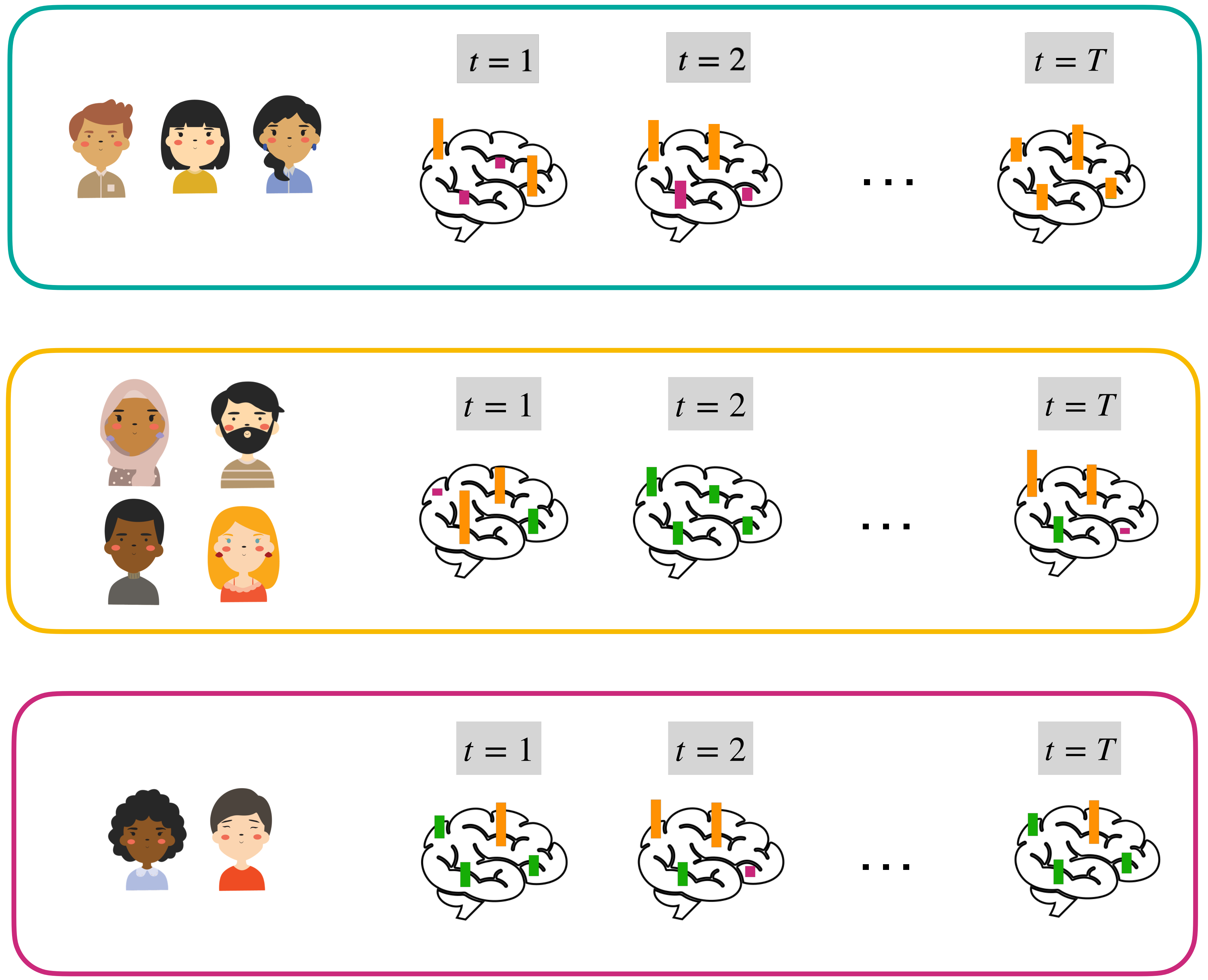
Bayesian temporal biclustering with applications to multi-subject neuroscience studies



Federica Zoe Ricci (UC Irvine)

July 5 2024, ISBA World Meeting

Collaborators

Jaylen Lee (UCI Center for Statistical Consulting)

Erik Sudderth (UCI Computer Science and Statistics)

Michele Guindani (UCLA Biostatistics)

Marina Vannucci (Rice University Statistics)

Megan Peters (UCI Cognitive Sciences)

Funding

HPI Hasso Plattner Institut
IT Systems Engineering | Universität Potsdam

Hasso Plattner Institute in Machine Learning and Data Science at UCI

Motivating data

Many studies observe **multiple units** over time and collect **a range of measurements** on each unit at **specific time intervals**

Application field	Units	Measurements
Climate studies	Locations	Meteorological variables
Economic studies	Countries	Economic indicators
Neuroscience studies	Subjects	Brain regions' signals
…	…	…

Many studies observe **multiple units** over time and collect **a range of measurements** on each unit at **specific time intervals**

Application field	Units	Measurements
Climate studies	Locations	Meteorological variables
Economic studies	Countries	Economic indicators
Neuroscience studies	Subjects	Brain regions' signals
…	…	…

Functional Magnetic Resonance Imaging (fMRI)

Electroencephalogram (EEG)

Motivating data (1)

- fMRI study conducted at UC Riverside Center for Advanced Neuroimaging
- Multiple subjects undergo 12.8 m experiment of alternating between
  - 18-sec squeeze blocks (SQ1-SQ5) when dominant hand brought to chest and ball is squeezed at maximum grip strength
  - resting blocks (RS1-RS5)

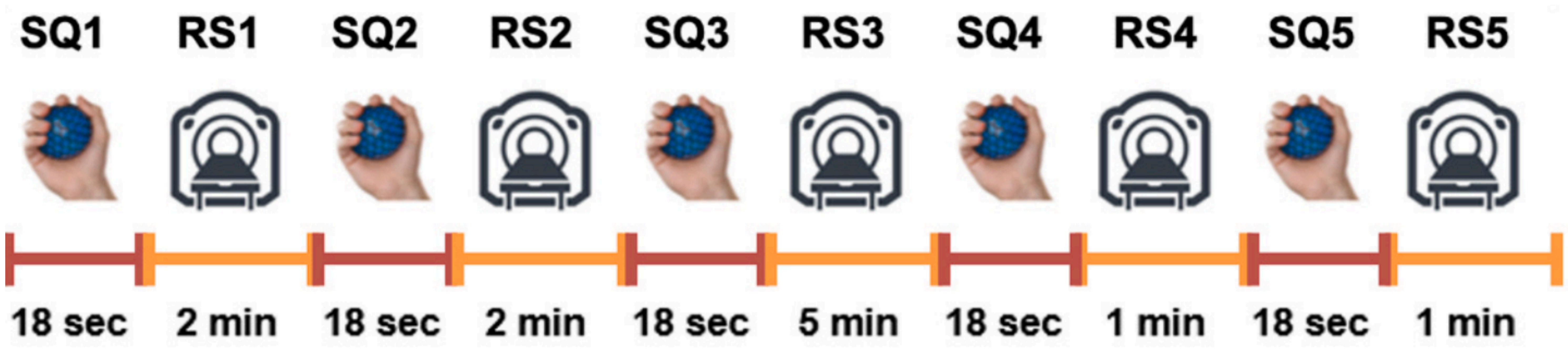


Image credit: Hussain, Sana, et al. (2023)

Motivating data (1)

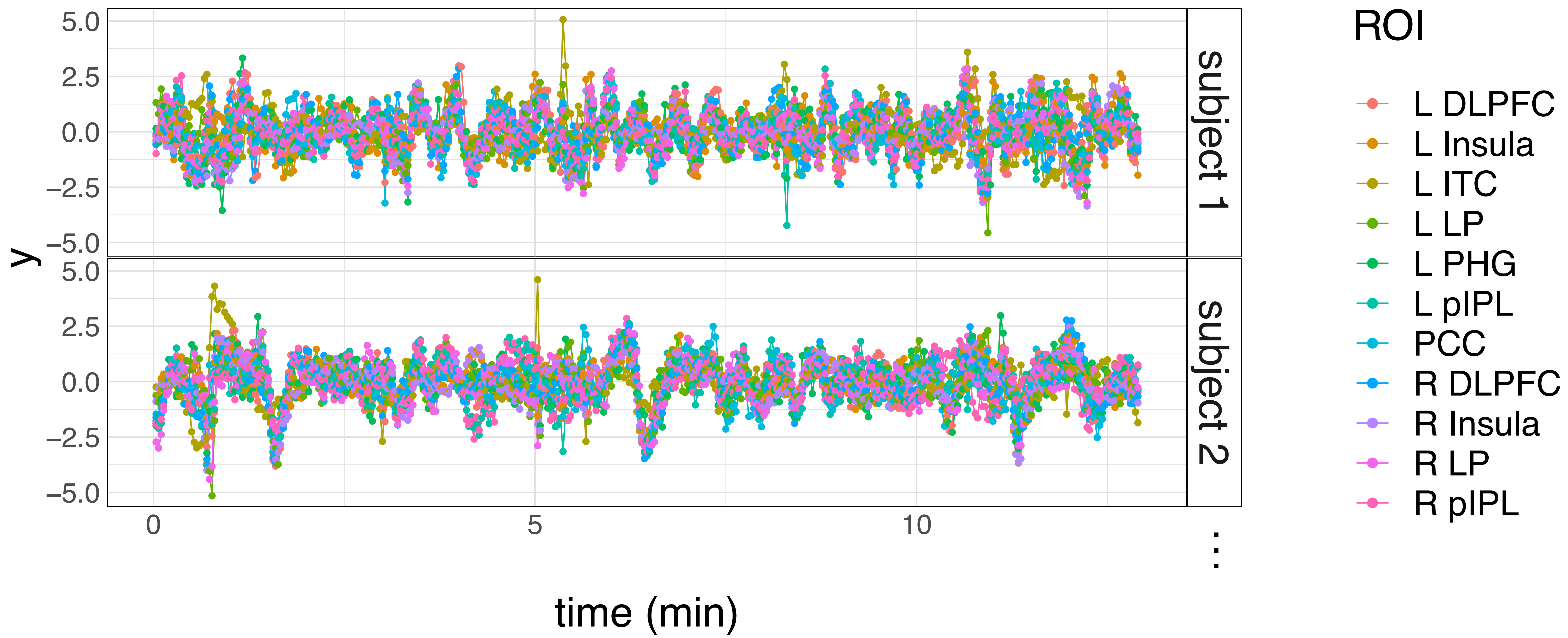
- Blood Oxygenation Level Dependent (BOLD) time-series from 23 subjects
- 11 regions of interest (ROIs), chosen from the Default Mode Network (associated with resting) or the Salience Network (allocating response to stimuli)

Y_{irt} : BOLD signal

subject $i = 1, \ldots, N$

ROI $r = 1, \ldots, R$

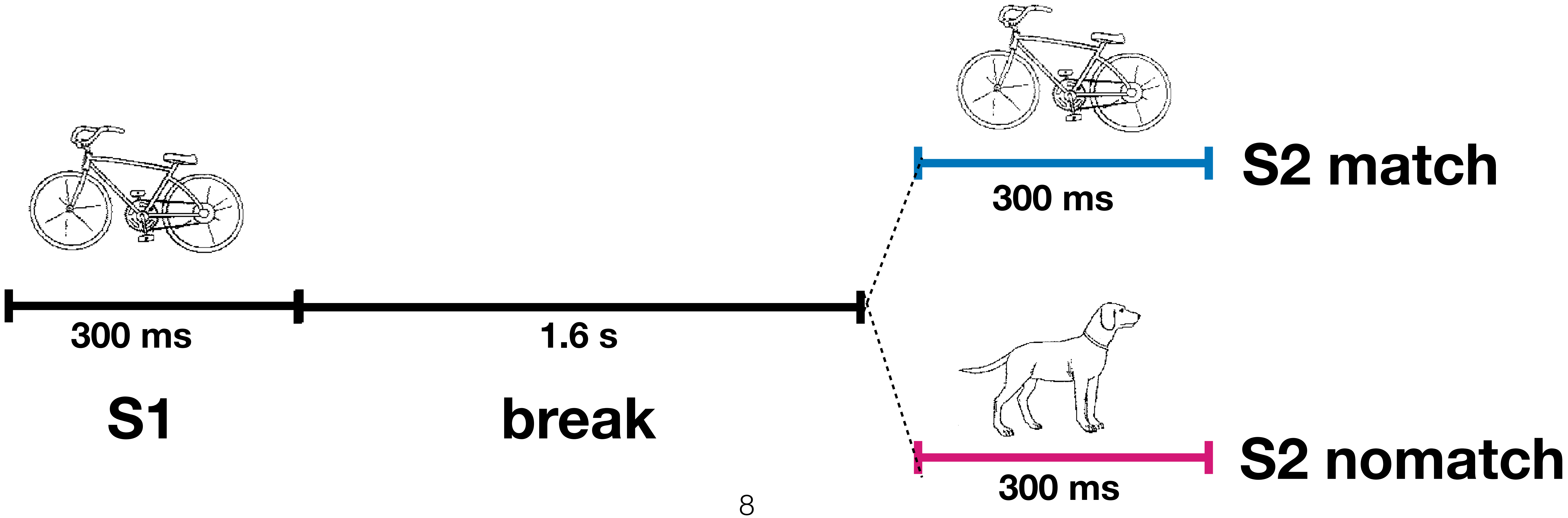
time $t = 1, \ldots, T$



Motivating data (2)

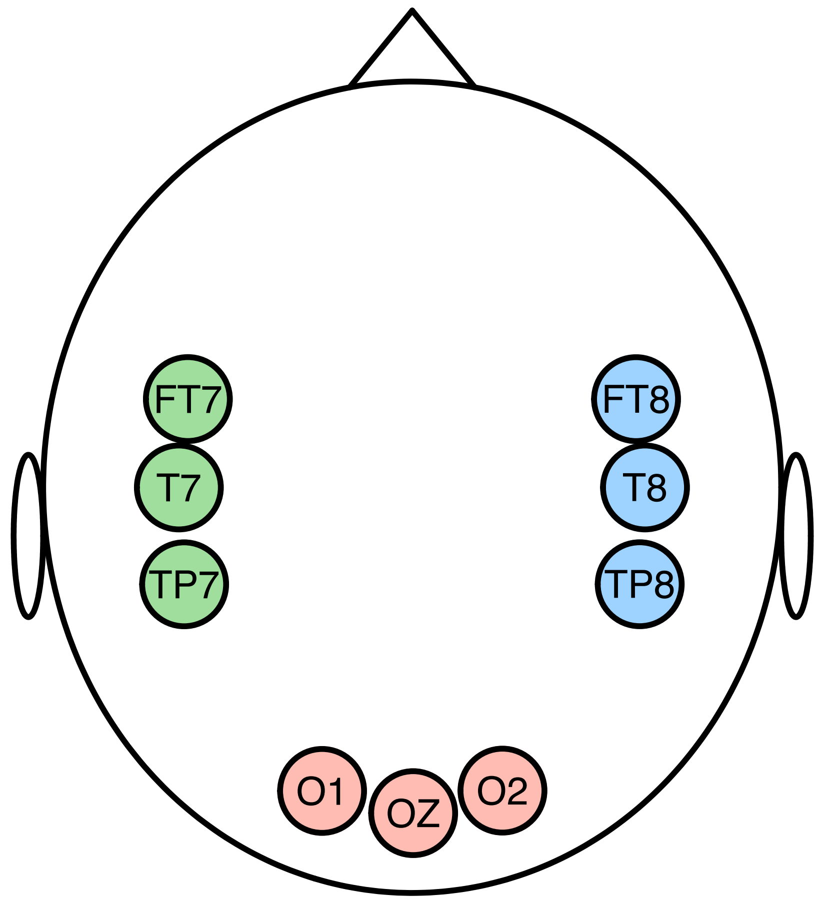
- EEG dataset publicly available on the UCI Machine Learning Repository
- Multiple subjects undergo multiple trials in which they are shown
  - a first picture (S1)
  - a picture identical to first (S2 match) or semantically different (S2 nomatch)

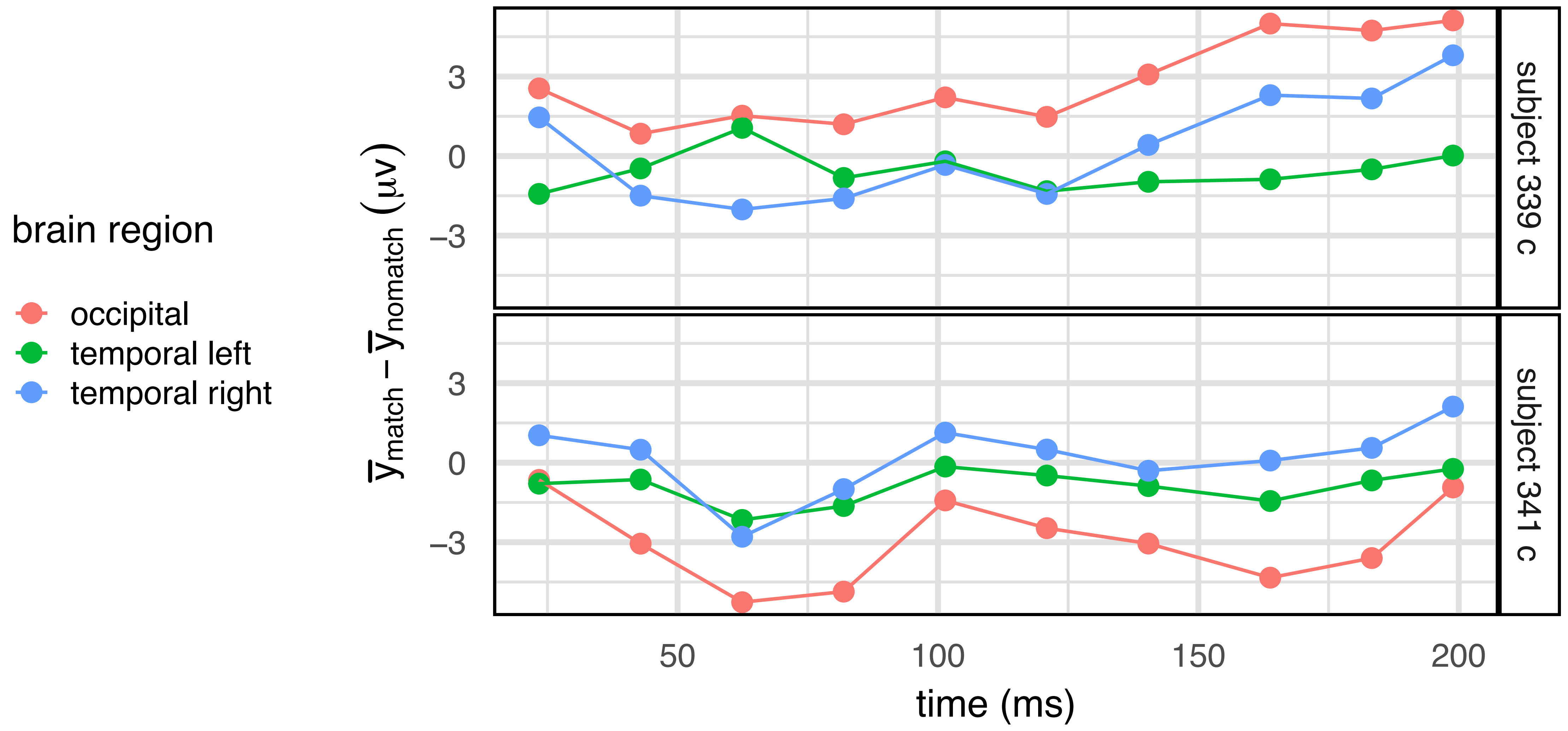
Images from Snodgrass & Vanderwart (1980)



Motivating data (2)

- Event Related Potentials (ERP) time-series (averaged across trials) from 121 subjects
- ERP averaged across electrodes in left temporal region and in right temporal region (memory) and occipital region (receptive to visual stimuli)
- ERP difference of S2 match and S2 nomatch conditions -> remove subject-specific variability





Research questions

- What patterns of brain-region activation are there in any given moment of the experiment? — *Brain region clusters*
- How do patterns of brain-region activation change during the experiment? — *Dynamic brain region clusters*
- How do patterns of brain-region activation vary across subjects? — *Subject clusters*

