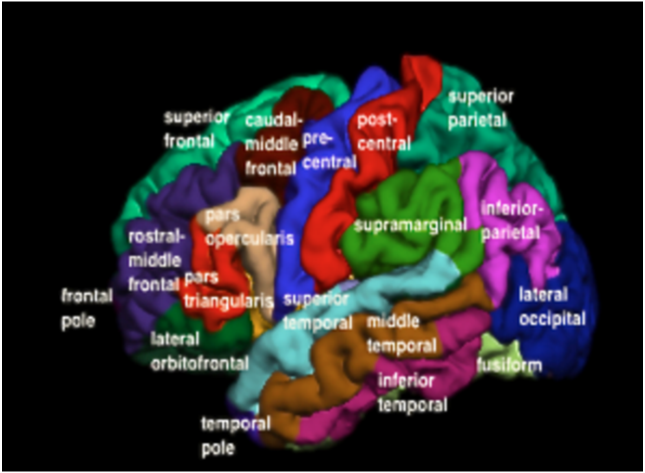
The adaptive incorporation of multiple sources of information in Brain Imaging via penalized optimization

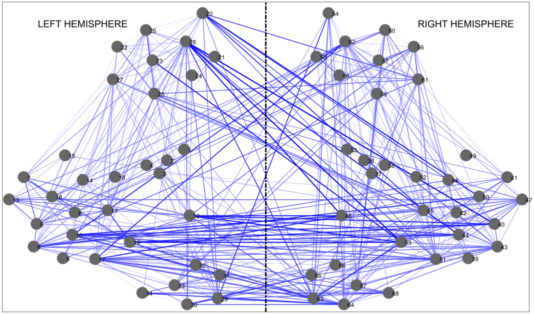
Damian Brzyski

Wrocław University of Science and Technology,
Wrocław, Poland

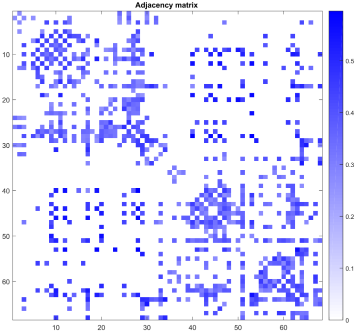
5 June 2020

Parcelation of the cortex



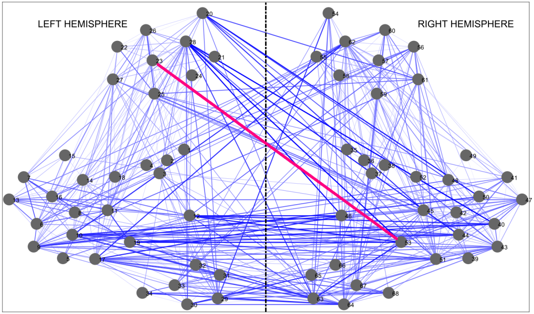


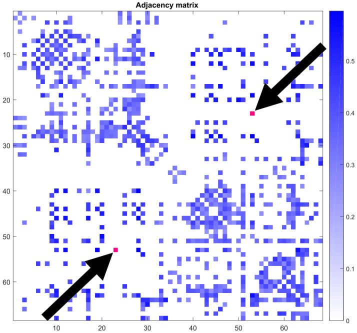
LEFT HEMISPHERE
RIGHT HEMISPHERE



Adjacency matrix

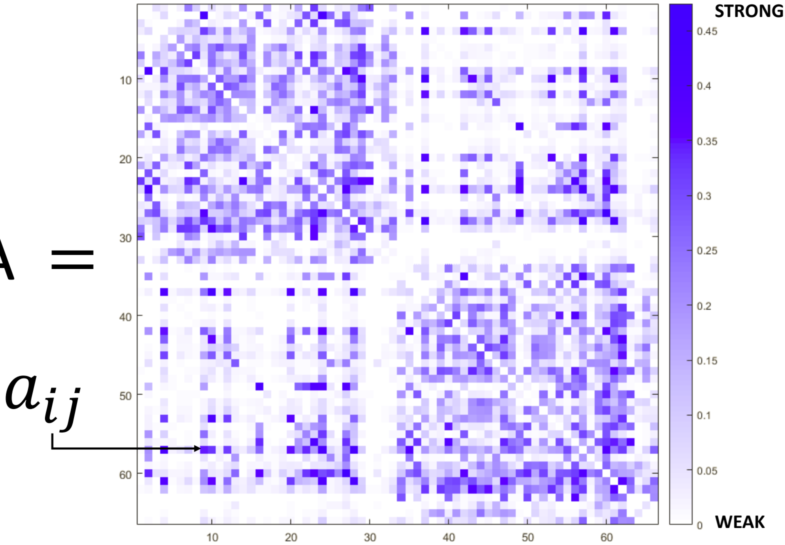
Connectivities in the brain





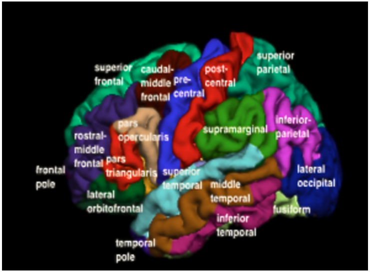
Connectivity information for population

$$A = $$

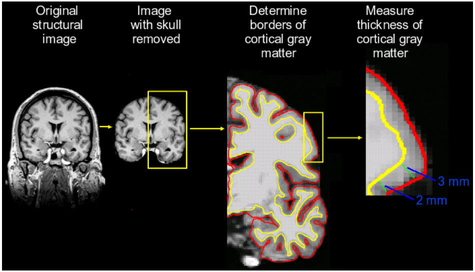


MRI-derived data: cortical thickness

1. We model the association between given response variable and average cortical thickness
2. We consider the parcellation of the brain into 66 regions

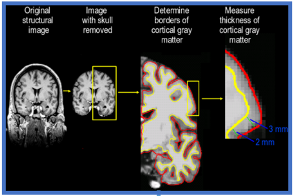


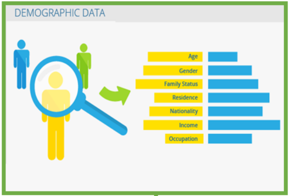
(a) Parcellation of the brain



(b) Cortical thickness

Statistical model

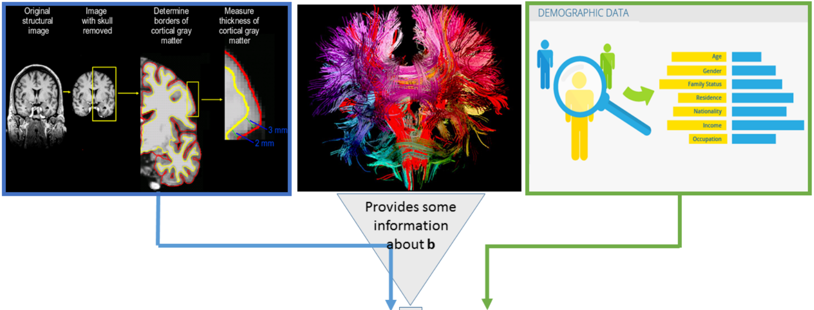




$$y = Zb + X\beta + \varepsilon$$

1. y is n-dimensional vector of considered responses
2. $Z \in \mathbb{R}^{n\times 66}$ and $X \in \mathbb{R}^{n\times m}$
3. $\varepsilon \sim \mathcal{N}\left(0, \sigma^2 I_n\right)$ for some unknown $\sigma^2 > 0$

Statistical model



$$y = Z\boxed{b} + X\beta + \varepsilon$$

1. y is n-dimensional vector of considered responses
2. $Z \in \mathbb{R}^{n \times 66}$ and $X \in \mathbb{R}^{n \times m}$
3. $\varepsilon \sim \mathcal{N}\left(0, \sigma^2 I_n\right)$ for some unknown $\sigma^2 > 0$

Penalized estimation

To find estimates of b and β we consider the optimization problem of the form

