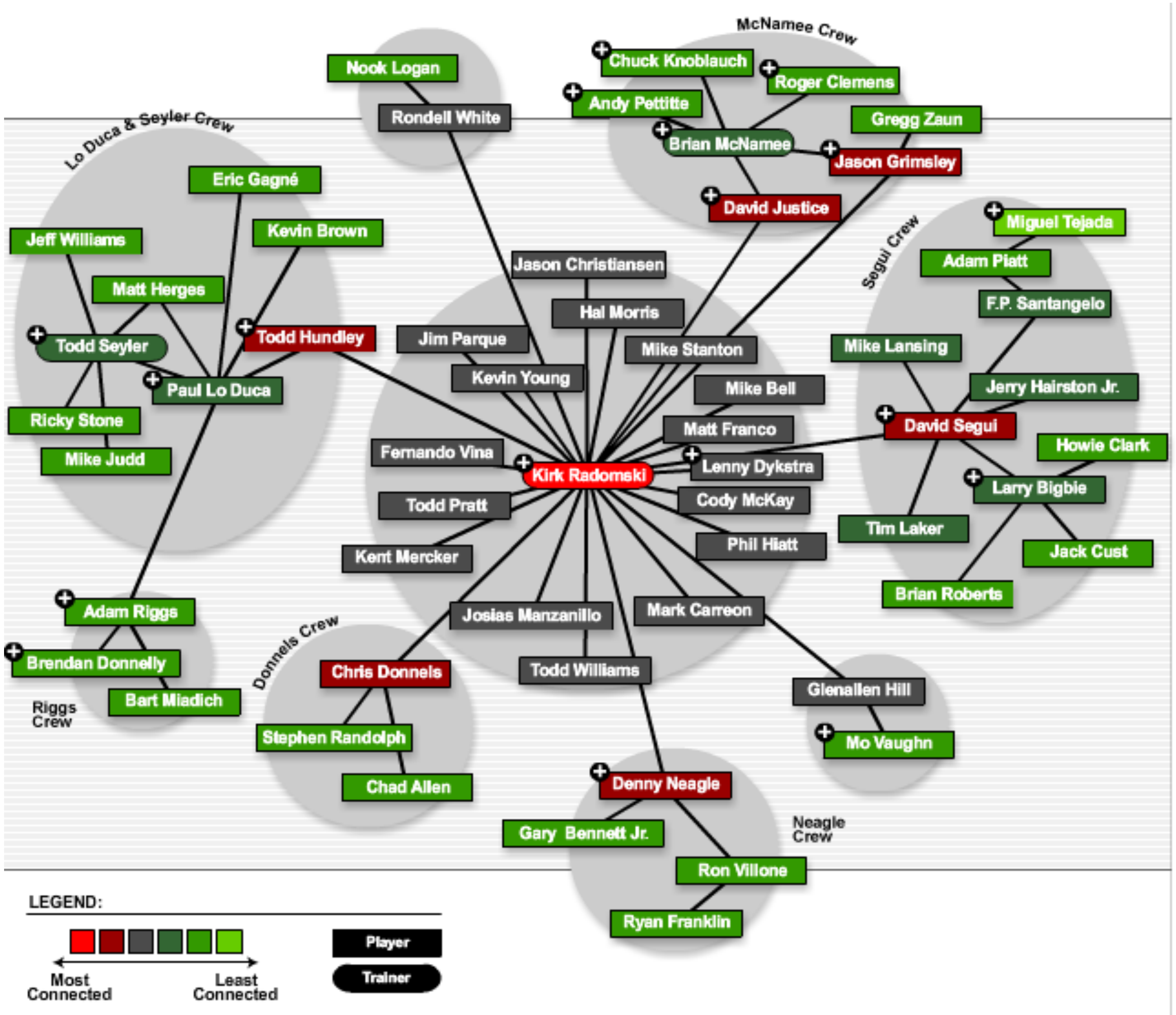
# Community Detection



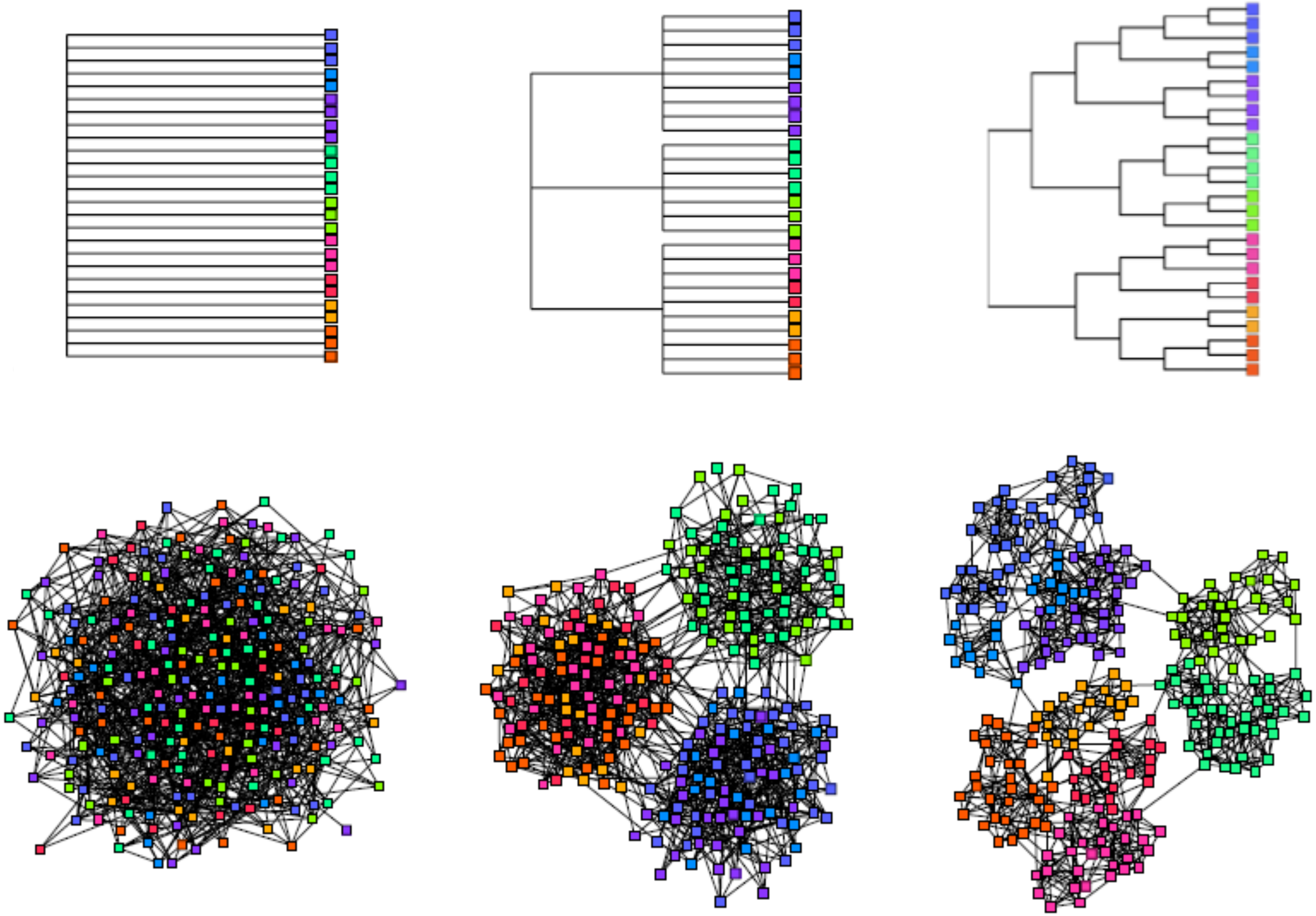


Puck Rombach

# Community Structure by hand?: Baseball Steroids Networks

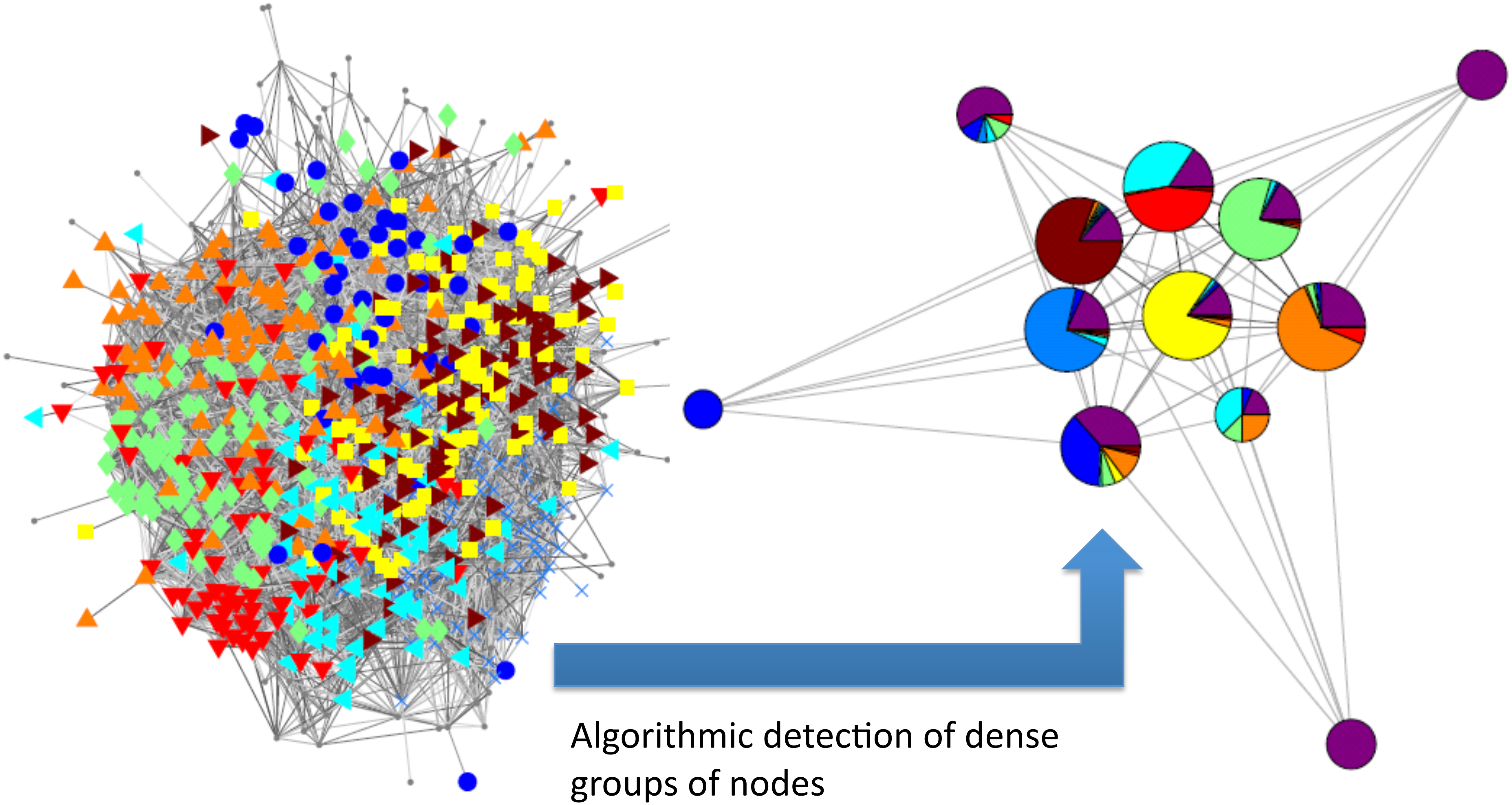


# Identifying Communities Algorithmically

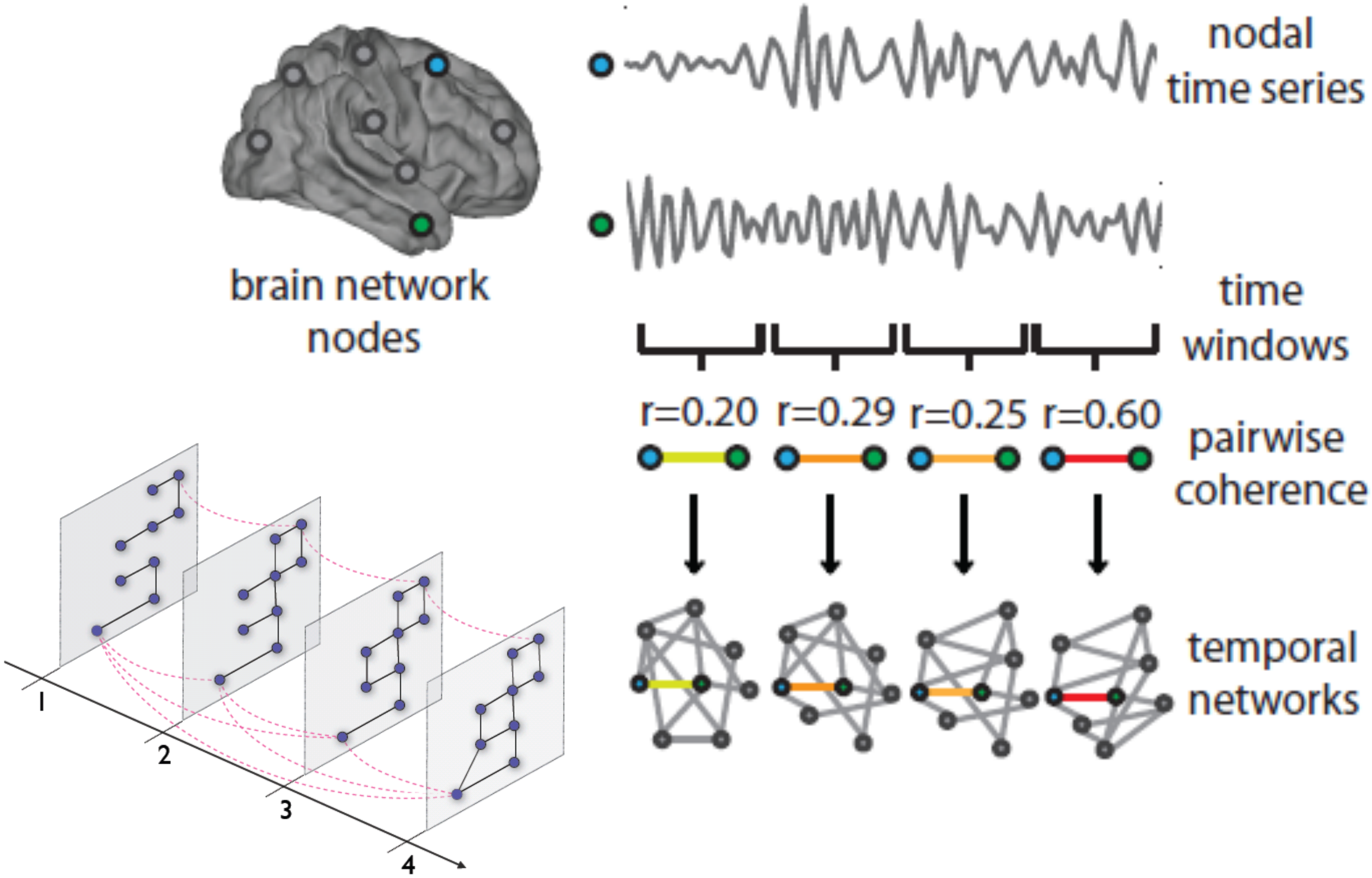


Images from A. Clauset, C. Moore, & M. E. J. Newman (*Nature*, 2008)

# Example: Facebook Friendship Networks



# Time-Dependent Networks
(e.g. from fMRI data)

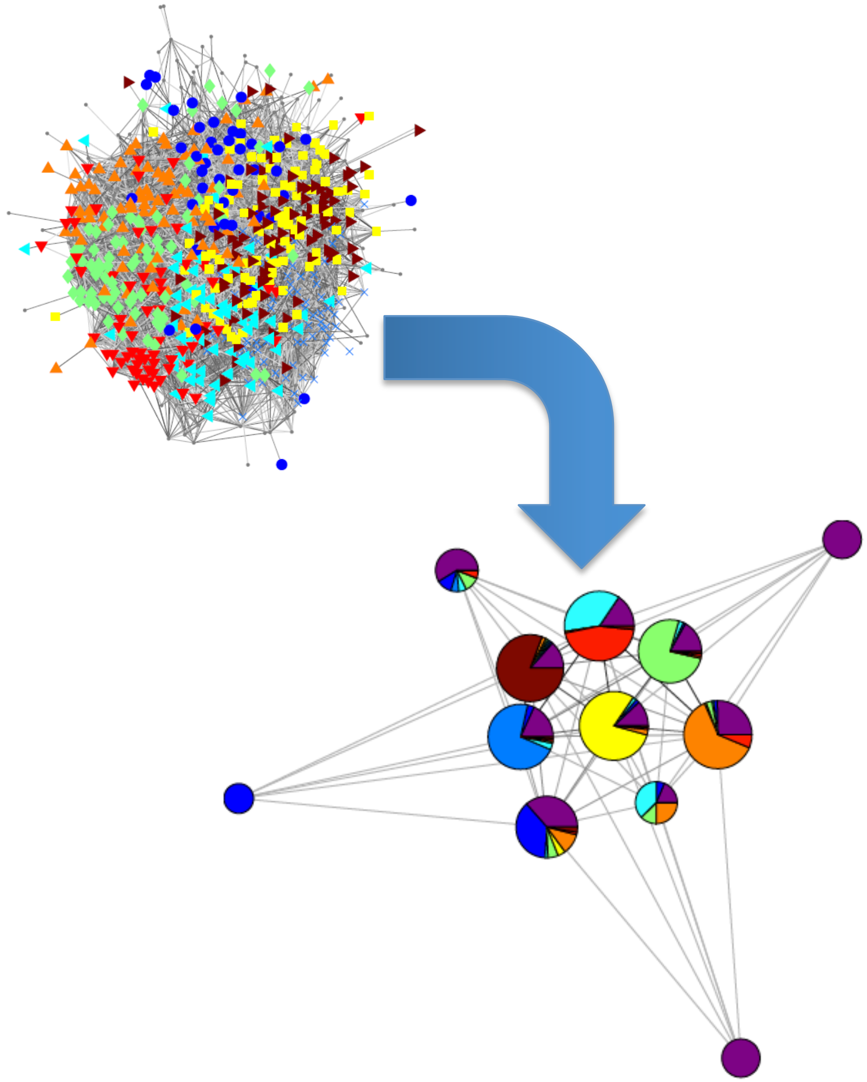


# Preliminaries

- “Hard/rigid” versus “soft/fuzzy/overlapping” clustering
- A *community* should describe a “cohesive group” of nodes
  - Tons of algorithms available
- Usual notion: more intra-community edges than one would expect at random
  - *But what does “at random” mean?*
- Review articles
  - “Communities in Networks,” M. A. Porter, J.-P. Onnela & P. J. Mucha, *Notices of the American Mathematical Society* **56**, 1082–1097 & 1164–1166 (2009).
  - “Community Detection in Graphs,” S. Fortunato, *Physics Reports* **486**, 75–174 (2010).

# Network Communities

- Communities = cohesive groups/modules/ mesoscopic structures
  - In stat phys, you try to derive macroscopic and mesoscopic insights from microscopic information
- Community structure consists of complicated interactions between modular (horizontal) and hierarchical (vertical) structures