Within subject *Across subjects*

Prior relevant work

Prior relevant work

BICLUSTERING

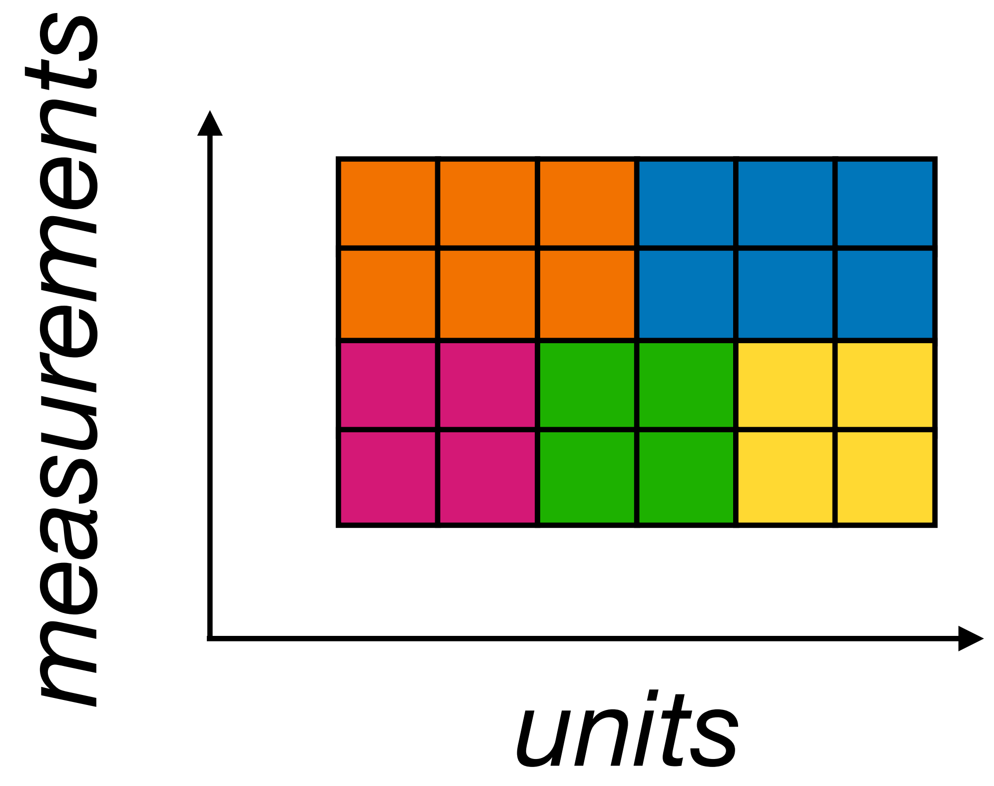
Nonparametric Bayesian Model for Local Clustering With Application to Proteomics (Lee et al. 2013, JASA)

+ *Murua and Quintana (2022, BA)*

Multiple units ✅

Multiple measurements ✅

Time ❌



Prior relevant work

BICLUSTERING

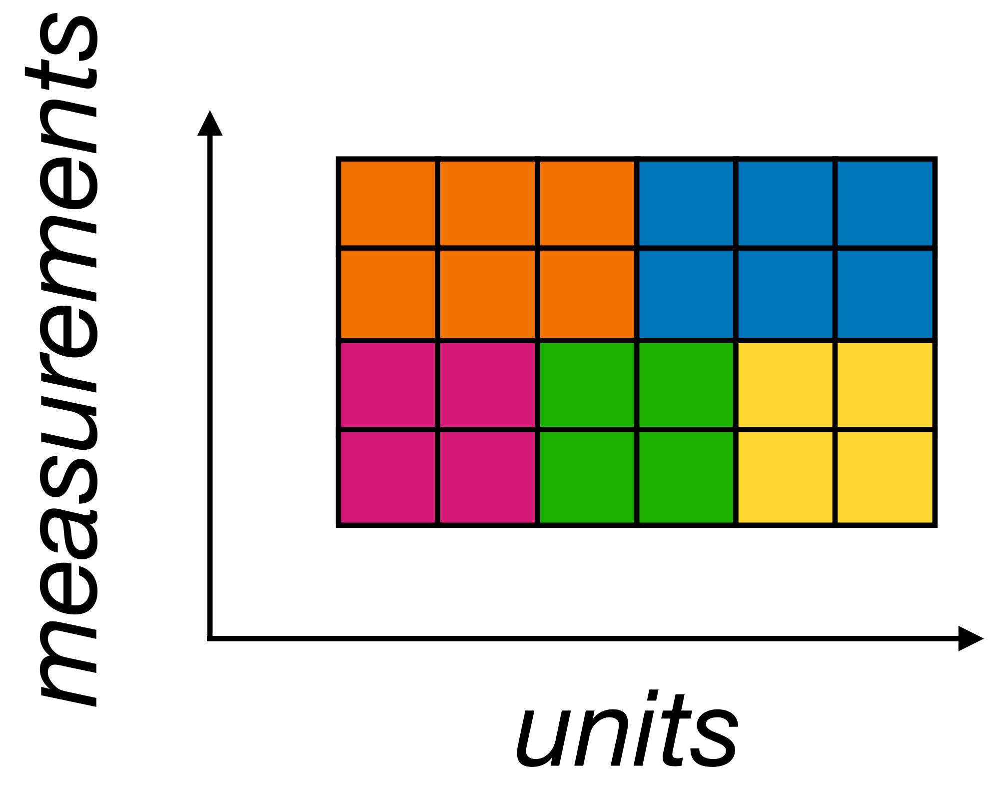
Nonparametric Bayesian Model for Local Clustering With Application to Proteomics (Lee et al. 2013, JASA)

+ *Murua and Quintana (2022, BA)*

Multiple units ✅

Multiple measurements ✅

Time ❌



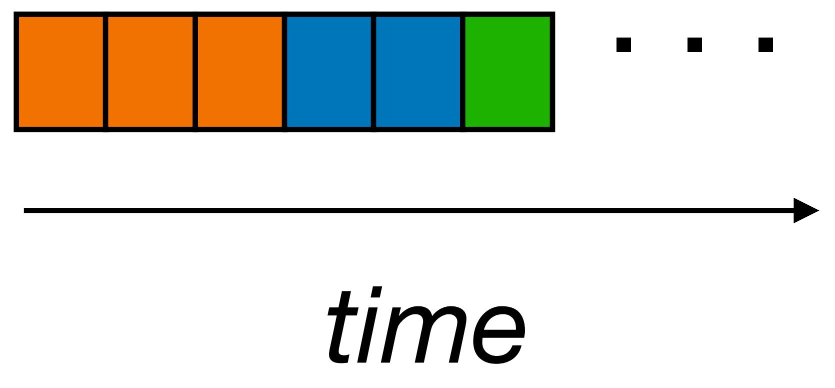
TIME-SERIES CLUSTERING

A Sticky HDP-HMM with Applications to Speaker Diarization (Fox et al. 2011, AOAS)

Multiple units ❌

Multiple measurements ❌

Time ✅



Prior relevant work

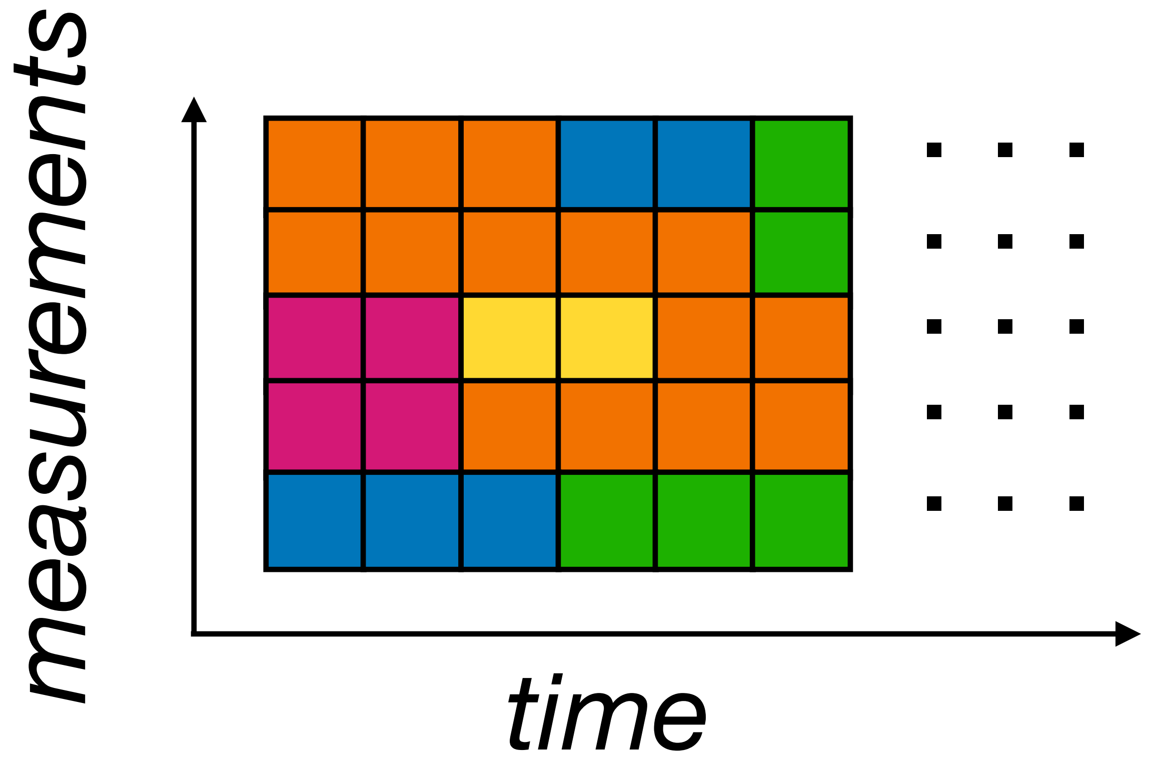
MULTIVARIATE TIME-SERIES CLUSTERING

Dependent Modeling of Temporal Sequences of Random Partitions (Page et al. 2022, JCGS)

Multiple units ❌

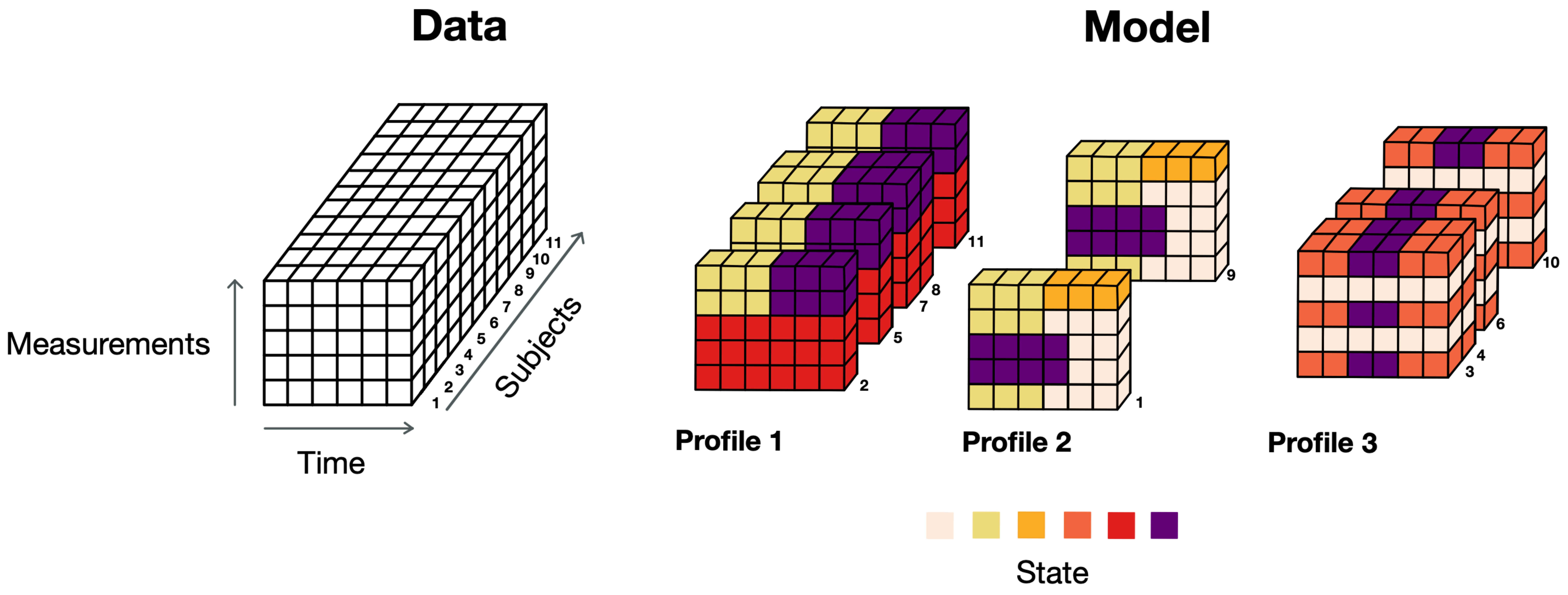
Multiple measurements ✅

Time ✅



Overview of our framework

MULTI-SUBJECT, MULTIVARIATE TIME-SERIES CLUSTERING

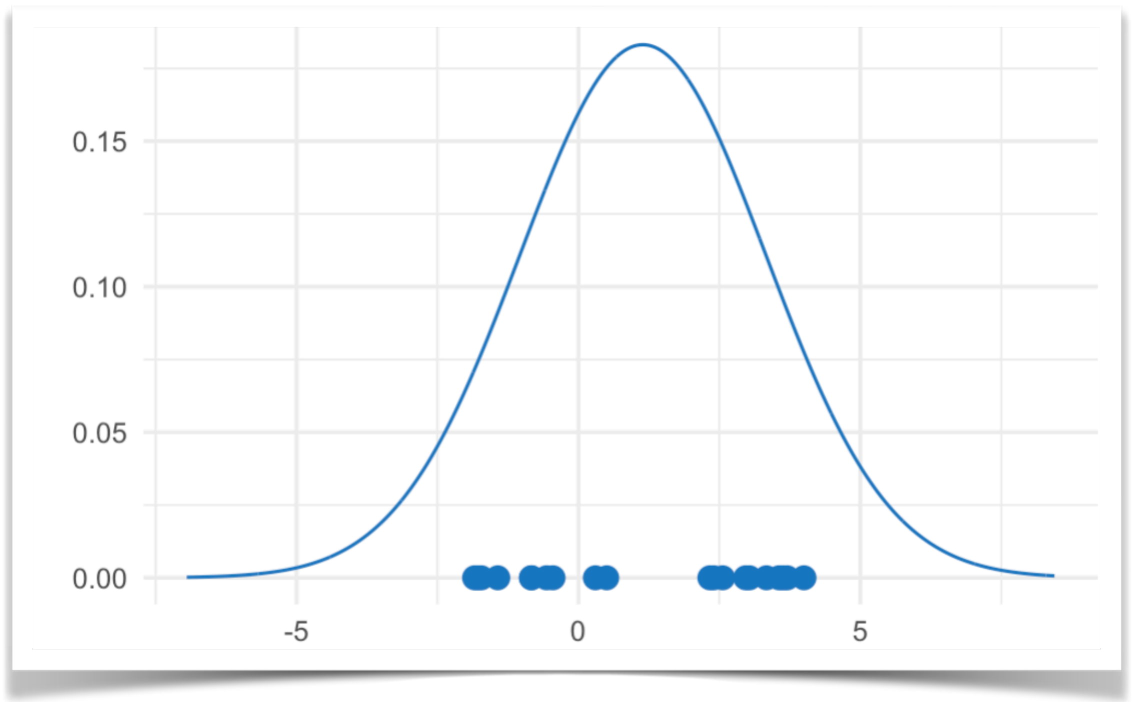


Temporal biclustering model

Clustering all brain regions for one subject and first time step

Data

Y_r for ROI $r = 1, \ldots, R$

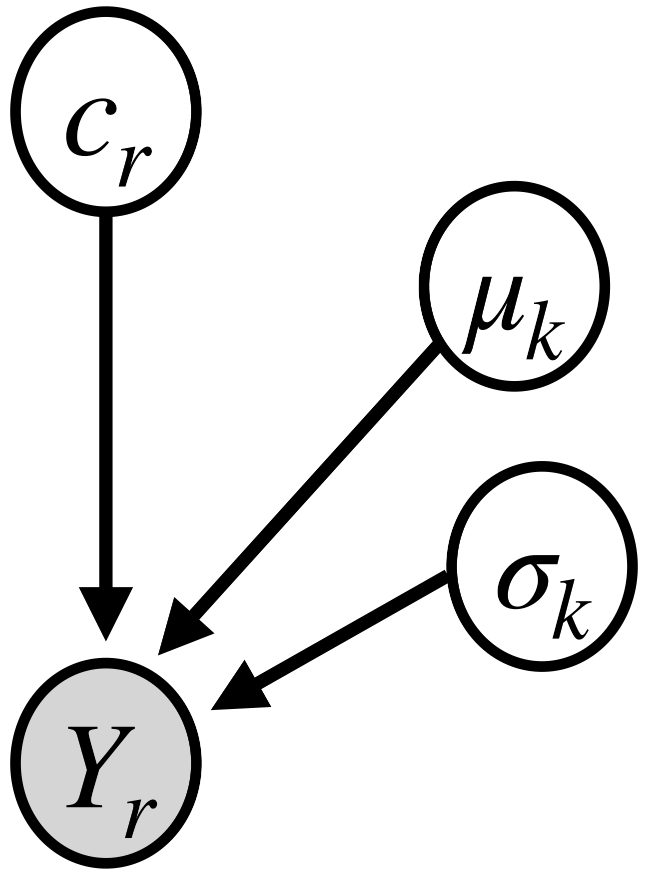


Clustering all brain regions for one subject and first time step

Data

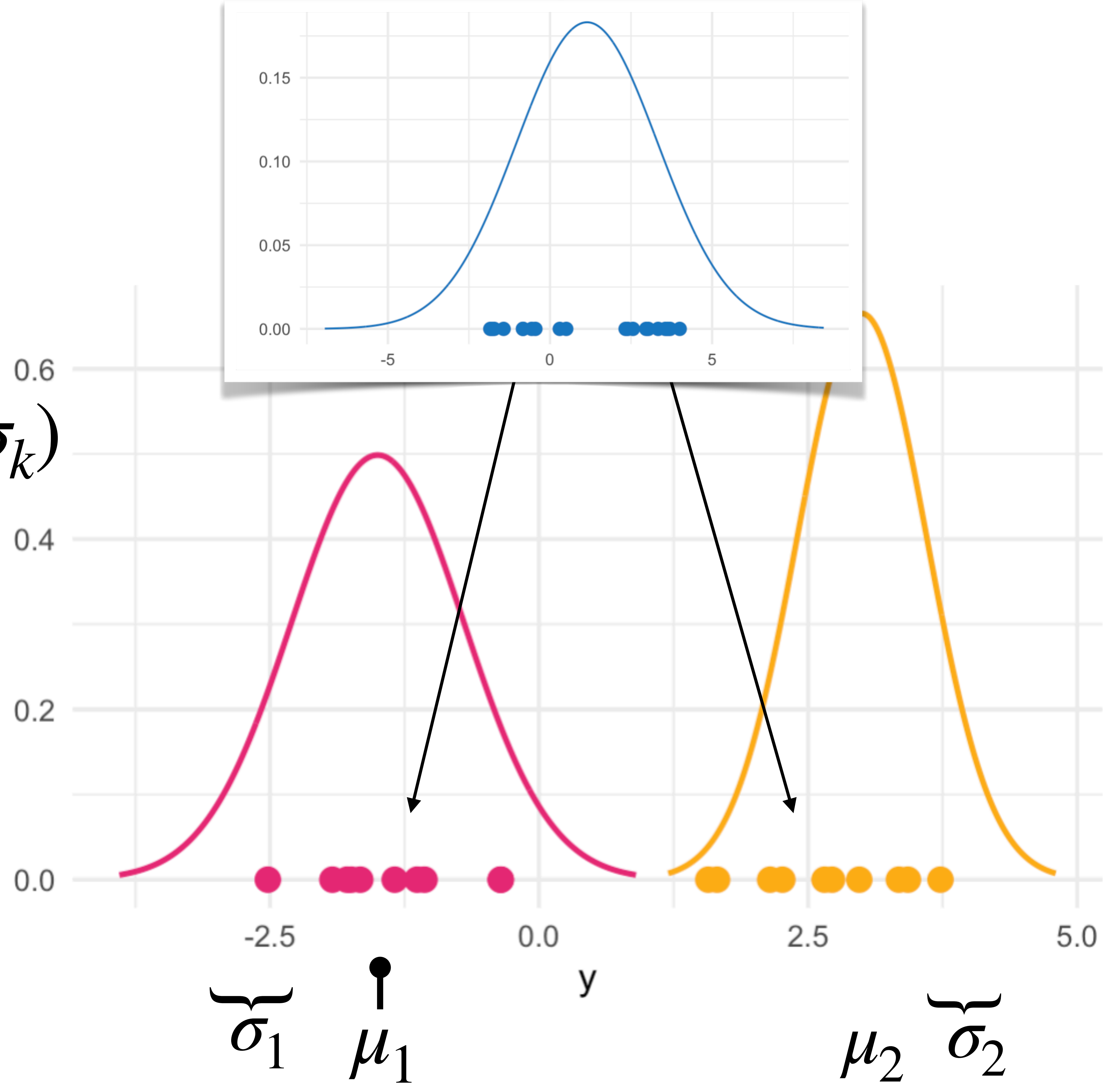
Y_r for ROI $r = 1, \ldots, R$

Statistical model



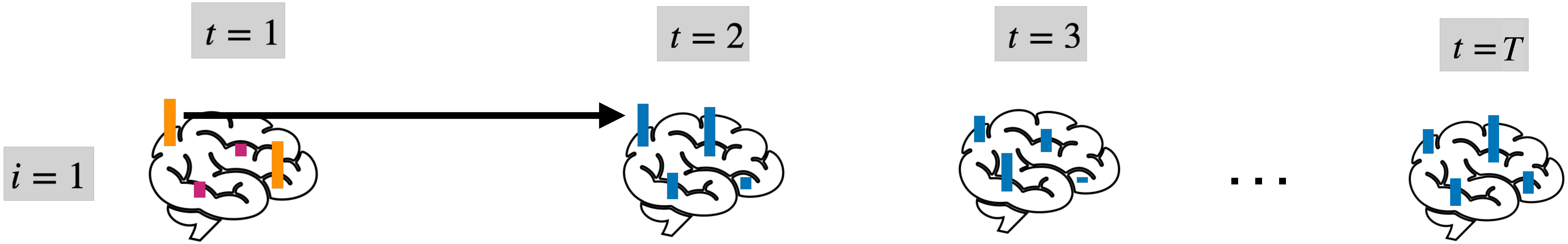
$Y_r \mid c_r = k \sim$ Student-t(μ_k, σ_k)

