A Hierarchical Random Effects State-Space Model for Modeling Brain Activities from EEG Data

Xingche Guo

Department of Biostatistics
Columbia University

JSM, August 2024

COLUMBIA UNIVERSITY | MAILMAN SCHOOL of PUBLIC HEALTH
BIOSTATISTICS

EEG Experiments

- Electroencephalogram (EEG): a non-invasive neuroimaging technique that measures changes in electrical voltage on the scalp induced by brain activities.

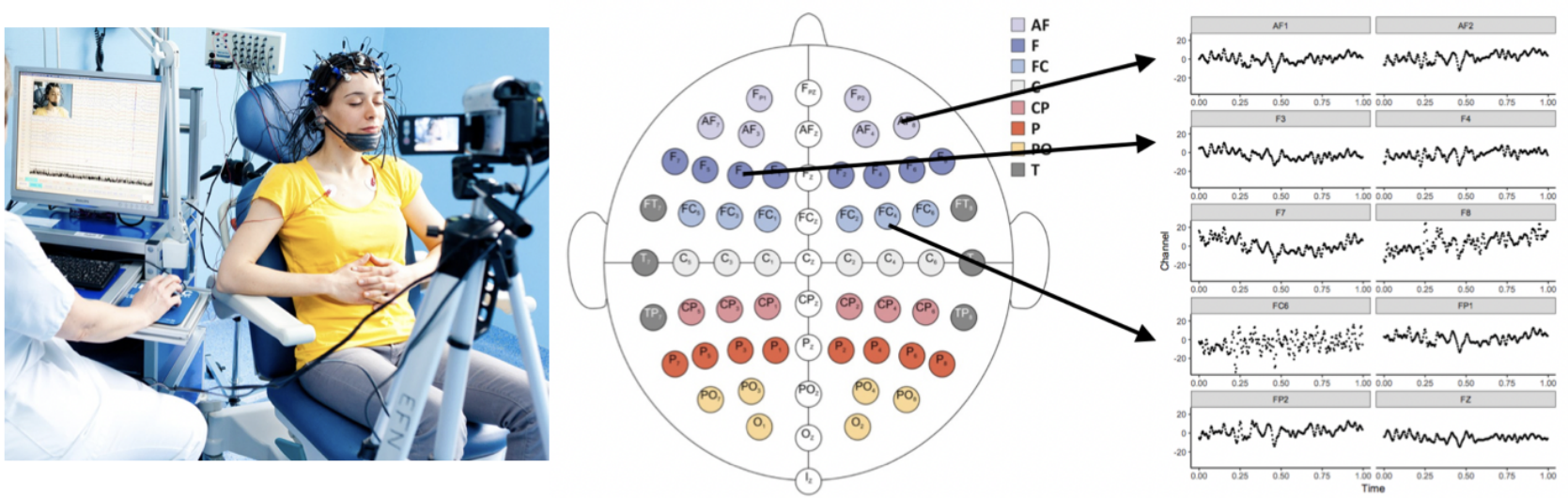


Fig. EEG scalp map and signals observed from each electrode channel

- Clinical applications: monitoring epilepsy and head injury
- Research applications: biomarkers for dementia, psychiatric disorders.

EEG Experiments

- Advantages of EEG experiments:
  - Non-invasive neuroimaging technique
  - High temporal resolution to capture dynamic changes of brain activities
- Challenges of EEG experiments:
  - Spatial resolution is low
  - Fast-changing brain activities with complex interactions
  - Substantial between-subject variability
  - Low signal-to-noise ratio
- Analytical challenge: multi-dimensional time-series data

Existing Methods for Modeling EEG Data

- Reduce EEG signals into frequency domain summary measures, e.g., band powers (Tong and Thankor, 2009)
- Brain sources reconstruction: using a head model, brain anatomy and physiology information on volume conduction (Baillet et al., 2001)
- Blind-source separation approaches: independent component analysis (ICA) (Huster et al., 2015)
- Ordinal differential equations (Zhang et al., 2017; Sun et al., 2020):
- State-space models (Li et al., 2019; Gao et al., 2020): single subject, model each channel separately
- Multi-subject state-space model (Wang et al., 2023): homogeneous spatial patterns, assume stationarity.

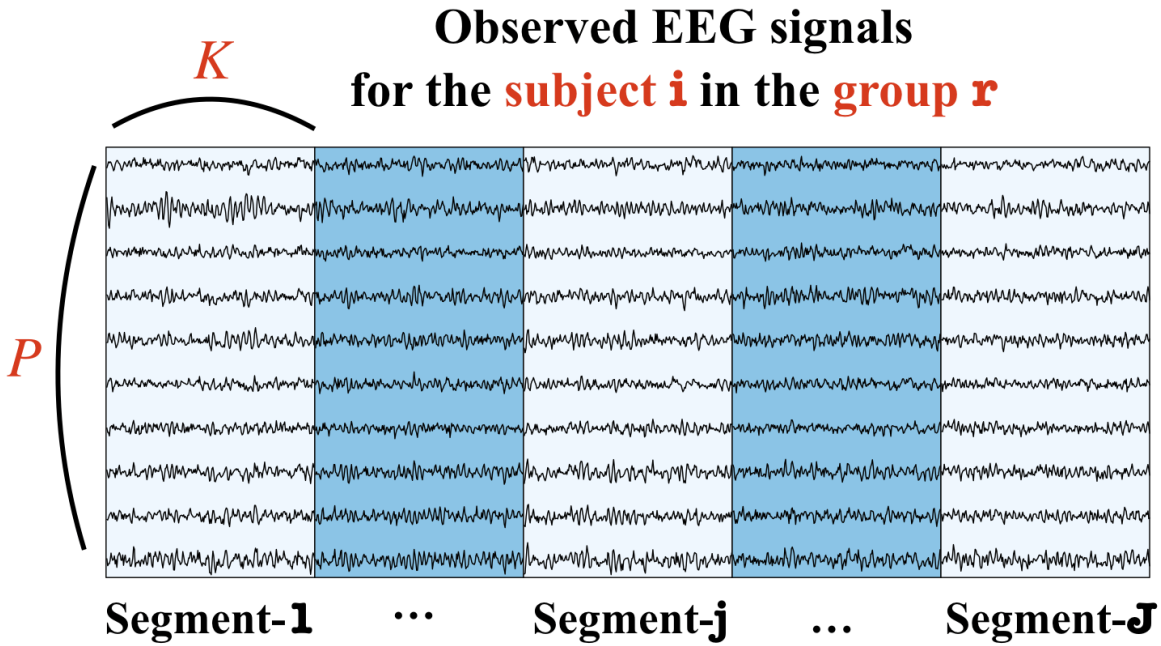
Challenges

- High between-subject variations observed in EEG data due to
  - difference in cortical connectivity (Hassan, 2018).
  - variability of the onset of neuronal processes in response to external stimulus or internal events (Makeig et al., 2004).
  - inter-individual variability of scalp topographies/brain anatomy (Huster et al., 2015).
- Resting-state EEG data without stimuli
- Our method is based on latent state space models.
  - Jointly analyzing multi-channel EEG recordings and learning dynamics of sources of cortical activities
  - Express observed recordings as combinations of concurrently active lower-dimensional latent brain sources
  - Accounting for heterogeneity of spatial mapping and non-stationarity using random effects