- Communities have denser set of internal links relative to some null model for what links are present at random
  - "Modularity"



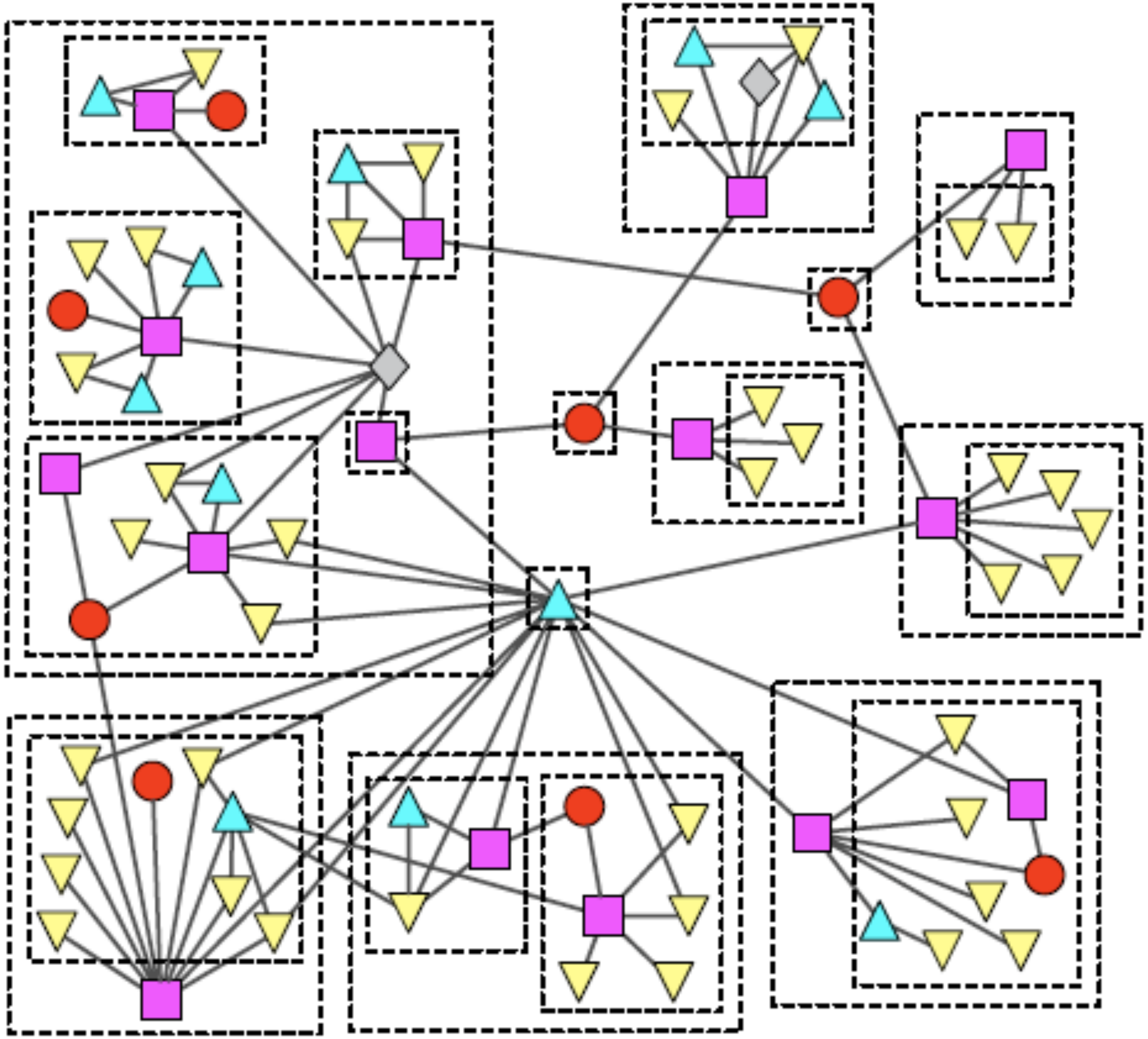
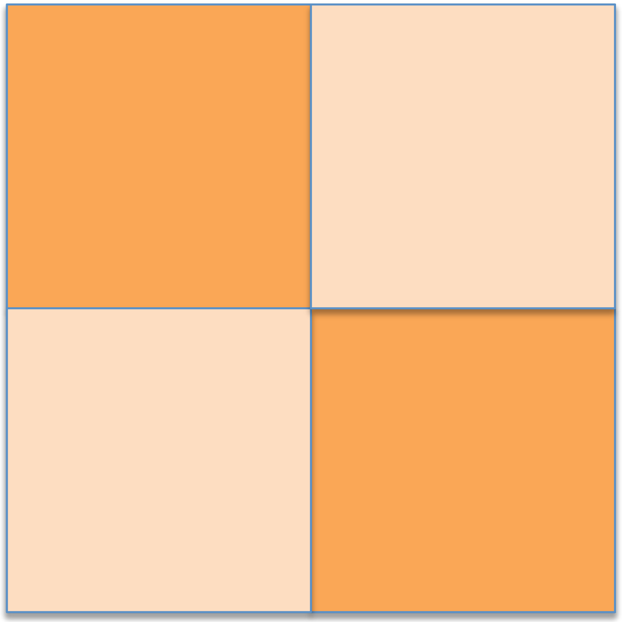


Image from A. Clauset, C. Moore, & M. E. J. Newman (*Nature*, 2008)

# Detecting Communities

- MAP, J.-P. Onnela, & P. J. Mucha [2009], *Notices of the American Mathematical Society* **56**(9): 1082–1097, 1164–1166
- Several types of methods
  - Agglomerative
  - Divisive
  - Local methods
  - Edge clustering
  - Etc.



# Quality / Modularity

- Popular approach: Use a "modularity" quality function

$$Q = \frac{1}{2W} \sum_{i,j} B_{ij} \delta(C_i, C_j) \, , \quad B_{ij} = A_{ij} - P_{ij}$$

where $\delta(C_i, C_j)$ indicates that the $B_{ij}$ components are only summed over cases in which nodes $i$ and $j$ are classified in the same community. The factor $W = \frac{1}{2}\sum_{ij} A_{ij}$ is the total edge strength in the network (equal to the total number of edges for unweighted networks), where $k_i$ again denotes the strength of node $i$. In (3.2), $P_{ij}$ denotes the components of a *null model* matrix, which specifies the relative value of intra-community edges in assessing when communities are closely connected [8, 77].

