Fast Bayesian spatial 3D priors for brain imaging

Per Sidén
Division of Statistics and Machine Learning
Linköping University
per.siden@liu.se

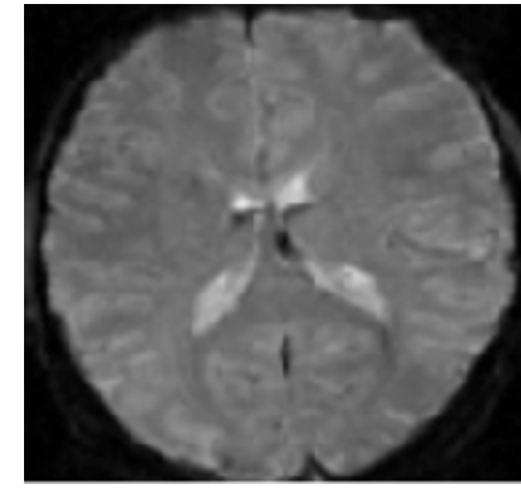
Joint work with Mattias Villani, David Bolin, Finn Lindgren and Anders Eklund

Bayes@Lund 2018

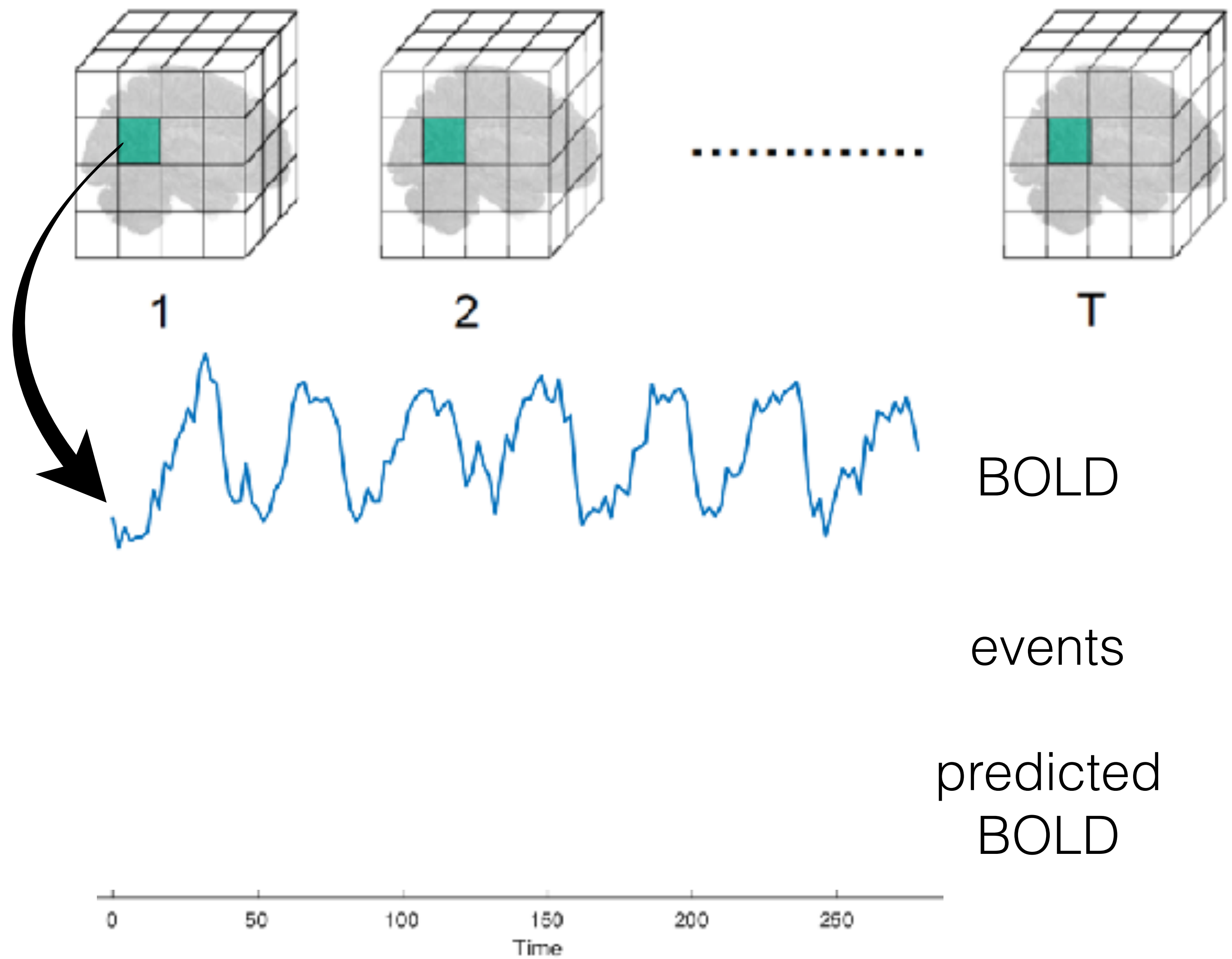
Talk overview

- fMRI basics
- Spatial 3D GMRF priors for brain activity
