Modelling Time-varying effective Brain Connectivity using Multiregression Dynamic Models

Thomas Nichols

University of Warwick

Dynamic Linear Model

$$Y_t = \beta_{0t} + \beta_{1t}x_{1t} + \cdots + \beta_{pt}x_{pt} + v_t \qquad v_t \sim \mathcal{N}(0, V_t)$$

$$\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} + \mathbf{w}_t \qquad \mathbf{w}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{W}_t)$$

- Bayesian time series model
  - Predictors $\{X_1,\ldots,X_p\}$
    - 'Exogenous' input variables, or
    - Lagged versions of Y, generalizing ARIMA models
  - Regression coefficients $\boldsymbol{\beta}_t$
  - Observation variance V_t
  - System (co)variance $\mathbf{W}_t$ (p × p)
  - All time varying!

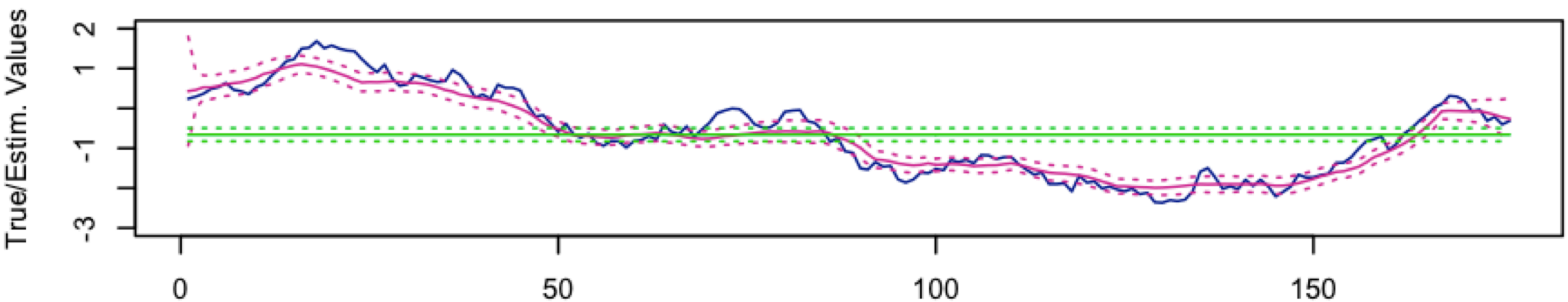
West, M., & Harrison, J. (1997). Bayesian forecasting and dynamic models. Springer-Verlag.

Dynamic Linear Model

$$Y_t = \beta_{0t} + \beta_{1t}x_{1t} + \cdots + \beta_{pt}x_{pt} + v_t \qquad v_t \sim \mathcal{N}(0, V_t)$$

$$\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} + \mathbf{w}_t \qquad \mathbf{w}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{W}_t)$$

- Example
  - "Local level model", X_1 = **1** (compared to static fit)



- Wait… this is totally over-determined model!!!
- How can this possibly work!?

West, M., & Harrison, J. (1997). Bayesian forecasting and dynamic models. Springer-Verlag.

Dynamic Linear Model

$$Y_t = \beta_{0t} + \beta_{1t}x_{1t} + \cdots + \beta_{pt}x_{pt} + v_t \qquad v_t \sim \mathcal{N}(0, V_t)$$

$$\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} + \mathbf{w}_t \qquad \mathbf{w}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{W}_t)$$

- Special sauce: Variance discounting
  - System variance fixed fraction of observation variance

$$\mathbf{W}_t = \mathrm{Var}(\boldsymbol{\beta}_{t-1}|Y_{t-1}) \ \frac{1-\delta}{\delta} \ V_t$$

Posterior variance of $\boldsymbol{\beta}$ at time t-1, i.e. the prior variance for $\boldsymbol{\beta}$ at time t

Effect of "Discount factor" δ

Observation variance

  - δ = 0 Static model, δ = ½ Random walk
  - 1-δ is loss of information at time t
    - e.g. for δ=0.95, 5% loss of information from t-1 to t

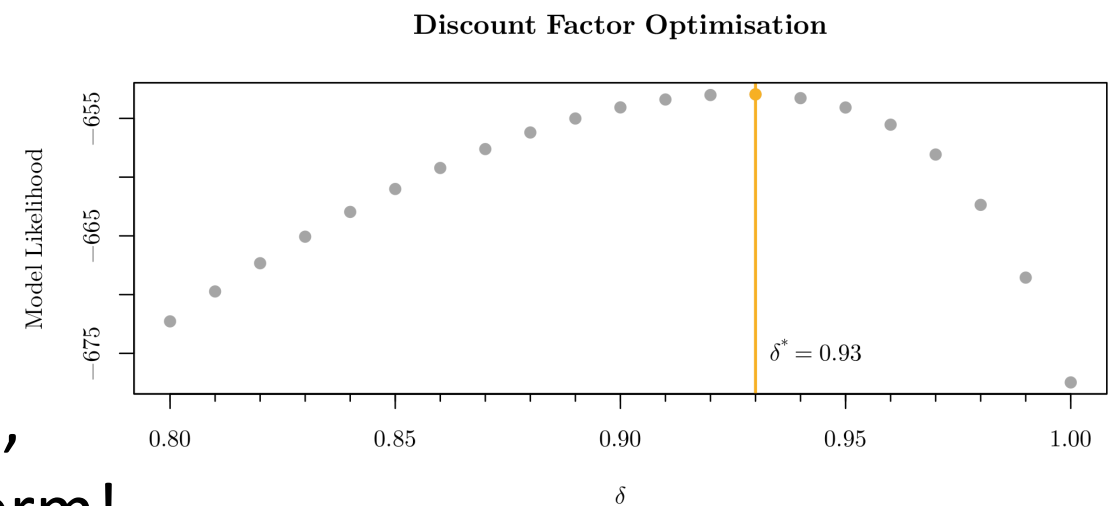
West, M., & Harrison, J. (1997). Bayesian forecasting and dynamic models. Springer-Verlag.

Dynamic Linear Model

$$Y_t = \beta_{0t} + \beta_{1t}x_{1t} + \cdots + \beta_{pt}x_{pt} + v_t \qquad v_t \sim \mathcal{N}\left(0, V_t\right)$$

$$\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} + \mathbf{w}_t \qquad \mathbf{w}_t \sim \mathcal{N}\left(\mathbf{0}, \mathbf{W}_t\right)$$

- Special sauce: Variance discounting
  - Estimate δ by maximum likelihood
  - Then model entirely conjugate!
  - Posterior, model evidence, etc, all closed form!



West, M., & Harrison, J. (1997). Bayesian forecasting and dynamic models. Springer-Verlag.