- GOAL: Assign nodes to communities to maximize Q.

# Example Null Models

## (aka: what does "at random" mean?)

- Erdös-Rényi (Bernoulli)

$$P_{ij} = p$$

- Newman-Girvan*

$$P_{ij} = \gamma \frac{k_i k_j}{2W}$$

- Leicht-Newman* (directed)

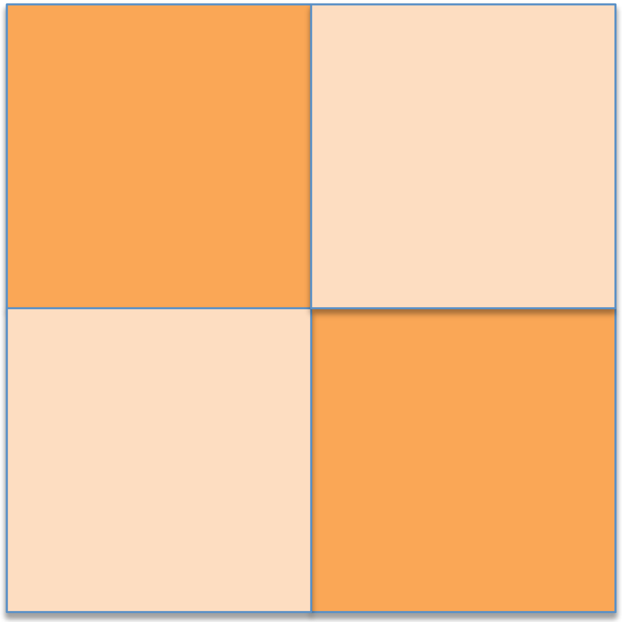
$$P_{ij} = \gamma \frac{k_i^{in} k_j^{out}}{W}$$

- Barber* (bipartite)

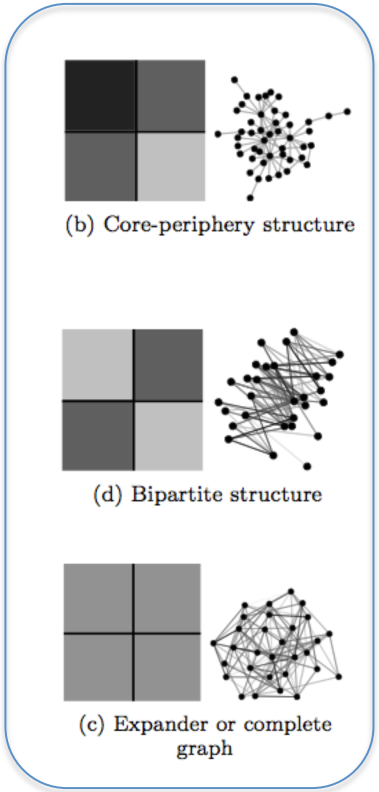
$$P_{ij} = \begin{cases} \gamma \frac{k_i d_j}{W} \\ 0 \end{cases}$$

** With additional resolution parameter γ*

# Platonic ideal of block structure for "traditional" Newman-Girvan choice of Q (nested version of this)



Community Structure

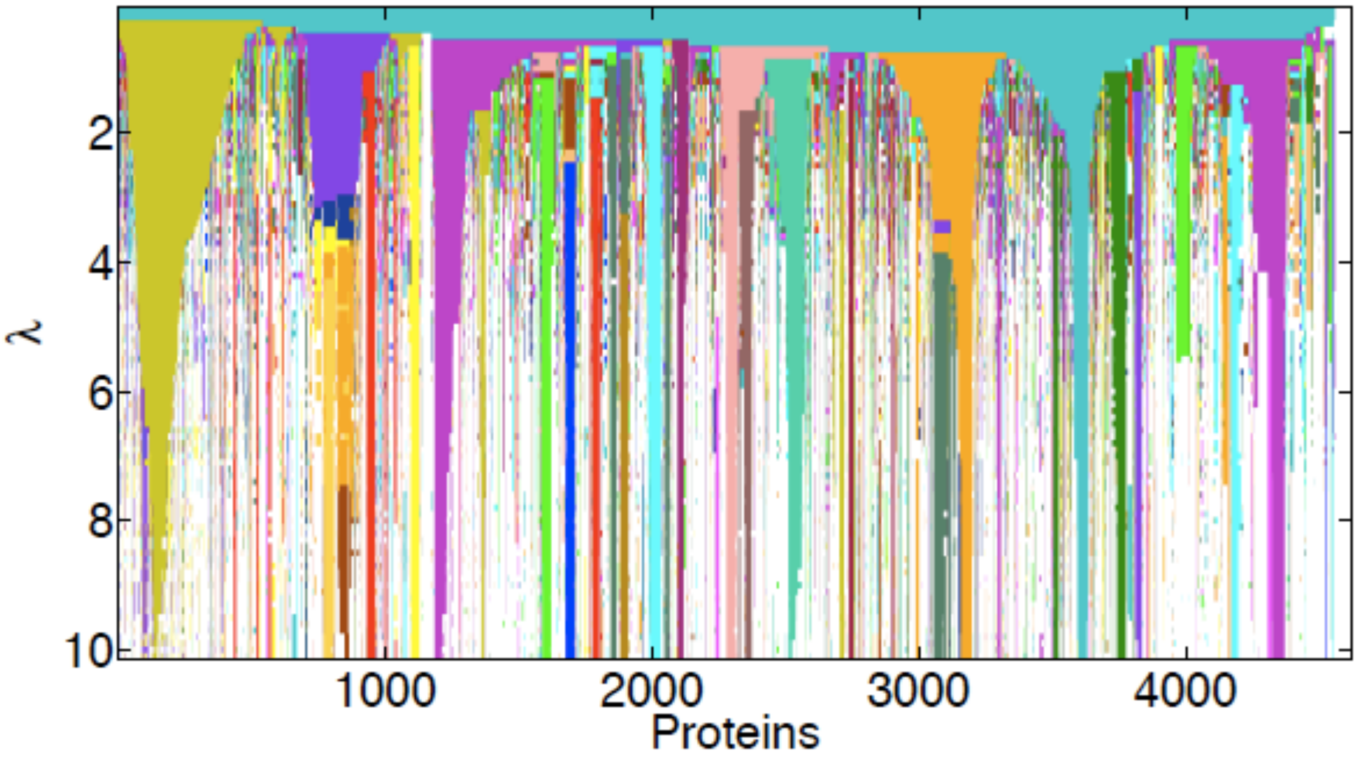


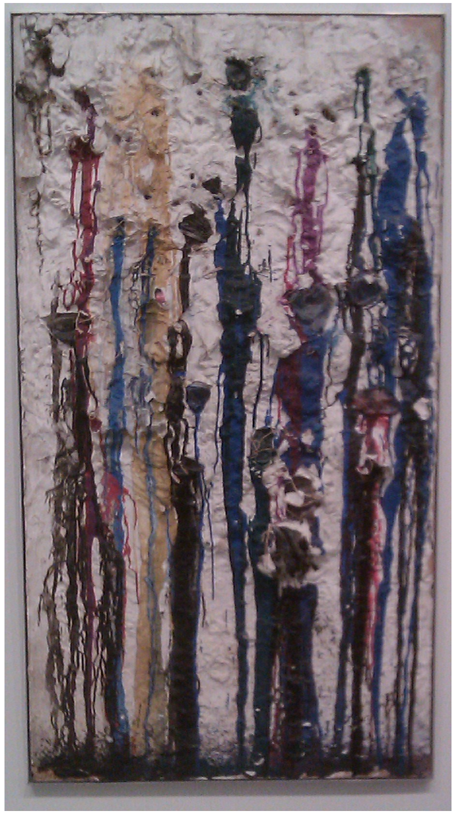

- This can be generalized, though the vast majority of methods have this in mind…
  - Note: I will focus on hard partitioning, but one can also think about overlapping communities in multilayer networks.

# Real Networks: Onion Peeling

## Example: Protein-Protein Interaction Networks

A. C. F. Lewis, N. S. Jones, MAP, & C. M. Deane, *BMC Systems Biology* **4**: 100 (2010)



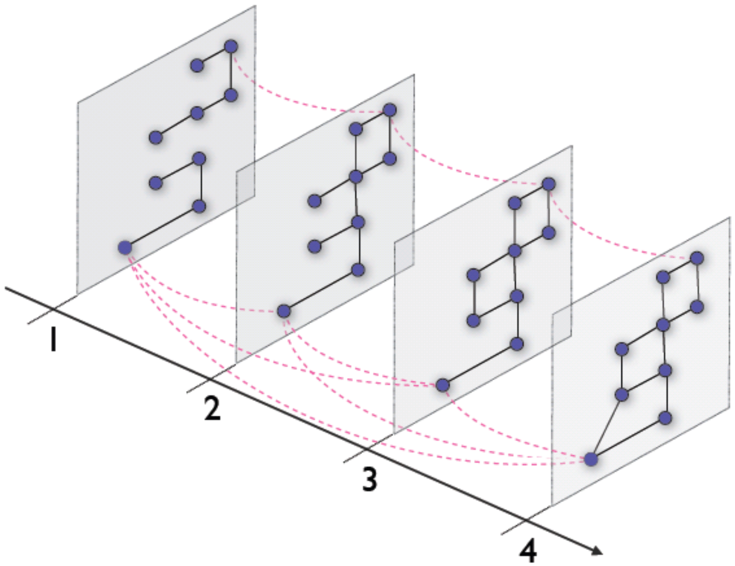



# Community Detection: Computational Heuristics

$$Q = \frac{1}{2W} \sum_{i,j} B_{ij} \delta(C_i, C_j)\,, \;\; B_{ij} = A_{ij} - P_{ij}$$

- Cannot guarantee optimal quality without full enumeration of possible partitions
  - NP-hard problem
  - Many algorithms available (spectral, Louvain, etc.)
  - Need to pick null model appropriate to problem
  - Extreme near-degeneracies in “good” local optima of Q
    - (B. H. Good, Y.-A. de Montjoye, & A. Clauset, *PRE*, 2010)

# "Multislice" Networks (Mucha et al, 2010)
## [a type of multilayer network]




- Traditional formulation for studying networks: Static networks with a single kind of edge and partitioned at a single spatial resolution
  - Also potentially sweep over multiple resolutions (or over multiple static snapshots) but in an ad hoc fashion
- Multislice framework: time-dependent, multiplex, and with communities at multiple scales
- Simple idea: Glue common brain regions across "slices" (i.e. "layers")

# *What is an appropriate null model?*

$$Q = \frac{1}{2W} \sum_{i,j} B_{ij} \delta(C_i, C_j), \; B_{ij} = A_{ij} - P_{ij}$$

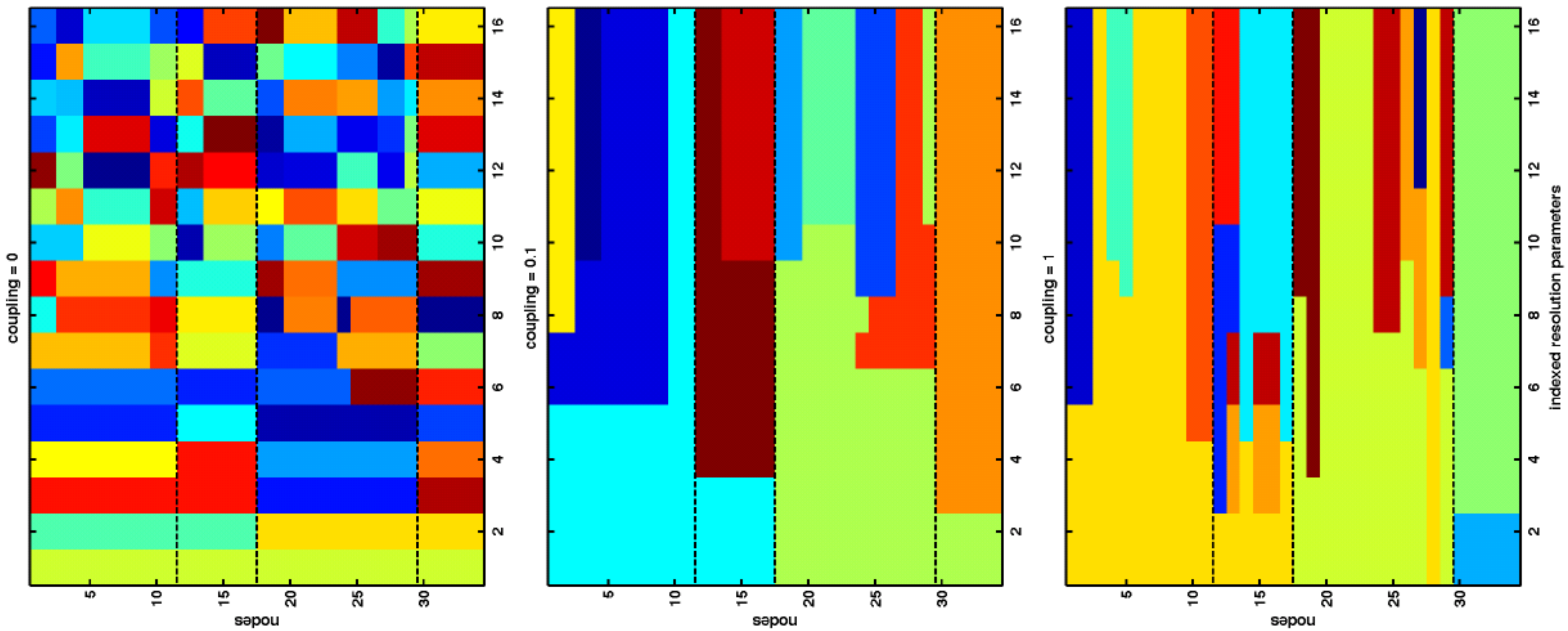
- Each layer is a network (static, single type of edge) with a specified spatial resolution of interest
- Different layers can mean: different value of resolution parameter, different time snapshot, different type of connection
- Have both intra-layer edges & inter-layer edges
- How to choose a null model?

