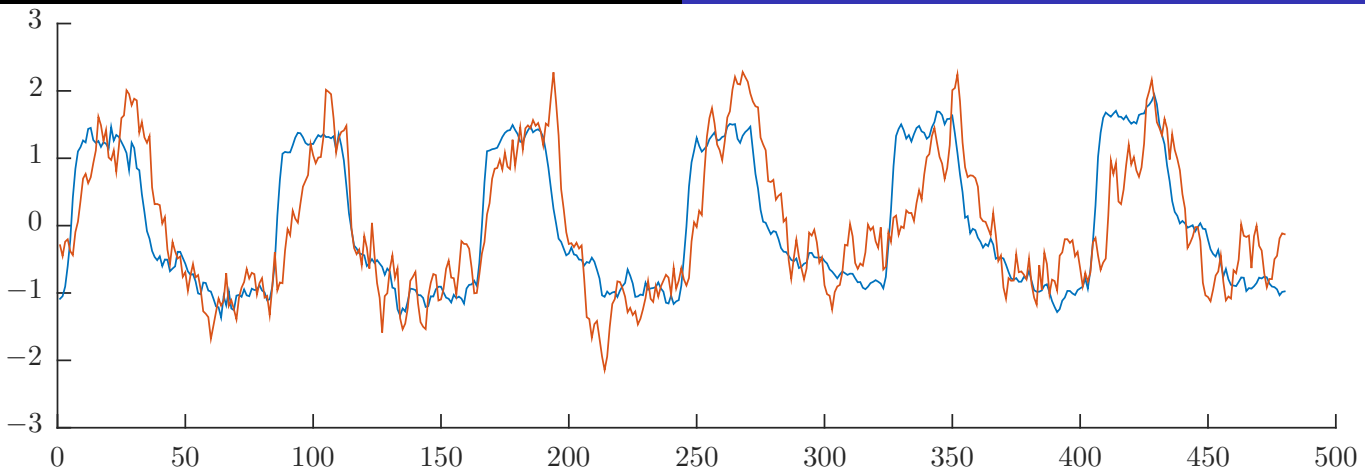
April 2017 update

Tue Herlau

Technical University of Denmark

tuhe@dtu.dk



- In fMRI we observe a signal $y_k(t)$ in region k at time t.
- The observed signal y_k is a convolution of a *true signal* s_k

$$y_k(t) = h^T \begin{bmatrix} s_k(t) & s_k(t-1) & \cdots, & s_k(t-L+1) \end{bmatrix}^T + \text{noise}$$

- The underlying signal is assumed to evolve as a first-order stochastic differential equation with Wiener noise

$$ds_k(t) = Cs_k(t)dt + dW(dt)$$

- Or in discretized form, allowing C to depend on time and introducing stimuli $u_k(t)$

$$s(t+1) = (I + dtC_t))s(t) + u(t) + \sqrt{D}\epsilon, \ \epsilon \sim \mathcal{N}(0,1)$$

$$s(t+1) = C_t s_t + U(t) + \sqrt{D_w}\epsilon_w$$
$$y_k(t) = h^T \begin{bmatrix} s_k(t) \\ \vdots \\ s_k(t-L+1) \end{bmatrix} + \sqrt{D_e}\epsilon_e$$

- It is furthermore assumed that $C_t = \sum_{j=1}^{J} v_{jt} C_j$.
- Basic idea: Put priors on above (unknown) quantities such as C, D_w, D_e, etc. etc. and compute(/estimate) density of C:

$$p(C|y) \propto \int dsdUd \cdots D_w p(y, C, s, \cdots)$$

Using either Markov-Chain Monte-Carlo (MCMC) sampling or Variational Bayes (standard MDS, also quite similar to DCM).

Three categories of headache

Complicated models The more complicated the model, the harder the parameters are to find

$$\text{(Simple, } \theta \text{ easy)} \quad x_i = \theta + \epsilon_i$$
$$\text{(Complicated, } \theta \text{ not so easy)} \quad x_i = f(w_i, \theta) + \epsilon_i$$

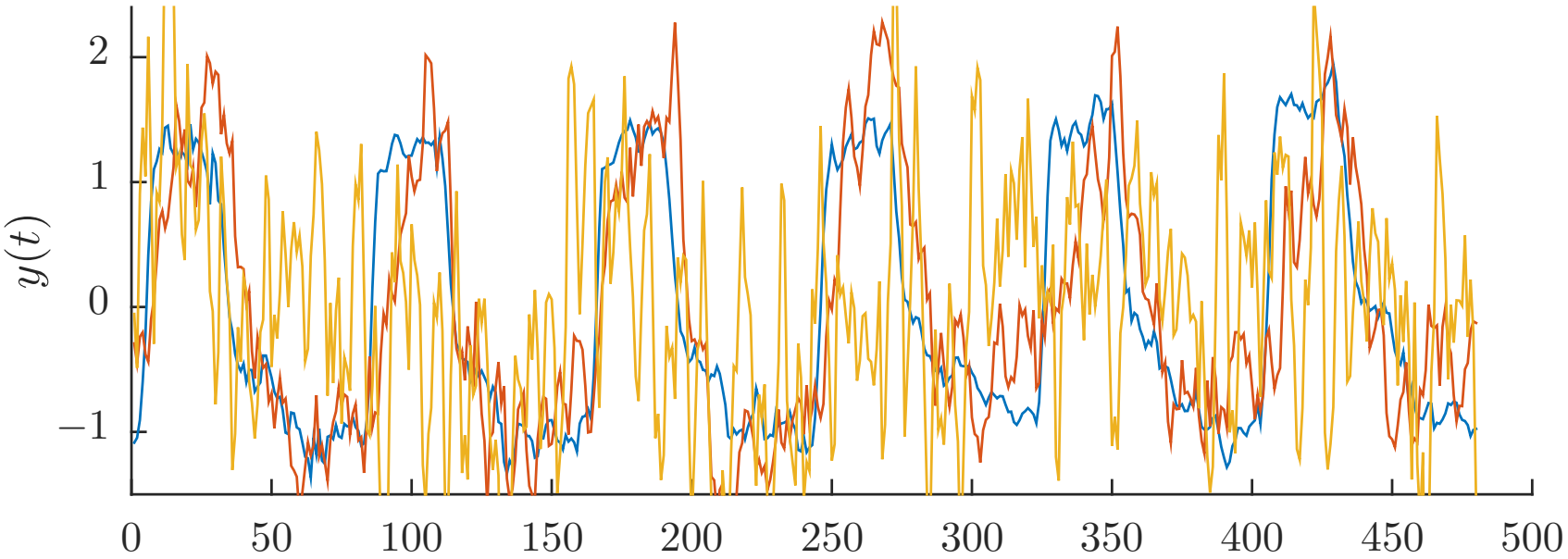
this is typically handled using regularization