Multiregression Linear Model

Observation equations - for ROI r

$$Y_t(r) = \mathbf{F}_t(r)'\boldsymbol{\theta}_t(r) + v_t(r), \qquad v_t(r) \sim \mathcal{N}(0, V_t(r));$$

System equation - for p-dimensional θ

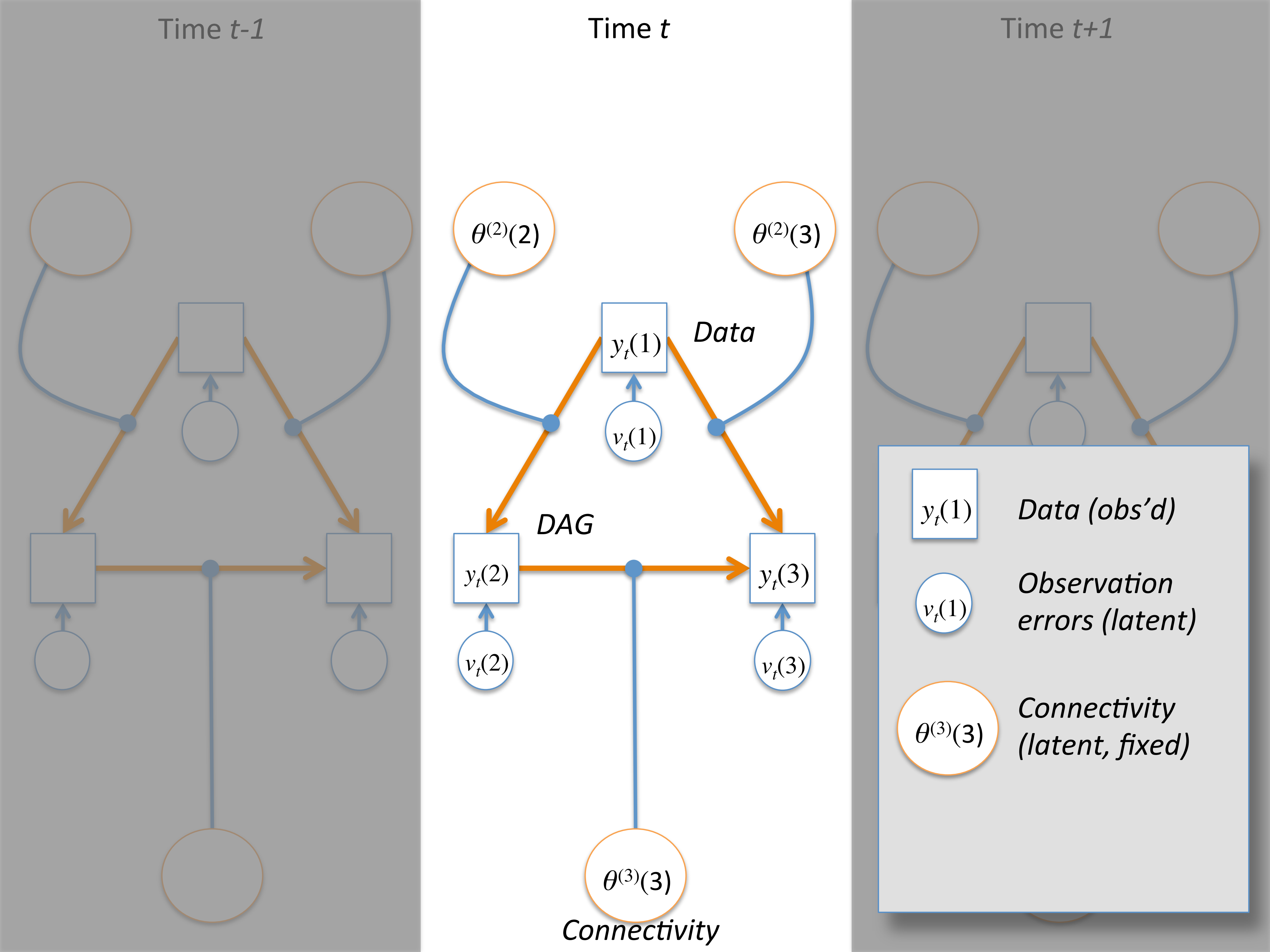
$$\boldsymbol{\theta}_t = \boldsymbol{\theta}_{t-1} + \mathbf{w}_t, \qquad \mathbf{w}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{W}_t) \text{ and } \mathbf{W}_t(r) = V_t(r)\mathbf{W}_t^*(r);$$

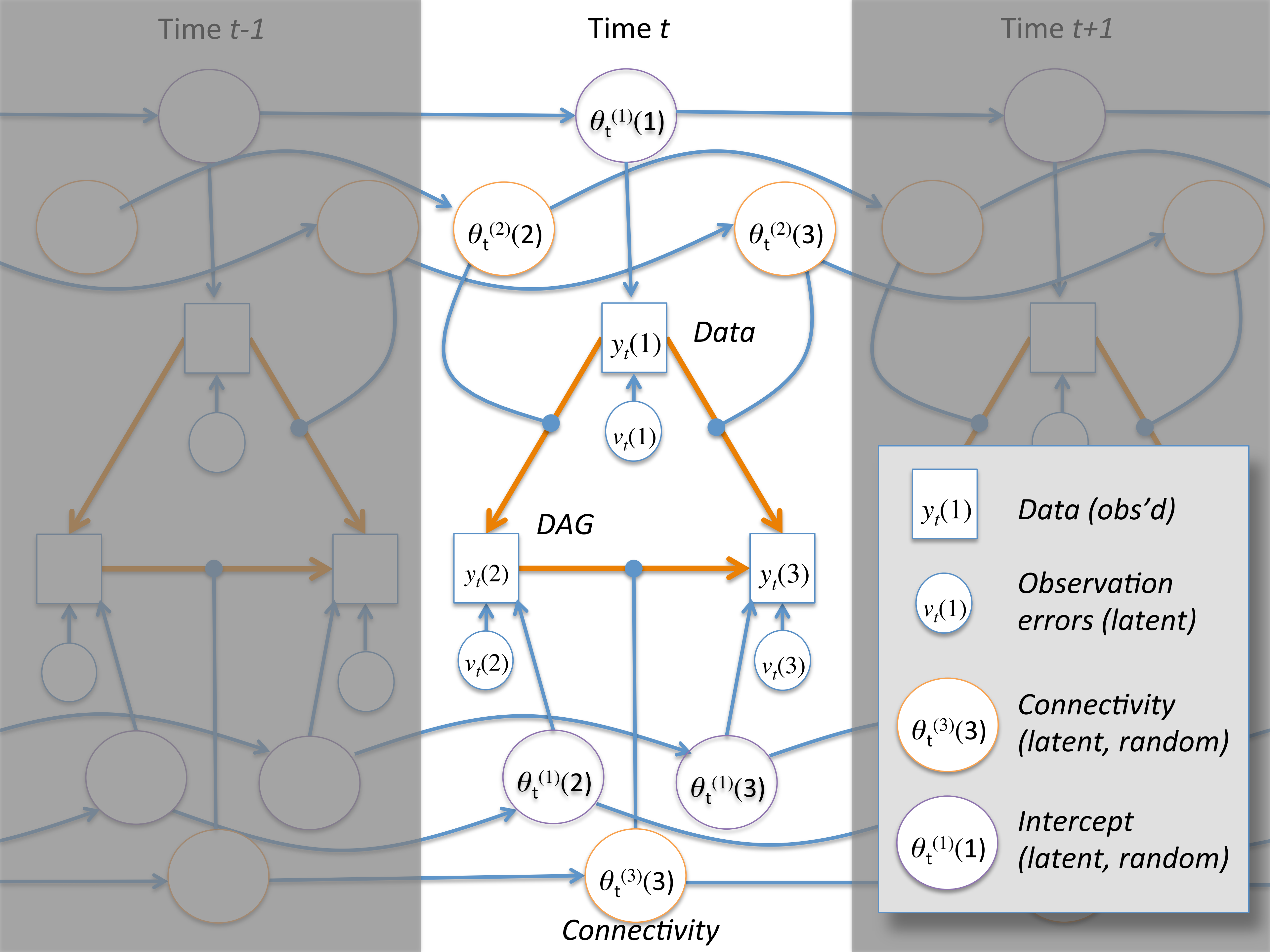
Initial information

$$(\boldsymbol{\theta}_0 | y_0) \sim \mathcal{N}(\mathbf{m}_0, \mathbf{C}_0) \text{ and } \mathbf{C}_0(r) = V_t(r)\mathbf{C}_0^*(r).$$