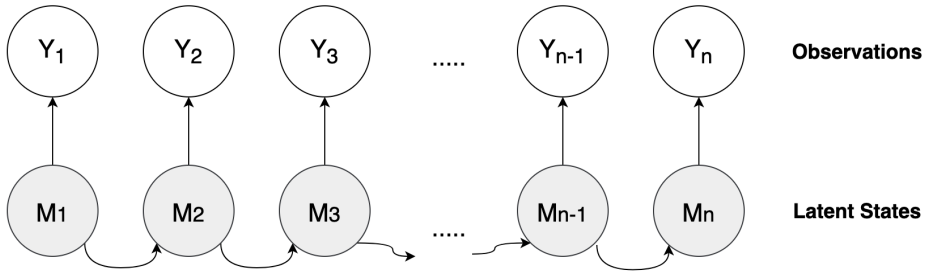
Methods

Notation

- Resting-state EEG data: $\mathbf{Y}_{rij}(t_k) = (Y_{rij1}(t_k), \ldots, Y_{rijP}(t_k))^\top$.
- P scalp electrodes ($P = 54$).
- Subgroup $r = 1, \ldots, R$ ($R = 4 : \{\text{CU, TX}\} \times \{\text{MDD, Control}\}$).
- Subject $i = 1, \ldots, N_r$ ($\sum_{r=1}^{4} N_r = 147$).
- Time segment $j = 1, \ldots, J_{ri}$ ($J_{ri} \equiv J = 40$).
- Observations $k = 1, \ldots, K$ ($K = 250$).



Random effects state-space models (RESSM)[1]



- The EEG signals (observed states) are often a mixture of unknown underlying brain source (latent states) with lower dimensions.
- Sensor model:

$$\mathbf{Y}_{rij}(t_k) = \mathbf{\Theta}_{rij}\mathbf{M}_{rij}(t_k) + \boldsymbol{\epsilon}_{rij}(t_k), \quad \boldsymbol{\epsilon}_{rij}(t_k) \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_{rij}).$$

- $\mathbf{M}_{rij}(t_k)$ is Q-dimensional latent EEG signals with $(Q < P)$.
- Vector autoregressive (VAR) model (with order m):

$$\mathbf{M}_{rij}(t_k) = \sum_{h=1}^{m} \mathbf{A}_{rijh}\mathbf{M}_{rij}(t_{k-h}) + \mathbf{W}_{rij}(t_k), \ \mathbf{W}_{rij}(t_k) \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_w).$$

- Locally stationary within each segment and non-stationary across the entire EEG sequence.

[1] Guo et al. (2024). arXiv:2310.03164.

Random effects state-space models (RESSM)

- $\mathbf{\Theta}_{rij}$: $P \times Q$ spatial mapping matrix.
- $\mathbf{A}_{rij} = [\mathbf{A}_{rij1}, \ldots, \mathbf{A}_{rijm}]$: $Q \times mQ$ temporal dynamical matrix.
- Define $\widetilde{\mathbf{Y}}_{rij}(t_k) = \mathbf{\Theta}_{rij}\mathbf{M}_{rij}(t_k)$, then:

$$\widetilde{\mathbf{Y}}_{rij}(t_k) = \sum_{h=1}^{m} \mathbf{B}_{rijh}\widetilde{\mathbf{Y}}_{rij}(t_{k-h}) + \mathbf{U}_{rij}(t_k), \ \mathbf{U}_{rij}(t_k) = \mathbf{\Theta}_{rij}\mathbf{W}_{rij}(t_k),$$

where

$$\mathbf{B}_{rijh} = \mathbf{\Theta}_{rij}\mathbf{A}_{rijh}\left(\mathbf{\Theta}_{rij}^{\top}\mathbf{\Theta}_{rij}\right)^{-1}\mathbf{\Theta}_{rij}^{\top}, \quad h = 1, \ldots, m.$$

- $\mathbf{B}_{rijh}$: $P \times P$ directional connectivity[2] matrix.

[2] Li et al (2021). Mapping epileptic directional brain networks using intracranial EEG data. *Biostatistics*, 22(3):613–628.

Model identifiability

- RESSM with an arbitrary covariance structure of $\boldsymbol{\Sigma}_w$ and $\boldsymbol{\Sigma}_{rij}$ can be transformed into an equivalent RESSM where $\boldsymbol{\Sigma}_w = \mathbf{I}$ and $\boldsymbol{\Sigma}_{rij}$ is diagonal.

$\Rightarrow$ We set $\boldsymbol{\Sigma}_w = \mathbf{I}$ and constraining $\boldsymbol{\Sigma}_{rij}$ as diagonal matrices.

- RESSM is only identifiable up to a rotation: for any orthogonal matrix $\mathbf{R}$, the following is an equivalent parameterization:

$$\mathbf{M}^*_{rij}(t_k) = \mathbf{R}\mathbf{M}_{rij}(t_k), \quad \boldsymbol{\Theta}^*_{rij} = \boldsymbol{\Theta}_{rij}\mathbf{R}^\top, \quad \mathbf{A}^*_{rijh} = \mathbf{R}\mathbf{A}_{rijh}\mathbf{R}^\top.$$

$\Rightarrow$ Using QR decomposition, we can force $\boldsymbol{\Theta}_{rij}$ to be lower-triangular matrices.

- RESSM is identifiable up to a sign.

