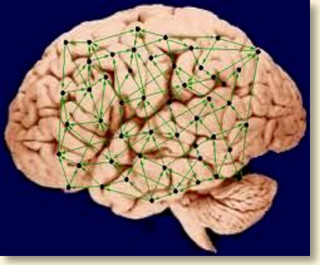
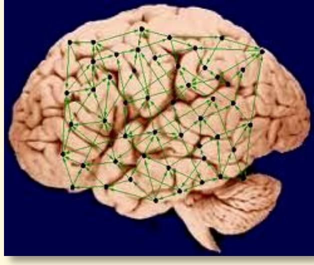
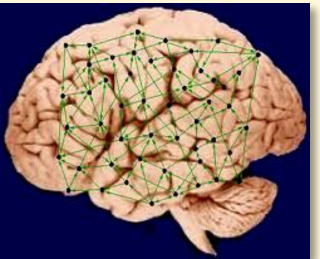
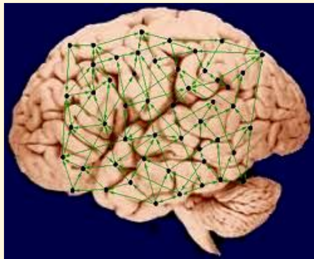



Multivariate Modeling and Inference for Brain Networks: ERGMs and Mixed Models

Sean L. Simpson
Dept. of Biostatistical Sciences
Wake Forest School of Medicine

Outline

I. Motivation

II. Brain Network Construction and Description

III. Multivariate Modeling and Inference: ERGMs

IV. Multivariate Modeling and Inference: Mixed Models

V. Summary

VI. Useful References

I. Motivation

- Connectivity and network analyses have exploded over the last decade, and hold potential in helping us understand normal and abnormal brain function.

I. Motivation

- Connectivity and network analyses have exploded over the last decade, and hold potential in helping us understand normal and abnormal brain function.

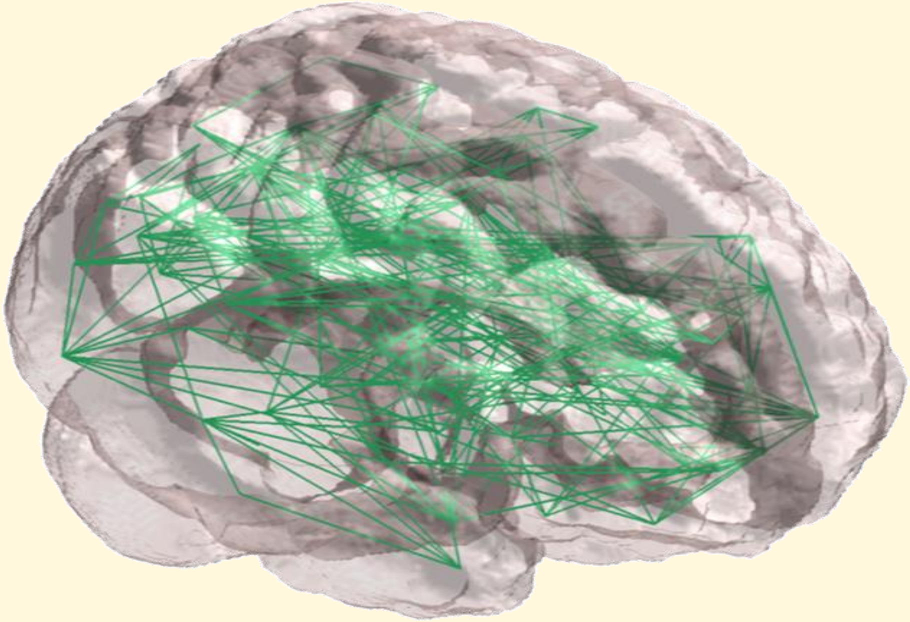
- FC analysis examines associations between time series in **specific** regions.

I. Motivation

- Connectivity and network analyses have exploded over the last decade, and hold potential in helping us understand normal and abnormal brain function.

- FC analysis examines associations between time series in **specific** regions.

- Network analysis quantifies associations between time series in **all** regions to create an interconnected representation of the brain (a brain network).

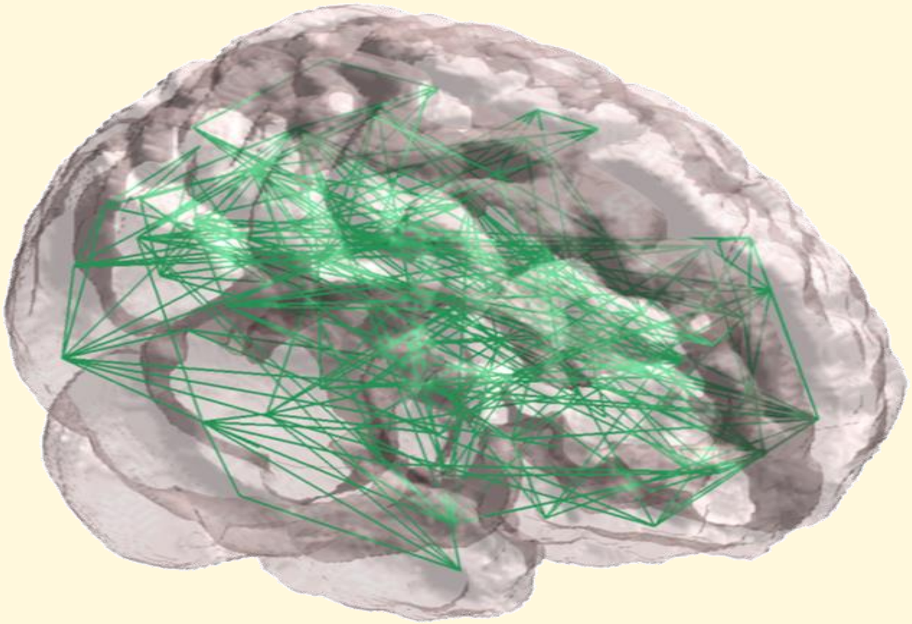


I. Motivation

- Connectivity and network analyses have exploded over the last decade, and hold potential in helping us understand normal and abnormal brain function.

- FC analysis examines associations between time series in **specific** regions.

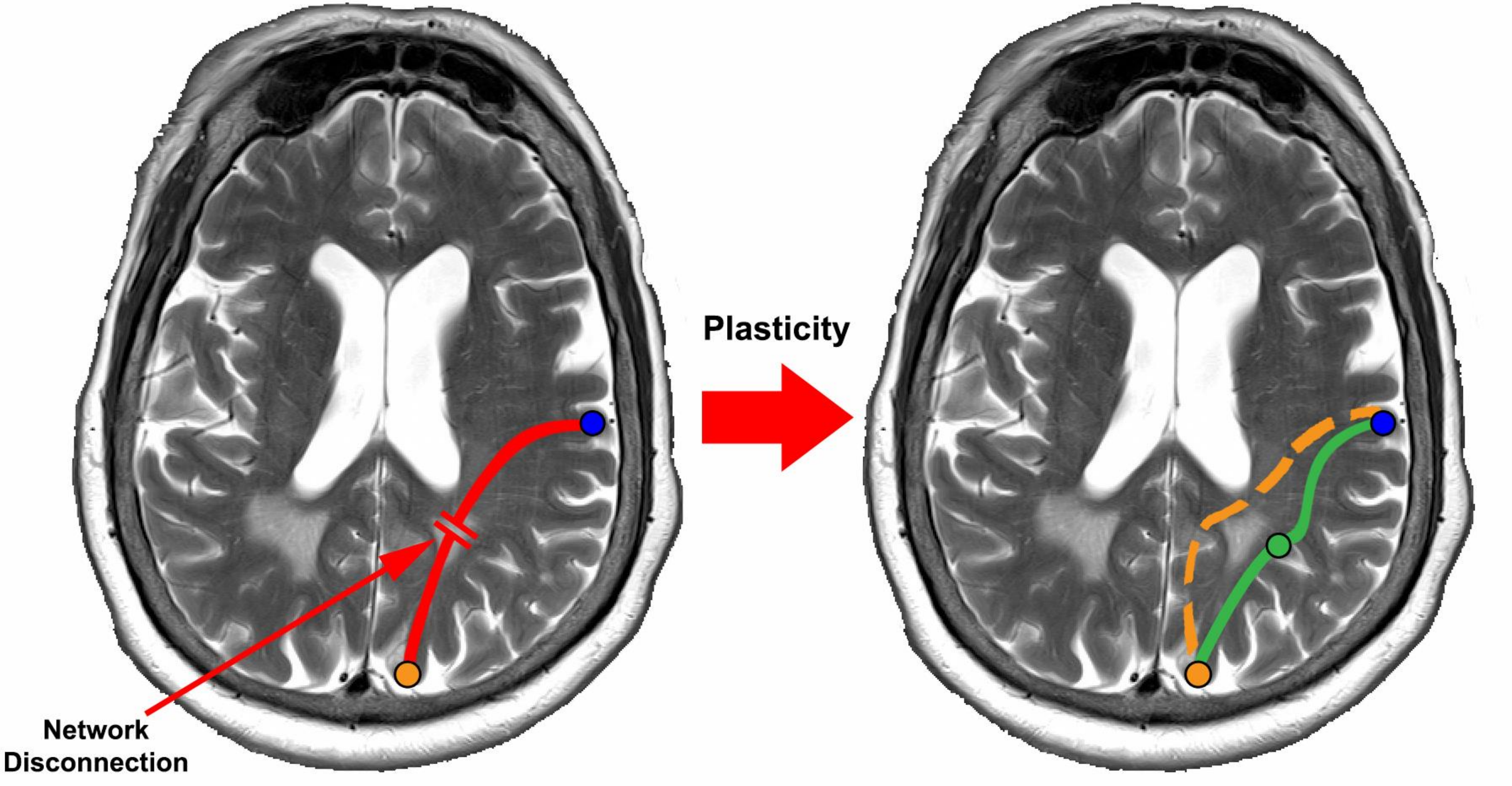
- Network analysis quantifies associations between time series in **all** regions to create an interconnected representation of the brain (a brain network).



- FC underlie network analyses, subtle distinction overlooked in the literature.

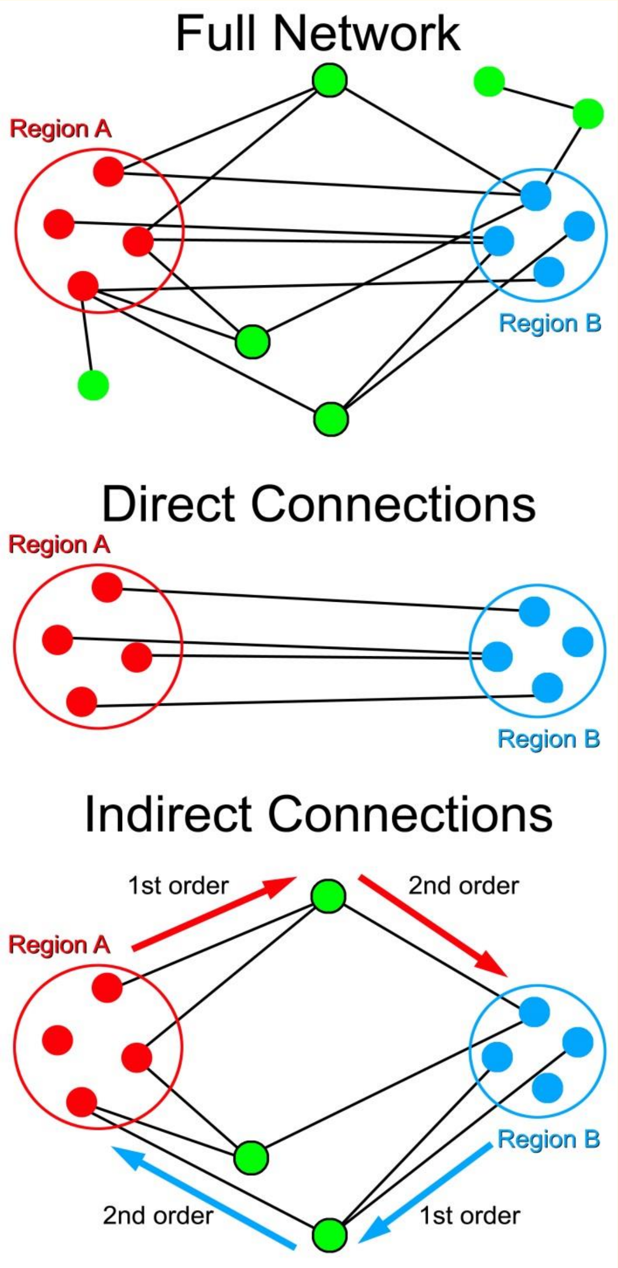
I. Motivation

- Systemic organization confers functional abilities as connections may be lost due to adverse health condition, but compensatory connections may develop to maintain organizational consistency and functional performance.



