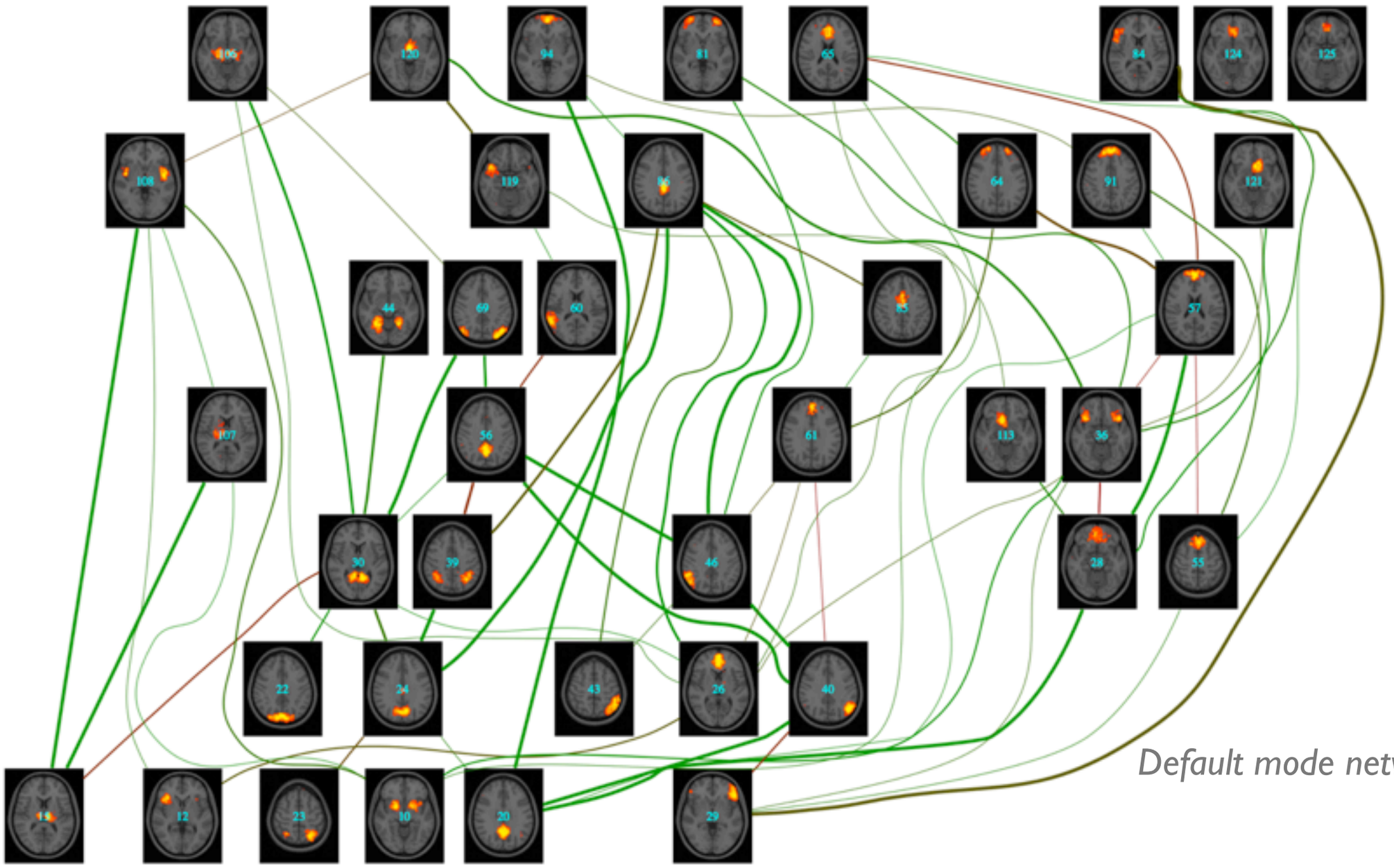
Don’t fall in love with your method

Steve Smith FMRIB Oxford

Don’t fall in love with your method

Steve Smith FMRIB Oxford

Detailed Network Analyses

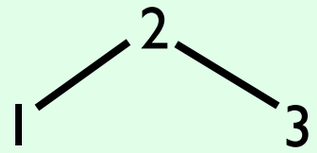


Default mode network

A network model comprises:

- "**Nodes**" distinct functional voxels/regions
- "**Edges**" connections between nodes

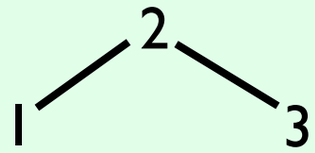
What do we want to estimate?



from \ to	1	2	3
1	0	1	0
2	1	0	1
3	0	1	0

- What are the network edges?

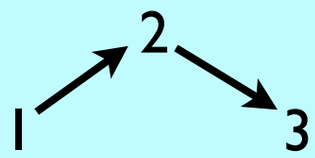
What do we want to estimate?



from \ to	1	2	3
1	0	1	0
2	1	0	1
3	0	1	0

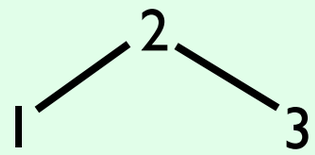
- What are the network edges?

- What is the dominant directionality of the edges?



from \ to	1	2	3
1	0	1	0
2	0	0	1
3	0	0	0

What do we want to estimate?



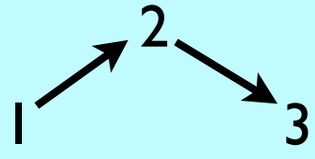
from \ to	1	2	3
1	0	1	0
2	1	0	1
3	0	1	0

- What are the network edges?

Monday

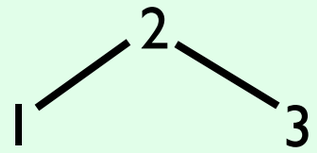
Wednesday

- What is the dominant directionality of the edges?



from \ to	1	2	3
1	0	1	0
2	0	0	1
3	0	0	0

What do we want to estimate?

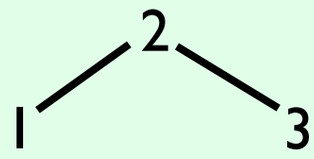


from \ to	1	2	3
1	0	1	0
2	1	0	1
3	0	1	0

- What are the network edges?

Direct vs. Indirect Connections

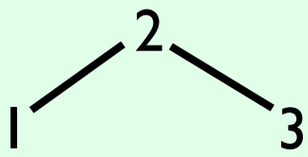
suppose the truth is:



from \ to	1	2	3
1	0	1	0
2	1	0	1
3	0	1	0

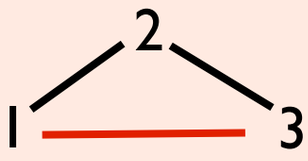
Direct vs. Indirect Connections

suppose the truth is:



from \ to	1	2	3
1	0	1	0
2	1	0	1
3	0	1	0

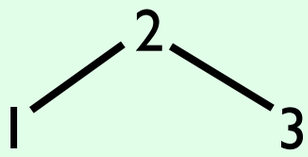
but 1 will correlate with 3, so if we estimate edges using (full) correlation, we will (wrongly) estimate:



from \ to	1	2	3
1	0	1	1
2	1	0	1
3	1	1	0

Direct vs. Indirect Connections

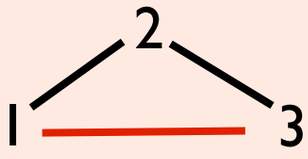
suppose the truth is:



from \ to	1	2	3
1	0	1	0
2	1	0	1
3	0	1	0

but 1 will correlate with 3, so if we estimate edges using (full) correlation, we will (wrongly) estimate:

1-3 is referred to as an "indirect connection"

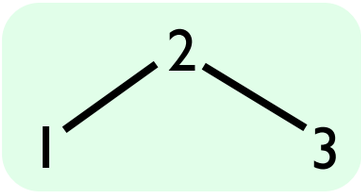


from \ to	1	2	3
1	0	1	1
2	1	0	1
3	1	1	0

Disambiguating direct vs. indirect connections

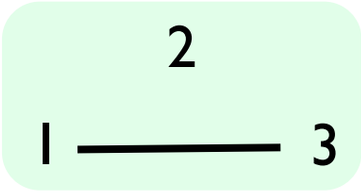
- Need to take into account multiple nodes' timecourses in order to estimate *direct* edges

- E.g., DCM, SEM, Bayes Nets, LiNGAM



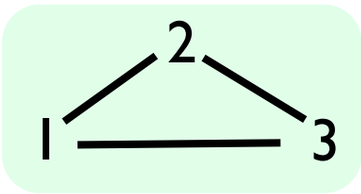
vs.

- In some cases (e.g. large numbers of nodes, no input timings known), "principled models" such as SEM, DCM are not estimable



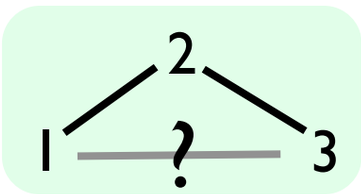
vs.

- One useful approximation to SEM is "partial correlation"
  - *Marrelec NeuroImage 2006*
  - *Fransson NeuroImage 2008*
  - *Varoquaux NIPS 2010*



Disambiguating direct vs. indirect connections using partial correlation

- Before correlating 1 and 3, first regress 2 out of both ("orthogonalise wrt 2")
- If 1 and 3 are *still* correlated, a direct connection exists



- More generally, first regress *all other nodes' timecourses* out of the pair in question
- Equivalent to the *inverse covariance* matrix

- Urgh! If you have 200 nodes and 100 timepoints, this is impossible!
- A problem of DoF - need large #timepoints - #nodes