$$\operatorname*{argmin}_{b,\beta} \Big\{ \underbrace{\|y - Zb - X\beta\|_2^2}_{\textbf{model fit term}} + \lambda \underbrace{g(b)}_{\textbf{penalty on } b} \Big\}.$$

Penalized estimation

To find estimates of b and β we consider the optimization problem of the form

$$\underset{b,\beta}{\operatorname{argmin}} \left\{ \underbrace{\|y - Zb - X\beta\|_2^2}_{\textbf{model fit term}} + \lambda \underbrace{g(b)}_{\textbf{penalty on } b} \right\}.$$

1. How to select the penalty term g?
2. How to select an optimal regularization parameter λ ?

Penalty selection

We want to get the property

"The stronger brain's regions i and j are connected based on the connectivity matrix entry a_{ij}, the closer coefficients $\hat{b}_i$ and $\hat{b}_j$ are to each other."

Penalty selection

We want to get the property

"The stronger brain's regions i and j are connected based on the connectivity matrix entry a_{ij}, the closer coefficients $\hat{b}_i$ and $\hat{b}_j$ are to each other."

1. The natural idea then is to consider the penalty of the form

$$g(b) = \sum_{i,j} a_{ij}\big(b_i - b_j\big)^2$$

Penalty selection

We want to get the property

"The stronger brain's regions i and j are connected based on the connectivity matrix entry a_{ij}, the closer coefficients $\hat{b}_i$ and $\hat{b}_j$ are to each other."

1. The natural idea then is to consider the penalty of the form

$$g(b) = \sum_{i,j} a_{ij}\left(b_i - b_j\right)^2$$

2. Note $\sum_{i,j} a_{ij}\left(b_i - b_j\right)^2 = b^T Q b$, where Q is the Laplacian of A defined as $Q := D - A$, for $D := diag\left(\sum_j a_{1j}, \ldots, \sum_j a_{pj}\right)$

Connection with linear mixed models (LMM)

Consequently, we get the following form of the objective function

$$\underset{b,\beta}{\operatorname{argmin}} \ \Big\{ \|y - Zb - X\beta\|_2^2 \ + \ \lambda b^T Q b \Big\},$$

where Q is a Laplacian of A.