- Fast inference in large spatial models using sampling (MCMC) and optimisation (VB)
- Efficient posterior covariance approximations

functional magnetic resonance imaging (fMRI)



- Large number (≈100,000) of “voxels”
- BOLD signal (oxygen level) time series in each voxel
- Experimental paradigm (task-fMRI)



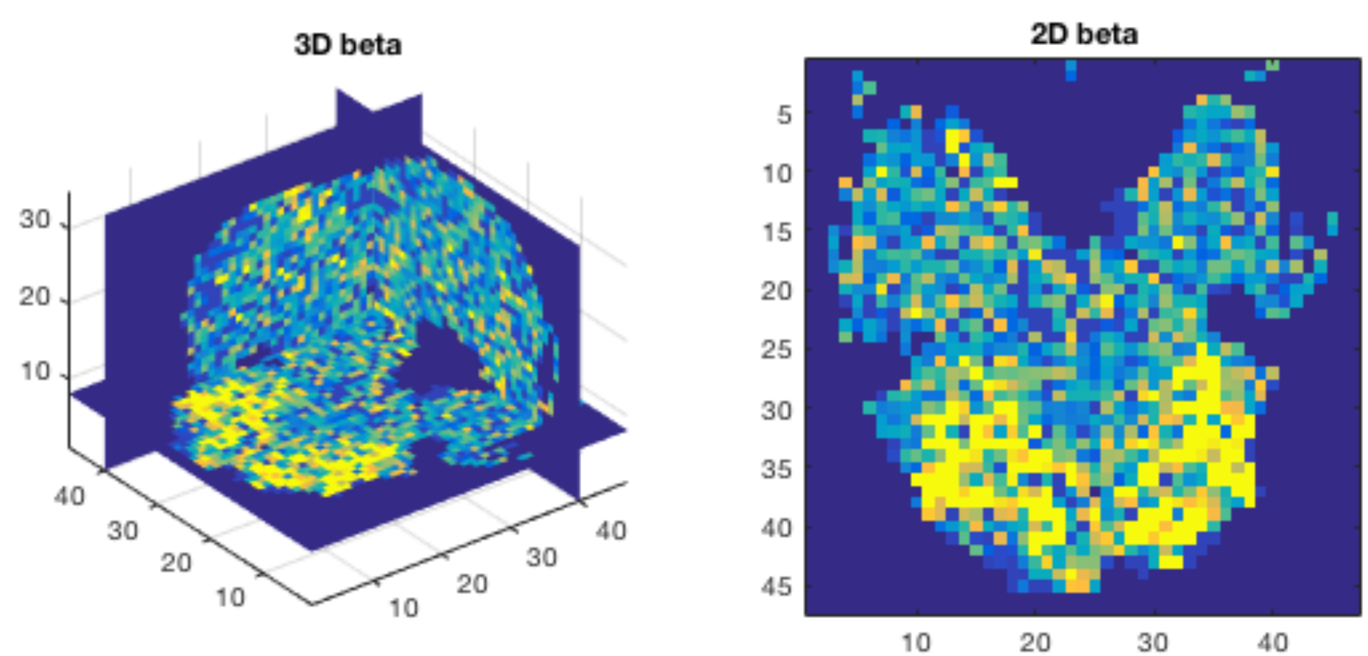
- Simple approach: fit a general linear regression model (GLM) in each voxel

