Spatial Bayesian Models for Task fMRI

OHBM 2023
Montreal, Canada

Amanda F. Mejia, Associate Professor
Department of Statistics
Indiana University, USA

StatMIND
Statistical Methods for Individualized Neuro-Discovery

A quick aside on icebergs


Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

ELSEVIER

Highlight results, don't hide them: Enhance interpretation, reduce biases and improve reproducibility

Paul A. Taylor[a,*], Richard C. Reynolds[a], Vince Calhoun[b], Javier Gonzalez-Castillo[c], Daniel A. Handwerker[c], Peter A. Bandettini[c,d], Amanda F. Mejia[e], Gang Chen[a]

[a] Scientific and Statistical Computing Core, NIMH, NIH, Bethesda, MD, USA
[b] Tri-institutional Center for Translational Research in Neuroimaging and Data Science (TReNDS), Georgia State, Georgia Tech, and Emory University, Atlanta, GA, USA
[c] Section on Functional Imaging Methods, NIMH, NIH, Bethesda, MD, USA
[d] Functional MRI Core Facility, NIMH, NIH, Bethesda, MD, USA
[e] Department of Statistics, Indiana University, IN, USA

ARTICLE INFO

Keywords:
FMRI
Results
Visualization
Reproducibility
Highlight

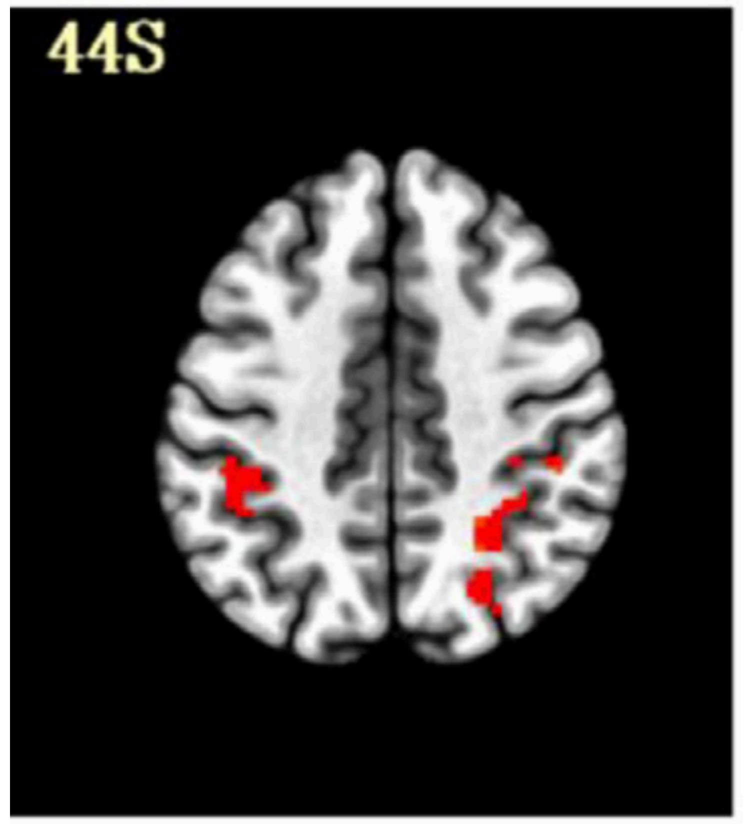
ABSTRACT

Most neuroimaging studies display results that represent only a tiny fraction of the collected data. While it is conventional to present "only the significant results" to the reader, here we suggest that this practice has several negative consequences for both reproducibility and understanding. This practice hides away most of the results of the dataset and leads to problems of selection bias and irreproducibility, both of which have been recognized as major issues in neuroimaging studies recently. Opaque, all-or-nothing thresholding, even if well-intentioned, places undue influence on arbitrary filter values, hinders clear communication of scientific results, wastes data, is antithetical to good scientific practice, and leads to conceptual inconsistencies. It is also inconsistent with

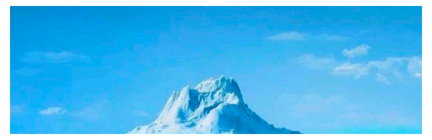
https://doi.org/10.1016/j.neuroimage.2023.120138

Check out Poster 731!

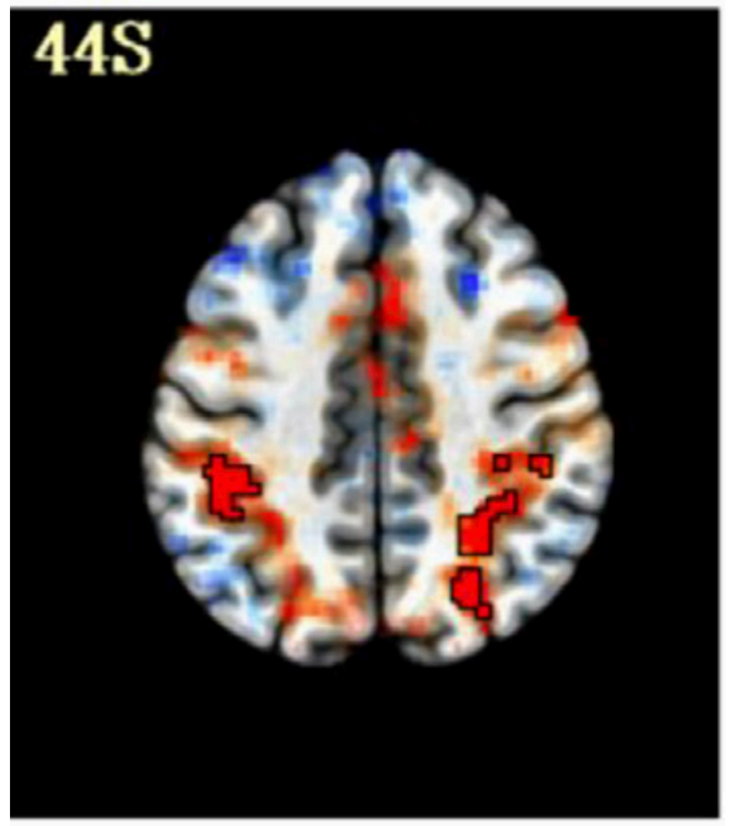
go from This



Tip of the iceberg



to This



The whole iceberg with the water level outlined

Outline

1. Bayesian statistics primer
2. Spatial priors
3. Spatial Bayesian GLM for task fMRI
4. Demo of BayesfMRI R package

Bayesian Statistics Primer

Frequentist vs. Bayesian Statistics

Frequentist Statistics

- Use only the observed data to estimate parameters, typically by maximizing the *likelihood*
- Usually fast & easy to compute 👍 (e.g., sample mean)
- Noise in the data is reflected in the estimates 👎
- Need to collect more data to reduce noise

Frequentist vs. Bayesian Statistics

Frequentist Statistics

- Use only the observed data to estimate parameters, typically by maximizing the *likelihood*
- Usually fast & easy to compute 👍 (e.g., sample mean)
- Noise in the data is reflected in the estimates 👎
- Need to collect more data to reduce noise

Bayesian Statistics

- Take into account *prior beliefs* about parameter via a *prior distribution*
- Estimation done by maximizing the *posterior distribution*, which reflects updated beliefs based on data
- Results in better (more efficient) estimates 👍
- Estimation can be difficult and/or slow 👎

What do we believe about task activation?

What do we believe about task activation?

1. Patterns of task activation should be fairly similar across subjects

What do we believe about task activation?

1. **Patterns of task activation should be fairly similar across subjects**

    - Encode this belief via a prior on the activation amplitude that reflects variability observed in the population

What do we believe about task activation?

1. **Patterns of task activation should be fairly similar across subjects**

    - Encode this belief via a prior on the activation amplitude that reflects variability observed in the population

    - *Hierarchical Bayesian models* use this kind of prior

What do we believe about task activation?

1. **Patterns of task activation should be fairly similar across subjects**
    - Encode this belief via a prior on the activation amplitude that reflects variability observed in the population
    - *Hierarchical Bayesian models* use this kind of prior

2. **Patterns of task activation should be similar for “proximal” locations**

What do we believe about task activation?

1. **Patterns of task activation should be fairly similar across subjects**
   - Encode this belief via a prior on the activation amplitude that reflects variability observed in the population
   - *Hierarchical Bayesian models* use this kind of prior

2. **Patterns of task activation should be similar for “proximal” locations**
   - Encode this belief via a multivariate prior across locations with covariance structure that represents the expected spatial dependence among “neighbors”

What do we believe about task activation?

1. **Patterns of task activation should be fairly similar across subjects**
    - Encode this belief via a prior on the activation amplitude that reflects variability observed in the population
    - *Hierarchical Bayesian models* use this kind of prior

2. **Patterns of task activation should be similar for "proximal" locations**
    - Encode this belief via a multivariate prior across locations with covariance structure that represents the expected spatial dependence among "neighbors"
    - *Spatial Bayesian models* use this kind of prior (a *spatial prior*)

Likelihood, Prior and Posterior

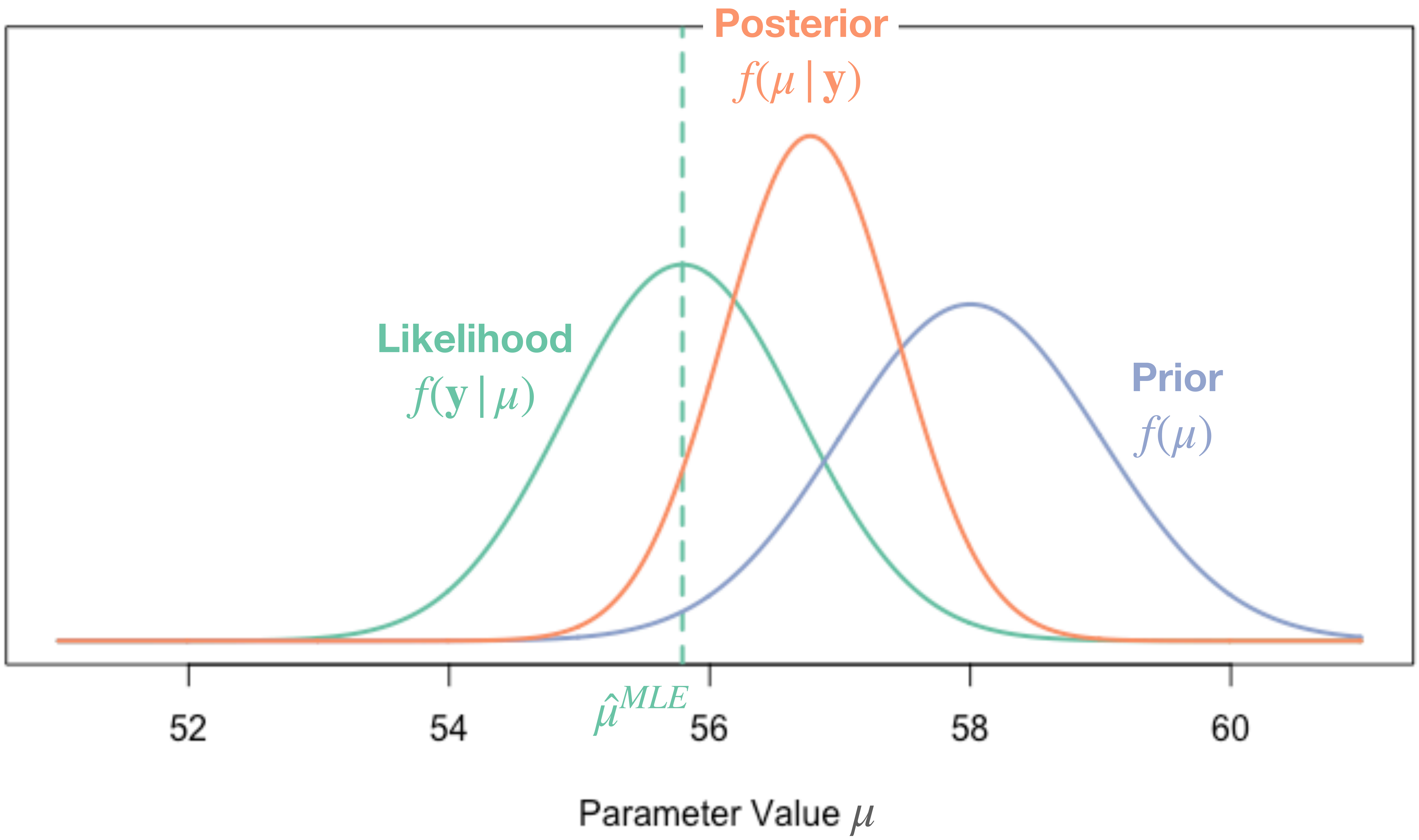
- The **likelihood** $f(\mathbf{y} \mid \mu)$ is the probability of observing the data $\mathbf{y}$, given a particular value of a parameter μ

- The **prior distribution** $f(\mu)$ reflects our original or prior beliefs, absent data, about a parameter μ

- The **posterior distribution** $f(\mu \mid \mathbf{y})$ reflects our updated beliefs about μ based on the data $\mathbf{y}$

$$\overbrace{f(\mu \mid \mathbf{y})}^{\text{Posterior}} \underset{\text{Bayes Rule}}{=} \frac{\overbrace{f(\mathbf{y} \mid \mu)}^{\text{Likelihood}}\,\overbrace{f(\mu)}^{\text{Prior}}}{\underbrace{f(\mathbf{y})}_{\text{Marginal Likelihood}}} \overset{\mu}{\propto} \overbrace{f(\mathbf{y} \mid \mu)}^{\text{Likelihood}}\,\overbrace{f(\mu)}^{\text{Prior}}$$

Likelihood, Prior and Posterior

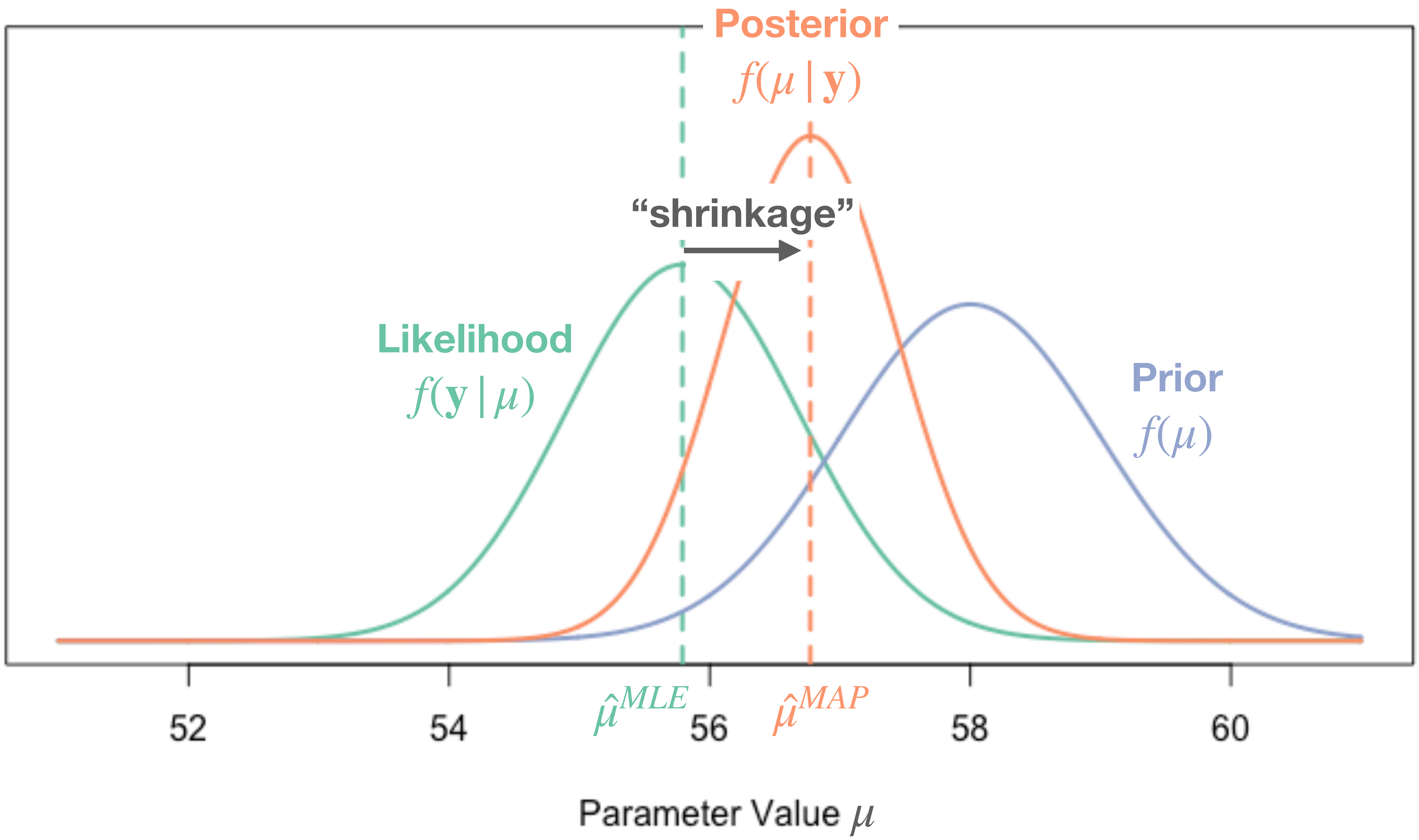


MLE = the value of μ that maximizes the *likelihood*

Maximum-a-posteriori (MAP) estimators

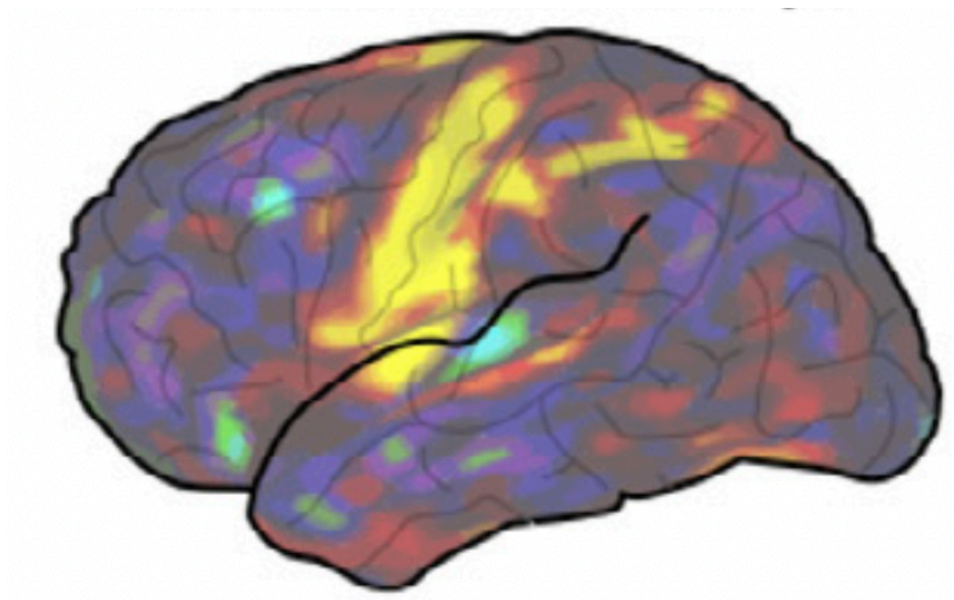
MAP = the value of μ that maximizes the *posterior*

MLE = the value of μ that maximizes the *likelihood*



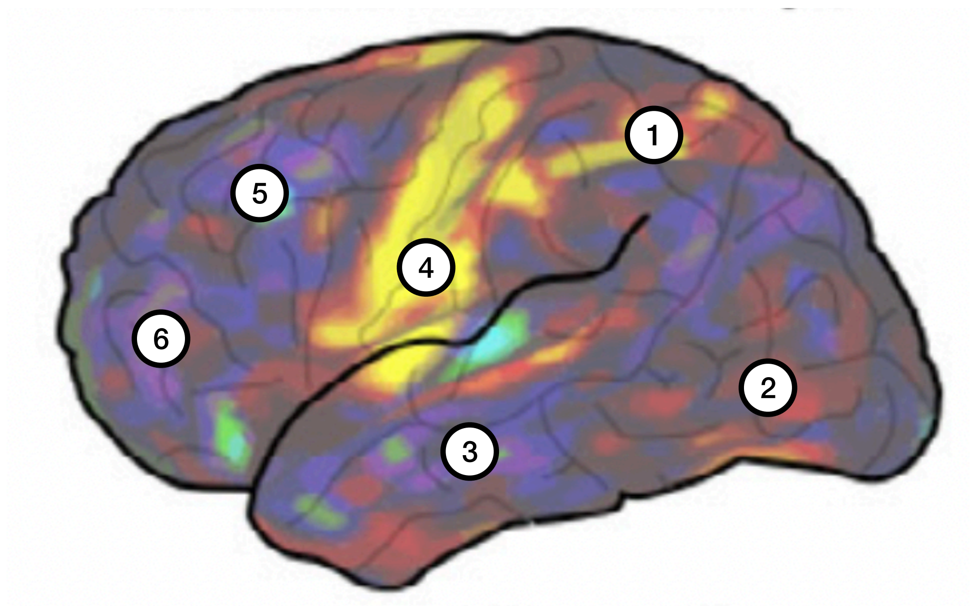
Spatial Priors

What is a spatial prior?