Connection with linear mixed models (LMM)

Consequently, we get the following form of the objective function

$$\underset{b,\beta}{\operatorname{argmin}} \left\{ \|y - Zb - X\beta\|_2^2 \ + \ \lambda b^T Q b \right\},$$

where Q is a Laplacian of A.

This problem is "equivalent" with LMM formulation

1. $y = Zb + X\beta + \varepsilon$, where β is a vector of fixed effects and b a vector of random effects,
2. $\varepsilon \sim \mathcal{N}\left(0, \sigma^2 I\right)$,
3. $b \sim \mathcal{N}\left(0, \sigma_b^2 Q^{-1}\right)$,
4. λ, σ and σ_b are tied via $\lambda = \sigma^2/\sigma_b^2$.

Selection of regularization parameter

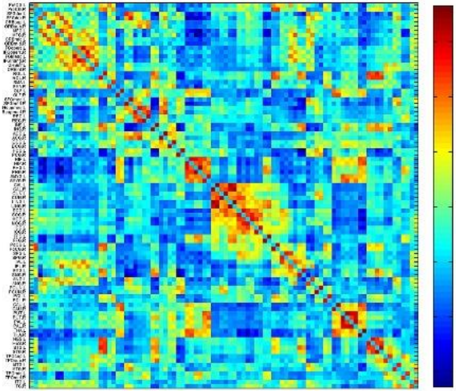
IDEA: define $\hat{\lambda}$ as $\hat{\lambda} = \frac{\hat{\sigma}^2}{\hat{\sigma}_b^2}$ where $\hat{\sigma}^2$ and $\hat{\sigma}_b^2$ are maximum likelihood estimates from the corresponding linear mixed model

Selection of regularization parameter

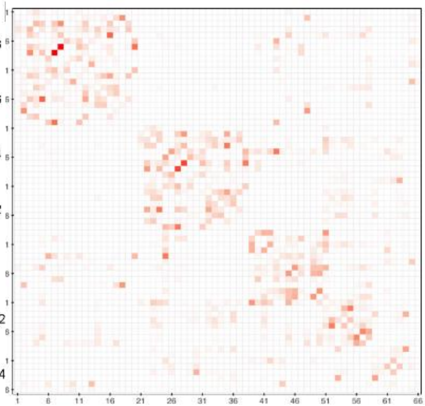
IDEA: define $\hat{\lambda}$ as $\hat{\lambda} = \frac{\hat{\sigma}^2}{\hat{\sigma}_b^2}$ where $\hat{\sigma}^2$ and $\hat{\sigma}_b^2$ are maximum likelihood estimates from the corresponding linear mixed model

PROBLEM: Neither Laplacian nor normalized Laplacian is an invertible matrix, which is required in computation