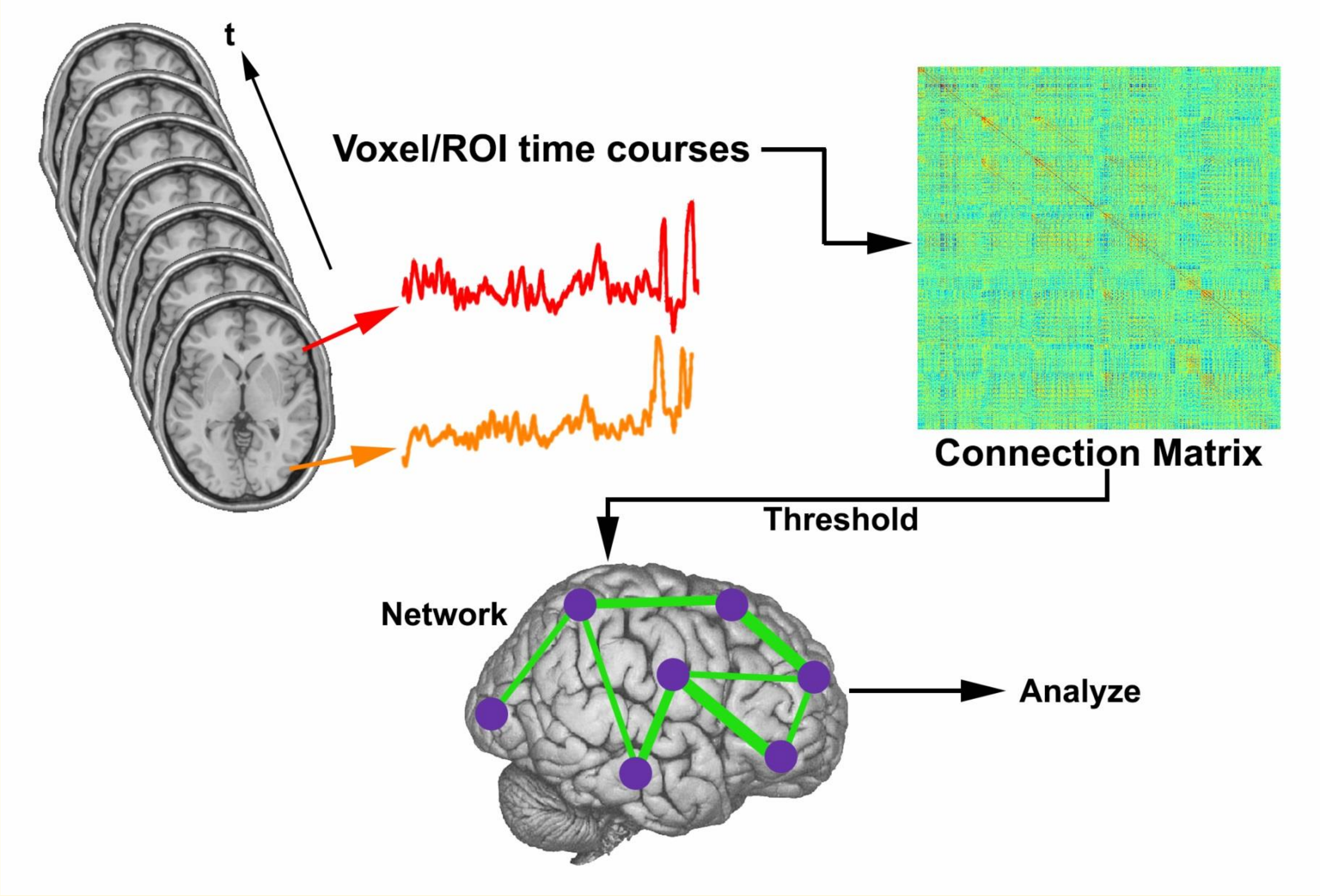
I. Motivation

- Also,...



II. Brain Network Construction and Description

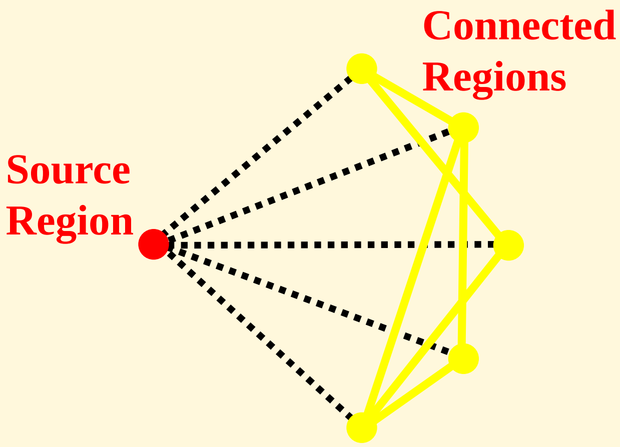
Schematic for generating network from fMRI time series



II. Brain Network Construction and Description

SMALL-WORLD METRICS

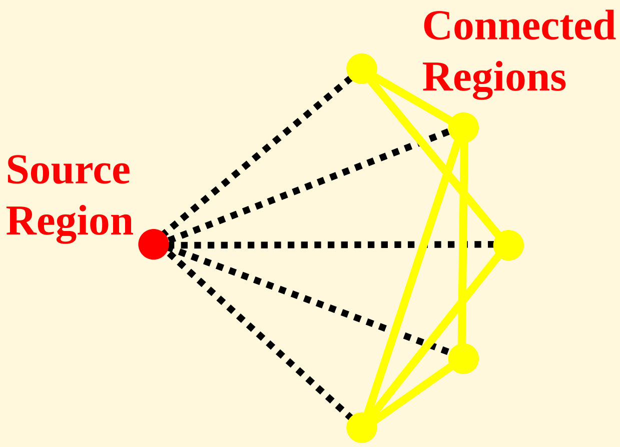
clustering coefficient (C)
Proportion of a region's connections that are connected to each other

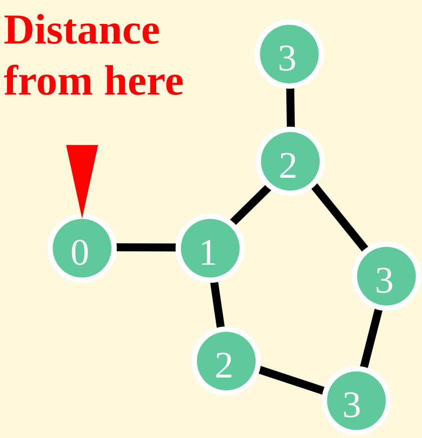


II. Brain Network Construction and Description

SMALL-WORLD METRICS

clustering coefficient (C)
Proportion of a region’s connections that are connected to each other

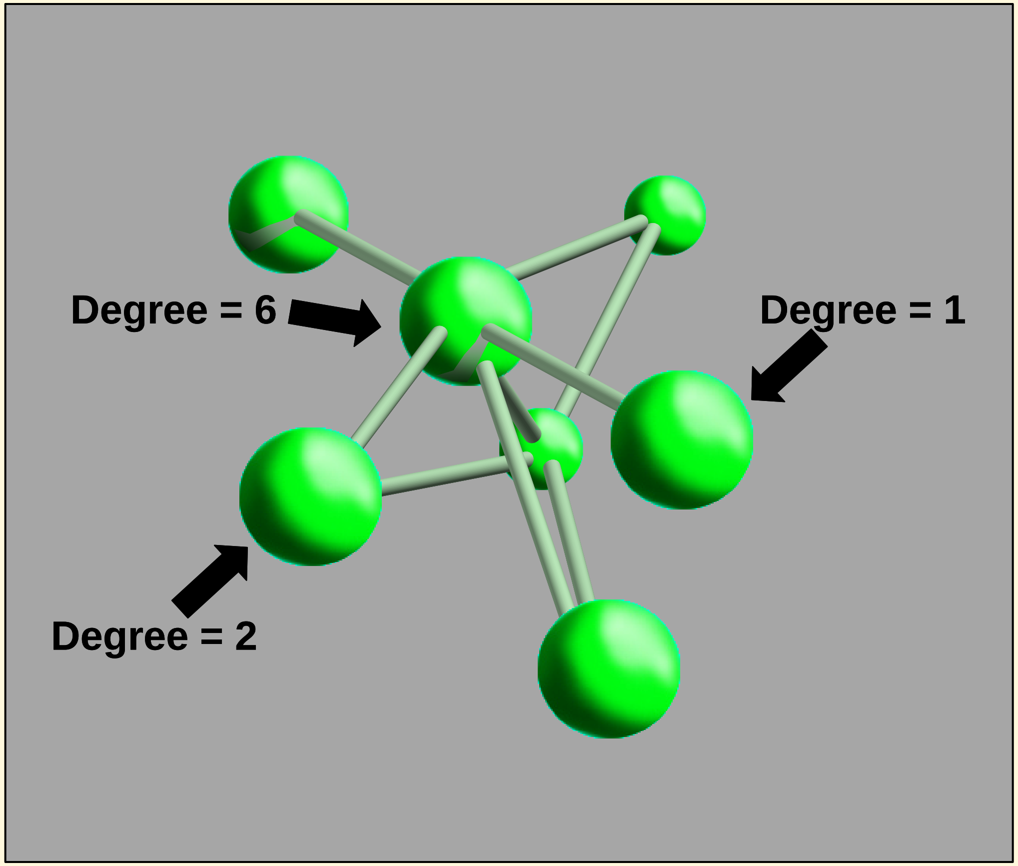




path length (L)

Average shortest distance between region pairs

Degree

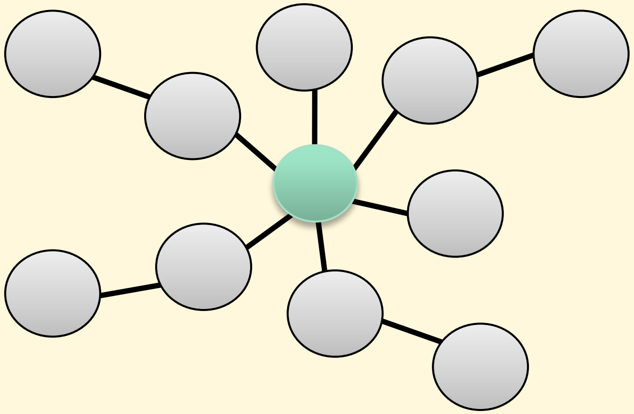


- Degree – *K*
  - Number of connections for each node
  - Distribution is assessed to evaluate network type/resilience properties
  - Assortativity is assessed to evaluate network type/resilience properties