Model misspecified When we push data into the wrong model class (for instance violating Gaussian noise assumptions, stationarity of noise) we get funky behaviour

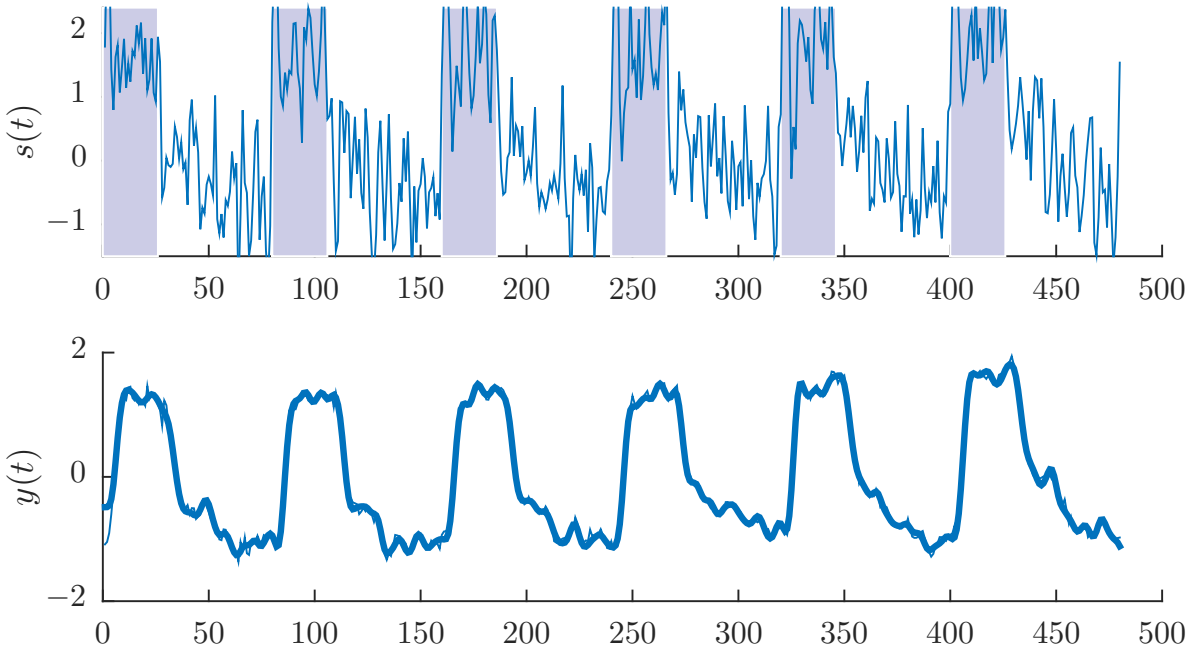
Inference problems Both MCMC and VB can get stuck in local minima if the likelihood function is not convex (it isn't). This problem is exacerbated when adding more regions (and possibly, more data).

- Model too complicated? Add regularization (model now misspecified, likelihood more likely to be non-convex making inference harder)
- Model misspecified? Either add additional "stuff"to model (model more complicated and inference harder), or restrict parameter ranges away from defective behaviour
- Inference problems? Cut down on number of parameters, acceptance

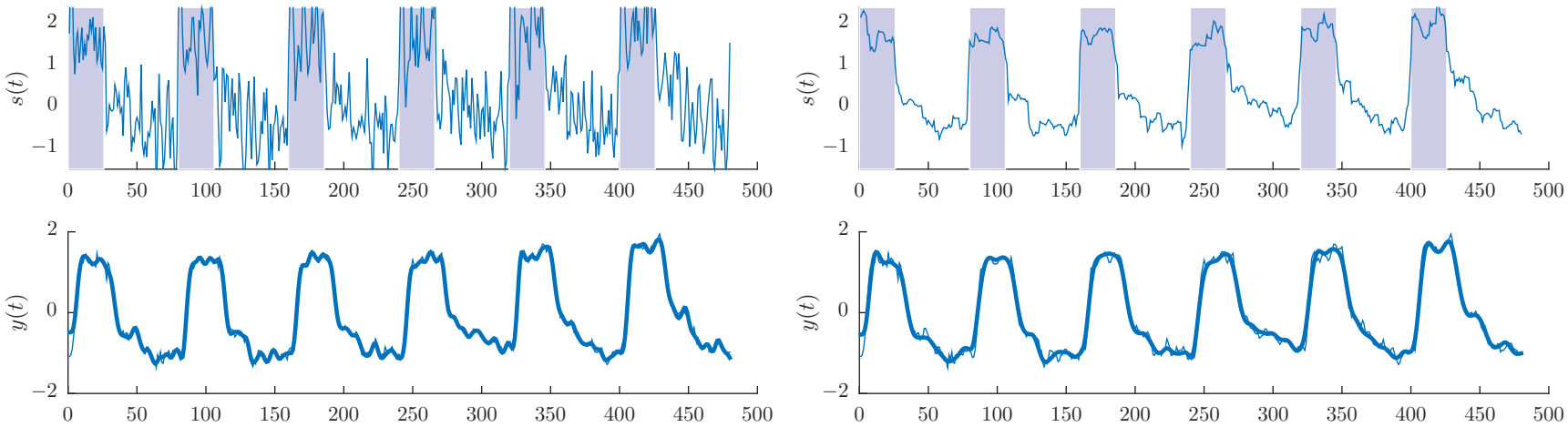
Case study: Rat dataset. 3 Rats, 3 ROI, ROI 1 (blue) stimulated in a block pattern and should influence ROI 2 (red) but not ROI 3 (yellow)



Focus on $s(t)$, $y(t)$ for 1 ROI.