Inferred activation pattern



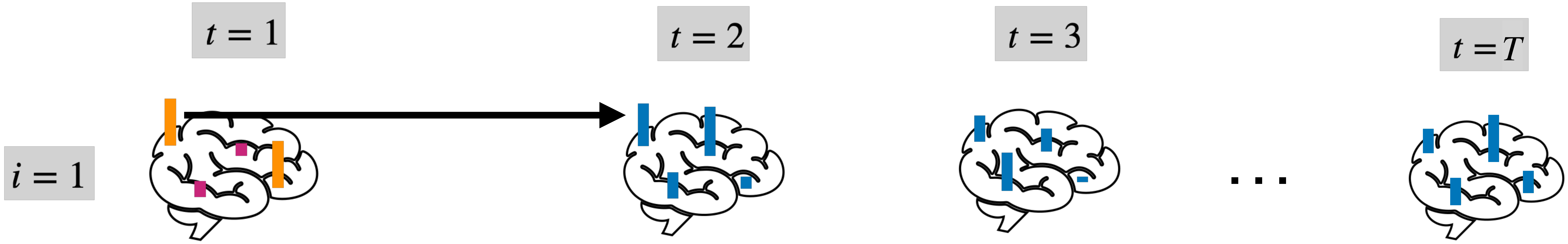
Brain region clusters (dynamic)

How can we model the way clustering changes over time?



Brain region clusters (dynamic)

How can we model the way clustering changes over time?



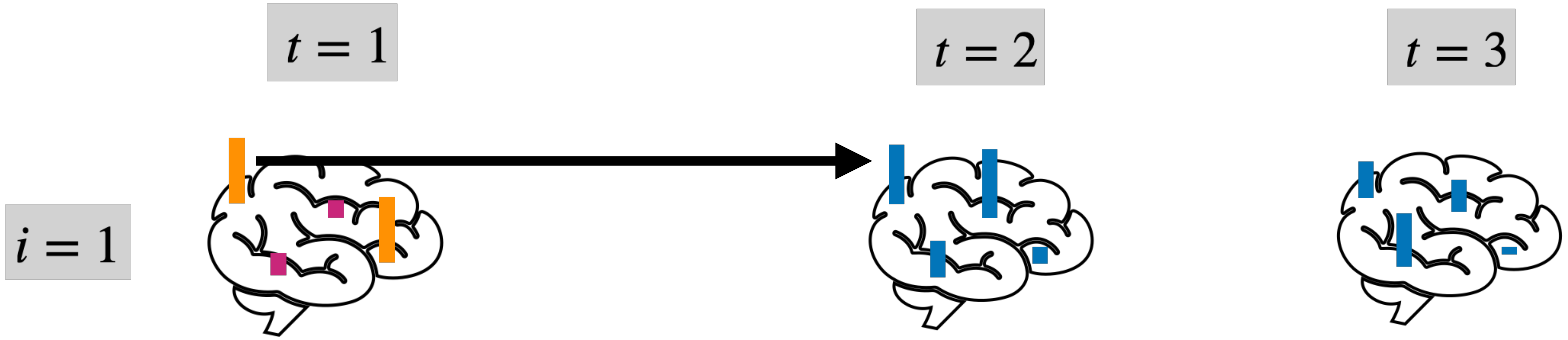
with probability α_2 : $\quad c_{r2} = c_{r1}$

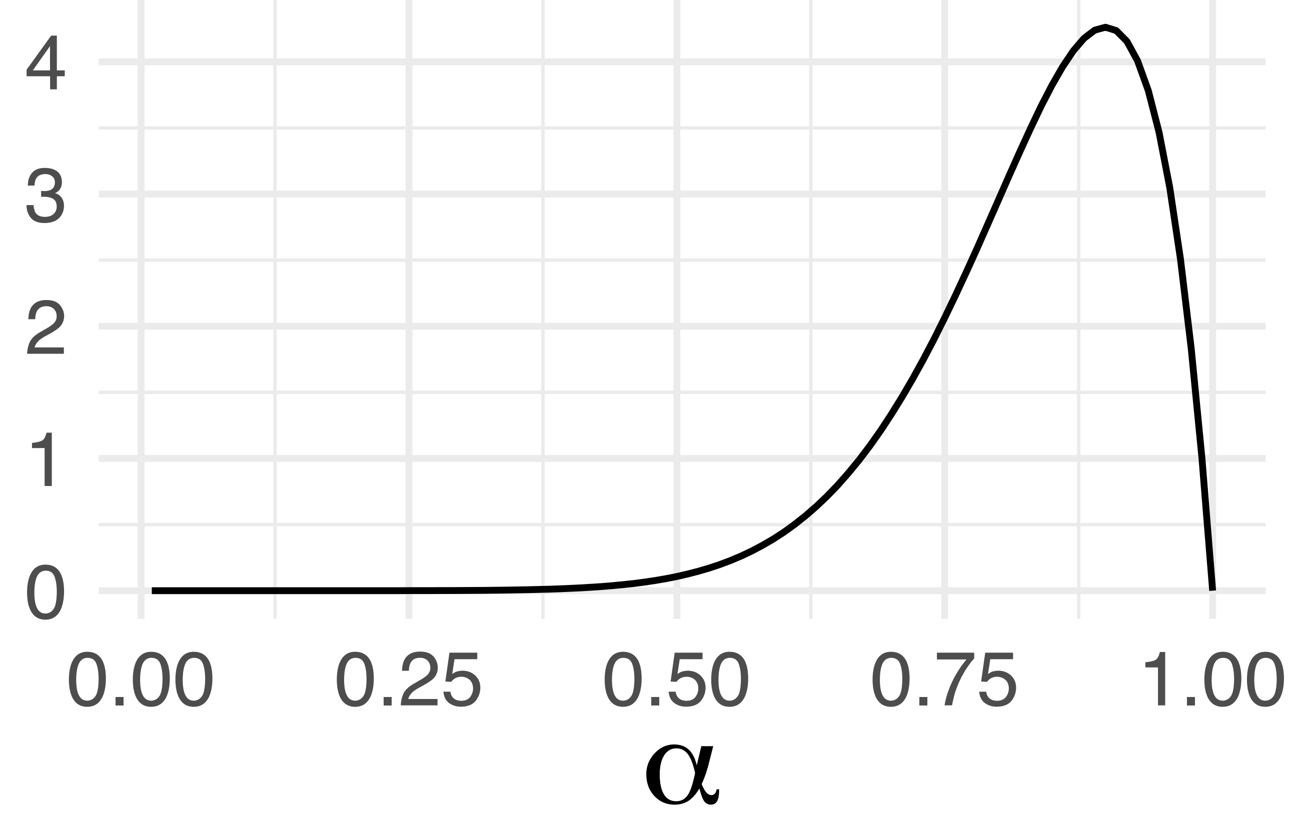
with probability $1 - \alpha_2$: $\quad c_{r2} \sim$ Categorical$(p_1, \dots, p_K)$

$$Y_{r2} \mid c_{r2} = k \sim \text{Student-t}(\mu_k, \sigma_k)$$

Brain region clusters (dynamic)

How can we model the way clustering changes over time?





with probability α_2 : $\quad c_{r2} = c_{r1}$

with probability $1 - \alpha_2$: $\quad c_{r2} \sim \text{Categorical}(p_1, \ldots, p_K)$

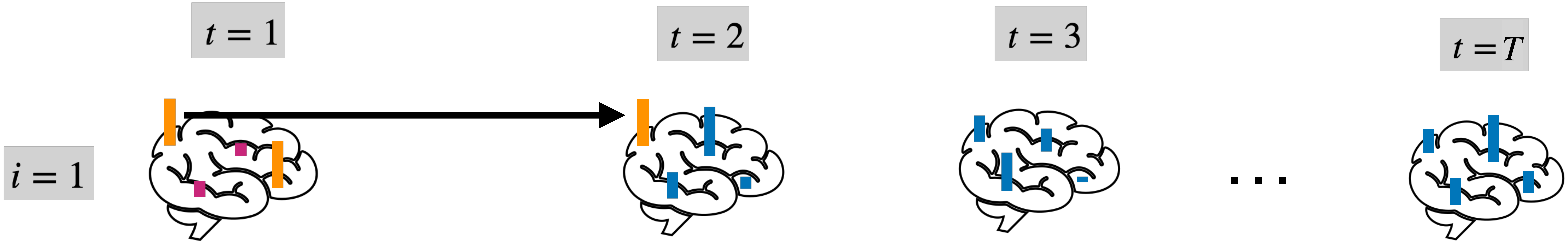
$$Y_{r2} \mid c_{r2} = k \sim \text{Student-t}(\mu_k, \sigma_k)$$

$$\alpha_t \sim \text{Beta}(10,2)$$

Large α_t encourages smooth dynamics!

Brain region clusters (dynamic)

How can we model the way clustering changes over time?



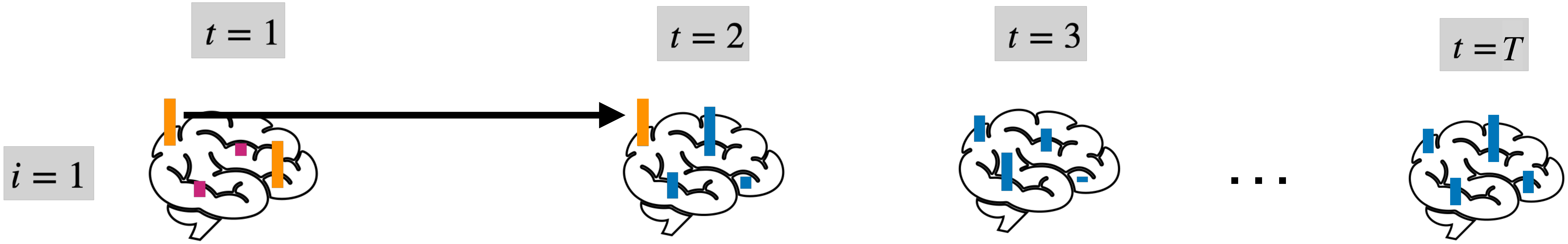
with probability α_2 : $\quad c_{r2} = c_{r1}$

with probability $1 - \alpha_2$: $\quad c_{r2} \sim$ Categorical$(p_1, \dots, p_K)$

$$Y_{r2} \mid c_{r2} = k \sim \text{Student-t}(\mu_k, \sigma_k)$$

Brain region clusters (dynamic)

How can we model the way clustering changes over time?



with probability α_2 : $\quad c_{r2} = c_{r1}$

with probability $1 - \alpha_2$: $\quad c_{r2} \sim$ Categorical$(p_1, \dots, p_K)$

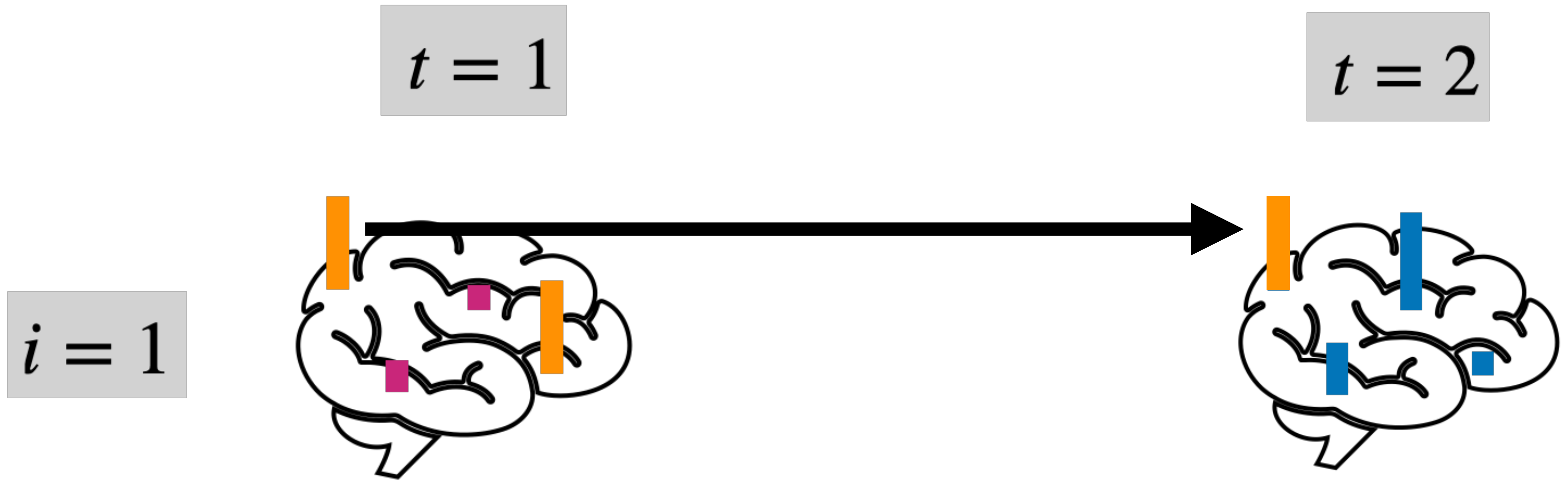
$Y_{r2} \mid c_{r2} = k \sim$ Student-t(μ_k, σ_k)

Prior on $\mathbf{p} = (p_1, \dots, p_K)$ *such that active number of clusters can:*

- *be learned from data*
- *differ across time and subjects*

Brain region clusters (dynamic)

How can we model the way clustering changes over time?



$$\boldsymbol{p} \mid \boldsymbol{p}_0 \sim \text{Dirichlet}(\phi\, \omega_{01}, \ldots, \phi\, p_{0K}),$$

$$\boldsymbol{p}_0 \mid \eta \sim \text{Dirichlet}\left(\frac{\eta}{K}, \ldots, \frac{\eta}{K}\right)$$

$$\eta \sim \text{Gamma}(d_1, d_2)$$

[Finite approx. of Hierarchical Dirichlet Process]

(Malsiner-Walli et al. 2016 Stat. Comput.)

with probability α_2 : $\quad c_{r2} = c_{r1}$

with probability $1 - \alpha_2$: $\quad c_{r2} \sim \text{Categorical}(p_1, \ldots, p_K)$

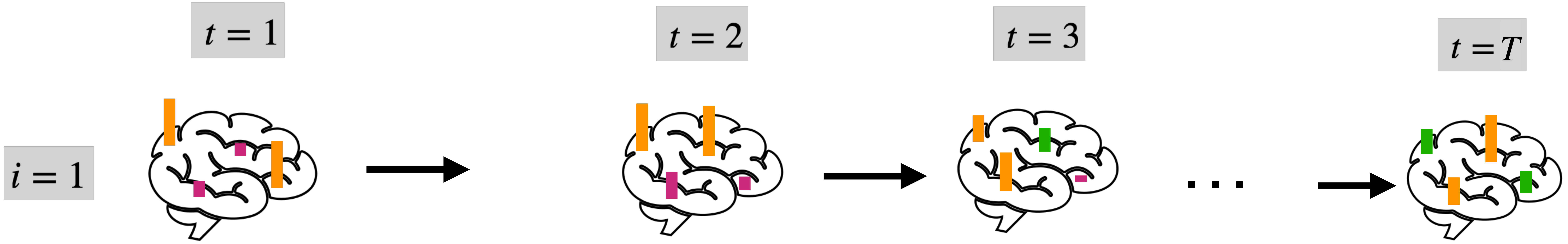
$$Y_{r2} \mid c_{r2} = k \sim \text{Student-t}(\mu_k, \sigma_k)$$

Prior on $\mathbf{p} = (p_1, \ldots, p_K)$ such that active number of clusters can:

- *be learned from data*
- *differ across time and subjects*

Brain region clusters (dynamic)

How can we model the way clustering changes over time?



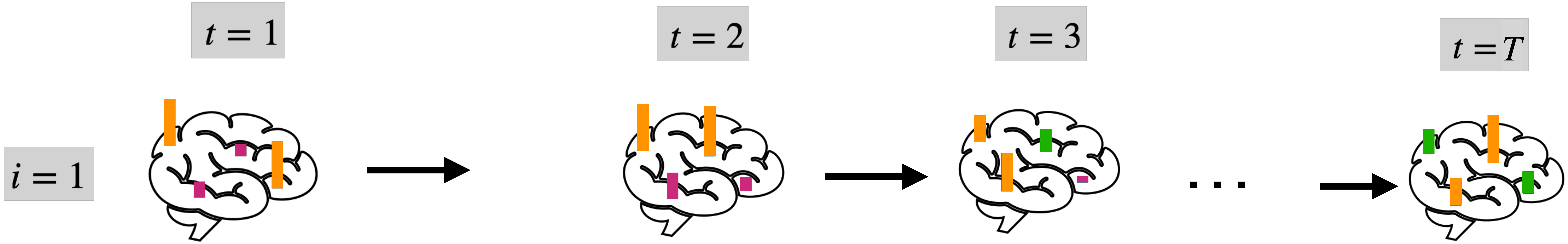
with probability α_T : $\quad c_{rT} = c_{r,T-1}$

with probability $1 - \alpha_T$: $\quad c_{rT} \sim$ Categorical$(p_1, \ldots, p_K)$

$$Y_{rT} \mid c_{rT} = k \sim \text{Student-t}(\mu_k, \sigma_k)$$

Brain region clusters (dynamic)

How can we model the way clustering changes over time?



with probability α_T : $\quad c_{rT} = c_{r,T-1}$

with probability $1 - \alpha_T$: $\quad c_{rT} \sim$ Categorical$(p_1, \ldots, p_K)$

$$Y_{rT} \mid c_{rT} = k \sim \text{Student-t}(\mu_k, \sigma_k)$$

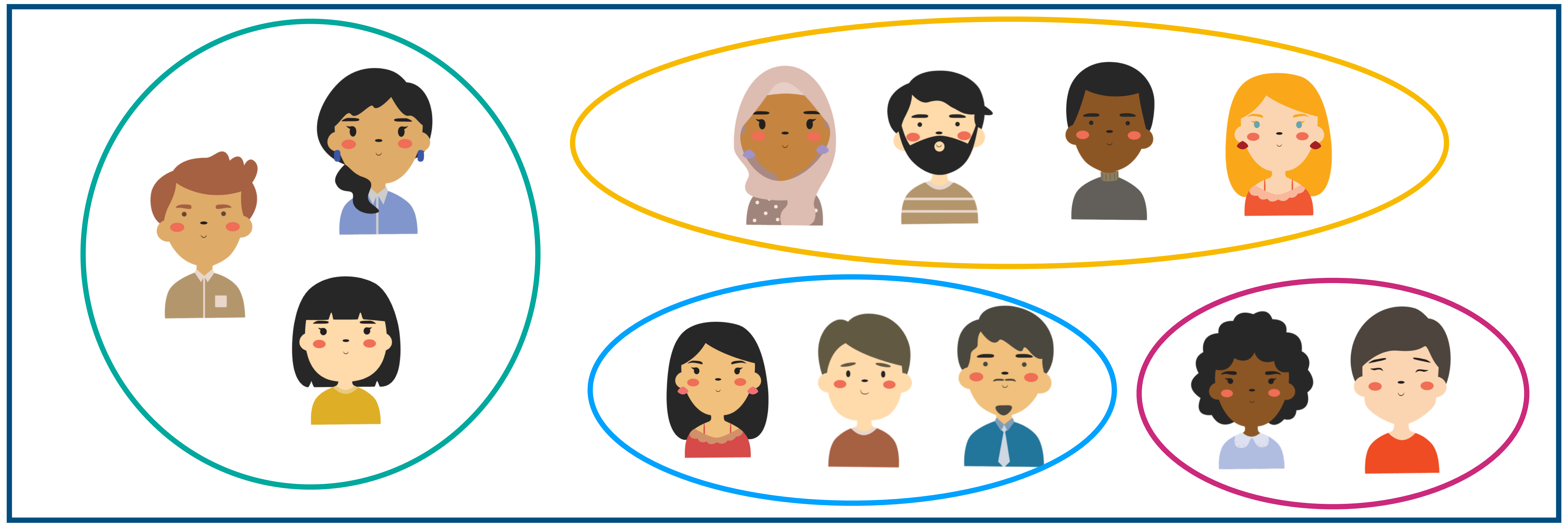
Adapted from the **Temporal Random Partition Model (Page et al. 2022, JCGS)**

Subject clusters

So far: brain regions clusters, over time.. for a single subject!