II. Brain Network Construction and Description

Graph Centrality and Information Flow

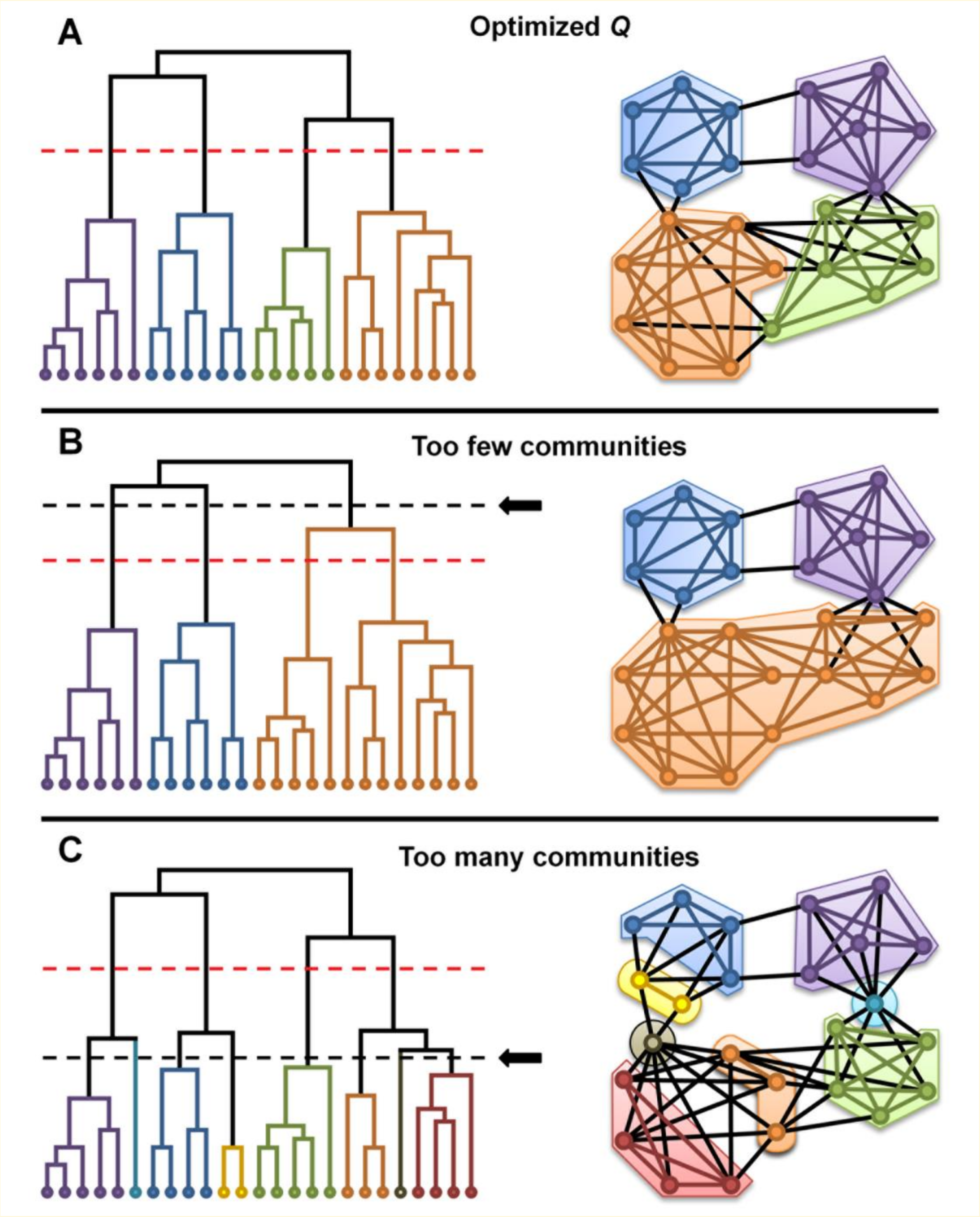
Leverage centrality (LC) *identifies nodes that have* ***high degree relative to neighbors***



Joyce et al. (2010)

II. Brain Network Construction and Description

Community Structure



II. Brain Network Construction and Description

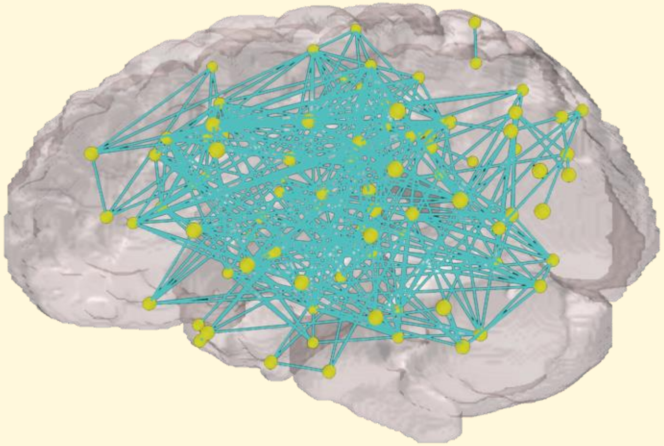
- Need a multivariate explanatory and predictive brain network model.

Data $\begin{cases} Y_i\text{: network of subject i} \\ X_i\text{: covariate information (network metrics, demographics, etc.)} \end{cases}$

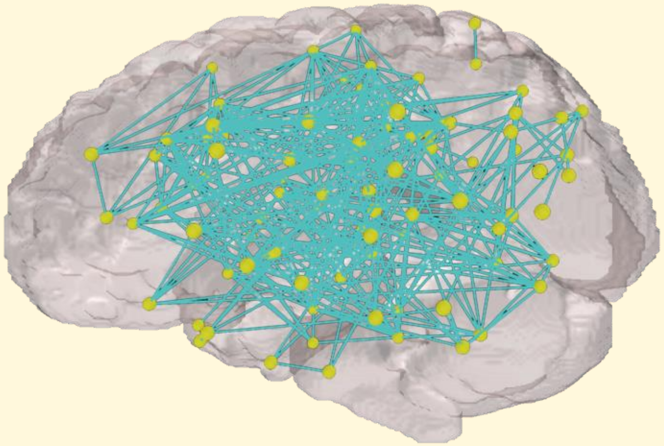
θ_i: parameters

Want $P(Y_i | X_i, \theta_i)$

II. Brain Network Construction and Description

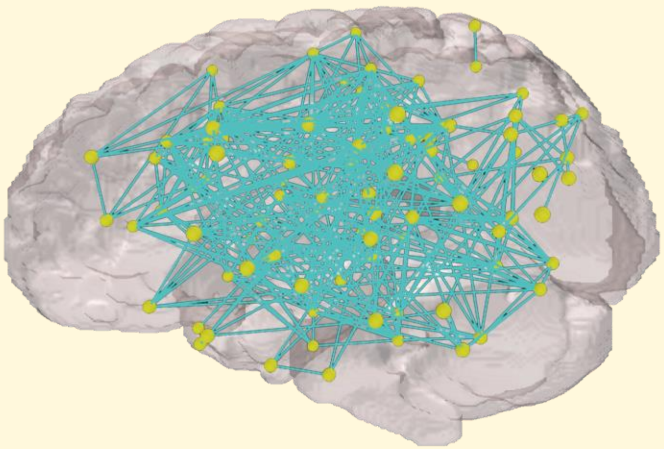


II. Brain Network Construction and Description

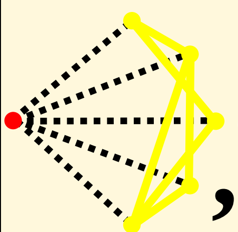
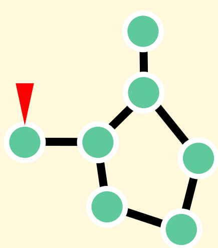
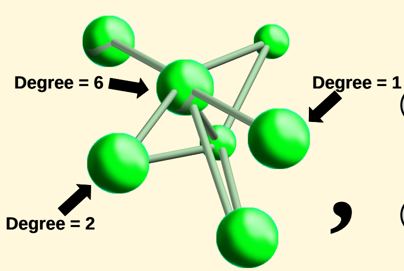


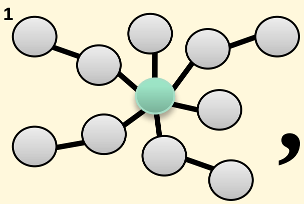
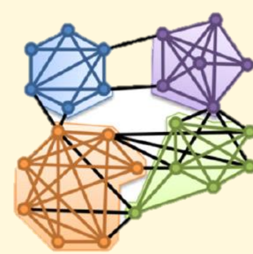
=

II. Brain Network Construction and Description



=

f [, , 

, , , X_i, θ_i]

III. Multivariate Modeling and Inference: ERGMs

Exponential random graph models have the following form:

$$P(\boldsymbol{Y} = \boldsymbol{y}) = \kappa(\boldsymbol{\theta})^{-1} \exp\{\boldsymbol{\theta}^{\mathrm{T}} \mathbf{g}(\boldsymbol{y})\} \tag{1}$$

where

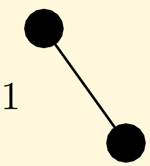
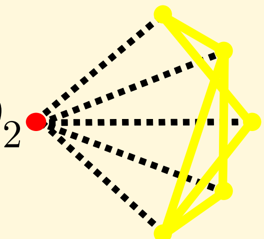
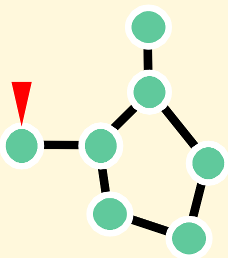
- $\boldsymbol{Y}$ — $n \times n$ (n nodes) random symmetric adjacency matrix, $Y_{ij} = 1$ if an edge exists between nodes i and j and $Y_{ij} = 0$ otherwise;
- $\mathbf{g}(\boldsymbol{y})$ — vector of prespecified network statistics (functions of network);
- $\boldsymbol{\theta}$ — vector of parameters associated with $\mathbf{g}(\boldsymbol{y})$ (importance, Δ log-odds);
- $\kappa(\boldsymbol{\theta})$ — normalizing constant ensuring probabilities sum to one.

Goal: Identify local metrics $\mathbf{g}(\boldsymbol{y})$ that concisely summarize the global (whole-brain) network structure.

III. Multivariate Modeling and Inference: ERGMs

- Once most appropriate statistics established, parameter profiles ($\boldsymbol{\theta}$) can be utilized to classify and compare whole-brain networks.

E.g., Best Model:

$P(\boldsymbol{Y} = \boldsymbol{y}) = \frac{1}{\kappa}\exp\Big\{\ \theta_1$  $+\ \theta_2$  $+\ \theta_3$  $\Big\}$

- Caveats: comparisons require use of a uniform set of predictors for all networks (due to predictor interdependencies) and balanced networks (same number of nodes for all networks) due to dependence of predictors on network size.

III. Multivariate Modeling and Inference: ERGMs

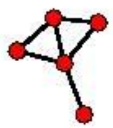
- Use graphical goodness-of-fit (GOF) approach (Hunter et al., 2008) to establish most appropriate set of explanatory metrics for each subject's brain network.

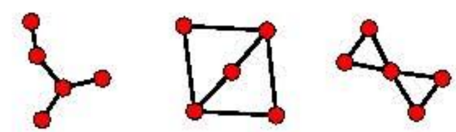
III. Multivariate Modeling and Inference: ERGMs

(Hunter, 2007)

Goodness of fit intuition

ERGM class		(approx) MLE		Fitted ERGM
$\exp\{\theta^t g(y)\}$	$\longrightarrow$	$\widehat{\theta}$	$\longrightarrow$	$\exp\{\widehat{\theta}^t g(y)\}$
	$\uparrow$			$\downarrow$
	y^{obs}			Randomly generated networks $\tilde{Y}_1, \tilde{Y}_2, \ldots$





- Question: How does y^{obs} "look" as a representative of the sample $\tilde{Y}_1, \tilde{Y}_2, \ldots$?