- Underlying neural signal $s(t)$ has very high variance
- Convolution ensures estimate $\hat{y}(t)$ match observed $y(t)$ well
- Variance of $s(t)$ is $\propto \frac{\sigma_w}{dt^2 C^2 - 2dtC}$
- Regularization (informative prior) on $\sigma_w = \sqrt{D_w}$ have little effect since signal must fluctate at certain time points
- Matters because $Cs(t) = s(t+1) - s(t) + \text{noise}$

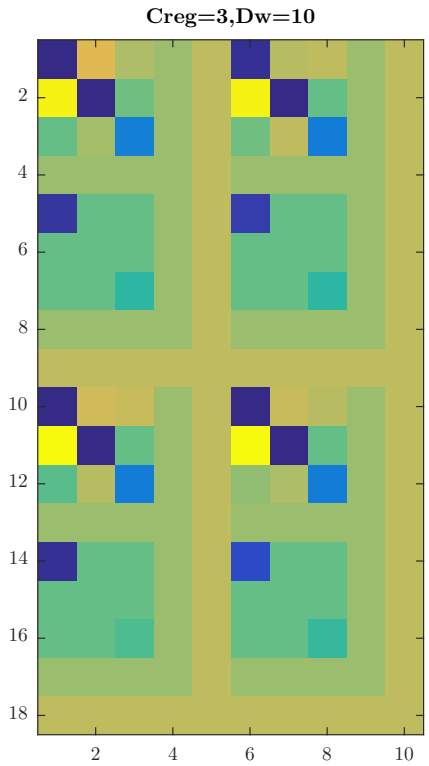


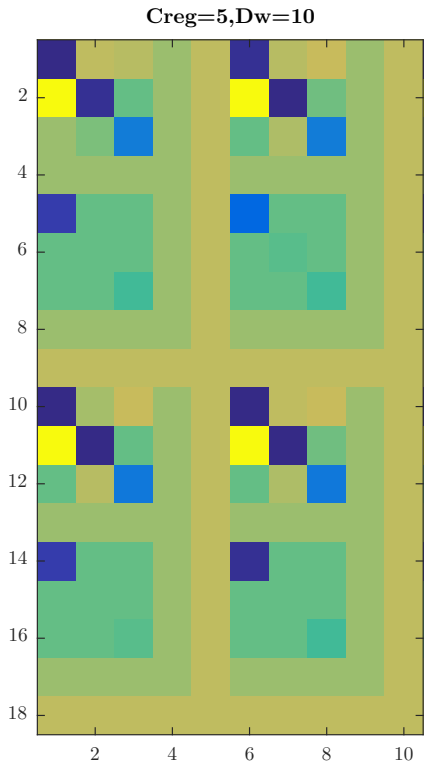
Fixes

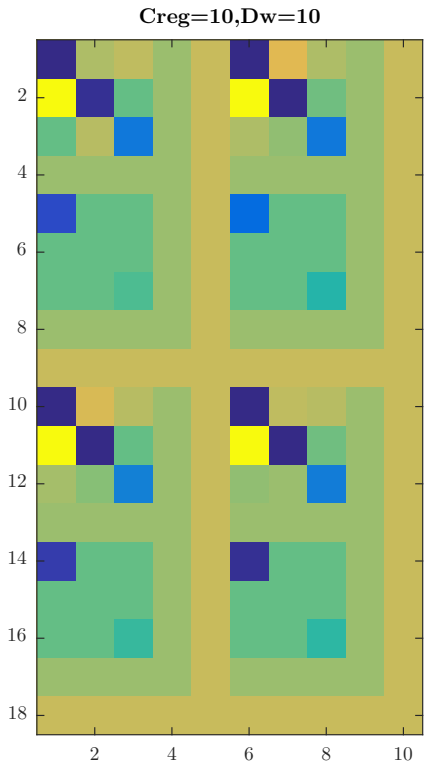
- Assume noise is temporal $\sigma_w \rightarrow \sigma_{wt}$
- Add constraints to C to ensure it is not 0 (recall variance of $s(t)$ $\propto \frac{\sigma_w}{dt^2 C^2 - 2dtC}$. (in general: implement any interval constraint on C_{ij})
- Above is shown effect of adding temporal variance σ_{wt}; same priors used.

Convergence of MDS-VB on larger datasets

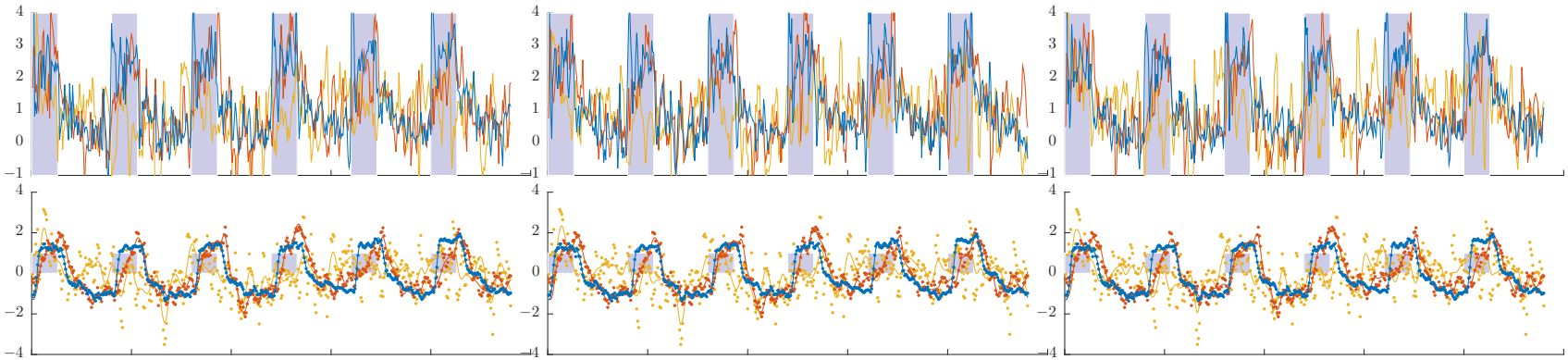
- Assume C-matrices C_1, C_2 corresponding stimuli condition on/off
- Vary hyper(parameters) for σ_{wt} (shown) and prior for C_{ij}
- Each subplot shows different hyperparameter $\sigma_C \sim \mathrm{Gam}(1, \beta_C^{-1})$
- each of the 4 panes show C_1, C_2 stacked (4 random restarts)