What about clusters of subjects?

Data

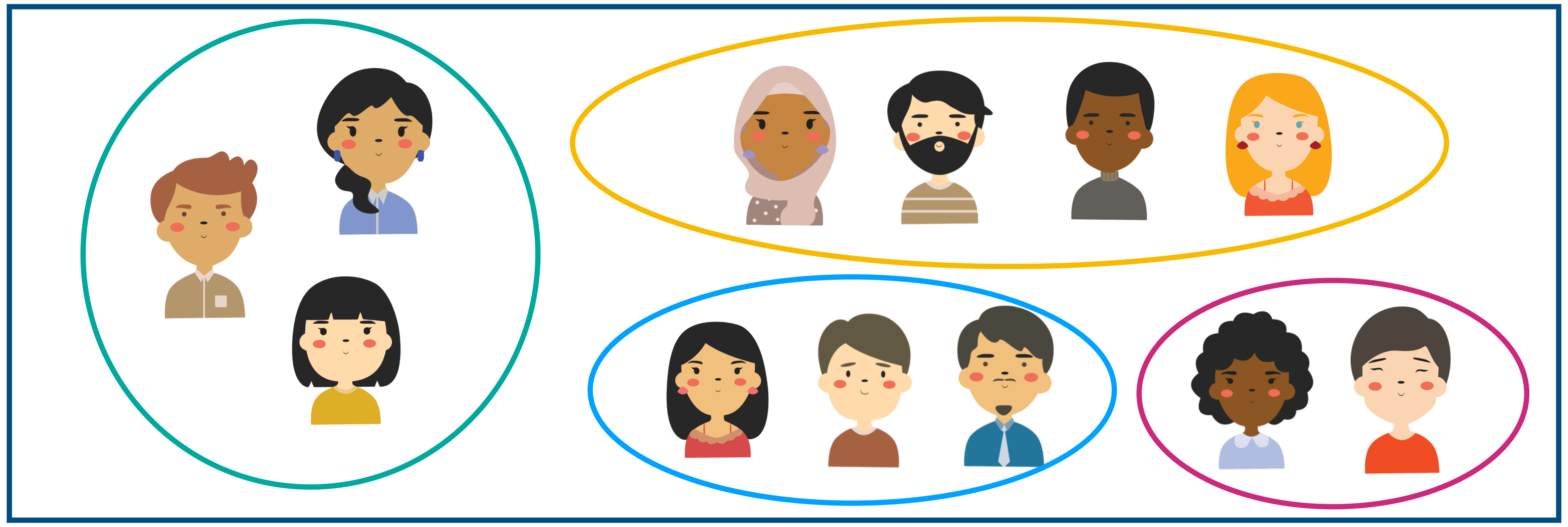


Subject clusters

So far: brain regions clusters, over time.. for a single subject!

What about clusters of subjects?

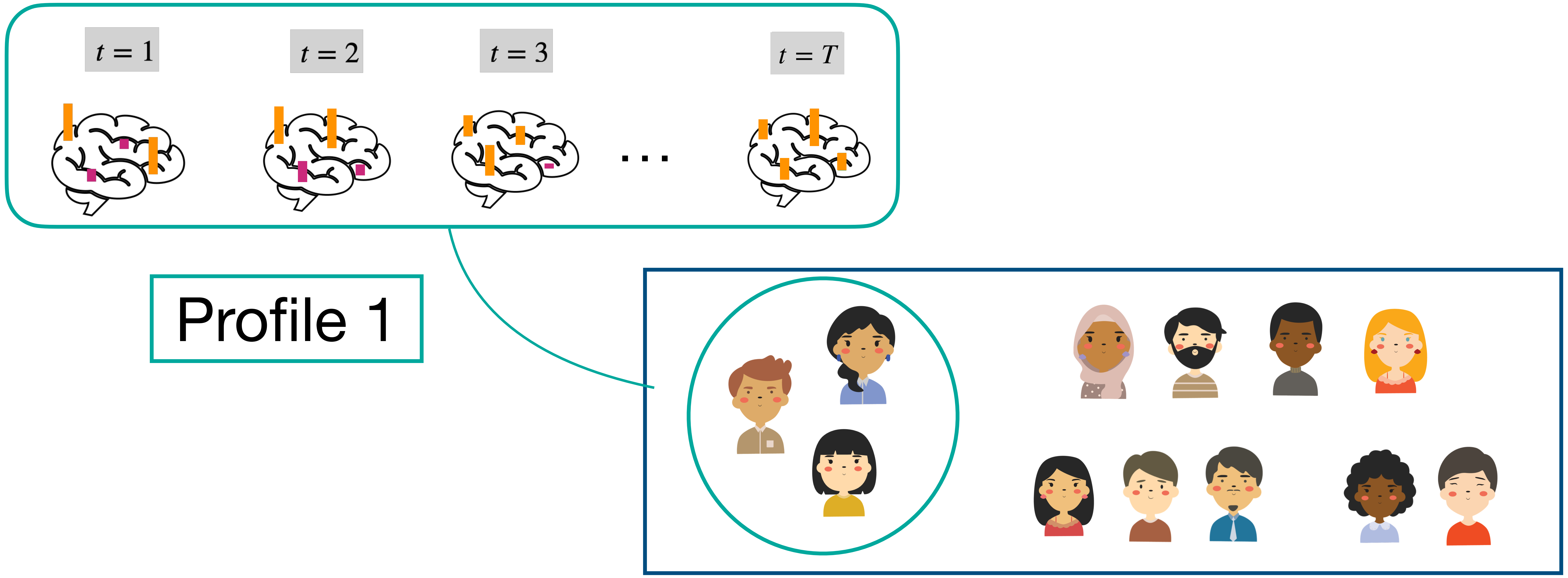
Data + **Inference on profiles** →



💡 **Profile**: specific **sequence of brain-region clusters** during the experiment

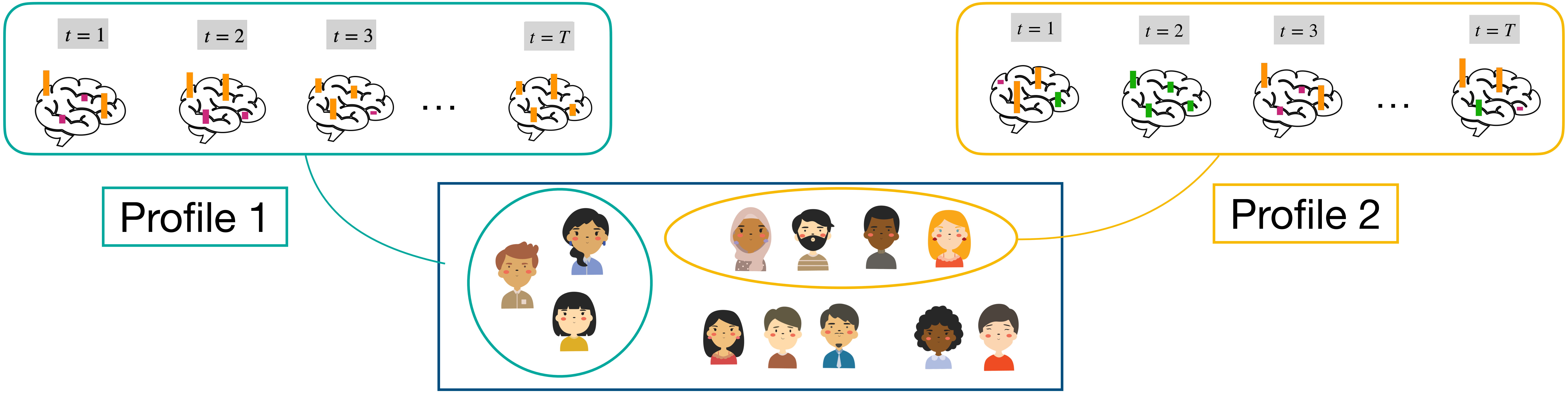
Subject clusters

💡 **Profile**: specific **sequence of brain-region clusters** during the experiment



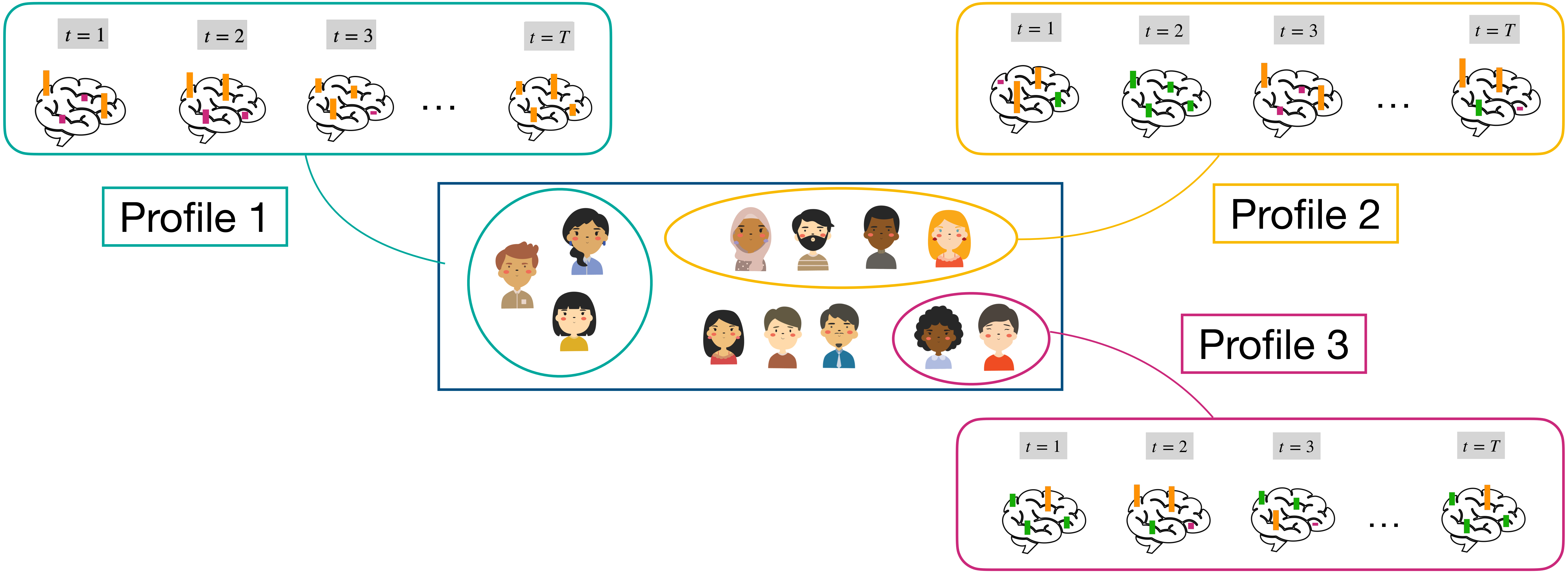
Subject clusters

💡 **Profile**: specific **sequence of brain-region clusters** during the experiment



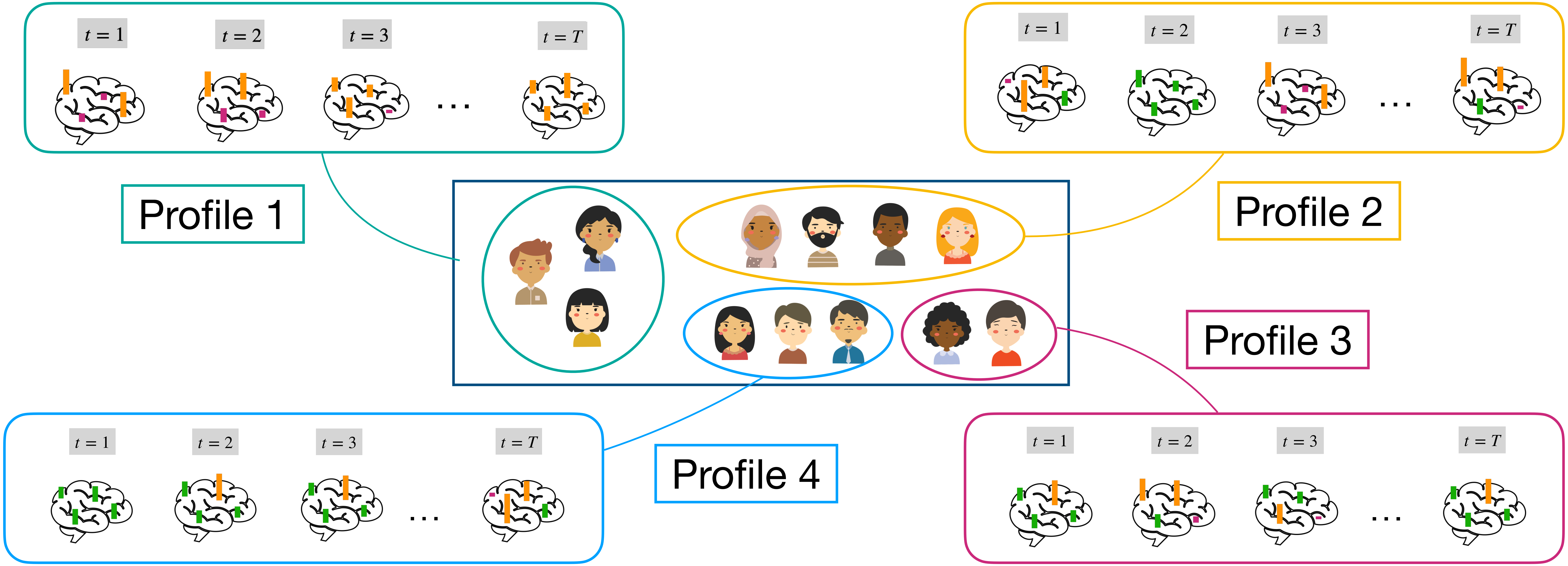
Subject clusters

💡 **Profile**: specific **sequence of brain-region clusters** during the experiment



Subject clusters

💡 **Profile**: specific **sequence of brain-region clusters** during the experiment



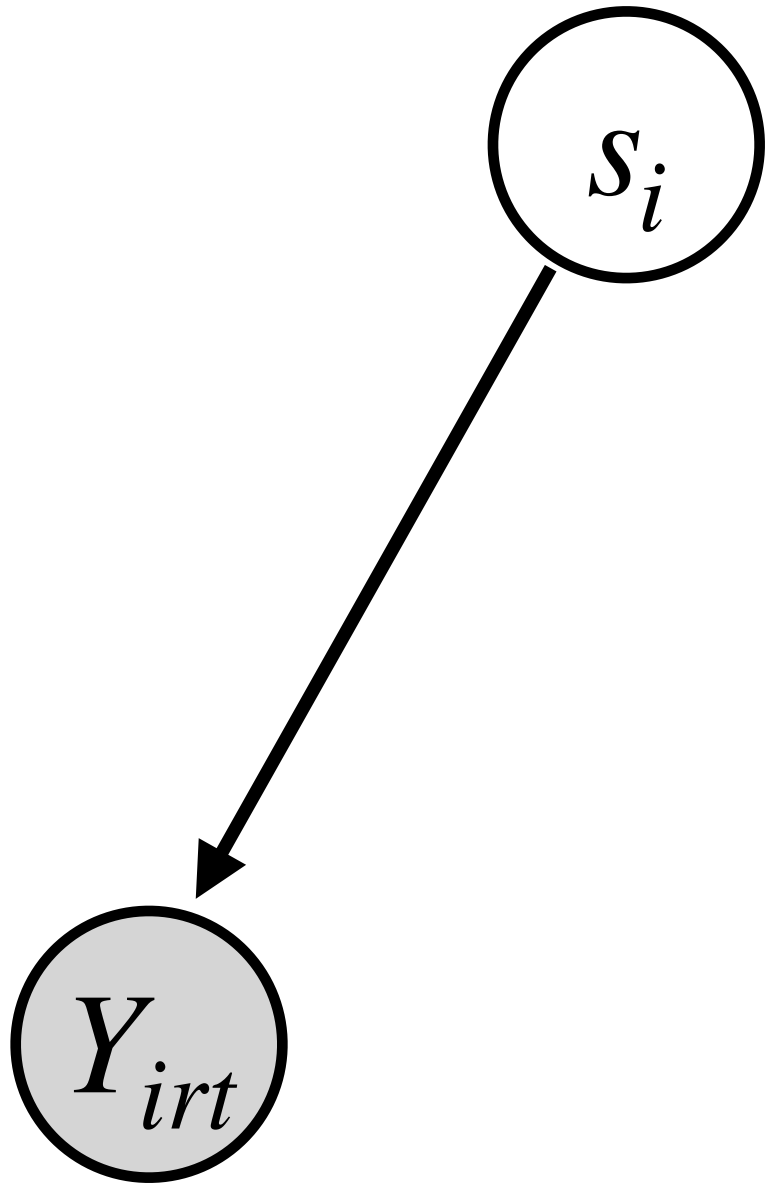
Full model

For subject i, brain region r and time t

Y_{irt}

Full model

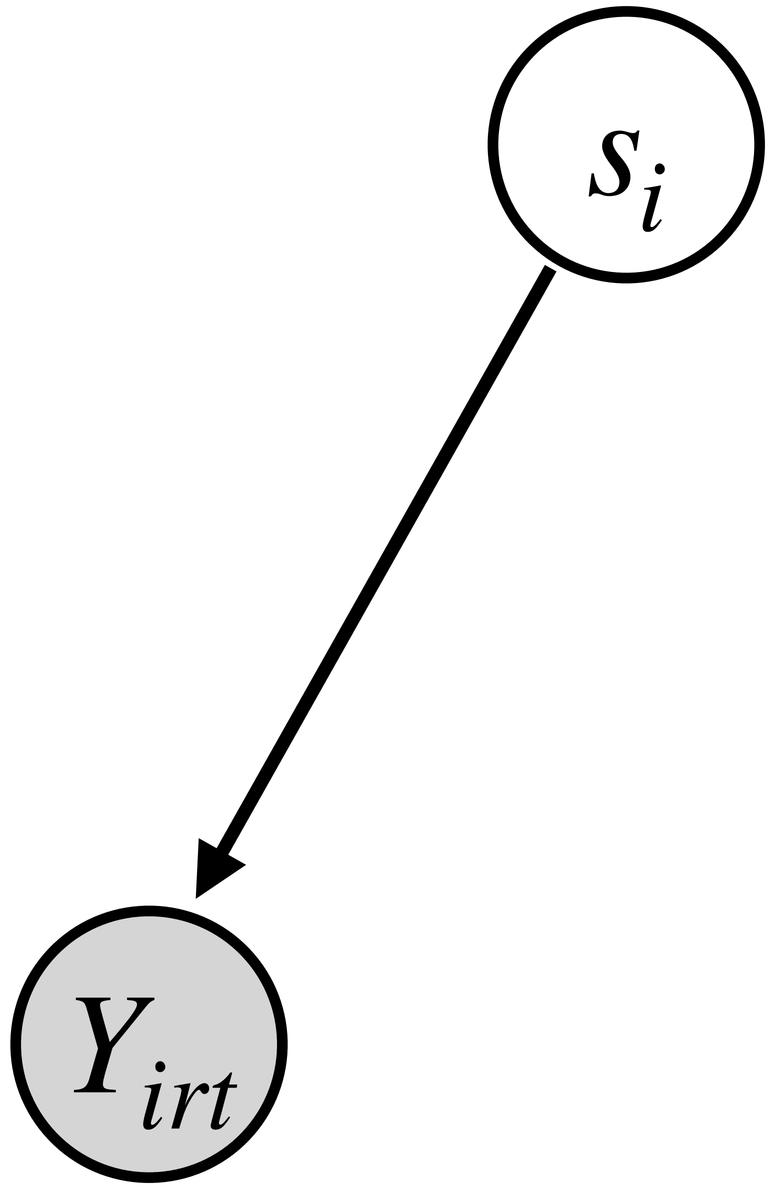
For subject i, brain region r and time t