- Multivariate, Bayes Net version of DLM
  - Regressors (here, $\boldsymbol{F}_t$ at time t) are other Y_t's, *contemporaneous* values at other regions

Queen & Smith (1993). Multiregression dynamic models. *JRSS-B*, 55(4), 849–870.


Time t-1
Time t
Time t+1
$\theta^{(2)}(2)$
$\theta^{(2)}(3)$
$y_t(1)$
Data
$v_t(1)$
DAG
$y_t(2)$
$y_t(3)$
$v_t(2)$
$v_t(3)$
$\theta^{(3)}(3)$
Connectivity
$y_t(1)$ Data (obs'd)
$v_t(1)$ Observation errors (latent)
$\theta^{(3)}(3)$ Connectivity (latent, fixed)


Time t-1
Time t
Time t+1
$\theta_t^{(1)}(1)$
$\theta_t^{(2)}(2)$
$\theta_t^{(2)}(3)$
$y_t(1)$
Data
$v_t(1)$
DAG
$y_t(2)$
$y_t(3)$
$v_t(2)$
$v_t(3)$
$\theta_t^{(1)}(2)$
$\theta_t^{(1)}(3)$
$\theta_t^{(3)}(3)$
Connectivity
$y_t(1)$ Data (obs'd)
$v_t(1)$ Observation errors (latent)
$\theta_t^{(3)}(3)$ Connectivity (latent, random)
$\theta_t^{(1)}(1)$ Intercept (latent, random)

Fully Conjugate Inference

$\mathcal{T}_\nu$
Multivariate T distribution, ν DF

- Posterior of regression coefficient θ_t

$$(\boldsymbol{\theta}_t(r)|\mathbf{y}^T) \sim \mathcal{T}_{n_T(r)}\left(\mathbf{sm}_t(r), \mathbf{sC}_t(r)\right)$$

- "Filtering" posterior – update as of time t

$$(\boldsymbol{\theta}_t(r)|\mathbf{y}^t) \sim \mathcal{T}_{n_t(r)}(\mathbf{m}_t(r), \mathbf{C}_t(r)), \text{ where } \mathbf{y}^t = (y_1, \ldots, y_t)$$

- Predictive distribution at time t

$$(Y_t(r)|\mathbf{y}^{t-1}, \mathbf{x}_t(r)) \sim \mathcal{T}_{n_{t-1}(r)}(f_t(r), Q_t(r)), \text{ where}$$
$$f_t(r) = \mathbf{F}_t'(r)\mathbf{m}_{t-1}(r) \text{ and } \mathbf{x}_t(r)' = (y_t(1), \ldots, y_t(r-1))$$

- Model evidence

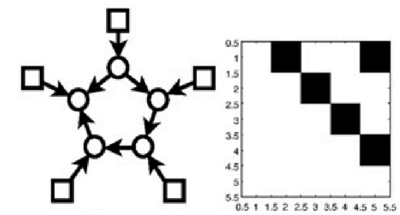
$$\log p_1(\mathbf{Y}(1)) + \ldots + \log p_n(\mathbf{Y}(n)|\mathbf{Y}(1), \ldots, \mathbf{Y}(n-1))$$
$$= \sum_{r=1}^n \sum_{t=1}^T \log p_{tr}(Y_t(r)|\mathbf{y}^{t-1}, \mathbf{x}_t(r));$$

- Log Bayes factor

$$\log(\text{BF}) = \text{LPL}(m_1) - \text{LPL}(m_0)$$

Multiregression Dynamic Model for fMRI

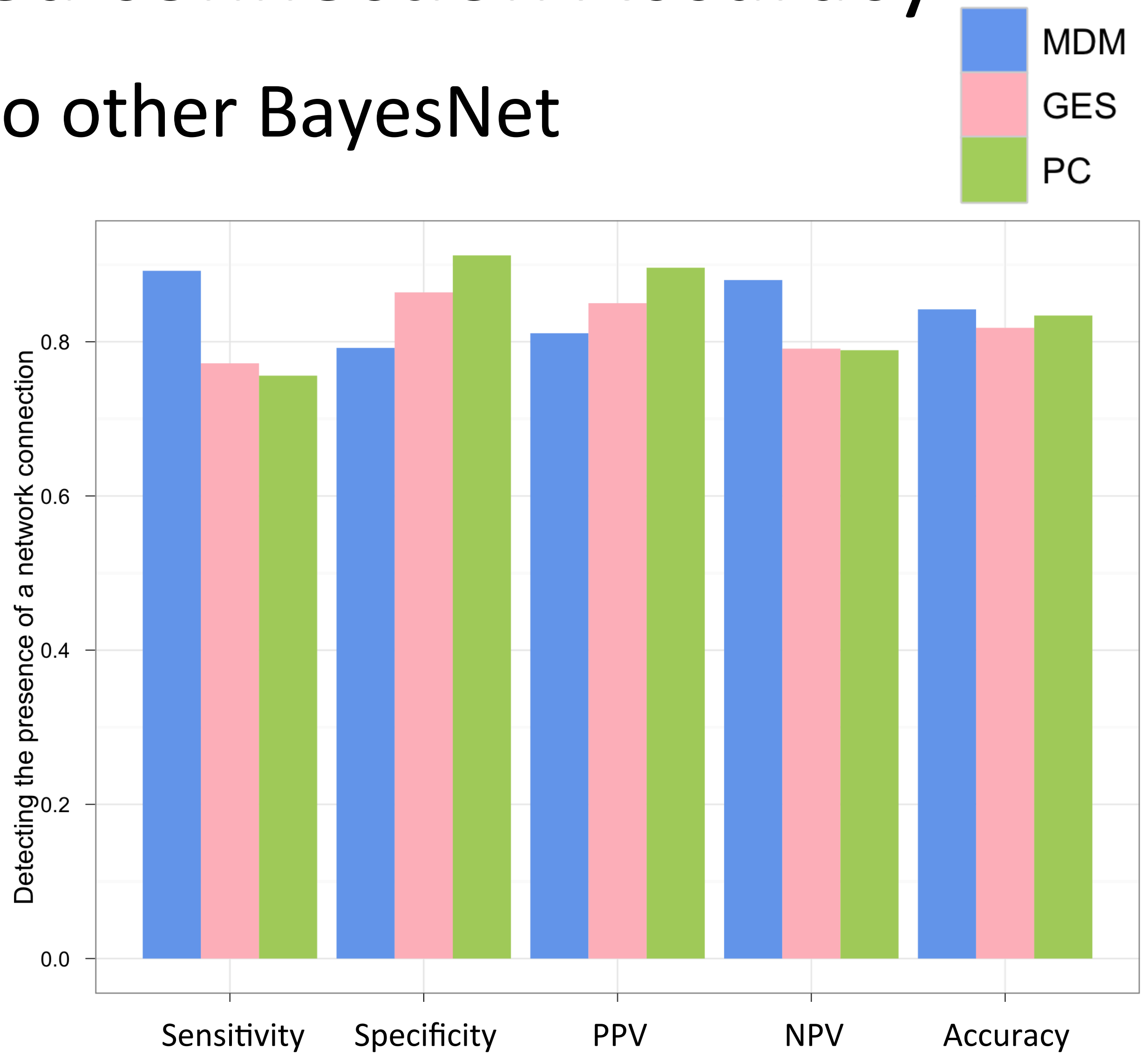
- As a Bayes Net, can estimate directionality
- Evaluate accuracy with Smith et al. (2011) 'NetSim' data
  - DCM forward model, 5 nodes
  - "sim22" dynamic simulation
    - Directed (neural) connections bistable, on/off, ≈20s/30s
    - 10 min sessions, TR = 3.00s, 50 realisations
  - Measured
    - 'c-sensitivity' – Connection (ignoring direction) accuracy
    - 'd-accuracy' – Proportion of directed edges correctly detected