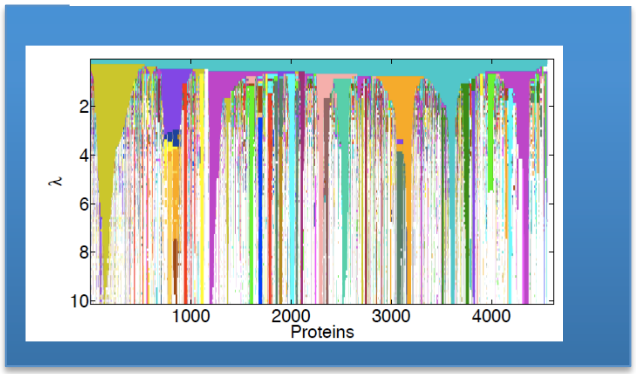
# Multislice Modularity

- Find communities algorithmically by optimizing "multislice modularity"
  - We derived this function in Mucha et al, 2010
    - Laplacian dynamics: find communities based on how long random walkers are trapped there. Exponentiate and then linearize to derive modularity.
    - Generalizes derivation of ordinary modularity from R. Lambiotte, J.-C. Delvenne, &. M Barahona, arXiv:0812.1770
  - Brain region *x* in layer *r* is a *different node* from brain region *x* in layer *s*
    - A *layer* could come from e.g. similarities between regions computed during some time window

$$Q_{\text{multislice}} = \frac{1}{2\mu} \sum_{ijsr} \left\{ \left( A_{ijs} - \gamma_s \frac{k_{is} k_{js}}{2m_s} \right) \delta_{sr} + \delta_{ij} C_{jsr} \right\} \delta(g_{is}, g_{jr})$$

# Example: Zachary Karate Club




$$C_{jsr} = \{0, \omega\}$$

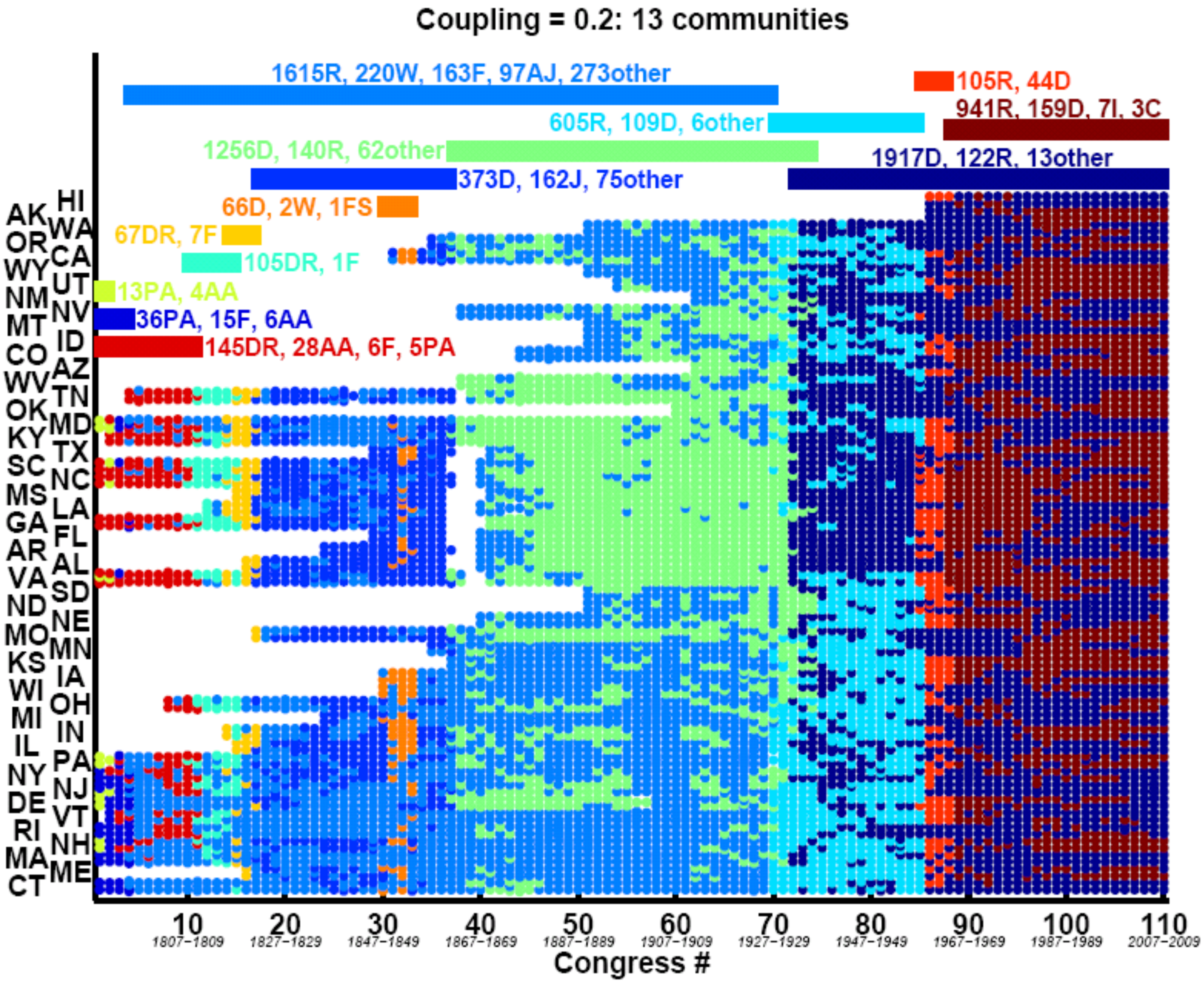
# Roll-Call Voting Networks

## (example to illustrate effect of parameter ω)

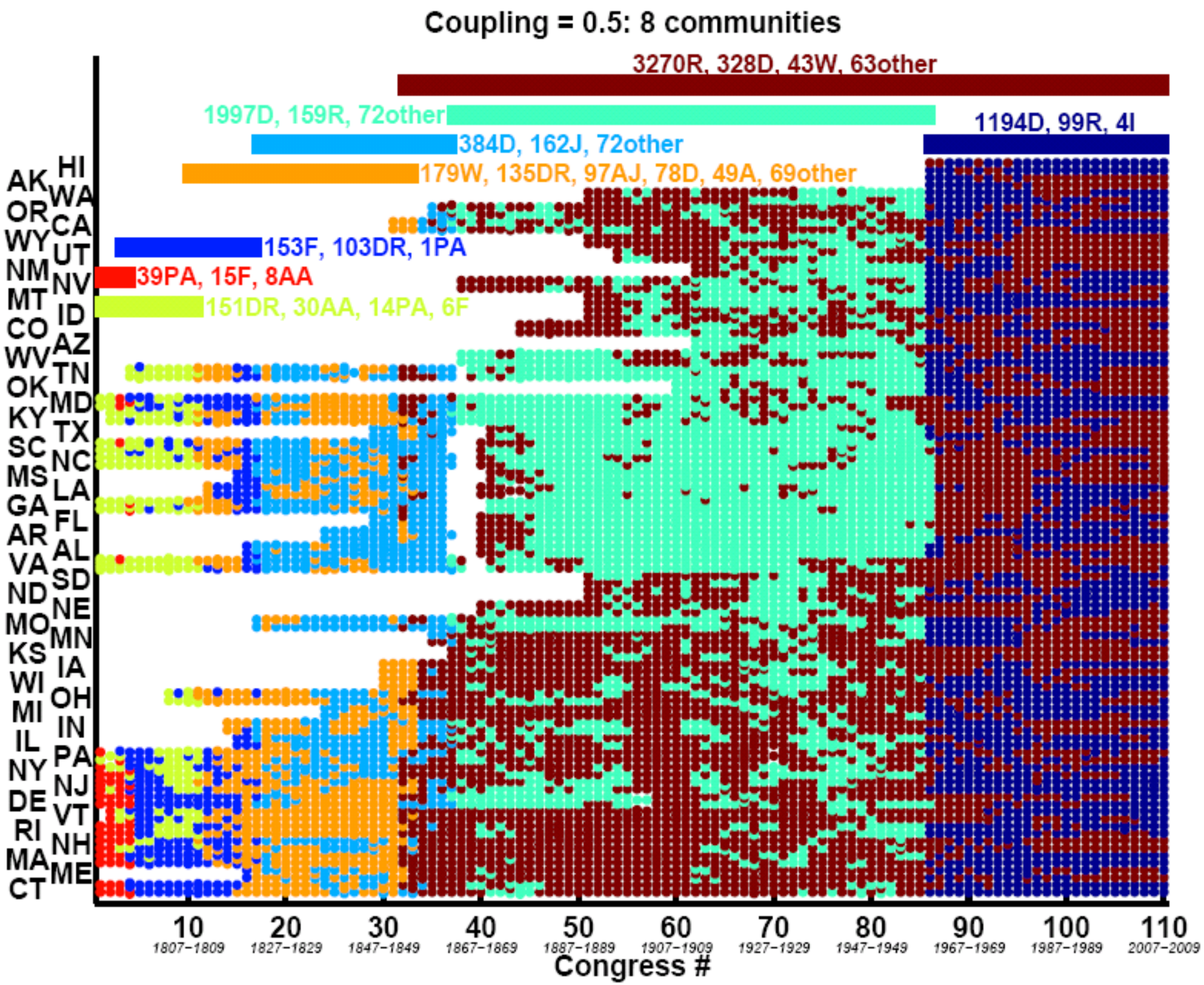
$$A_{ij} = \frac{1}{b_{ij}} \sum_k \alpha_{ijk} \,, \tag{1}$$

where $\alpha_{ijk}$ equals 1 if legislators $i$ and $j$ voted the same on bill $k$ and 0 otherwise and $b_{ij}$ is the total number of bills on which both legislators voted. The matrix $A$ encodes a network of weighted affiliations between legislators, with weights determined by the similarity of their roll-call records