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

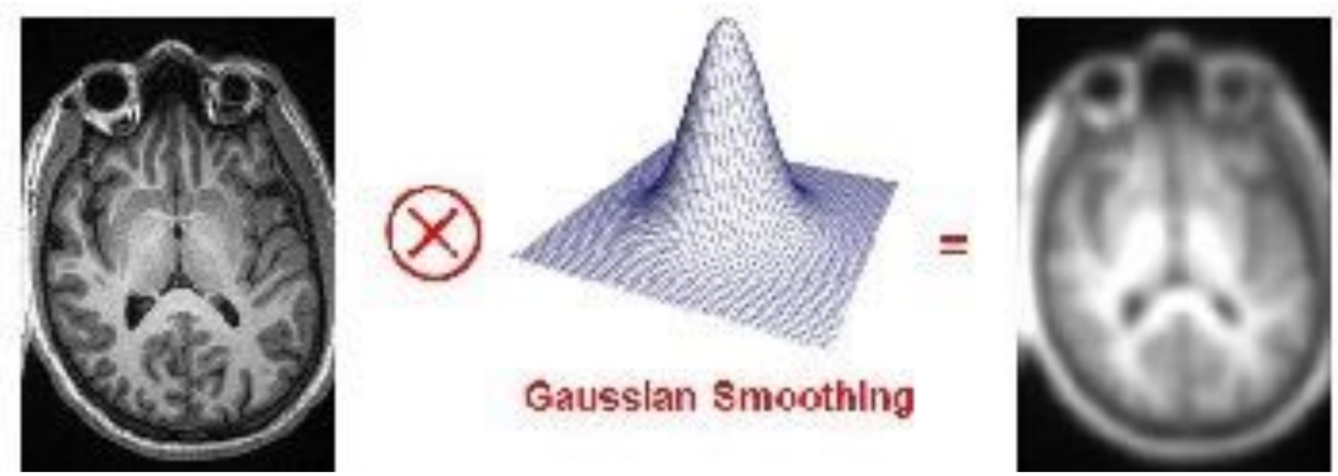
$$\varepsilon_t - \rho\varepsilon_{t-1} \sim N\left(0, 1/\lambda\right)$$

- Do t-test or posterior probability maps (PPMs)

