Methods Update

Tue Herlau

Technical University of Denmark

herlau@stanford.edu

Basic MDSMC:

$$s_{t+1} = \sum_{j=1}^{J} C_j v_{jt} s_t + U u_{t+1} + \epsilon_t$$
$$y_{kt} = B_k^T \Phi \begin{bmatrix} s_{k,t} & s_{k,t-1} & \cdots , & s_{k,t-L} \end{bmatrix}^T$$

- Code available on `citron.stanford.edu`
- For $J = 1$, method seems to works better than vanilla MDS, DCM
- Focus has been on the $J > 1$ case, i.e. dynamic connectivity. A much harder problem!

Validation (consider $J = 1$ for simplicity and data from S subjects)

- The ideal way to use the method is to estimate a single C matrix from S subjects.
- We write this estimate as $C(1:S)$
- However, since a given method always produce a result, this tells us little in terms of how good the estimate actually is
- Accordingly, we consider similarity of estimates from single subjects $C(s)$, or small groups of subjects $C(S_k)$, and compare these
- Finally, C is a matrix with continuous entries, and we need a way to turn it into an actual graph (later)

Consider $J = 1$:

- Consider estimates from S subjects $C(1), \ldots, C(S)$
- "Working"can be defined as how "similar"these matrices are
- We consider similarity of order statistic as estimated by Kendalls tau:

$$K = \sum_{s_1=1}^{S} \sum_{s_2=s_1+1}^{S} K_\tau(C(s_1), C(s_2))$$

- Interpretation: "If the method is run independently on S subjects, how similar will the C-matrices be"?

Data and results in turns of Kendalls tau:

Name	Timepoints	Nr.ROIs	Subjects	TR	Regions
HCP	405	5	122	0.72	lAI, rAI, rMFG, lFEF, rFEF
Yale	295	3	59	2	rACC, rAI, rDLPFC, rIFG, rPPC
StopGo	756	3	14	0.49	lM1, lPut, lSMA, lVis, rVis
Opto-rats	480	3	3	0.75	M1, Thalamus, Insula
Oddball	200	5	15	2	rAI, rdACC, rVLPFC, rPPC, rDLPFC

Dataset	Maximum (debug)			i,j fixed			Bootstrap		
	DCM	MDS	MDS-MC	DCM	MDS	MDS-MC	DCM	MDS	MDS-MC
oddball ($M=3$)	0.11	0.21	**0.32**	0.11	0.21	**0.32**	0.11	0.21	**0.22**
oddball ($M=5$)	0.12	0.14	**0.23**	0.12	0.14	**0.23**	0.12	0.14	**0.23**
Yale ($M=3$)	0.21	0.16	**0.27**	0.21	0.16	**0.27**	0.21	0.16	**0.26**
Yale ($M=5$)	**0.17**	0.091	0.13	**0.17**	0.091	0.11	**0.17**	0.091	0.092
stopgo ($M=3$)	0.18	0.15	**0.27**	0.18	0.15	**0.27**	0.18	0.15	**0.26**
stopgo ($M=5$)	0.2	0.07	**0.4**	0.2	0.07	**0.38**	0.2	0.07	**0.38**
HCP ($M=3$)	0.1	0.16	**0.36**	0.1	0.16	**0.36**	0.1	0.16	**0.24**
HCP ($M=5$)	0.028	0.081	**0.14**	0.028	0.081	**0.13**	0.028	0.081	**0.11**

Dynamic connectivity

- Estimating dynamic connectivity is J times harder than static
- With the added problem that for $C_{j>1}$ we only have few time points (when condition is active)
- We will consider a $J = 3$ study using data from

ORIGINAL ARTICLE

Causal Interactions Within a Frontal-Cingulate-Parietal Network During Cognitive Control: Convergent Evidence from a Multisite–Multitask Investigation

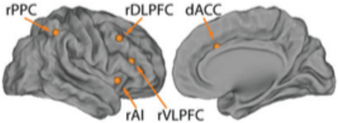
Weidong Cai[1], Tianwen Chen[1], Srikanth Ryali[1], John Kochalka[1], Chiang-Shan R. Li[4], and Vinod Menon[1,2,3]

Data will consist of the following three datasets, each having two

Table 1 Task 1: Task information

Task	Number of subjects	Task design	TR (s)
SST1[a]	19	fast-jitter (average ITI = 1 s)	2
Flanker[b]	25	slow-jitter (average ITI = 12 s)	2
SST2[c]	59	jittered fixation = 1–5 s ITI = 2 s	2