Costa, Smith, Nichols, Cussens, Duff, Makin. (2015). Searching Multiregression Dynamic Models of Resting-State fMRI Networks Using Integer Programming. *Bayesian Analysis*, 10(2), 441–478.

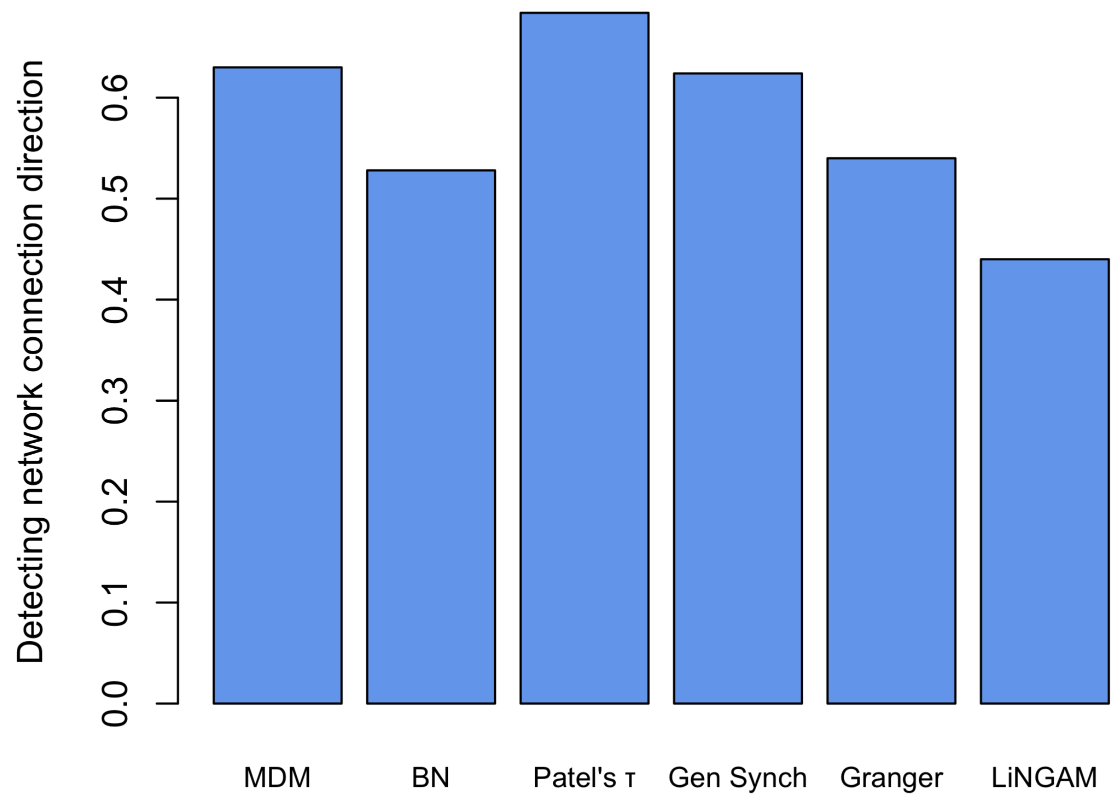
MDM for fMRI: NetSim Results Undirected connection Accuracy

- Comparable to other BayesNet estimation methods
  - GES, PC
  - Though note other methods not dynamic



MDM for fMRI: NetSim Results Directed connection Accuracy

- Nearly as accurate as best, Patel's τ
  - But, again, all other methods static
  - MDM is achieving this accuracy while estimating dynamic connectivity