Degrees of freedom and estimability when using partial correlation

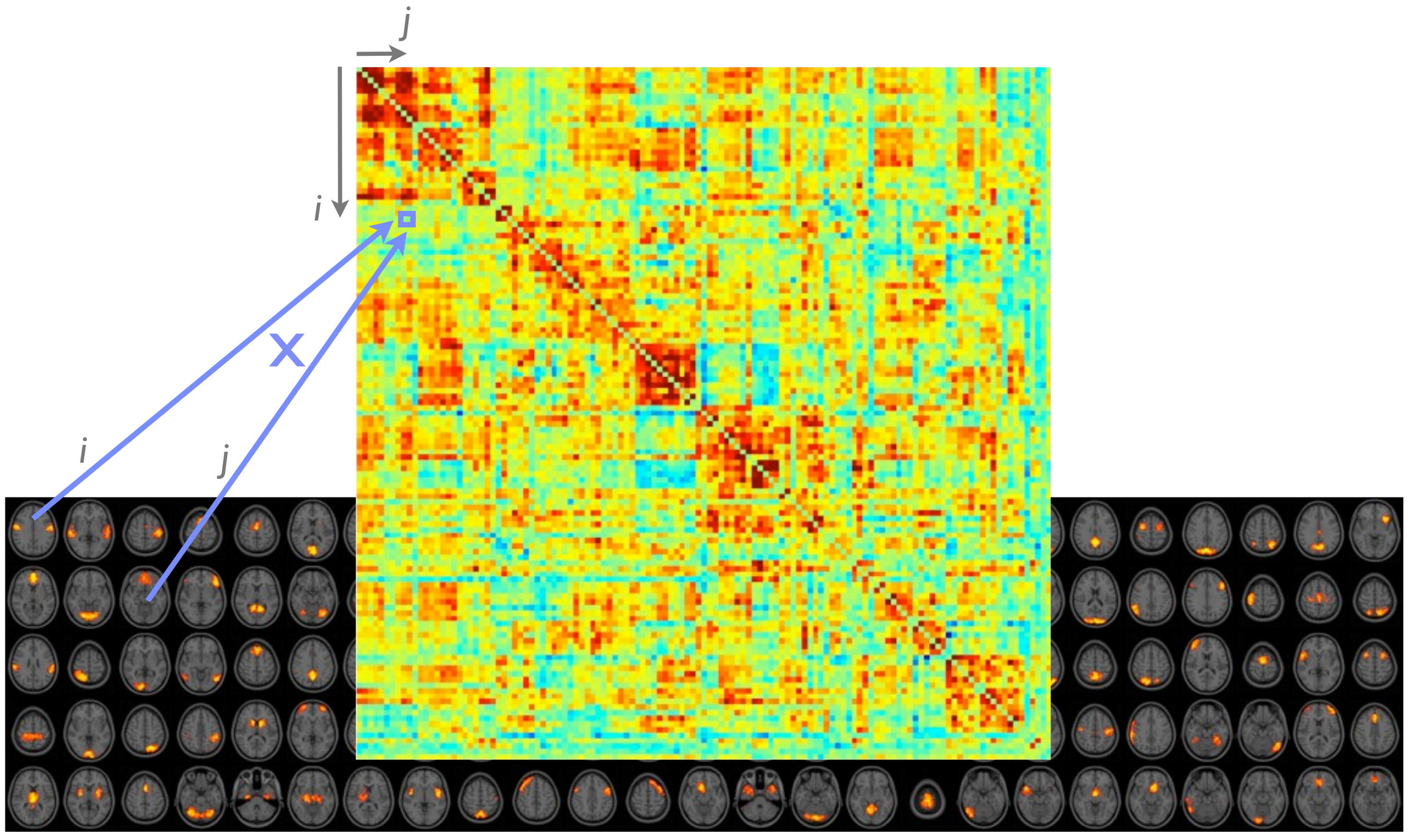
- When inverting a "rank-deficient" matrix it is common to aid this with some mathematical conditioning, e.g. force it to be sparse (force low values that are poorly estimated to zero)

- E.g. "ICOV" (regularised inverse covariance)
  - *Friedman Biostat 2008*

- Can improve further when analysing multiple subjects
  - *Varoquaux NIPS 2010*

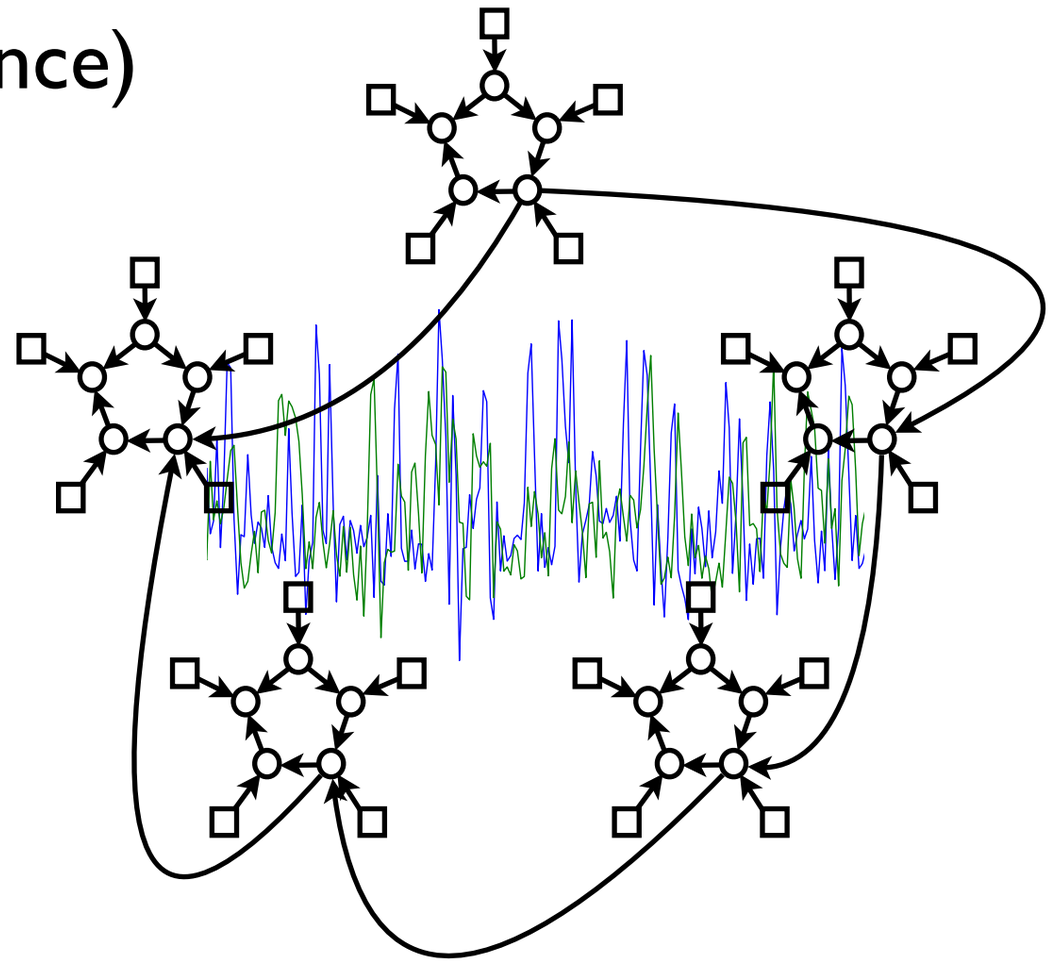
- But still important to maximise temporal DoF

So: what is the right method for estimating connectivity, based on FMRI timeseries?



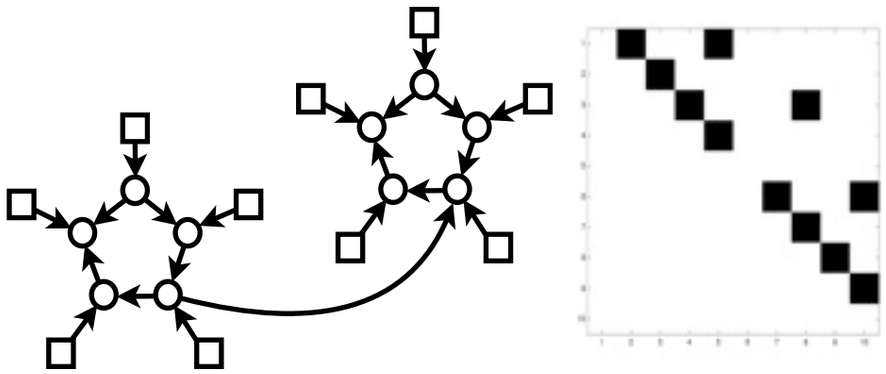
Network Modelling Methods for FMRI

- Ground-truth networks used to simulate BOLD timeseries
- 28 scenarios (e.g. TR=0.25s vs 3s)
- Compare 38 network modelling methods for estimating:
  - **direct** connections (edge presence)
  - edge **directionality** (causality)