$$PPM_n = P\left(\beta_n > \gamma | \mathbf{Y}\right)$$

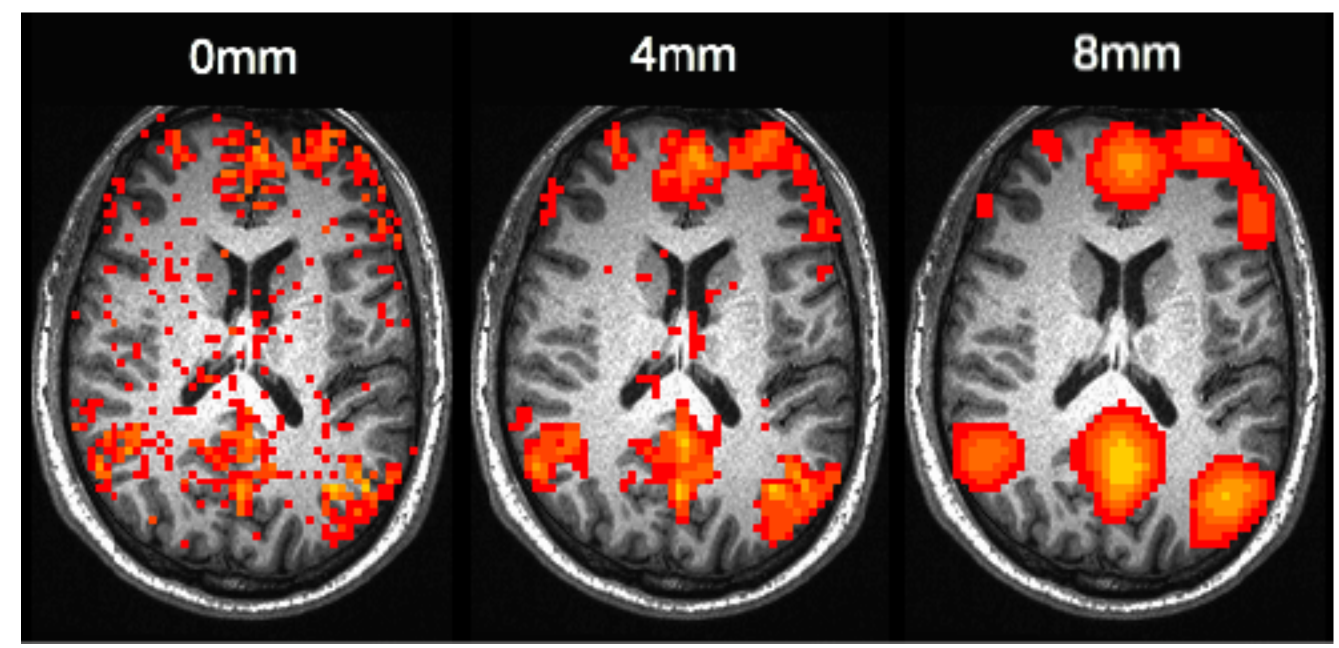


- Classic approach: Smooth the data as in pre-processing



- Results in smoother brain brain activity maps

- Not good enough!



What we want

- A Bayesian hierarchical model with **spatial priors**
- Estimate the spatial dependence from the data
- Problem: roughly 10^7 data points,
  10^6 model parameters

Spatial priors on brain activity

- Multi-voxel GLM:

$$\underset{T\times N}{\mathbf{Y}} = \underset{T\times K}{\mathbf{X}}\underset{K\times N}{\mathbf{W}} + \underset{T\times N}{\mathbf{E}},$$

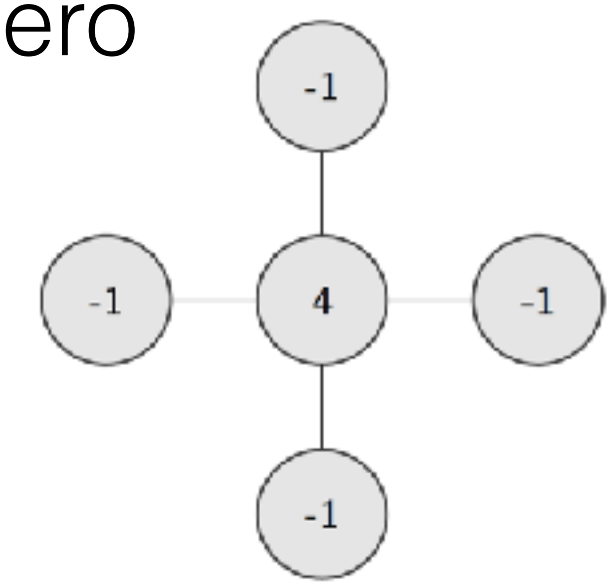
  where **E** is Gaussian order P AR noise with voxel-specific precisions $\boldsymbol{\lambda}$ and AR parameters **A**.

- Spatial Gaussian Markov random field (GMRF) priors on the regression coefficients **W**

Penny et al. "Bayesian fMRI time series analysis with spatial priors", NeuroImage (2005).

GMRFs

- Random walk (ICAR) prior: difference between neighbouring voxels i and j has mean zero