[a] Xue et al. 2008.
[b] Kelly et al. 2008.
[c] Zhang and Li 2012.



sessions: (I will include HPC)

Observations:

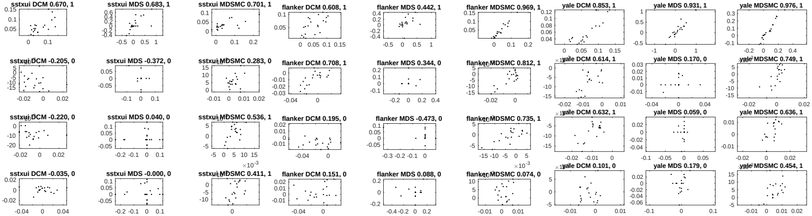
- Estimated C-matrices are very noisy; subtle things like pre-processing can bias estimates giving false appearance of positive results. Typical failure mode is not $C = 0$ but "something"
- I block subjects into groups of 5 and estimate a C matrix for each such block $C_j(S_k)$, $k = 1, \dots, \frac{S}{5}$
- I will focus on replicability in two ways:
- (i) Can C_j be estimated above chance level based on this data. I.e. does the method find anything
- (ii) Is the resulting estimate of C_j (averaged across block) consistently found between sessions?
- (iii) Do the networks look reasonable?
- (iiii) When C_j is binarized to actual functional networks, are results consistent between sessions?
- This analyzes should be done both for C matrices and their binarized counterparts

(i) Consistency results both within sessions and between as measured by Kendalls tau:

Dataset	$C_{J=1}$			$C_{J=2}$			$C_{J=3}$		
	ses_1	ses_2	ses_{12}	ses_1	ses_2	ses_{12}	ses_1	ses_2	ses_{12}
yale, DCM	0.21	0.17	0.2	0.05	0.04	0.07	0.01	0.04	0.05
yale, MDS	0.3	0.37	0.33	0.05	0.0	0.04	-0.01	-0.01	-0.0
yale, MDSMC	**0.37**	**0.59**	**0.48**	0.04	**0.14**	0.09	0.05	**0.14**	**0.12**
sstxui, DCM	0.21	0.05	0.2	0.12	-0.04	0.08	0.22	0.1	0.11
sstxui, MDS	0.25	**0.43**	0.26	0.12	0.0	0.03	0.1	0.11	-0.05
sstxui, MDSMC	0.41	0.26	0.36	0.02	-0.01	0.05	0.03	-0.06	0.03
flanker, DCM	0.03	-0.06	0.1	0.04	0.04	0.12	0.07	-0.0	-0.02
flanker, MDS	-0.03	-0.05	0.07	0.09	-0.04	0.05	-0.02	0.06	-0.02
flanker, MDSMC	**0.42**	**0.28**	**0.39**	0.21	**0.26**	**0.24**	**0.28**	0.09	**0.18**

"When C_j is estimated from blocks of 5 subjects, MDS-MC tends to find more consistent results and these results are more often (but not always) consistent across sessions"

(ii) Similarity between sessions. Suppose we compute estimate $C_j = \frac{5}{S} C_j(S_k)$. We can then plot values of C_j for each session as a scatter plot:



A way to sumarize this is as a table of which correlations are significant:

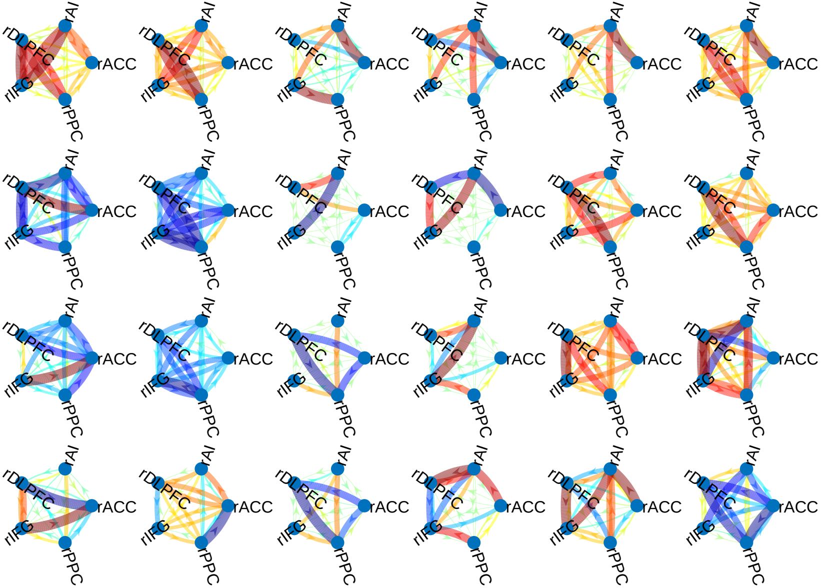
Dataset	$C_{J=1}$	$C_{J=2}$	$C_{J=3}$	$C_{J=3} - C_{J=2}$
sstxui DCM	**0.001**	0.807	0.825	0.559
sstxui MDS	**0.0**	0.947	0.433	0.501
sstxui MDSMC	**0.0**	0.114	**0.007**	**0.036**
flanker DCM	**0.002**	**0.0**	0.205	0.262
flanker MDS	**0.025**	0.069	0.982	0.356
flanker MDSMC	**0.0**	**0.0**	**0.0**	0.379
yale DCM	**0.0**	**0.002**	**0.001**	0.336
yale MDS	**0.0**	0.236	0.402	0.225
yale MDSMC	**0.0**	**0.0**	**0.001**	**0.022**

"There is (?) a tendency towards higher correlation between C matrices between sessions for MDS-MC"

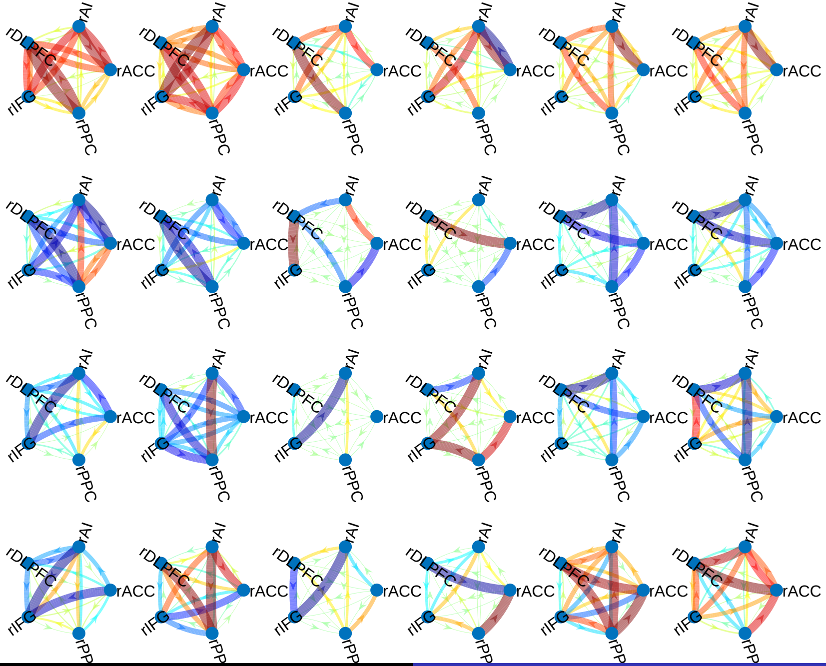
(iii) Visual evaluation of C matrices (brace yourself!)

- I am just going to plot the "raw" C-matrices (averaged) to begin with. Recall there are 3 methods (DCM, MDS, MDS-MC) and two sessions giving 6 columns and $J = 3$ rows plus a 4'th row corresponding to $C_3 - C_2$.