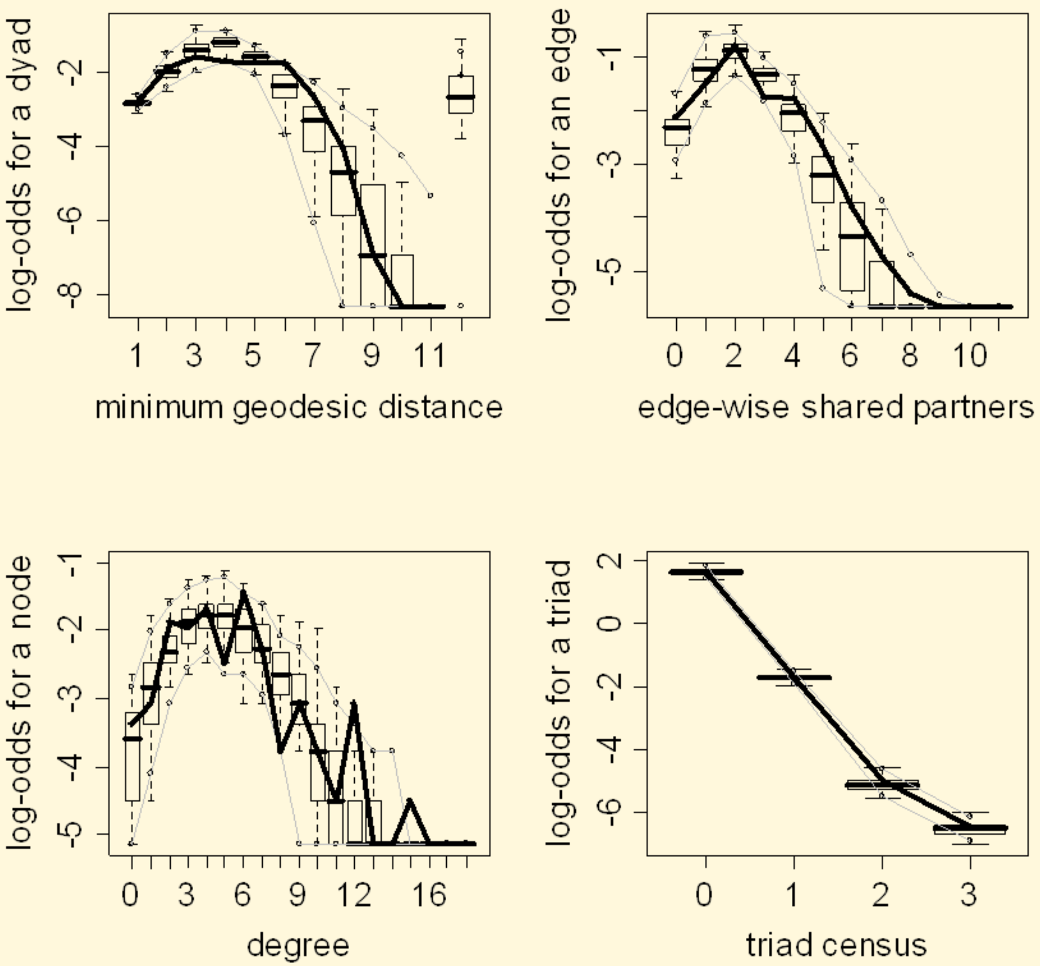
III. Multivariate Modeling and Inference: ERGMs

- Use graphical goodness-of-fit (GOF) approach (Hunter et al., 2008) to establish most appropriate set of explanatory metrics for each subject's brain network.

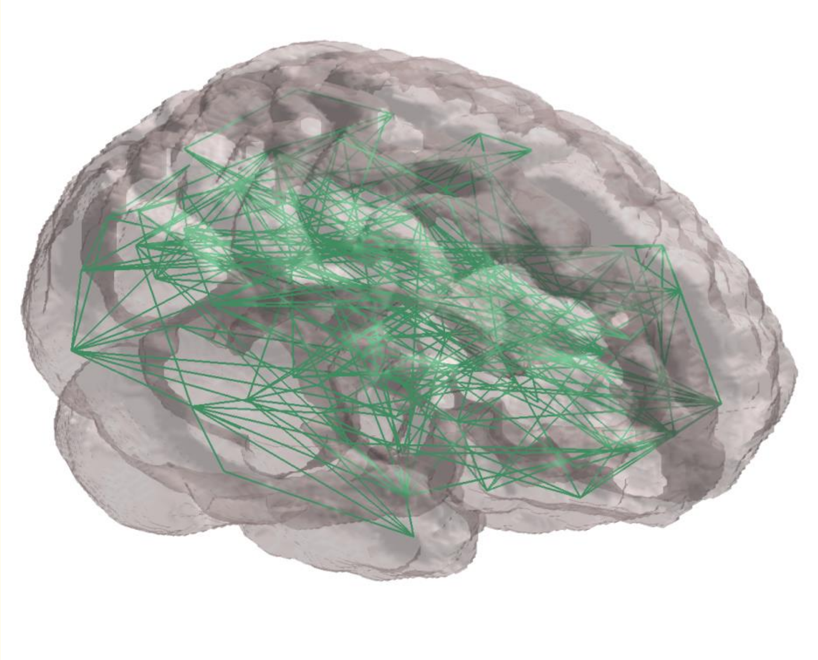
- POC: ERGMs fitted to networks from 10 normal subjects (Simpson et al., 2011)
  - Several R packages available: ergm, ergm.count, GERGM, Bergm, btergm, tergm xergm, xergm.common, blkergm, hergm.

III. Multivariate Modeling and Inference: ERGMs

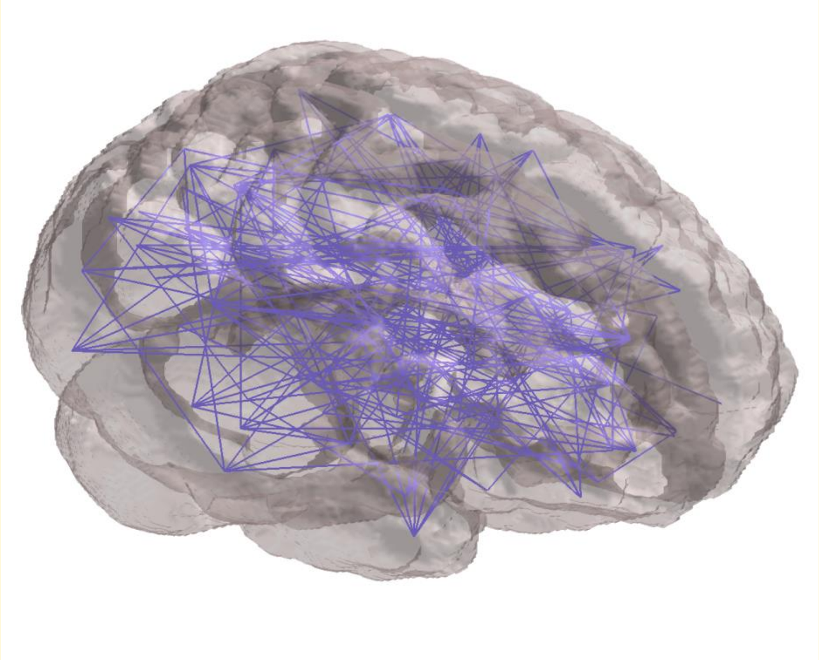
Final ERGMs (composed of most informative explanatory metrics) for each subject provided a good fit to the data as evidenced by graphical GOF plots.