Hierarchical structure

- Temporal dynamical matrices:

$$[\text{vec}(\mathbf{A}_{rij}) \mid \text{vec}(\mathbf{A}_{ri})] \sim \mathcal{N}\left(\text{vec}(\mathbf{A}_{ri}), \mathbf{\Sigma}_{v,r}\right),$$

$$[\text{vec}(\mathbf{A}_{ri}) \mid \text{vec}(\mathbf{A}_{r})] \sim \mathcal{N}\left(\text{vec}(\mathbf{A}_{r}), \mathbf{\Sigma}_{\gamma,r}\right),$$

$$[\text{vec}(\mathbf{A}_{r}) \mid \text{vec}(\mathbf{A})] \sim \mathcal{N}\left(\text{vec}(\mathbf{A}), \mathbf{\Sigma}_{a}\right)$$

- Spatial mapping matrices

$$[\text{low}(\mathbf{\Theta}_{rij}) \mid \text{low}(\mathbf{\Theta}_{ri})] \sim \mathcal{N}\left(\text{low}(\mathbf{\Theta}_{ri}), \mathbf{\Sigma}_{u,r}\right),$$

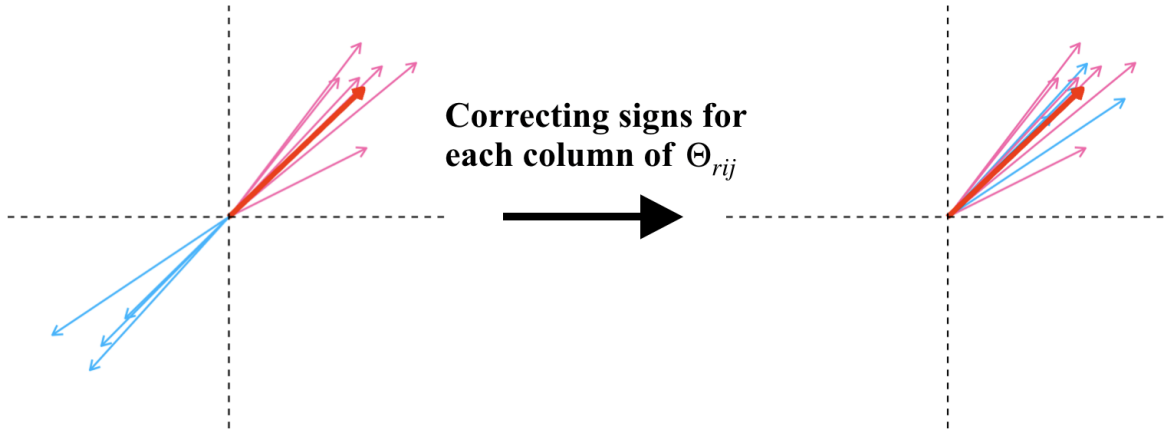
$$[\text{low}(\mathbf{\Theta}_{ri}) \mid \text{low}(\mathbf{\Theta}_{r})] \sim \mathcal{N}\left(\text{low}(\mathbf{\Theta}_{r}), \mathbf{\Sigma}_{\psi,r}\right),$$

$$[\text{low}(\mathbf{\Theta}_{r}) \mid \text{low}(\mathbf{\Theta})] \sim \mathcal{N}\left(\text{low}(\mathbf{\Theta}), \mathbf{\Sigma}_{\theta}\right)$$

- Full Bayesian analysis coupled with an efficient Gibbs sampler
  - $\sim$6000 spatial matrices, each 54×5, leverage block structure of covariances

Sign identifiability

- Switching the sign of the q-th element of $\mathbf{M}_{rij}(t_k)$ and the sign of the q-th column of $\mathbf{\Theta}_{rij}$ won't change the model.
- Need sign identifiability of $\mathbf{\Theta}_{rij}$ for
  - a meaning computation of their parent parameters (i.e. $\mathbf{\Theta}_{ri}$ and $\mathbf{\Theta}_{r}$).
  - a valid MCMC inference.
- A two-stage procedure to achieve sign identifiability:
  - (Initialization stage): Apply the MCMC algorithm to a simplified RESSM (with $\mathbf{\Theta}_{rij} \equiv \mathbf{\Theta}_0$) as initial value for all levels of $\mathbf{\Theta}$.
  - (Sign-tracking stage): Monitor the sign for each column of $\mathbf{\Theta}$ at each MCMC iteration and adjust if different from their parent level.
  - A cosine correlation is used to check the sign agreement.



Selecting the number of latent states Q and VAR order m

- Complete DIC (cDIC; Celeux et al. 2006): replace the observed likelihood $p(\mathbf{Y})$ in DIC with the complete likelihood $p(\mathbf{Y}, \mathbf{M} \mid \mathbf{\Lambda})$

$$\begin{aligned}\mathrm{cDIC} = &-4\mathbb{E}_{\mathbf{\Lambda},\mathbf{M}|\mathbf{Y}}\left[\log p(\mathbf{Y}, \mathbf{M} \mid \mathbf{\Lambda})\right] \\ &+ 2\mathbb{E}_{\mathbf{M}|\mathbf{Y}}\left[\log p\left(\mathbf{Y}, \mathbf{M} \;\middle|\; \mathbb{E}_{\mathbf{\Lambda}|\mathbf{Y},\mathbf{M}}[\mathbf{\Lambda}]\right)\right]\end{aligned}$$

  $\mathbf{\Lambda}$: collection of all parameters.

- cDIC is effective in selecting the number of latent states Q, not sensitive to the VAR order m.
- The model's performance is robust and not significantly affected by the VAR order m.

Simulation Studies

Simulation study

- $P = 54,\ Q = 2,\ m = 2,$ K=250 (125Hz), AR(2), Rep $= 100$.
- Scenario 1: $n_1 = n_2 = 75$; Scenario 2: $n_1 = 75, n_2 = 20$