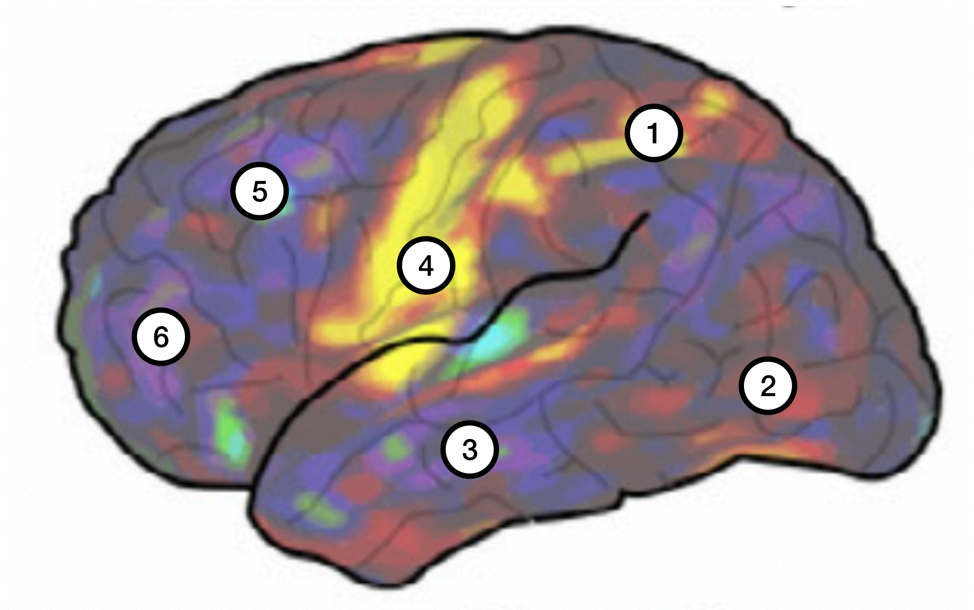
latent spatial map $\boldsymbol{\beta}$

What is a spatial prior?



latent spatial map $\boldsymbol{\beta}$

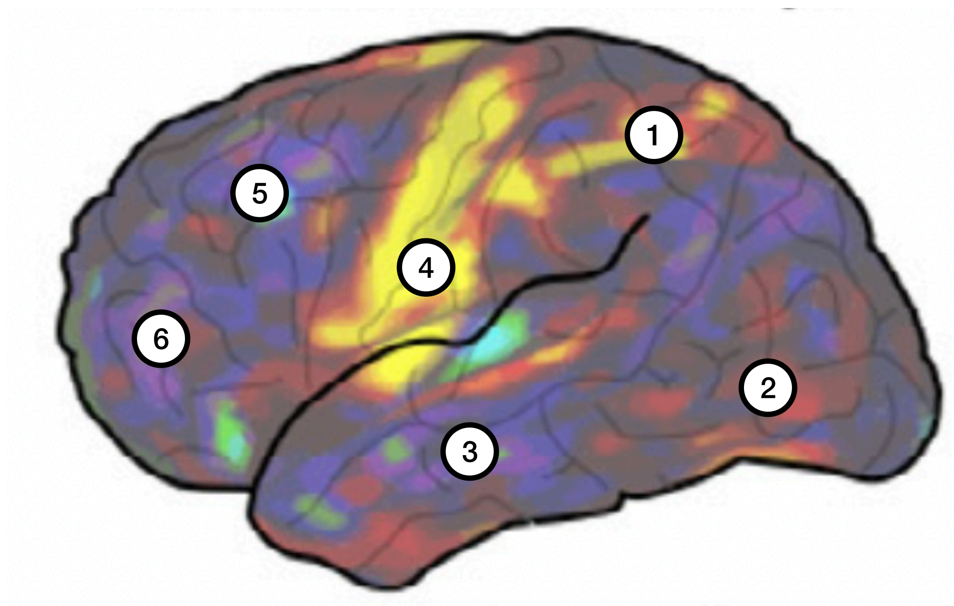
What is a spatial prior?



latent spatial map $\boldsymbol{\beta}$

$$\boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \\ \beta_6 \end{pmatrix} \sim (\underset{(6x1)}{\boldsymbol{\mu}_0}, \underset{(6x6)}{\boldsymbol{\Sigma}_0})$$

What is a spatial prior?



latent spatial map $\boldsymbol{\beta}$

$$\boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \\ \beta_6 \end{pmatrix} \sim (\underset{(6\text{x}1)}{\boldsymbol{\mu}_0}, \underset{(6\text{x}6)}{\boldsymbol{\Sigma}_0})$$

One example of a spatial prior:

Multivariate Gaussian prior with mean zero and sparse inverse covariance

$$\boldsymbol{\beta} \sim N(\mathbf{0}, \mathbf{Q}^{-1})$$

$\mathbf{Q}$ contains zeros everywhere except along diagonal & in cells corresponding to "neighbors"

Toy fMRI example

Toy fMRI example

Consider a single task, and for simplicity let the design column $\mathbf{x}$ be all 1's

Toy fMRI example

Consider a single task, and for simplicity let the design column $\mathbf{x}$ be all 1's

- Likelihood at location $v = 1,\ldots,V$: $\quad \underset{(T \times 1)}{\mathbf{y}(v)} = \beta(v)\underset{(T \times 1)}{\mathbf{x}} + \underset{(T \times 1)}{\mathbf{e}(v)}, \quad \mathbf{e}(v) \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_T)$

Toy fMRI example

Consider a single task, and for simplicity let the design column $\mathbf{x}$ be all 1's

- Likelihood at location $v = 1,\ldots,V$: $\underset{(T \times 1)}{\mathbf{y}(v)} = \beta(v)\underset{(T \times 1)}{\mathbf{x}} + \underset{(T \times 1)}{\mathbf{e}(v)}, \quad \mathbf{e}(v) \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_T)$
- Likelihood across all locations: $\underset{(TV \times 1)}{\mathbf{y}} = \mathbf{X}\underset{(V \times 1)}{\boldsymbol{\beta}} + \mathbf{e}, \quad \mathbf{e} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_{TV})$

$$\underset{(TV \times V)}{\mathbf{X} = \begin{pmatrix} \mathbf{x} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{x} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{x} \end{pmatrix}}$$

Toy fMRI example

Consider a single task, and for simplicity let the design column $\mathbf{x}$ be all 1's

- Likelihood at location $v = 1, \ldots, V$: $\underset{(T \times 1)}{\mathbf{y}(v)} = \beta(v)\underset{(T \times 1)}{\mathbf{x}} + \underset{(T \times 1)}{\mathbf{e}(v)}, \quad \mathbf{e}(v) \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_T)$
- Likelihood across all locations: $\underset{(TV\times 1)}{\mathbf{y}} = \mathbf{X}\underset{(V \times 1)}{\boldsymbol{\beta}} + \mathbf{e}, \quad \mathbf{e} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_{TV})$
- Prior on $\boldsymbol{\beta}$: $\underset{(V \times 1)}{\boldsymbol{\beta}} \sim N(\underset{(V \times 1)}{\mathbf{0}}, \underset{(V \times V)}{\mathbf{Q}^{-1}})$

$$\underset{(TV \times V)}{\mathbf{X} = \begin{pmatrix} \mathbf{x} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{x} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{x} \end{pmatrix}}$$

Toy fMRI example

Consider a single task, and for simplicity let the design column $\mathbf{x}$ be all 1's

- Likelihood at location $v = 1, \ldots, V$: $\mathbf{y}(v) = \beta(v)\mathbf{x} + \mathbf{e}(v), \quad \mathbf{e}(v) \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_T)$

  (T x 1) (T x 1) (T x 1)

- Likelihood across all locations: $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}, \quad \mathbf{e} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_{TV})$

  (TVx1) (V x 1)

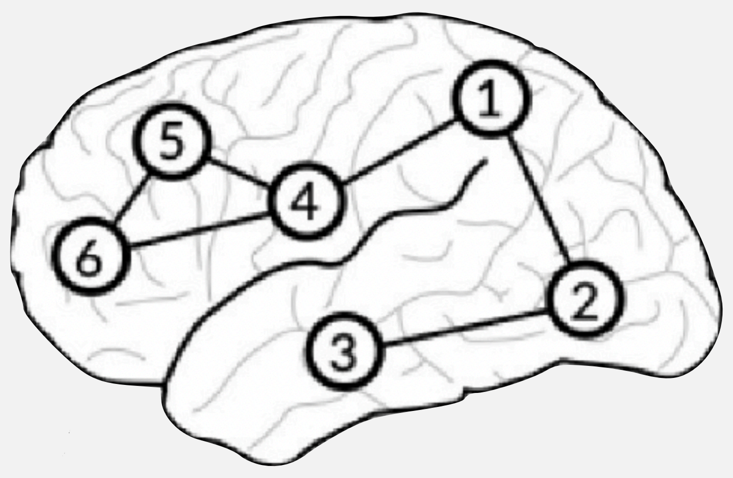
- Prior on $\boldsymbol{\beta}$: $\boldsymbol{\beta} \sim N(\mathbf{0}, \mathbf{Q}^{-1})$

  (V x 1) (V x 1) (V x V)

$$\mathbf{X} = \begin{pmatrix} \mathbf{x} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{x} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{x} \end{pmatrix}$$

(TV x V)

$\mathbf{Q}$ represents our beliefs about the spatial dependence of task activation amplitudes in $\boldsymbol{\beta}$.



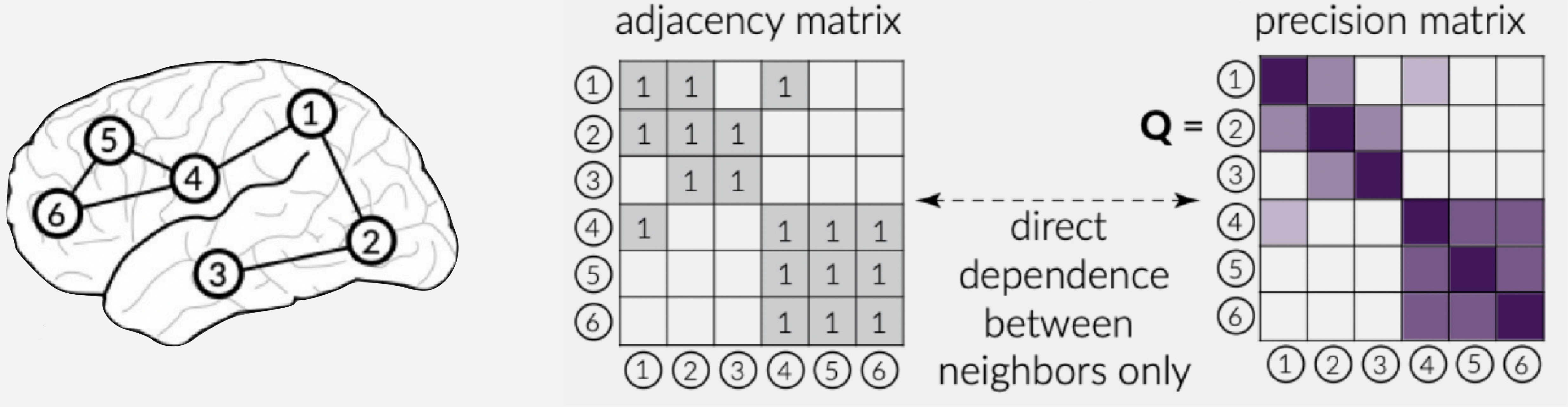
Toy fMRI example

Consider a single task, and for simplicity let the design column $\mathbf{x}$ be all 1's

- Likelihood at location $v = 1,\ldots,V$: $\quad \underset{(T \times 1)}{\mathbf{y}(v)} = \beta(v)\underset{(T \times 1)}{\mathbf{x}} + \underset{(T \times 1)}{\mathbf{e}(v)}, \quad \mathbf{e}(v) \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_T)$
- Likelihood across all locations: $\quad \underset{(TVx1)}{\mathbf{y}} = \mathbf{X}\underset{(V \times 1)}{\boldsymbol{\beta}} + \mathbf{e}, \quad \mathbf{e} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_{TV})$
- Prior on $\boldsymbol{\beta}$: $\quad \underset{(V \times 1)}{\boldsymbol{\beta}} \sim N(\underset{(V \times 1)}{\mathbf{0}}, \underset{(V \times V)}{\mathbf{Q}^{-1}})$

$$\underset{(TV \times V)}{\mathbf{X}} = \begin{pmatrix} \mathbf{x} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{x} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{x} \end{pmatrix}$$

$\mathbf{Q}$ represents our beliefs about the spatial dependence of task activation amplitudes in $\boldsymbol{\beta}$.



Toy fMRI example

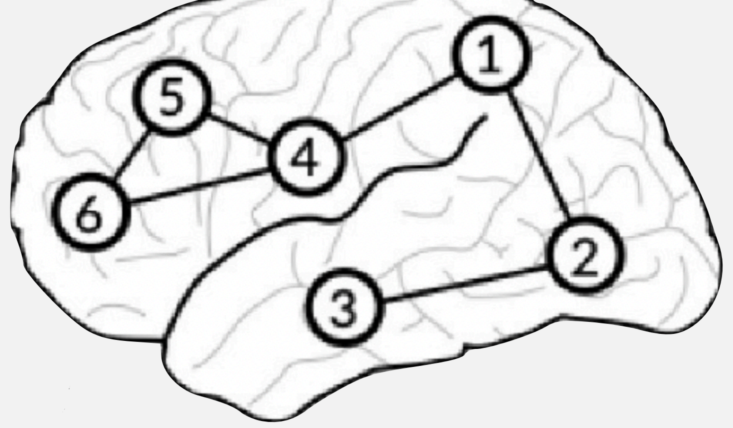
Consider a single task, and for simplicity let the design column $\mathbf{x}$ be all 1's

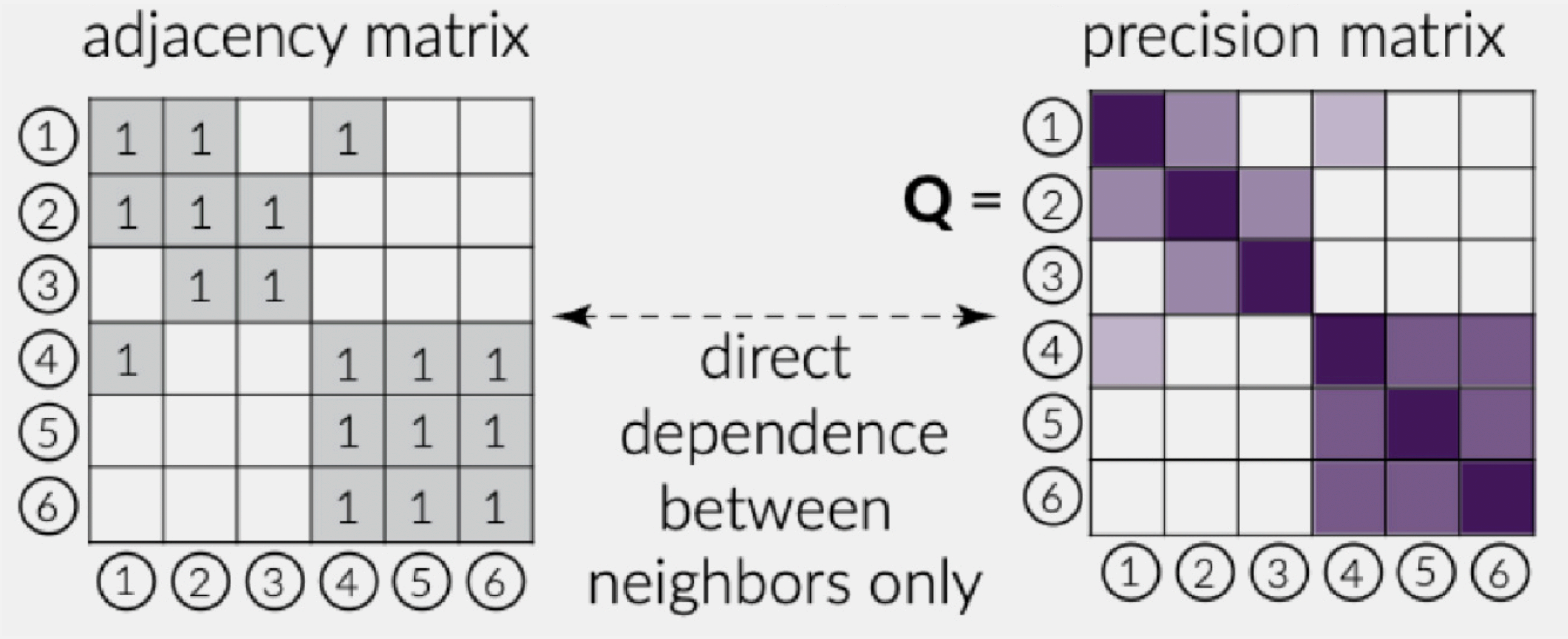
- Likelihood at location $v = 1, \ldots, V$: $\underset{(T \times 1)}{\mathbf{y}(v)} = \beta(v)\underset{(T \times 1)}{\mathbf{x}} + \underset{(T \times 1)}{\mathbf{e}(v)}, \quad \mathbf{e}(v) \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_T)$
- Likelihood across all locations: $\underset{(TV \times 1)}{\mathbf{y}} = \mathbf{X}\underset{(V \times 1)}{\boldsymbol{\beta}} + \mathbf{e}, \quad \mathbf{e} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_{TV})$
- Prior on $\boldsymbol{\beta}$: $\underset{(V \times 1)}{\boldsymbol{\beta}} \sim N(\underset{(V \times 1)}{\mathbf{0}}, \underset{(V \times V)}{\mathbf{Q}^{-1}})$

$$\mathbf{X} = \begin{pmatrix} \mathbf{x} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{x} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{x} \end{pmatrix}$$

