A Constrained, Weighted-$\ell 1$ Minimization Approach for Joint Discovery of Heterogeneous Neural Connectivity Graphs

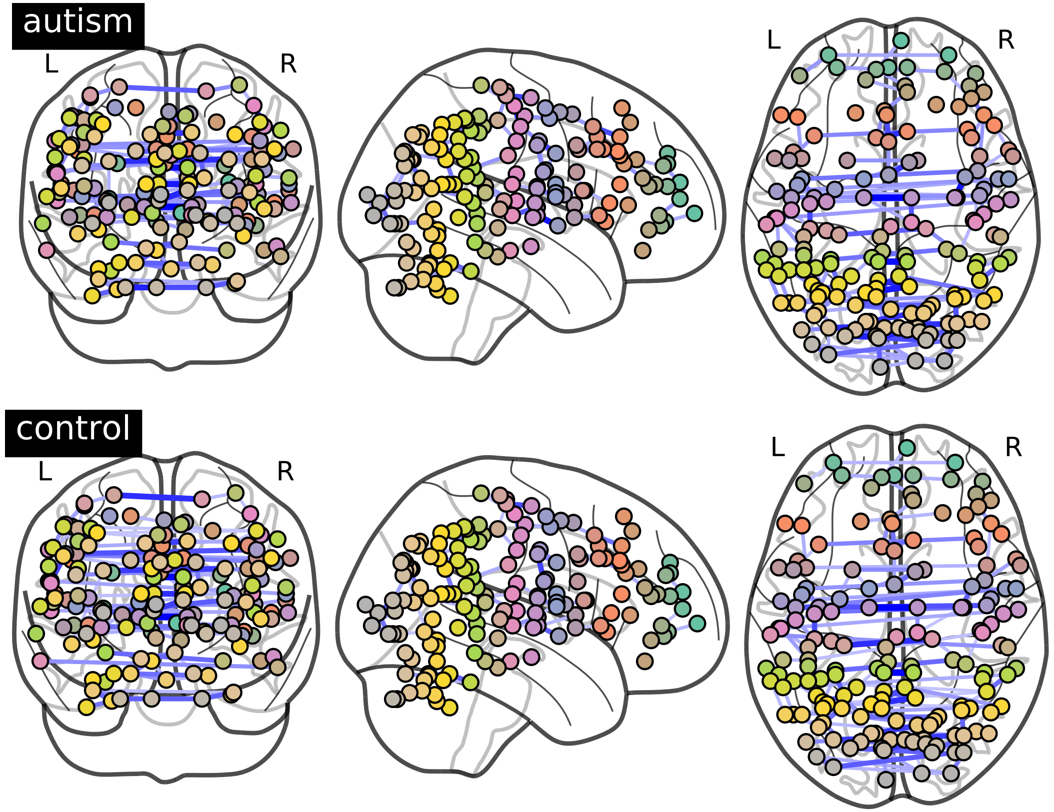
PRESENTER: BEILUN WANG

ADVISOR: PROFESSOR YANJUN QI

JOINT WORK WITH CHANDAN SINGH (NOW IN UC BERKELEY)