Connectivity information types

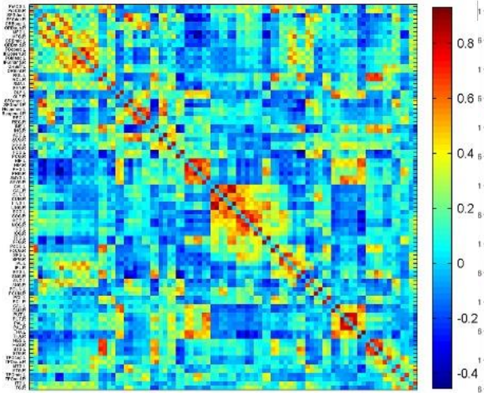


Functional Connectivity

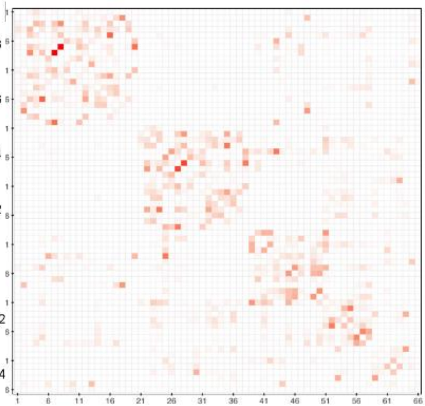


Structural Connectivity

Connectivity information types



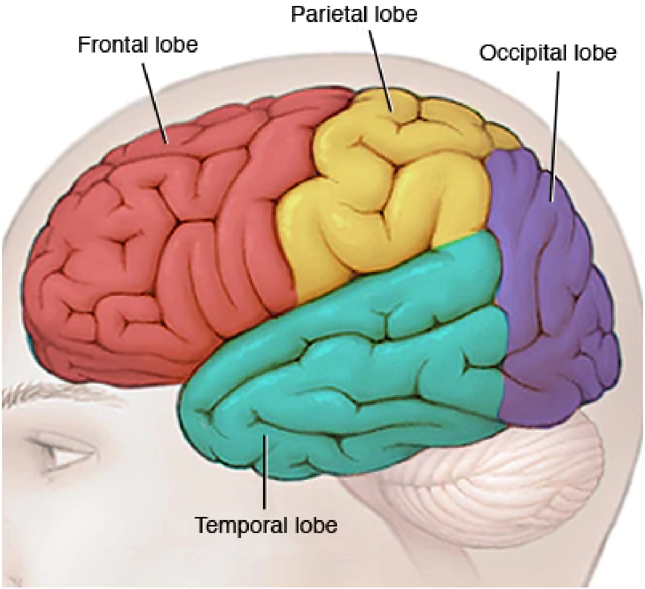
Functional Connectivity



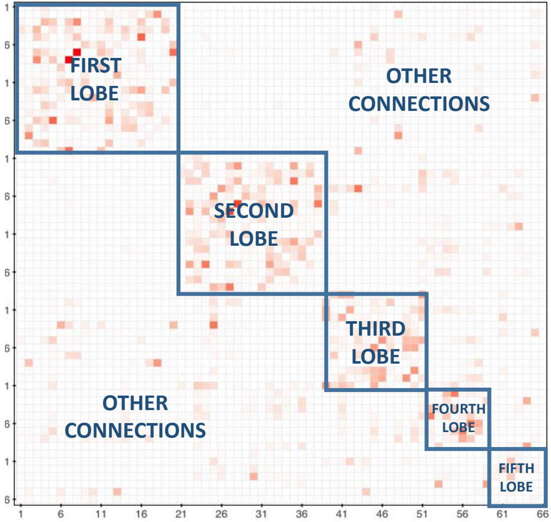
Structural Connectivity

Which connectivity matrix should we use to define Q?

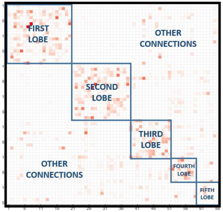
Brain lobes



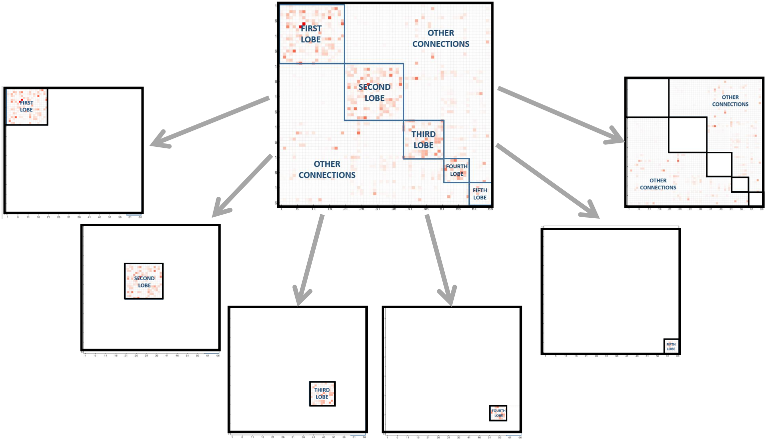
Brain lobes


FIRST LOBE
SECOND LOBE
THIRD LOBE
FOURTH LOBE
FIFTH LOBE
OTHER CONNECTIONS
OTHER CONNECTIONS
1
6
11
16
21
26
31
36
41
46
51
56
61
66