(TV x V)

$\mathbf{Q}$ represents our beliefs about the spatial dependence of task activation amplitudes in $\boldsymbol{\beta}$.





some example numbers

$$\mathbf{Q} = \begin{pmatrix} 5 & -1 & 0 & -1 & 0 & 0 \\ -1 & 5 & -1 & 0 & 0 & 0 \\ 0 & -1 & 5 & 0 & 0 & 0 \\ -1 & 0 & 0 & 5 & -2 & -2 \\ 0 & 0 & 0 & -2 & 5 & -2 \\ 0 & 0 & 0 & -2 & -2 & 5 \end{pmatrix}$$

(Inverse covariance)

$$\mathbf{Q}^{-1} = \begin{pmatrix} 0.23 & 0.05 & 0.01 & 0.10 & 0.07 & 0.07 \\ 0.05 & 0.22 & 0.04 & 0.02 & 0.01 & 0.01 \\ 0.01 & 0.04 & 0.21 & 0.00 & 0.00 & 0.00 \\ 0.10 & 0.02 & 0.00 & 0.47 & 0.31 & 0.31 \\ 0.07 & 0.01 & 0.00 & 0.31 & 0.45 & 0.30 \\ 0.07 & 0.01 & 0.00 & 0.31 & 0.30 & 0.45 \end{pmatrix}$$

(covariance matrix)

Toy example

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}, \quad \mathbf{e} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_{TV})$$

<u>Frequentist estimate of $\boldsymbol{\beta}$</u>

(MLE equivalent to OLS)

$$\hat{\boldsymbol{\beta}}^{MLE} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = (\bar{y}_1, \cdots \bar{y}_V)'$$

Toy example

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}, \quad \mathbf{e} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_{TV})$$

<u>Frequentist estimate of $\boldsymbol{\beta}$</u>

(MLE equivalent to OLS)

$$\hat{\boldsymbol{\beta}}^{MLE} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = (\bar{y}_1, \cdots \bar{y}_V)'$$

$\mathbf{x} = (1,1,\ldots,1)'$

$$\mathbf{X} = \begin{pmatrix} \mathbf{x} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{x} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{x} \end{pmatrix} \quad \mathbf{X}'\mathbf{X} = \begin{pmatrix} T & & \\ & \ddots & \\ & & T \end{pmatrix} \quad \mathbf{X}'\mathbf{y} = \begin{pmatrix} \sum_t y_{t,1} \\ \vdots \\ \sum_t y_{t,V} \end{pmatrix}$$

Toy example

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}, \quad \mathbf{e} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_{TV})$$

<u>Frequentist estimate of $\boldsymbol{\beta}$</u>

(MLE equivalent to OLS)

$$\hat{\boldsymbol{\beta}}^{MLE} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = (\bar{y}_1, \cdots \bar{y}_V)'$$

$\mathbf{x} = (1,1,\ldots,1)'$

$$\mathbf{X} = \begin{pmatrix} \mathbf{x} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{x} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{x} \end{pmatrix} \quad \mathbf{X}'\mathbf{X} = \begin{pmatrix} T & & \\ & \ddots & \\ & & T \end{pmatrix} \quad \mathbf{X}'\mathbf{y} = \begin{pmatrix} \sum_t y_{t,1} \\ \vdots \\ \sum_t y_{t,V} \end{pmatrix}$$

<u>Bayesian estimate of $\boldsymbol{\beta}$:</u> $\hat{\boldsymbol{\beta}}^{MAP} = \left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{X} + \mathbf{Q}\right)^{-1}\left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{y}\right)$

Toy example

Bayesian estimate of $\boldsymbol{\beta}$: $\hat{\boldsymbol{\beta}}^{MAP} = \left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{X} + \mathbf{Q}\right)^{-1}\left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{y}\right)$

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}, \quad \mathbf{e} \sim N(\mathbf{0}, \sigma^2\mathbf{I}_{TV})$$

Frequentist estimate of $\boldsymbol{\beta}$

(MLE equivalent to OLS)

$$\hat{\boldsymbol{\beta}}^{MLE} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = (\bar{y}_1, \cdots \bar{y}_V)'$$

$\mathbf{x} = (1,1,\ldots,1)'$

$$\mathbf{X} = \begin{pmatrix} \mathbf{x} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{x} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{x} \end{pmatrix} \quad \mathbf{X}'\mathbf{X} = \begin{pmatrix} T & & \\ & \ddots & \\ & & T \end{pmatrix} \quad \mathbf{X}'\mathbf{y} = \begin{pmatrix} \sum_t y_{t,1} \\ \vdots \\ \sum_t y_{t,V} \end{pmatrix}$$

Toy example

Bayesian estimate of $\boldsymbol{\beta}$: $\hat{\boldsymbol{\beta}}^{MAP} = \left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{X} + \mathbf{Q}\right)^{-1}\left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{y}\right)$

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}, \quad \mathbf{e} \sim N(\mathbf{0}, \sigma^2\mathbf{I}_{TV})$$

Frequentist estimate of $\boldsymbol{\beta}$

(MLE equivalent to OLS)

$$\hat{\boldsymbol{\beta}}^{MLE} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = (\bar{y}_1, \cdots \bar{y}_V)'$$

$\mathbf{x} = (1,1,\ldots,1)'$

$$\mathbf{X} = \begin{pmatrix} \mathbf{x} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{x} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{x} \end{pmatrix} \quad \mathbf{X}'\mathbf{X} = \begin{pmatrix} T & & \\ & \ddots & \\ & & T \end{pmatrix} \quad \mathbf{X}'\mathbf{y} = \begin{pmatrix} \sum_t y_{t,1} \\ \vdots \\ \sum_t y_{t,V} \end{pmatrix}$$

What if $\mathbf{Q} \rightarrow$ a matrix of zero's?

(Prior var $\rightarrow \infty$)

$$\hat{\boldsymbol{\beta}}^{MAP} \rightarrow \left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{X}\right)^{-1}\left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{y}\right) = \hat{\boldsymbol{\beta}}^{MLE}$$

Toy example

Bayesian estimate of $\boldsymbol{\beta}$: $\hat{\beta}^{MAP} = \left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{X} + \mathbf{Q}\right)^{-1}\left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{y}\right)$

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}, \quad \mathbf{e} \sim N(\mathbf{0}, \sigma^2\mathbf{I}_{TV})$$

Frequentist estimate of $\boldsymbol{\beta}$

(MLE equivalent to OLS)

$$\hat{\beta}^{MLE} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = (\bar{y}_1, \cdots \bar{y}_V)'$$

$\mathbf{x} = (1,1,\ldots,1)'$

$$\mathbf{X} = \begin{pmatrix} \mathbf{x} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{x} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{x} \end{pmatrix} \quad \mathbf{X}'\mathbf{X} = \begin{pmatrix} T & & \\ & \ddots & \\ & & T \end{pmatrix} \quad \mathbf{X}'\mathbf{y} = \begin{pmatrix} \sum_t y_{t,1} \\ \vdots \\ \sum_t y_{t,V} \end{pmatrix}$$

What if $\mathbf{Q}$ equals a diagonal matrix, $\frac{1}{\sigma_0^2}\mathbf{I}_V$?

(no dependence information, prior var = σ_0^2)

Toy example

<u>Bayesian estimate of $\boldsymbol{\beta}$:</u> $\hat{\boldsymbol{\beta}}^{MAP} = \left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{X} + \mathbf{Q}\right)^{-1}\left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{y}\right)$

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}, \quad \mathbf{e} \sim N(\mathbf{0}, \sigma^2\mathbf{I}_{TV})$$

<u>Frequentist estimate of $\boldsymbol{\beta}$</u>

(MLE equivalent to OLS)

$$\hat{\boldsymbol{\beta}}^{MLE} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = (\bar{y}_1, \cdots \bar{y}_V)'$$

$\mathbf{x} = (1,1,\ldots,1)'$

$$\mathbf{X} = \begin{pmatrix} \mathbf{x} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{x} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{x} \end{pmatrix} \quad \mathbf{X}'\mathbf{X} = \begin{pmatrix} T & & \\ & \ddots & \\ & & T \end{pmatrix} \quad \mathbf{X}'\mathbf{y} = \begin{pmatrix} \sum_t y_{t,1} \\ \vdots \\ \sum_t y_{t,V} \end{pmatrix}$$

What if $\mathbf{Q}$ equals a diagonal matrix, $\frac{1}{\sigma_0^2}\mathbf{I}_V$?

(no dependence information, prior var = σ_0^2)

$$\left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{X} + \mathbf{Q}\right) = \left(\frac{T}{\sigma^2} + \frac{1}{\sigma_0^2}\right)\mathbf{I}_V$$

Toy example

Bayesian estimate of $\boldsymbol{\beta}$: $\hat{\boldsymbol{\beta}}^{MAP} = \left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{X} + \mathbf{Q}\right)^{-1}\left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{y}\right)$

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}, \quad \mathbf{e} \sim N(\mathbf{0}, \sigma^2\mathbf{I}_{TV})$$

Frequentist estimate of $\boldsymbol{\beta}$

(MLE equivalent to OLS)

$$\hat{\boldsymbol{\beta}}^{MLE} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = (\bar{y}_1, \cdots \bar{y}_V)'$$

$\mathbf{x} = (1,1,\ldots,1)'$

$$\mathbf{X} = \begin{pmatrix} \mathbf{x} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{x} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{x} \end{pmatrix} \quad \mathbf{X}'\mathbf{X} = \begin{pmatrix} T & & \\ & \ddots & \\ & & T \end{pmatrix} \quad \mathbf{X}'\mathbf{y} = \begin{pmatrix} \sum_t y_{t,1} \\ \vdots \\ \sum_t y_{t,V} \end{pmatrix}$$

What if $\mathbf{Q}$ equals a diagonal matrix, $\frac{1}{\sigma_0^2}\mathbf{I}_V$?

(no dependence information, prior var = σ_0^2)

$$\left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{X} + \mathbf{Q}\right) = \left(\frac{T}{\sigma^2} + \frac{1}{\sigma_0^2}\right)\mathbf{I}_V$$

$$\frac{1}{\sigma^2}\mathbf{X}'\mathbf{y} = \frac{T}{\sigma^2}(\bar{y}_1, \cdots, \bar{y}_V)' = \frac{T}{\sigma^2}\hat{\boldsymbol{\beta}}^{MLE}$$

Toy example

Bayesian estimate of $\boldsymbol{\beta}$: $\hat{\boldsymbol{\beta}}^{MAP} = \left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{X} + \mathbf{Q}\right)^{-1}\left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{y}\right)$

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}, \quad \mathbf{e} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_{TV})$$

Frequentist estimate of $\boldsymbol{\beta}$

(MLE equivalent to OLS)

$$\hat{\boldsymbol{\beta}}^{MLE} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = (\bar{y}_1, \cdots \bar{y}_V)'$$

$\mathbf{x} = (1,1,\ldots,1)'$

$$\mathbf{X} = \begin{pmatrix} \mathbf{x} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{x} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{x} \end{pmatrix} \quad \mathbf{X}'\mathbf{X} = \begin{pmatrix} T & & \\ & \ddots & \\ & & T \end{pmatrix} \quad \mathbf{X}'\mathbf{y} = \begin{pmatrix} \sum_t y_{t,1} \\ \vdots \\ \sum_t y_{t,V} \end{pmatrix}$$

What if $\mathbf{Q}$ equals a diagonal matrix, $\frac{1}{\sigma_0^2}\mathbf{I}_V$?

(no dependence information, prior var = σ_0^2)

$$\left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{X} + \mathbf{Q}\right) = \left(\frac{T}{\sigma^2} + \frac{1}{\sigma_0^2}\right)\mathbf{I}_V$$

$$\frac{1}{\sigma^2}\mathbf{X}'\mathbf{y} = \frac{T}{\sigma^2}(\bar{y}_1, \cdots, \bar{y}_V)' = \frac{T}{\sigma^2}\hat{\boldsymbol{\beta}}^{MLE}$$

$$\Rightarrow \hat{\boldsymbol{\beta}}^{MAP} = \frac{\sigma_0^2}{\sigma_0^2 + \sigma^2/T}\hat{\boldsymbol{\beta}}^{MLE} = \lambda\hat{\boldsymbol{\beta}}^{MLE}$$

Toy example

Bayesian estimate of $\boldsymbol{\beta}$: $\hat{\boldsymbol{\beta}}^{MAP} = \left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{X} + \mathbf{Q}\right)^{-1}\left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{y}\right)$

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}, \quad \mathbf{e} \sim N(\mathbf{0}, \sigma^2\mathbf{I}_{TV})$$

Frequentist estimate of $\boldsymbol{\beta}$

(MLE equivalent to OLS)

$$\hat{\boldsymbol{\beta}}^{MLE} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = (\bar{y}_1, \cdots \bar{y}_V)'$$

$\mathbf{x} = (1,1,\ldots,1)'$

$$\mathbf{X} = \begin{pmatrix} \mathbf{x} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{x} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{x} \end{pmatrix} \quad \mathbf{X}'\mathbf{X} = \begin{pmatrix} T & & \\ & \ddots & \\ & & T \end{pmatrix} \quad \mathbf{X}'\mathbf{y} = \begin{pmatrix} \sum_t y_{t,1} \\ \vdots \\ \sum_t y_{t,V} \end{pmatrix}$$

What if $\mathbf{Q}$ equals a diagonal matrix, $\frac{1}{\sigma_0^2}\mathbf{I}_V$?

(no dependence information, prior var = σ_0^2)

$$\left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{X} + \mathbf{Q}\right) = \left(\frac{T}{\sigma^2} + \frac{1}{\sigma_0^2}\right)\mathbf{I}_V$$

$$\frac{1}{\sigma^2}\mathbf{X}'\mathbf{y} = \frac{T}{\sigma^2}(\bar{y}_1, \cdots, \bar{y}_V)' = \frac{T}{\sigma^2}\hat{\boldsymbol{\beta}}^{MLE}$$

$$\Rightarrow \hat{\boldsymbol{\beta}}^{MAP} = \frac{\sigma_0^2}{\sigma_0^2 + \sigma^2/T}\hat{\boldsymbol{\beta}}^{MLE} = \lambda\hat{\boldsymbol{\beta}}^{MLE}$$

If $T = 100$, $\sigma^2 = 1$, $\sigma_0^2 = 1/5$, then $\lambda = 0.988$

Toy example

Bayesian estimate of $\boldsymbol{\beta}$: $\hat{\boldsymbol{\beta}}^{MAP} = \left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{X} + \mathbf{Q}\right)^{-1}\left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{y}\right)$

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}, \quad \mathbf{e} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_{TV})$$

Frequentist estimate of $\boldsymbol{\beta}$

(MLE equivalent to OLS)

$$\hat{\boldsymbol{\beta}}^{MLE} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = (\bar{y}_1, \cdots \bar{y}_V)'$$

$\mathbf{x} = (1,1,\ldots,1)'$

$$\mathbf{X} = \begin{pmatrix} \mathbf{x} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{x} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{x} \end{pmatrix} \quad \mathbf{X}'\mathbf{X} = \begin{pmatrix} T & & \\ & \ddots & \\ & & T \end{pmatrix} \quad \mathbf{X}'\mathbf{y} = \begin{pmatrix} \sum_t y_{t,1} \\ \vdots \\ \sum_t y_{t,V} \end{pmatrix}$$

What if $\mathbf{Q}$ equals a diagonal matrix, $\frac{1}{\sigma_0^2}\mathbf{I}_V$?