UVA MACHINE LEARNING AND BIOMEDICINE GROUP

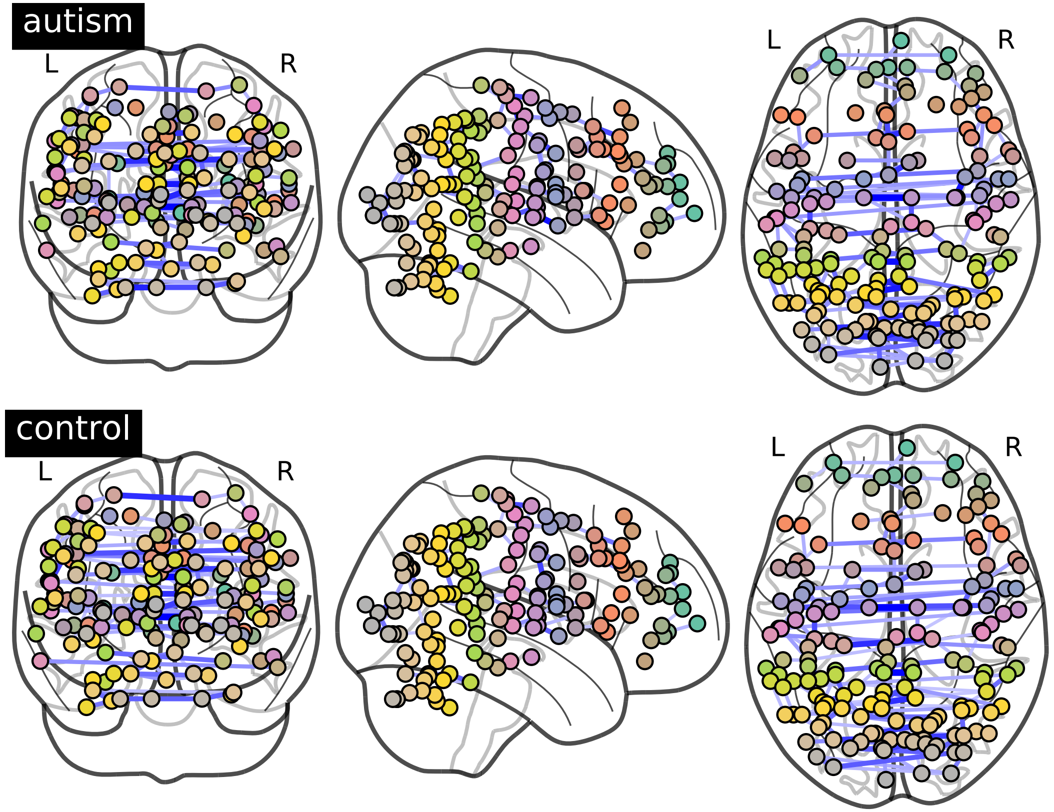
Connectomics: mapping the brain



- Brain's connectivity is largely unknown
  - 6 regions: $2^6 = 64$ possible connectomes
  - 160 regions: $2^{160} \approx 10^{78}$ possibilities

    10^{66} Trillions
  - Impossible to test out one by one
  - Need other tools/methods

Connectomics: mapping the brain

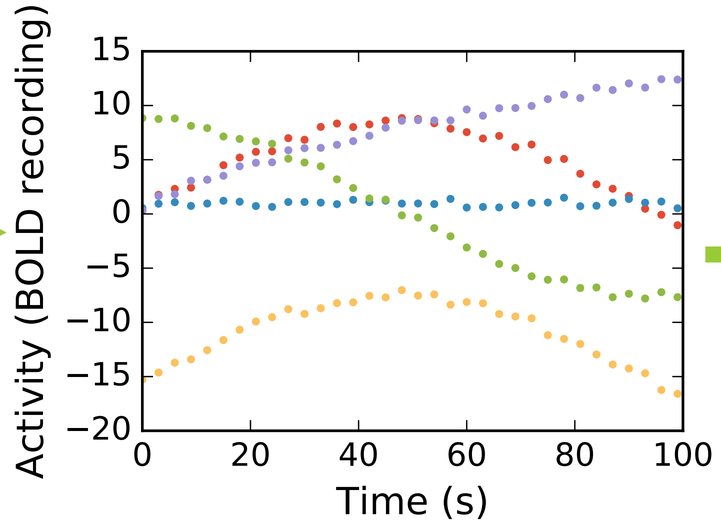


- Find patterns (connections) from brain-imaging data through machine learning methods

- A very interesting task:
  - Find differences in the brains of people with diseases, *e.g.* Autism, Alzheimer's
    - Used for understanding
    - Used for diagnosis

Background: from data to connectome

Many human brains are scanned with fMRI

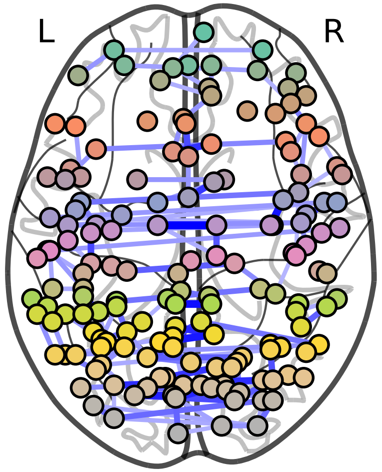


Resulting signals of activity

1	0.25	0.117	0.34	0.436
0.3	1	0.89	0.36	0.46
0.6	0.25	1	0.78	0.2
0.37	0.74	0.02	1	0.37
0.74	0.85	0.26	0.64	1