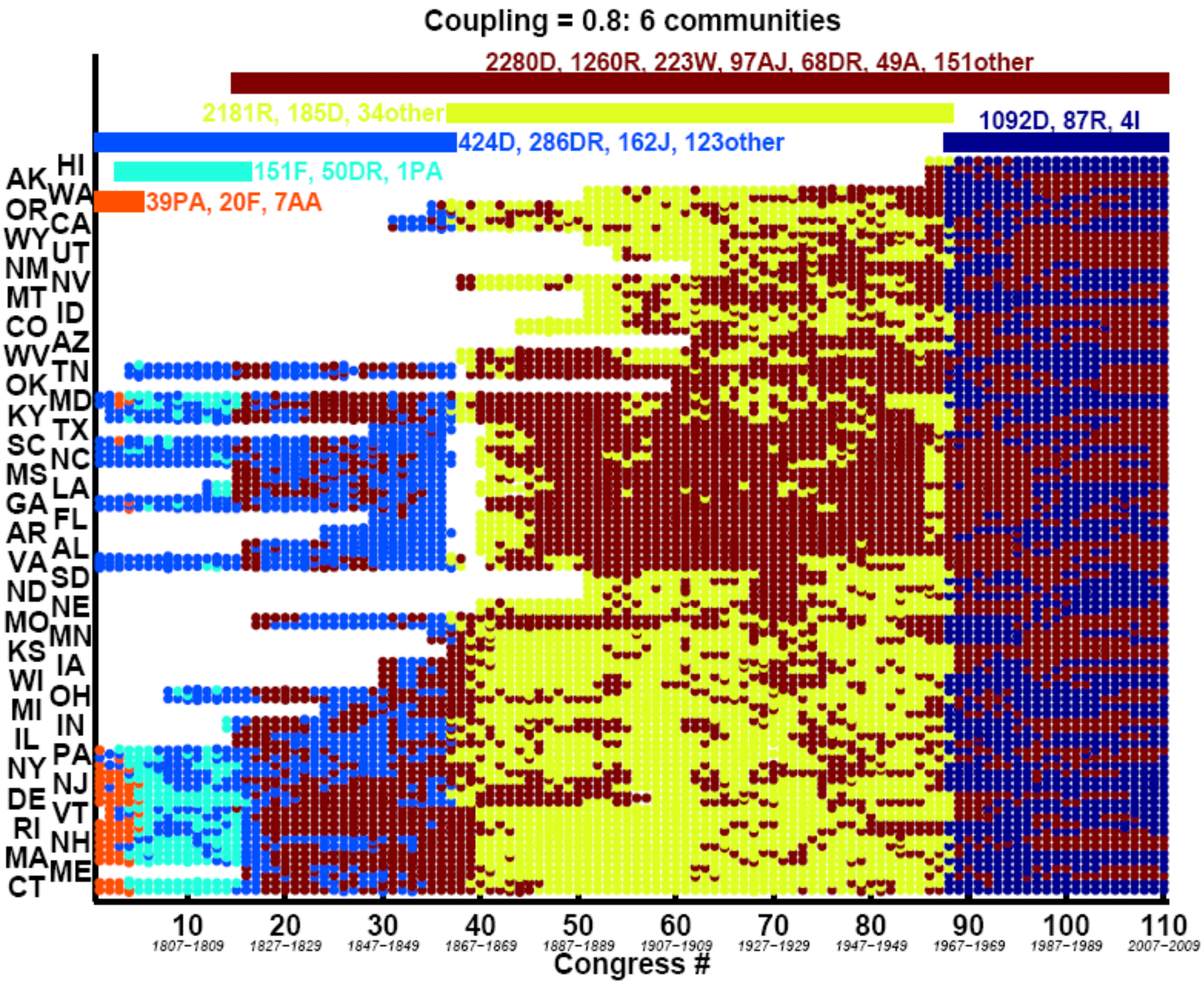
- A. S. Waugh, L. Pei, J. H. Fowler, P. J. Mucha, & M. A. Porter [2012], *arXiv: 0907.3509* (without multilayer formulation)
- Modularity Q as a measure of polarization
- Can calculate how closely each legislator is tied to their community (e.g. by looking at magnitude of corresponding component of leading eigenvector of modularity matrix if using a spectral optimization method)
- Medium levels of optimized modularity as a predictor of majority turnover
  - By contrast, leading political science measure doesn't give statistically significant indication
- One network slice for each two-year Congress



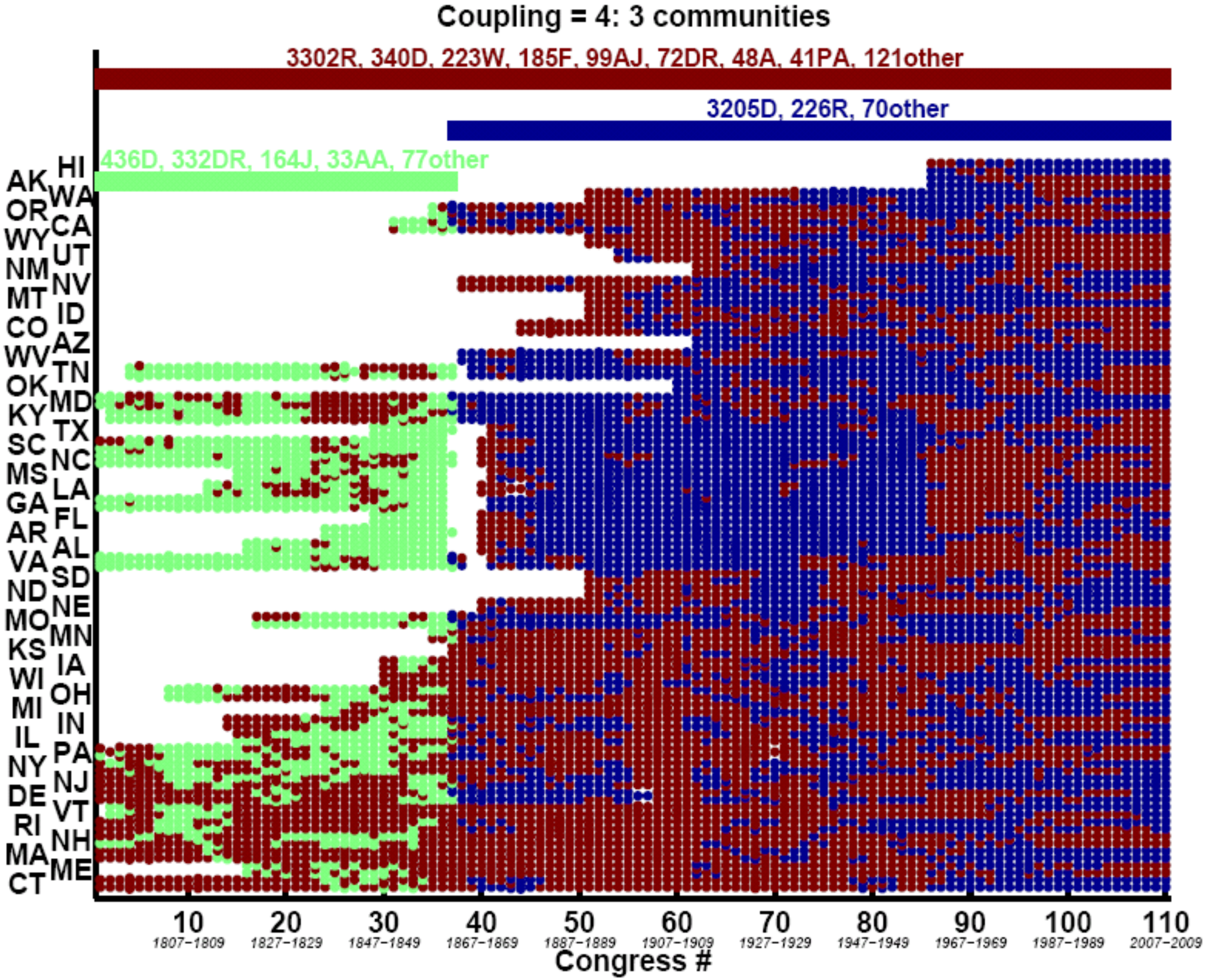

P. J. Mucha & M. A. Porter, *Chaos*, Vol. 20, No. 4, 041108 (2010)



Coupling = 0.5: 8 communities
3270R, 328D, 43W, 63other
1997D, 159R, 72other
1194D, 99R, 4I
384D, 162J, 72other
179W, 135DR, 97AJ, 78D, 49A, 69other
153F, 103DR, 1PA
39PA, 15F, 8AA
151DR, 30AA, 14PA, 6F
AK HI
OR WA
WY CA
NM UT
MT NV
CO ID
WV AZ
OK TN
KY MD
SC TX
MS NC
GA LA
AR FL
VA AL
ND SD
MO NE
KS MN
WI IA
MI OH
IL IN
NY PA
DE NJ
RI VT
MA NH
CT ME
10 1807−1809
20 1827−1829
30 1847−1849
40 1867−1869
50 1887−1889
60 1907−1909
70 1927−1929
80 1947−1949
90 1967−1969
100 1987−1989
110 2007−2009
Congress #

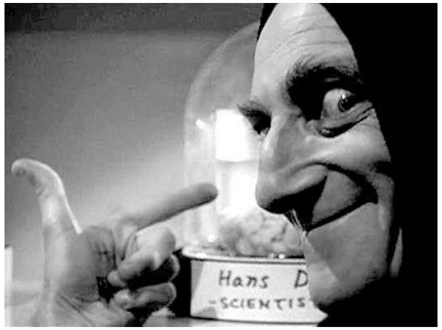


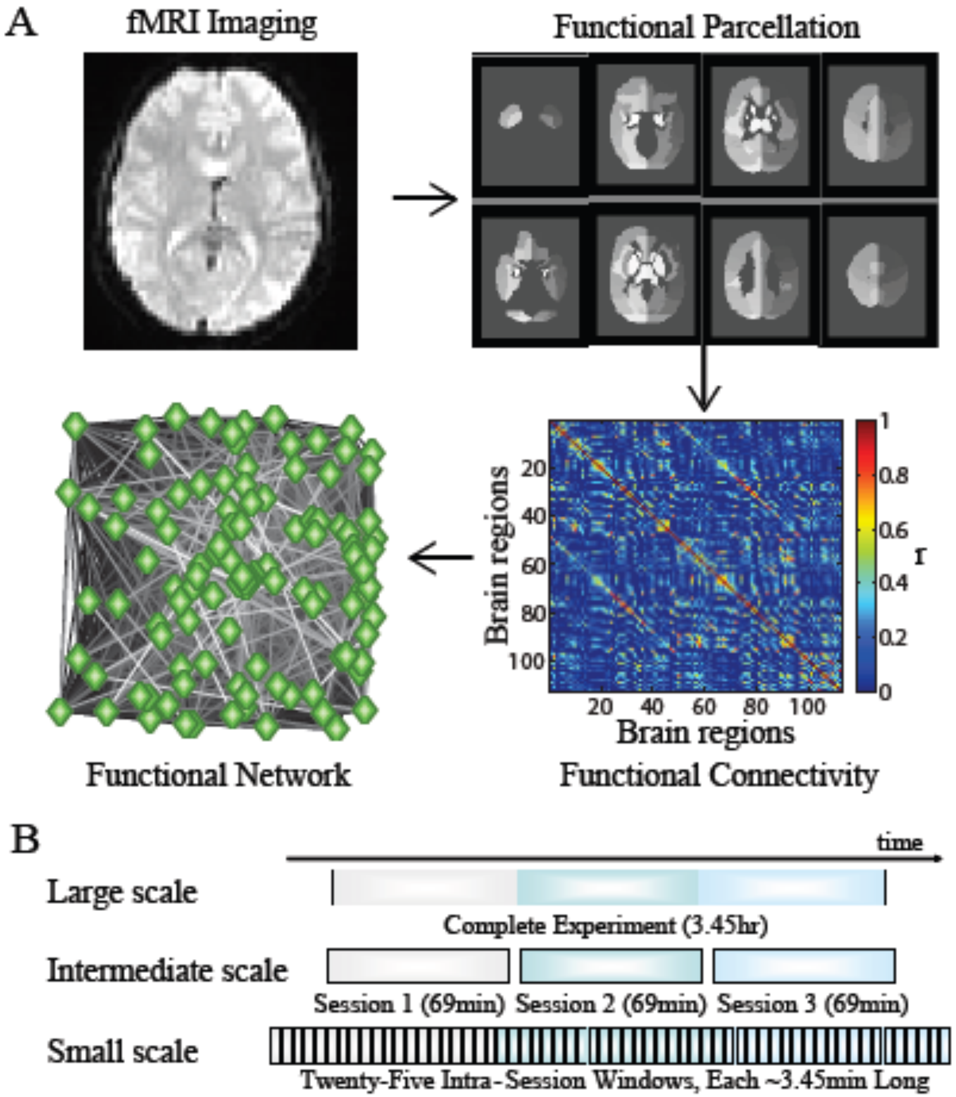
Coupling = 0.8: 6 communities
2280D, 1260R, 223W, 97AJ, 68DR, 49A, 151other
2181R, 185D, 34other
1092D, 87R, 4I
424D, 286DR, 162J, 123other
151F, 50DR, 1PA
39PA, 20F, 7AA
HI AK WA OR CA WY UT NM NV MT ID CO AZ WV TN OK MD KY TX SC NC MS LA GA FL AR AL VA SD ND NE MO MN KS IA WI OH MI IN IL PA NY NJ DE VT RI NH MA ME CT
10 20 30 40 50 60 70 80 90 100 110
1807−1809 1827−1829 1847−1849 1867−1869 1887−1889 1907−1909 1927−1929 1947−1949 1967−1969 1987−1989 2007−2009
Congress #