Dynamic Graphic Model

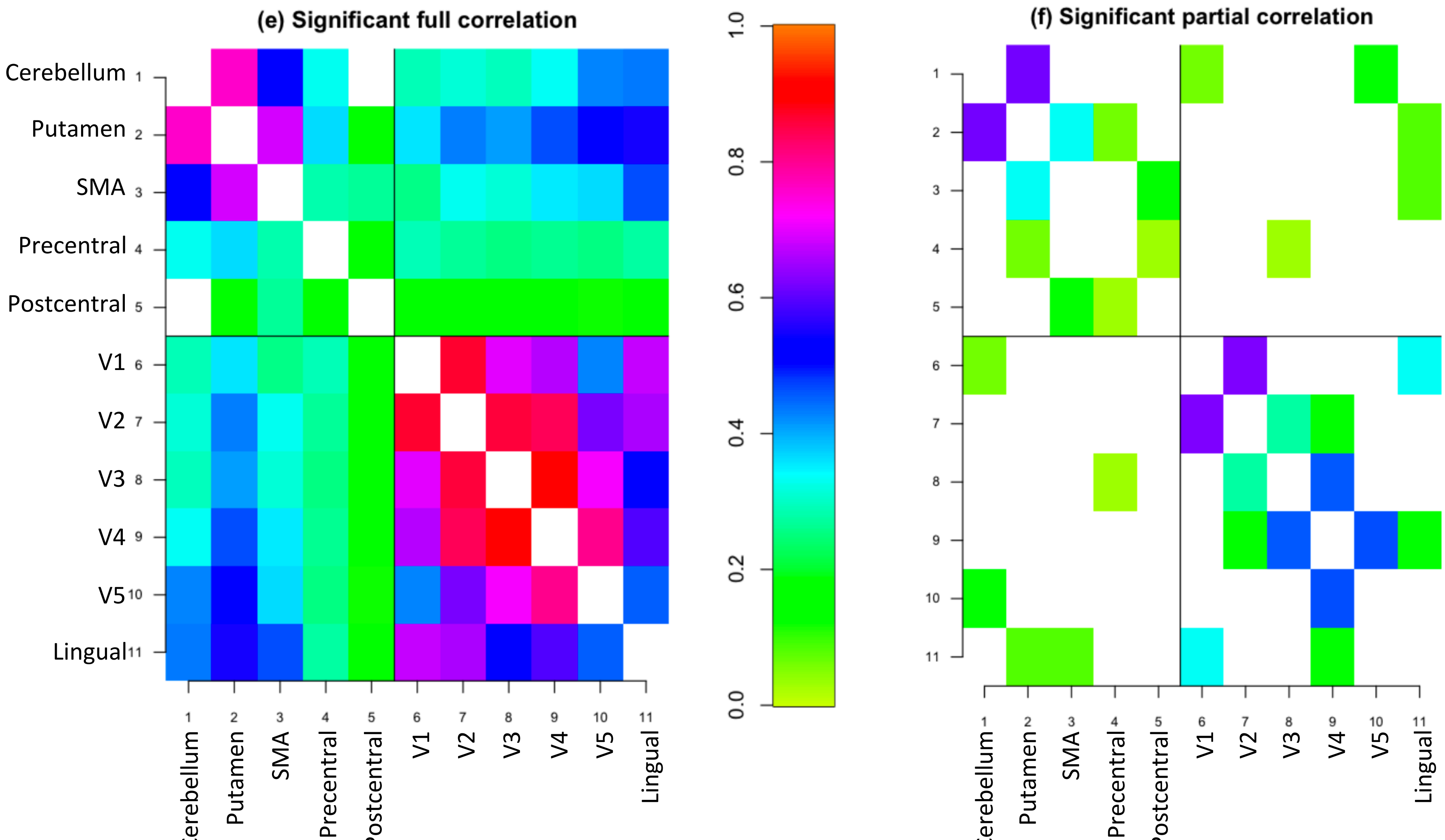
- MDM & NetSim truth are DAG models
  - Directed Acyclic Graphs
- The brain isn't a DAG
  - Cyclical connections rule rather exception
- Solution
  - Instead of finding optimal Bayes Net,
  - Find optimal set of parents for each node
    - Allows cycles, yet still provides a graphical model
  - Yields a set of *R* directed models (*R* ROIs)
    - "Dynamic Graphical Model"

Application 1
Resting fMRI Motor-Visual Network

- 15 subjects
- 230 time points, TR=1.3s
- 11 ROIs
  - 5 motor, nodes 1,…,5
    - Cerebellum, Putamen, Supplementary Motor Area (SMA), Precentral Gyrus and Postcentral Gyrus
  - 6 visual, nodes 6,…,11
    - V1, V2, V3, V4, V5 and Lingual
- Assessed consistent edges over subjects
  - Binomial test
  - Ho: Edges randomly distributed over all subjects, edges