Correlations between signals

Our Focus

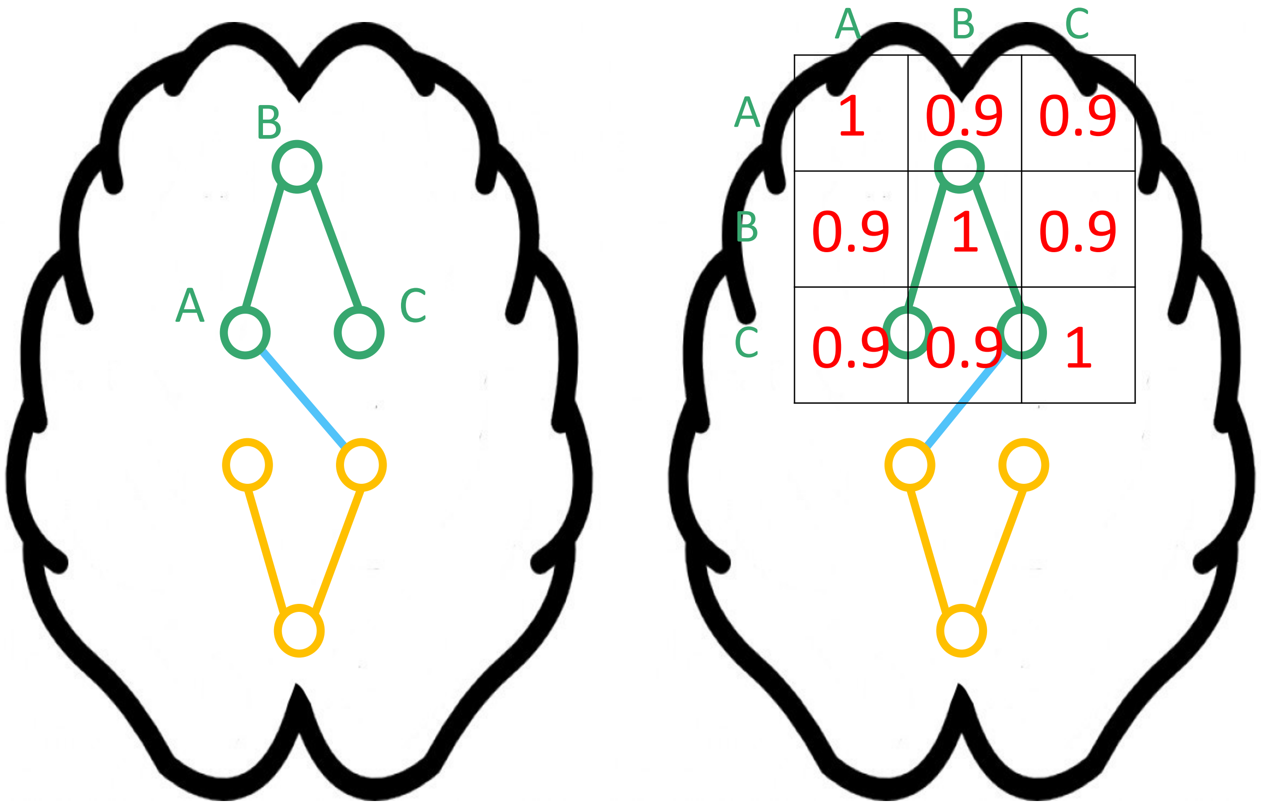


Connectome

The problem with correlations

- Correlations cannot find actual connections
  - E.g., A and C should not be connected

- Correlations cannot find the connections about differences
  - E.g., blue connection between two brains



The Solution: sparse Gaussian Graphical Model (sGGM)

1.05	-0.23	0.05	-0.02	0.05
-0.23	1.45	-0.25	0.10	-0.25
0.05	-0.25	1.10	-0.24	0.10
-0.02	0.10	-0.24	1.10	-0.24
0.05	-0.25	0.10	-0.24	1.10

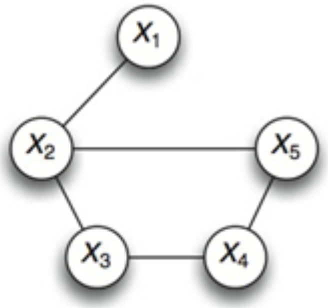
Covariances between signals (Correlation is normalized covariance)

sGGM →

1	0.2	0	0	0
0.2	1	0.2	0	0.2
0	0.2	1	0.2	0
0	0	0.2	1	0.2
0	0.2	0	0.2	1

Sparse inversion (Precision Matrix)