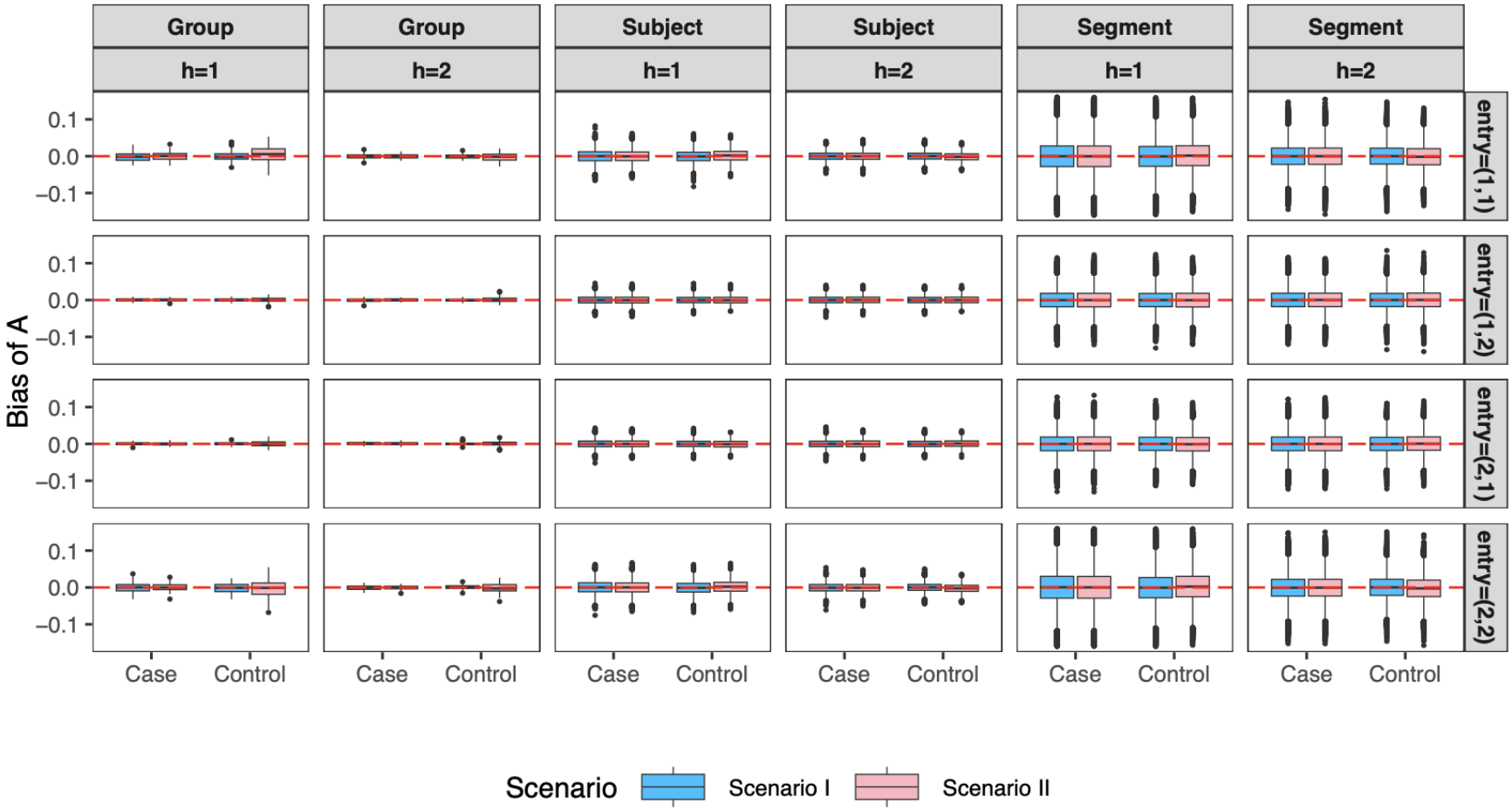


Figure 1: Bias of the temporal dynamic matrix **A**

Simulation study

- $P = 54$, $Q = 2$, $m = 2$, K=250 (125Hz), AR(2), Rep = 100.
- Scenario 1: $n_1 = n_2 = 75$; Scenario 2: $n_1 = 75, n_2 = 20$

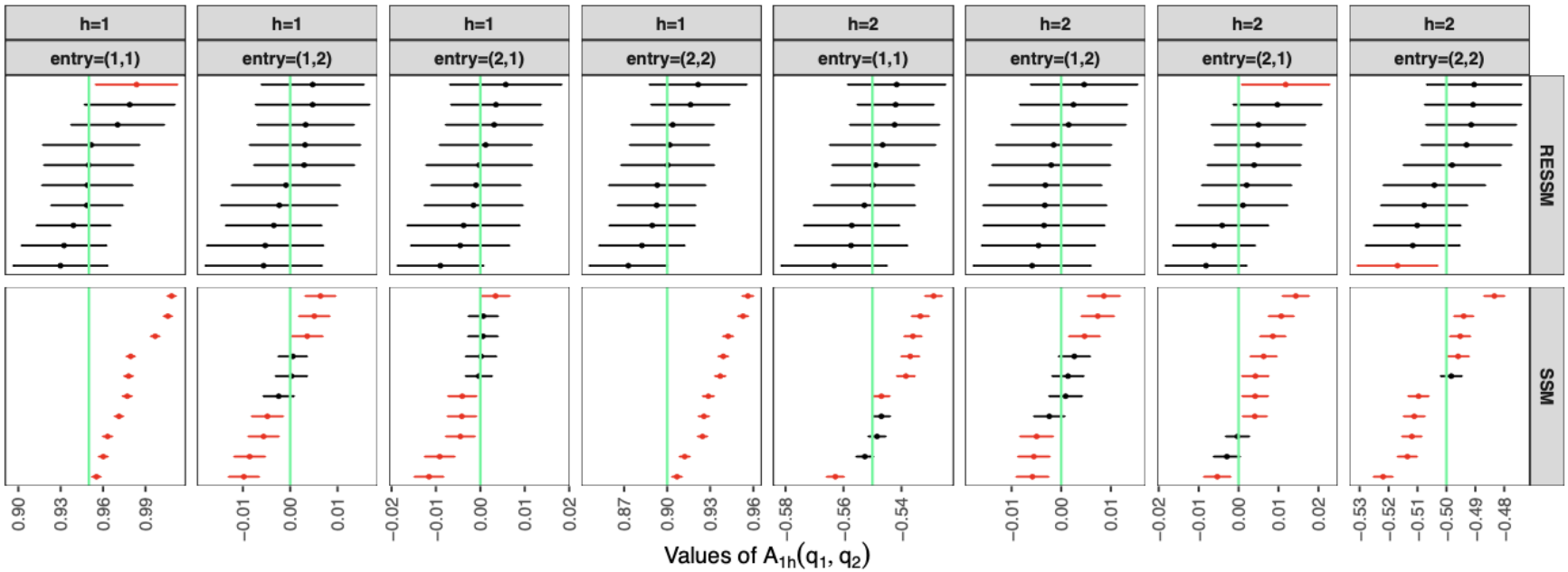


Figure 2: Comparison with state-space model ignoring heterogeneity in **A**

Simulation study (model selection)

- Vary the number of latent states Q from 2 to 4.
- Vary the number of latent states m from 1 to 3.
- Predicted latent states are robust to AR order (correlation>0.99).

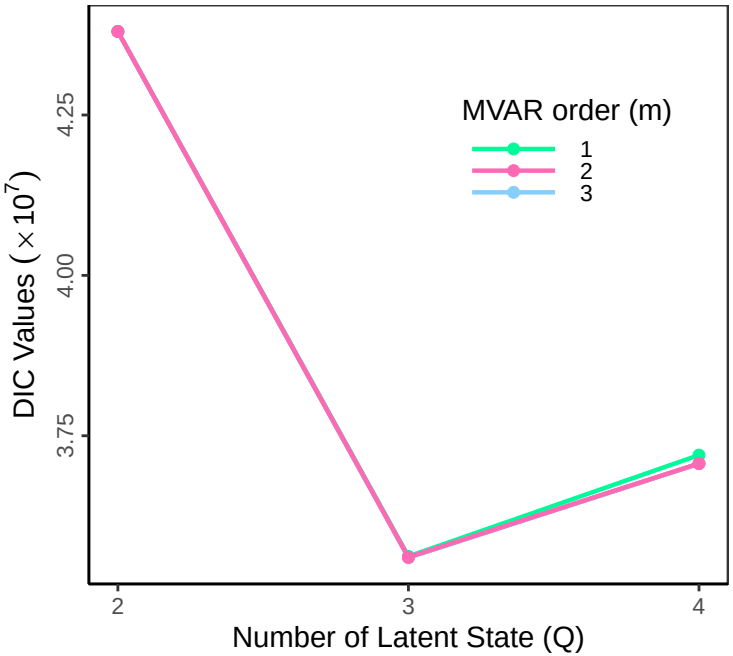


Figure 3: The cDIC values for the 9 models, where Q varies in $\{2,3,4\}$ and m varies in $\{1,2,3\}$, with the true model being $(Q=3, m=2)$.