(no dependence information, prior var = σ_0^2)

$$\left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{X} + \mathbf{Q}\right) = \left(\frac{T}{\sigma^2} + \frac{1}{\sigma_0^2}\right)\mathbf{I}_V$$

$$\frac{1}{\sigma^2}\mathbf{X}'\mathbf{y} = \frac{T}{\sigma^2}(\bar{y}_1, \cdots, \bar{y}_V)' = \frac{T}{\sigma^2}\hat{\boldsymbol{\beta}}^{MLE}$$

$$\Rightarrow \hat{\boldsymbol{\beta}}^{MAP} = \frac{\sigma_0^2}{\sigma_0^2 + \sigma^2/T}\hat{\boldsymbol{\beta}}^{MLE} = \lambda\hat{\boldsymbol{\beta}}^{MLE}$$

If $T = 100$, $\sigma^2 = 1$, $\sigma_0^2 = 1/5$, then $\lambda = 0.988$

Toy example

Bayesian estimate of $\boldsymbol{\beta}$: $\hat{\beta}^{MAP} = \left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{X} + \mathbf{Q}\right)^{-1}\left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{y}\right)$

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}, \quad \mathbf{e} \sim N(\mathbf{0}, \sigma^2\mathbf{I}_{TV})$$

Frequentist estimate of $\boldsymbol{\beta}$

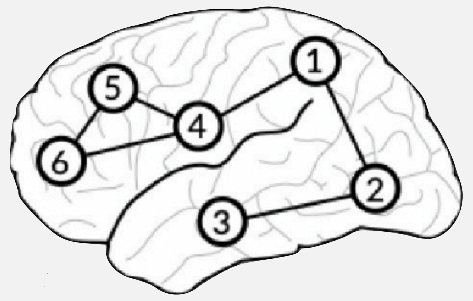
(MLE equivalent to OLS)

$$\hat{\beta}^{MLE} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = (\bar{y}_1, \cdots \bar{y}_V)'$$

$\mathbf{x} = (1,1,\ldots,1)'$

$$\mathbf{X} = \begin{pmatrix} \mathbf{x} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{x} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{x} \end{pmatrix} \quad \mathbf{X}'\mathbf{X} = \begin{pmatrix} T & & \\ & \ddots & \\ & & T \end{pmatrix} \quad \mathbf{X}'\mathbf{y} = \begin{pmatrix} \sum_t y_{t,1} \\ \vdots \\ \sum_t y_{t,V} \end{pmatrix}$$

What if $\mathbf{Q} = \begin{pmatrix} 5 & -1 & 0 & -1 & 0 & 0 \\ -1 & 5 & -1 & 0 & 0 & 0 \\ 0 & -1 & 5 & 0 & 0 & 0 \\ -1 & 0 & 0 & 5 & -2 & -2 \\ 0 & 0 & 0 & -2 & 5 & -2 \\ 0 & 0 & 0 & -2 & -2 & 5 \end{pmatrix}$ **?**



Toy example

Bayesian estimate of $\boldsymbol{\beta}$: $\hat{\boldsymbol{\beta}}^{MAP} = \left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{X} + \mathbf{Q}\right)^{-1}\left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{y}\right)$

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}, \quad \mathbf{e} \sim N(\mathbf{0}, \sigma^2\mathbf{I}_{TV})$$

Frequentist estimate of $\boldsymbol{\beta}$

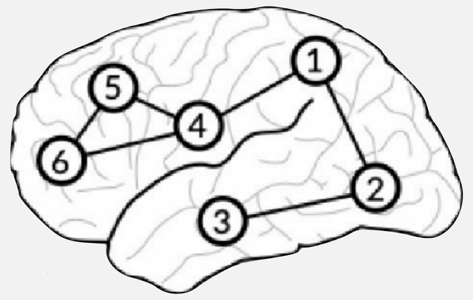
(MLE equivalent to OLS)

$$\hat{\boldsymbol{\beta}}^{MLE} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = (\bar{y}_1, \cdots \bar{y}_V)'$$

$\mathbf{x} = (1,1,\ldots,1)'$

$$\mathbf{X} = \begin{pmatrix} \mathbf{x} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{x} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{x} \end{pmatrix} \quad \mathbf{X}'\mathbf{X} = \begin{pmatrix} T & & \\ & \ddots & \\ & & T \end{pmatrix} \quad \mathbf{X}'\mathbf{y} = \begin{pmatrix} \sum_t y_{t,1} \\ \vdots \\ \sum_t y_{t,V} \end{pmatrix}$$

What if $\mathbf{Q} = \begin{pmatrix} 5 & -1 & 0 & -1 & 0 & 0 \\ -1 & 5 & -1 & 0 & 0 & 0 \\ 0 & -1 & 5 & 0 & 0 & 0 \\ -1 & 0 & 0 & 5 & -2 & -2 \\ 0 & 0 & 0 & -2 & 5 & -2 \\ 0 & 0 & 0 & -2 & -2 & 5 \end{pmatrix}$ **?**



We just showed that

$$\hat{\boldsymbol{\beta}}^{MAP} = \frac{T}{\sigma^2}\left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{X} + \mathbf{Q}\right)^{-1}\hat{\boldsymbol{\beta}}^{MLE} = \boldsymbol{\Lambda}\hat{\boldsymbol{\beta}}^{MLE}.$$

Toy example

Bayesian estimate of $\boldsymbol{\beta}$: $\hat{\boldsymbol{\beta}}^{MAP} = \left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{X} + \mathbf{Q}\right)^{-1}\left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{y}\right)$

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}, \quad \mathbf{e} \sim N(\mathbf{0}, \sigma^2\mathbf{I}_{TV})$$

Frequentist estimate of $\boldsymbol{\beta}$

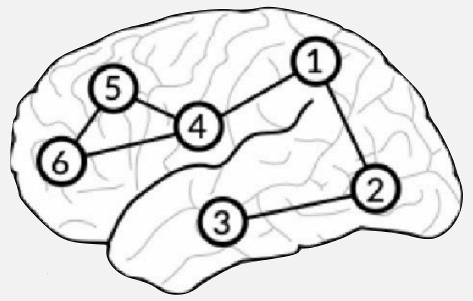
(MLE equivalent to OLS)

$$\hat{\boldsymbol{\beta}}^{MLE} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = (\bar{y}_1, \cdots \bar{y}_V)'$$

$\mathbf{x} = (1,1,\ldots,1)'$

$$\mathbf{X} = \begin{pmatrix} \mathbf{x} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{x} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{x} \end{pmatrix} \quad \mathbf{X}'\mathbf{X} = \begin{pmatrix} T & & \\ & \ddots & \\ & & T \end{pmatrix} \quad \mathbf{X}'\mathbf{y} = \begin{pmatrix} \sum_t y_{t,1} \\ \vdots \\ \sum_t y_{t,V} \end{pmatrix}$$

What if $\mathbf{Q} = \begin{pmatrix} 5 & -1 & 0 & -1 & 0 & 0 \\ -1 & 5 & -1 & 0 & 0 & 0 \\ 0 & -1 & 5 & 0 & 0 & 0 \\ -1 & 0 & 0 & 5 & -2 & -2 \\ 0 & 0 & 0 & -2 & 5 & -2 \\ 0 & 0 & 0 & -2 & -2 & 5 \end{pmatrix}$ **?**



We just showed that

$$\hat{\boldsymbol{\beta}}^{MAP} = \frac{T}{\sigma^2}\left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{X} + \mathbf{Q}\right)^{-1}\hat{\boldsymbol{\beta}}^{MLE} = \boldsymbol{\Lambda}\hat{\boldsymbol{\beta}}^{MLE}.$$

If $T = 100$ and $\sigma^2 = 1$, $\boldsymbol{\Lambda} = \begin{pmatrix} 0.95 & 0.01 & 0 & 0.01 & 0 & 0 \\ 0.01 & 0.95 & 0.01 & 0 & 0 & 0 \\ 0 & 0.01 & 0.95 & 0 & 0 & 0 \\ 0.01 & 0 & 0 & 0.95 & 0.02 & 0.02 \\ 0 & 0 & 0 & 0.02 & 0.95 & 0.02 \\ 0 & 0 & 0 & 0.02 & 0.02 & 0.95 \end{pmatrix}$

Toy example

<u>Bayesian estimate of $\boldsymbol{\beta}$:</u> $\hat{\boldsymbol{\beta}}^{MAP} = \left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{X} + \mathbf{Q}\right)^{-1}\left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{y}\right)$

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}, \quad \mathbf{e} \sim N(\mathbf{0}, \sigma^2\mathbf{I}_{TV})$$

<u>Frequentist estimate of $\boldsymbol{\beta}$</u>

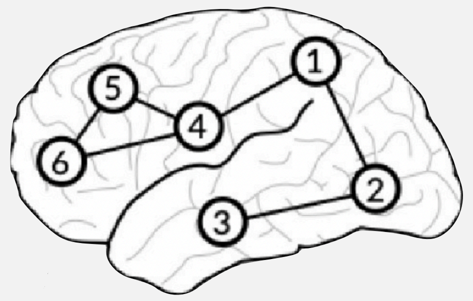
(MLE equivalent to OLS)

$$\hat{\boldsymbol{\beta}}^{MLE} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = (\bar{y}_1, \cdots \bar{y}_V)'$$

$\mathbf{x} = (1,1,\ldots,1)'$

$$\mathbf{X} = \begin{pmatrix} \mathbf{x} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{x} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{x} \end{pmatrix} \quad \mathbf{X}'\mathbf{X} = \begin{pmatrix} T & & \\ & \ddots & \\ & & T \end{pmatrix} \quad \mathbf{X}'\mathbf{y} = \begin{pmatrix} \sum_t y_{t,1} \\ \vdots \\ \sum_t y_{t,V} \end{pmatrix}$$

What if $\mathbf{Q} = \begin{pmatrix} 5 & -1 & 0 & -1 & 0 & 0 \\ -1 & 5 & -1 & 0 & 0 & 0 \\ 0 & -1 & 5 & 0 & 0 & 0 \\ -1 & 0 & 0 & 5 & -2 & -2 \\ 0 & 0 & 0 & -2 & 5 & -2 \\ 0 & 0 & 0 & -2 & -2 & 5 \end{pmatrix}$ **?**



We just showed that

$$\hat{\boldsymbol{\beta}}^{MAP} = \frac{T}{\sigma^2}\left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{X} + \mathbf{Q}\right)^{-1}\hat{\boldsymbol{\beta}}^{MLE} = \boldsymbol{\Lambda}\hat{\boldsymbol{\beta}}^{MLE}.$$

If $T = 100$ and $\sigma^2 = 1$, $\boldsymbol{\Lambda} = \begin{pmatrix} 0.95 & 0.01 & 0 & 0.01 & 0 & 0 \\ 0.01 & 0.95 & 0.01 & 0 & 0 & 0 \\ 0 & 0.01 & 0.95 & 0 & 0 & 0 \\ 0.01 & 0 & 0 & 0.95 & 0.02 & 0.02 \\ 0 & 0 & 0 & 0.02 & 0.95 & 0.02 \\ 0 & 0 & 0 & 0.02 & 0.02 & 0.95 \end{pmatrix}$

Toy example

Bayesian estimate of $\boldsymbol{\beta}$: $\hat{\boldsymbol{\beta}}^{MAP} = \left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{X} + \mathbf{Q}\right)^{-1}\left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{y}\right)$

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}, \quad \mathbf{e} \sim N(\mathbf{0}, \sigma^2\mathbf{I}_{TV})$$

Frequentist estimate of $\boldsymbol{\beta}$

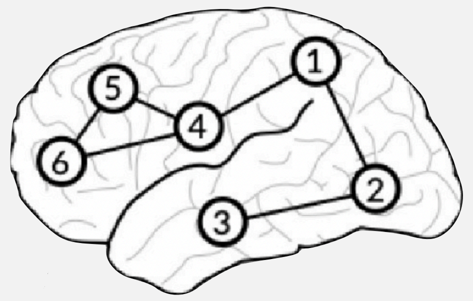
(MLE equivalent to OLS)

$$\hat{\boldsymbol{\beta}}^{MLE} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = (\bar{y}_1, \cdots \bar{y}_V)'$$

$\mathbf{x} = (1,1,\ldots,1)'$

$$\mathbf{X} = \begin{pmatrix} \mathbf{x} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{x} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{x} \end{pmatrix} \quad \mathbf{X}'\mathbf{X} = \begin{pmatrix} T & & \\ & \ddots & \\ & & T \end{pmatrix} \quad \mathbf{X}'\mathbf{y} = \begin{pmatrix} \sum_t y_{t,1} \\ \vdots \\ \sum_t y_{t,V} \end{pmatrix}$$

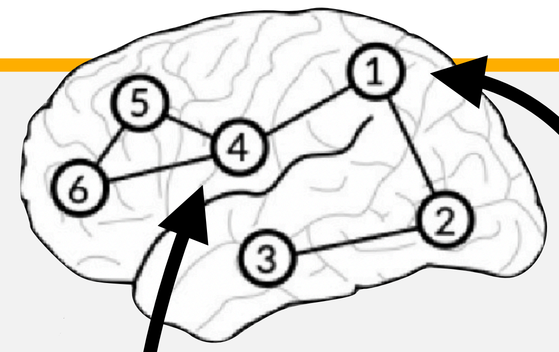
What if $\mathbf{Q} = \begin{pmatrix} 5 & -1 & 0 & -1 & 0 & 0 \\ -1 & 5 & -1 & 0 & 0 & 0 \\ 0 & -1 & 5 & 0 & 0 & 0 \\ -1 & 0 & 0 & 5 & -2 & -2 \\ 0 & 0 & 0 & -2 & 5 & -2 \\ 0 & 0 & 0 & -2 & -2 & 5 \end{pmatrix}$ **?**



We just showed that

$$\hat{\boldsymbol{\beta}}^{MAP} = \frac{T}{\sigma^2}\left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{X} + \mathbf{Q}\right)^{-1}\hat{\boldsymbol{\beta}}^{MLE} = \boldsymbol{\Lambda}\hat{\boldsymbol{\beta}}^{MLE}.$$

If $T = 100$ and $\sigma^2 = 1$, $\boldsymbol{\Lambda} = \begin{pmatrix} 0.95 & 0.01 & 0 & 0.01 & 0 & 0 \\ 0.01 & 0.95 & 0.01 & 0 & 0 & 0 \\ 0 & 0.01 & 0.95 & 0 & 0 & 0 \\ 0.01 & 0 & 0 & 0.95 & 0.02 & 0.02 \\ 0 & 0 & 0 & 0.02 & 0.95 & 0.02 \\ 0 & 0 & 0 & 0.02 & 0.02 & 0.95 \end{pmatrix}$



For example:

$$\hat{\beta}_1^{MAP} = 0.95\hat{\beta}_1^{MLE} + 0.01\hat{\beta}_2^{MLE} + 0.01\hat{\beta}_4^{MLE}$$

$$\hat{\beta}_4^{MAP} = 0.95\hat{\beta}_4^{MLE} + 0.01\hat{\beta}_1^{MLE} + 0.02\hat{\beta}_5^{MLE} + 0.02\hat{\beta}_6^{MLE}$$

What about the variance?

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}, \quad \mathbf{e} \sim N(\mathbf{0}, \sigma^2\mathbf{I}_{TV})$$

<u>Frequentist estimate of $\boldsymbol{\beta}$</u>

$$\hat{\boldsymbol{\beta}}^{MLE} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = (\bar{y}_1, \cdots \bar{y}_V)'$$

$$Cov(\hat{\boldsymbol{\beta}}^{MLE}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$$

<u>Bayesian estimate of $\boldsymbol{\beta}$</u>

$$\hat{\boldsymbol{\beta}}^{MAP} = \left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{X} + \mathbf{Q}\right)^{-1}\left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{y}\right)$$

$$Cov(\hat{\boldsymbol{\beta}}^{MLE}) = \left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{X} + \mathbf{Q}\right)^{-1}$$

What about the variance?

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}, \quad \mathbf{e} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_{TV})$$

Let $T = 100, \sigma^2 = 1, \sigma_0^2 = 1/5$

Frequentist estimate of $\boldsymbol{\beta}$

$$\hat{\boldsymbol{\beta}}^{MLE} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = (\bar{y}_1, \cdots \bar{y}_V)'$$

$$Cov(\hat{\boldsymbol{\beta}}^{MLE}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$$

$$Var(\hat{\beta}^{MLE}_{(v)}) = 0.01$$

Bayesian estimate of $\boldsymbol{\beta}$

$$\hat{\boldsymbol{\beta}}^{MAP} = \left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{X} + \mathbf{Q}\right)^{-1}\left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{y}\right)$$

$$Cov(\hat{\boldsymbol{\beta}}^{MLE}) = \left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{X} + \mathbf{Q}\right)^{-1}$$

What about the variance?

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}, \quad \mathbf{e} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_{TV})$$

Let $T = 100$, $\sigma^2 = 1$, $\sigma_0^2 = 1/5$

<u>Frequentist estimate of $\boldsymbol{\beta}$</u>

$$\hat{\boldsymbol{\beta}}^{MLE} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = (\bar{y}_1, \cdots \bar{y}_V)'$$

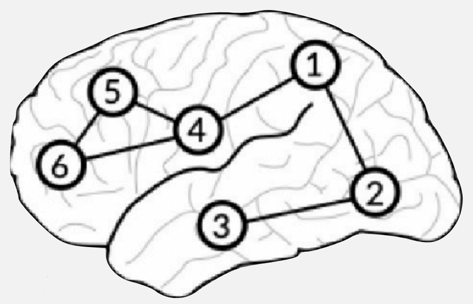
$$Cov(\hat{\boldsymbol{\beta}}^{MLE}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$$

$$Var(\hat{\beta}^{MLE}_{(v)}) = 0.01$$

<u>Bayesian estimate of $\boldsymbol{\beta}$</u>

$$\hat{\boldsymbol{\beta}}^{MAP} = \left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{X} + \mathbf{Q}\right)^{-1}\left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{y}\right)$$

$$Cov(\hat{\boldsymbol{\beta}}^{MLE}) = \left(\frac{1}{\sigma^2}\mathbf{X}'\mathbf{X} + \mathbf{Q}\right)^{-1}$$



$$\mathbf{Q} = \begin{pmatrix} 5 & -1 & 0 & -1 & 0 & 0 \\ -1 & 5 & -1 & 0 & 0 & 0 \\ 0 & -1 & 5 & 0 & 0 & 0 \\ -1 & 0 & 0 & 5 & -2 & -2 \\ 0 & 0 & 0 & -2 & 5 & -2 \\ 0 & 0 & 0 & -2 & -2 & 5 \end{pmatrix}$$

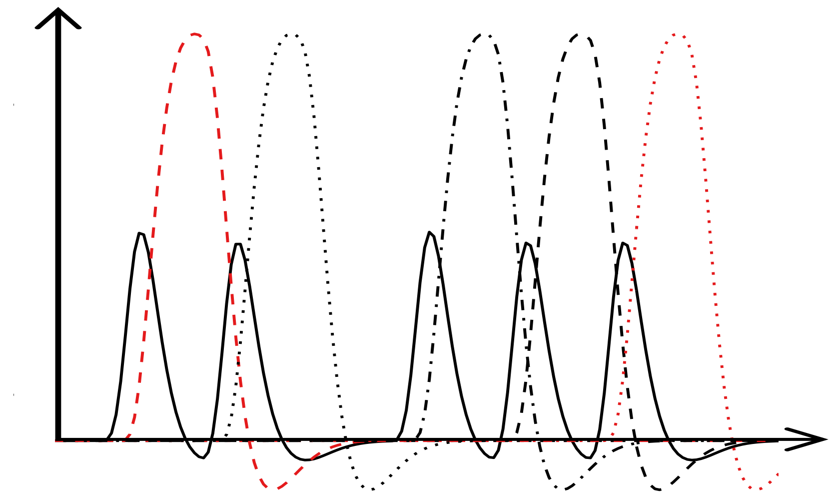
$$Cov(\hat{\boldsymbol{\beta}}^{MAP}) = \begin{pmatrix} 0.0095 & 0.0001 & 0 & 0.0001 & 0 & 0 \\ 0.01 & 0.0095 & 0.0001 & 0 & 0 & 0 \\ 0 & 0.0001 & 0.0095 & 0 & 0 & 0 \\ 0.0001 & 0 & 0 & 0.0095 & 0.0002 & 0.0002 \\ 0 & 0 & 0 & 0.0002 & 0.0095 & 0.0002 \\ 0 & 0 & 0 & 0.0002 & 0.0002 & 0.0095 \end{pmatrix}$$