Decode → Sparsity pattern



Connectome

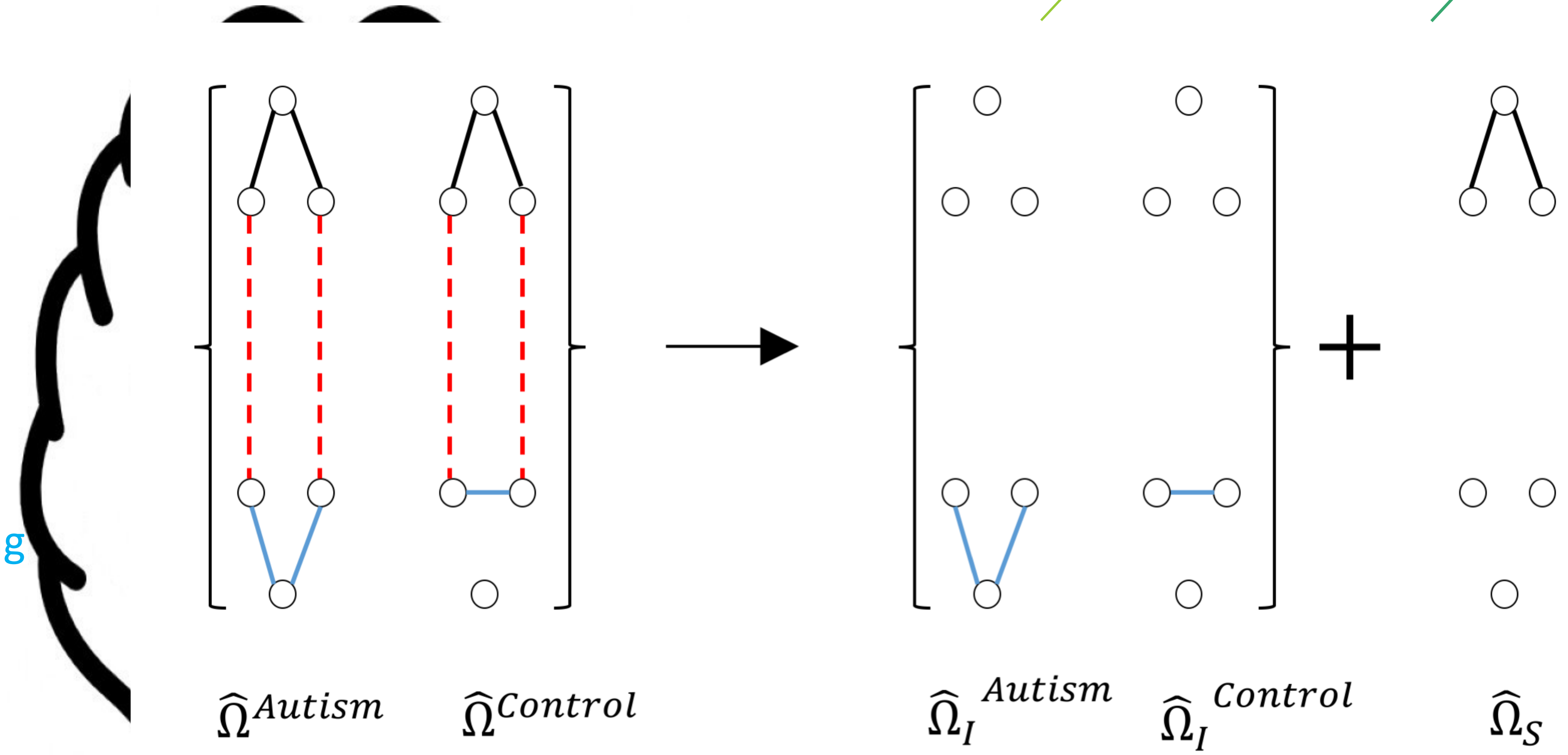
CLIME: $$\underset{\Omega}{\operatorname{argmin}} |\Omega|_1 \quad \text{subject to: } |\widehat{\Sigma}\Omega - I|_\infty \leq \lambda$$

The Solution: W-SIMULE

$$\widehat{\Omega}_I^{(1)}, \ldots, \widehat{\Omega}_I^{(k)}, \widehat{\Omega}_S = \sum_i \underset{\Omega_I^{(i)}, \Omega_S}{\operatorname{argmin}} ||W \cdot \Omega_I^{(i)}||_1 + \epsilon k ||W \cdot \Omega_S||_1$$

$$\text{Subject to: } ||\Sigma^{(i)}(\Omega_I^{(i)} + \Omega_S) - I||_\infty \le \lambda, i = 1, \ldots, k.$$