III. Multivariate Modeling and Inference: ERGMs



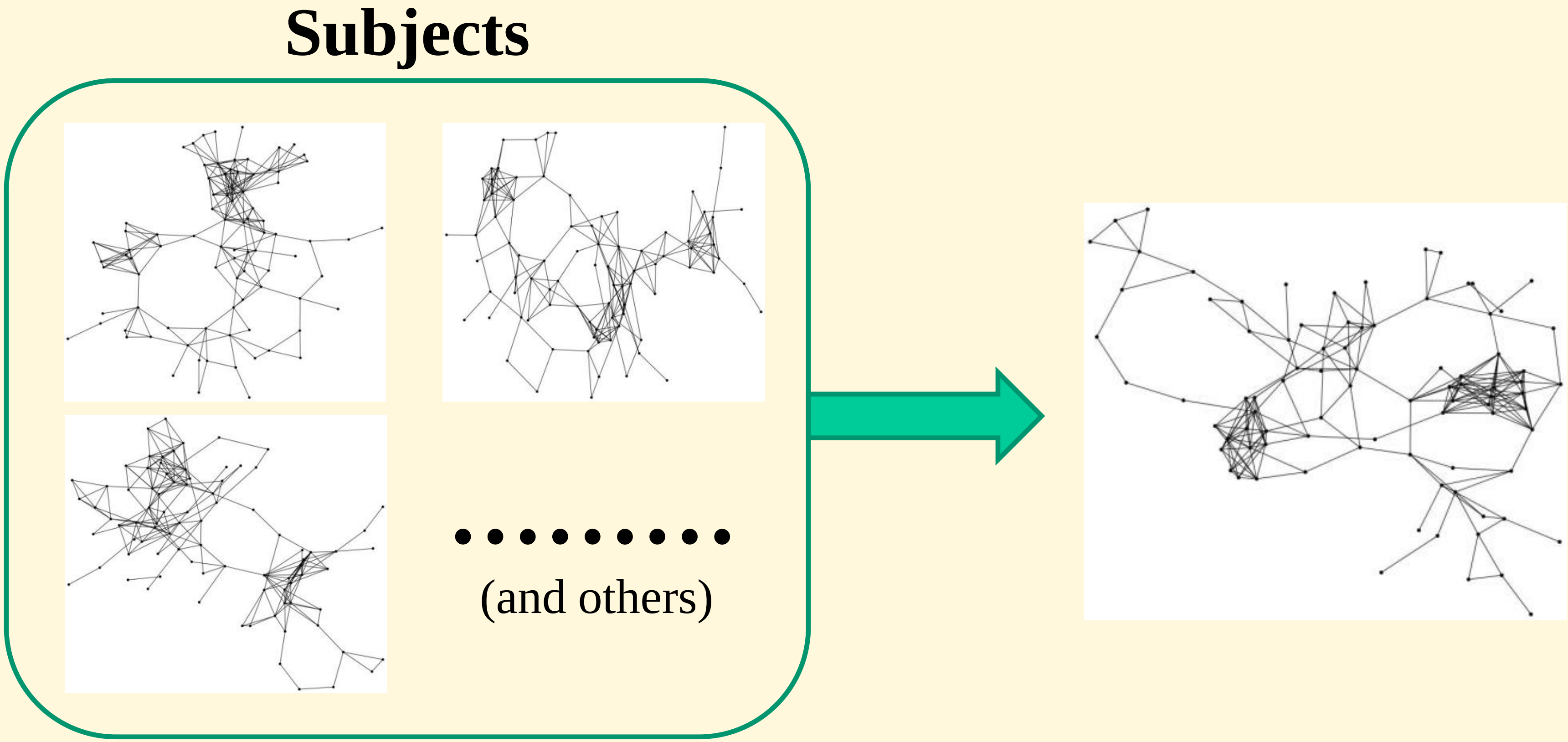
Observed Network



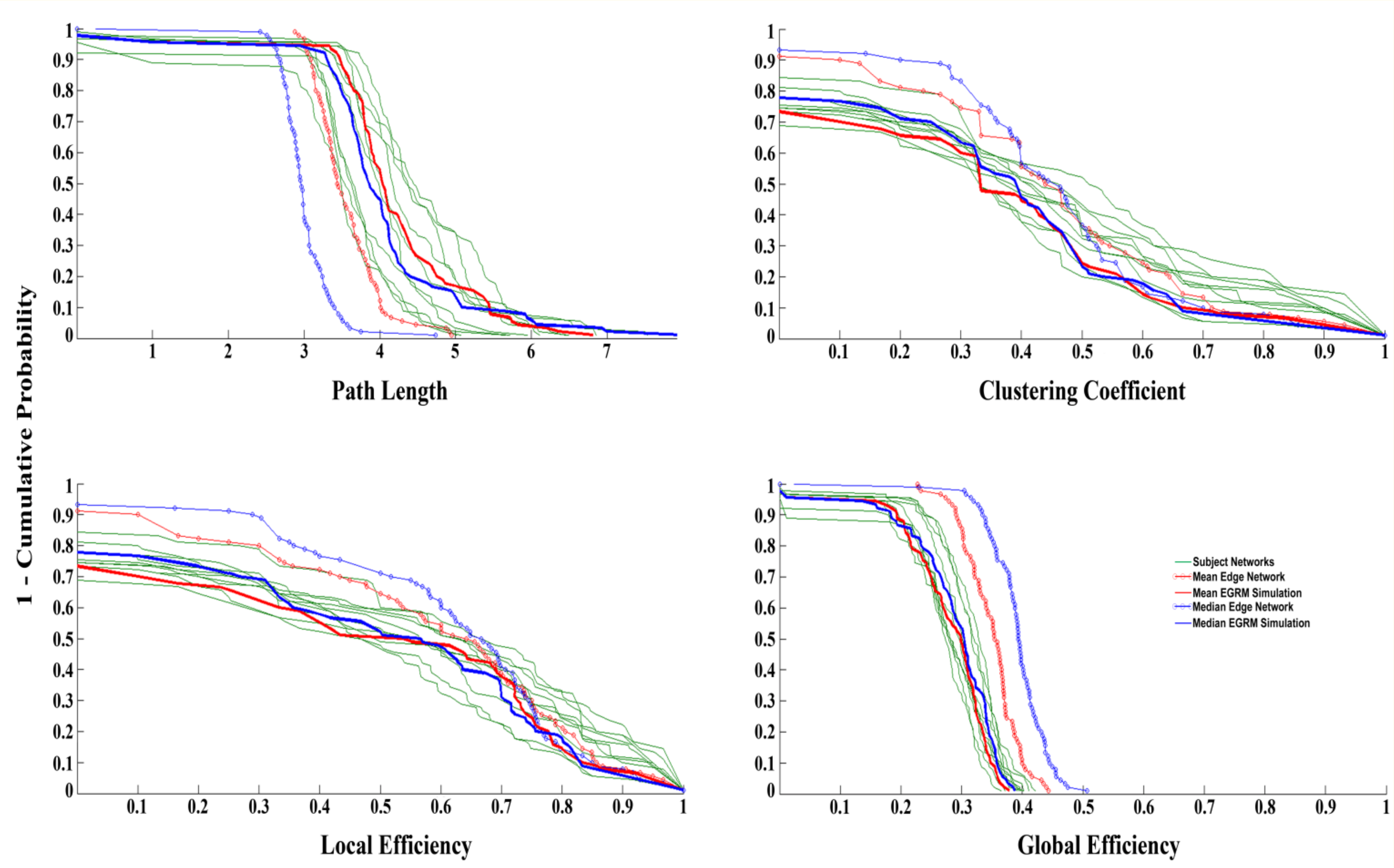
Simulated Network

III. Multivariate Modeling and Inference: ERGMs

- Create group "representative" networks via simulation (Simpson et al., 2012).
  - Traditional mean/median networks are edge-based and topologically differ greatly.



III. Multivariate Modeling and Inference: ERGMs



III. Multivariate Modeling and Inference: ERGMs

Advantages

- Statistically principled approach to topologically modeling, analyzing and simulating complex brain networks.

III. Multivariate Modeling and Inference: ERGMs

Advantages

- Statistically principled approach to topologically modeling, analyzing and simulating complex brain networks.

- Greatest appeal lies in ability to efficiently represent complex network data and allow examining way in which a network's global structure and function depend on its local structure.

III. Multivariate Modeling and Inference: ERGMs

Limitations

- Not well-suited for local examinations.

III. Multivariate Modeling and Inference: ERGMs

Limitations

- Not well-suited for local examinations.

- Multiple-subject comparisons can pose problems.
  - Each subject fitted individually.

III. Multivariate Modeling and Inference: ERGMs

Limitations

- Not well-suited for local examinations.

- Multiple-subject comparisons can pose problems.
  - Each subject fitted individually.

- Difficulty in incorporating novel metrics (more rooted in biology).
  - Due to degeneracy issues that may arise.

III. Multivariate Modeling and Inference: ERGMs

Limitations

- Not well-suited for local examinations.

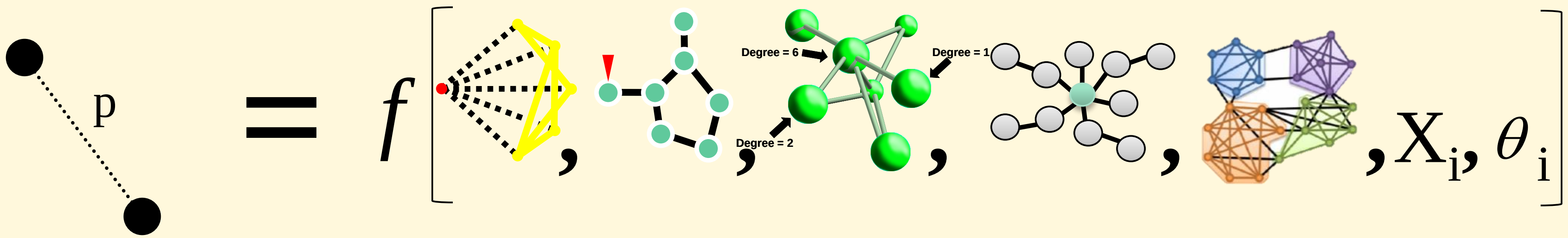
- Multiple-subject comparisons can pose problems.
  - Each subject fitted individually.

- Difficulty in incorporating novel metrics (more rooted in biology).
  - Due to degeneracy issues that may arise.

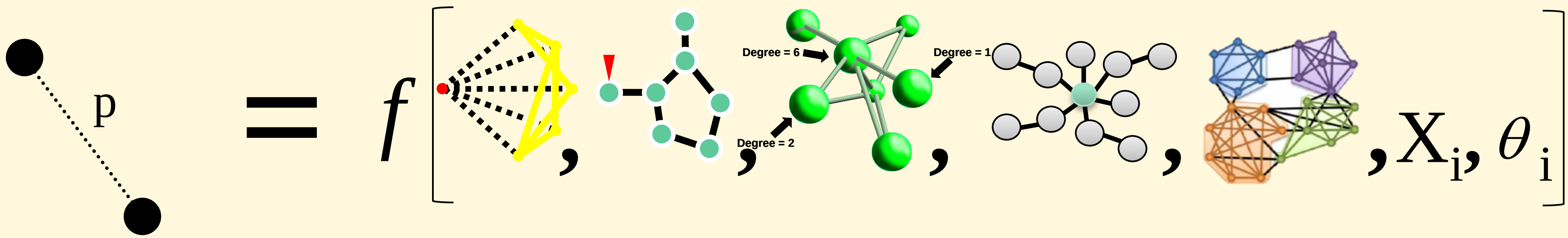
- Developed for static binary networks.
  - Development for longitudinal and weighted networks in infancy.

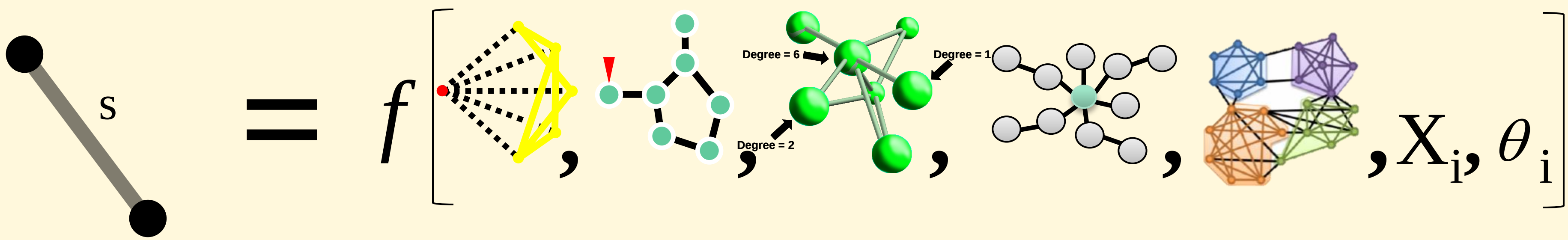
IV. Multivariate Modeling and Inference: Mixed Models

IV. Multivariate Modeling and Inference: Mixed Models



IV. Multivariate Modeling and Inference: Mixed Models





Simpson and Laurienti (2015)

IV. Multivariate Modeling and Inference: Mixed Models

Presence:

$$logit(p_{ijk}) = \boldsymbol{X'}_{ijk,1}\boldsymbol{\beta}_{Net} + \boldsymbol{X'}_{ijk,2}\boldsymbol{\beta}_{COI,Con,Int} + \theta_{ijk}$$

Strength:

$$FZT(S_{ijk}) = \boldsymbol{X'}_{ijk,1}\boldsymbol{\beta}_{Net} + \boldsymbol{X'}_{ijk,2}\boldsymbol{\beta}_{COI,Con,Int} + \theta_{ijk}$$

IV. Multivariate Modeling and Inference: Mixed Models

$$\boldsymbol{\theta}_{pi} = \mathbf{Z}'_{ijk}\boldsymbol{b}_{pi} = \mathbf{Z}'_{ijk}[\, b_{pi,0} \;\; \boldsymbol{b}_{pi,net} \; \boldsymbol{b}_{pi,dist} \, \boldsymbol{\delta}_{pi,j} \; \boldsymbol{\delta}_{pi,k}]'$$

$$\boldsymbol{\theta}_{si} = \mathbf{Z}'_{ijk}\boldsymbol{b}_{si} + e_{ijk} = \mathbf{Z}'_{ijk}[\, b_{si,0} \;\; \boldsymbol{b}_{si,net} \; \boldsymbol{b}_{si,dist} \, \boldsymbol{\delta}_{si,j} \; \boldsymbol{\delta}_{si,k}]' + e_{ijk}$$

$b_{i,0}$	deviation of subject-specific intercepts (from population)
$\boldsymbol{b}_{i,net}$	deviation of subject-specific metric-edge relationships
$\boldsymbol{b}_{i,dist}$	deviation of subject-specific spatial distance-edge relationships
$\boldsymbol{\delta}_{i,j/k}$	propensity for node j/k (of given dyad) to be connected and magnitude of its connections

IV. Multivariate Modeling and Inference: Mixed Models

1) Explain: quantifies relationship between Net/COI/Con and probability/strength of connections.

IV. Multivariate Modeling and Inference: Mixed Models

1) Explain: quantifies relationship between Net/COI/Con and probability/strength of connections.

2) Compare: statistically compares connectivity, network structure, and edge properties by COI (e.g., between groups).

IV. Multivariate Modeling and Inference: Mixed Models

1) Explain: quantifies relationship between Net/COI/Con and probability/strength of connections.

2) Compare: statistically compares connectivity, network structure, and edge properties by COI (e.g., between groups).

3) Predict: predicts connectivity and topology based on participant characteristics, and network structure and its variability via simulations.

IV. Multivariate Modeling and Inference: Mixed Models

1) **Explain:** quantifies relationship between Net/COI/Con and probability/strength of connections.

2) **Compare:** statistically compares connectivity, network structure, and edge properties by COI (e.g., between groups).

3) **Predict:** predicts connectivity and topology based on participant characteristics, and network structure and its variability via simulations.

4) **Threshold:** leverages group-level data to better distinguish between "true" weak connections and noise in individual-level networks.

IV. Multivariate Modeling and Inference: Mixed Models

1) Explain: quantifies relationship between Net/COI/Con and probability/strength of connections.

2) Compare: statistically compares connectivity, network structure, and edge properties by COI (e.g., between groups).

3) Predict: predicts connectivity and topology based on participant characteristics, and network structure and its variability via simulations.

4) Threshold: leverages group-level data to better distinguish between "true" weak connections and noise in individual-level networks.

5) Simulate: simulates group- and individual-level networks useful for model GOF assessments, representative network creation, and network variability assessment.

IV. Multivariate Modeling and Inference: Mixed Models

- **Aging Brain:** assess neurological underpinnings of cognitive decline by examining effects of aging on integration of sensory information.

- Young Adults: 27 $\pm$ 5.8 y/o (n=20) Older Adults: 73 $\pm$ 6.6 y/o (n=19)

- Three separate conditions of fMRI scans:
  - Rest
  - Visual (viewing of a silent movie)
  - Multisensory (MS) (visual and auditory – movie with sound)

- 90 node AAL atlas based networks constructed for each participant.