Application to EMBARC

EMBARC study

- A clinical trial for constructing biomarker signatures of antidepressant treatment response for Major Depressive Disorder[3] (MDD).
- Baseline measures:
  - Demographical data (e.g., age, gender, education, etc).
  - Clinical data (e.g., HAMD-17 scores).
  - Neuroimaging data (i.e., task/resting-state EEG/fMRI data).
  - Human behavioral data (e.g., probabilistic reward task, etc).
- Experimental design
  - Recruited at four study centers: TX, CU, MG, UM.
  - Pre-treatment: MDD vs healthy Control.
  - MDD group randomized to SSRI antidepressants or Placebo.

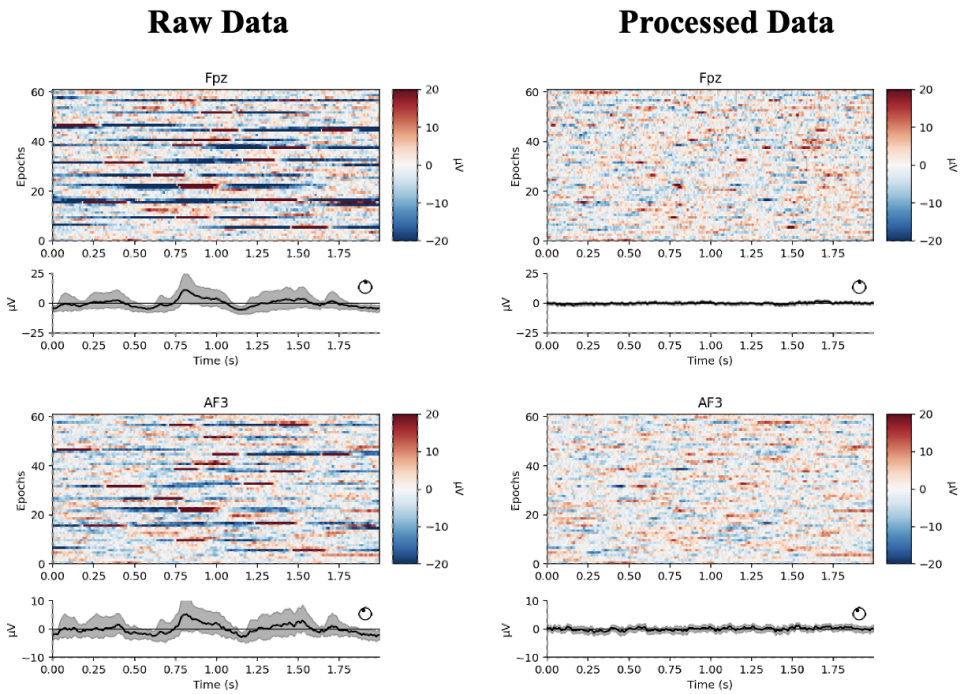
[3] Trivedi et al. (2016). Establishing moderators and biosignatures of antidepressant response in clinical care (EMBARC): Rationale and design. *Journal of Psychiatric Research*, 78:11–23

Resting-state EEG data in the EMBARC study

- 54 common EEG channels.
- Four 2-minute blocks at 125Hz measured with eyes open/close/close/open.
- EEG preprocessing procedure proposed in Yang et al.[4] (2024).

Preprocessing

- high/low-pass filtering.
- remove artifacts (e.g. eye blink).
- denoise.
- remove problematic time segments.
- standardized.



[4] Yang et al. (2024). Learning optimal biomarker-guided treatment policy for chronic disorders. *Statistics in Medicine*. In press.

EMBARC analysis details

- Subgroups: $\{\text{CU, TX}\} \times \{\text{MDD, Control}\}$.
- Total of 147 subjects (128 MDD, 19 controls).
- First 2-minute block (eyes open) is used.
- Each time segment has length 2 seconds.
- MCMC parameters
  - Total of $10,000$ MCMC iterations.
  - First $3,000$ iterations as burn-in.
  - Thinning by 10.
- Set $m = 1$ (model performance not sensitive to m).
- Selected $Q = 5$ using cDIC (varying from 2 to 6).

Spatial mapping matrix $\mathbf{\Theta}$

- No significant difference in mean $\mathbf{\Theta}_r$ between MDD and controls.
- Each latent state corresponds to specific brain regions.