1. Nonparanormal assumption
2. Sparse
3. Imposes prior
4. Multi-task learning
5. Parallelizable



The Solution: W-SIMULE

$$\widehat{\Omega}_I^{(1)},\dots,\widehat{\Omega}_I^{(k)},\widehat{\Omega}_S = \sum_i \operatorname*{argmin}_{\Omega_I^{(i)},\Omega_S} ||W\cdot\Omega_I^{(i)}||_1 + \epsilon k||W\cdot\Omega_S||_1$$

Subject to: $||\Sigma^{(i)}(\Omega_I^{(i)}+\Omega_S)-I||_\infty \le \lambda, i=1,\dots,k.$

Column-wise

$$\operatorname*{argmin}_{\beta^{(i)},\beta^s}\sum_i ||W_{,j}\cdot\beta^{(i)}||_1+\epsilon K||W_{,j}\cdot\beta^s||_1$$

Subject to: $||\Sigma^{(i)}(\beta^{(i)}+\beta^s)-e_j||_\infty\le\lambda, i=1,\dots,K$

Re-formulation

$$\operatorname*{argmin}_{\theta} ||W_{,j}\cdot\theta||_1$$

Subject to: $|\mathbf{A}^{(i)}\theta-b|_\infty\le c,\ i=1,\dots,K$

Where $\mathbf{A}^{(i)}=[0,\dots,0,\Sigma^{(i)},0,\dots,0,\frac{1}{\epsilon K}\Sigma^{(i)}],$

$\theta=[\beta^{(1)^T},\dots,\beta^{(K)^T},\epsilon K(\beta^s)^T]^T,$

$\boldsymbol{b}=\mathbf{e}_j, c=\lambda$

Linear Programming

$$\operatorname*{argmin}_{\theta^+,\theta^-} W_{,j}\cdot\theta^+ + W_{,j}\cdot\theta^-$$

Subject to :

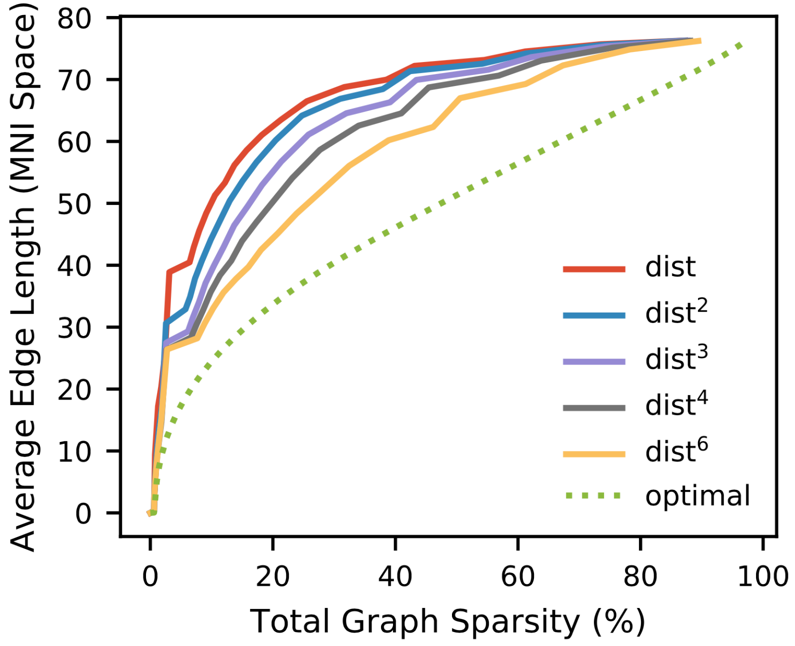
$$\begin{pmatrix}\mathbf{A}^{(i)} & -\mathbf{A}^{(i)}\\ -\mathbf{A}^{(i)} & \mathbf{A}^{(i)}\end{pmatrix}\begin{pmatrix}\theta^+\\ \theta^-\end{pmatrix}\le\begin{pmatrix}c+b\\ c-b\end{pmatrix}$$