$Var(\hat{\beta}^{MLE}_{(4)}) = 0.0095$ ➔ **5% lower variance!**

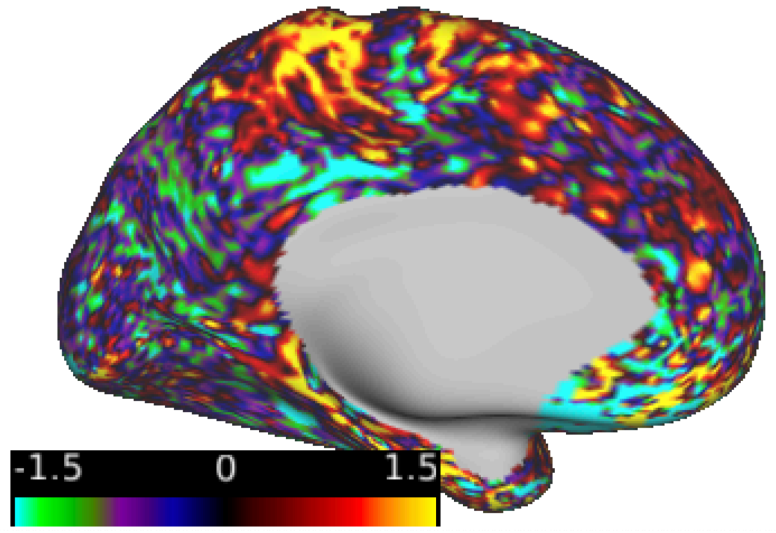
➔ more powerful inference (especially if we take into account the covariances between locations)

Spatial Bayesian GLM

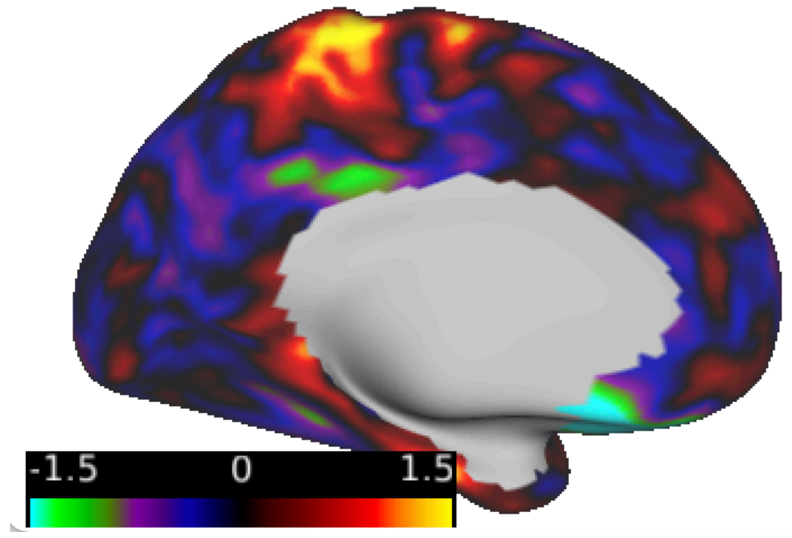
Spatial Bayesian GLM



GLM (General Linear Model): Estimate amplitude and areas of activation during a task or stimulus



Classical GLM: Fit a separate linear model at every voxel or vertex ("massive univariate") to estimate activation amplitude



Spatial Bayesian GLM: Fit a single Bayesian model accounting for spatial dependence via spatial priors for greater accuracy & power

The Classical GLM

$$\underset{(T \times 1)}{\mathbf{y}(v)} = \sum_{k=1}^{K} \beta_k(v)\underset{(T \times 1)}{\mathbf{x}_k} + \underset{(T \times 1)}{\boldsymbol{\epsilon}(v)},\ \boldsymbol{\epsilon}(v) \sim N(\mathbf{0}, \sigma^2\mathbf{I})$$

The Classical GLM

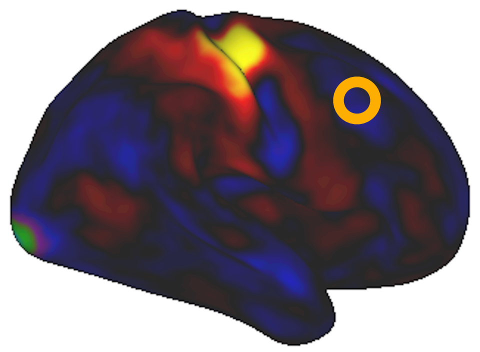
Observed BOLD signal at location v

$$\mathbf{y}(v) = \sum_{k=1}^{K} \beta_k(v)\mathbf{x}_k + \boldsymbol{\epsilon}(v),\ \boldsymbol{\epsilon}(v) \sim N(\mathbf{0}, \sigma^2\mathbf{I})$$

(T x 1) (T x 1) (T x 1)

The Classical GLM

Observed
BOLD signal
at location v

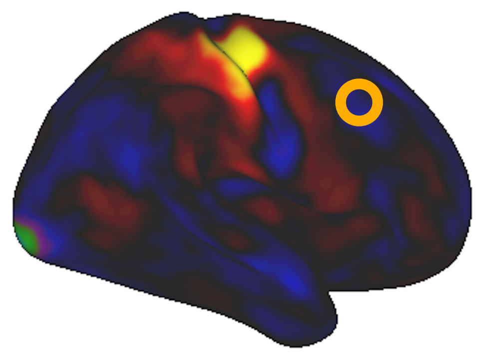


Amplitude of
activation for task k
at location v

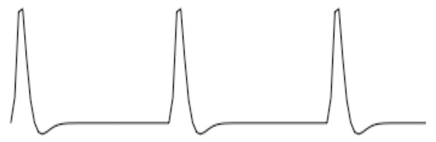
$$\underset{\text{(T x 1)}}{\mathbf{y}(v)} = \sum_{k=1}^{K} \beta_k(v)\underset{\text{(T x 1)}}{\mathbf{x}_k} + \underset{\text{(T x 1)}}{\boldsymbol{\epsilon}(v)},\ \boldsymbol{\epsilon}(v) \sim N(\mathbf{0}, \sigma^2\mathbf{I})$$

The Classical GLM

Observed BOLD signal at location v



Amplitude of activation for task k at location v

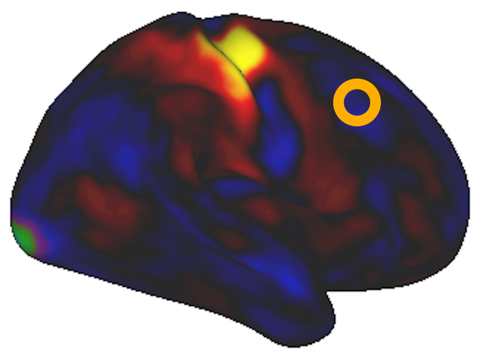


Expected BOLD response to task k

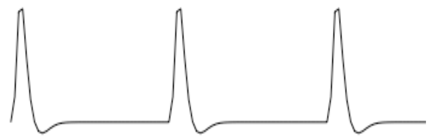
$$\underset{(\mathrm{T} \times 1)}{\mathbf{y}(v)} = \sum_{k=1}^{K} \beta_k(v) \underset{(\mathrm{T} \times 1)}{\mathbf{x}_k} + \underset{(\mathrm{T} \times 1)}{\boldsymbol{\epsilon}(v)}, \ \boldsymbol{\epsilon}(v) \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$$

The Classical GLM

Observed BOLD signal at location v



Amplitude of activation for task k at location v



Expected BOLD response to task k

$$\underset{(T \times 1)}{\mathbf{y}(v)} = \sum_{k=1}^{K} \beta_k(v) \underset{(T \times 1)}{\mathbf{x}_k} + \underset{(T \times 1)}{\epsilon(v)}, \ \epsilon(v) \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$$

No baseline field (data & design matrix have been centered)

No nuisance signals (Have been regressed from data & design)

Homoskedastic, uncorrelated residuals (data are prewhitened)

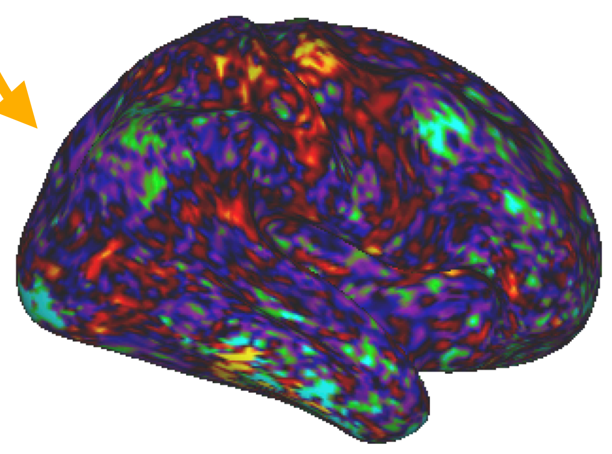
Several assumptions being made for simplicity

Classical GLM Activations

$$\mathbf{y}(1) = \sum_{k=0}^{K} \beta_k(1)\mathbf{x}_k + \boldsymbol{\epsilon}(1), \quad \boldsymbol{\epsilon}(1) \sim MVN(\mathbf{0}, \sigma^2\mathbf{I})$$

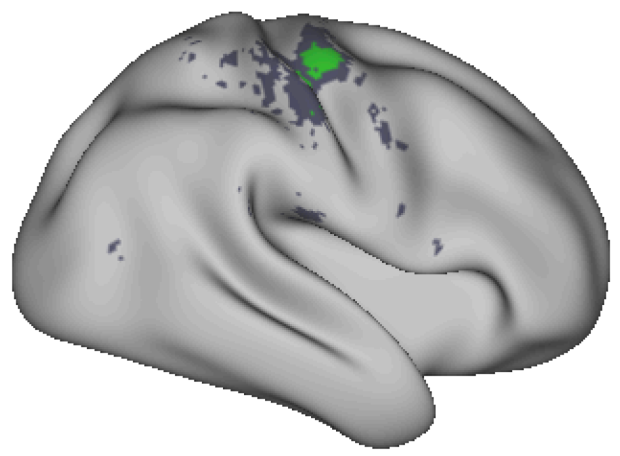
$$\vdots \qquad\qquad \vdots \qquad\qquad\qquad\qquad \vdots \qquad\qquad\qquad \vdots$$

$$\mathbf{y}(V) = \sum_{k=0}^{K} \beta_k(V)\mathbf{x}_k + \boldsymbol{\epsilon}(V), \ \boldsymbol{\epsilon}(V) \sim MVN(\mathbf{0}, \sigma^2\mathbf{I})$$



Hypothesis Testing +

Multiple Comparisons Correction



The Spatial Bayesian GLM

The classical GLM,
concatenated
across all locations

$$\boxed{\underset{(TV \times 1)}{\mathbf{y}} = \sum_{k=1}^{K} \underset{(TV \times V)}{\mathbf{X}_k} \underset{(V \times 1)}{\boldsymbol{\beta}_k} + \boldsymbol{\epsilon},\ \boldsymbol{\epsilon} \sim MVN(\mathbf{0}, \sigma^2 \mathbf{I}_{TV})}$$

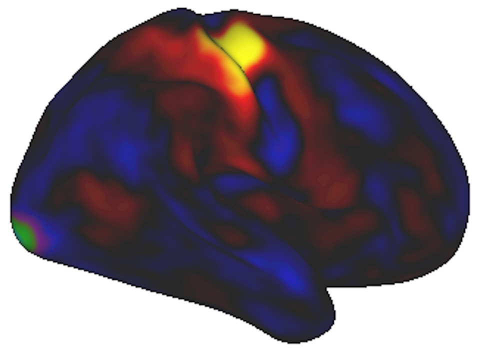
The Spatial Bayesian GLM

The classical GLM, concatenated across all locations

Observed BOLD signal, concatenated across all locations

$$\boxed{\underset{(TV \times 1)}{\mathbf{y}} = \sum_{k=1}^{K} \underset{(TV \times V)}{\mathbf{X}_k} \underset{(V \times 1)}{\boldsymbol{\beta}_k} + \boldsymbol{\epsilon},\ \boldsymbol{\epsilon} \sim MVN(\mathbf{0}, \sigma^2 \mathbf{I}_{TV})}$$

The Spatial Bayesian GLM



The classical GLM, concatenated across all locations

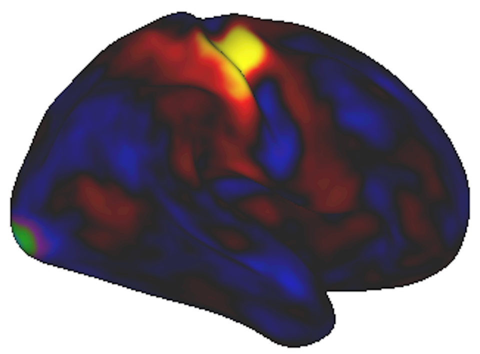
Observed BOLD signal, concatenated across all locations

Activation amplitude for task k, concatenated across all locations

$$\underset{(TV \times 1)}{\mathbf{y}} = \sum_{k=1}^{K} \underset{(TV \times V)}{\mathbf{X}_k} \underset{(V \times 1)}{\boldsymbol{\beta}_k} + \boldsymbol{\epsilon}, \; \boldsymbol{\epsilon} \sim MVN(\mathbf{0}, \sigma^2 \mathbf{I}_{TV})$$

The Spatial Bayesian GLM

$$\mathbf{X}_k = \begin{pmatrix} \mathbf{x}_k & 0 & \cdots & 0 \\ 0 & \mathbf{x}_k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{x}_k \end{pmatrix}$$



Observed BOLD signal, concatenated across all locations

Expected BOLD response to task k

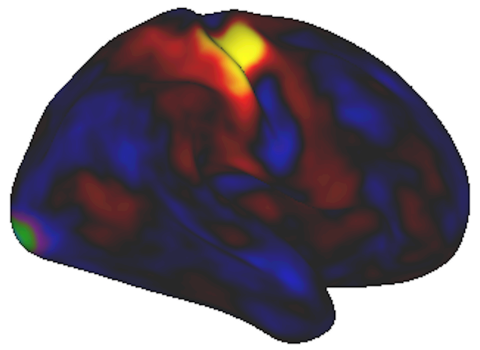
Activation amplitude for task k, concatenated across all locations

The classical GLM, concatenated across all locations

$$\underbrace{\mathbf{y}}_{(TV \times 1)} = \sum_{k=1}^{K} \underbrace{\mathbf{X}_k}_{(TV \times V)} \underbrace{\boldsymbol{\beta}_k}_{(V \times 1)} + \epsilon,\ \epsilon \sim MVN(\mathbf{0}, \sigma^2 \mathbf{I}_{TV})$$

The Spatial Bayesian GLM

$$\mathbf{X}_k = \begin{pmatrix} \mathbf{x}_k & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{x}_k & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{x}_k \end{pmatrix}$$



Observed BOLD signal, concatenated across all locations

Expected BOLD response to task k

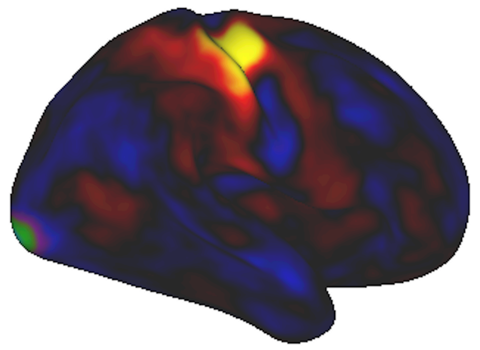
Activation amplitude for task k, concatenated across all locations

$$\underset{(TV \times 1)}{\mathbf{y}} = \sum_{k=1}^{K} \underset{(TV \times V)}{\mathbf{X}_k} \underset{(V \times 1)}{\boldsymbol{\beta}_k} + \boldsymbol{\epsilon}, \ \boldsymbol{\epsilon} \sim MVN(\mathbf{0}, \sigma^2 \mathbf{I}_{TV})$$

$$\boldsymbol{\beta}_k \sim N(\mathbf{0}, \mathbf{Q}_k^{-1})$$

The Spatial Bayesian GLM

$$\mathbf{X}_k = \begin{pmatrix} \mathbf{x}_k & 0 & \cdots & 0 \\ 0 & \mathbf{x}_k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{x}_k \end{pmatrix}$$



Observed BOLD signal, concatenated across all locations

Expected BOLD response to task k

Activation amplitude for task k, concatenated across all locations

$$\underset{(TV \times 1)}{\mathbf{y}} = \sum_{k=1}^{K} \underset{(TV \times V)}{\mathbf{X}_k} \underset{(V \times 1)}{\boldsymbol{\beta}_k} + \boldsymbol{\epsilon},\ \boldsymbol{\epsilon} \sim MVN(\mathbf{0}, \sigma^2 \mathbf{I}_{TV})$$

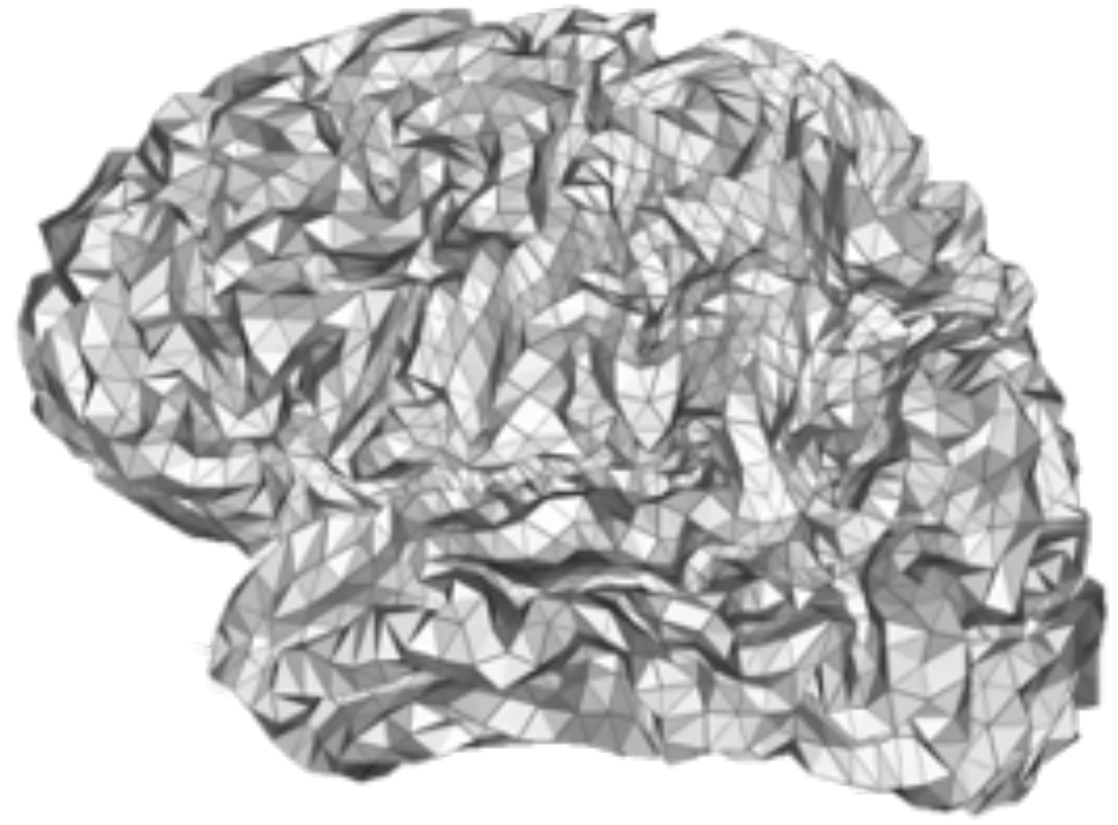
$$\boldsymbol{\beta}_k \sim N(\mathbf{0}, \mathbf{Q}_k^{-1})$$

spatial Bayesian GLM

Spatial priors for cortical surface task fMRI

We need/want a spatial prior that:

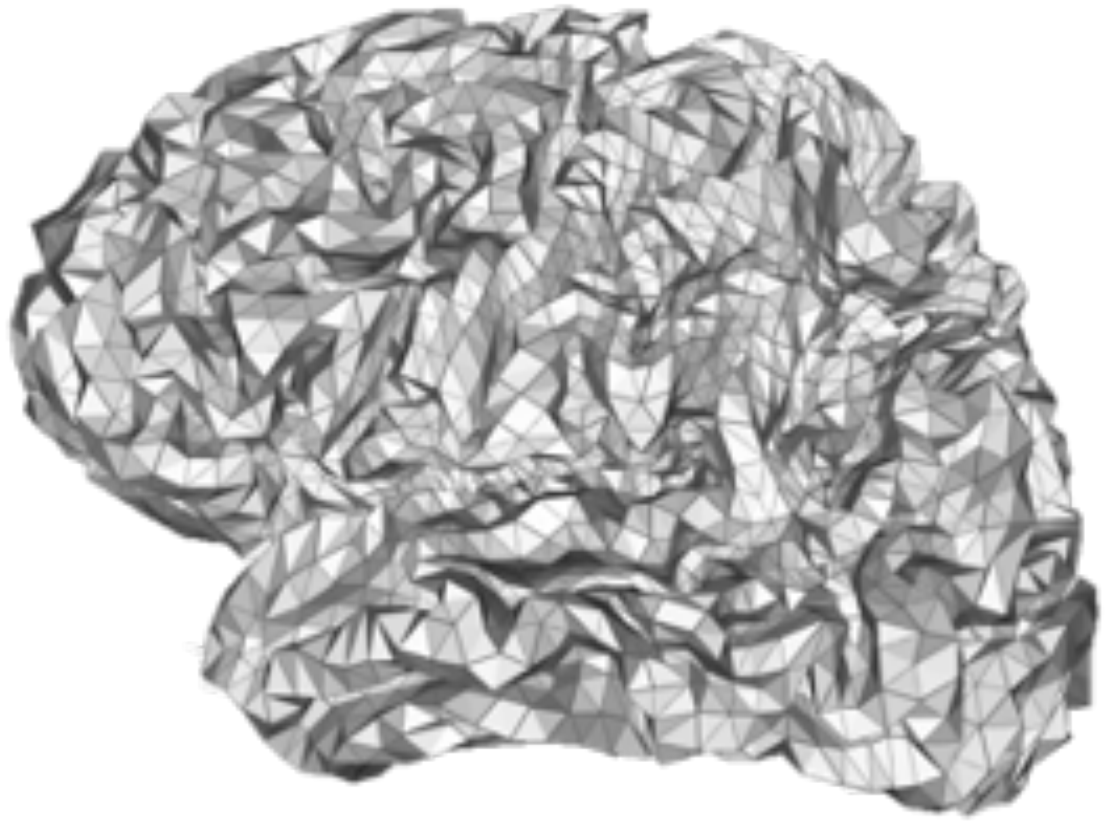
- Have sparse inverse covariance
- Is applicable to triangular mesh data
- Has a Gaussian distribution



Stochastic partial differential equation (SPDE) priors (Lindgren et al. 2011) check all the boxes!

Lindgren, et al. (2011). "An explicit link between Gaussian fields and Gaussian Markov random fields: the SPDE approach." *Journal of the Royal Statistical Society Series B: Statistical Methodology*.

SPDE Prior



$$\boldsymbol{\beta} \sim MVN(\mathbf{0}, \mathbf{Q}_{\kappa,\tau}^{-1})$$

$$\mathbf{Q}_{\kappa,\tau} = \tau^2(\kappa^4\mathbf{F} + 2\kappa^2\mathbf{G} + \mathbf{G}\mathbf{F}^{-1}\mathbf{G})$$

A fixed diagonal matrix

A fixed sparse matrix with non-zero entries for neighboring locations in mesh

κ controls the spatial correlation (must be estimated)

τ controls the variance (must be estimated)

Lindgren, et al. (2011). "An explicit link between Gaussian fields and Gaussian Markov random fields: the SPDE approach." *Journal of the Royal Statistical Society Series B: Statistical Methodology*.

Areas of Activation via Excursion Sets

Bolin, D., & Lindgren, F. (2015). Excursion and contour uncertainty regions for latent Gaussian models. *Journal of the Royal Statistical Society Series B: Statistical Methodology.*
Mejia, A. F., Yue, Y., Bolin, D., Lindgren, F., & Lindquist, M. A. (2020). A Bayesian general linear modeling approach to cortical surface fMRI data analysis. *Journal of the American Statistical Association.*

Areas of Activation via Excursion Sets

For each task k, we use the ***joint posterior distribution*** **of** $\boldsymbol{\beta}_k$
to identify areas of statistically significant activation above a minimum effect size of interest

Bolin, D., & Lindgren, F. (2015). Excursion and contour uncertainty regions for latent Gaussian models. *Journal of the Royal Statistical Society Series B: Statistical Methodology.*
Mejia, A. F., Yue, Y., Bolin, D., Lindgren, F., & Lindquist, M. A. (2020). A Bayesian general linear modeling approach to cortical surface fMRI data analysis. *Journal of the American Statistical Association.*

Areas of Activation via Excursion Sets

For each task k, we use the ***joint posterior distribution* of** $\boldsymbol{\beta}_k$
to identify areas of statistically significant activation above a minimum effect size of interest

This approach ***avoids the need to correct for multiple comparisons***, which would be required if we used the marginal posterior distribution at each location to identify areas of activation

Bolin, D., & Lindgren, F. (2015). Excursion and contour uncertainty regions for latent Gaussian models. *Journal of the Royal Statistical Society Series B: Statistical Methodology.*
Mejia, A. F., Yue, Y., Bolin, D., Lindgren, F., & Lindquist, M. A. (2020). A Bayesian general linear modeling approach to cortical surface fMRI data analysis. *Journal of the American Statistical Association.*

Areas of Activation via Excursion Sets

For each task k, we use the ***joint posterior distribution* of $\boldsymbol{\beta}_k$**
to identify areas of statistically significant activation above a minimum effect size of interest

This approach ***avoids the need to correct for multiple comparisons***, which would be required if we used the marginal posterior distribution at each location to identify areas of activation

There is a ***dramatic gain in power*** over massive univariate modeling, facilitating the use of a scientifically meaningful minimum effect size

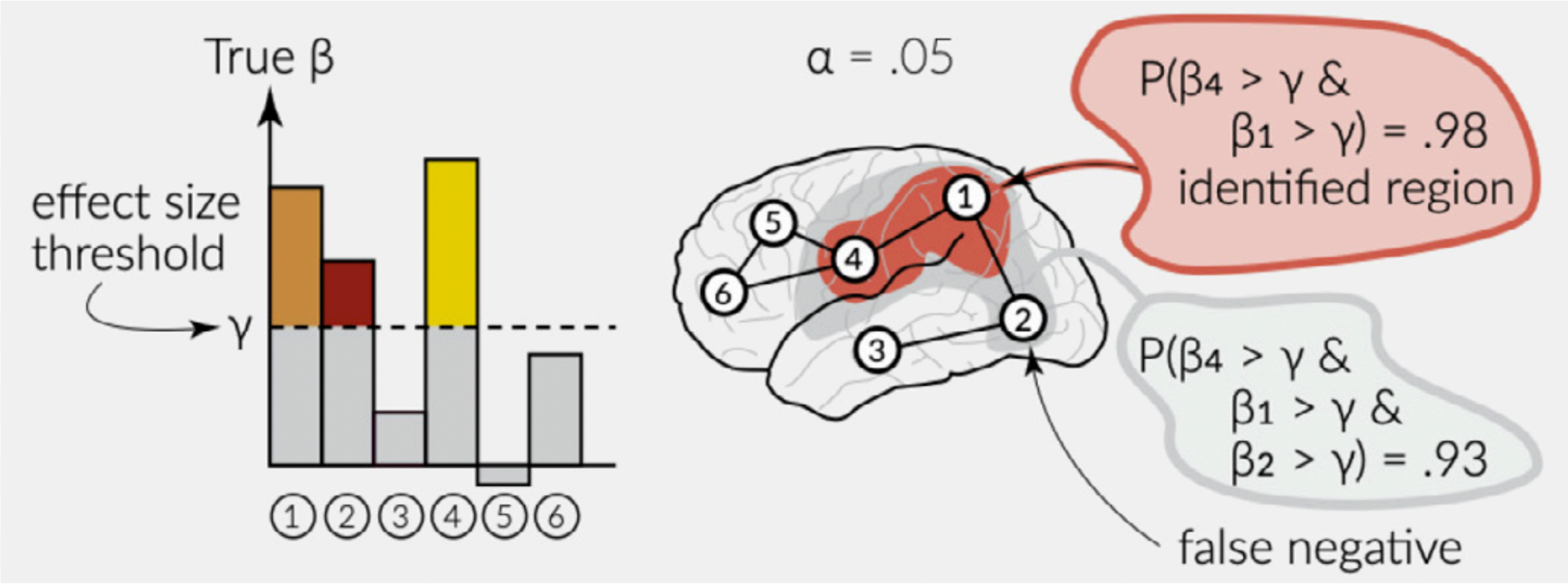
Bolin, D., & Lindgren, F. (2015). Excursion and contour uncertainty regions for latent Gaussian models. *Journal of the Royal Statistical Society Series B: Statistical Methodology.*
Mejia, A. F., Yue, Y., Bolin, D., Lindgren, F., & Lindquist, M. A. (2020). A Bayesian general linear modeling approach to cortical surface fMRI data analysis. *Journal of the American Statistical Association.*

Areas of Activation via Excursion Sets

For each task k, we use the ***joint posterior distribution of*** $\boldsymbol{\beta}_k$ to identify areas of statistically significant activation above a minimum effect size of interest

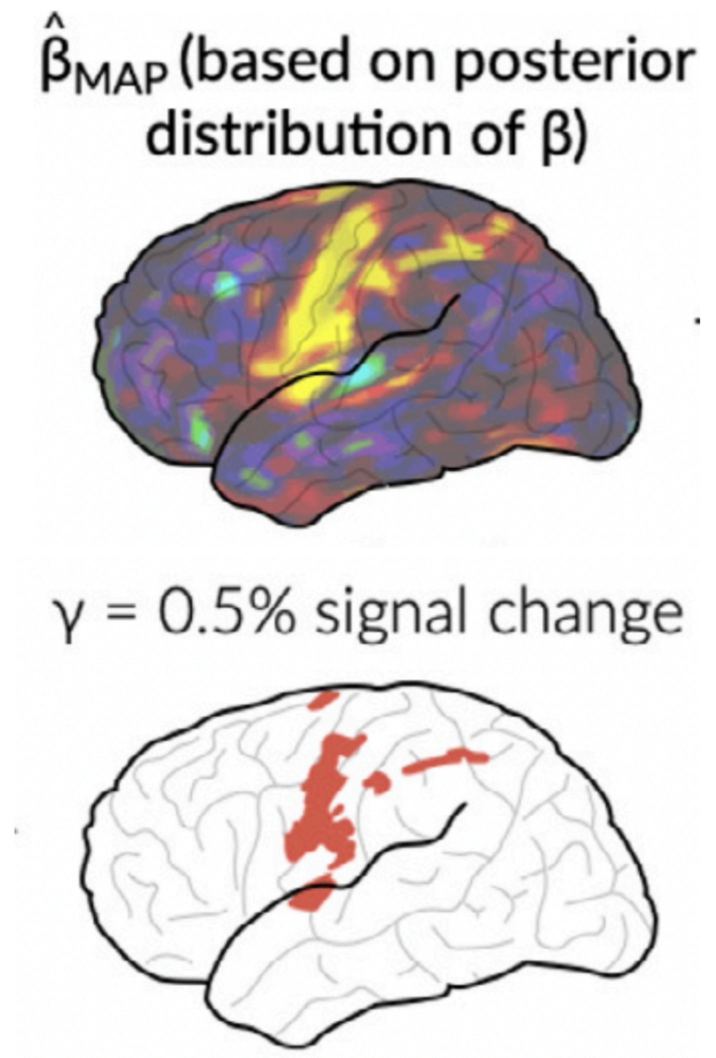
This approach ***avoids the need to correct for multiple comparisons***, which would be required if we used the marginal posterior distribution at each location to identify areas of activation

There is a ***dramatic gain in power*** over massive univariate modeling, facilitating the use of a scientifically meaningful minimum effect size



Bolin, D., & Lindgren, F. (2015). Excursion and contour uncertainty regions for latent Gaussian models. *Journal of the Royal Statistical Society Series B: Statistical Methodology.*
Mejia, A. F., Yue, Y., Bolin, D., Lindgren, F., & Lindquist, M. A. (2020). A Bayesian general linear modeling approach to cortical surface fMRI data analysis. *Journal of the American Statistical Association.*

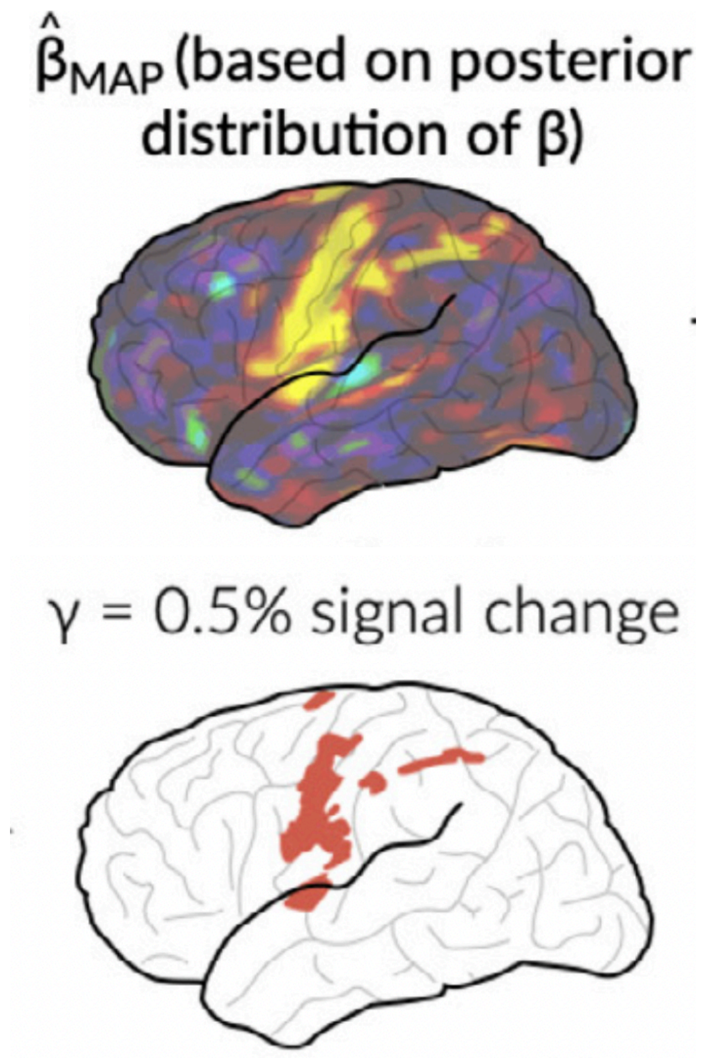
Excursions Set Procedure



Mejia, A. F., Yue, Y., Bolin, D., Lindgren, F., & Lindquist, M. A. (2020). A Bayesian general linear modeling approach to cortical surface fMRI data analysis. *Journal of the American Statistical Association.*

Excursions Set Procedure

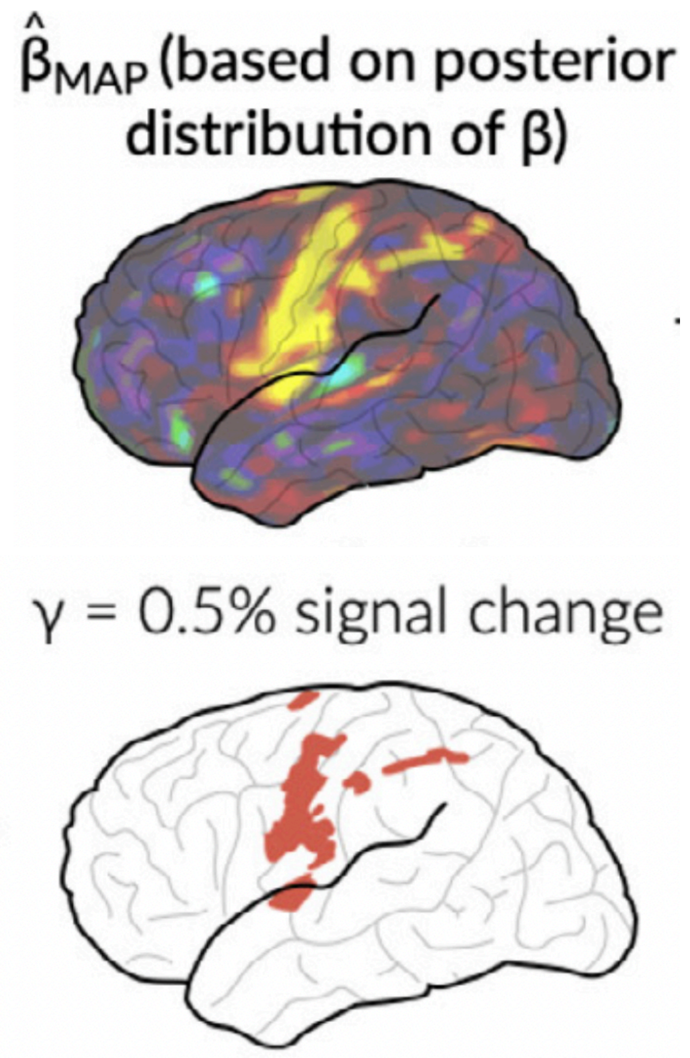
- A location v is defined as *truly activated* by task *k* if $\beta_k(v) > \gamma$