IV. Multivariate Modeling and Inference: Mixed Models

Here,

$$\boldsymbol{\beta}_{Net} = [\,\beta_{C_{\text{avg}}}\ \ \beta_{Eglob_{\text{avg}}}\ \ \beta_{K_{\text{diff}}}\ \ \beta_{LC_{\text{avg}}}\ \ \beta_{Q}]'.$$

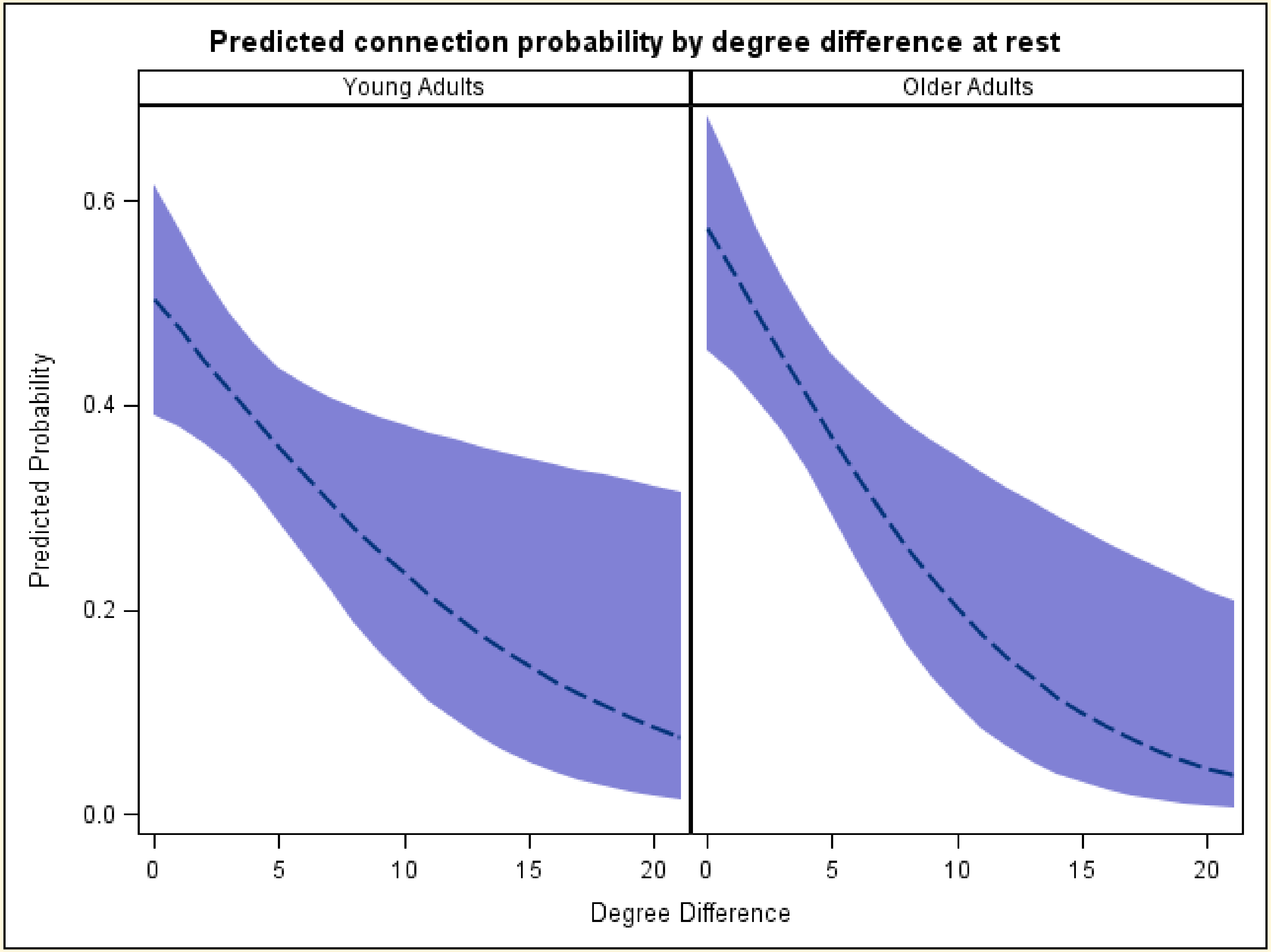
$$\beta_{COI} = \beta_{age}.$$

$$\boldsymbol{\beta}_{Con} = [\,\beta_{sex}\ \ \beta_{educ}\ \ \beta_{dist}\ \ \beta_{dist^2}]'.$$

$$\boldsymbol{\beta}_{Int} = [\,\beta_{age\times C}\ \ \beta_{age\times Eglob}\ \ \beta_{age\times K}\ \ \beta_{age\times LC}\ \ \beta_{age\times Q}\ \ \beta_{age\times sex}]'.$$

IV. Multivariate Modeling and Inference: Mixed Models

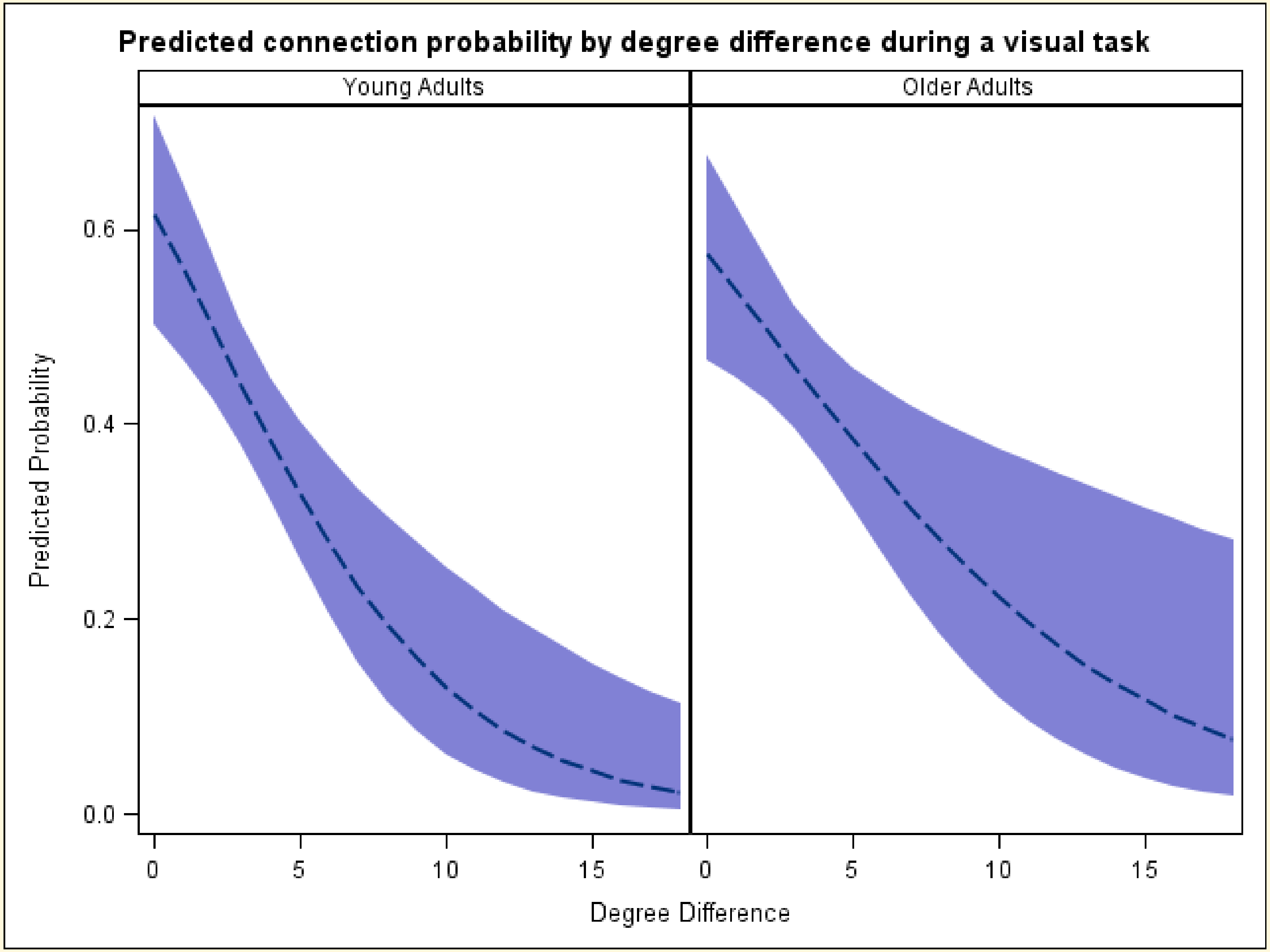
Predict:



Prediction intervals for connection probability as a function of degree difference in young and older participants at rest.

IV. Multivariate Modeling and Inference: Mixed Models

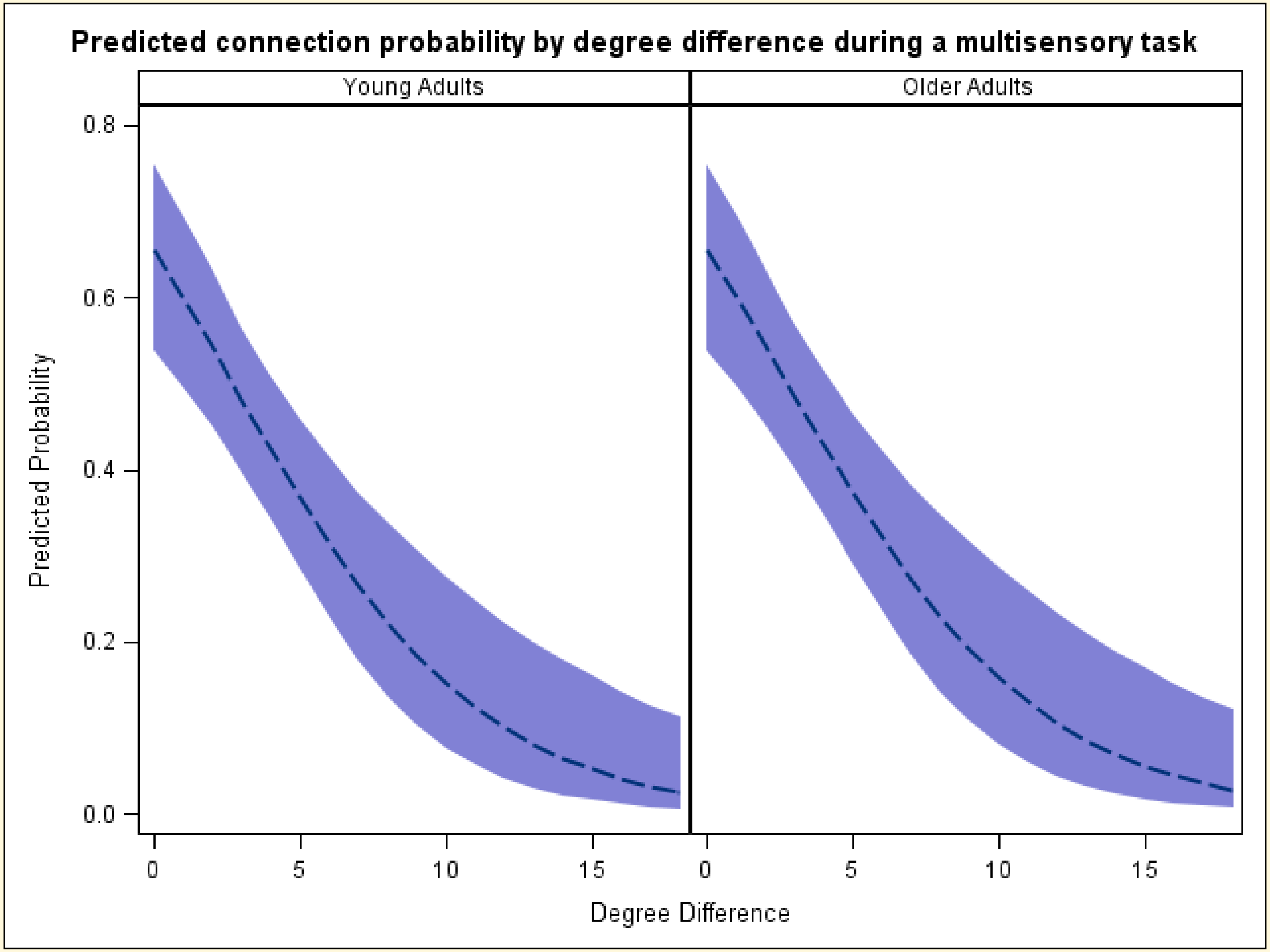
Predict:



Prediction intervals for connection probability as a function of degree difference in young and older participants during a visual task.