$$\begin{pmatrix}\theta^+\\ \theta^-\end{pmatrix}\ge 0$$

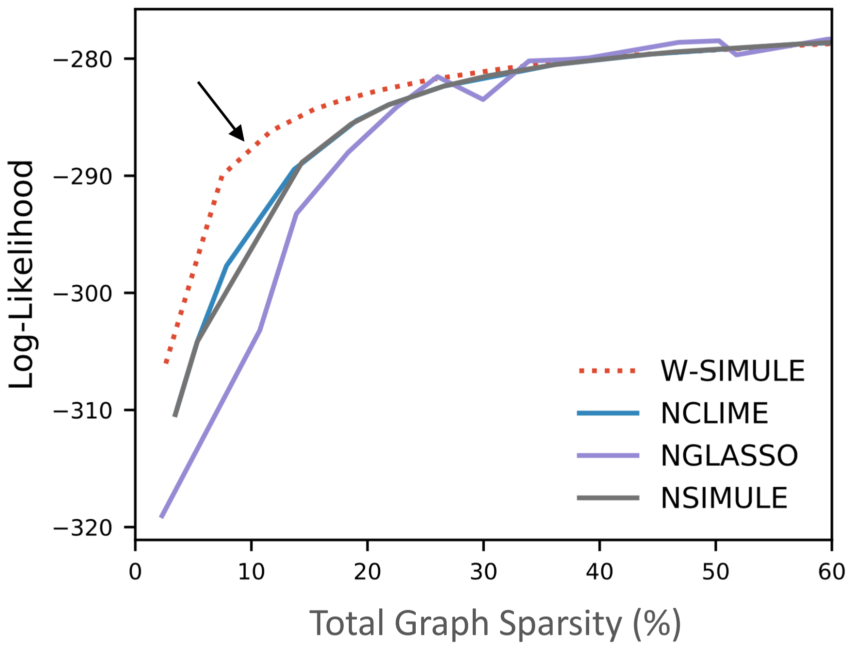
Apply lp solver

Results: validation

Connection lengths decrease



Fits data we have seen



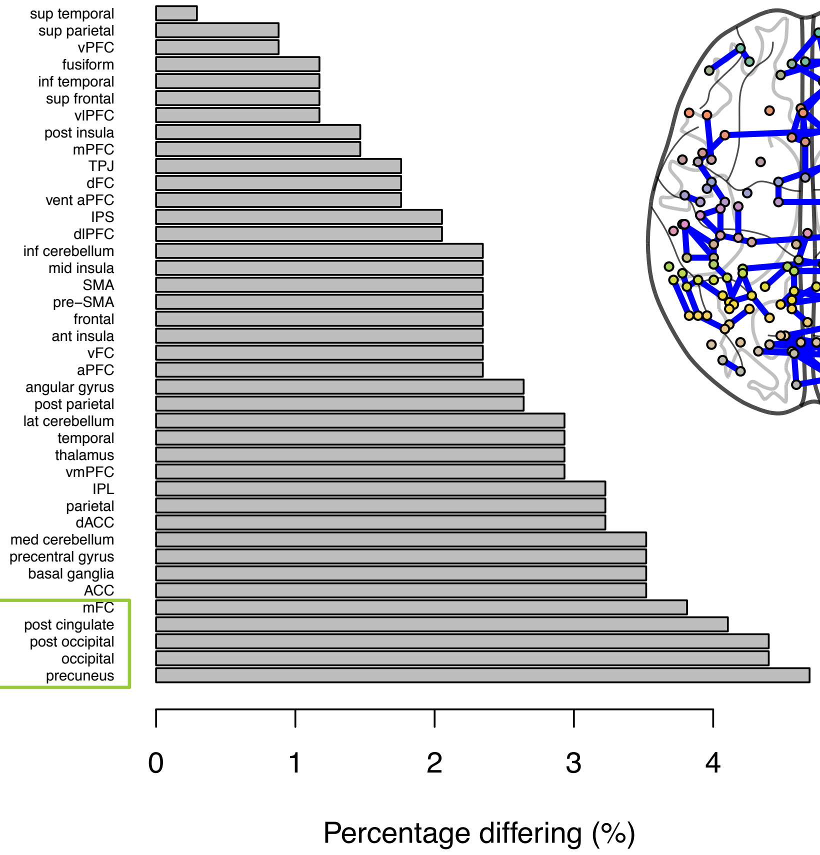
Finds differences between groups
Apply QDA