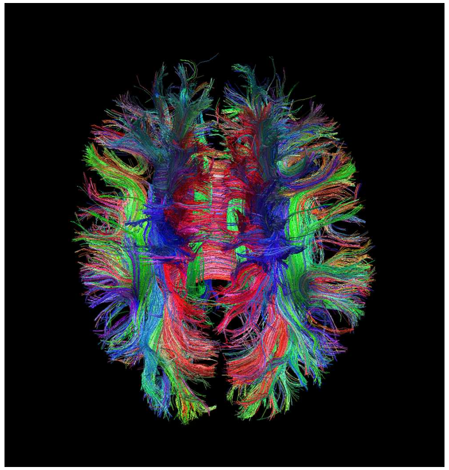


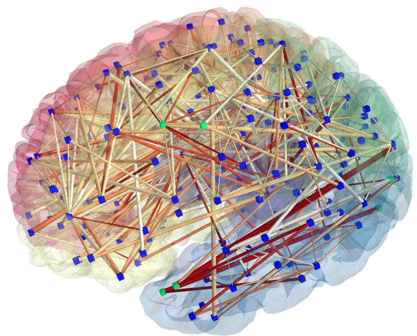
Coupling = 4: 3 communities
3302R, 340D, 223W, 185F, 99AJ, 72DR, 48A, 41PA, 121other
3205D, 226R, 70other
436D, 332DR, 164J, 33AA, 77other
HI
AK
WA
OR
CA
WY
UT
NM
NV
MT
ID
CO
AZ
WV
TN
OK
MD
KY
TX
SC
NC
MS
LA
GA
FL
AR
AL
VA
SD
ND
NE
MO
MN
KS
IA
WI
OH
MI
IN
IL
PA
NY
NJ
DE
VT
RI
NH
MA
ME
CT
10
1807–1809
20
1827–1829
30
1847–1849
40
1867–1869
50
1887–1889
60
1907–1909
70
1927–1929
80
1947–1949
90
1967–1969
100
1987–1989
110
2007–2009
Congress #

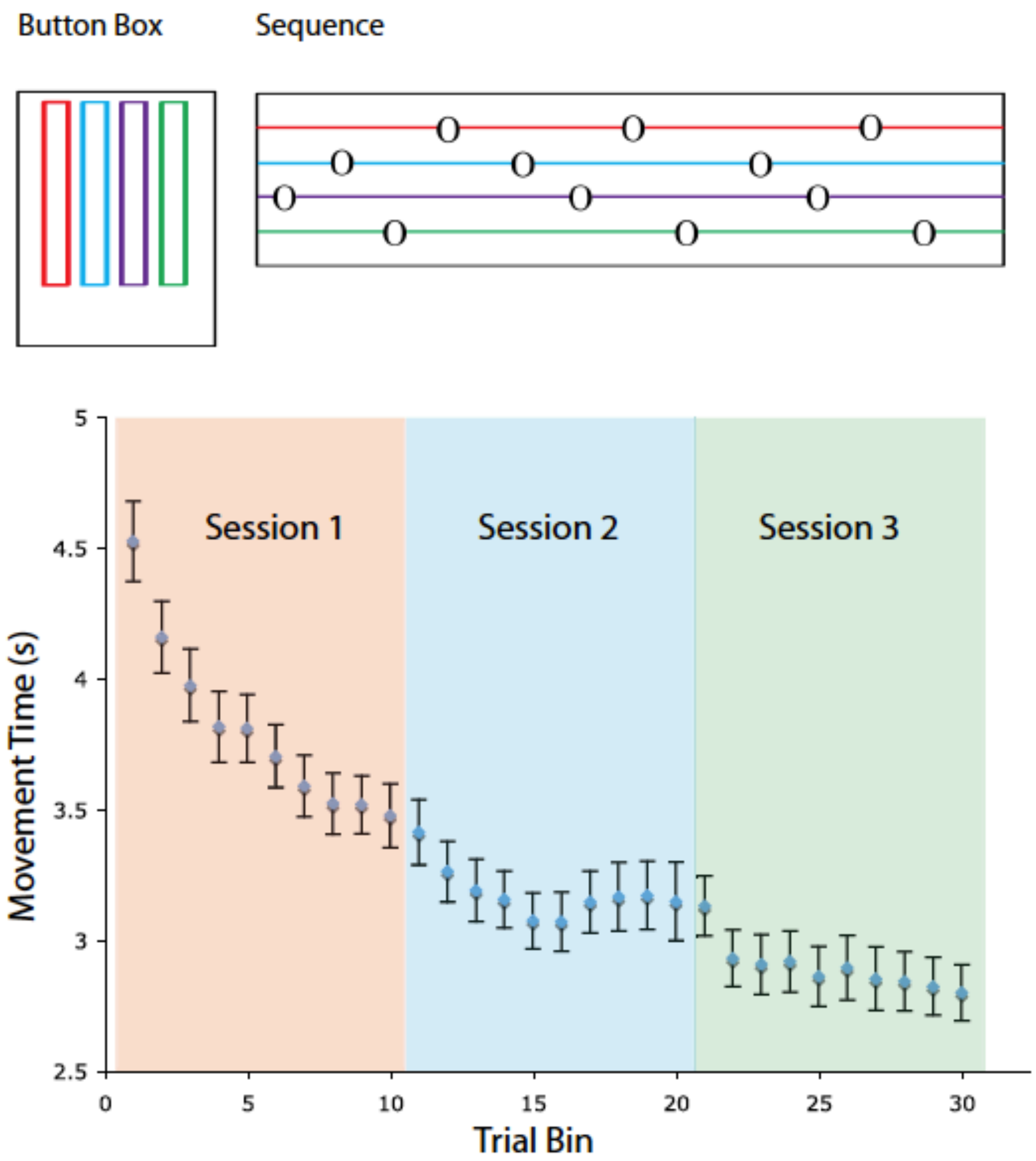
# Braiiiiiiiiiiiiins



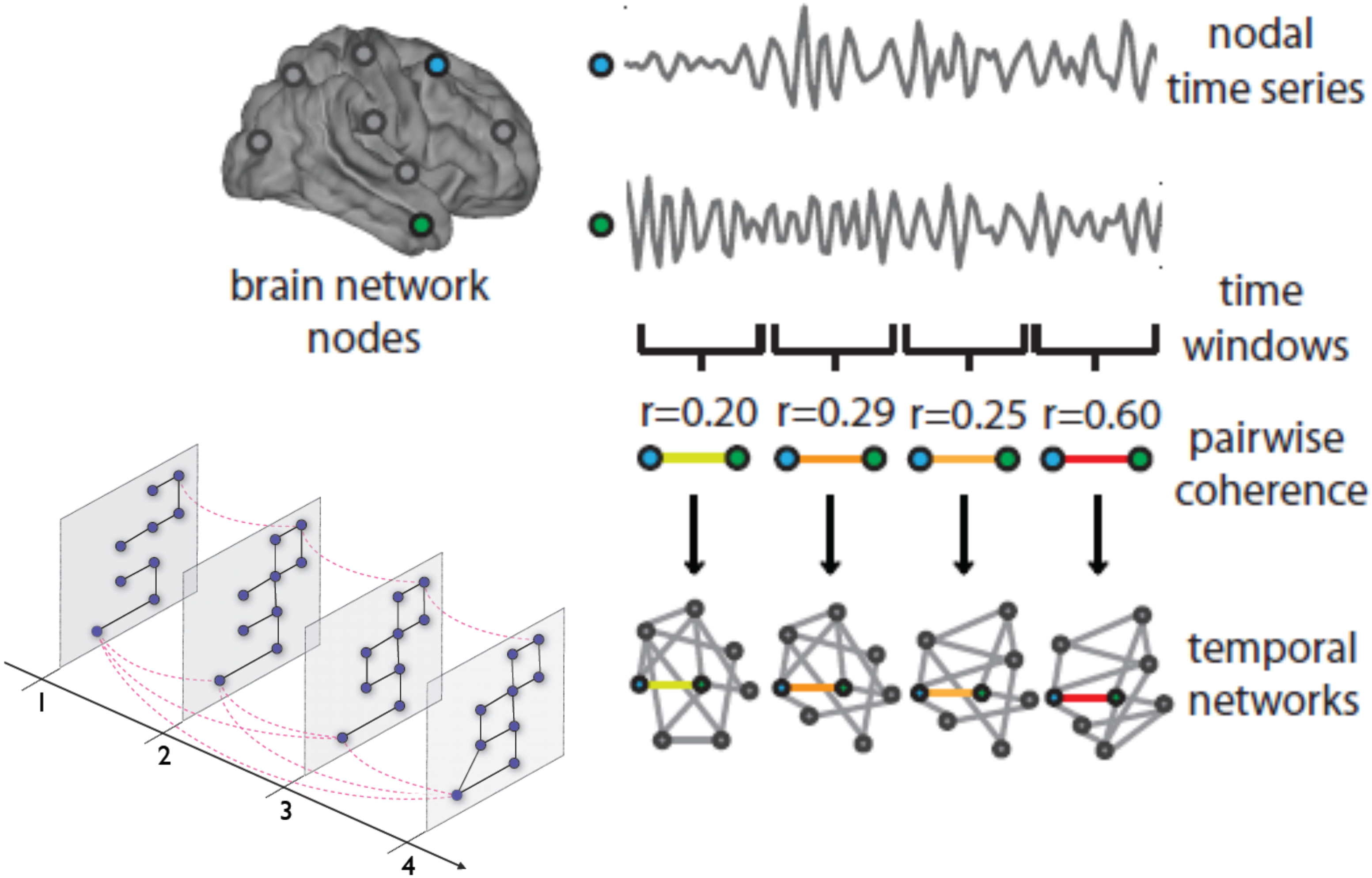







Button Box
Sequence
Session 1
Session 2
Session 3
Movement Time (s)
Trial Bin
5
4.5
4
3.5
3
2.5
0
5
10
15
20
25
30

# Constructing Time-Dependent Networks



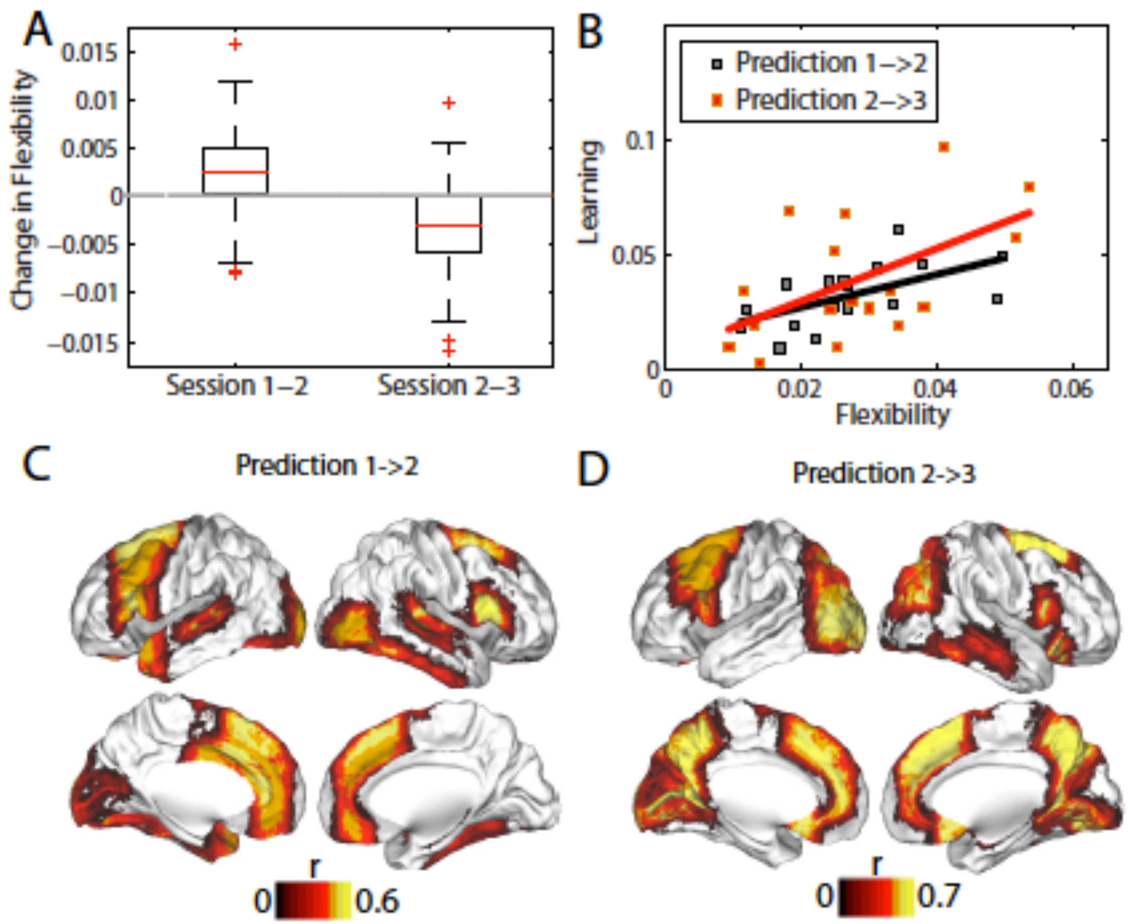
# Dynamic Reconfiguration of Human Brain Networks During Learning

(Bassett et al, *PNAS*, 2011)

- fMRI data: network from correlated time series

- Examine role of modularity in human learning by identifying dynamic changes in modular organization over multiple time scales

- Main result: flexibility, as measured by allegiance of nodes to communities, in one session predicts amount of learning in subsequent session