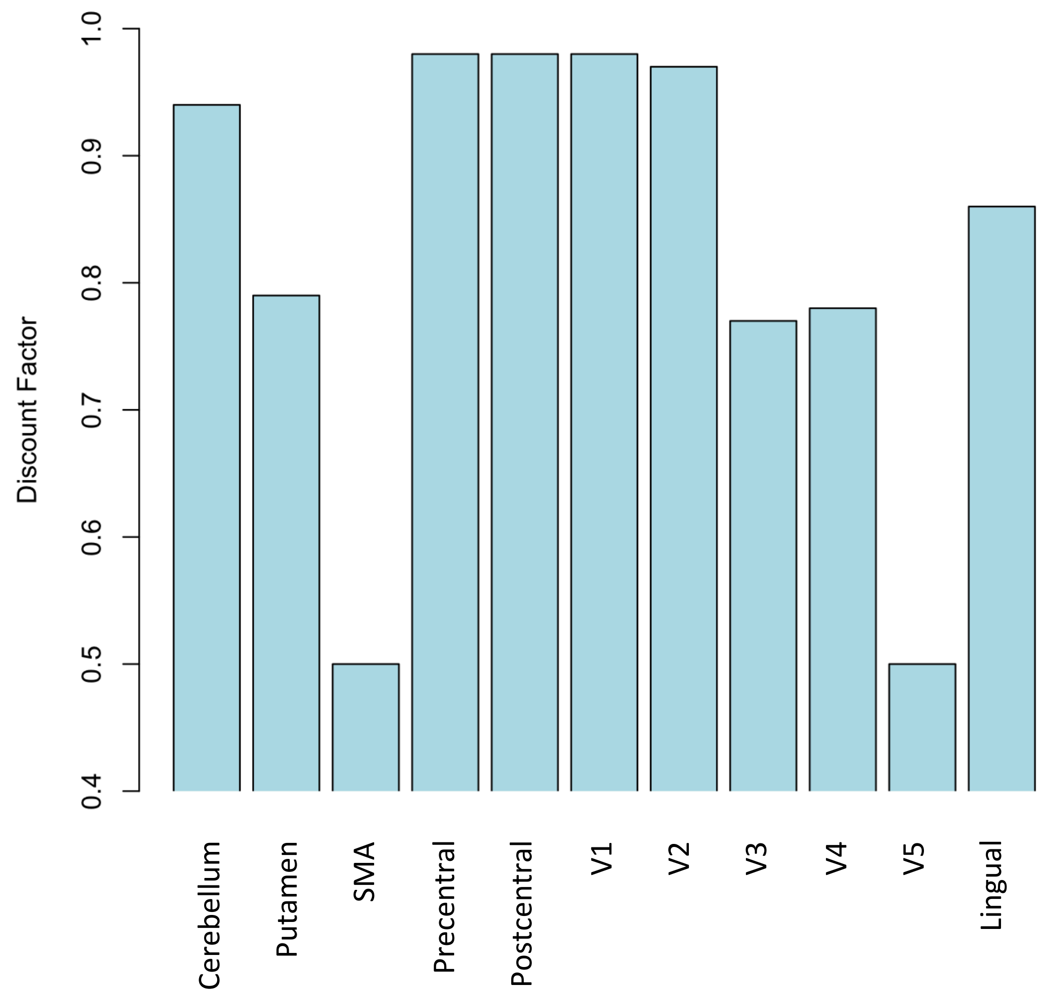
Application 1: Results

- For reference: Full & Partial correlation

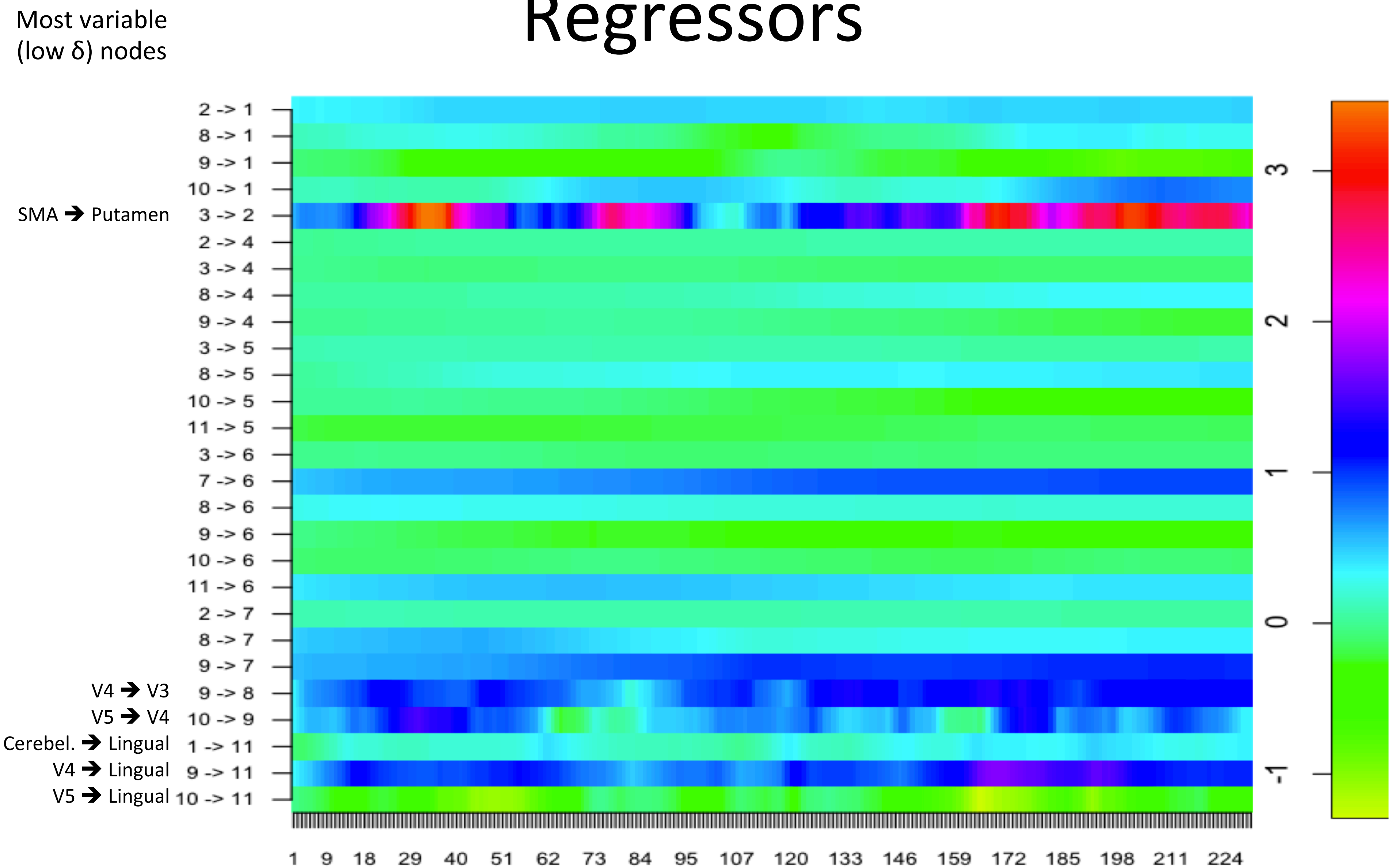


Application 1: Results Discount Factors

- SMA, V5 most variable
- Precentral, Postcenral, V1, V2 most stable
  - Also Putamen, V3, V4, Lingual