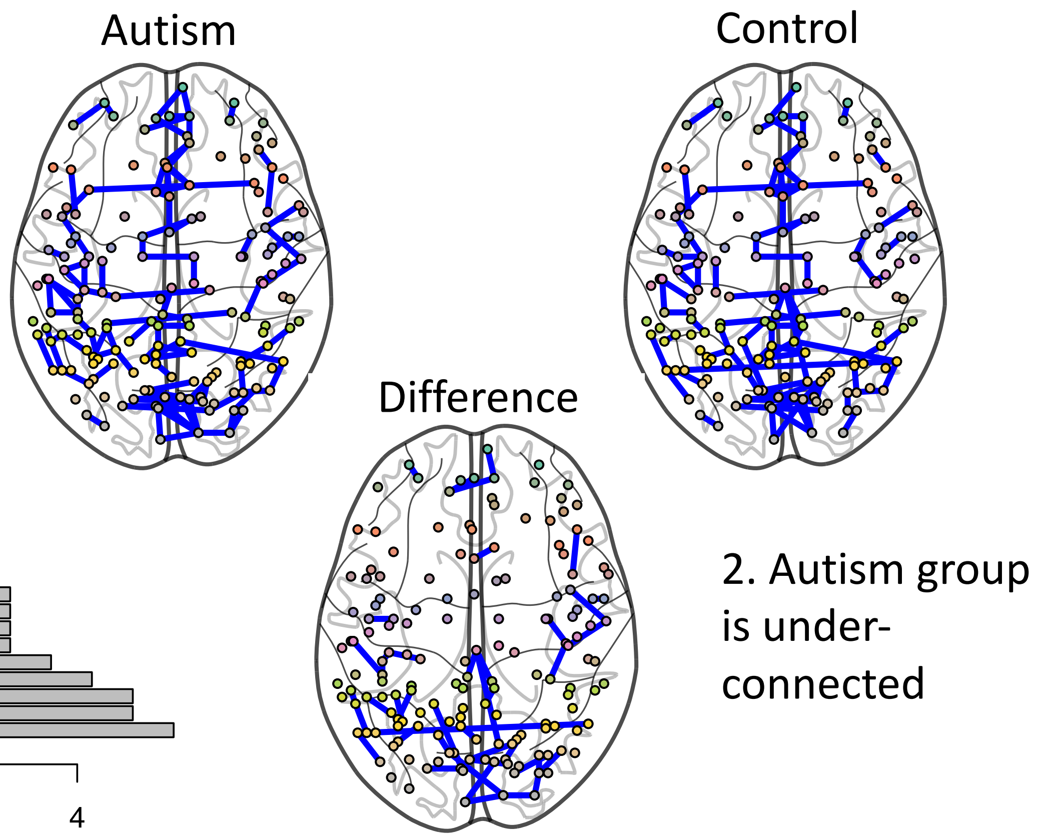
Method	Accuracy (%)
W-SIMULE	**58.62%**
CLIME	46.55%
GLASSO	53.71%
SIMULE	57.96%
JGL (fused)	56.90%
SIMONE	53.71%

Conclusions

1. Areas differing in autism

- Medial frontal cortex
- Post cingulate
- Post occipital
- Occipital lobe
- Precuneus