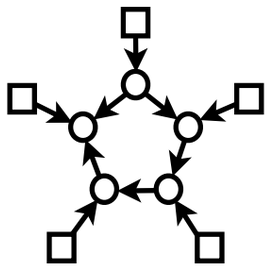


- *Smith NeuroImage 2011*

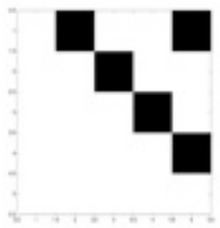
The simulated networks



S10

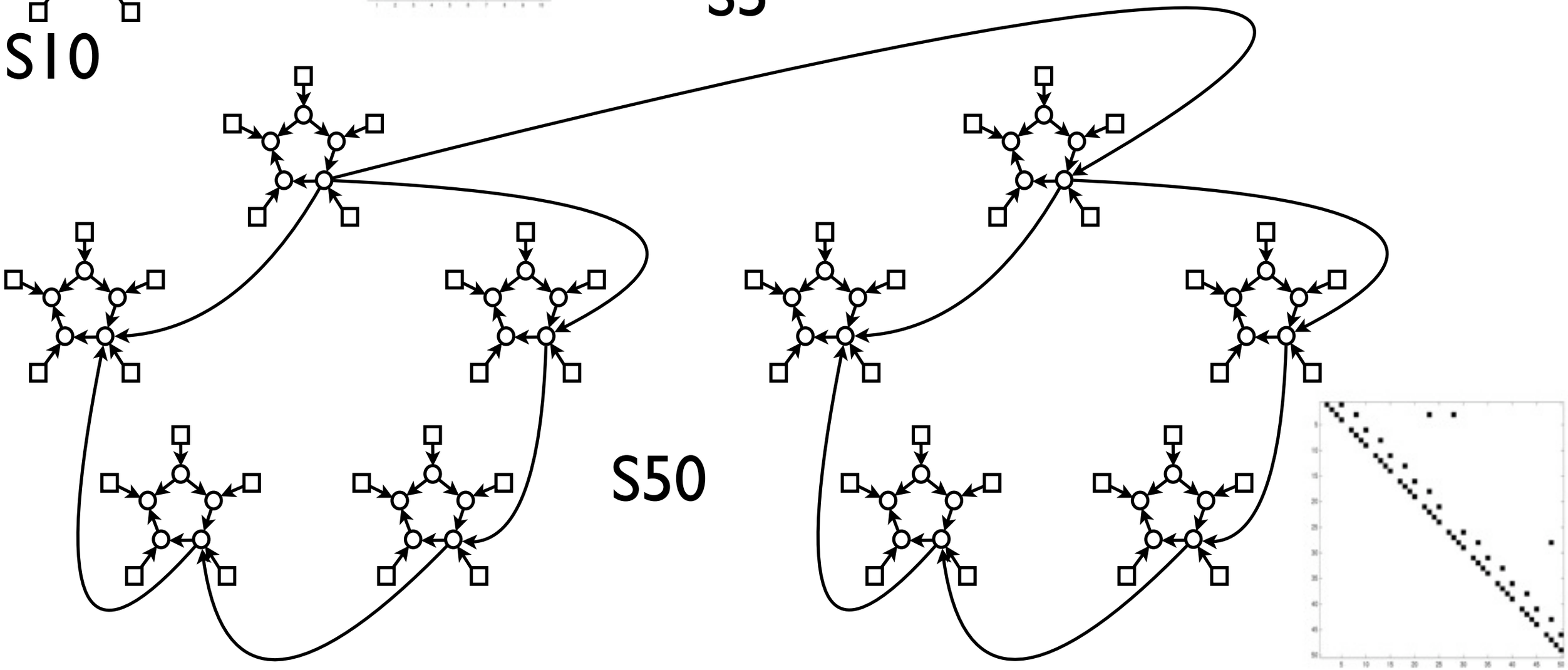


S5

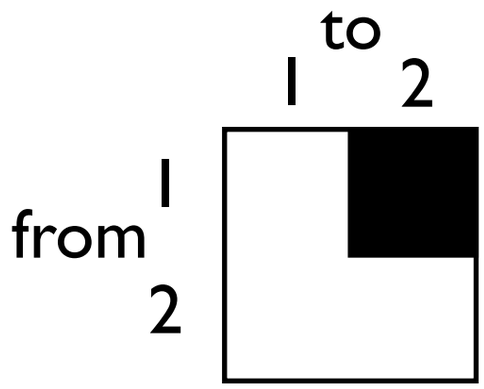


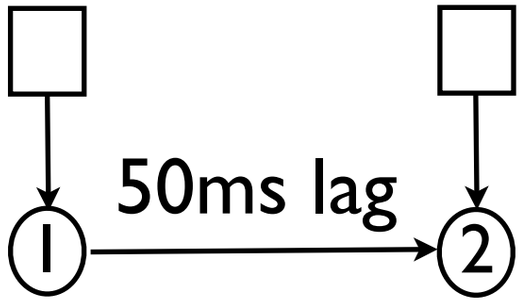
external input

network node (viewed data)



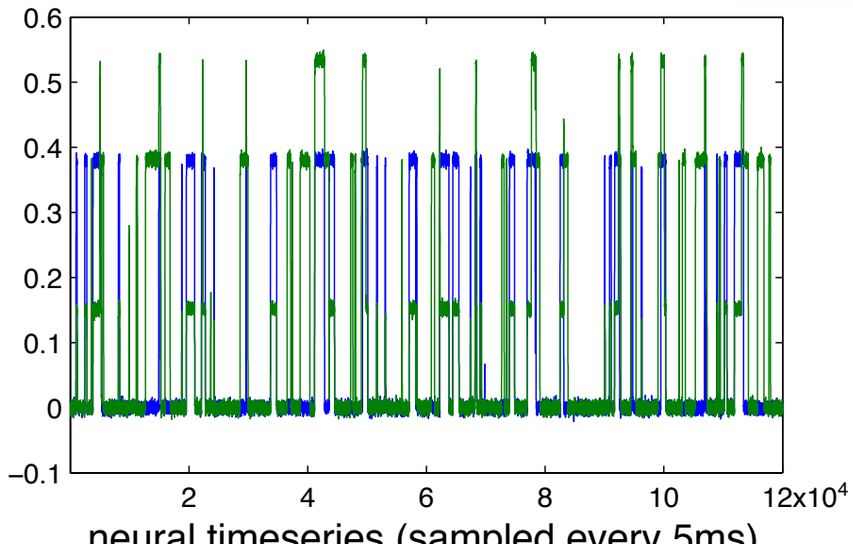
S50



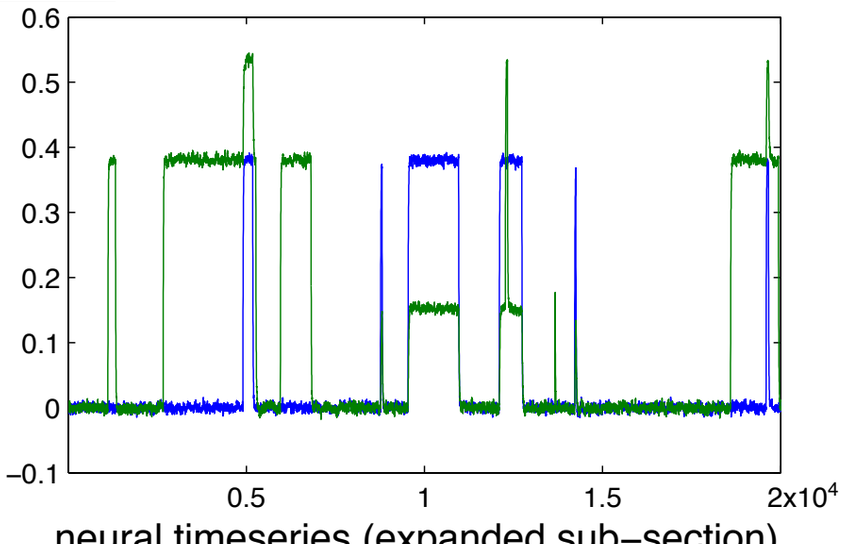


external input

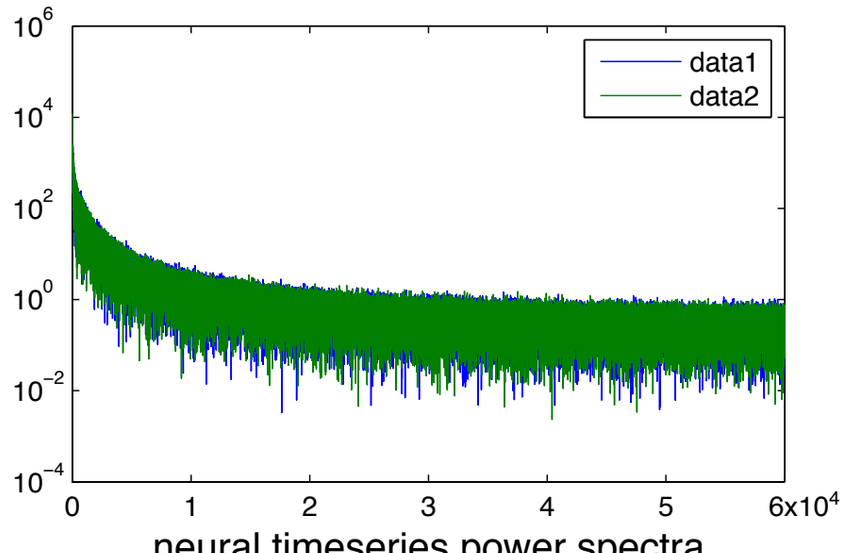
network node
(viewed data)



neural timeseries (sampled every 5ms)



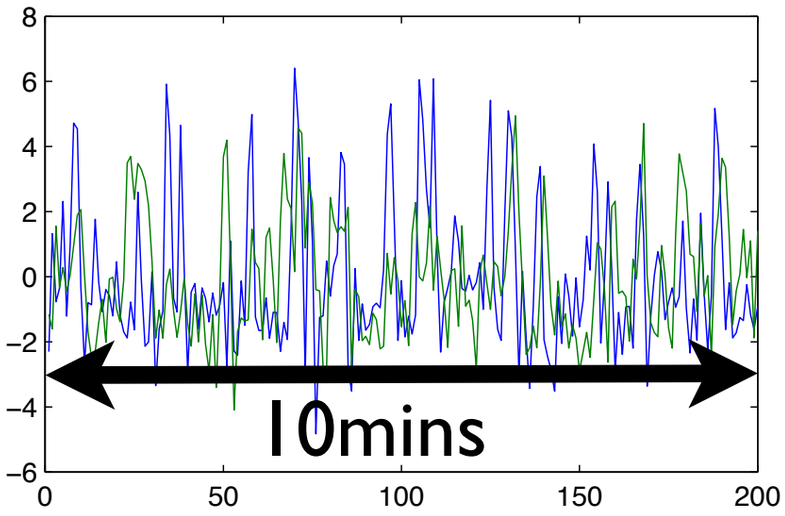
neural timeseries (expanded sub−section)



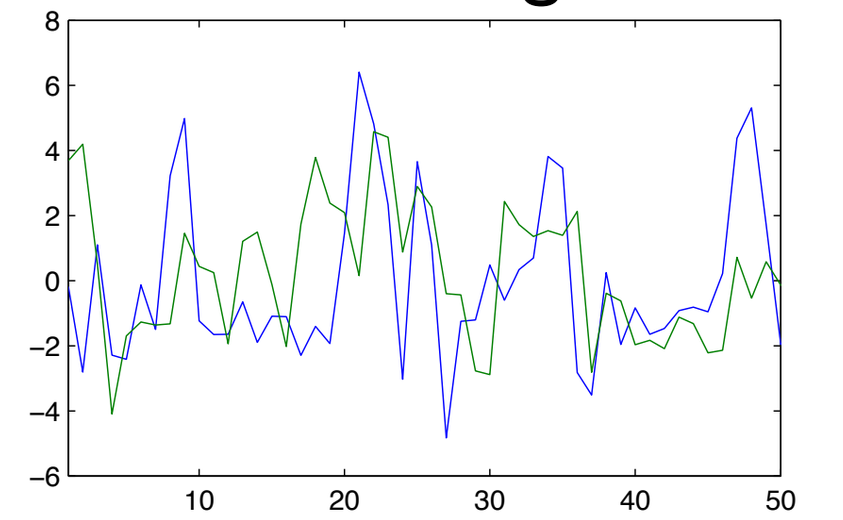
neural timeseries power spectra

Balloon nonlinear model for neural to haemodynamics

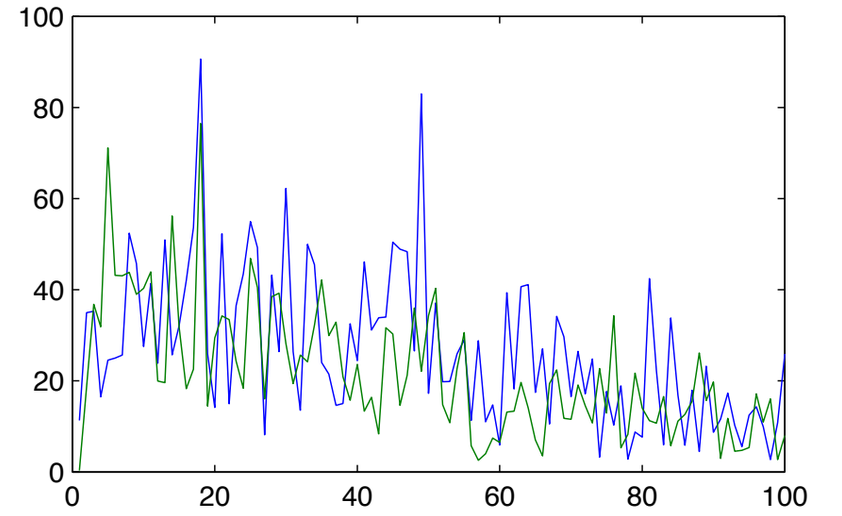
- variable HRF lag (±0.5s) across nodes
- TR = 0.25s - 3s
- add noise 0.1-1% to BOLD signal