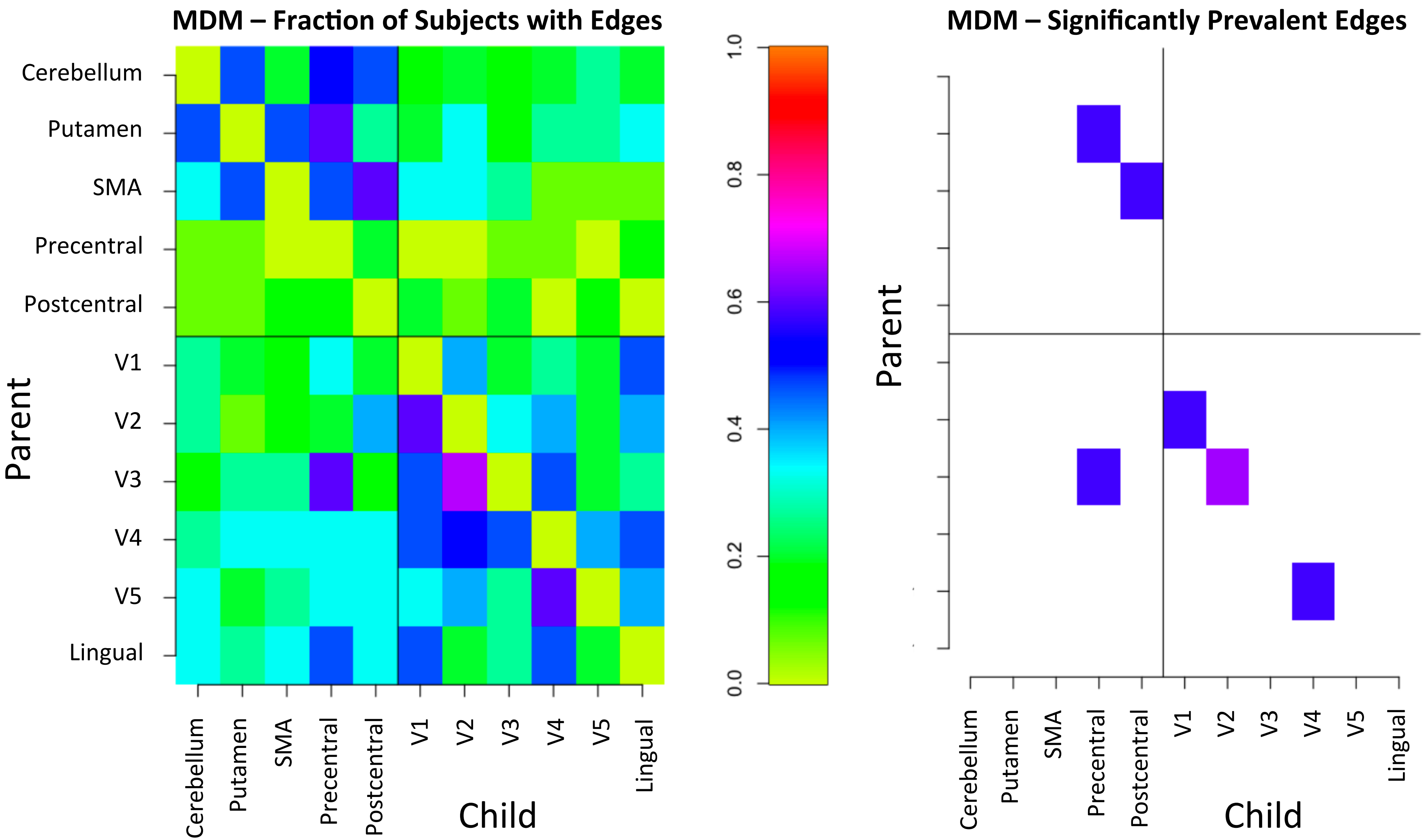


Application 1: Results
Regressors



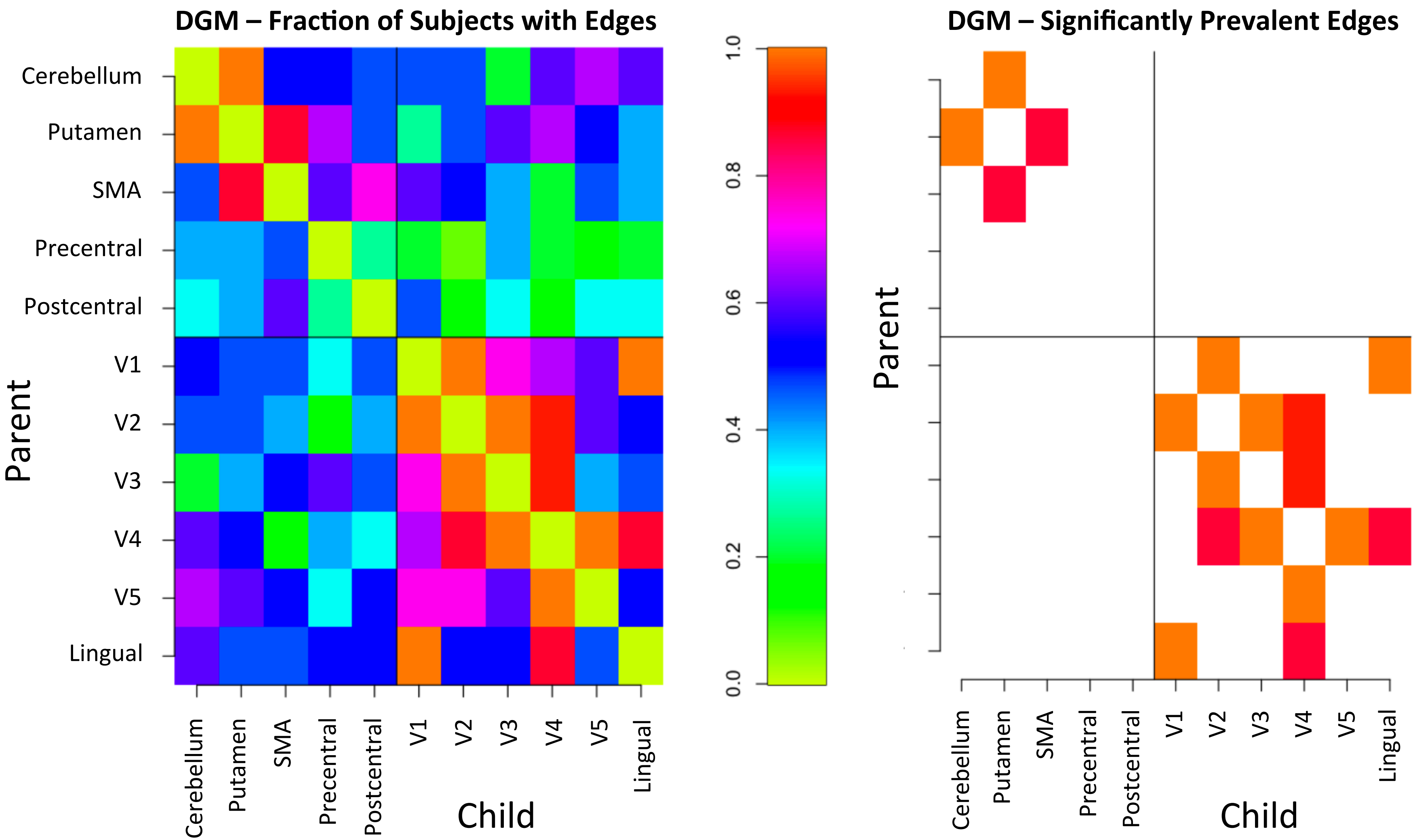
Application 1: Results

- MDM (DAG) results
  - – Clear directionality – but poor consistency



Application 1: Results

- DGM results
  - Clear evidence for bi-directial connections
  - Excellent consistency



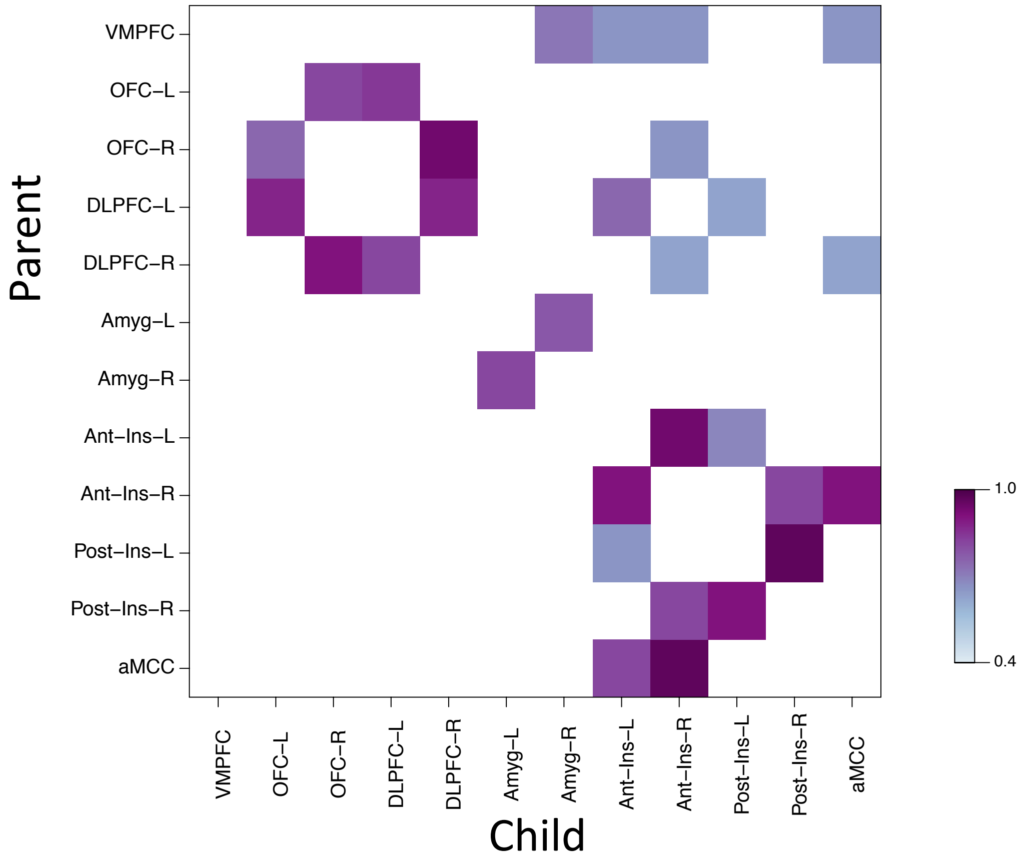
Application 2

- 32 subjects
- 15min , TR=1.1ss
- 12ROIs defined functionally & anatomically
  - Ventromedial prefrontal cortex (VMPFC), Amygdala (L,R), orbitofrontal cortex (OFC, L,R), dorsolateral prefrontal cortex (DLPFC, L,R), Anterior Insula (AntIns, L,R), Posterior Insula (PostIns, L,R), and anterior midcingulate cortex (aMCC).
- “Safe” resting data from Bijsterbosch et al., (2015).
  - Bijsterbosch et al. (2015) Functional Connectivity under Anticipation of Shock: Correlates of Trait Anxious Effect versus Induced Anxiety, Journal of Cognitive Neuroscience