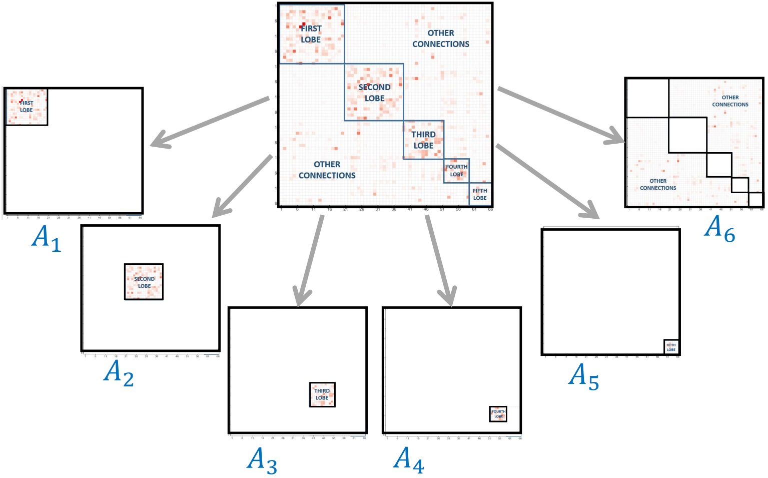
Brain Lobes


FIRST LOBE
OTHER CONNECTIONS
SECOND LOBE
THIRD LOBE
OTHER CONNECTIONS
FOURTH LOBE
FIFTH LOBE

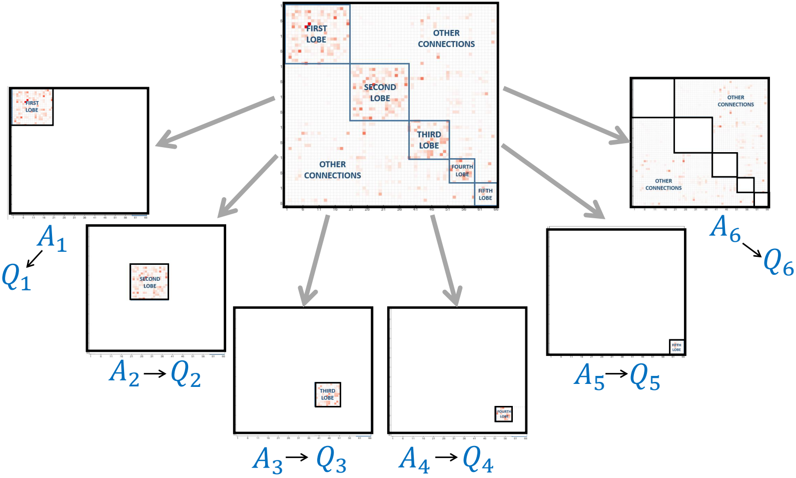
Brain Lobes


FIRST
LOBE
SECOND
LOBE
THIRD
LOBE
FOURTH
LOBE
FIFTH
LOBE
OTHER
CONNECTIONS
OTHER
CONNECTIONS

Brain Lobes


FIRST LOBE
OTHER CONNECTIONS
SECOND LOBE
THIRD LOBE
FOURTH LOBE
FIFTH LOBE
OTHER CONNECTIONS
A_1
A_2
A_3
A_4
A_5
A_6

Brain Lobes


FIRST LOBE
SECOND LOBE
THIRD LOBE
FOURTH LOBE
FIFTH LOBE
OTHER CONNECTIONS
OTHER CONNECTIONS
A_1
Q_1
$A_2 \rightarrow Q_2$
$A_3 \rightarrow Q_3$
$A_4 \rightarrow Q_4$
$A_5 \rightarrow Q_5$
A_6
Q_6

Merging a few sources of information

Suppose that we have m positive semidefinite matrices,

$$Q_1, \ldots Q_m$$

considered as potential candidates for the penalty.

Merging a few sources of information

Suppose that we have m positive semidefinite matrices,

$$Q_1, \ldots Q_m$$

considered as potential candidates for the penalty.

AIMER: (Adaptive Information Merging Estimator for Regression) is the solution to

$$\underset{b,\beta}{\operatorname{argmin}} \left\{ \|y - Zb - X\beta\|_2^2 + \sum_{i=1}^{m} \lambda_i b^T Q_i b \right\},$$

Merging a few sources of information

Suppose that we have m positive semidefinite matrices,

$$Q_1, \dots Q_m$$

considered as potential candidates for the penalty.

AIMER: (Adaptive Information Merging Estimator for Regression) is the solution to

$$\underset{b,\beta}{\operatorname{argmin}} \left\{ \|y - Zb - X\beta\|_2^2 + \sum_{i=1}^{m} \lambda_i b^T Q_i b \right\},$$

where $\lambda_1, \dots \lambda_n$ are tuning parameters.

Connection with linear mixed models

$$\underset{b,\beta}{\operatorname{argmin}} \left\{ \|y - Zb - X\beta\|_2^2 + \sum_{i=1}^{m} \lambda_i b^T Q_i b \right\},$$

Connection with linear mixed models

$$\underset{b,\beta}{\operatorname{argmin}} \left\{ \left\| y - Zb - X\beta \right\|_2^2 + \sum_{i=1}^{m} \lambda_i b^T Q_i b \right\},$$

This problem is "equivalent" with LMM formulation

1. $y = Zb + X\beta + \varepsilon$, where β is a vector of fixed effects and b a vector of random effects,
2. $\varepsilon \sim \mathcal{N}\left(0, \sigma^2 I\right)$,
3. $b \sim \mathcal{N}\left(0, \left[\sum_{i=1}^{m} \omega_i Q_i\right]^{-1}\right)$,
4. λ_i's, ω_i's and σ^2 are tied via $\lambda_i = \omega_i \cdot \sigma^2$.

Connection with linear mixed models

$$\underset{b,\beta}{\operatorname{argmin}} \left\{ \|y - Zb - X\beta\|_2^2 + \sum_{i=1}^{m} \lambda_i b^T Q_i b \right\},$$

This problem is “equivalent” with LMM formulation

1. $y = Zb + X\beta + \varepsilon$, where β is a vector of fixed effects and b a vector of random effects,
2. $\varepsilon \sim \mathcal{N}\big(0, \sigma^2 I\big)$,
3. $b \sim \mathcal{N}\Big(0, \big[\sum_{i=1}^m \omega_i Q_i\big]^{-1}\Big)$,
4. λ_i's, ω_i's and σ^2 are tied via $\lambda_i = \omega_i \cdot \sigma^2$.

Noninvertability problem is also addressed!