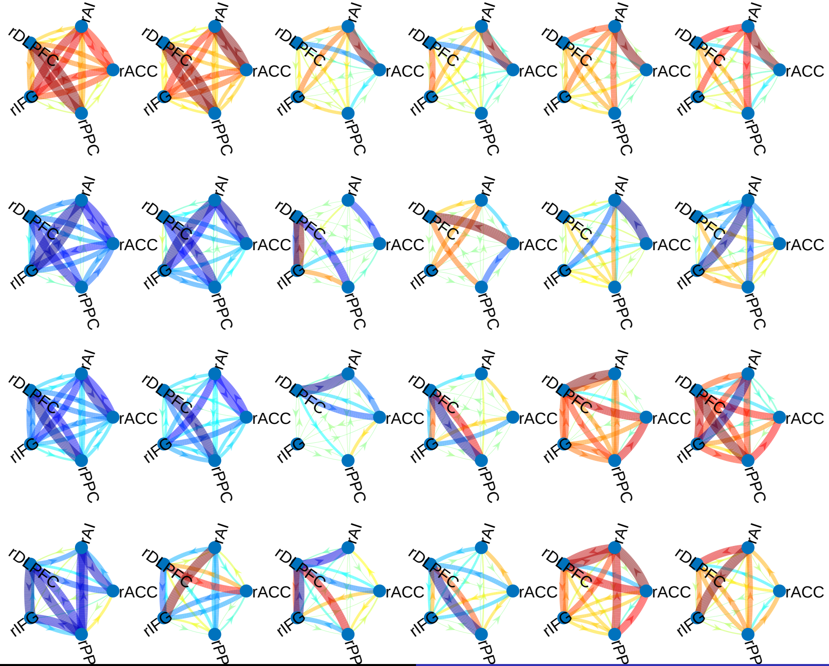
sstxui: (Each pair of columns are scan sessions and should be the same)



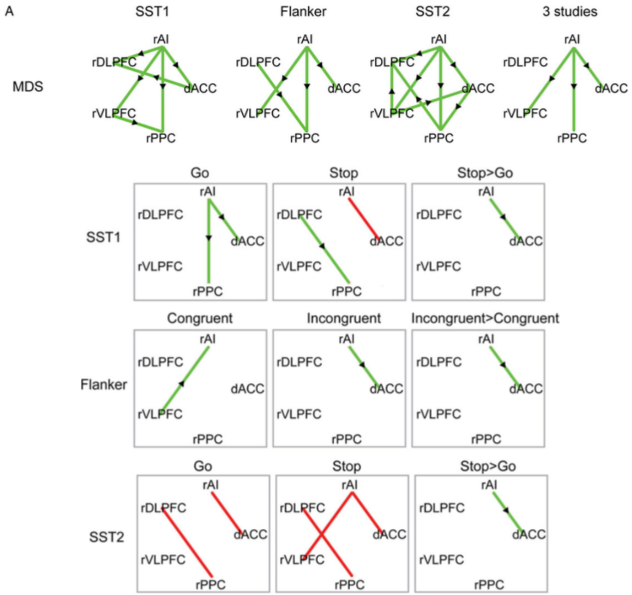
flanker: (Each pair of columns are scan sessions and should be the same)



yale: (Each pair of columns are scan sessions and should be the same)



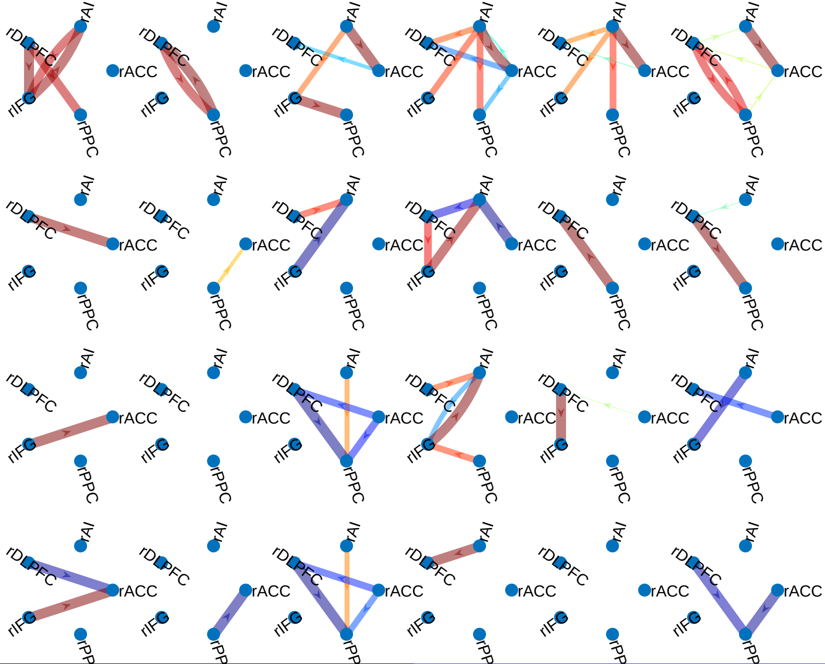
From original reference:



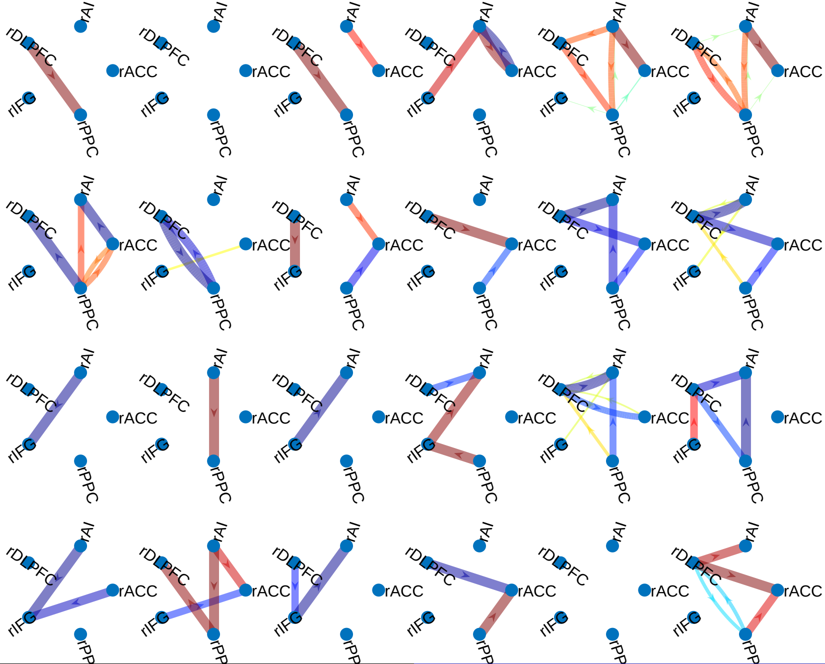
(iiii) Produce binarized graphs and compare

- MANY ways to binarize these graphs and they all produce different results, with some having no links and some having far too many links.
- The big choice is to use sampler (or VB) estimate of edge variance. I have chosen not to in order of being able to use same binarizing method for any method that produce a C matrix.
- A choice which appear to work is for fixed j, take each $C_j(S_k)$, randomly permute elements and use these to obtain estimate of "null"distribution of the means $\hat{C} = \frac{5}{S} C_j(S_k)$, then plot links in top/bottom 5th percentile

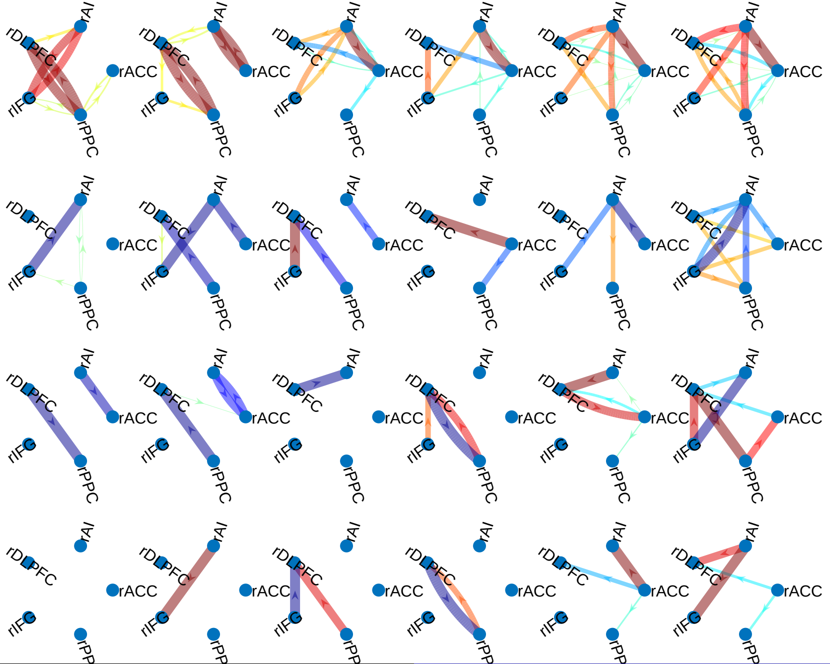
sstxui: (Each pair of columns are scan sessions and should be the same)



flanker: (Each pair of columns are scan sessions and should be the same)



yale: (Each pair of columns are scan sessions and should be the same)



- Quantitative evaluation of these graphs are hard because we are comparing binary graphs and have very little data to work with.
- I will consider each graph to be binary, and simply ask what the chance is to get this at least this much overlap of edges by chance, given the number of edges.