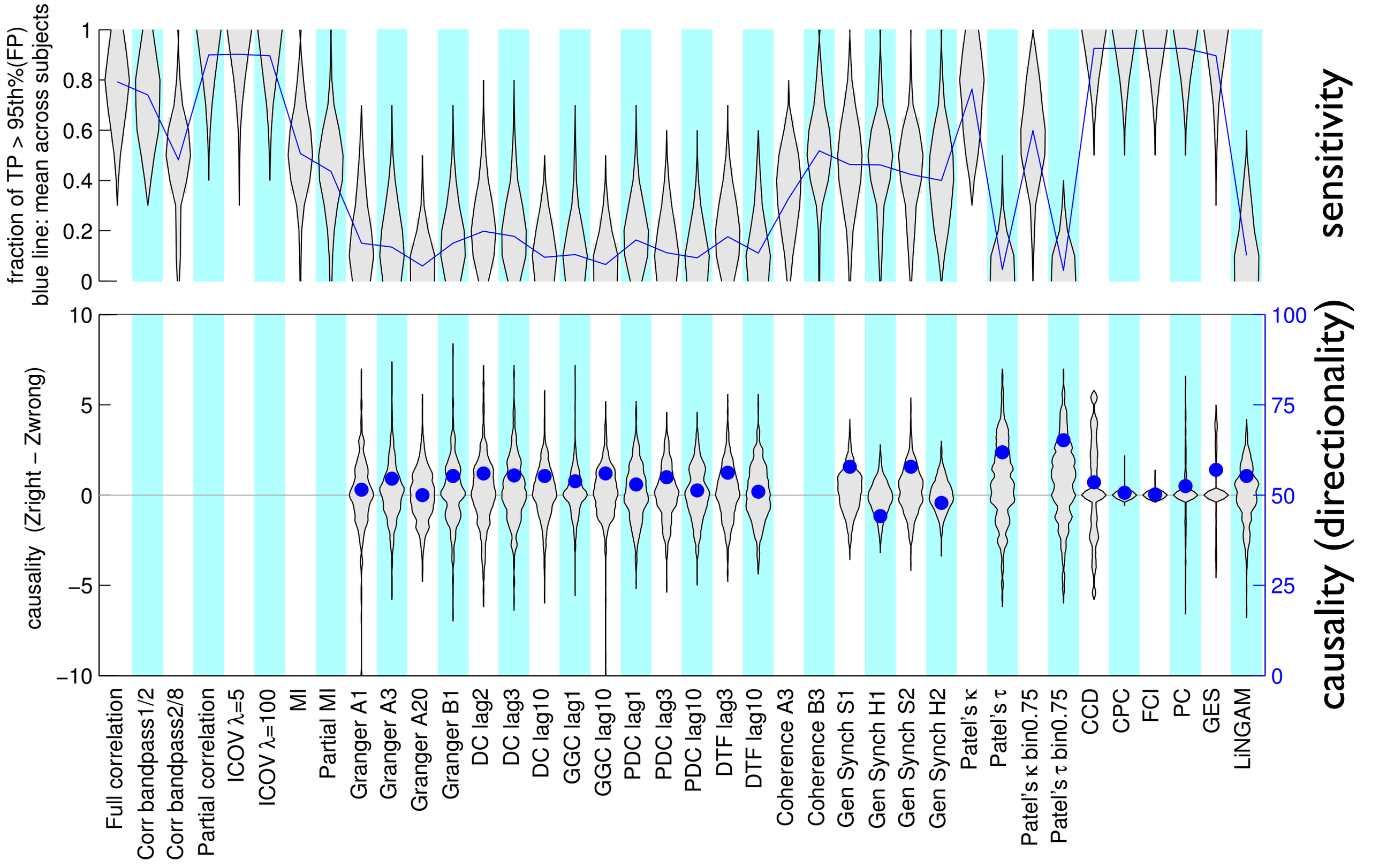
BOLD FMRI timeseries (sampled every 3s)



BOLD FMRI timeseries (expanded sub−section)



BOLD FMRI timeseries power spectra



10 nodes, 10-minute sessions, TR=3s, noise=1%

fraction of TP > 95th%(FP)
blue line: mean across subjects

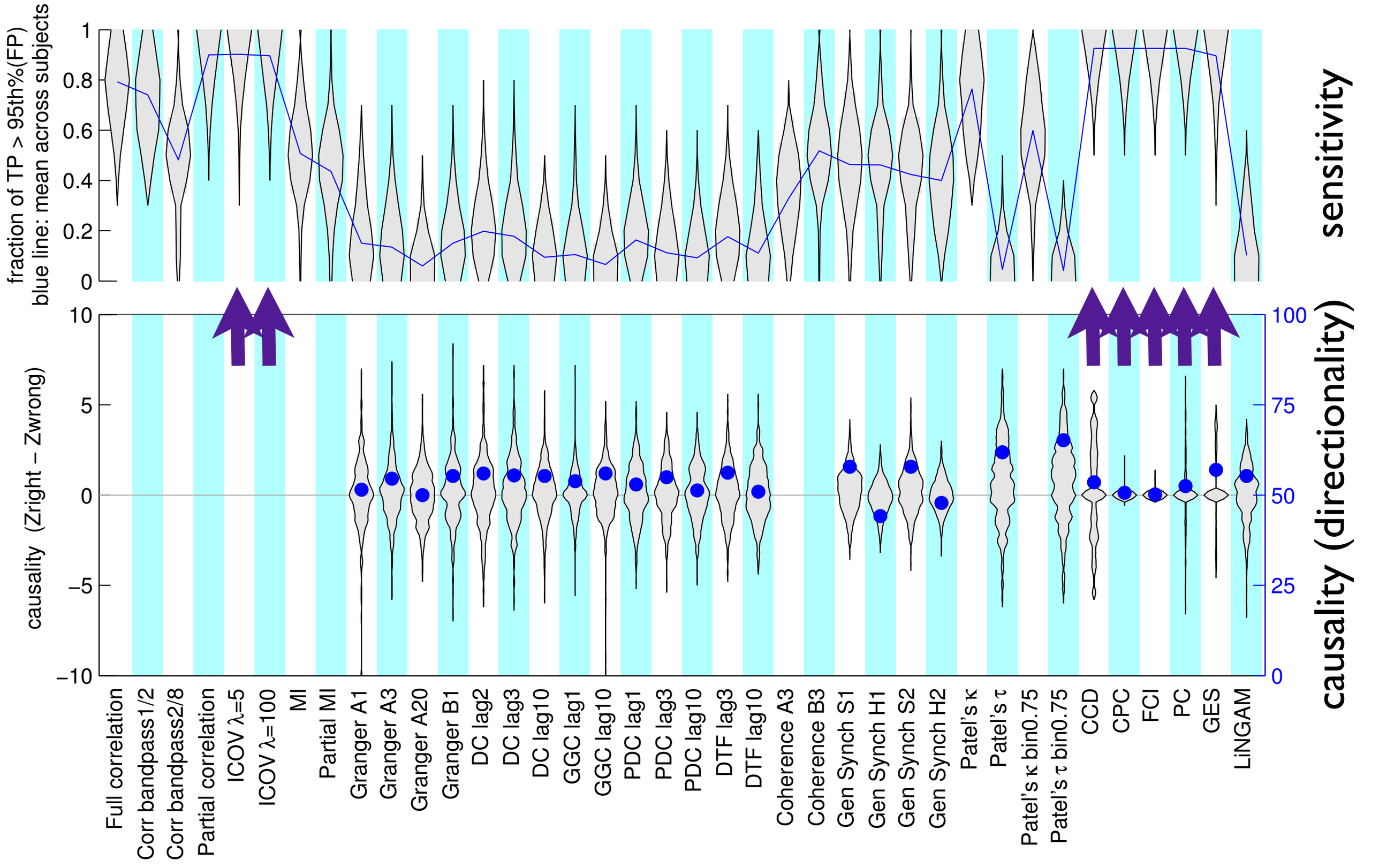
sensitivity

causality (Zright – Zwrong)

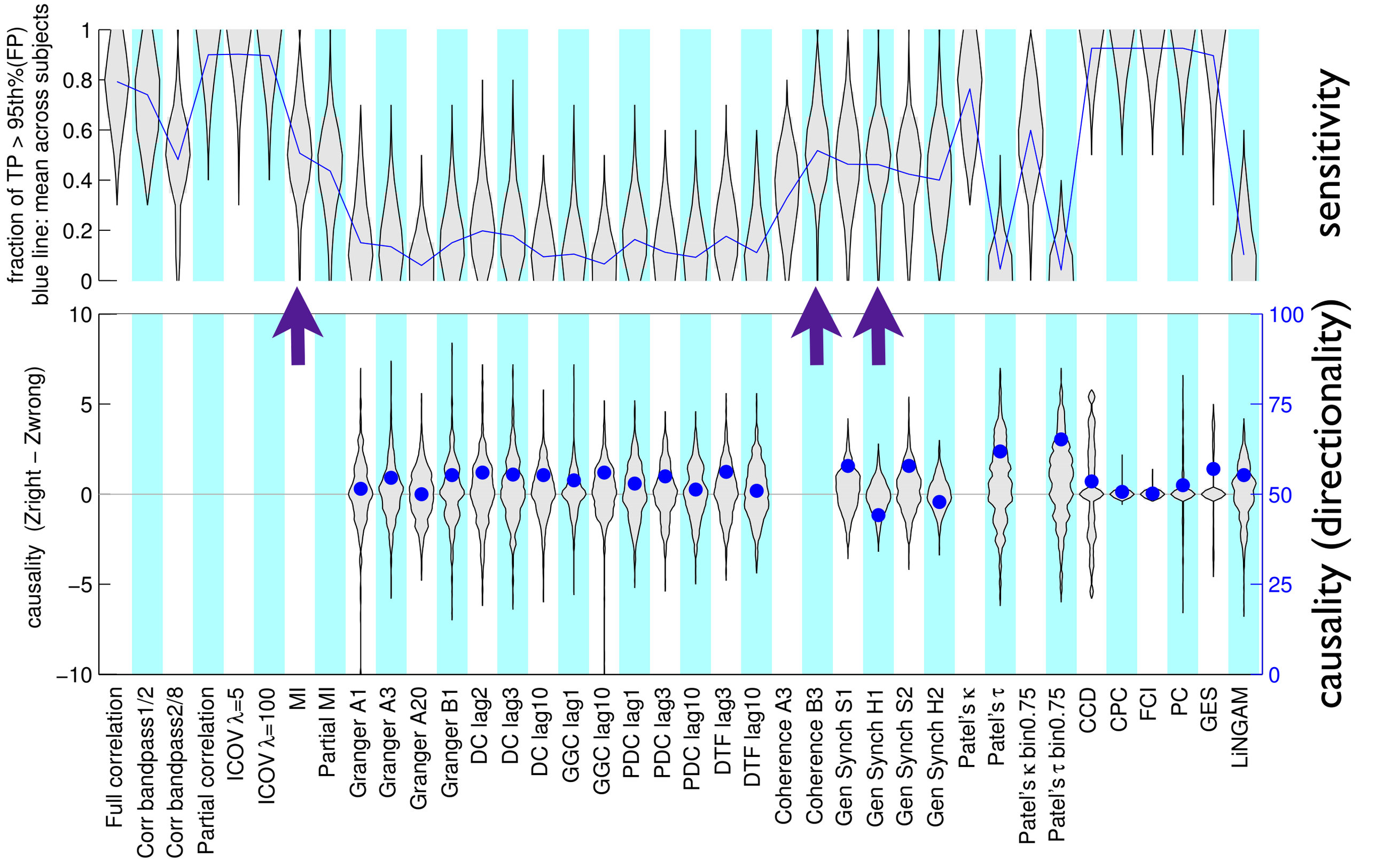
causality (directionality)