The example

- For $m = 2$ define Q_1 as the Laplacian and Q_2 as identity matrix

The example

- For $m = 2$ define Q_1 as the Laplacian and Q_2 as identity matrix
- The penalty becomes $\lambda_1 b^\mathsf{T} Q_1 b + \lambda_2 \|b\|_2^2$

The example

- For $m = 2$ define Q_1 as the Laplacian and Q_2 as identity matrix
- The penalty becomes $\lambda_1 b^\mathsf{T} Q_1 b + \lambda_2 \|b\|_2^2$
- Method reduces to *riPEER*

Simulation scheme

Two methods compared:

1. ridge: $\lambda_Q := 0$ (connectivity information is not used)
2. riPEER (both lambdas are selected in an adaptive way)

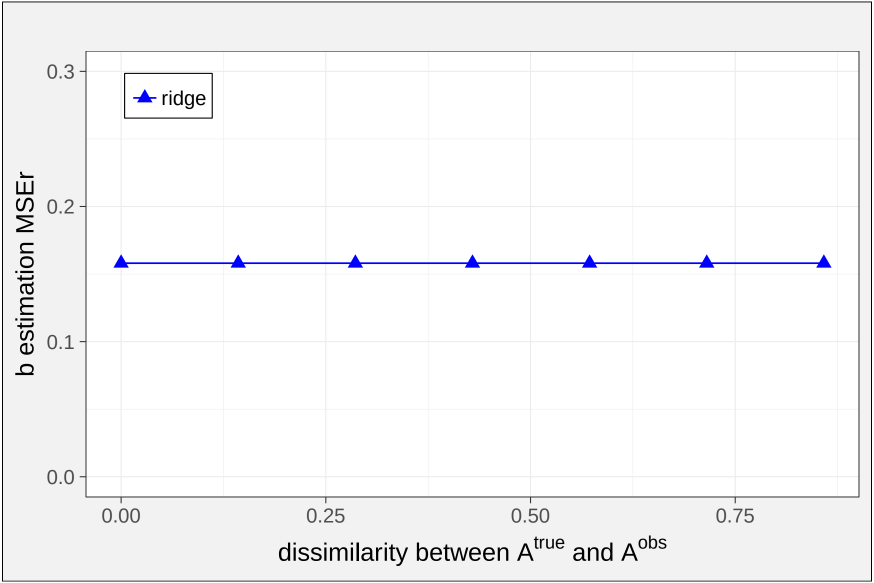
Axis of the plot

1. X axis:

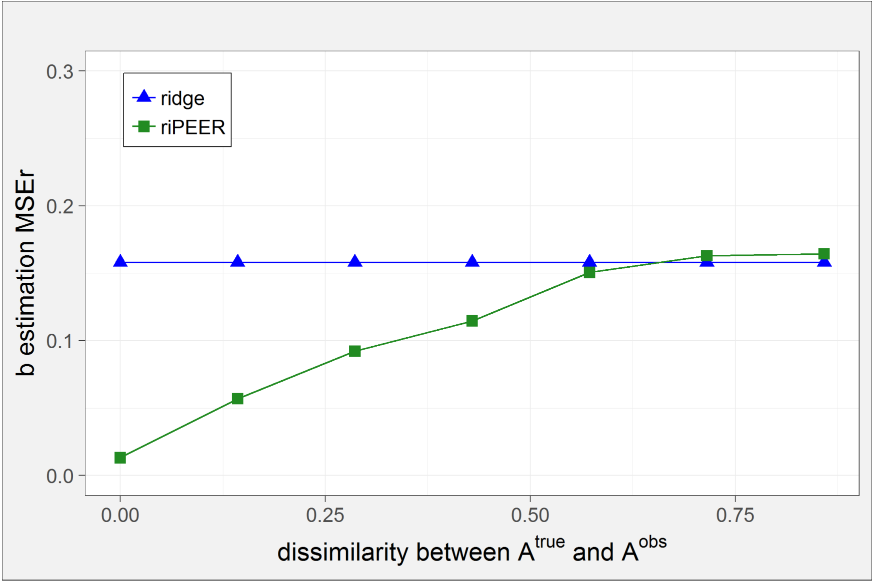
$$diss(A^{true}, A^{obs}) := \frac{\text{number of removed/added connections}}{\text{number of all nonzero connections in } A^{true}}$$

2. Y axis: $MSEr := \mathbb{E}\left[\frac{\|\hat{b}-b^{true}\|_2^2}{\|b^{true}\|_2^2}\right]$.

Simulation results



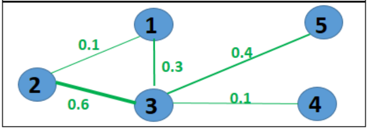
Simulation results



Simulation scheme

SIMULATED SIGNAL

Graph given by adjacency matrix A^{true}



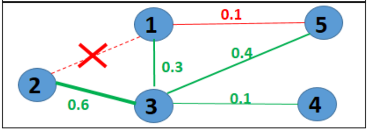
Normalized Laplacian : Q_{true}

„Invertible Normalized Laplacian" :
$\tilde{Q}_{true} := Q_{true} + 0.001 \cdot I$