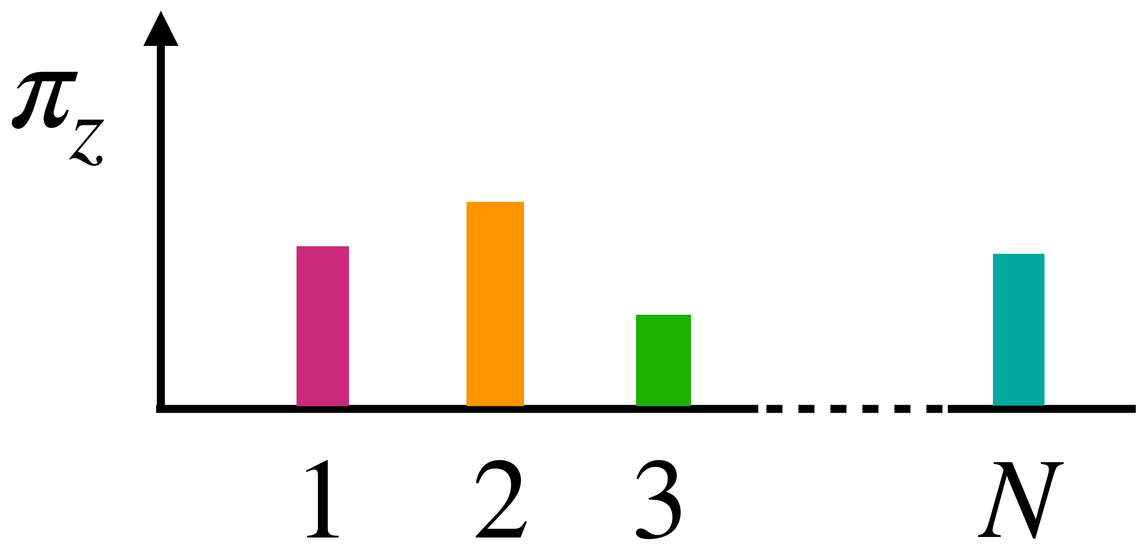
Profile of subject i

Full model

For subject i, brain region r and time t



Profile of subject i



$$s_i \sim \text{Categorical}(\pi_1, \ldots, \pi_N)$$

$$\pi \sim \text{Dirichlet}\left(\frac{\varepsilon}{N}, \ldots, \frac{\varepsilon}{N}\right)$$

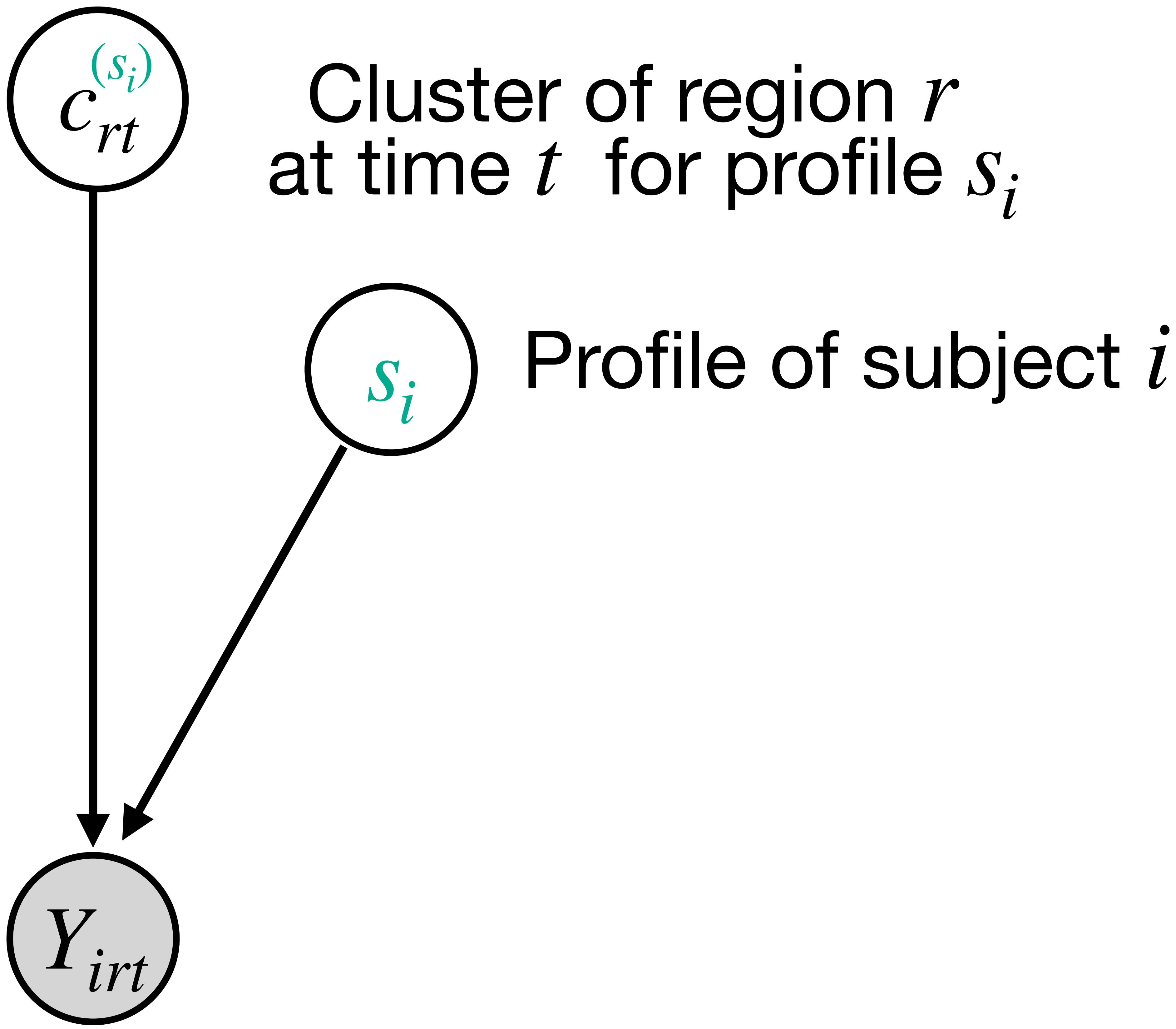
$$\varepsilon \sim \text{Gamma}(b_1, b_2)$$

Sparse Finite Mixture

[Finite approximation of Dirichlet Process]

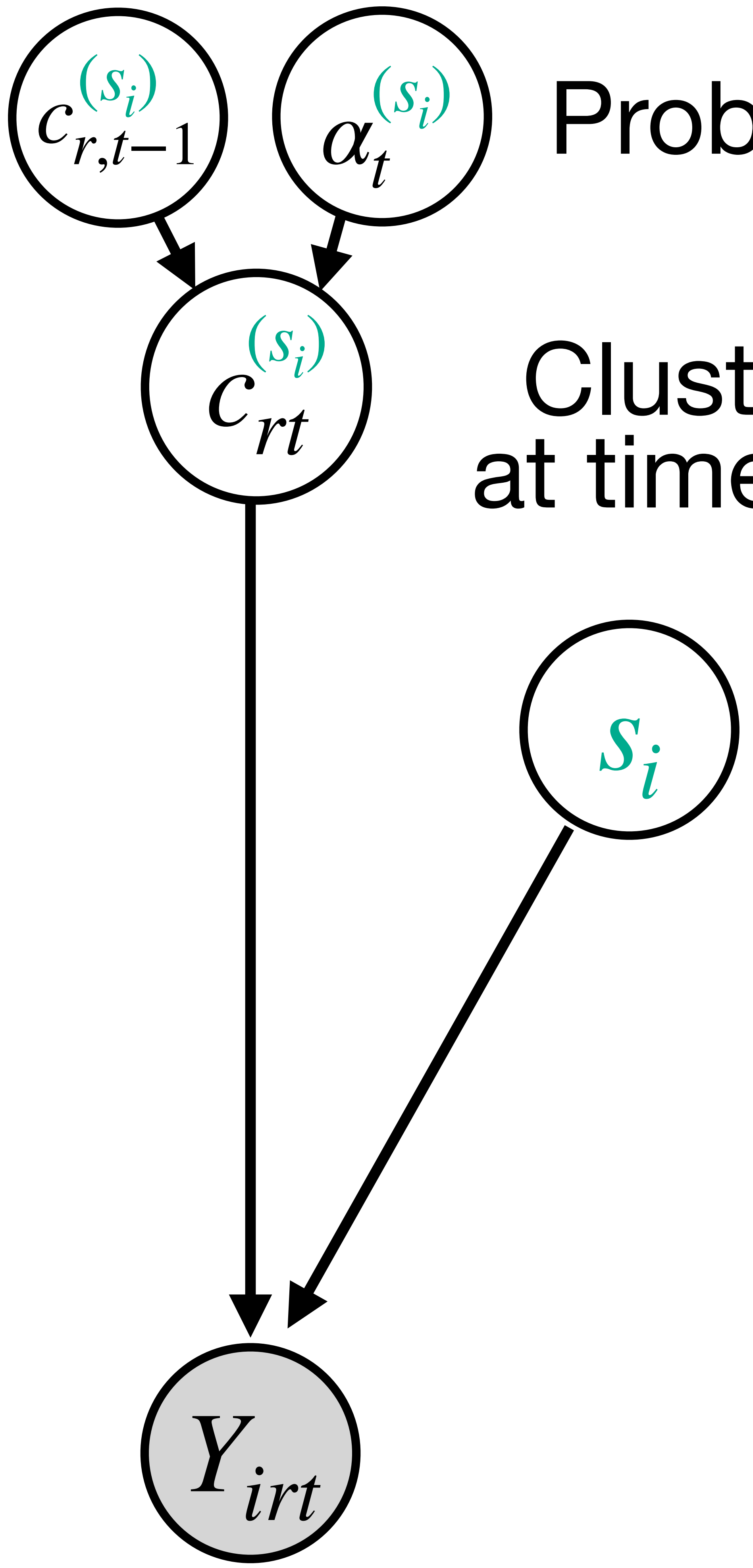
Full model

For subject i, brain region r and time t



Full model

For subject i, brain region r and time t



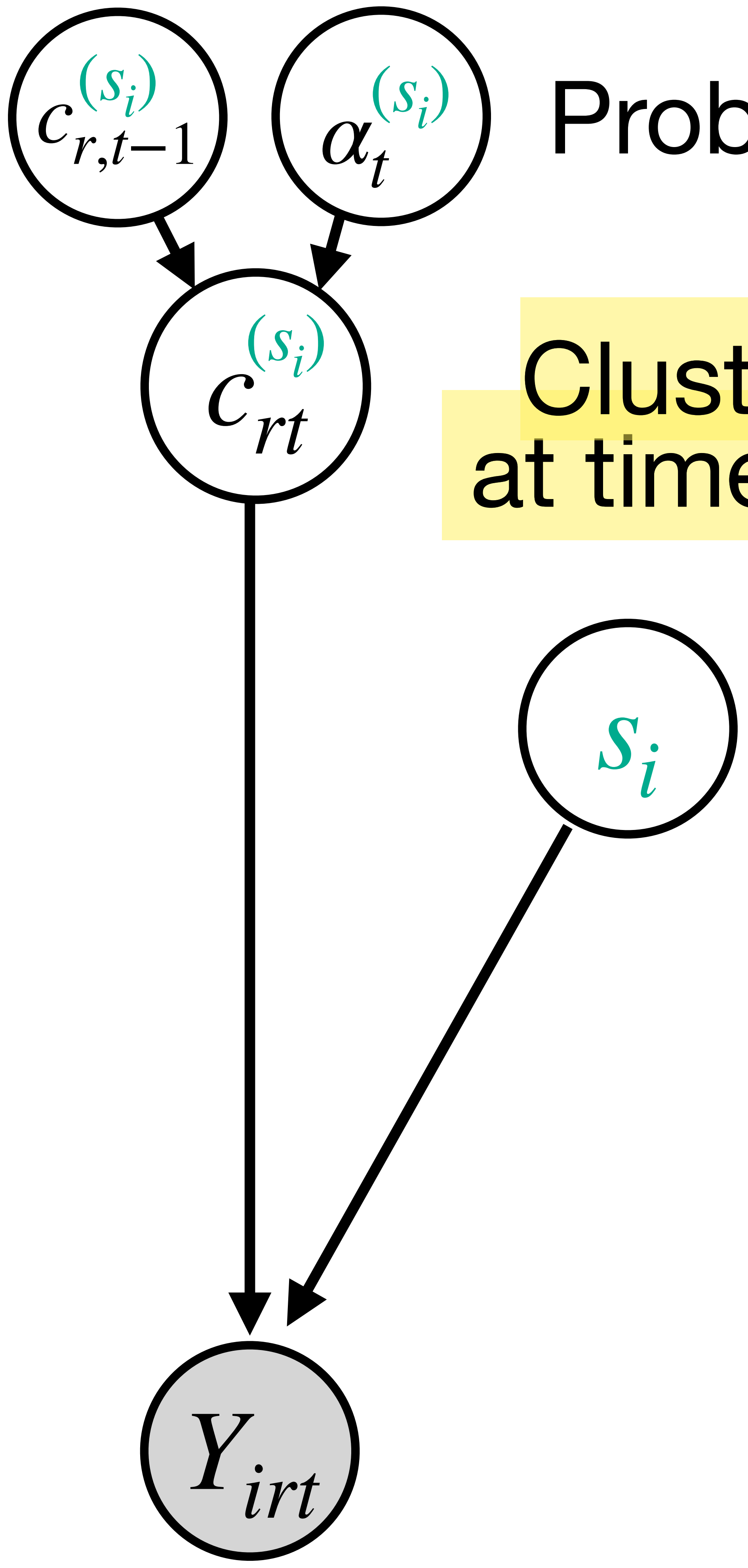
Probability that a cluster persists over time

Cluster of region r at time t for profile s_i

Profile of subject i

Full model

For subject i, brain region r and time t



Probability that a cluster persists over time

Cluster of region r at time t for profile s_i

$$\begin{cases} \text{with probability } \alpha_t^{(s_i)} & c_{r,t}^{(s_i)} = c_{r,t-1}^{(s_i)} \\ \text{with probability } 1 - \alpha_t^{(s_i)} & c_{r,t}^{(s_i)} \sim \text{Categ}\left(p_1^{(s_i)}, \ldots, p_K^{(s_i)}\right) \end{cases}$$

Profile of subject i

$$\boldsymbol{p}^{(s_i)} \mid \boldsymbol{p}_0 \sim \text{Dirichlet}(\phi\, \omega_{01}, \ldots, \phi\, p_{0K}),$$

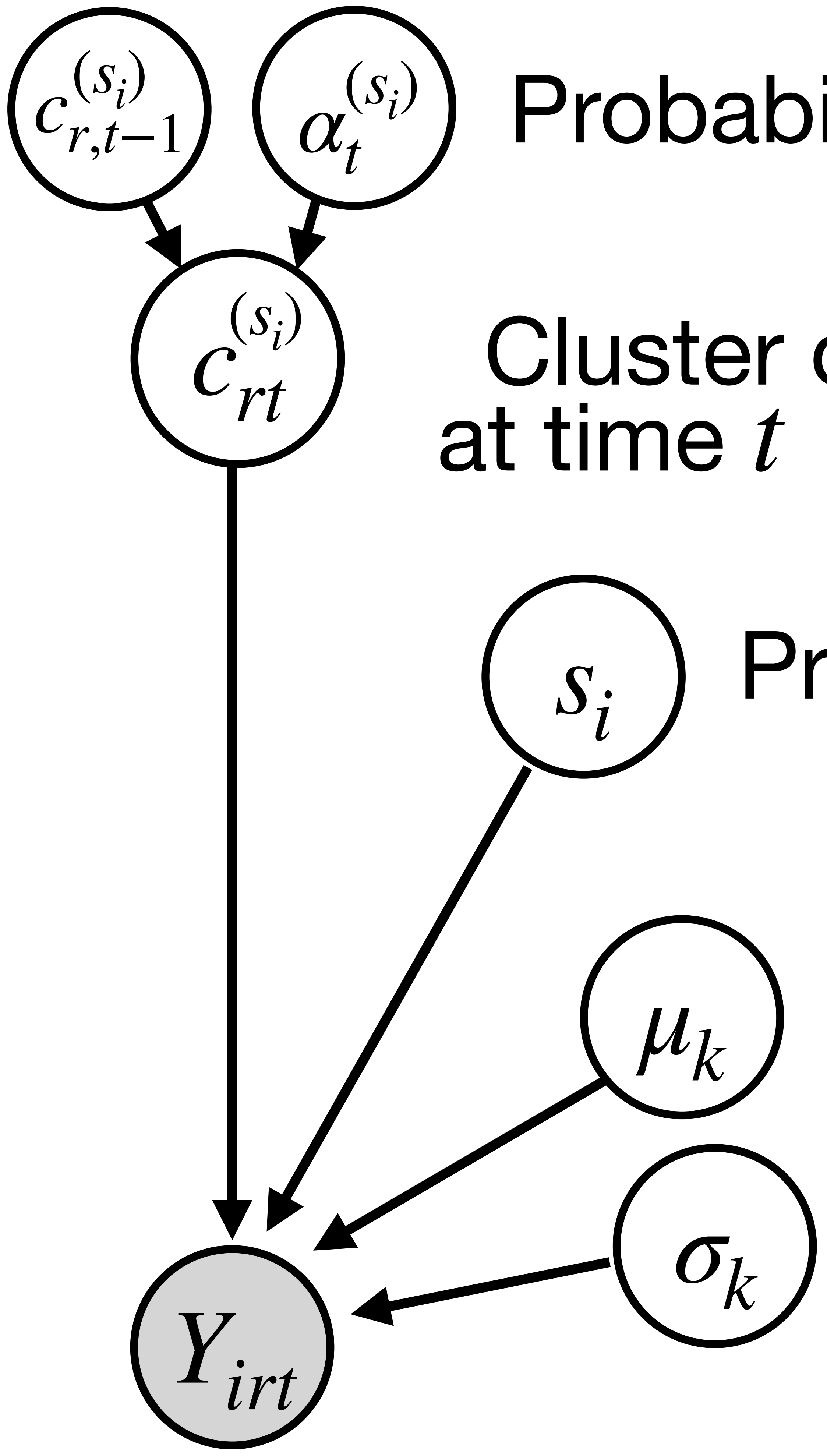
$$\boldsymbol{p}_0 \mid \eta \sim \text{Dirichlet}\left(\frac{\eta}{K}, \ldots, \frac{\eta}{K}\right)$$

$$\eta \sim \text{Gamma}(d_1, d_2)$$

[Finite approx. of Hierarchical Dirichlet Process]