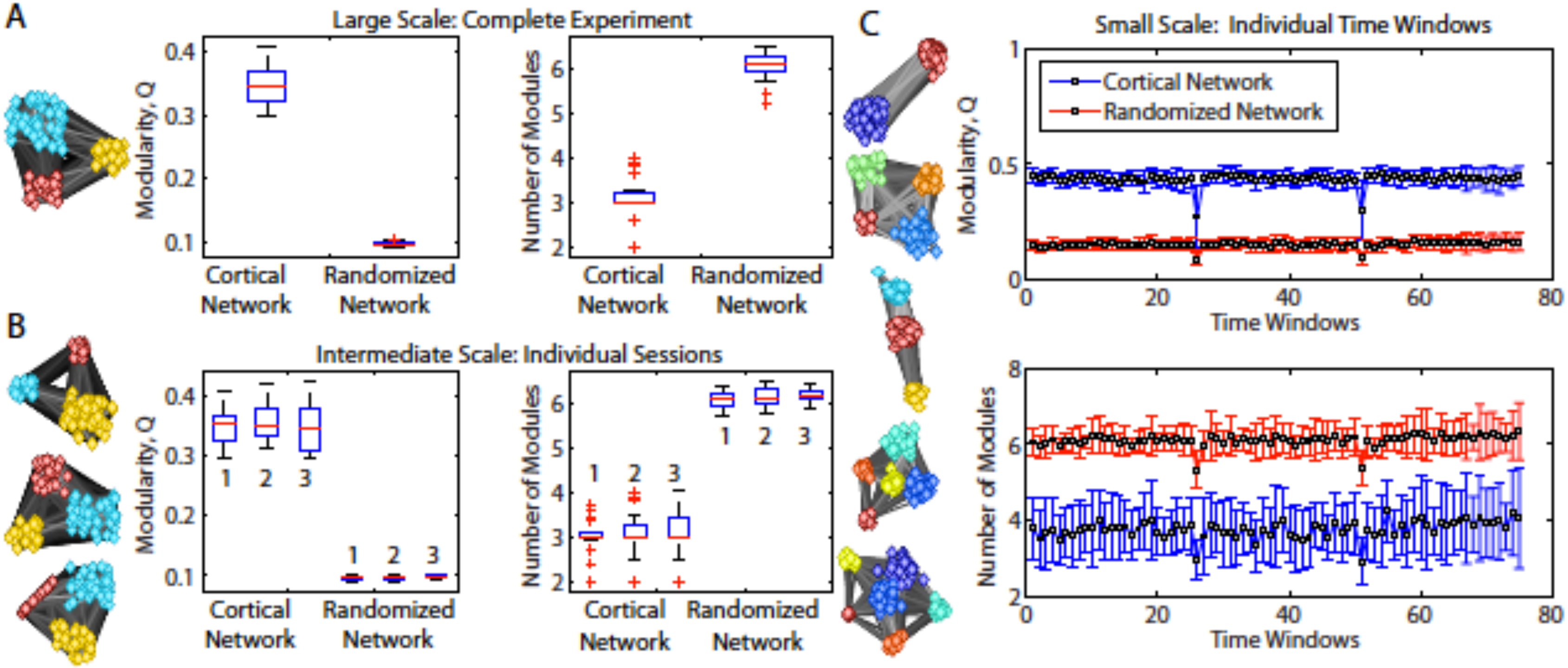
# Stationarity and Flexibility

- Community stationarity ζ (autocorrelation over time of community membership):

$$U(t, t+m) \equiv \frac{|G(t) \cap G(t+m)|}{|G(t) \cup G(t+m)|} \qquad \zeta \equiv \frac{\sum_{t=t_0}^{t'-1} U(t, t+1)}{t' - t_0 - 1}$$

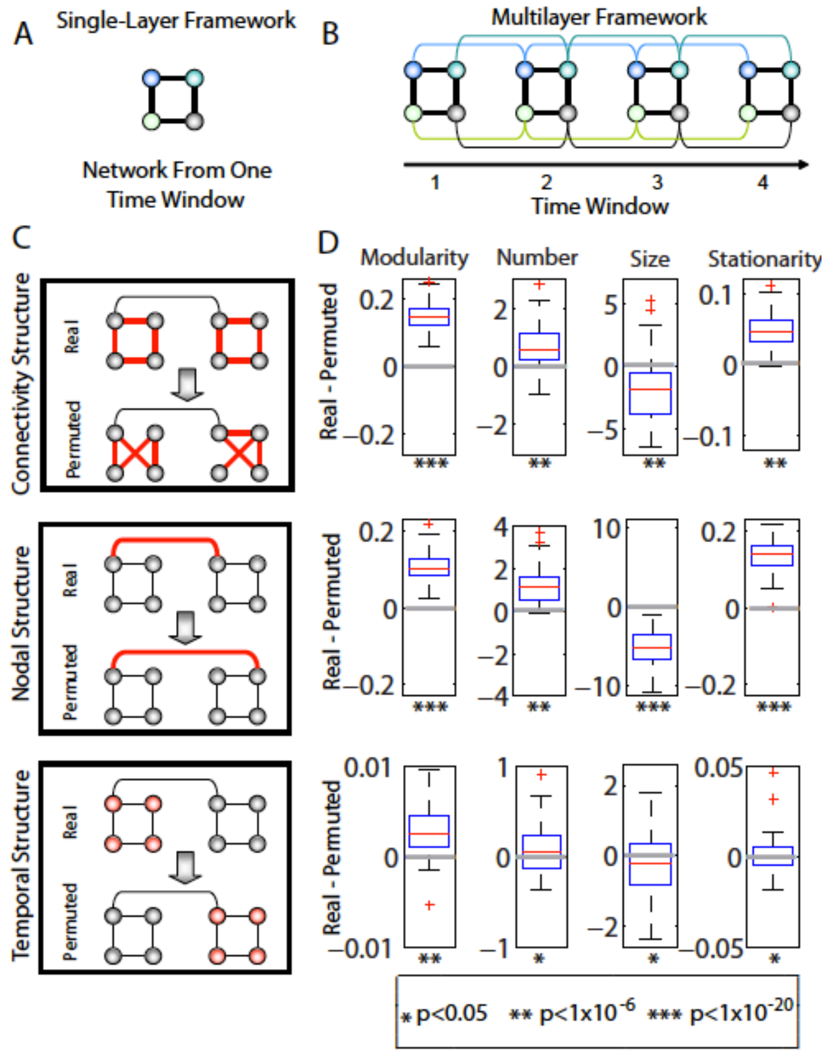
- Node flexibility:
  – $f_i$ = number of times node i changed communities divided by total number of possible changes
  – Flexibility $f = \langle f_i \rangle$

# Time Evolution of Static Communities



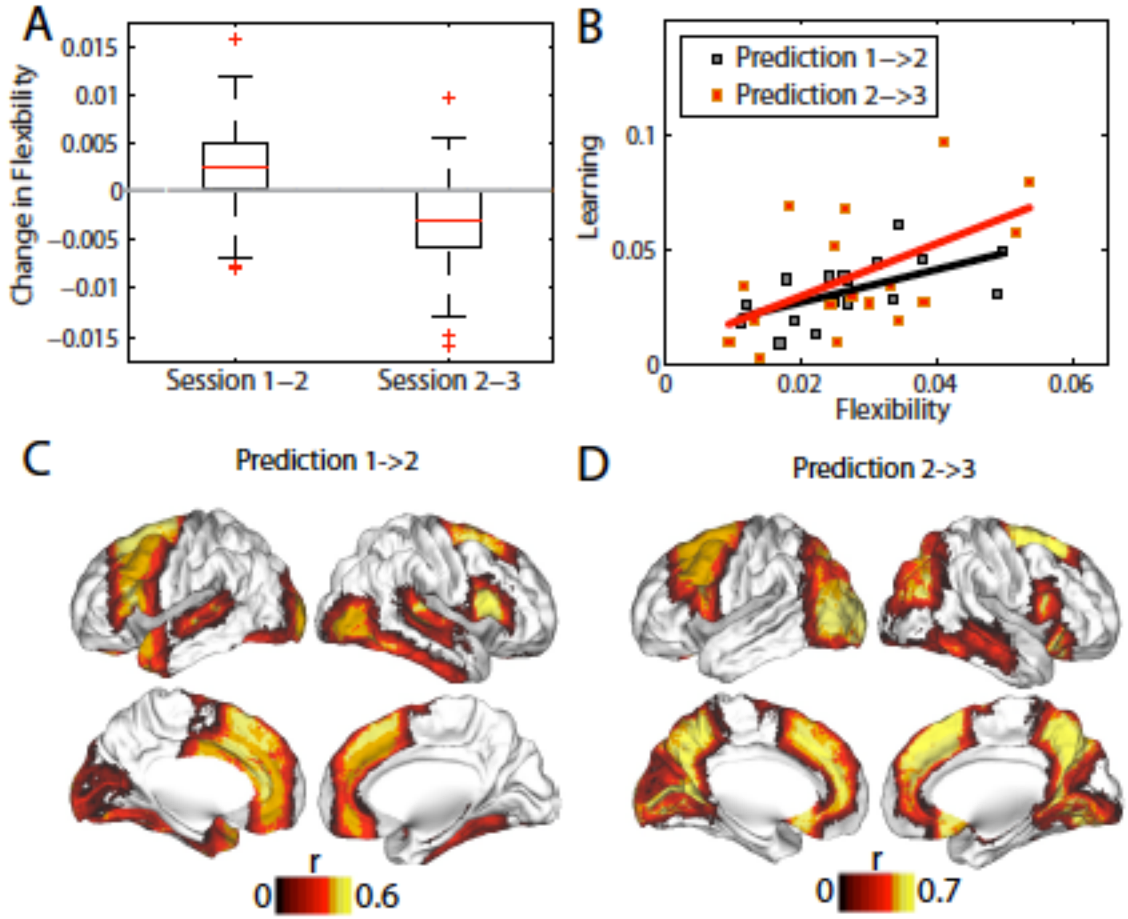
# Dynamic Community Structure

- Investigating community structure in a multilayer framework requires consideration of new null models
- Many more details!
  - E.g., Robustness of results to choice of size of time window, size of inter-slice coupling, particular definition of flexibility, complicated modularity landscape, definition of 'similarity' of time series, etc.



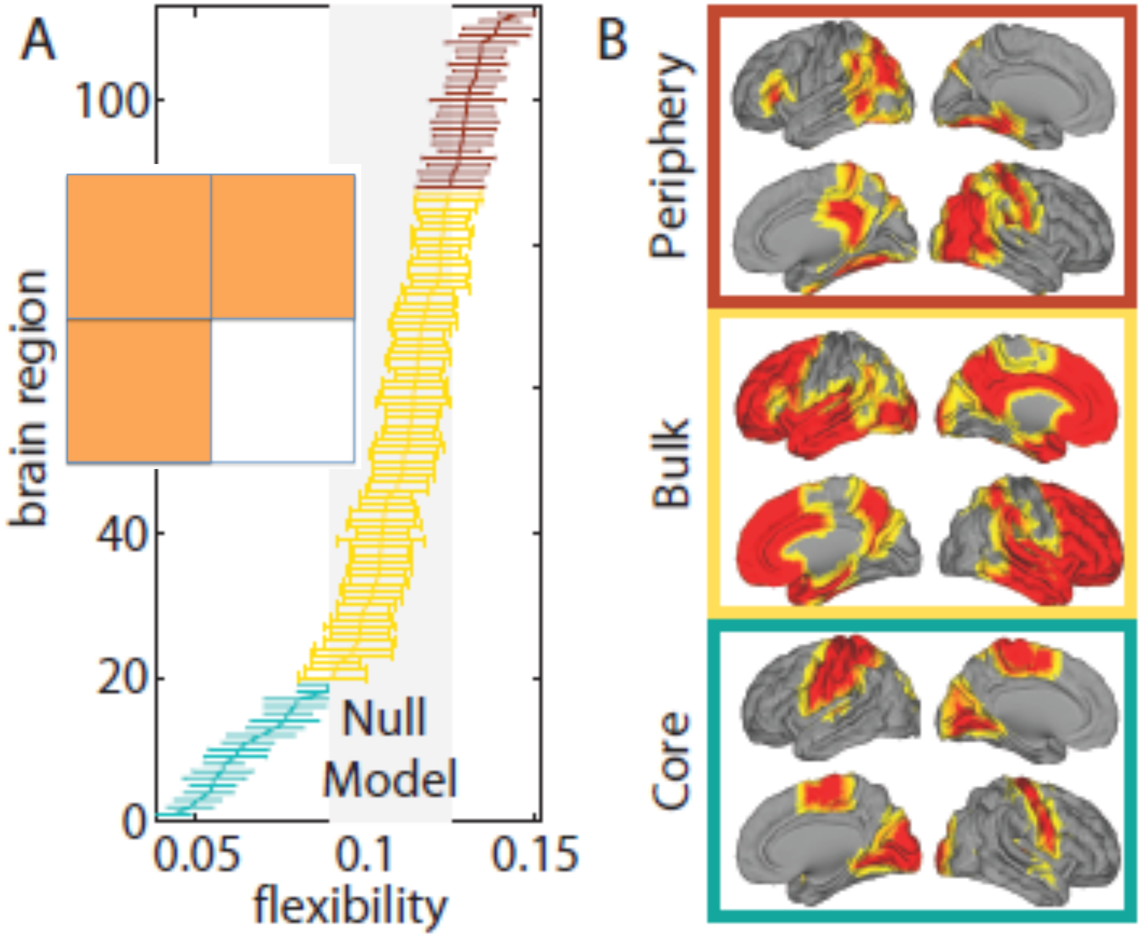
# Dynamic Reconfiguration of Human Brain Networks During Learning
(Bassett et al, *PNAS*, 2011)