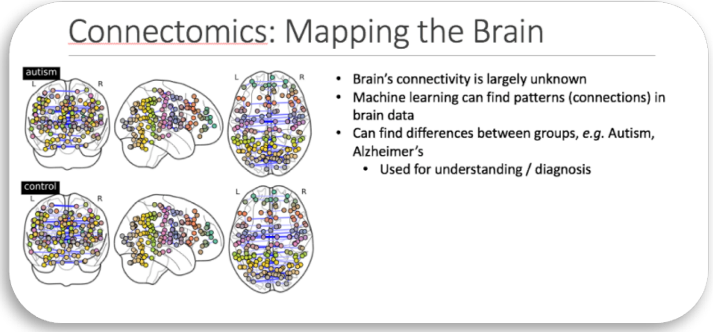


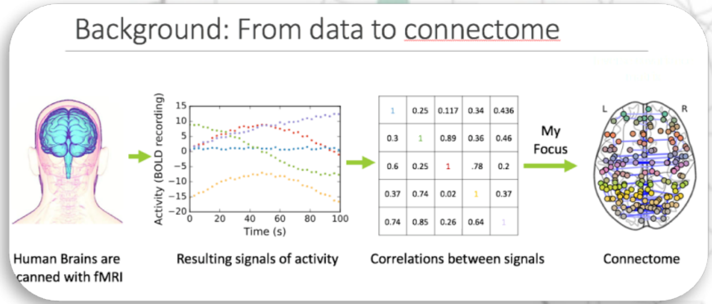
2. Autism group is under-connected

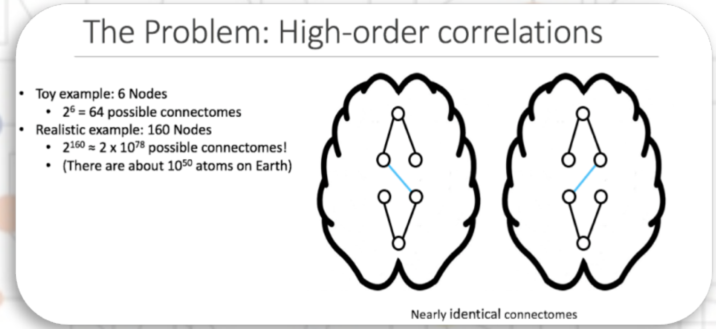
Conclusions

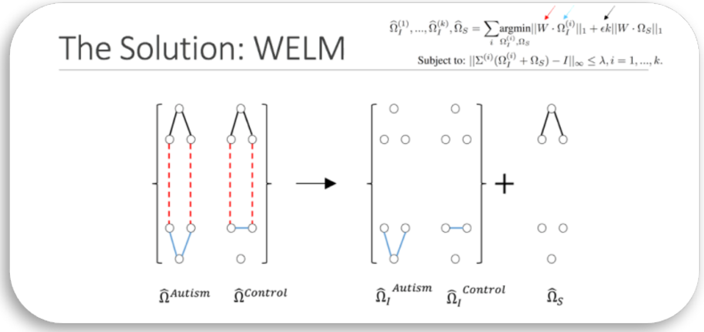
- The project website: http://jointggm.org/
- R package “simule”
  - Install.packages(“simule”)
  - Demo(wsimuleDemo)
  - https: //cran.r-project.org/web/packages/simule/index.html

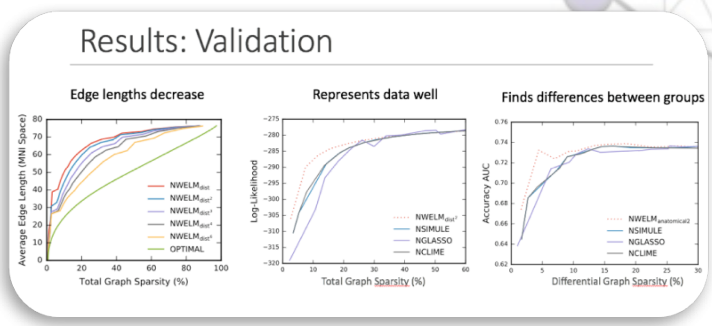
Appendix

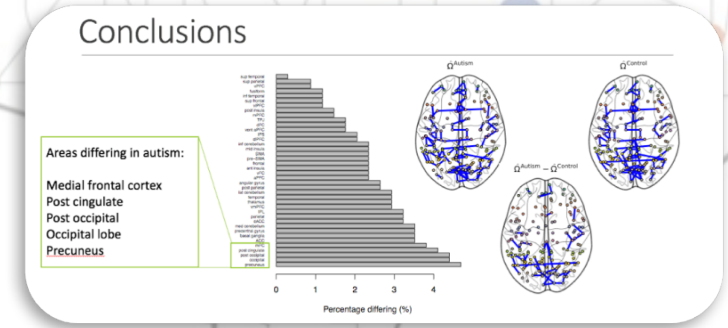

Connectomics: Mapping the Brain
autism
control
Brain's connectivity is largely unknown
Machine learning can find patterns (connections) in brain data
Can find differences between groups, e.g. Autism, Alzheimer's
Used for understanding / diagnosis


Background: From data to connectome
Human Brains are scanned with fMRI
Activity (BOLD recording)
Time (s)
Resulting signals of activity
Correlations between signals
My Focus
Connectome


The Problem: High-order correlations
Toy example: 6 Nodes
2^6 = 64 possible connectomes
Realistic example: 160 Nodes
2^160 ≈ 2 x 10^78 possible connectomes!
(There are about 10^50 atoms on Earth)
Nearly identical connectomes


The Solution: WELM


Results: Validation
Edge lengths decrease
Represents data well
Finds differences between groups
Average Edge Length (MNI Space)
Total Graph Sparsity (%)
Log-Likelihood
Accuracy AUC
Differential Graph Sparsity (%)
NSIMULE
NGLASSO
NCLIME
OPTIMAL


Conclusions
Areas differing in autism:
Medial frontal cortex
Post cingulate
Post occipital
Occipital lobe
Precuneus
Percentage differing (%)