Mejia, A. F., Yue, Y., Bolin, D., Lindgren, F., & Lindquist, M. A. (2020). A Bayesian general linear modeling approach to cortical surface fMRI data analysis. *Journal of the American Statistical Association.*

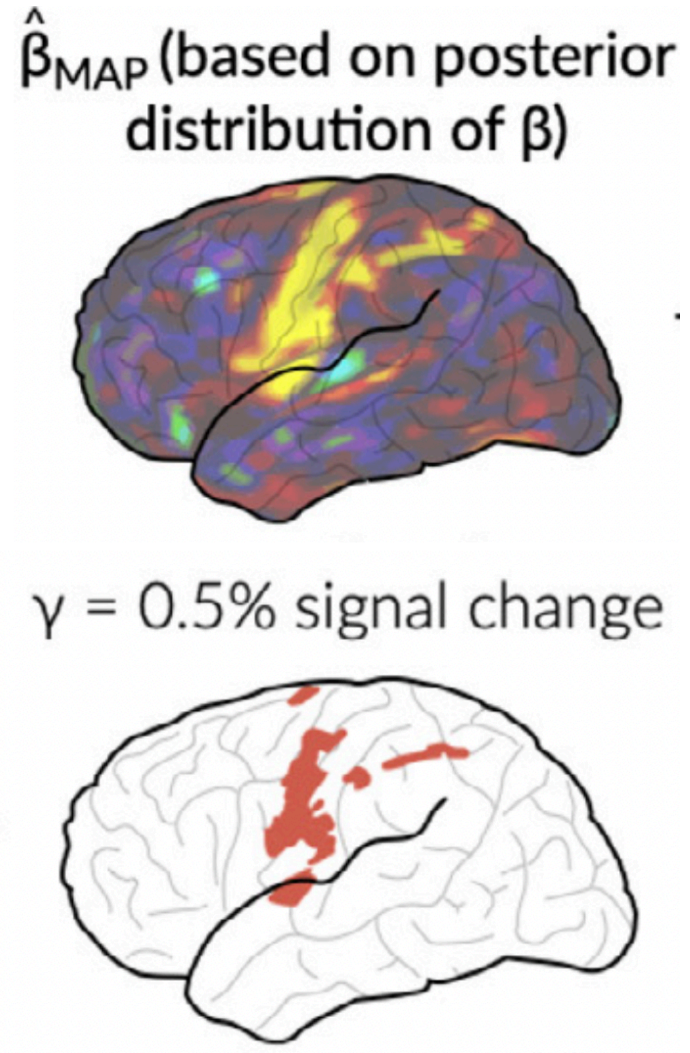
Excursions Set Procedure

- A location v is defined as *truly activated* by task *k* if $\beta_k(v) > \gamma$
- We aim to identify the *excursion region* $E_{\gamma,\alpha}$, the largest set of locations where with joint posterior probability $1 - \alpha$, every location is truly activated (above γ)



Mejia, A. F., Yue, Y., Bolin, D., Lindgren, F., & Lindquist, M. A. (2020). A Bayesian general linear modeling approach to cortical surface fMRI data analysis. *Journal of the American Statistical Association.*

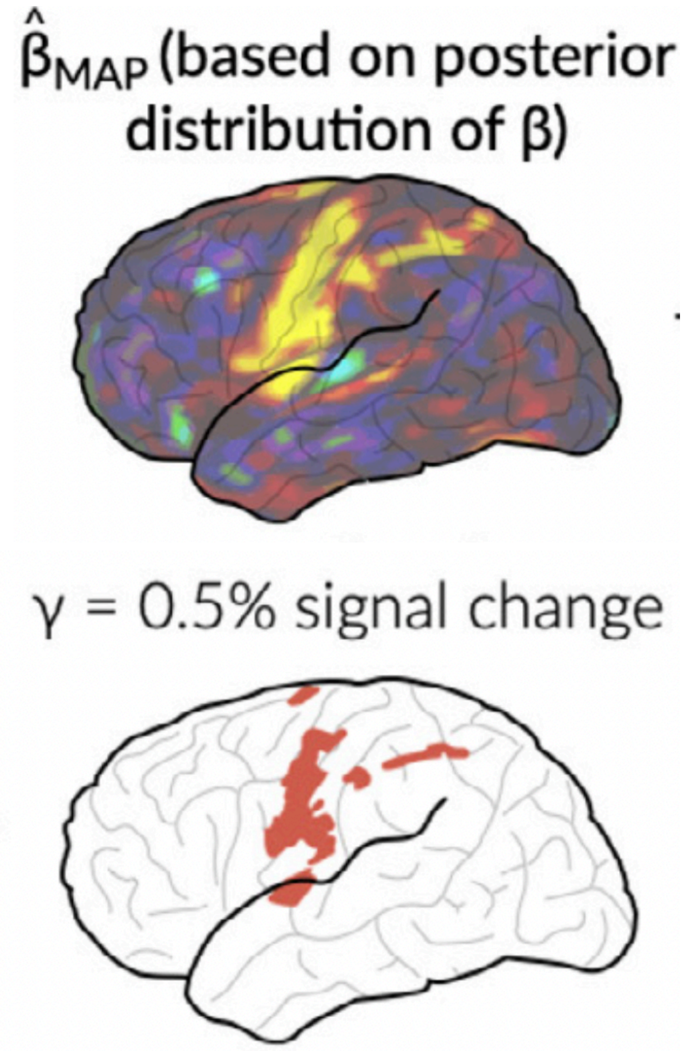
Excursions Set Procedure



- A location v is defined as *truly activated* by task *k* if $\beta_k(v) > \gamma$
- We aim to identify the *excursion region* $E_{\gamma,\alpha}$, the largest set of locations where with joint posterior probability $1 - \alpha$, every location is truly activated (above γ)
- We estimate the *excursion function*, the largest value of $1 - \alpha$ for which a location v would be included in $E_{\gamma,\alpha}$ (assuming locations are added in order of their marginal posterior probabilities)

Mejia, A. F., Yue, Y., Bolin, D., Lindgren, F., & Lindquist, M. A. (2020). A Bayesian general linear modeling approach to cortical surface fMRI data analysis. *Journal of the American Statistical Association.*

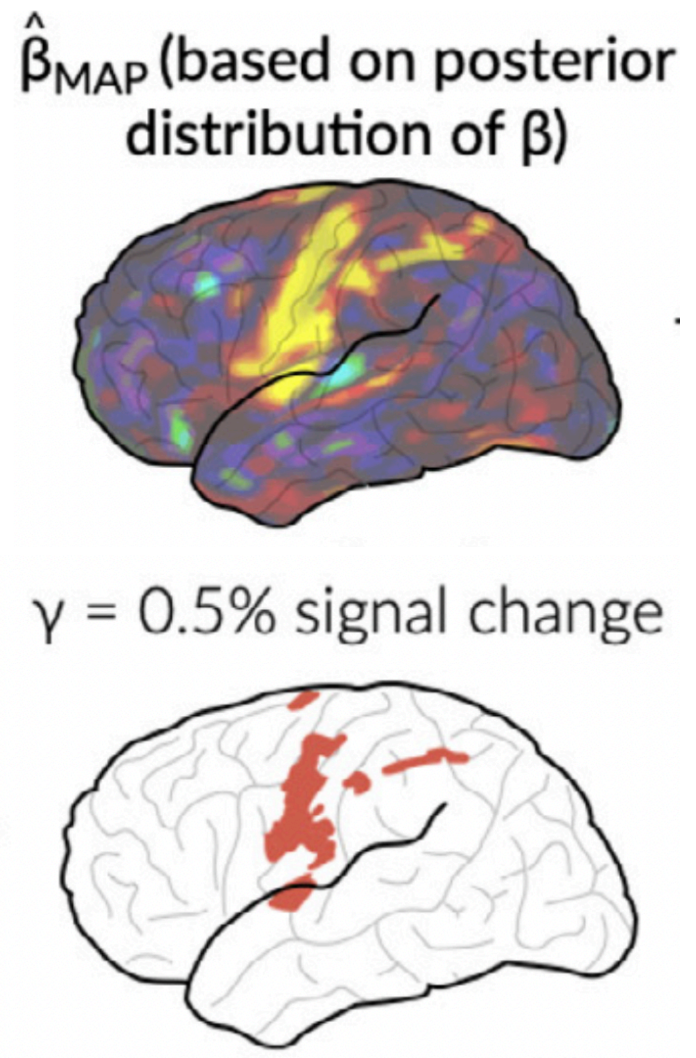
Excursions Set Procedure



- A location v is defined as *truly activated* by task *k* if $\beta_k(v) > \gamma$
- We aim to identify the *excursion region* $E_{\gamma,\alpha}$, the largest set of locations where with joint posterior probability $1-\alpha$, every location is truly activated (above γ)
- We estimate the *excursion function*, the largest value of $1-\alpha$ for which a location v would be included in $E_{\gamma,\alpha}$ (assuming locations are added in order of their marginal posterior probabilities)
- Thresholding the excursion function at, e.g., 0.95 yields the excursion region for $\alpha = 0.05$

Mejia, A. F., Yue, Y., Bolin, D., Lindgren, F., & Lindquist, M. A. (2020). A Bayesian general linear modeling approach to cortical surface fMRI data analysis. *Journal of the American Statistical Association.*

Excursions Set Procedure

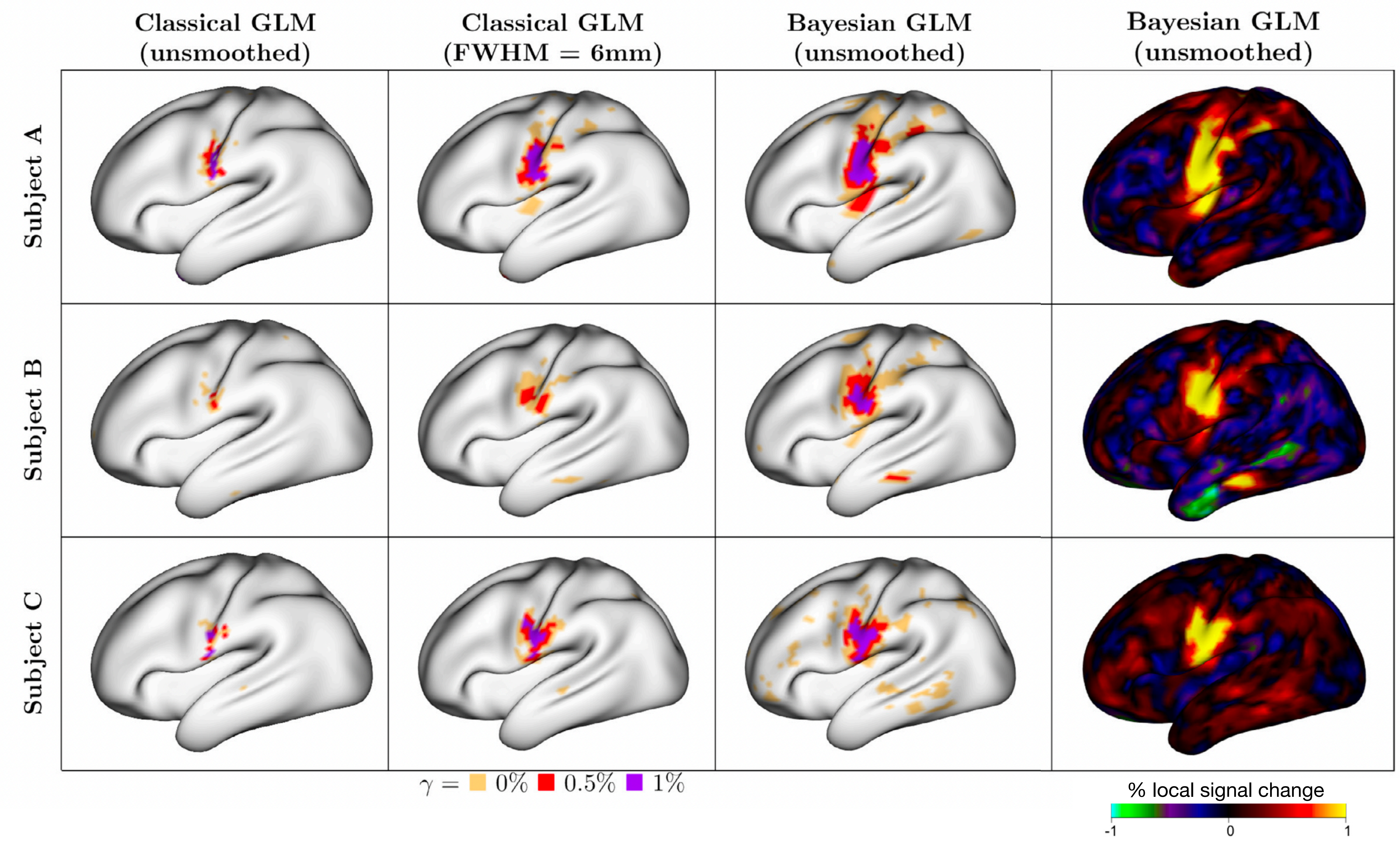


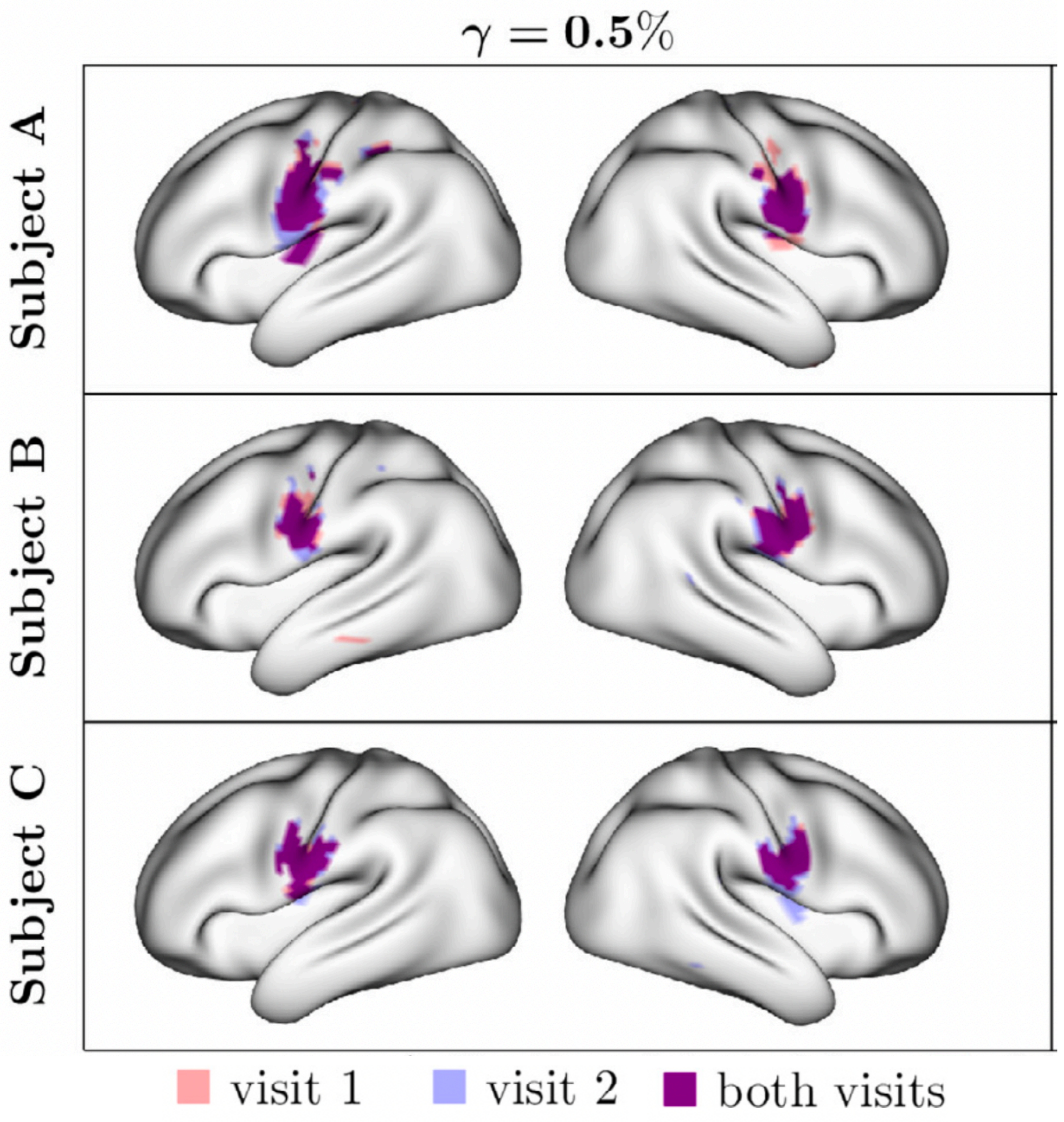
- A location v is defined as *truly activated* by task *k* if $\beta_k(v) > \gamma$
- We aim to identify the *excursion region* $E_{\gamma,\alpha}$, the largest set of locations where with joint posterior probability $1-\alpha$, every location is truly activated (above γ)
- We estimate the *excursion function*, the largest value of $1-\alpha$ for which a location v would be included in $E_{\gamma,\alpha}$ (assuming locations are added in order of their marginal posterior probabilities)
- Thresholding the excursion function at, e.g., 0.95 yields the excursion region for $\alpha = 0.05$
- **<u>This approach controls the FWER</u>**, since the probability that one or more locations in the set is *not* activated (there is at least one false positive) is α

Mejia, A. F., Yue, Y., Bolin, D., Lindgren, F., & Lindquist, M. A. (2020). A Bayesian general linear modeling approach to cortical surface fMRI data analysis. *Journal of the American Statistical Association.*

Bayesian GLM is powerful in individuals

Tongue task activations in the motor task fMRI study for 3 individuals in the HCP.
The LR + RL runs for each session were analyzed together for each subject.





Spencer, Yue, Bolin, Ryan,& Mejia (2022). Spatial Bayesian GLM on the cortical surface produces reliable task activations in individuals and groups. *NeuroImage*, 249, 118908.