True signal used in simulation:
$b^{true} \sim N(0,\ \sigma_b^2\, \tilde{Q}_{true}^{-1})$

ESTIMATION

Rewired graph which gives A^{obs}



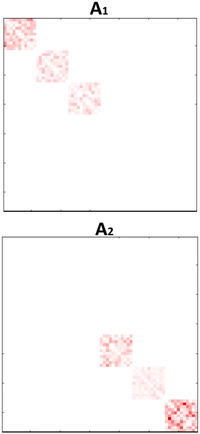
Normalized Laplacian of rewired graph was used to find the estimate, $\hat{b}$

MSEr defined as

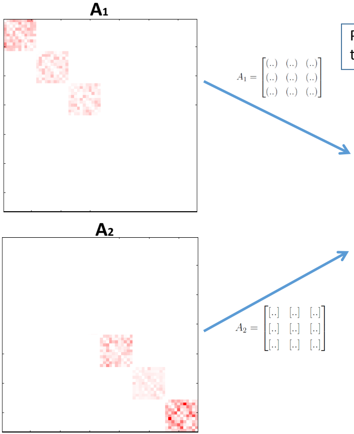
$$\text{MSEr:} = \mathbb{E}\left[\frac{\|\hat{b} - b_{true}\|_2^2}{\|b_{true}\|_2^2}\right]$$