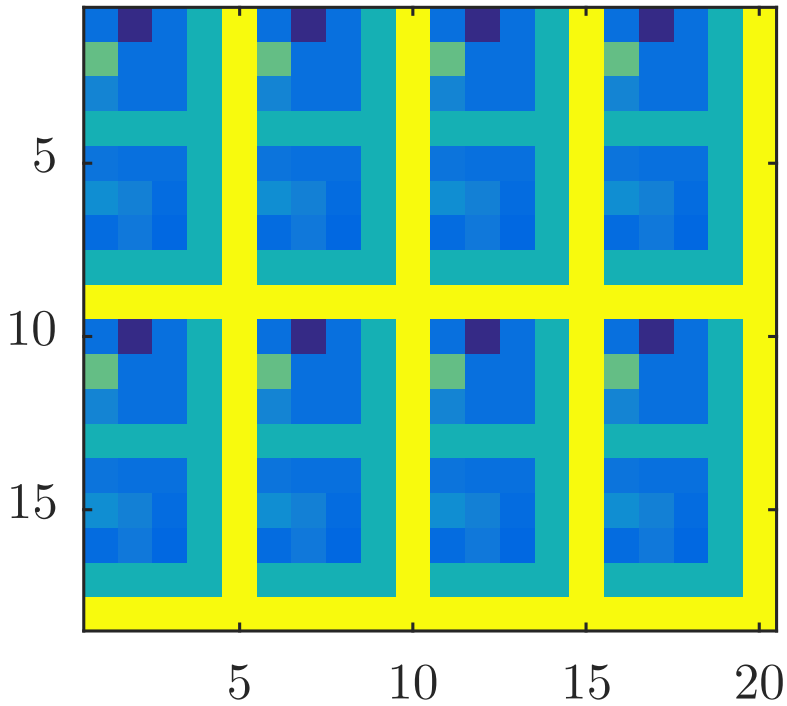


Same but showing $s(t)$, and $y(t)$

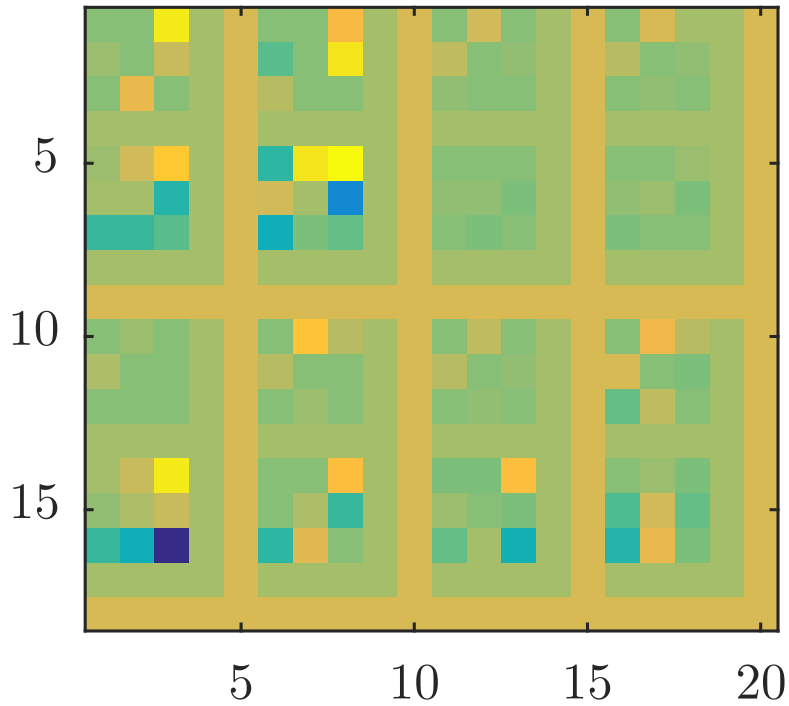


- Application of MDS-VB to rat dataset.
- For each subplot, each of the 8 panes represent the two C-matrices stacked on top of each other.
- Left without random initialization, right with random initialization.