Full correlation
Corr bandpass1/2
Corr bandpass2/8
Partial correlation
ICOV λ=5
ICOV λ=100
MI
Partial MI
Granger A1
Granger A3
Granger A20
Granger B1
DC lag2
DC lag3
DC lag10
GGC lag1
GGC lag10
PDC lag1
PDC lag3
PDC lag10
DTF lag3
DTF lag10
Coherence A3
Coherence B3
Gen Synch S1
Gen Synch H1
Gen Synch S2
Gen Synch H2
Patel's κ
Patel's τ
Patel's κ bin0.75
Patel's τ bin0.75
CCD
CPC
FCI
PC
GES
LiNGAM

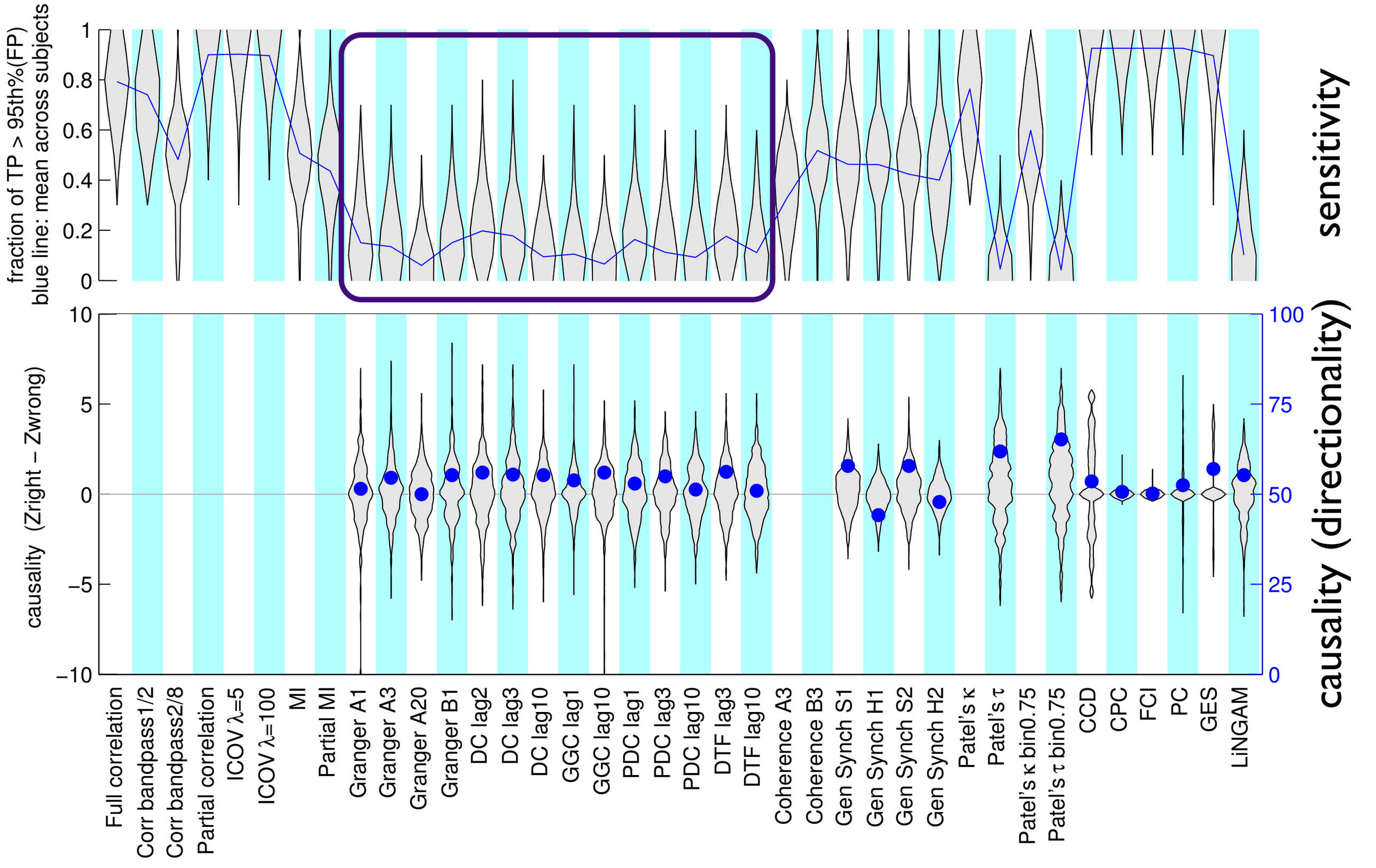
Partial corr better than *Full corr*



Also *ICOV* and *Bayes net* methods very good



MI, Coherence, Gen Synch significantly worse

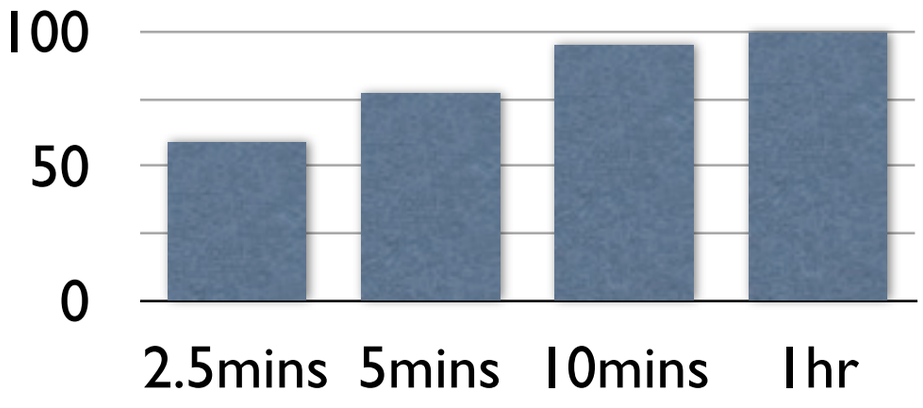


Lag-based (*Granger, PDC, DTF*) very bad

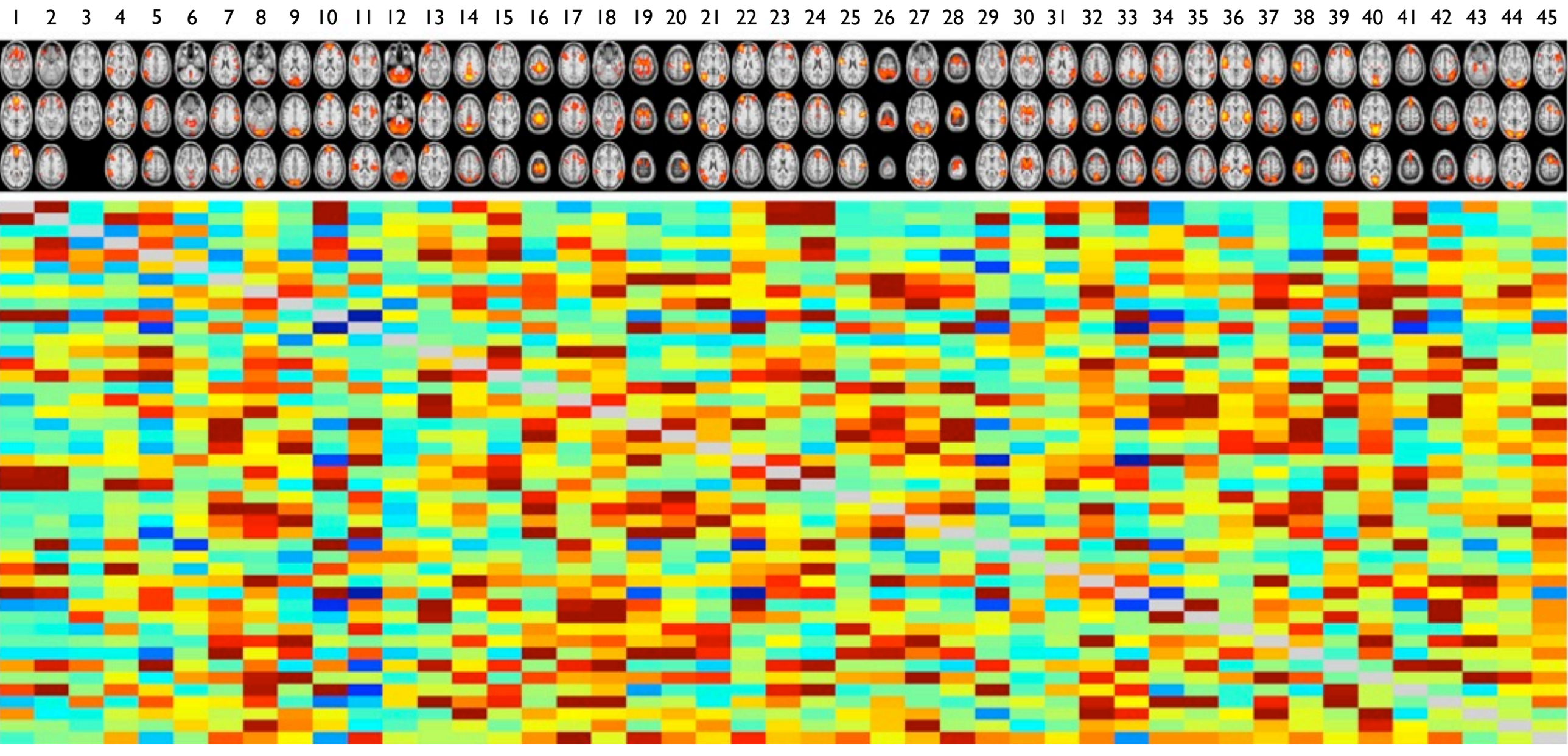
Summary: Estimating Network Connections

- Wide range of simulations show that in FMRI data, covariance/correlation-based methods perform the best
- Suggests that data is not *strongly* non-Gaussian
  (see also *Hlinka, NeuroImage 2010*)
- However, as data quality improves, more sophisticated measures (e.g. MI) may start to be more useful