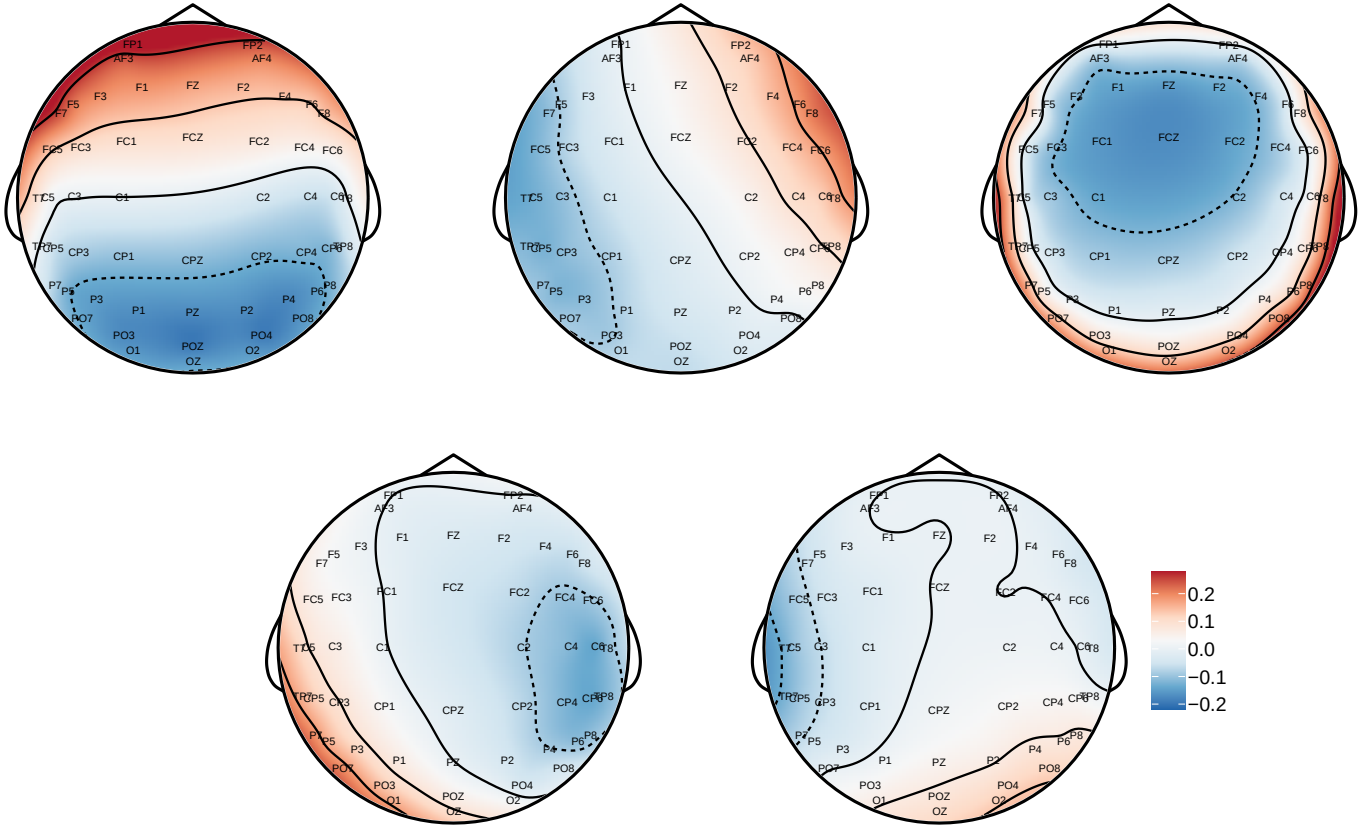


Figure 4: Topographies of the posterior means of the population-level spatial mapping matrix $\mathbf{\Theta}$ obtained from the analysis of EMBARC.

Spatial mapping matrix $\boldsymbol{\Theta}_{ri}$

- Larger between-subject heterogeneity for the MDD group.

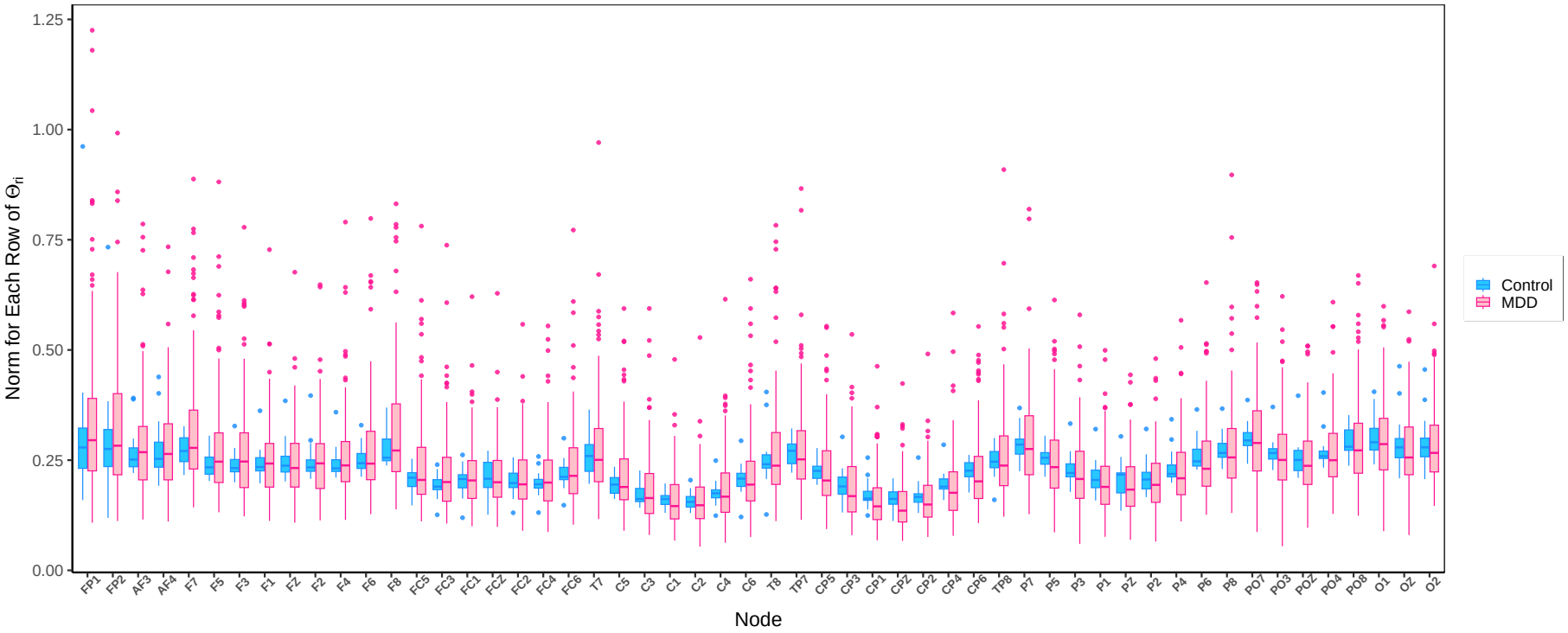


Figure 5: Boxplots of the norms of 54 rows within the subject-level spatial mapping matrices.