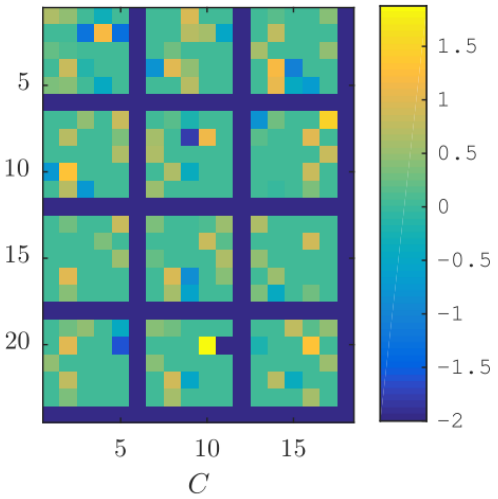
Default Initialization

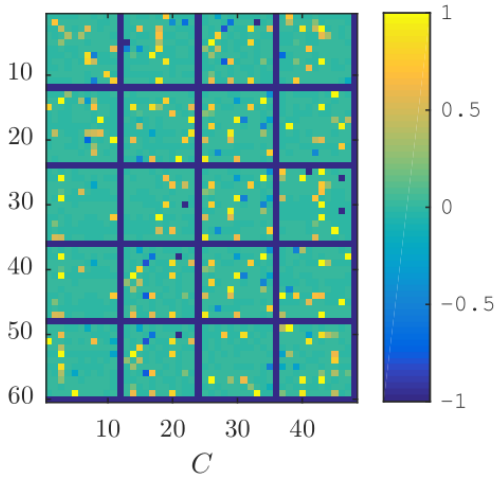


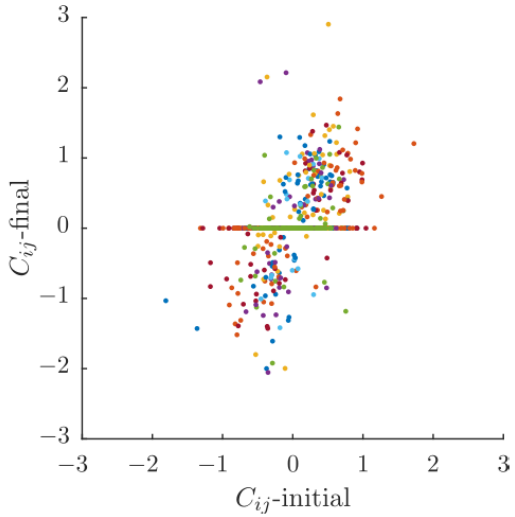
Random Initialization

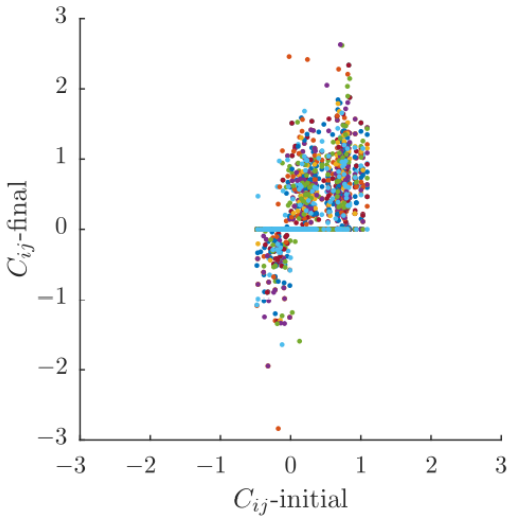


Results on Open-fMRI and HW datasets





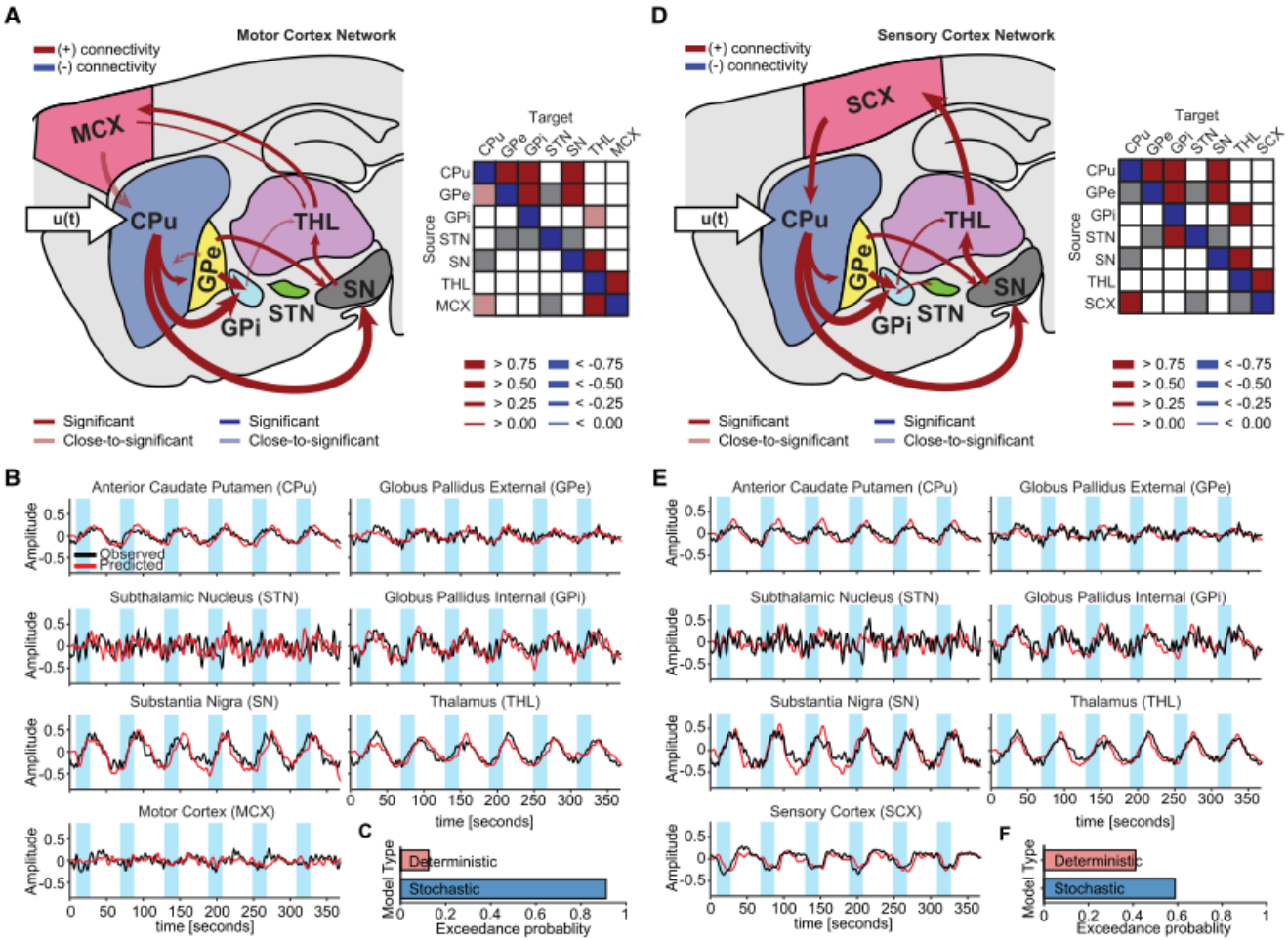