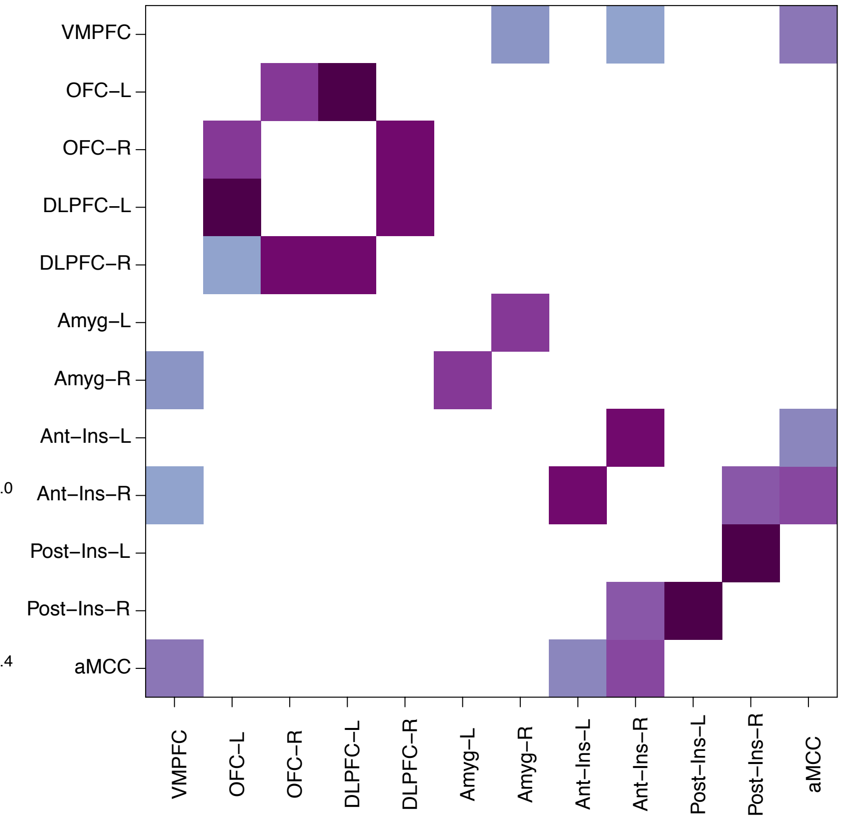
Consistent Edges

- Similar to partial correlation
  - But picks up DLPFC ➔ Insular, VMPFC ➔ (several)

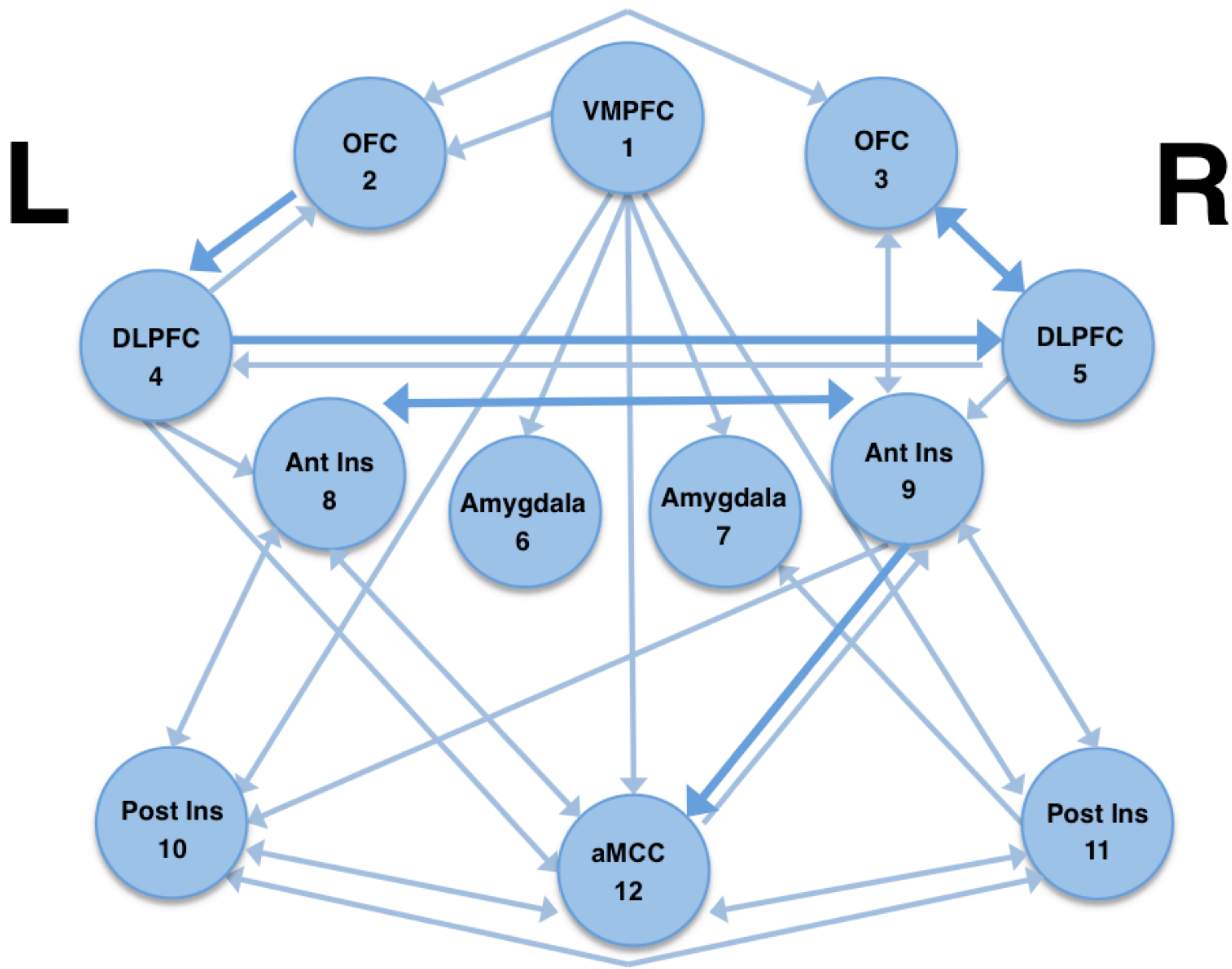
DGM – Fraction of Subjects with Edges
Significantly Prevalent Edges Only



Partial Correlation
Fraction of Subjects with Edges
Significantly Prevalent Edges Only



Consistent Network Edges



Conclusions

- Computationally efficient yet flexible framework for dynamic connectivity
- Here, dynamic aspect regarded as nuisance
  - Still interpreting dynamic model
- Great potential for understanding directed causal structure in the brain
- Thanks
  - Jim Q. Smith, Lilia Costa, Ruth Harbord – MDM/DGM
  - Steve Smith – Data & NetSim help
  - Sonia Bishop, Janine Bijsterbosch – Data
  - Eugene Duff, Tamar Makin – Data