Full model

For subject i, brain region r and time t



Probability that a cluster persists over time

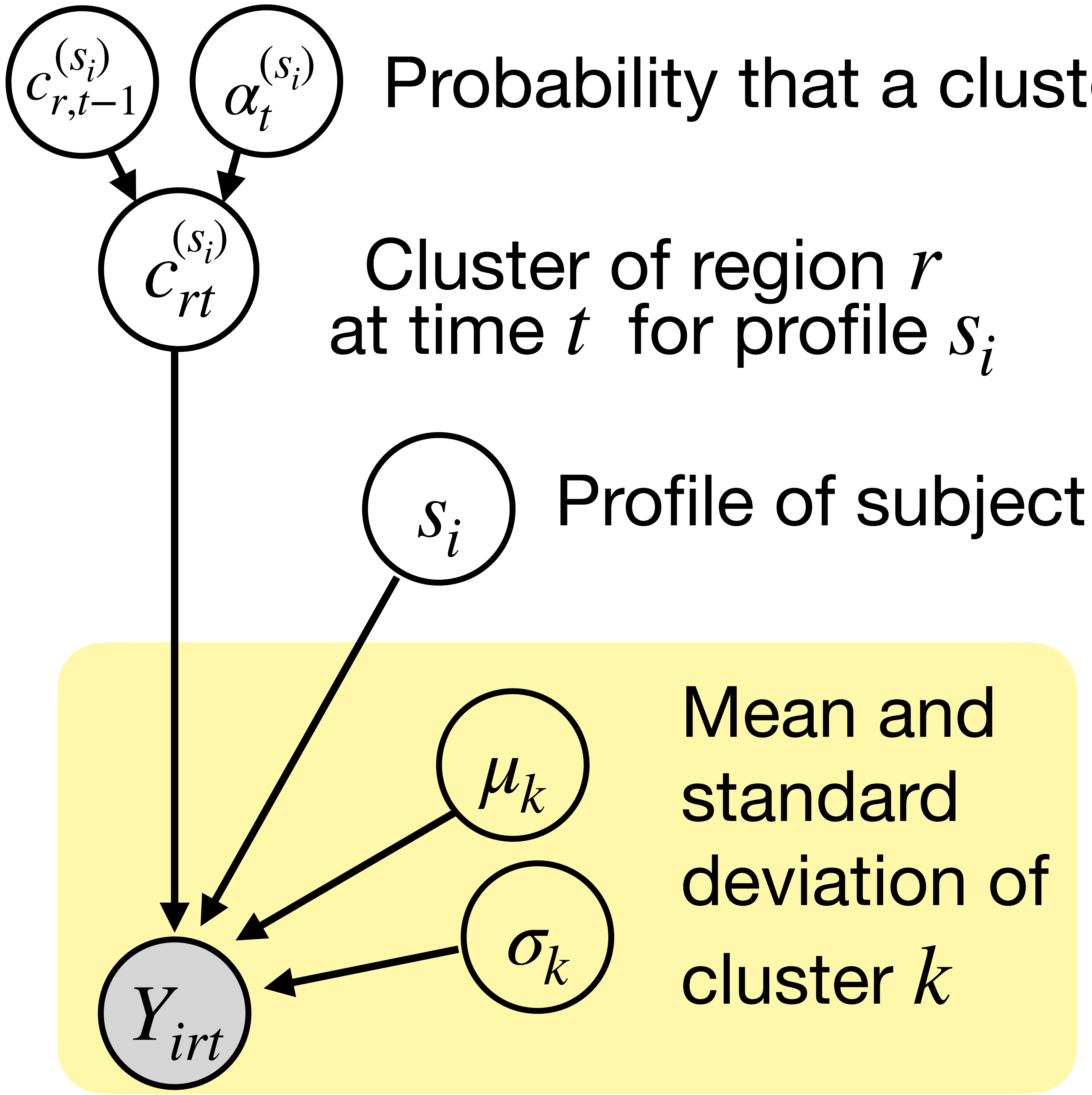
Cluster of region r at time t for profile s_i

Profile of subject i

Mean and standard deviation of cluster k

Full model

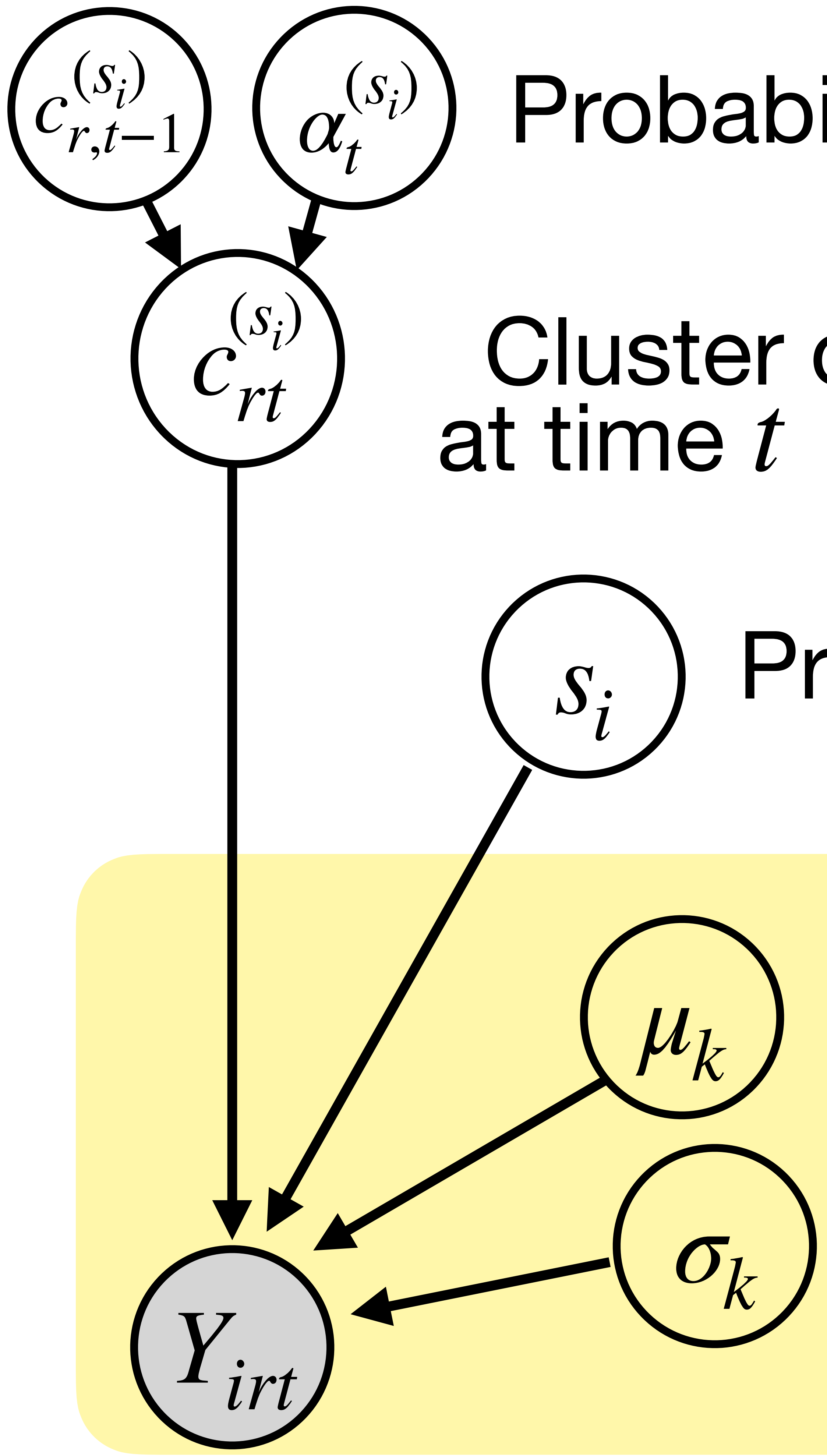
For subject i, brain region r and time t



$$Y_{irt} \mid s_i = z, c_{rt}^{(z)} = k, \mu_k, \sigma_k \sim \text{Student-t}(\mu_k, \sigma_k)$$

Full model

For subject i, brain region r and time t



Probability that a cluster persists over time

Cluster of region r
at time t for profile s_i

Profile of subject i

Mean and standard deviation of cluster k

$$Y_{irt} \mid s_i = z, c_{rt}^{(z)} = k, \mu_k, \sigma_k \sim \text{Student-t}(\mu_k, \sigma_k)$$

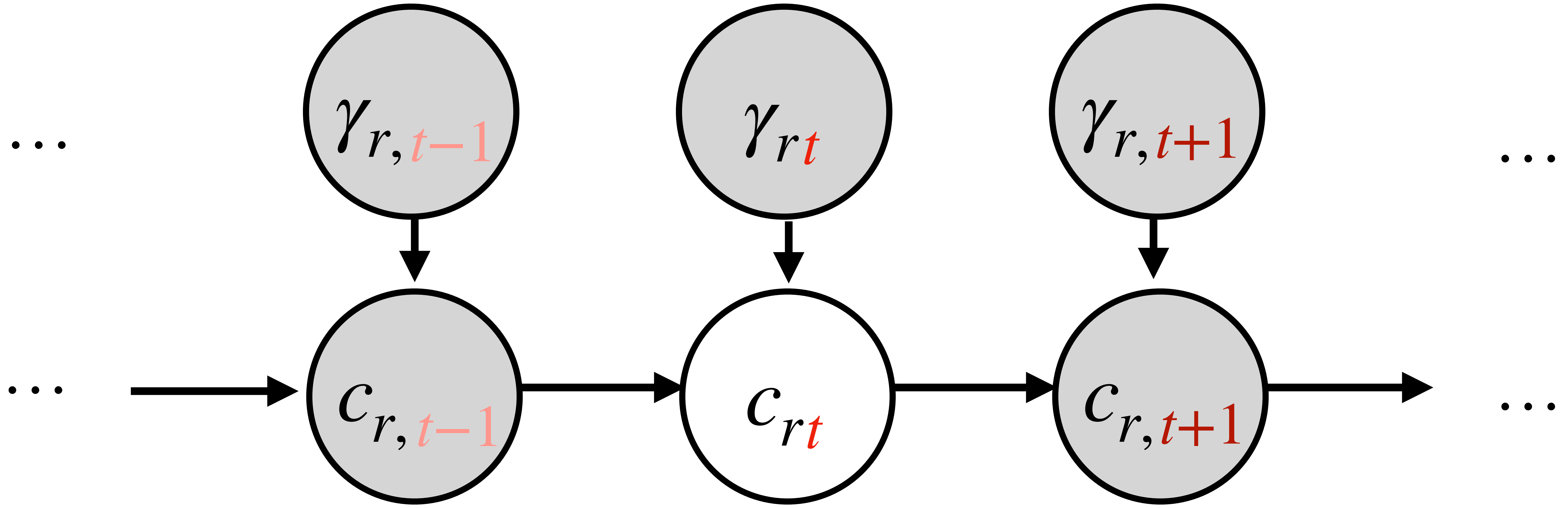
$$\Downarrow$$

$$Y_{i,r,t} \mid c_{r,t}^{(s_i)} = k, \mu_k, V_{i,r,t} \overset{\text{ind}}{\sim} \text{Normal}(\mu_k, V_{i,r,t})$$

$$V_{i,r,t} \mid c_{r,t}^{(s_i)} = k, \sigma_k^2 \overset{\text{iid}}{\sim} \text{Inv-}\chi^2(\nu, \sigma_k^2)\,.$$

Posterior Inference

- We design a MCMC for posterior inference, mostly using Gibbs updates
- Crucial step is the update of cluster-assignment sequence $(c_{r,1}^{(z)}, \ldots, c_{r,T}^{(z)})$ for each profile z and region r
- For the case with no profiles, Page et al. (2022) propose a marginal sampler
  - Let $\gamma_{r,t} = 1$ with probability α_t (so $\gamma_{r,t}$ is an indicator of cluster persistence)
  - Marginal updates are conditional on past, present and future persistence indicators and cluster assignments

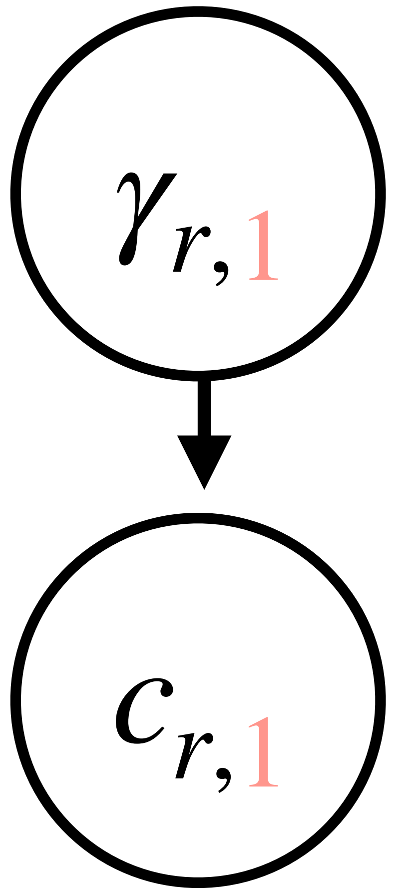


Block update of region clusters

- We design update of cluster-assignment sequences *in block:*
  - Update persistence indicators and cluster assignments together and sequentially, only conditioned on the past

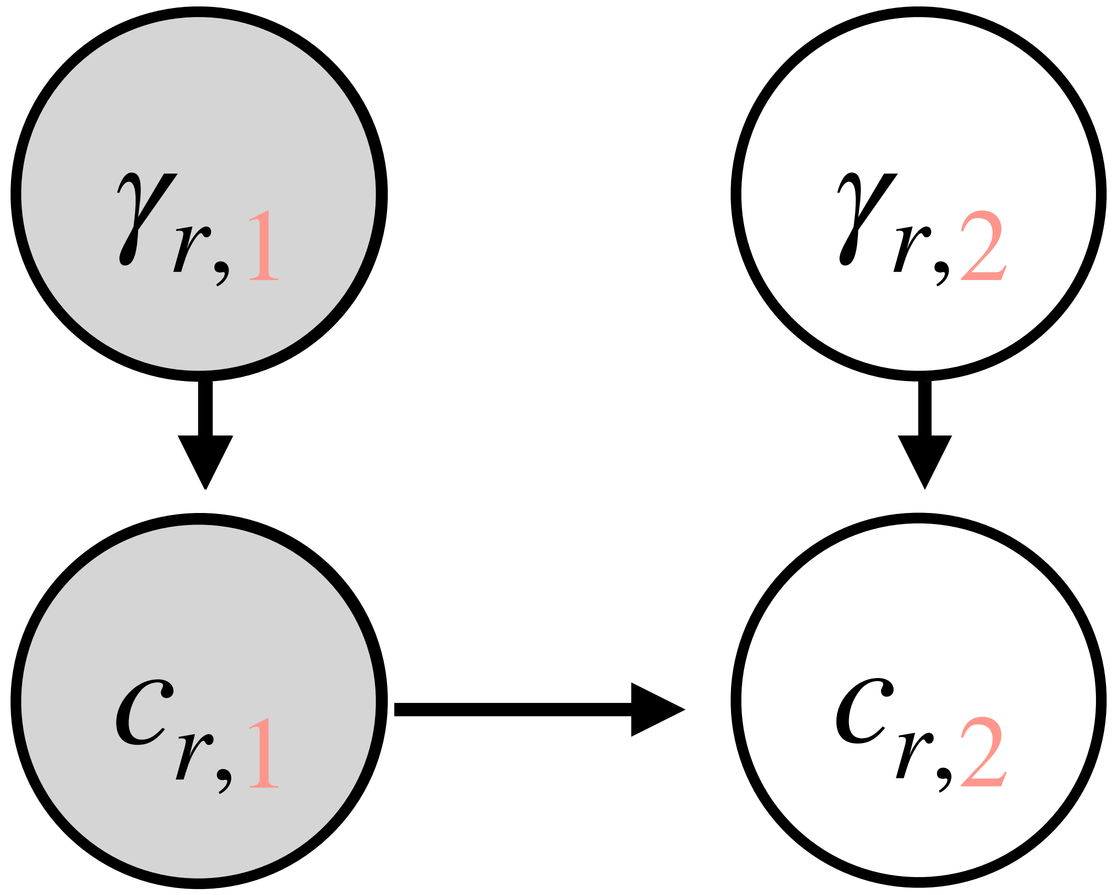
Block update of region clusters

- We design update of cluster-assignment sequences *in block:*
  - Update persistence indicators and cluster assignments together and sequentially, only conditioned on the past



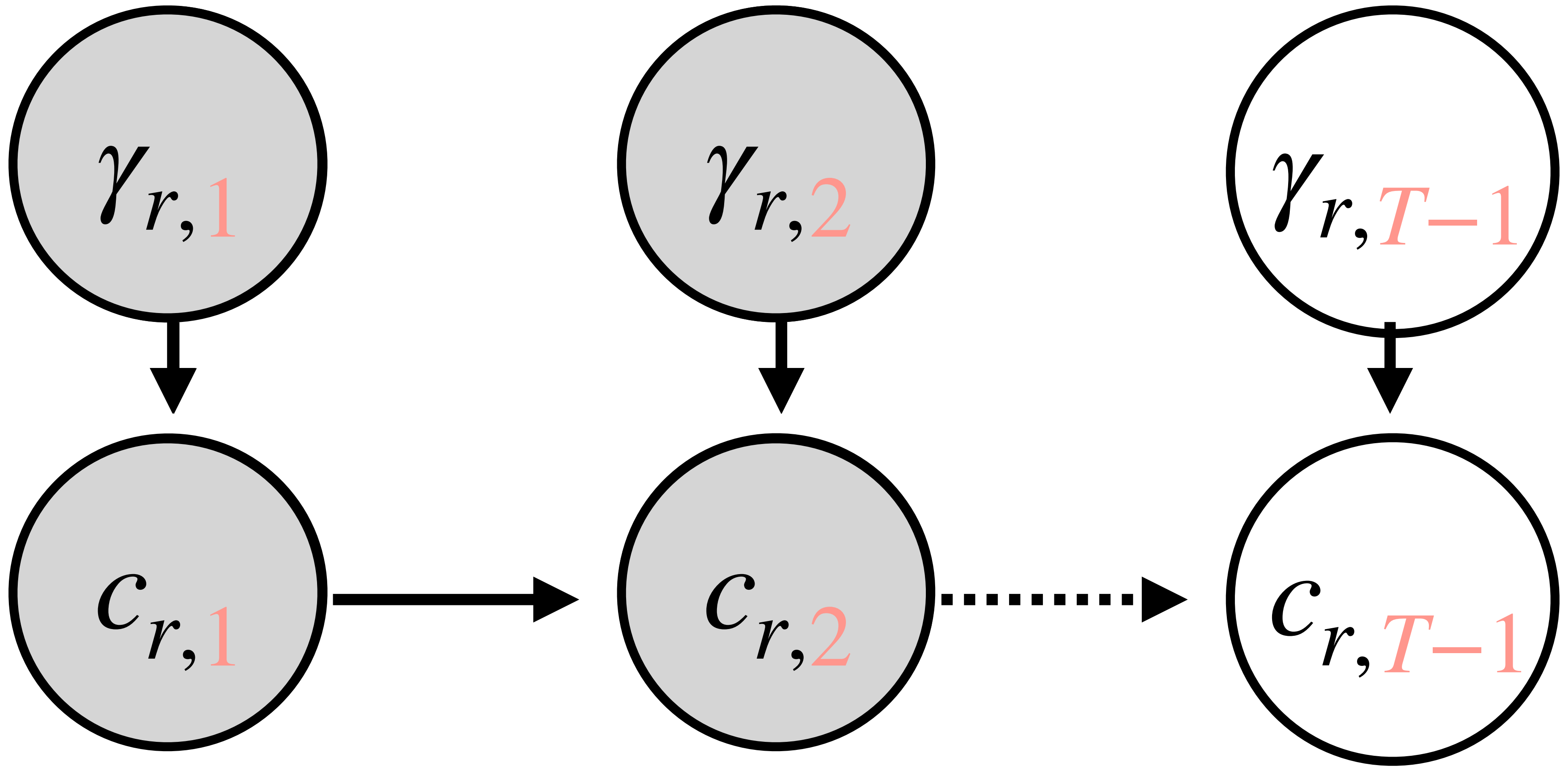
Block update of region clusters

- We design update of cluster-assignment sequences *in block:*
  - Update persistence indicators and cluster assignments together and sequentially, only conditioned on the past