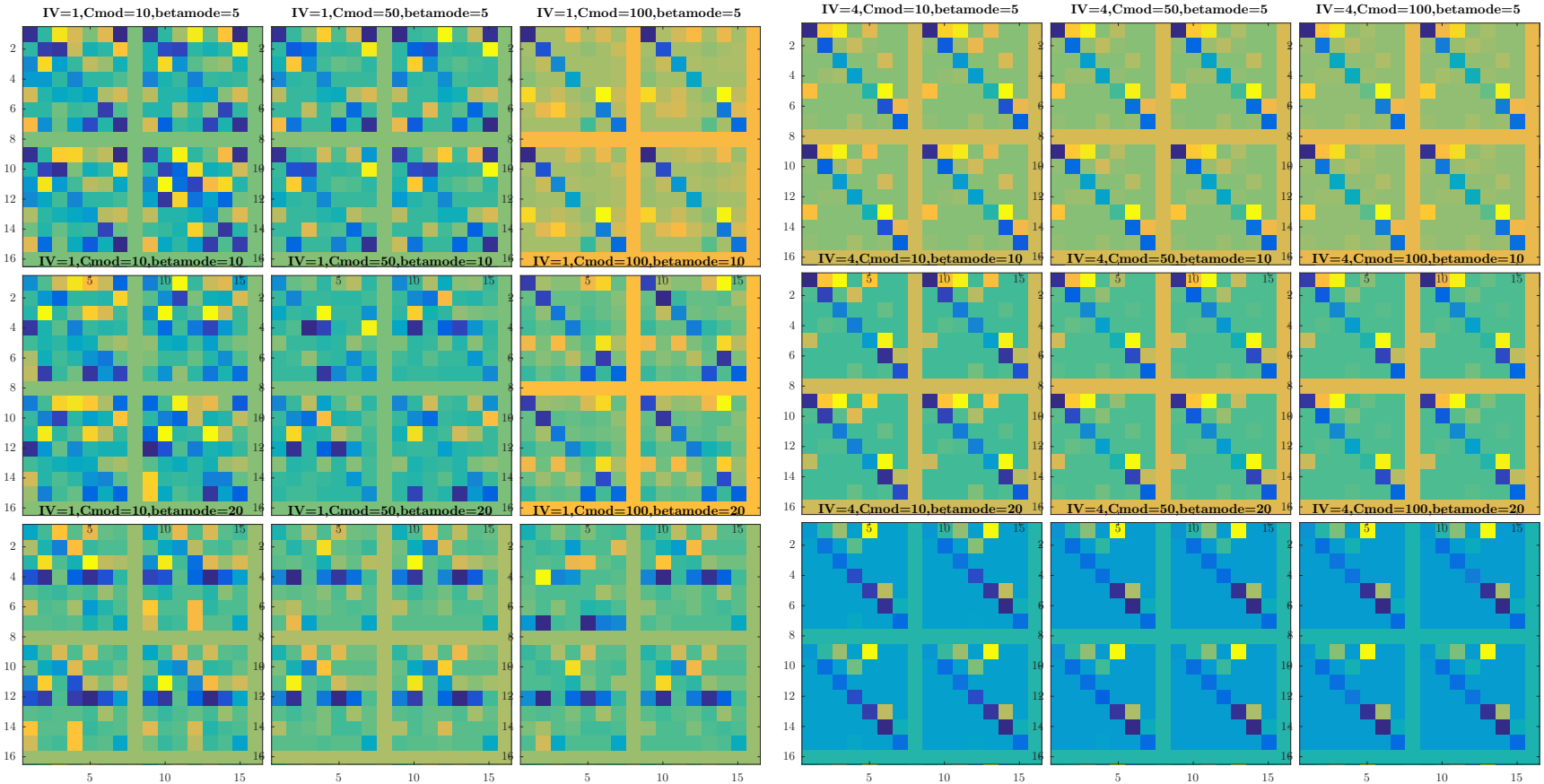
Recall the MMC dataset from the journal club a few weeks back

Studying Brain Circuit Function with Dynamic Causal Modeling for Optogenetic fMRI

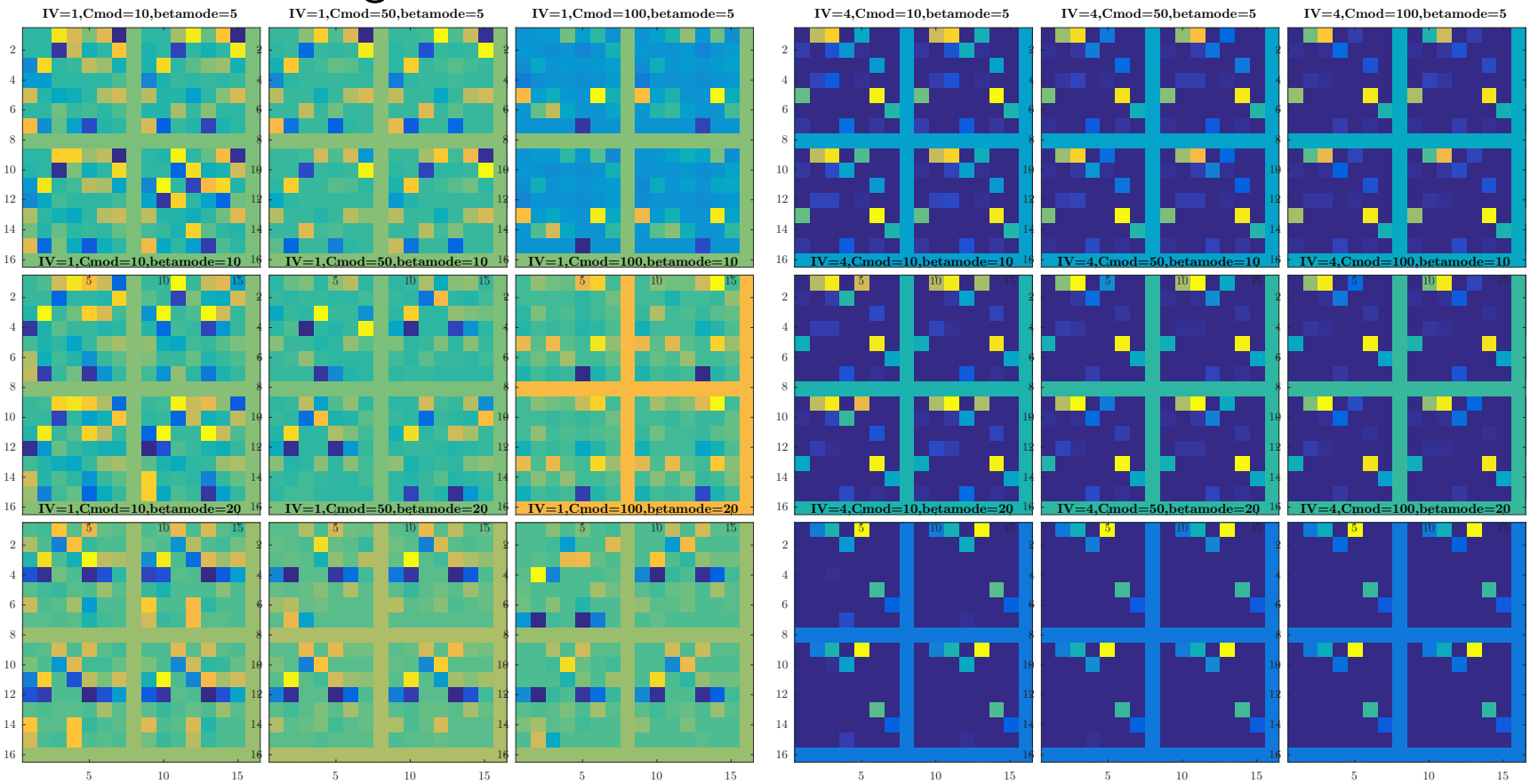
David Bernal-Casas,[1] Hyun Joo Lee,[1] Andrew J. Weitz,[1,2] and Jin Hyung Lee[1,2,3,4,5,*]