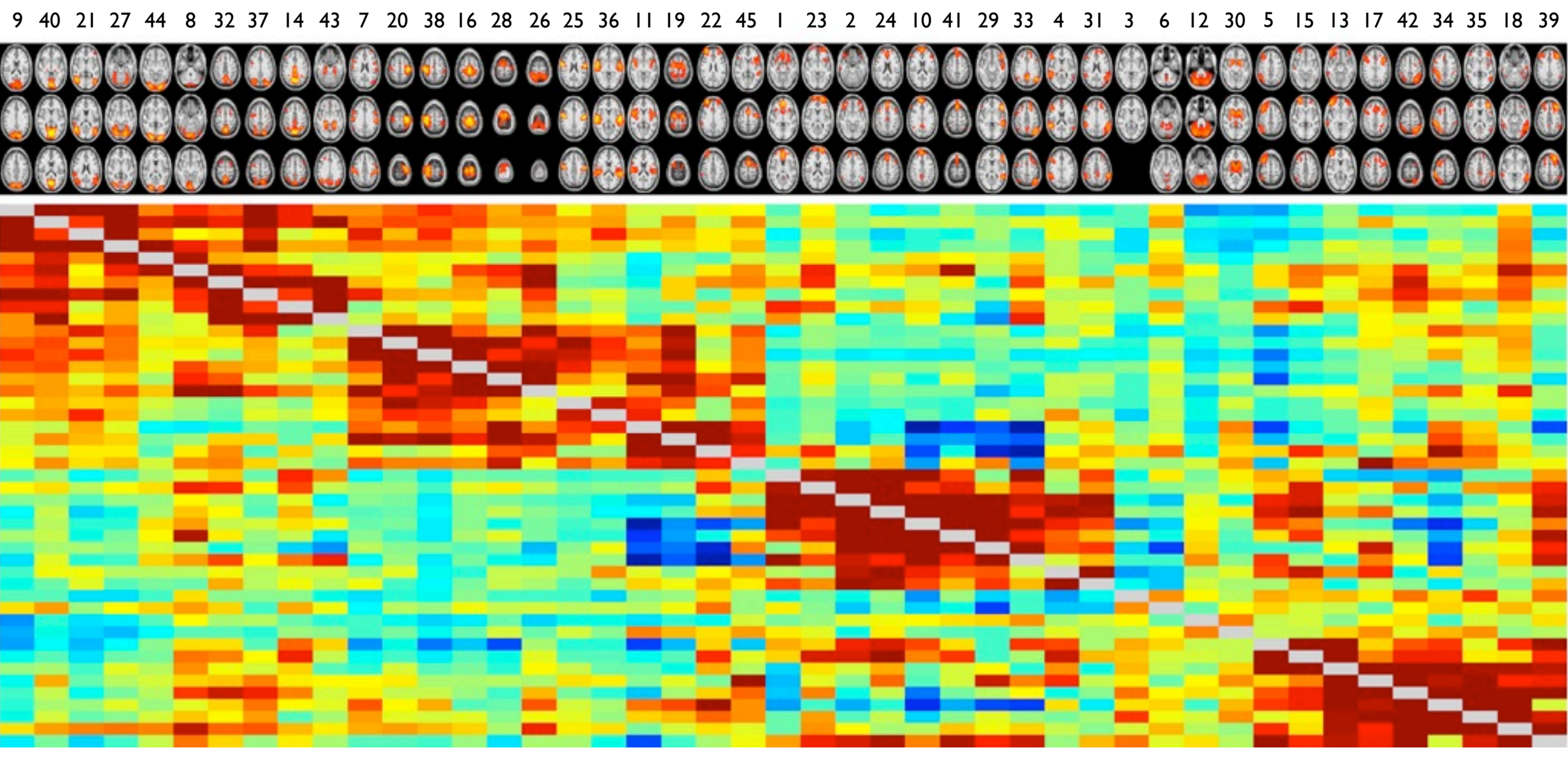
- From the covariance-based methods, partial correlation and ICOV better than full correlation (at distinguishing *direct* from *indirect* connections)



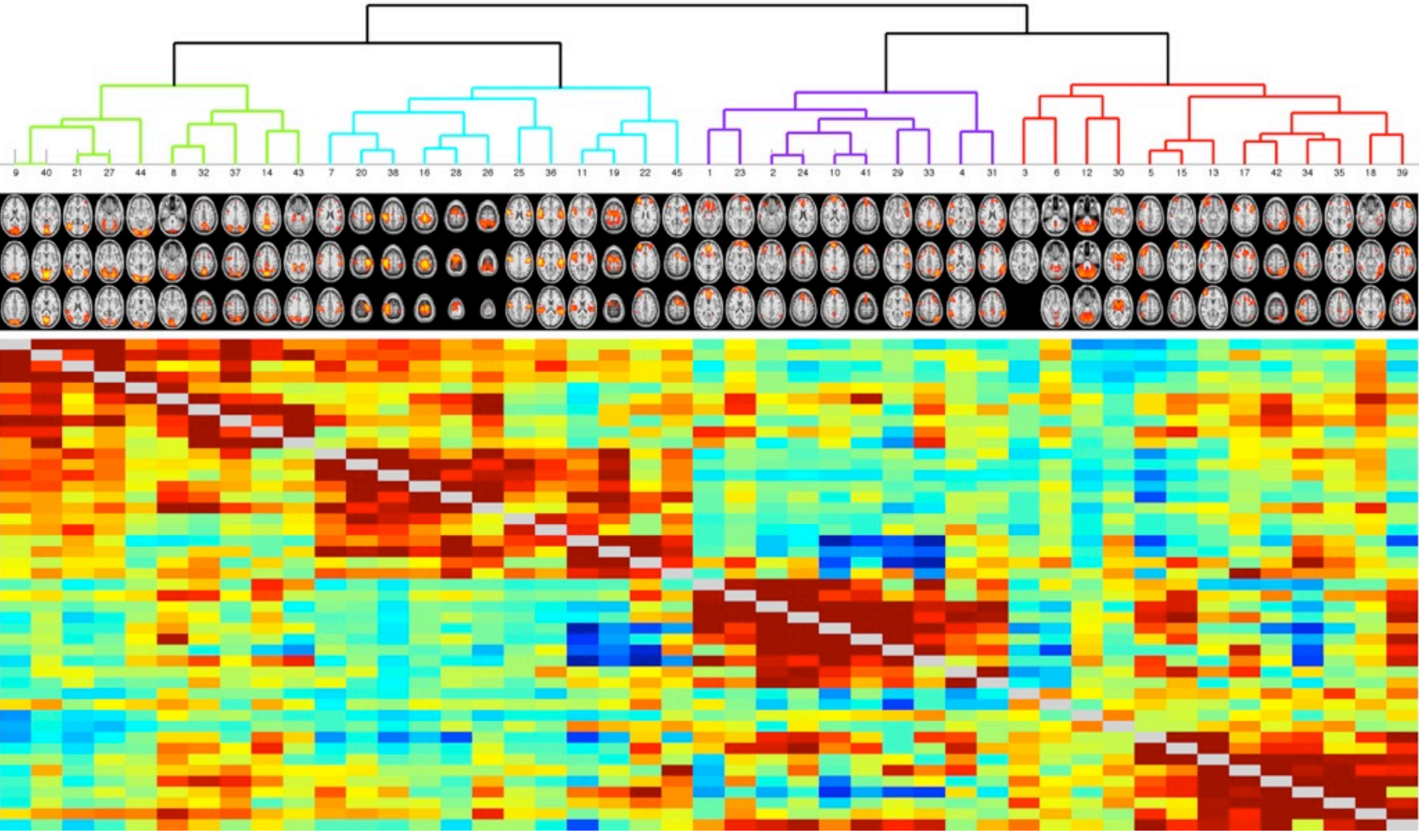
so we get a *Nodes x Nodes* network matrix