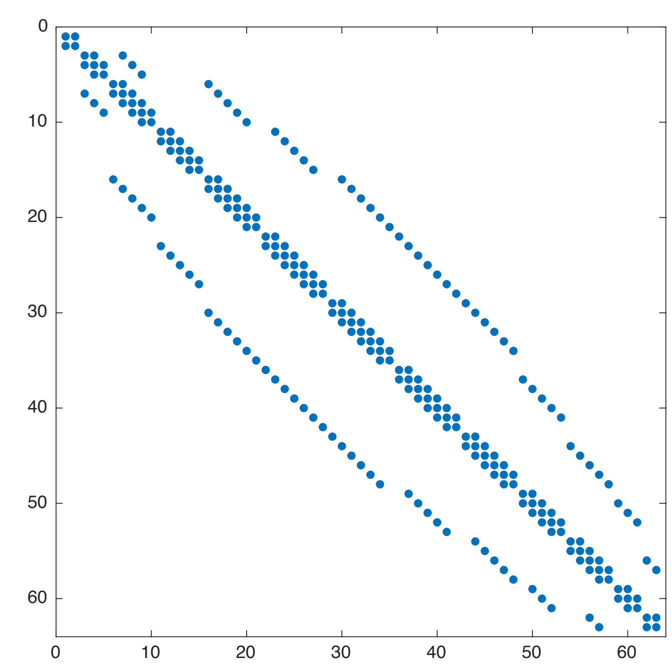


$$x_i - x_j | \alpha \overset{\text{"}iid\text{"}}{\sim} \mathcal{N}\left(0, \alpha^{-1}\right)$$

$$\Leftrightarrow \mathbf{x}|\alpha \sim \mathcal{N}\left(\mathbf{0}, (\alpha \mathbf{D})^{-1}\right),$$

$$\mathbf{D}_{i,j} = \begin{cases} n_i & \text{if } i = j \\ -1 & \text{if } i \sim j \\ 0 & \text{otherwise} \end{cases}$$