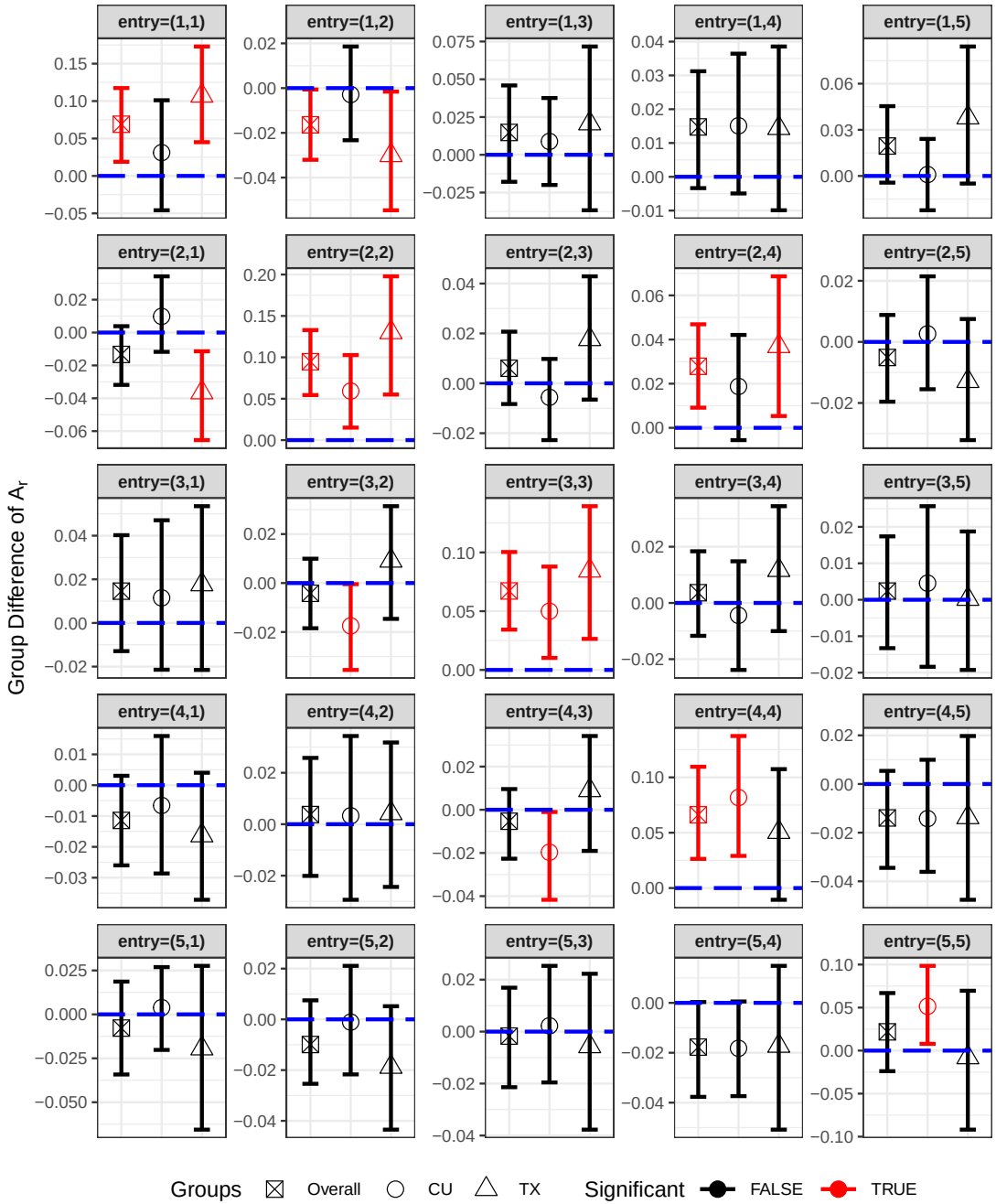
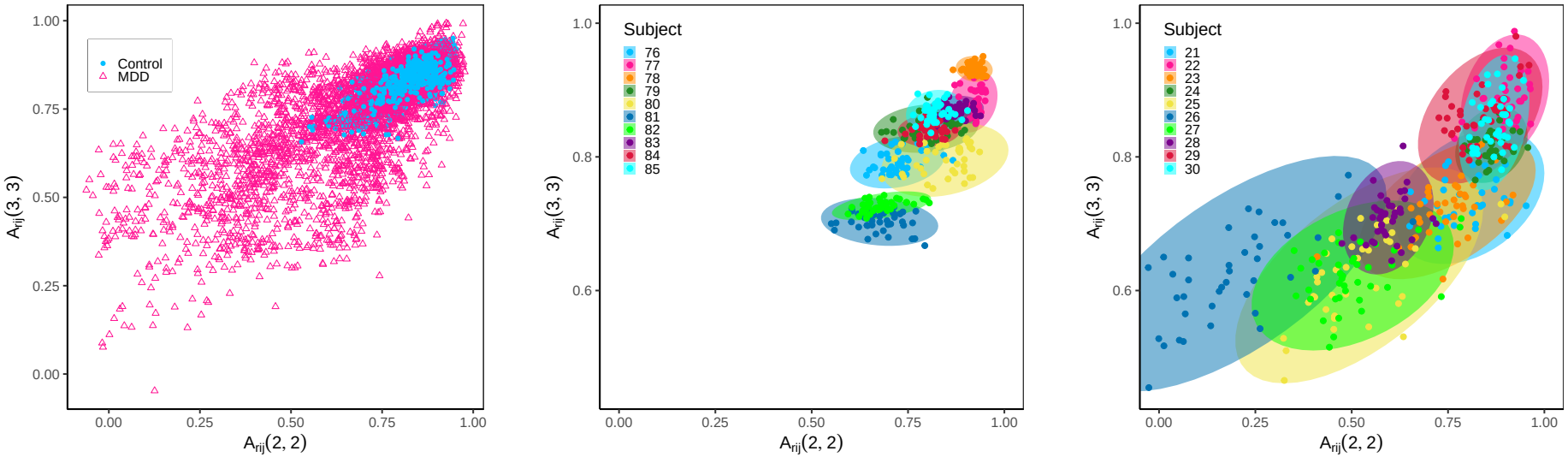


Figure 6: Inference on group differences of $\mathbf{A}_r$.

Visualize $\mathbf{A}_{rij}$



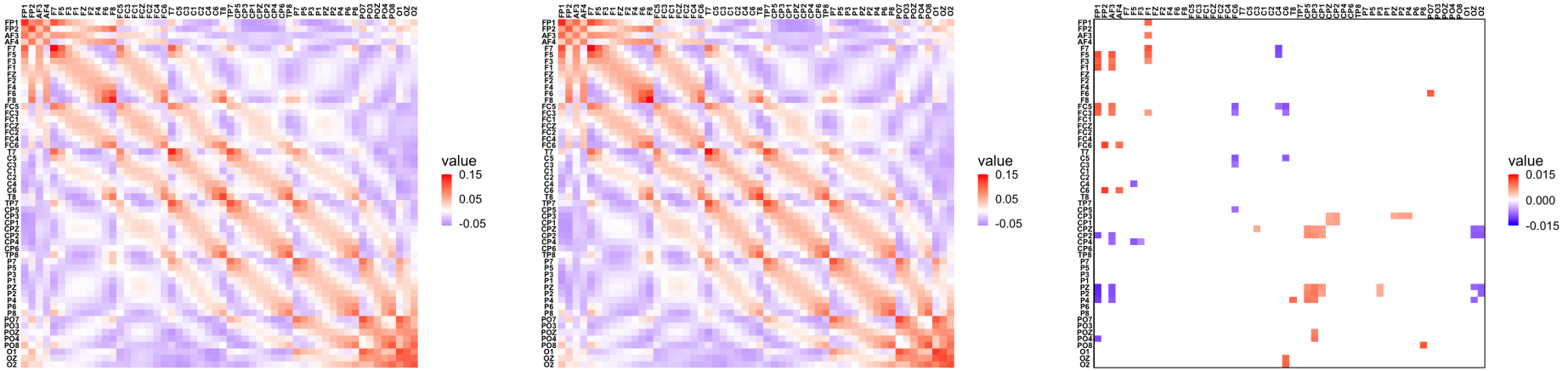
(a) All participants (b) 10 Control participants (c) 10 MDD participants

Figure 7: Scatterplots of the posterior means of segment-level $A_{rij}(3,3)$ versus $A_{rij}(2,2)$. Colors represent different subjects in subfigure (b) and (c).