reorder node ordering to find clusters

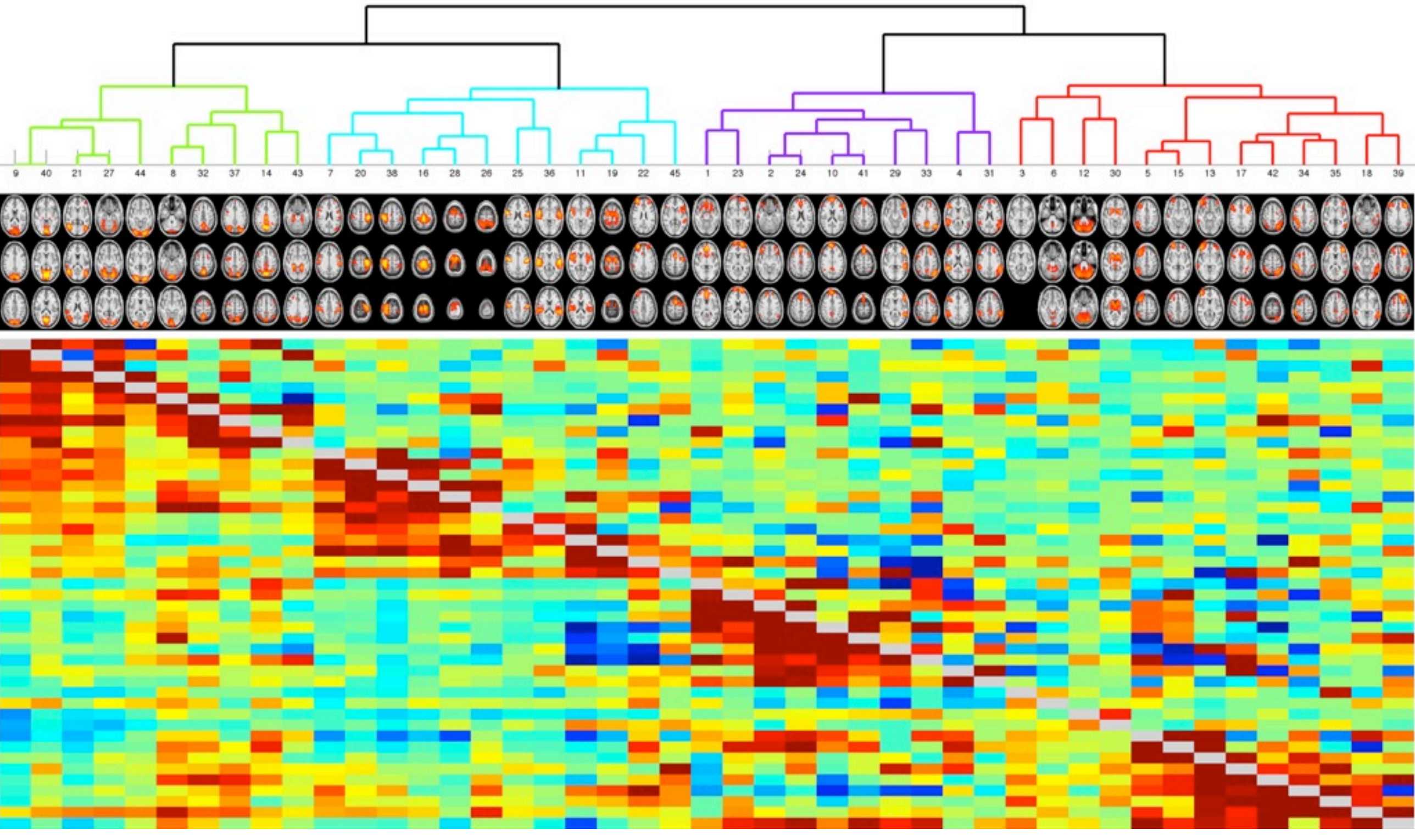


can view hierarchy of clusters



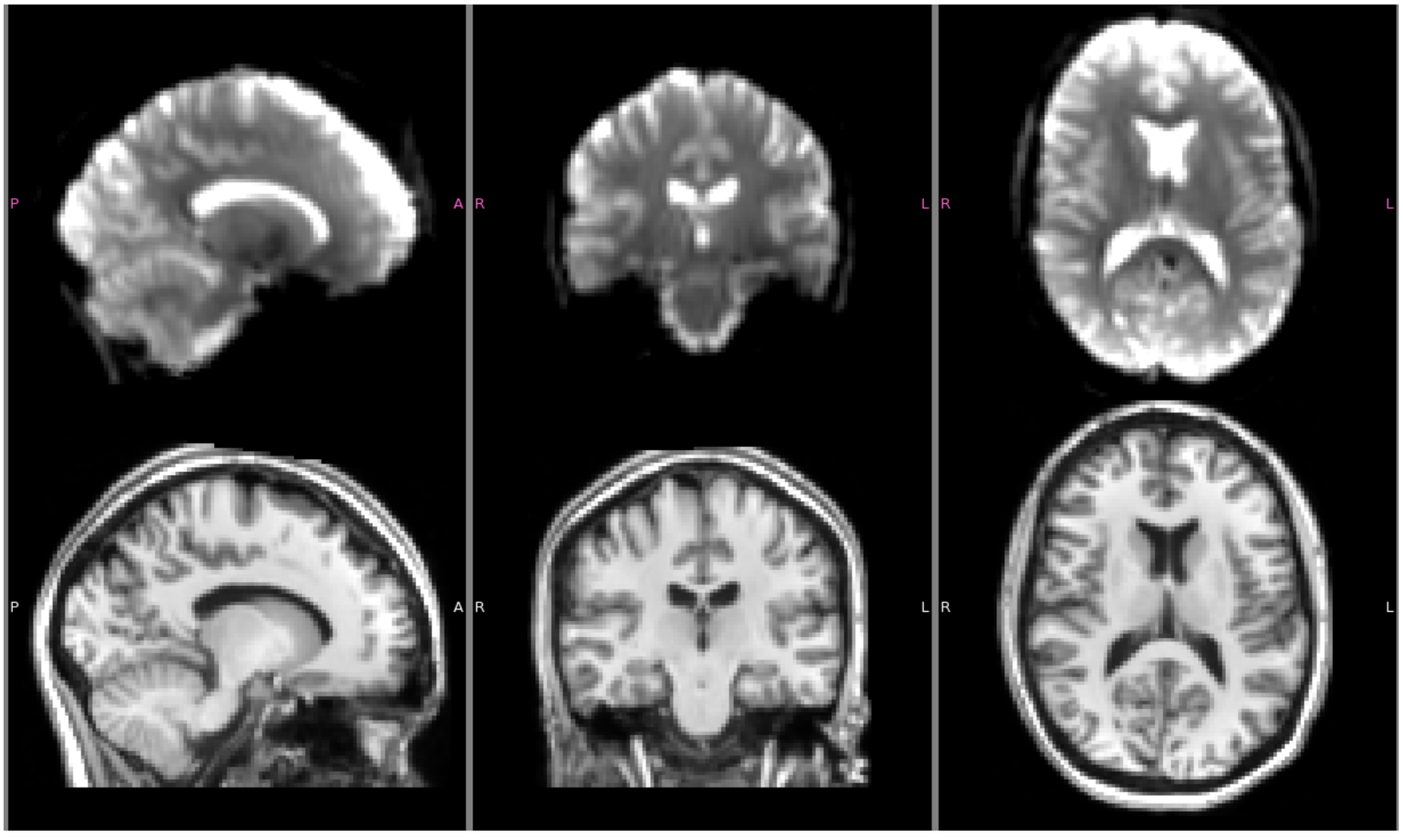
- *Cordes MRI 2002*
- *Salvador Cerebral Cortex 2005*

partial correlation matrix sparser than full



HCP - Pushing Spatial & Temporal Resolution

UMinn: Steen Moeller, Gordon Xu, Essa Yacoub, Kamil Ugurbil. UCB/AMRIT: David Feinberg

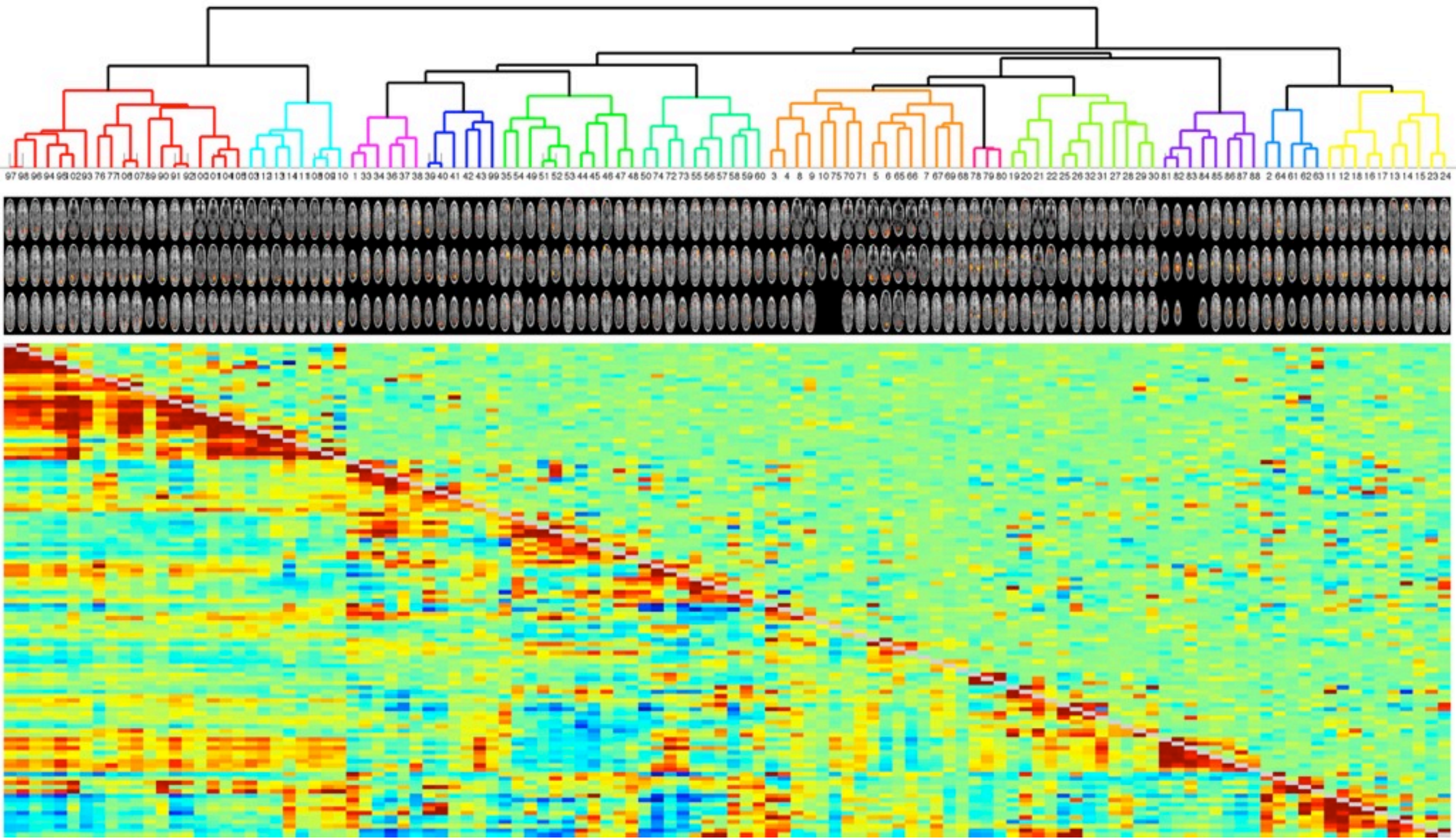


- UMinn 3T Trio, 32-channel head-coil, 3x3x3mm, whole-head
- 6 x 10-minute runs in one session
- SIR=1, MB=3, TR=0.8s (> 4000 timepoints)