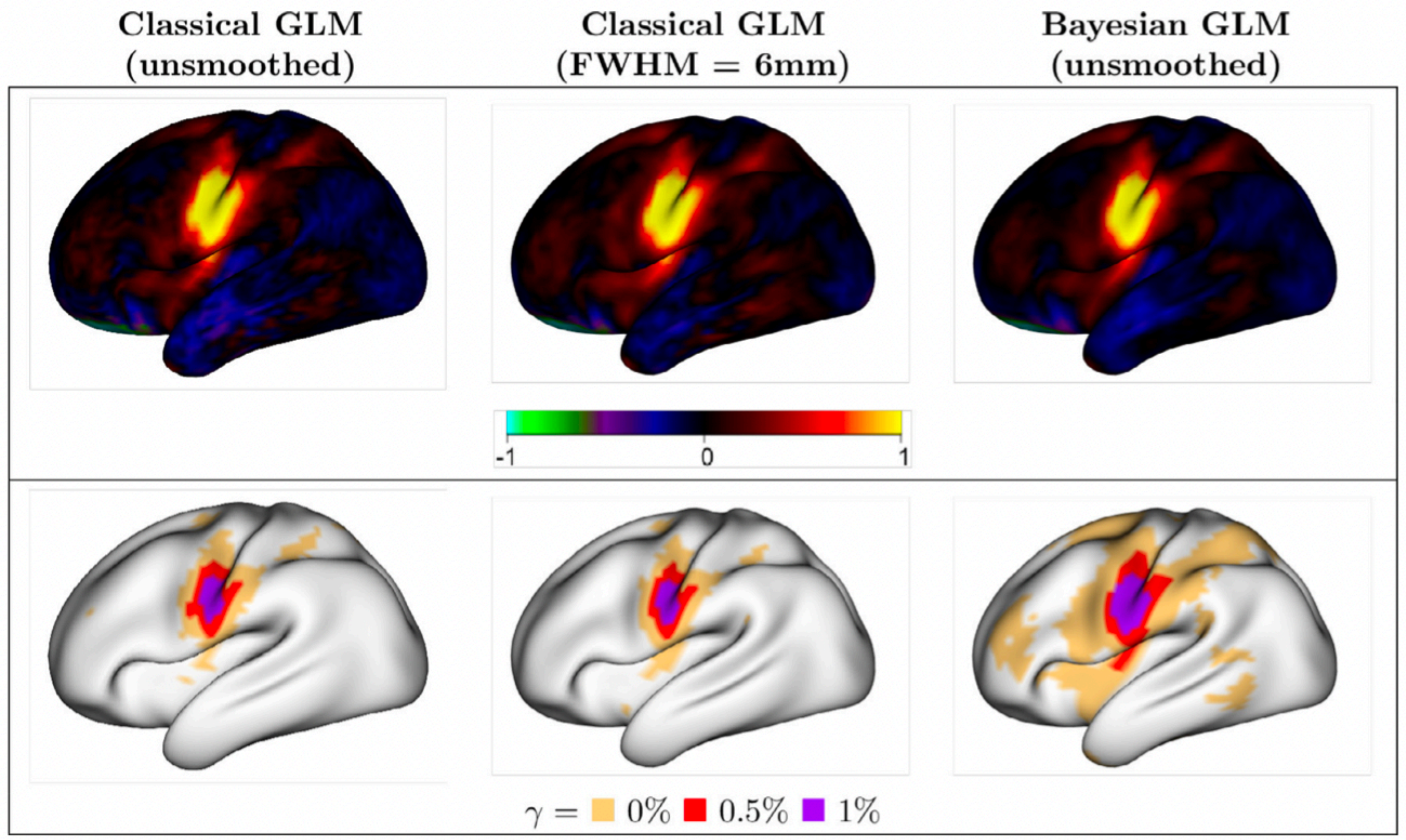
Bayesian Group Averages and Contrasts

We can analyze group averages and contrasts in a principled way after fitting all of the subject-level models, for **scalable & powerful** multi-subject analysis.

Mejia, A. F., Yue, Y., Bolin, D., Lindgren, F., & Lindquist, M. A. (2020). A Bayesian general linear modeling approach to cortical surface fMRI data analysis. *Journal of the American Statistical Association.*
Spencer, Yue, Bolin, Ryan,& Mejia (2022). Spatial Bayesian GLM on the cortical surface produces reliable task activations in individuals and groups. *NeuroImage*, 249, 118908.

Bayesian Group Averages and Contrasts

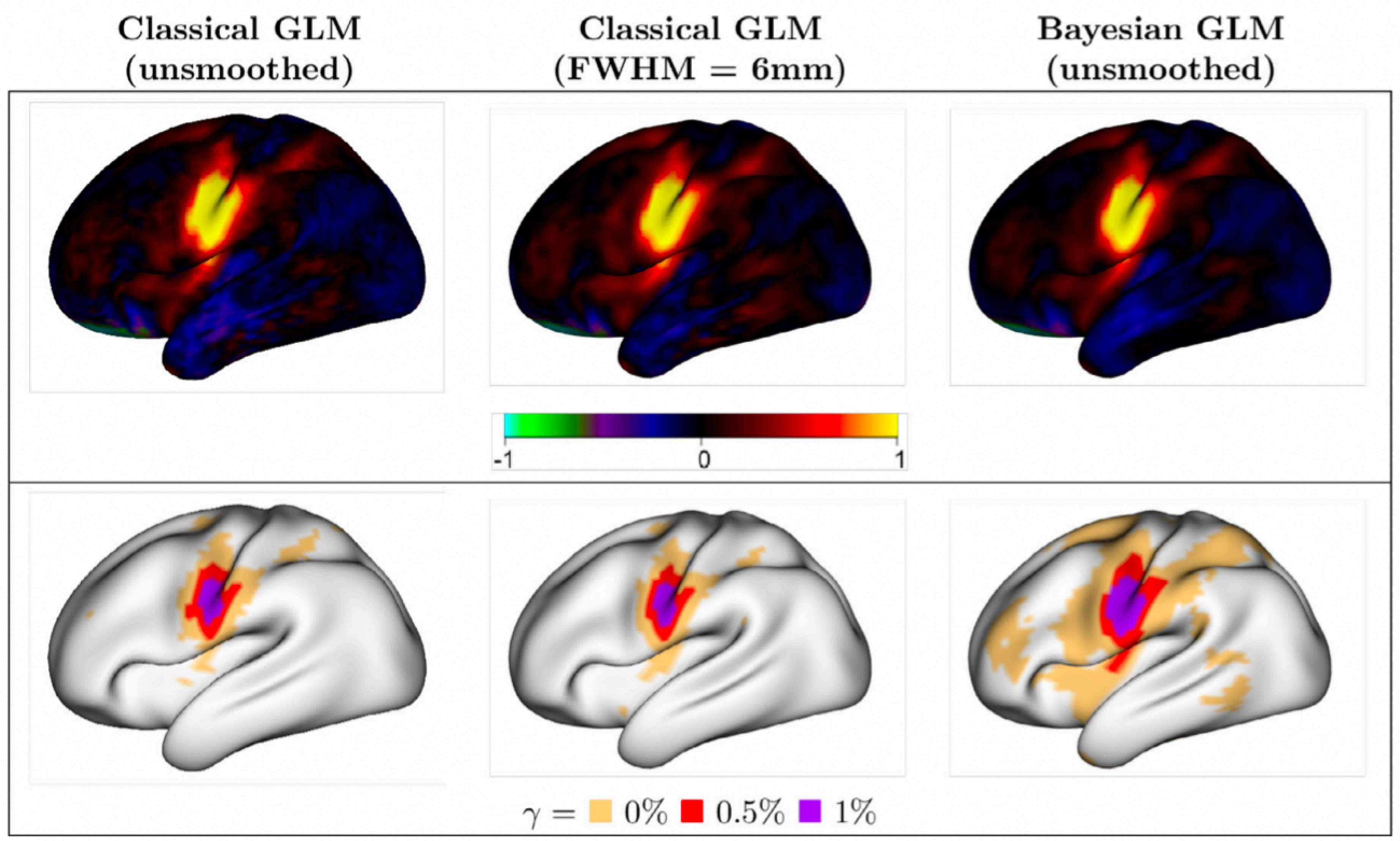
We can analyze group averages and contrasts in a principled way after fitting all of the subject-level models, for **scalable & powerful** multi-subject analysis.

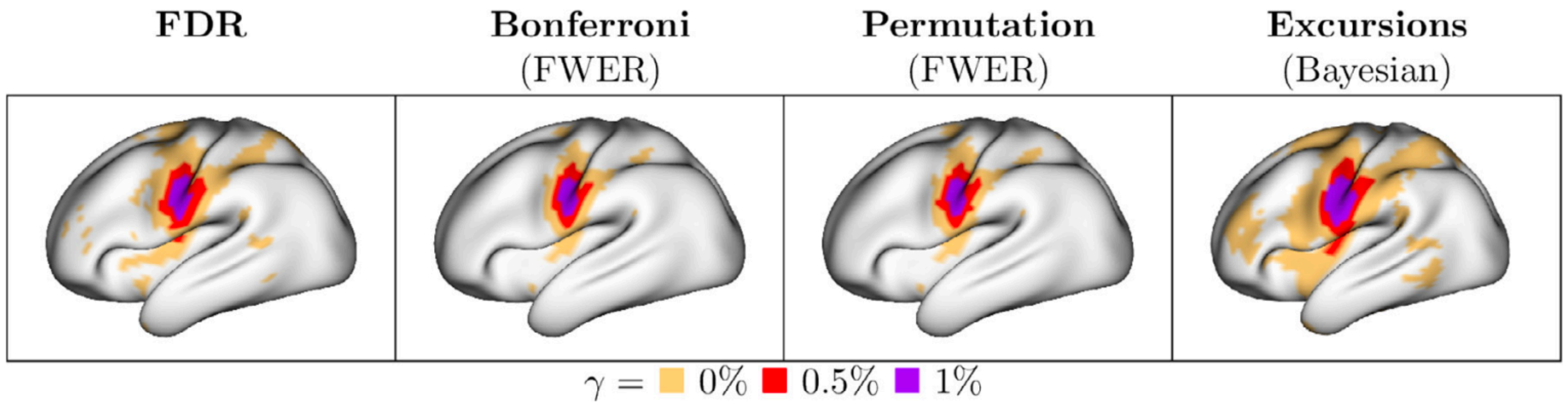


Mejia, A. F., Yue, Y., Bolin, D., Lindgren, F., & Lindquist, M. A. (2020). A Bayesian general linear modeling approach to cortical surface fMRI data analysis. *Journal of the American Statistical Association.*
Spencer, Yue, Bolin, Ryan,& Mejia (2022). Spatial Bayesian GLM on the cortical surface produces reliable task activations in individuals and groups. *NeuroImage*, 249, 118908.

Bayesian Group Averages and Contrasts

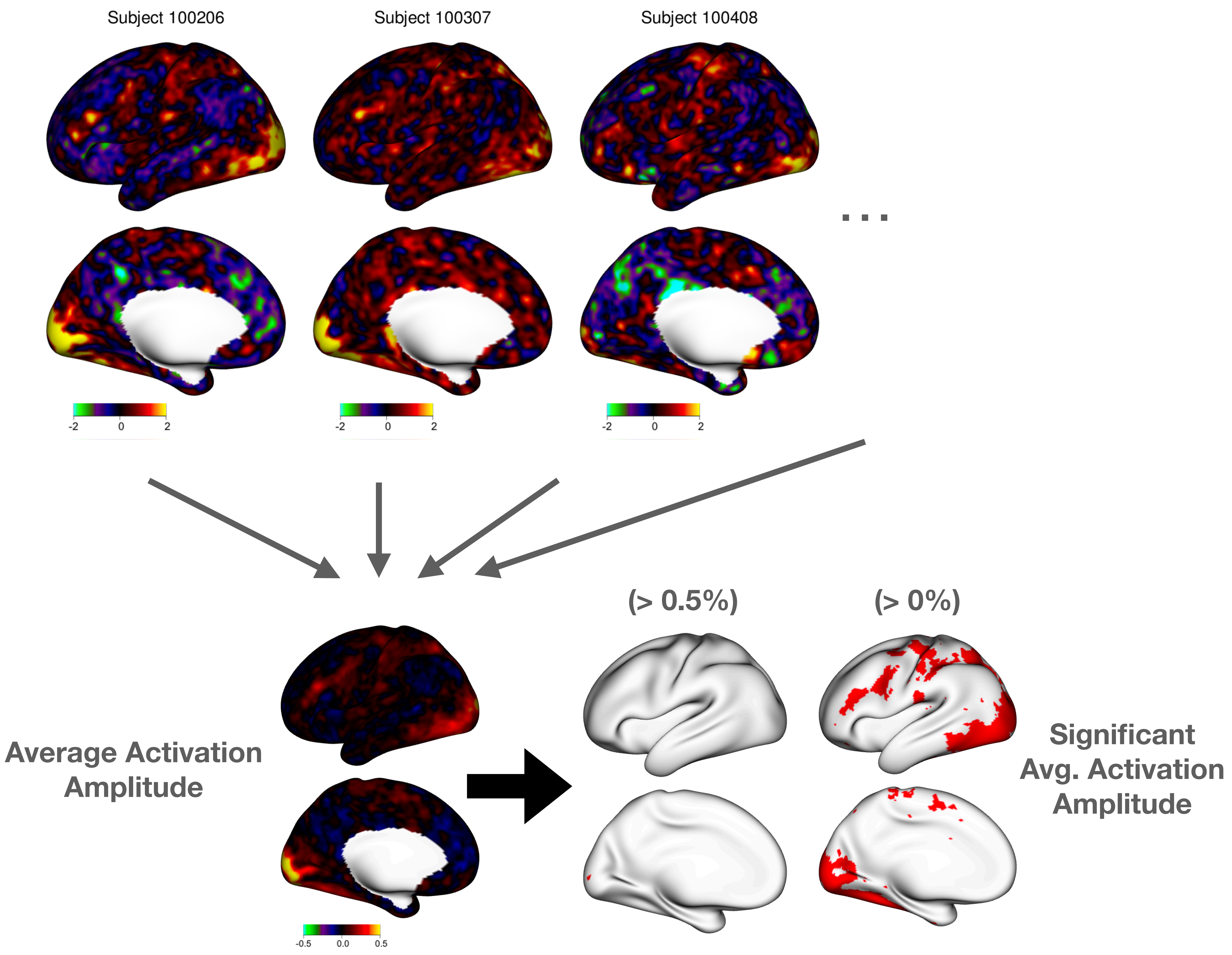
We can analyze group averages and contrasts in a principled way after fitting all of the subject-level models, for **scalable & powerful** multi-subject analysis.



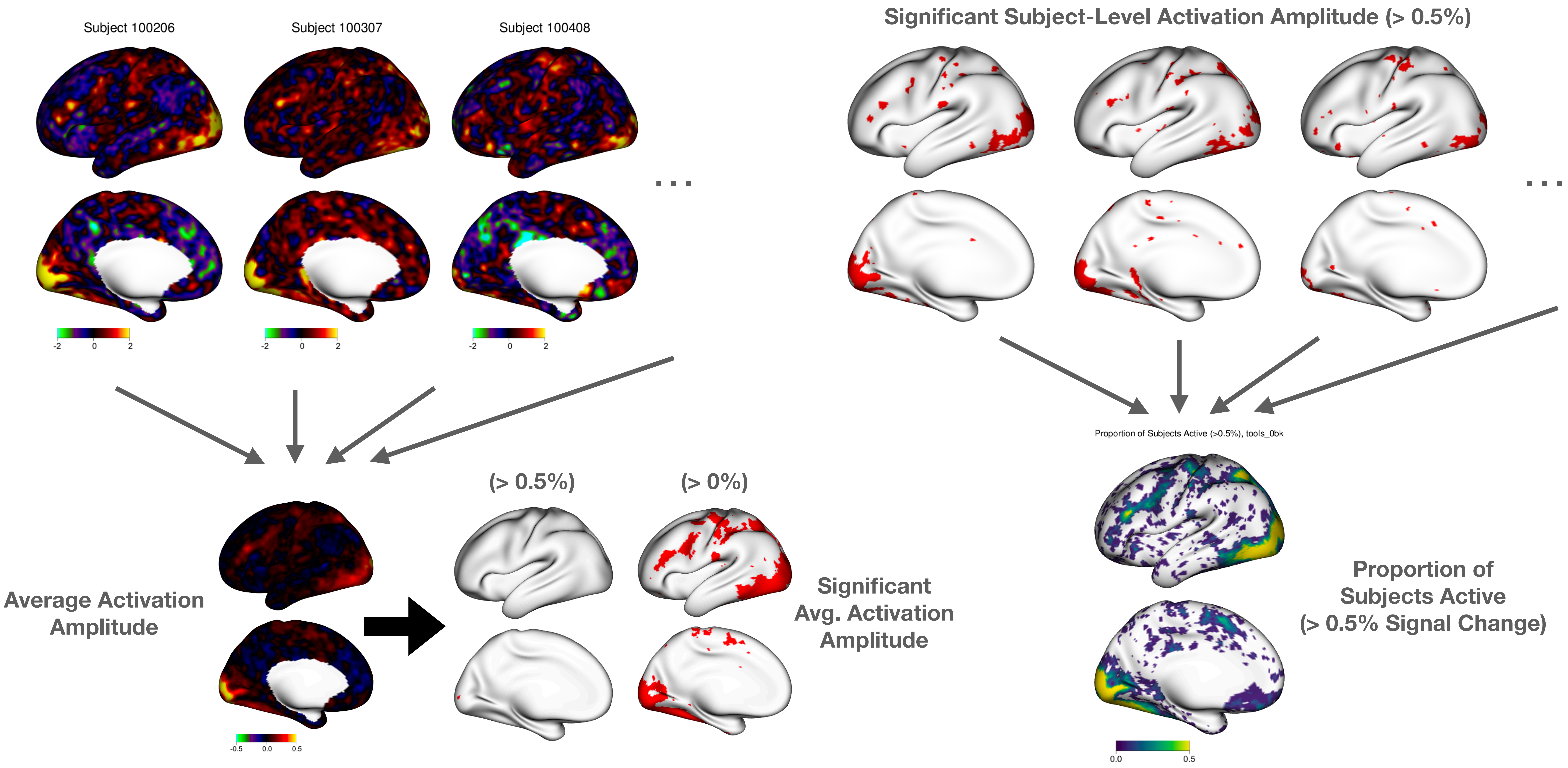


Mejia, A. F., Yue, Y., Bolin, D., Lindgren, F., & Lindquist, M. A. (2020). A Bayesian general linear modeling approach to cortical surface fMRI data analysis. *Journal of the American Statistical Association.*
Spencer, Yue, Bolin, Ryan,& Mejia (2022). Spatial Bayesian GLM on the cortical surface produces reliable task activations in individuals and groups. *NeuroImage*, 249, 118908.

From Averages to Proportions



From Averages to Proportions



Demo in R

BayesfMRI R Package