- Weaker brain temporal dynamics (autocorrelation) in the MDD group.
- Consistent with that MDD patients exhibit reduced levels of neurotransmitters.

Directional connectivities

$$\mathbf{B}_{rij} = \mathbf{\Theta}_{rij}\mathbf{A}_{rij}\left(\mathbf{\Theta}_{rij}^{\top}\mathbf{\Theta}_{rij}\right)^{-1}\mathbf{\Theta}_{rij}^{\top}$$

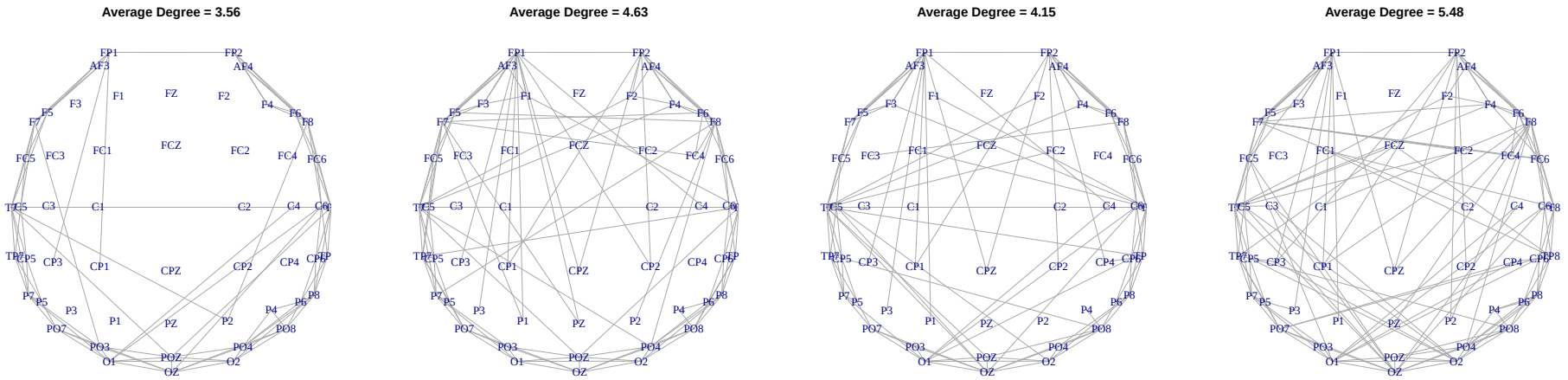


(a) MDD (b) Control (c) Control - MDD

Figure 8: Posterior means of directional connectivity matrices in MDD (a), control (b), and their differences (c).

- Symmetric patterns: mostly bi-directional connections.
- Block-banded patterns: positive between closer electrodes; negative between the opposite sides of the brain or farther brain regions (e.g., frontal/fronto-central; posterior/parieto-occipital).

Network view of group directional connectivities



(a) MDD (TX) (b) Control (TX) (c) MDD (CU) (d) Control (CU)

Figure 9: Networks of group-level directional connectivity matrices.

- Patients with MDD exhibit reduced connectivity (e.g., lower average degree) compared to the control groups.

RESSM features as biomarkers for predicting HTE

Evaluating predictiveness of EEG biomarkers on conditional average treatment effect[5] (CATE):

$$Y_i^* = \frac{Y_i(A_i - p_i)}{p_i(1-p_i)}, \quad E(Y_i^*|X_i) = E(Y_i(1) - Y_i(0)|X_i).$$

Table 1: Comparing utilities of RESSM-extracted EEG features on external treatment response outcomes in EMBARC study analysis.

		Outcome Measures		
		response at exit	remission status	change of HAMD score
RMSE[1]	clinical + RESSM[2]	1.317 (0.0446)	1.261 (0.0463)	21.114 (0.609)
	clinical + band power[3]	1.322 (0.0335)	1.263 (0.0366)	21.359 (0.588)
	clinical variables only	1.326 (0.0353)	1.266 (0.0364)	21.387 (0.579)
AUC or R-squared[4]		AUC	AUC	R-squared
	clinical + RESSM	0.677 (0.036)	0.654 (0.039)	17.9% (4.6%)
	clinical + band power	0.621 (0.022)	0.584 (0.041)	12.0% (3.7%)
	clinical variables only	0.575 (0.045)	0.517 (0.029)	8.6% (2.3%)

- RESSM features (e.g., mean, sd of $A_{rij}(s,s)$) have better predictive performance for the HTE compared to the EEG frequency band power.

[5] Tian, et al (2014). A simple method for estimating interactions between a treatment and a large number of covariates. *JASA* 109(508):1517–1532.

Disucssion

Discussion

- Propose a random effects latent state space model to simultaneously model multi-subject, multi-channel, resting EEG signals.
  - Meaningful lower-dimensional latent states represent a large number of observed electrodes.
  - Incorporate subject-, segment-level dynamic temporal matrices.
  - Substantial between-subject and between-segment heterogeneity both in spatial mappings and temporal dynamics.
  - MDD control group difference in temporal dynamics, not spatial mapping.
  - Variability of temporal dynamics may be useful biomarkers for CATE.
- Extensions
  - Model the temporal transitions over time to better account for dynamic directional connectivities (e.g., switching-state SSM).
  - Localization of brain regions (ROI-based analysis).

Acknowledgments

- References:
  Guo, X., Yang, B., Loh, J. M., Wang, Q., and Wang, Y. (2024+). A hierarchical random effects state-space model for modeling brain activities from electroencephalogram data. arXiv preprint arXiv:2310.03164. (Under revision at *Biometrics*)

  Wang Q, Loh J, He X, Wang Y. (2023). A Latent State Space Model for Estimating Brain Dynamics from Electroencephalogram (EEG) Data. *Biometrics.* 79(3): 2444-2457.

- Advisor: Prof. Yuanjia Wang.
- Research support: NS073671, GM124104, and MH123487

Thank You!