- *Feinberg PLoS ONE 2010*

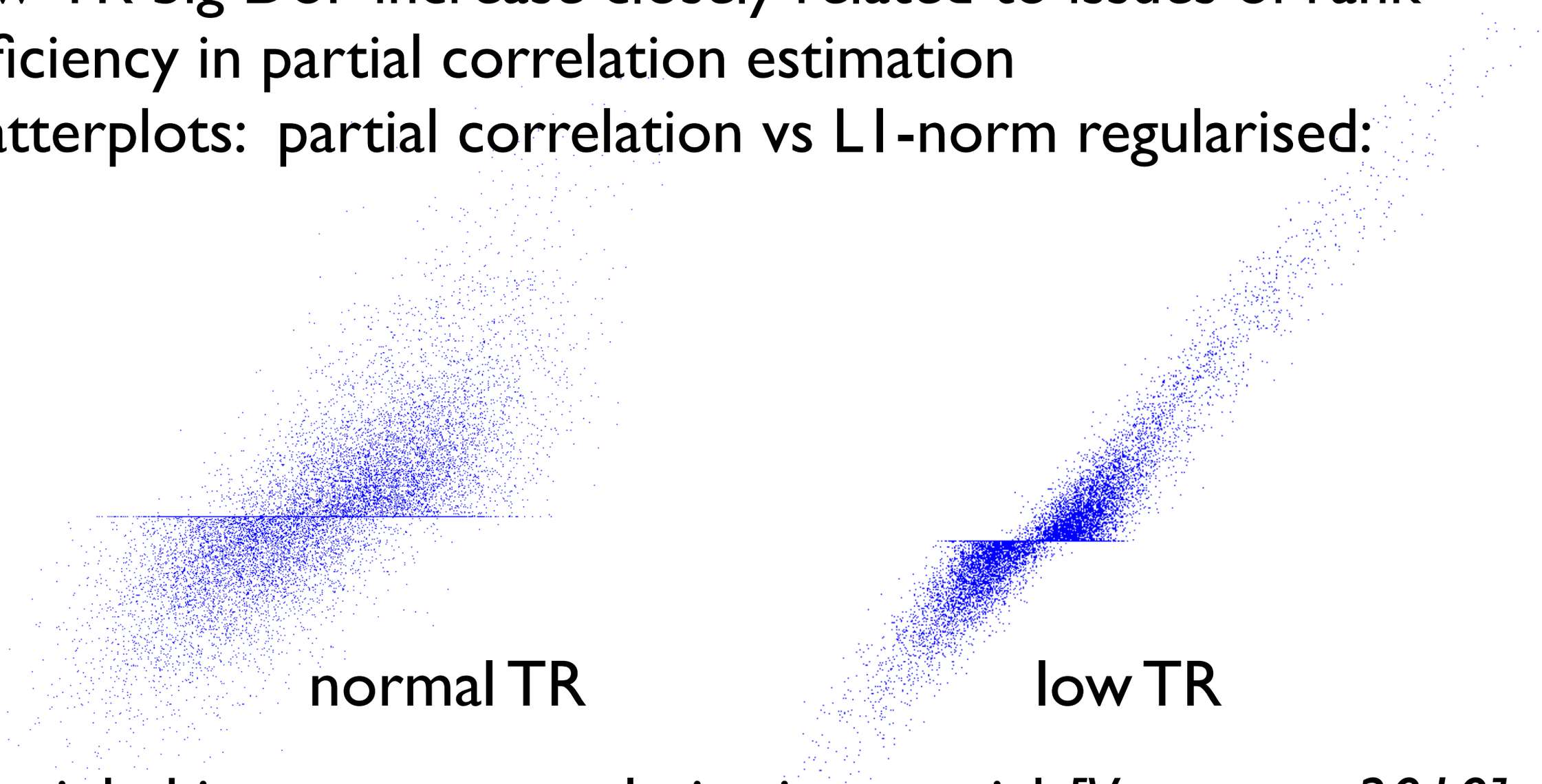
Human Connectome Project: Pushing Spatial & Temporal Resolution

3T, 3x3x3mm, TR=0.8s, 1hour scan, single subject, 110 RSN "nodes"



Temporal DoF and Rank Deficiency

- Low-TR big DoF increase closely related to issues of rank deficiency in partial correlation estimation
- Scatterplots: partial correlation vs L1-norm regularised:



normal TR low TR

- Principled inverse cov. regularisation crucial *[Varoquaux 2010]*
  - Relate regularisation to MVN likelihood (c.f. SEM)
  - Extend within-subject regularisation to also be cross-subject
- But low-TR also helps a lot

Gotcha #1
Inappropriate ROIs gives a bad network

Inappropriate node definition will blur node timeseries together

This give bad network matrix for all network modelling methods

E.g. use of structural atlases with large ROIs, such as AAL and Harvard-Oxford

Gotcha #2
Functional Connectivity is not "quantitative"

Full / partial correlation often referred to as "Functional connectivity"

These are **not quantitative measures:**

In addition to telling us about

- changes in network connection strength

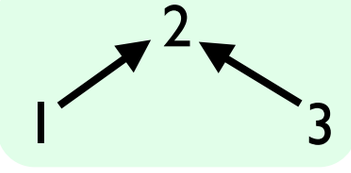
it is **also** sensitive to

- changes in noise level
- changes in input signal level
- signals passing round other parts of the network

- *Friston BC 2011*

But - BOLD FMRI isn't quantitative anyway.....so maybe not get **too** hung up on this? But be careful when interpreting group differences, etc.....

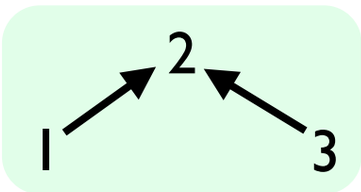
Gotcha #3
Partial correlation isn't perfect



Gotcha #3
Partial correlation isn't perfect

- Partial correlation is **not** perfect!

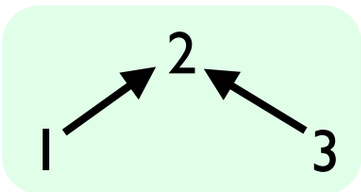
- For example, "Berkson's paradox"
  - 1 and 3 feed into 2
  - Their full correlation is zero

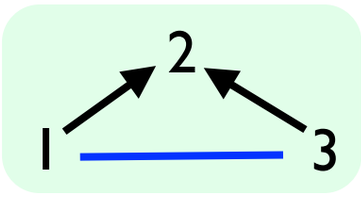


Gotcha #3
Partial correlation isn't perfect

- Partial correlation is **not** perfect!

- For example, "Berkson's paradox"
  - 1 and 3 feed into 2
  - Their full correlation is zero
  - *But* after regressing 2 out of 1 and 3.... their partial correlation is negative!!

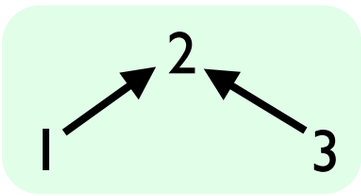


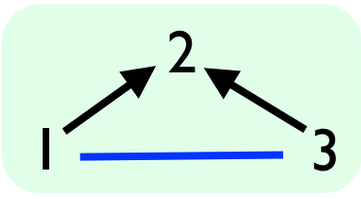


Gotcha #3
Partial correlation isn't perfect

- Partial correlation is **not** perfect!

- For example, "Berkson's paradox"
  - 1 and 3 feed into 2
  - Their full correlation is zero
  - *But* after regressing 2 out of 1 and 3....
    their partial correlation is negative!!

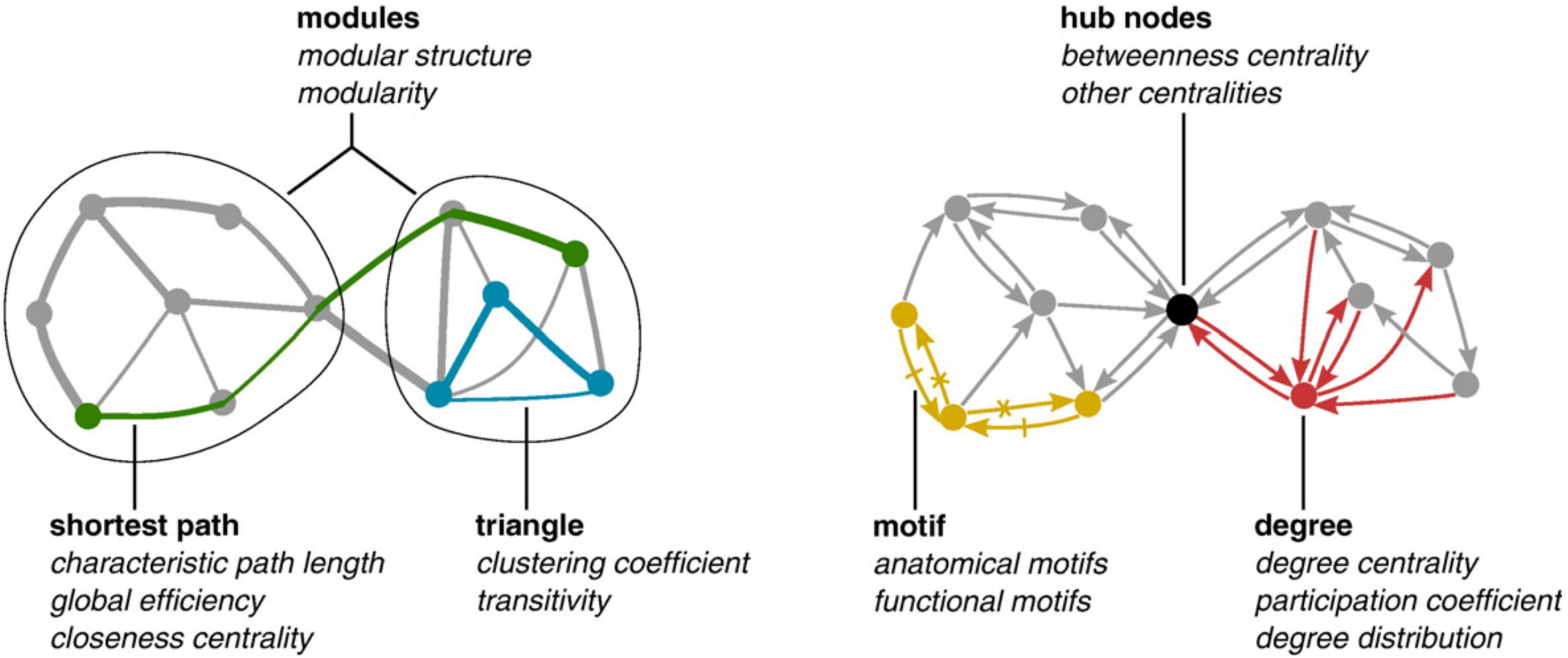




- Maybe Bayes Nets, DCM, etc. can do better!

Gotcha #4

Graph theory won't save a bad network matrix!



- *Rubinov NeuroImage 2010*

Danger ! Unless the network matrix is "correct" these measures are meaningless

Don’t fall in love with your method

Don’t fall in love with your method

it may come back to bite you!