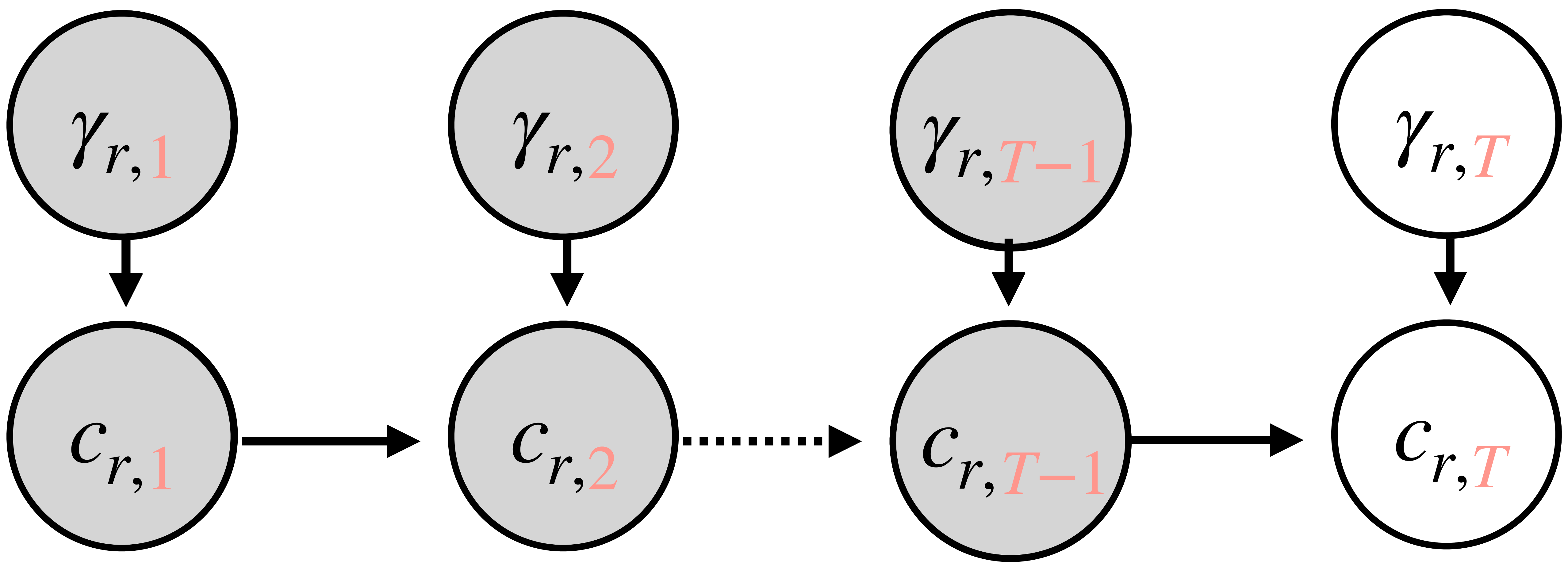
Block update of region clusters

- We design update of cluster-assignment sequences *in block:*
  - Update persistence indicators and cluster assignments together and sequentially, only conditioned on the past

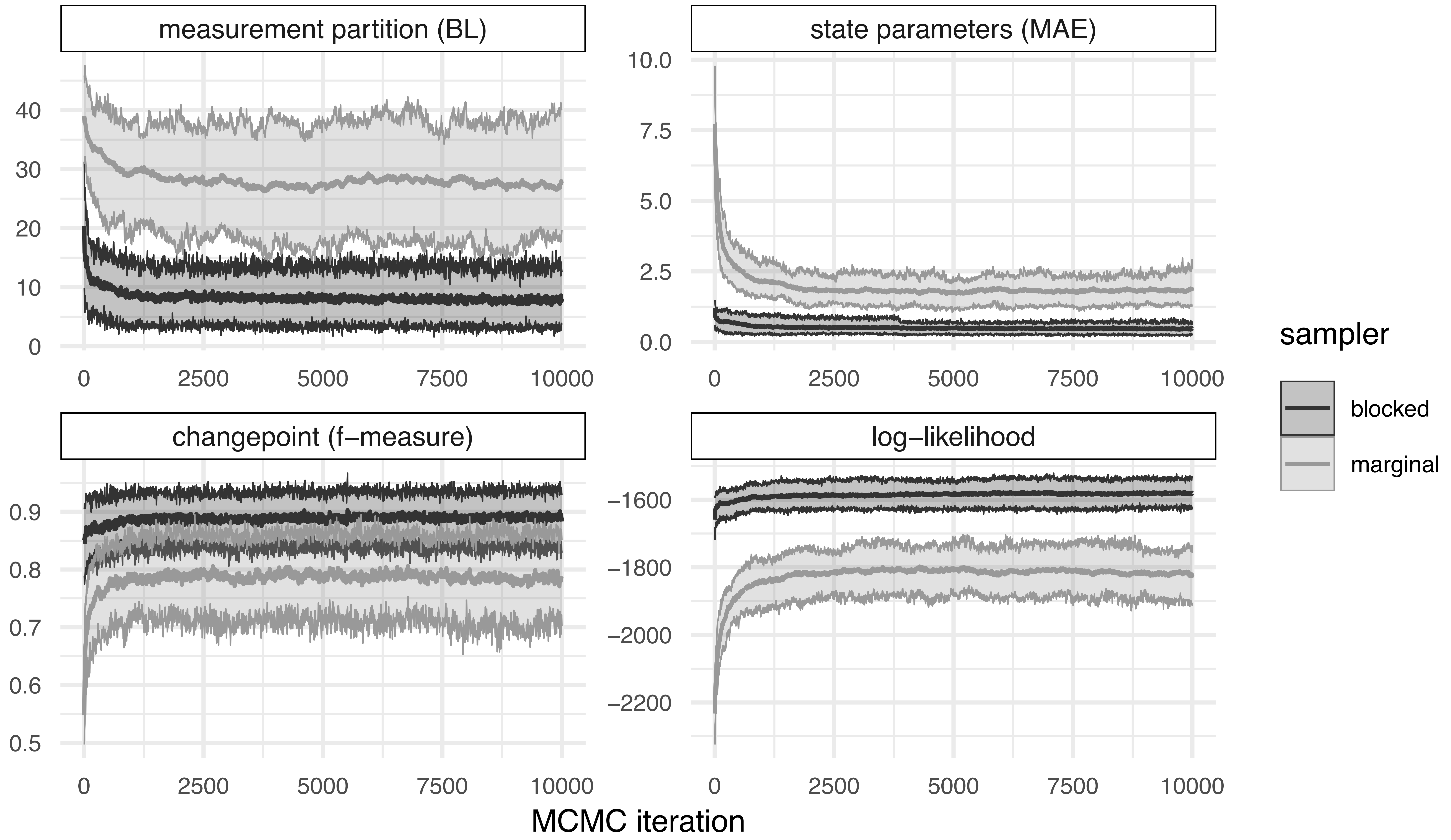


Block update of region clusters

- We design update of cluster-assignment sequences *in block:*
  - Update persistence indicators and cluster assignments together and sequentially, only conditioned on the past

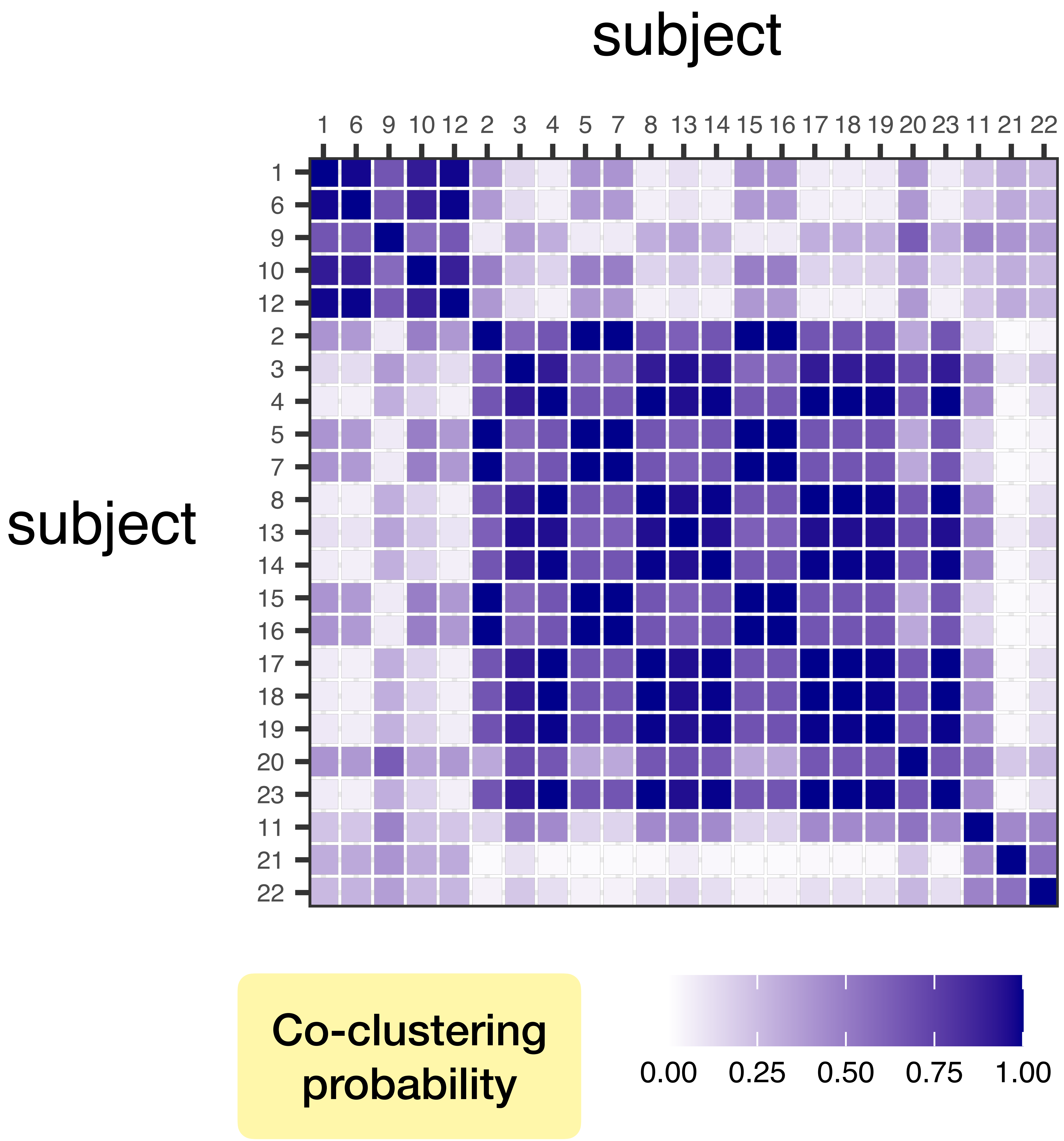


Blocked vs. Marginal sampler

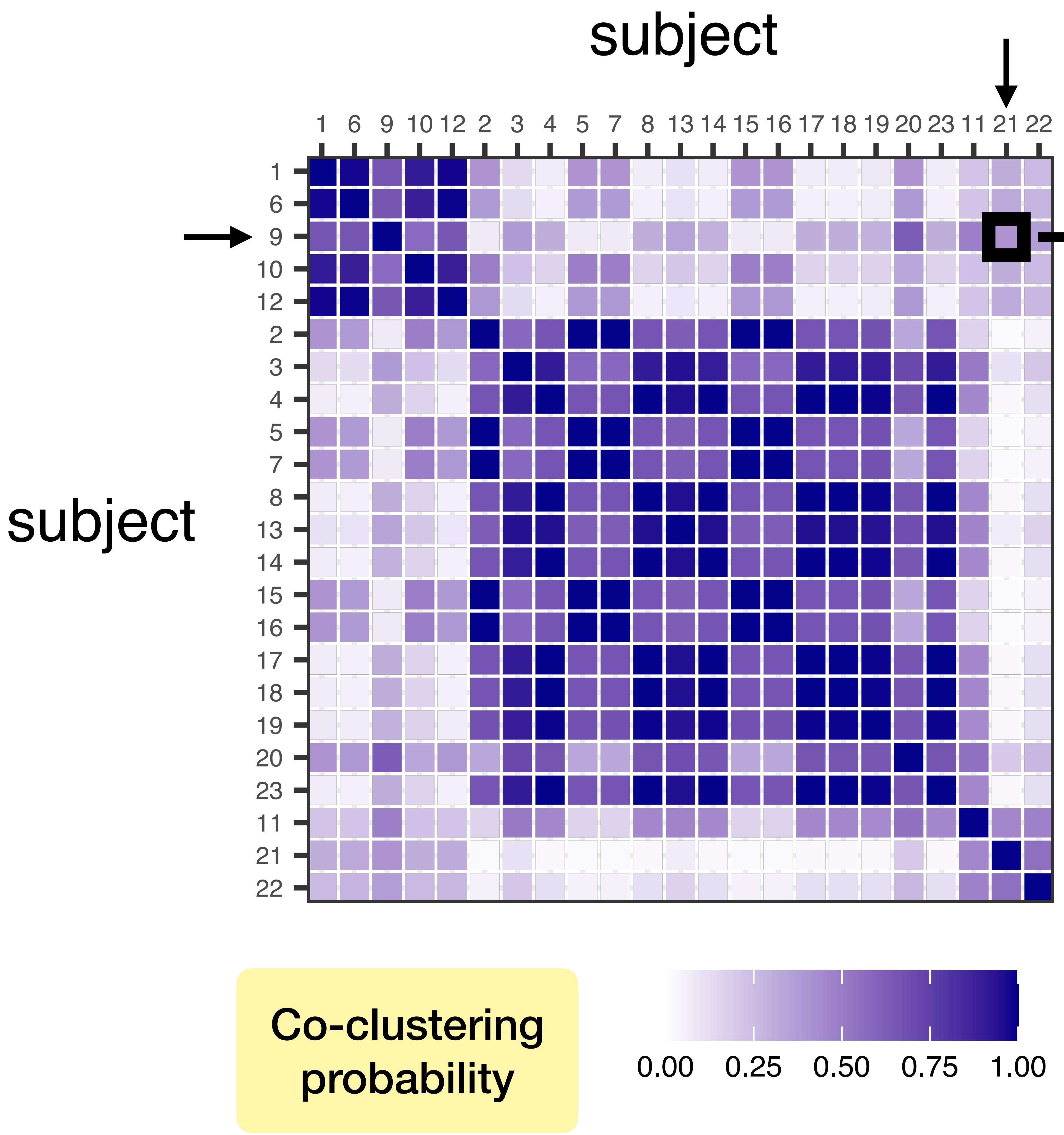


Application to neuroscience studies

fMRI data - profiles

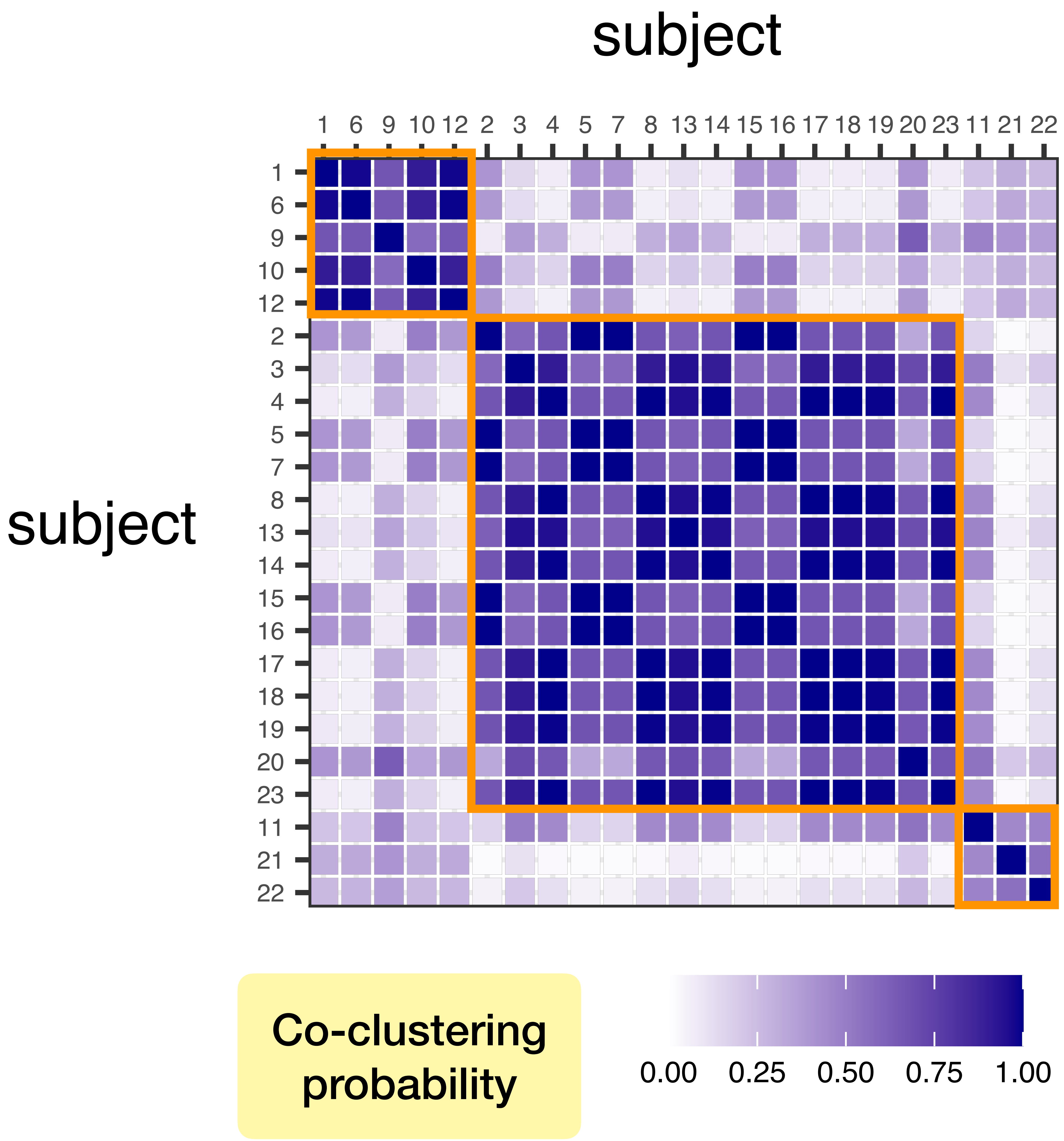


fMRI data - profiles



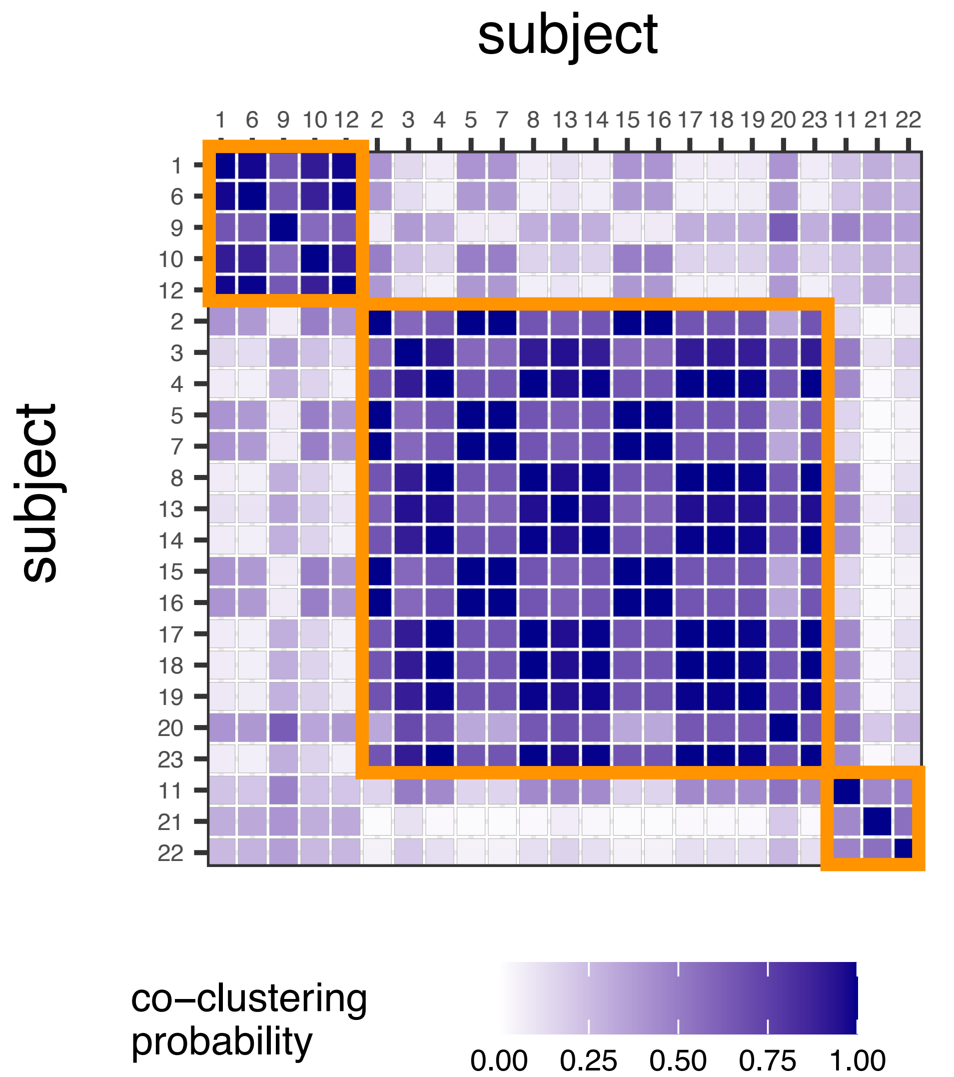
Probability that subject 9 and 21 have the same profile