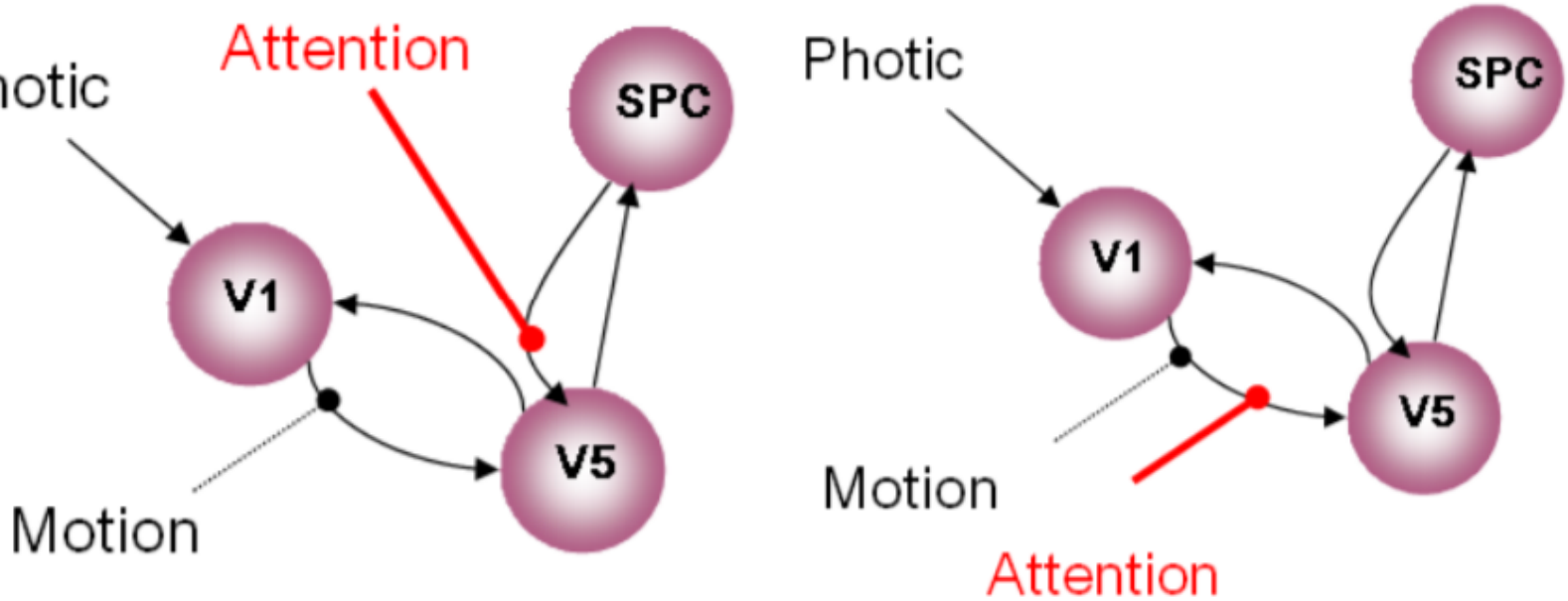
Applying MDS-MCMC using no structural assumptions (left) and using sparsity pattern $C_{ij} = 0$ as in paper (right):