- fMRI data: network from correlated time series
- Examine role of modularity in human learning by identifying dynamic changes in modular organization over multiple time scales
- Main result: flexibility, as measured by allegiance of nodes to communities, in one session predicts amount of learning in subsequent session

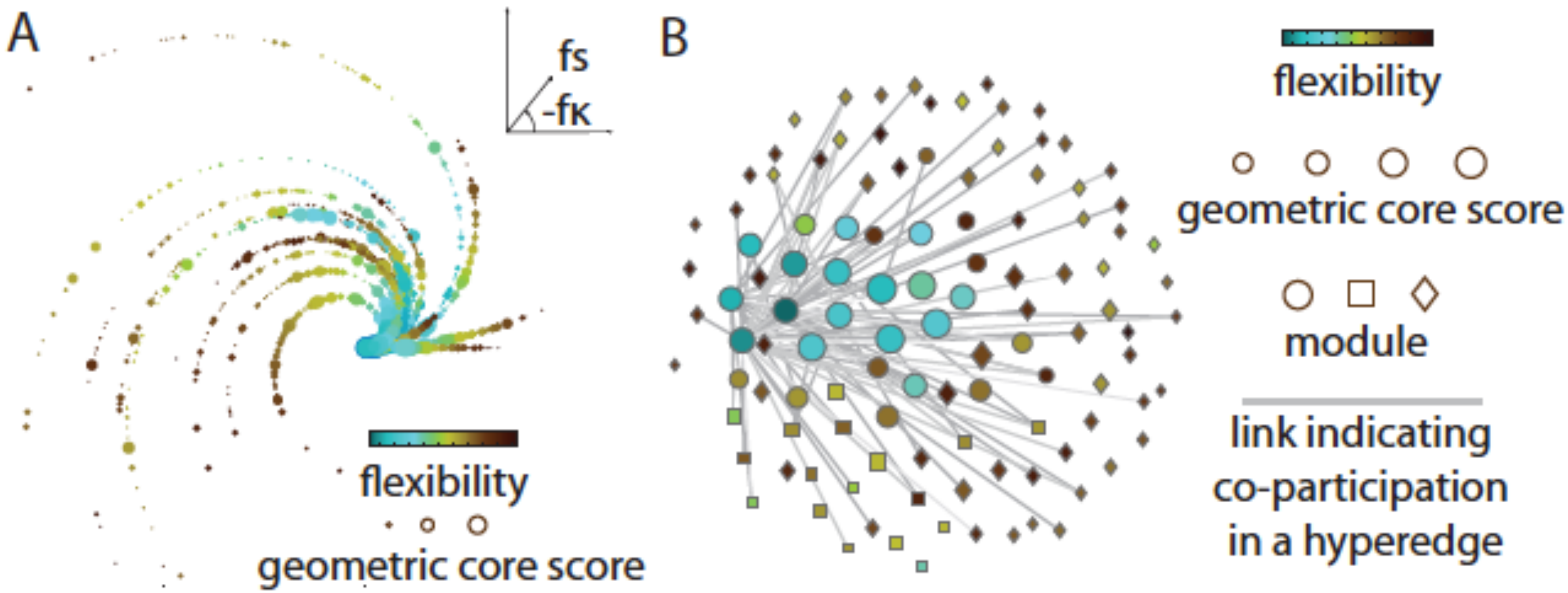


# Which Brain Regions are "Flexible"?

- D. S. Bassett, N. F. Wymbs, M. P. Rombach, MAP, P. J. Mucha, & S. T. Grafton, *PLoS Comp. Bio.* **9**(9): 1003171 (2013)
- Flexible nodes are consistently in a "periphery" as computed for static networks encompassing given time windows
- Nodes that are not flexible (call them "stiff") are consistently in a structural core in these static networks
- Note: I have not discussed our methodology for computing core-periphery structure
  - M. P. Rombach, MAP, J. H. Fowler, & P. J. Mucha, *SIAM J. App. Math.*, in press (2014); arXiv:1202.2684



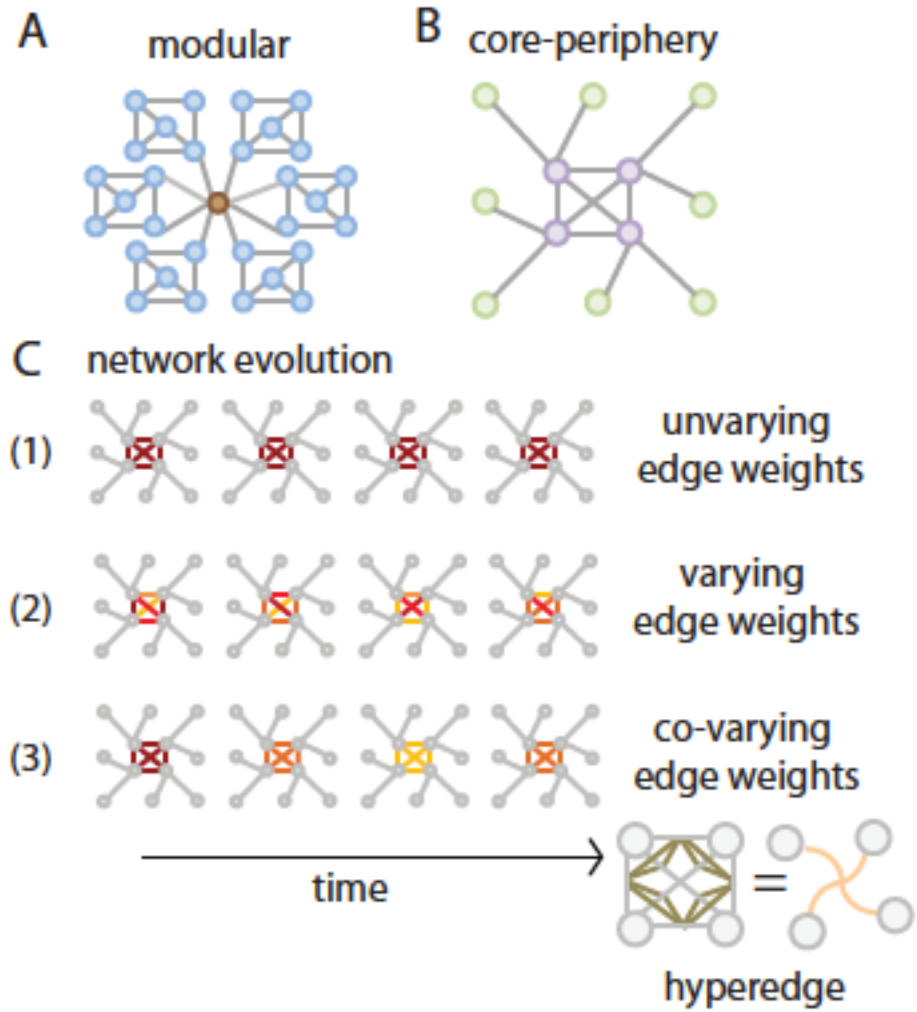
# Temporal Core ≈ Structural Core!




Temporal core-periphery organization ≈ Structural core-periphery organization!

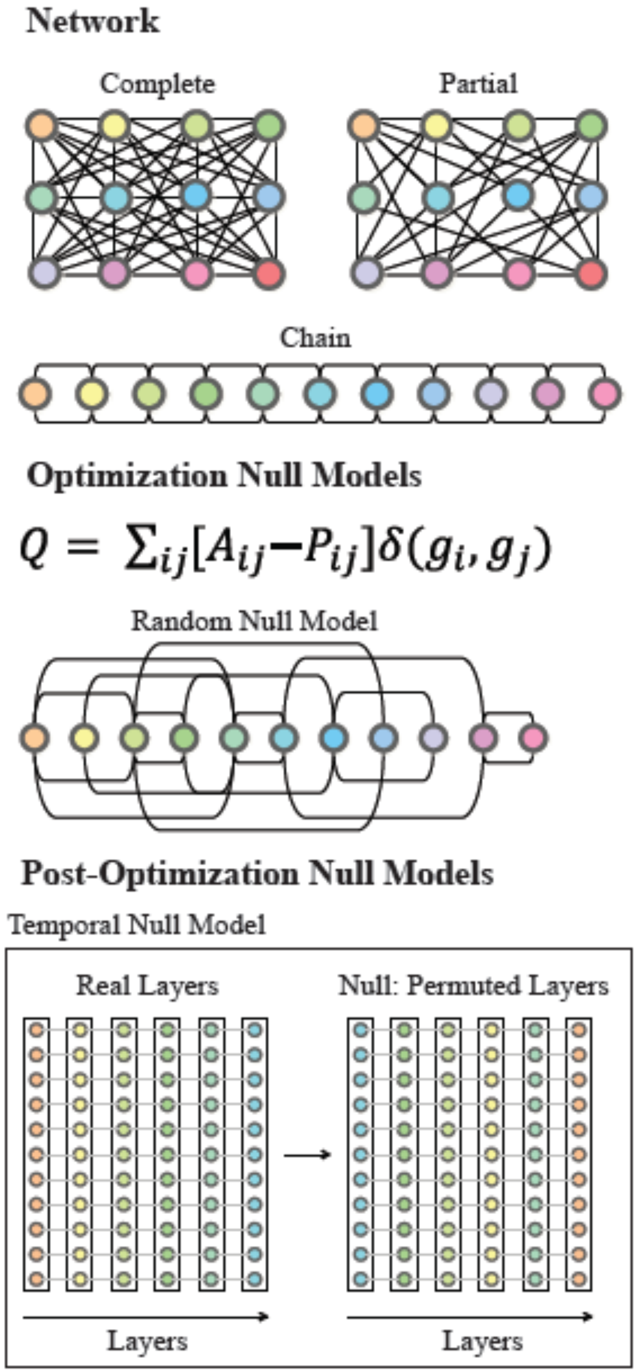
# Cross-Links

- D. S. Bassett, N. F. Wymbs, MAP, P. J. Mucha, & S. T. Grafton, *Chaos*, **24**(1): 013112 (2014)
- Cross-links connect time-dependent edges to each other based on the similarity of their time series
  - Yield hyperedges that connect the associated nodes
- Try to discern which network evolution scenario occurs
  - Most hyperedges involve core (i.e. stiff) regions



# Development of Null Models for Multilayer Networks

- D. S. Bassett, M. A. Porter, N. F. Wymbs, S. T. Grafton, J. M. Carlson, & P. J. Mucha, *Chaos*, **23**(1): 013142 (2013)
- Additional structure in adjacency tensors gives more freedom (and responsibility) for choosing null models.
- Null models that incorporate information about a system
  - E.g. chain null model fixes network topology but randomizes network "geometry" (edge weights)
- Also: Examine null models from shuffling time series directly (before turning into a network)
- Structural (γ) versus temporal resolution parameter (ω)
  - More generally, how to choose inter-layer (off-diagonal) terms $C_{jrs}$



# Conclusions

- Multilayer networks and adjacency tensors: their time has come
  - Review article: *Journal of Complex Networks*, in press (arXiv:1309.7233)
- Mesoscale structure of networks can be very insightful
  - E.g. community structure, core-periphery structure
- Generalization of community structure to multilayer networks allows investigation of more realistic situations while throwing away less data
- Insights on both brain and behavioral data
  - Dynamic reconfiguration of human brain networks during learning
    - Flexibility of nodes predicts simple motor learning
  - Good correspondence between structure and dynamics: flexible nodes in network periphery, and stiff nodes in network core
- Code available:
  - Code for Louvain optimization method for multislice modularity: http://netwiki.amath.unc.edu/GenLouvain
  - Code for visualizing networks: http://netwiki.amath.unc.edu/VisComms
  - Code for visualization of multilayer networks and some other calculations with multilayer networks: http://www.plexmath.eu/?page_id=327
- Thanks: James S. McDonnell Foundation, EPSRC, FET-Proactive project "PLEXMATH"