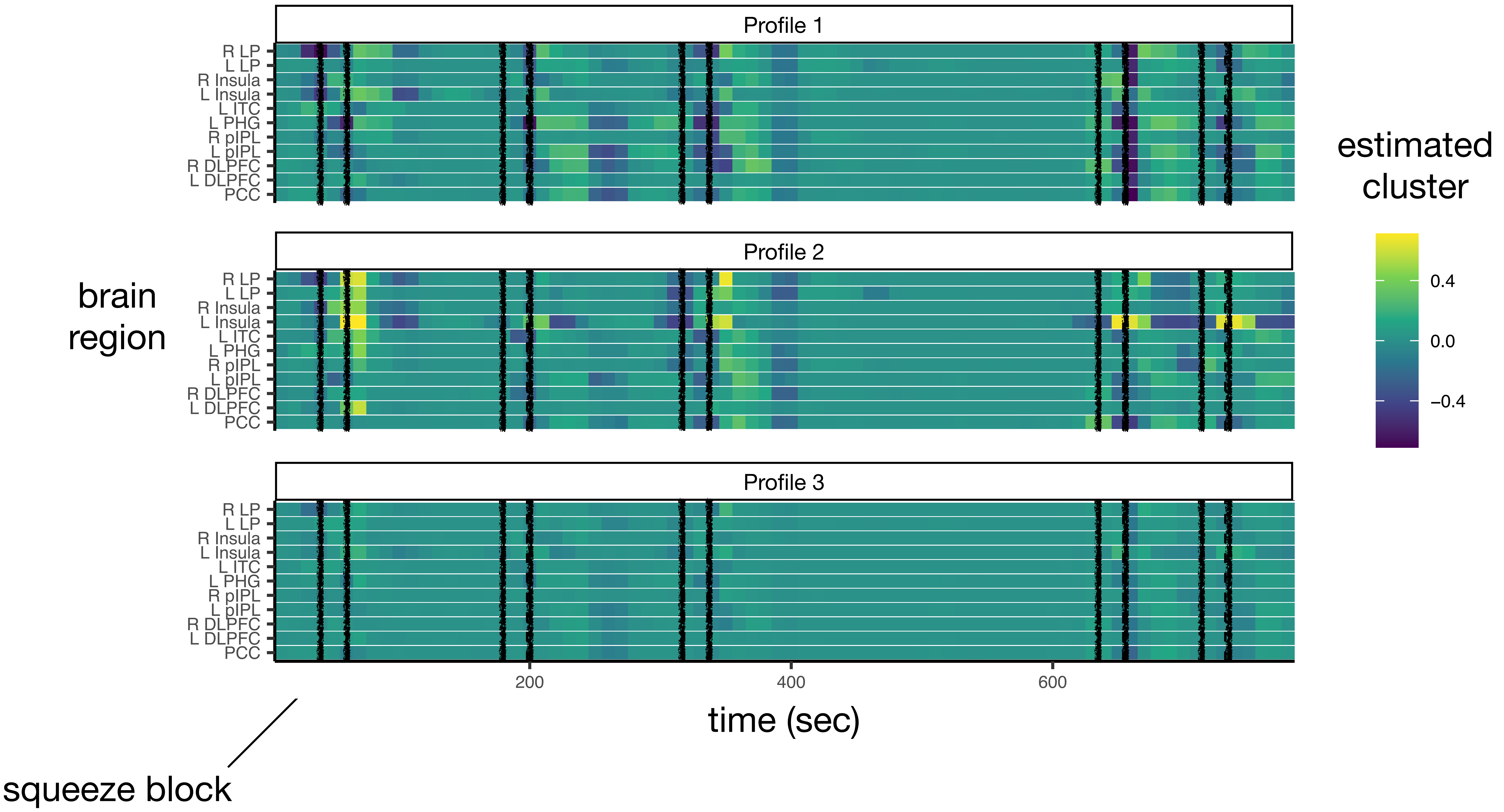
fMRI dataset - profiles



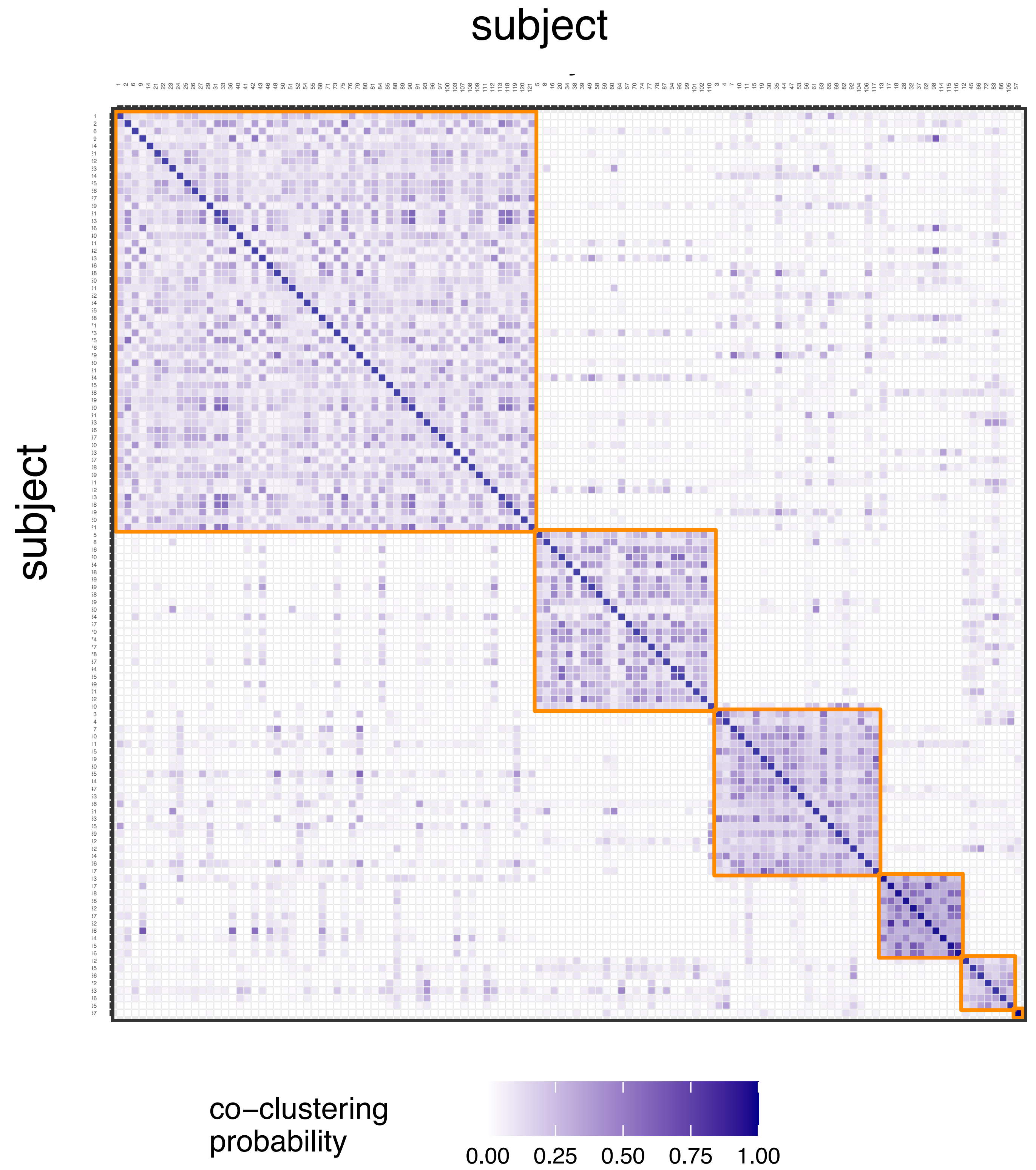
Estimate 3 profiles of subjects

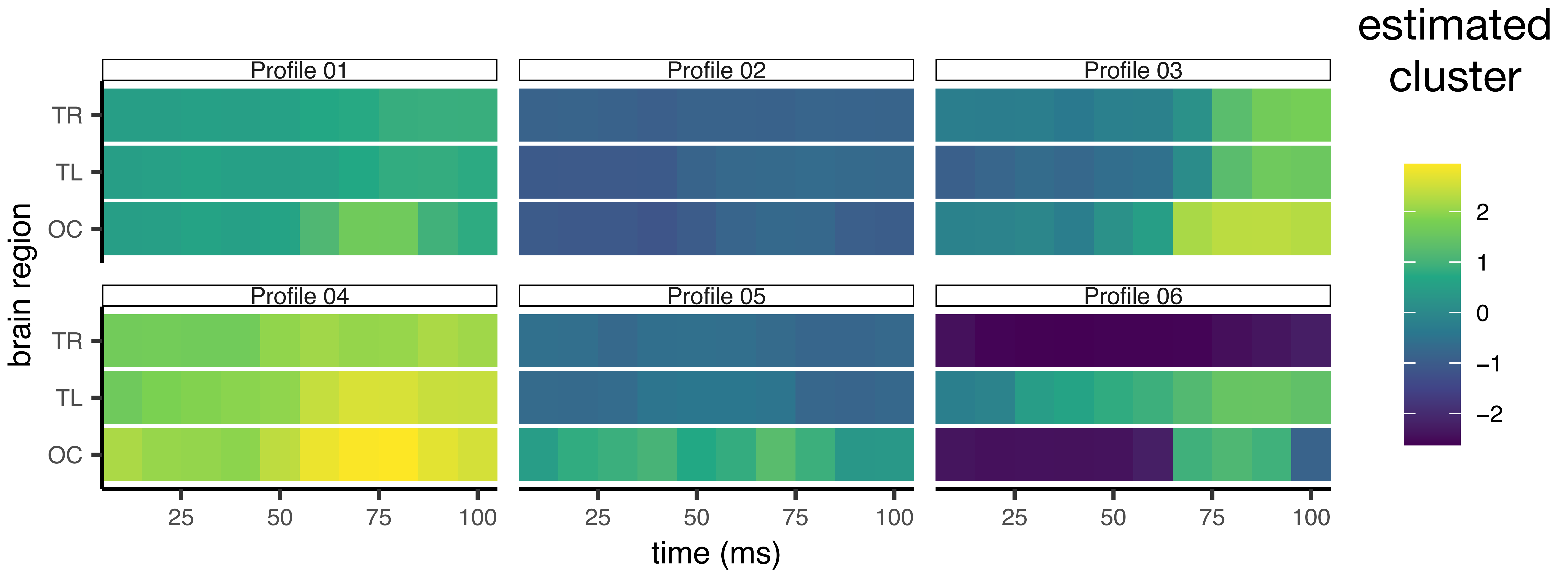
fMRI dataset - brain region clusters



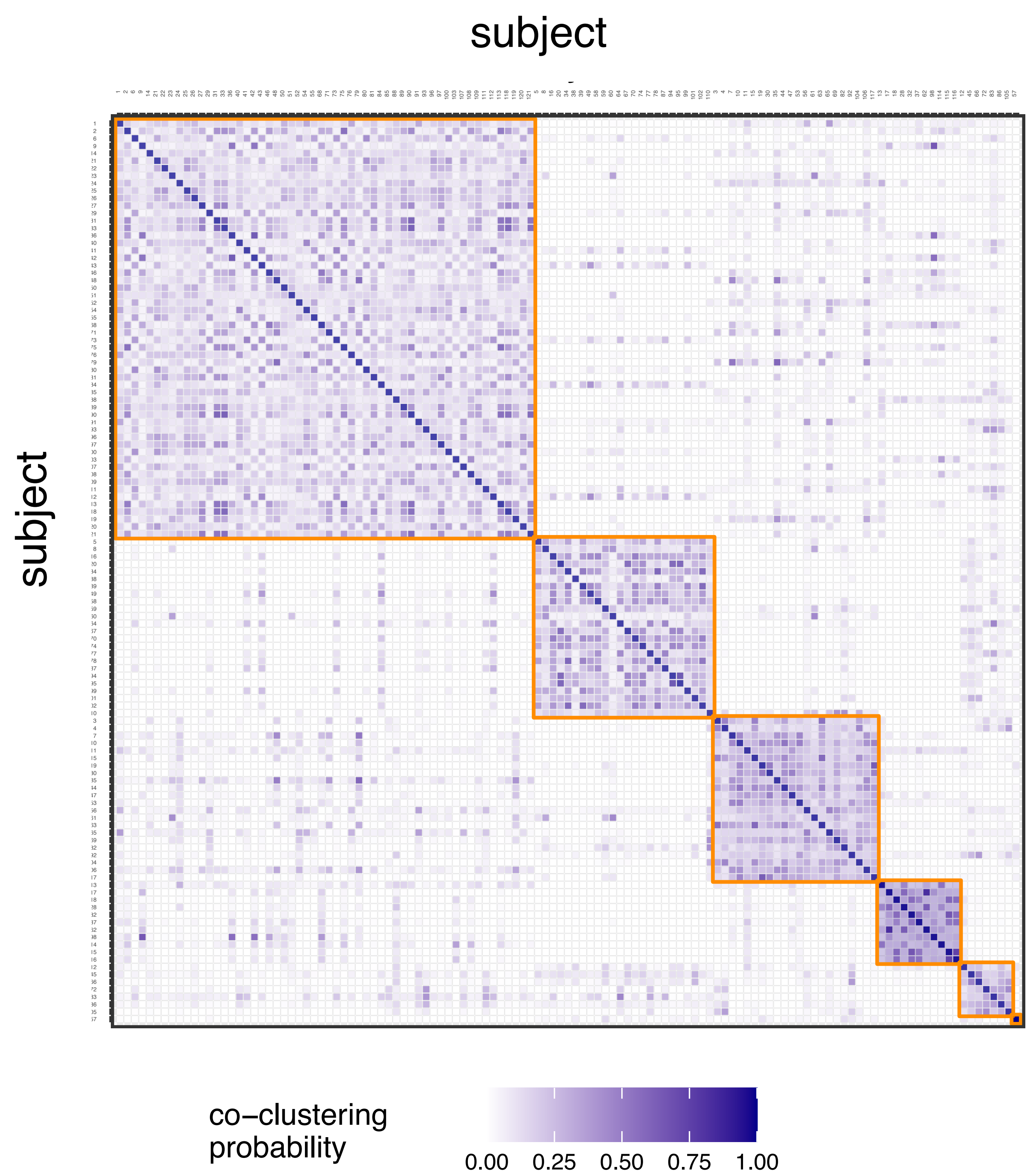


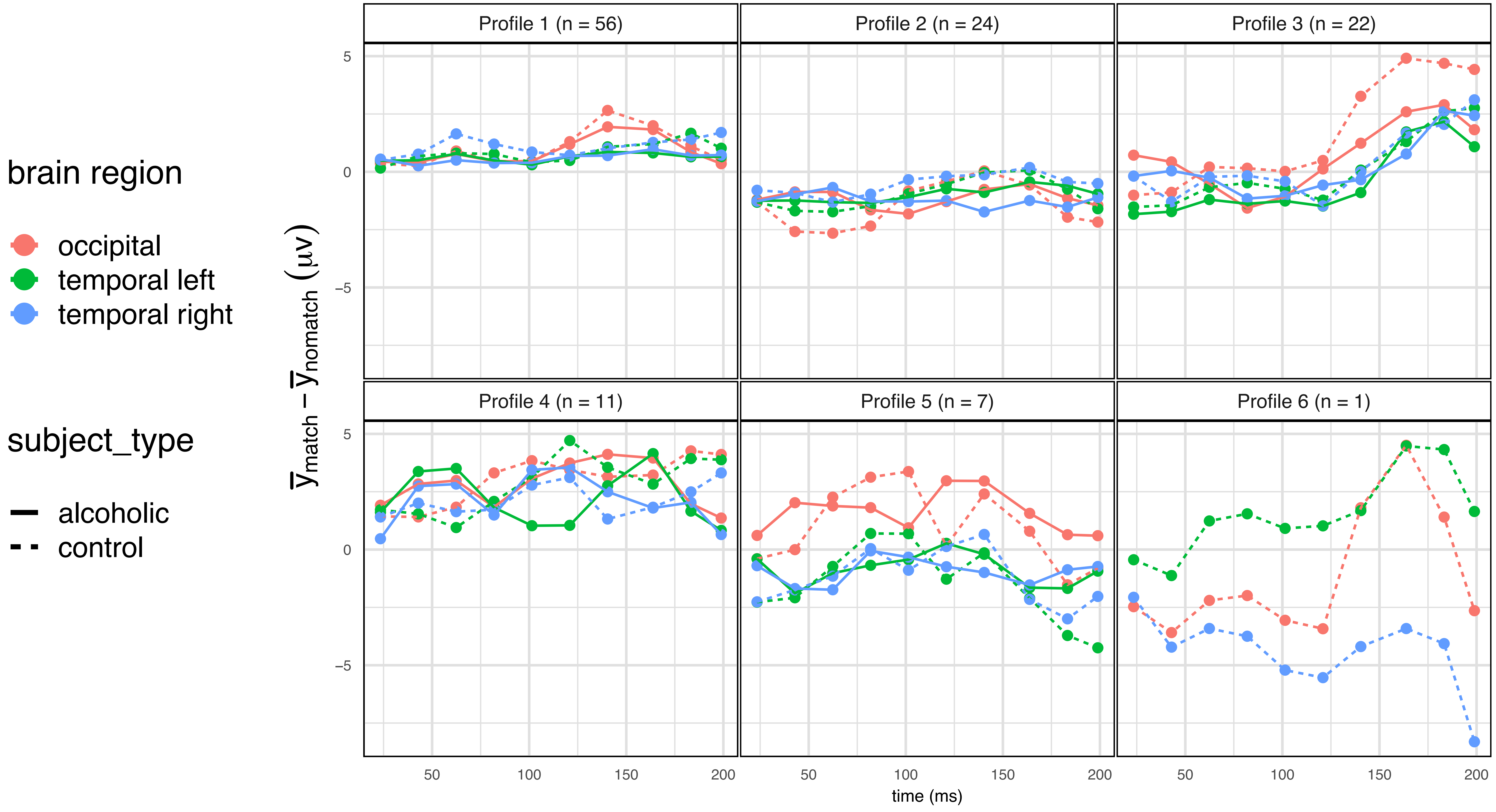
EEG dataset - brain region clusters





EEG dataset - brain region clusters





Discussion

Contributions

- Define a framework for time-varying partitions of measurements nested within time-invariant partitions of subjects

- Derive a MCMC block update for efficient exploration of state-sequence assignments

- Neuroscience applications: model identifies potential sub-populations within a given cohort based on neural profiles

Future directions

- Incorporate time-invariant covariates (e.g. subjects' characteristics)

- Allow for measurements to be observed at different time intervals for different subjects

- Include more measurements: using random effects (Kim and Smyth 2006), allowing there to be sets of irrelevant measurements (Lee et al. 2013) or using variable selection methods (Tadesse and Vannucci 2021)

- Allow subjects to change profile, e.g. multi-view clustering (Franzolini et al, 2023)

References

- Hussain, S., Menchaca, I., Shalchy, M. A., Yaghoubi, K., Langley, J., Seitz, A. R., ... & Peters, M. A. (2023). *Locus coeruleus integrity predicts ease of attaining and maintaining neural states of high attentiveness*. Brain Research Bulletin

- Lee, J., Hussain, S., Warnick, R., Vannucci, M., Menchaca, I., Seitz, A. R., ... & Guindani, M. (2024). *A predictor-informed multi-subject bayesian approach for dynamic functional connectivity.* Plos one

- Page, G. L., Quintana, F. A., & Dahl, D. B. (2022). *Dependent modeling of temporal sequences of random partitions*. Journal of Computational and Graphical Statistics

- Malsiner-Walli, G., Frühwirth-Schnatter, S., & Grün, B. (2016). *Model-based clustering based on sparse finite Gaussian mixtures*. Statistics and computing

- Lin, Q., Rebaudo, G., & Mueller, P. (2021). *Separate exchangeability as modeling principle in Bayesian nonparametrics*. arXiv preprint arXiv:2112.07755

Preprint: https://arxiv.org/abs/2406.17131

Questions?

Federica Zoe Ricci (UC Irvine)

fzricci@uci.edu