IV. Multivariate Modeling and Inference: Mixed Models

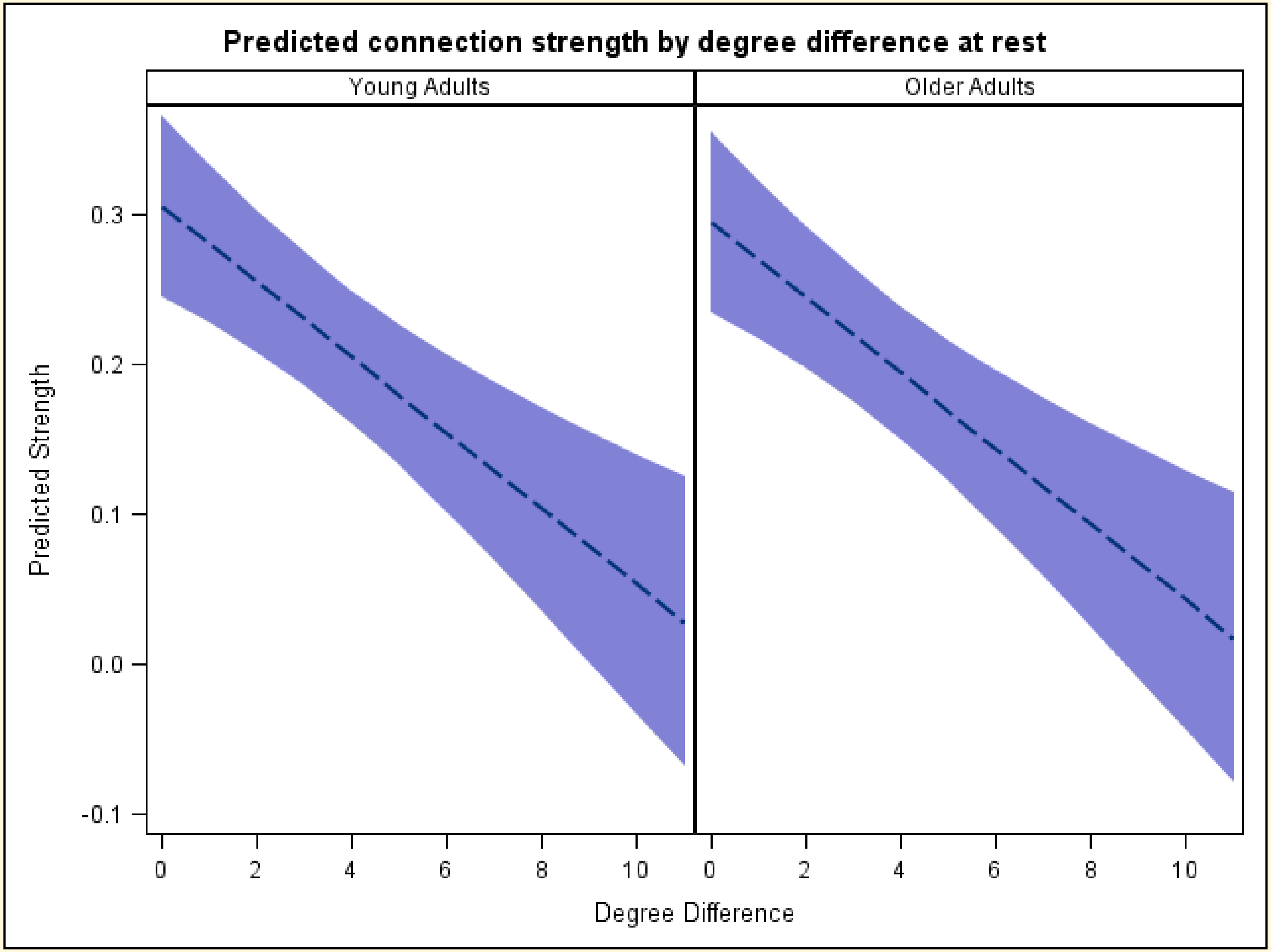
Predict:



Prediction intervals for connection probability as a function of degree difference in young and older participants during a multisensory task.

IV. Multivariate Modeling and Inference: Mixed Models

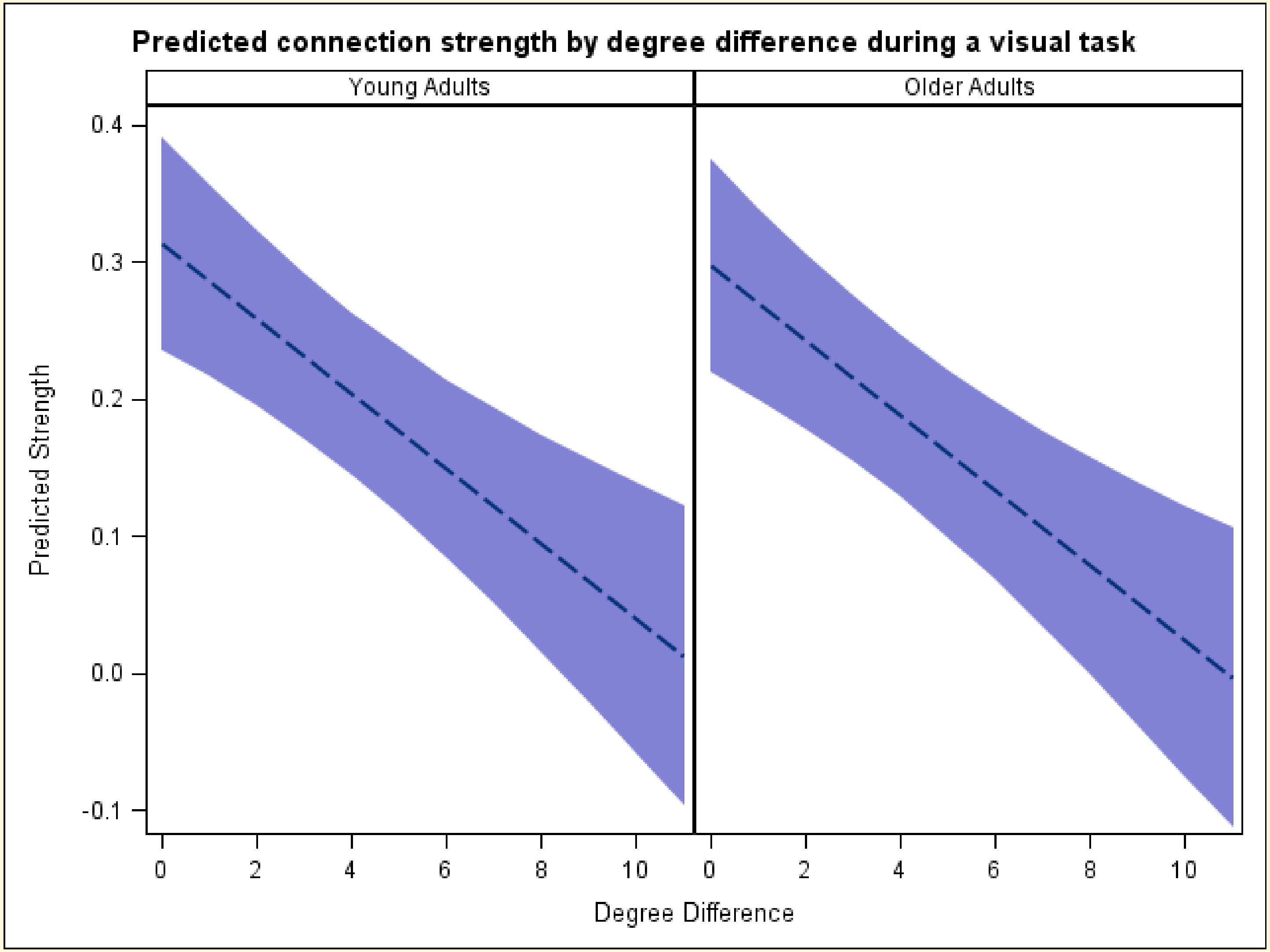
Predict:



Prediction intervals for connection strength as a function of degree difference in young and older participants at rest.

IV. Multivariate Modeling and Inference: Mixed Models

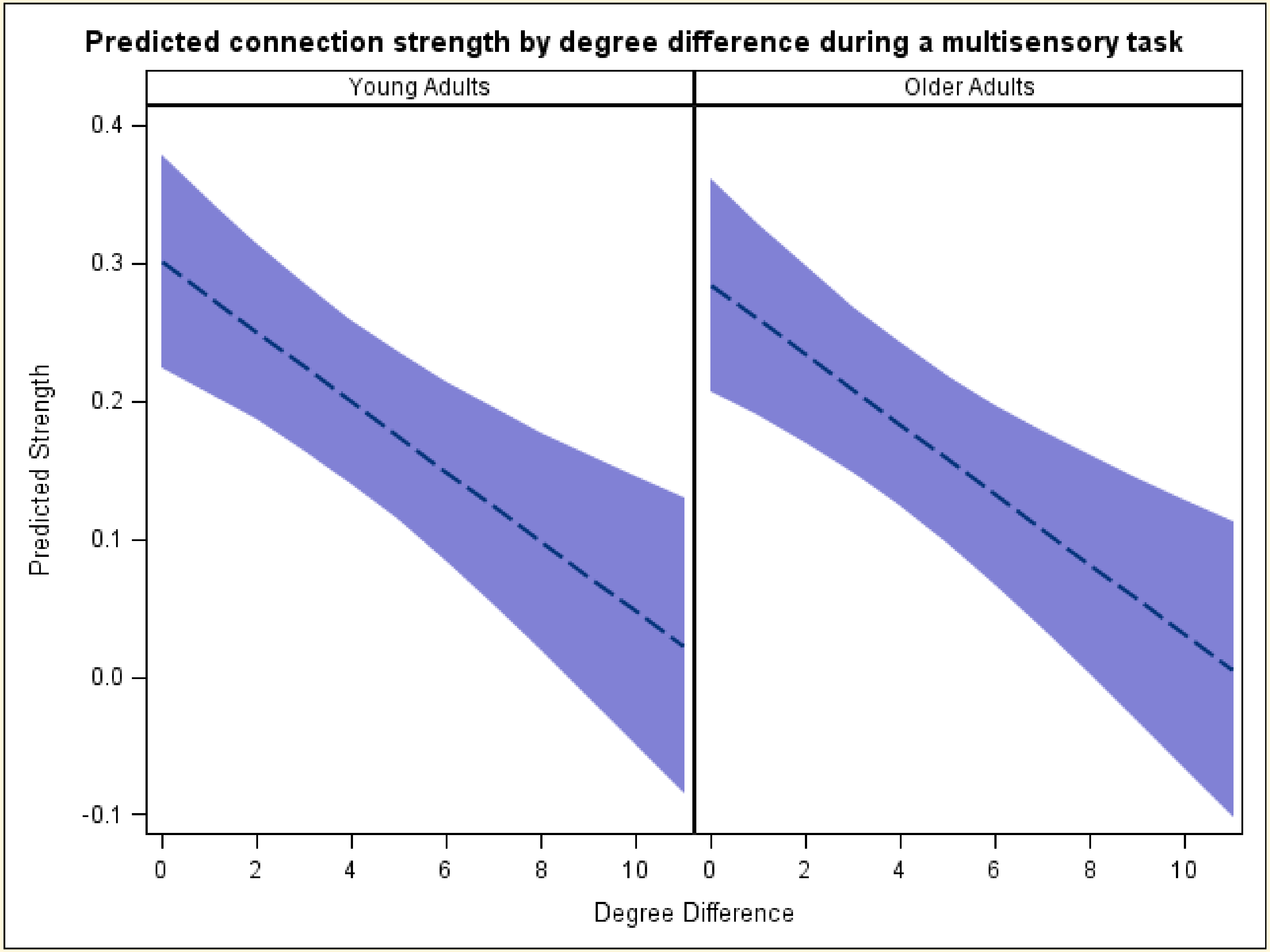
Predict:



Prediction intervals for connection strength as a function of degree difference in young and older participants during a visual task.