as a measure of estimation accuracy

Simulation scheme: multiple sources

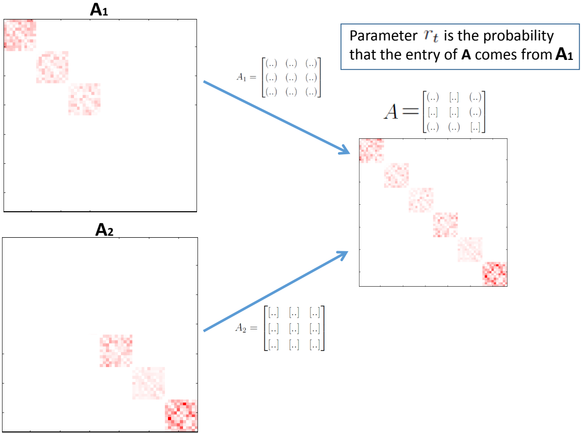


Simulation scheme: multiple sources

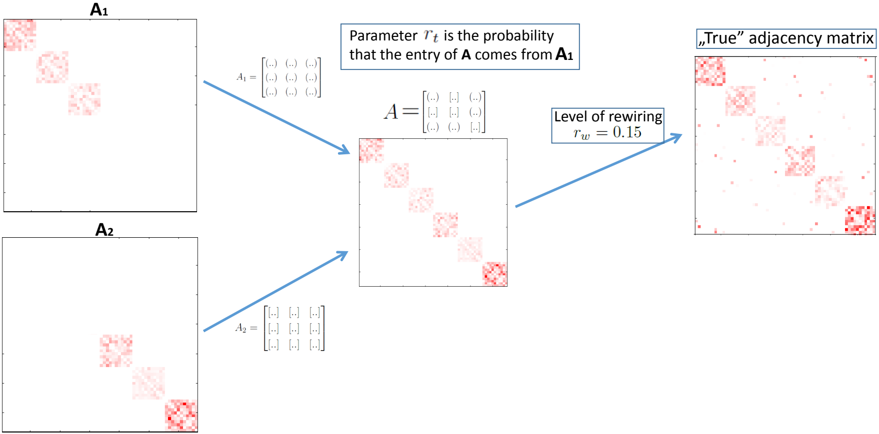


Parameter r_t is the probability that the entry of **A** comes from **A1**

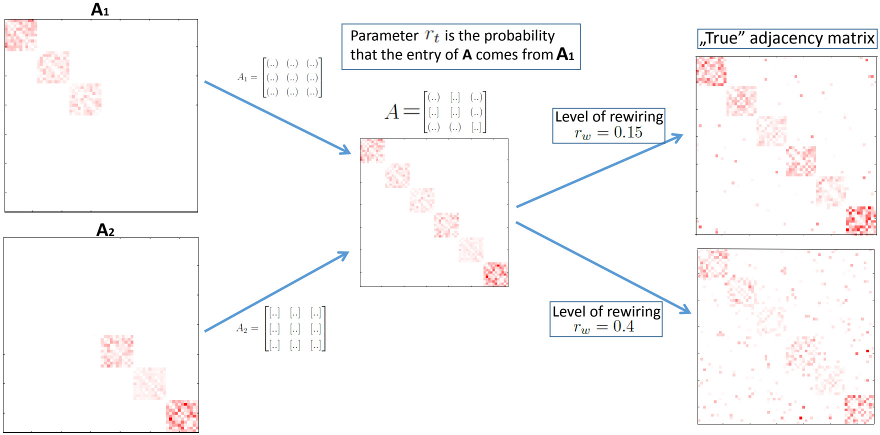
Simulation scheme: multiple sources



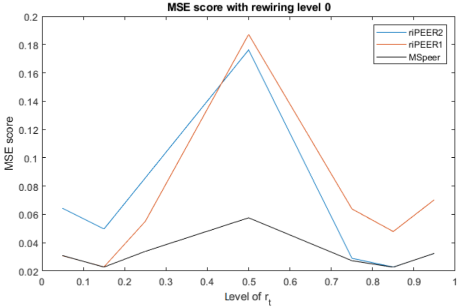
Simulation scheme: multiple sources

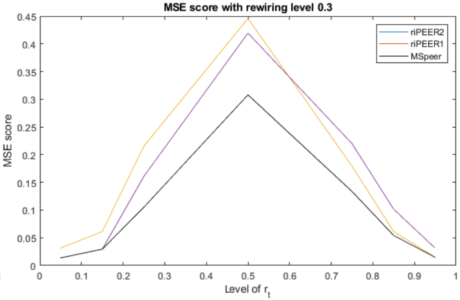


Simulation scheme: multiple sources

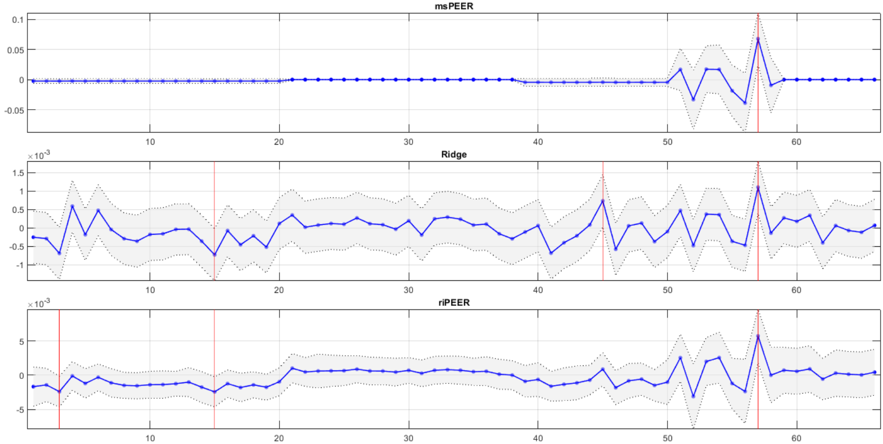


Simulation results

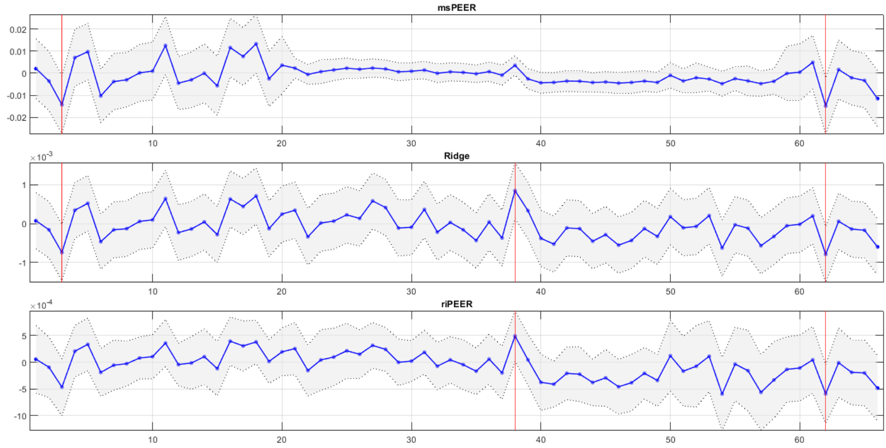




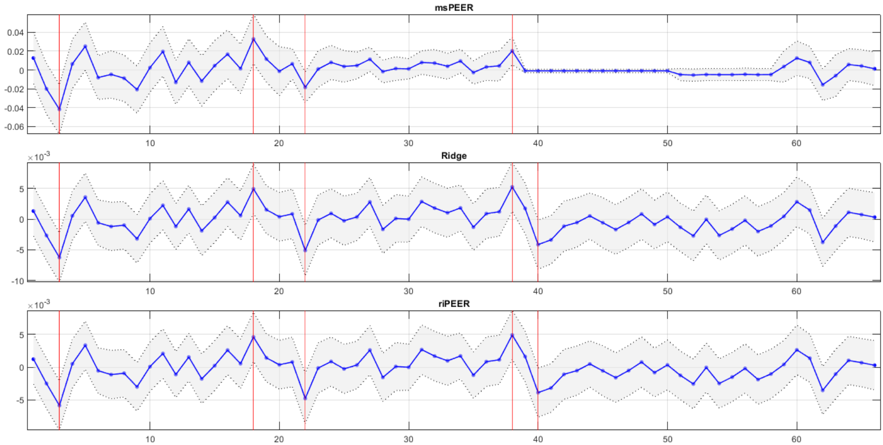
Real data analysis: some examples of estimates



Real data analysis: some examples of estimates



Real data analysis: some examples of estimates



Real data analysis: some examples of estimates