Same without diagonal

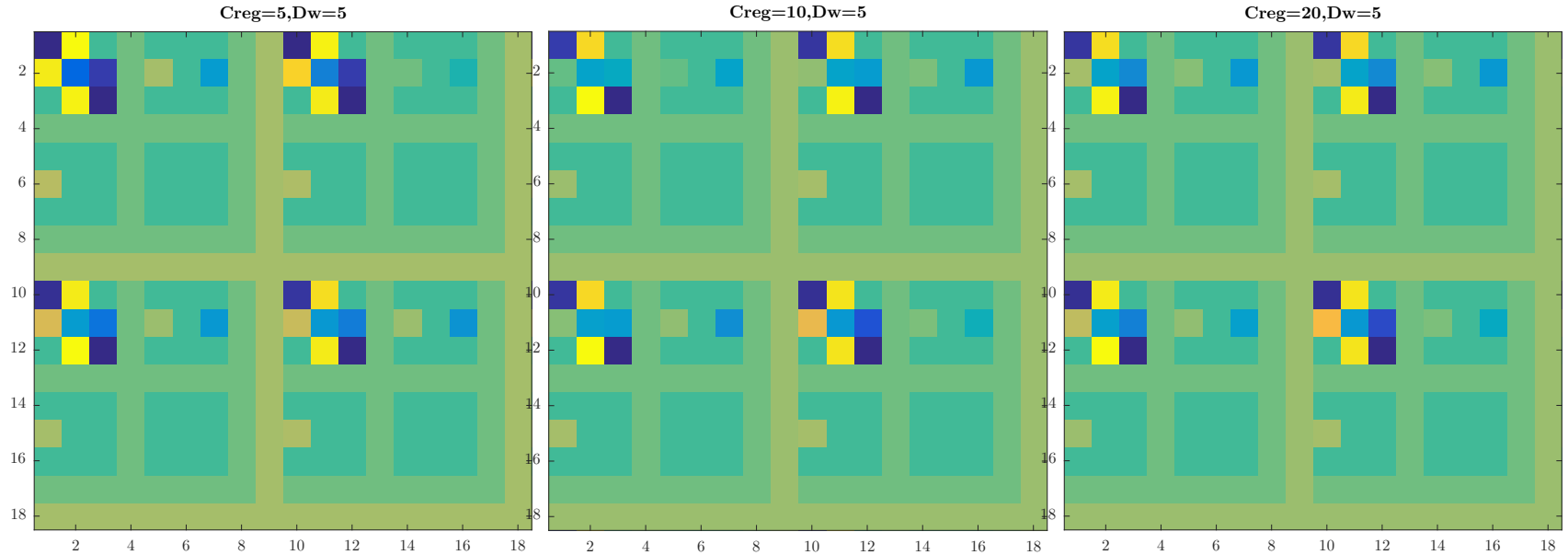


The attention dataset is the example included in the SPM manual. It consists of a simple 3 ROI test where the subject is shown different types of stimuli. Goal is to determine which of these models are better (SPM prefers model 2)



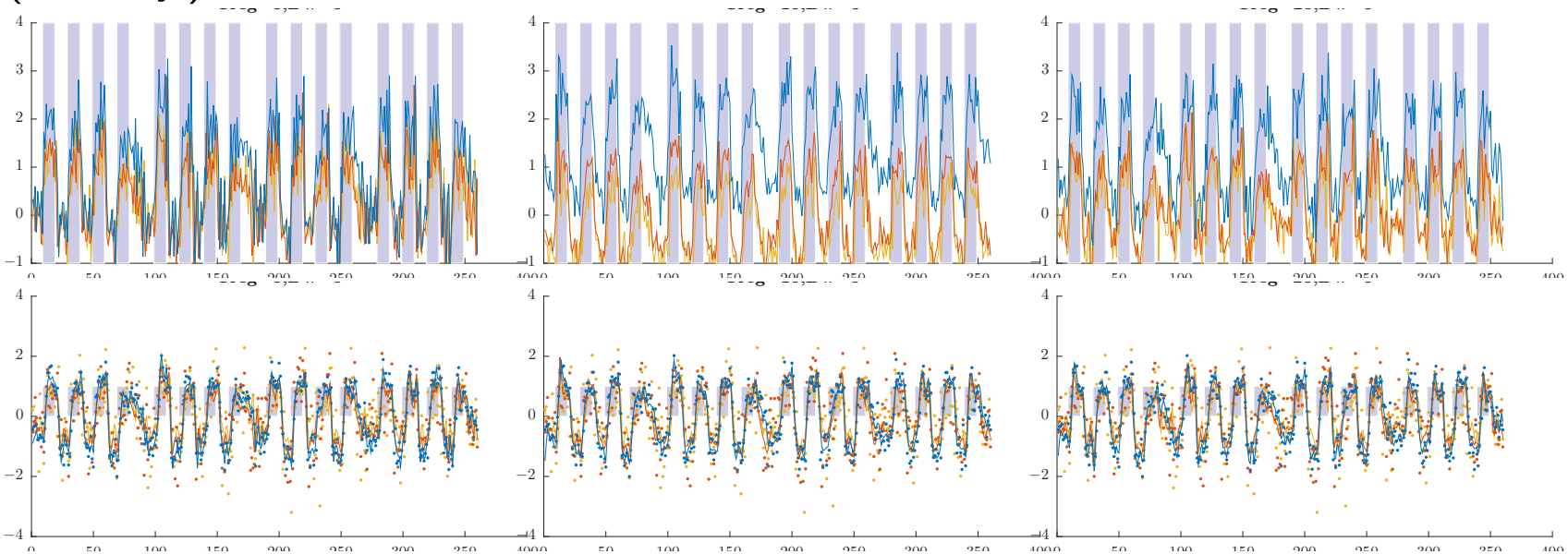
V1 : ROI 1, V5: ROI 2, SPC: ROI 3

Results (implement stimuli patterns, different C-matrices)



Panes corresponds to different hyper parameters; each pane represents 4 simulations, each simulation shows the 3 interaction matrices C_1, C_2, C_3. Upper-right is the matrix corresponding to the attention-modulation.

(s and y)



- MDS-VB greatly depends on initialization. This has implications on how MDS-VB results are interpreted: A $p < 0.05$ result can just as well mean that the initialization was non-zero with $p < 0.05$....
- Should we even accept that results are a consequence of initialization and not the model? (danger of story-telling, *"ROI 1 influence ROI 2 assuming you initialize your model as thus..?"*)
- MDS-MCMC also has severe difficulties forgetting initialization
- These problems are very unlikely to go away with more data
- Two partial solutions: 1) Add "plausible"limitations to model parameters, i.e. $C_{ii} < c_0 < 0$, etc. obtained by considering bounds on variance and autocorrelation time. 2) add equality constraints similar to $C_{ij} = 0$.

- Using DCM-type sparsity restrictions & other tweeks makes inference feasible on more problems (rats ok; attention/MMC: not crazy at least). Slow but steady progress
- No evidence we can blindly apply a MDS-type model to datasets of 5 or more ROI and automatically find something valid. Is there any evidence this is possible?
- Path-to-publication:
- (A) Lower the bar by documenting if these problems also occur in standard DCM (if not, why?). Present MDS-MCMC as a solution to these problems.
- (B) Automate constraining of model by finding $C_{ij} = 0$ constraints using independence testing and ideas from classical approaches to Causal discovery
- (C) Any publication must be build around comparisons against DCM
- (D) Feature-complete in comparison with DCM (I think!). Will be **very** happy to re-analyze studies where DCM has been used.