IV. Multivariate Modeling and Inference: Mixed Models

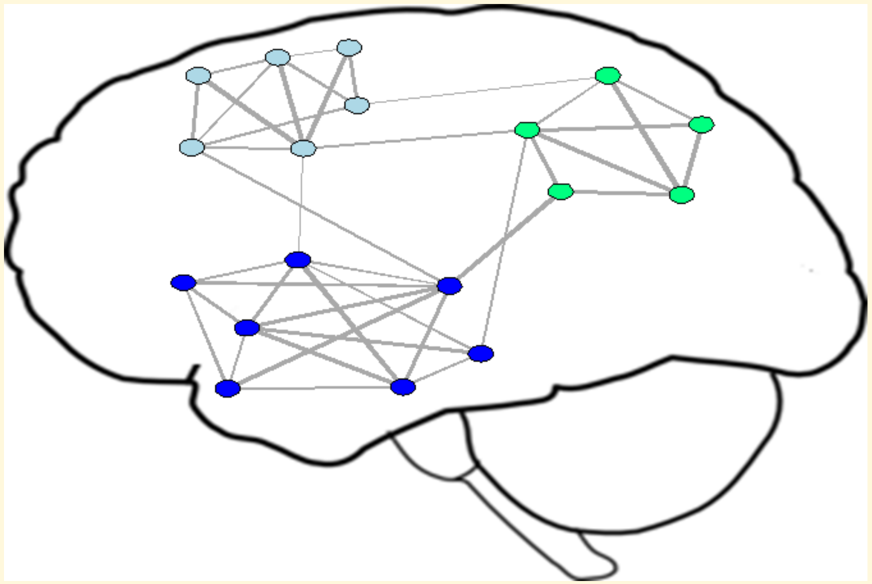
Predict:



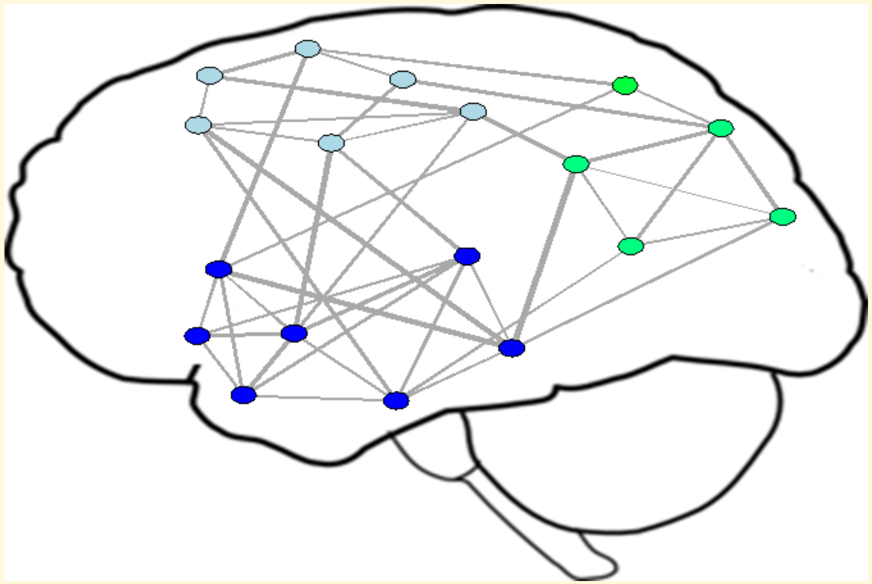
Prediction intervals for connection strength as a function of degree difference in young and older participants during a multisensory task.

IV. Multivariate Modeling and Inference: Mixed Models

- Another example: Used to examine the impacts of pesticide and nicotine exposures on farmworkers' functional brain networks.



Farmworkers

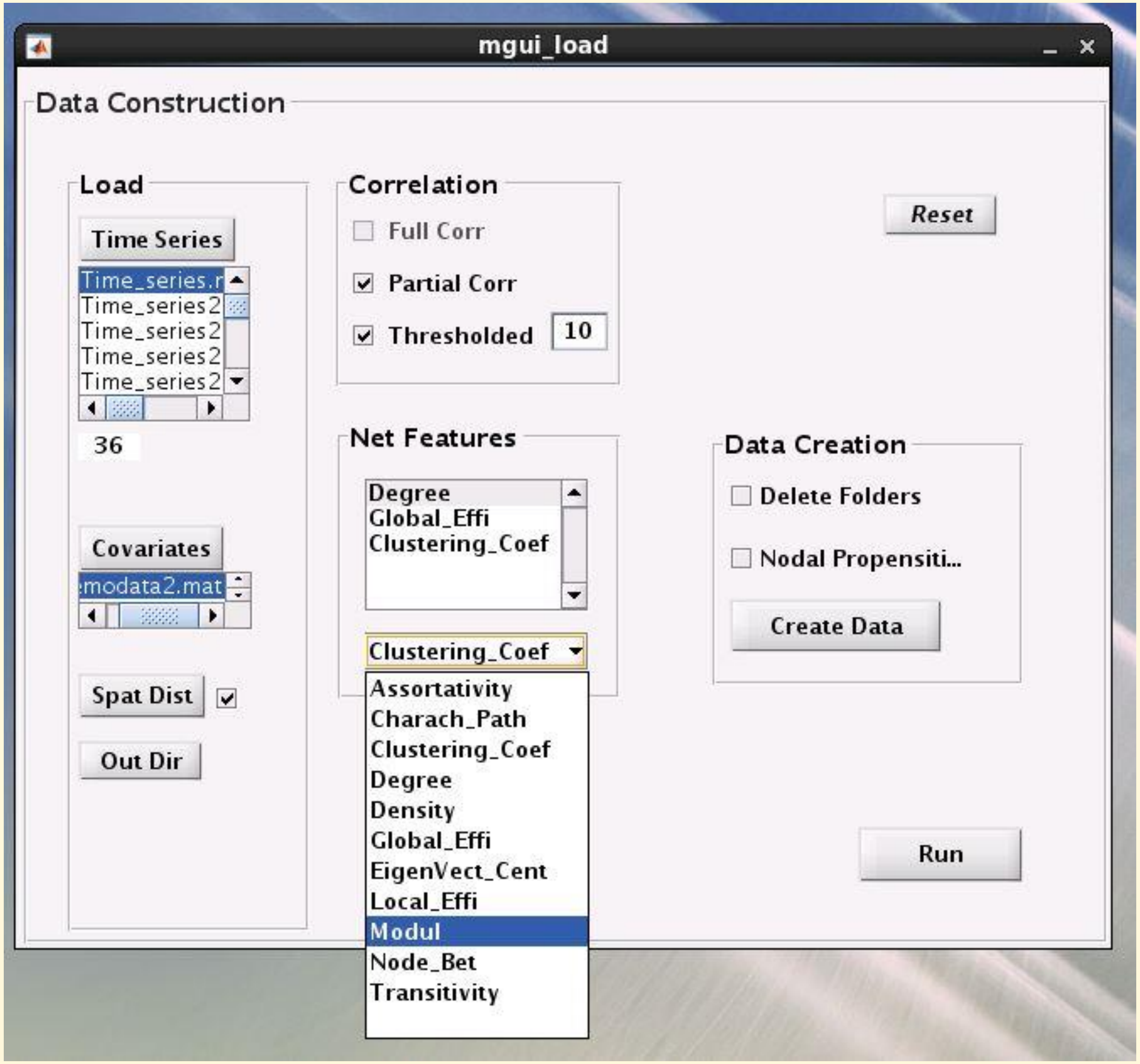


Non-Farmworkers

- FW: More modularly organized with higher functional specificity and lower inter-modular integrity

IV. Multivariate Modeling and Inference: Mixed Models

- Matlab GUI interface coming soon!



V. Summary

ERGMs vs. Mixed Models

- Provide complementary multivariate approaches for analyzing at network level.
    - I.e., assessing systemic infrastructural properties of networks as opposed to properties of specific nodes or connections

V. Summary

ERGMs vs. Mixed Models

- Provide complementary multivariate approaches for analyzing at network level.
  - I.e., assessing systemic infrastructural properties of network as opposed to properties of specific nodes or connections

ERGMs

- Efficiently represent network data by modeling global structure as function of local substructural (network) properties.

- Not well-suited for examining specific connections, comparing groups, or assessing network-phenotype relationships.

V. Summary

ERGMs vs. Mixed Models

Mixed Models

- Well-suited for examining specific connections, group comparisons, and network-phenotype relationship assessment.

- Limited in ability to capture inherent complex dependence structure of networks.
    - Simpson and Laurienti (2015) adapt to brain network context and account for dependence structure.

V. Summary

ERGMs vs. Mixed Models

Mixed Models

- Well-suited for examining specific connections, group comparisons, and network-phenotype relationship assessment.

- Limited in ability to capture inherent complex dependence structure of networks.
  - Simpson & Laurienti (2015) adapt to brain network context and account for dependence structure.

- Rudimentary connectivity/network hybrid method (Simpson & Laurienti, 2016).

- May provide machinery to develop needed advanced hybrid methods.

- Will at least be beneficial in joint network/connectivity analyses in conjunction with an appropriate connectivity method.

ACKNOWLEDGEMENTS

Collaborators

Paul J. Laurienti

F. DuBois Bowman

Satoru Hayasaka

Robert G. Lyday

Malaak N. Moussa

Mohsen Bahrami

Funding

Wake Forest Translational Science Institute

NIBIB K25 EB012236

Others

Members of the Laboratory for Complex Brain Networks

VI. Useful References

ERGMs

Hunter DR, Handcock MS, Butts CT, Goodreau SM, Morris M (2008). ergm: A package to fit, simulate and diagnose exponential-family models for networks. Journal of statistical software 24(3).

Simpson SL, Bowman FD, Laurienti PJ (2013). Analyzing complex functional brain networks: fusing statistics and network science to understand the brain. Statistics Surveys 7: 1-36.

Simpson SL, Hayasaka S, Laurienti PJ (2011) Exponential random graph modeling for complex brain networks. PLoS ONE 6: e20039.

Simpson SL, Moussa MN, Laurienti PJ (2012). An exponential random graph modeling approach to creating group-based representative whole-brain connectivity networks. NeuroImage 60: 1117-1126.

Sinke, MR, Dijkhuizen RM, Caimo A, Stam CJ, Otte WM (2016). Bayesian exponential random graph modeling of whole-brain structural networks across lifespan. NeuroImage 135: 79-91.

Mixed Models

Simpson SL, Laurienti PJ (2015). A two part mixed-effects modeling framework for analyzing whole-brain network data. NeuroImage 113: 310-319.

Simpson SL, Laurienti PJ (2016). Disentangling brain graphs: a note on the conflation of network and connectivity analyses. Brain Connectivity 6(2): 95-98.

Bahrami M, Laurienti PJ, Quandt SA, Talton J, Pope CN, Summers P, Burdette JH, Chen H, Liu J, Howard TD, Arcury TA, Simpson SL (2017). The impacts of pesticide and nicotine exposures on latino farmworkers' functional brain networks. Neurotoxicology, In Press.