$\mathbf{D}$

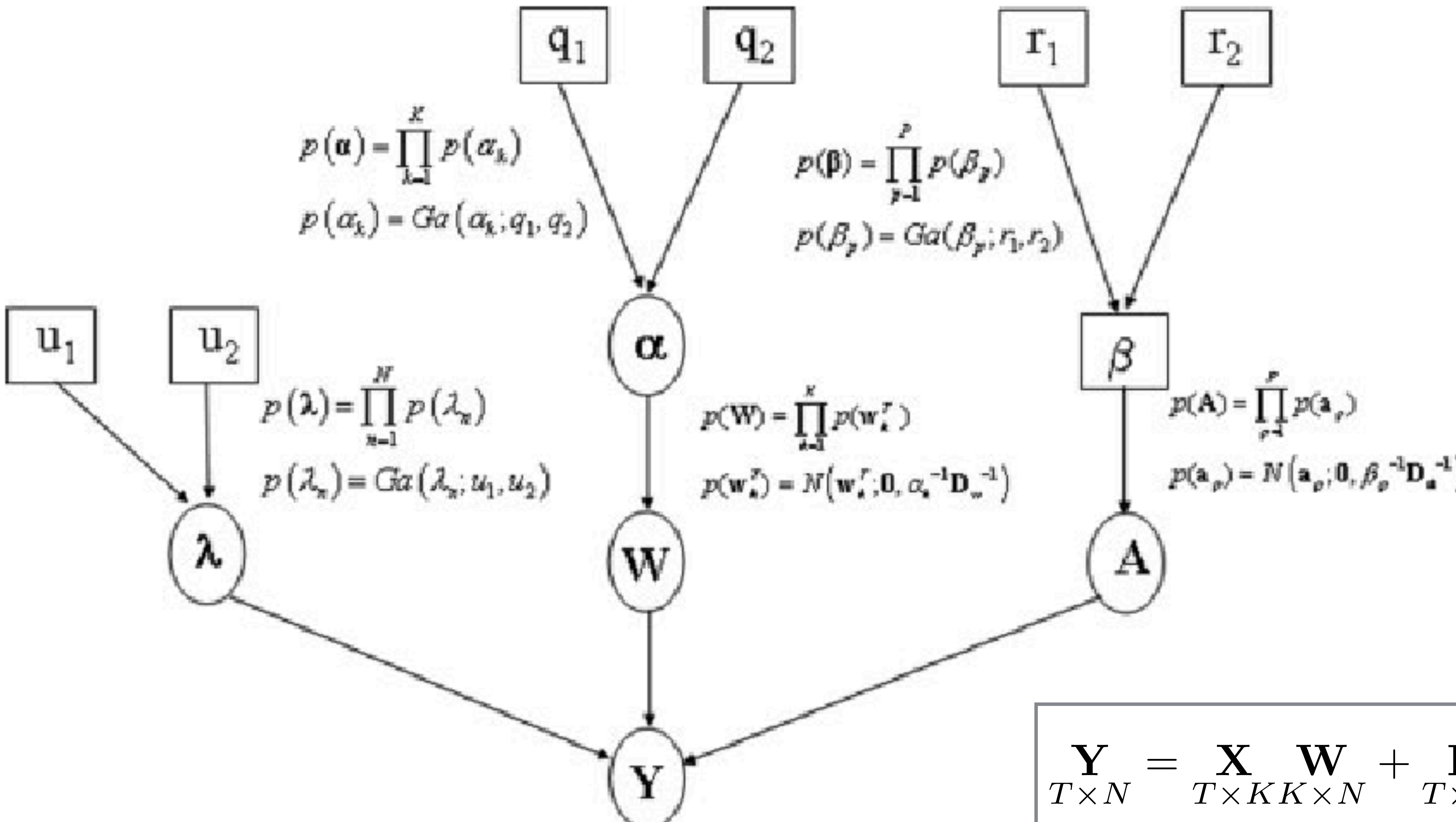


- This gives **sparsity** in the precision matrix!

Rue and Held “Gaussian Markov Random Fields: Theory and Applications”, CRC Press (2005).

Spatial model for brain activity

- Use similar spatial prior for AR coefficients **A**. Conjugate gamma priors for λ_n and α_k.



$$\underset{T\times N}{\mathbf{Y}} = \underset{T\times K}{\mathbf{X}}\,\underset{K\times N}{\mathbf{W}} + \underset{T\times N}{\mathbf{E}}$$

Penny et al. “Bayesian comparison of spatially regularised general linear models”, Human Brain Mapping (2007).

Spatial model for brain activity

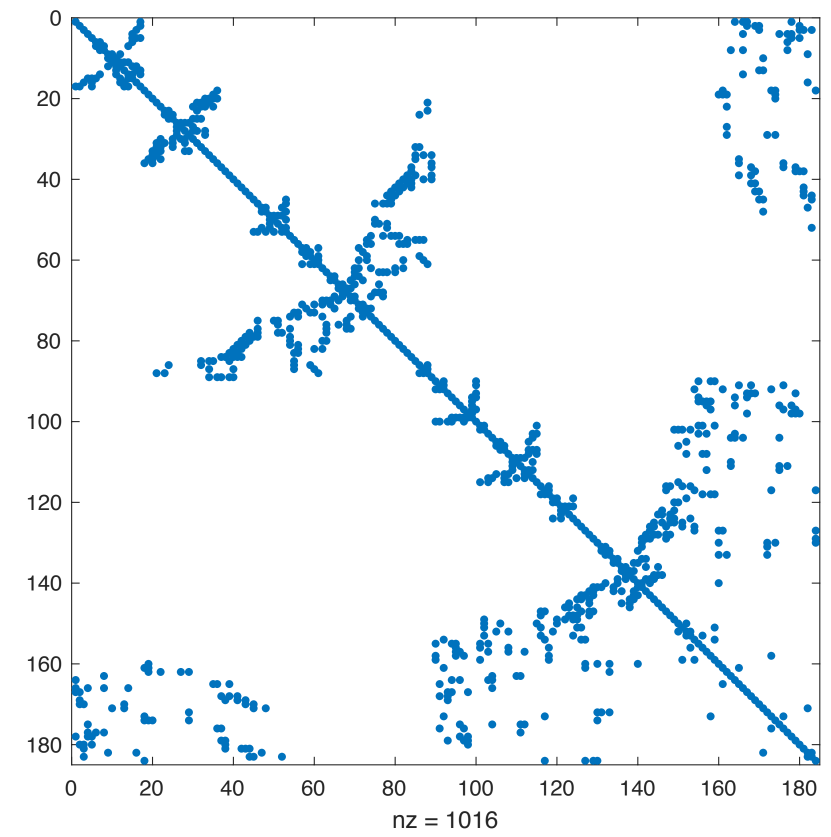
Our contributions:

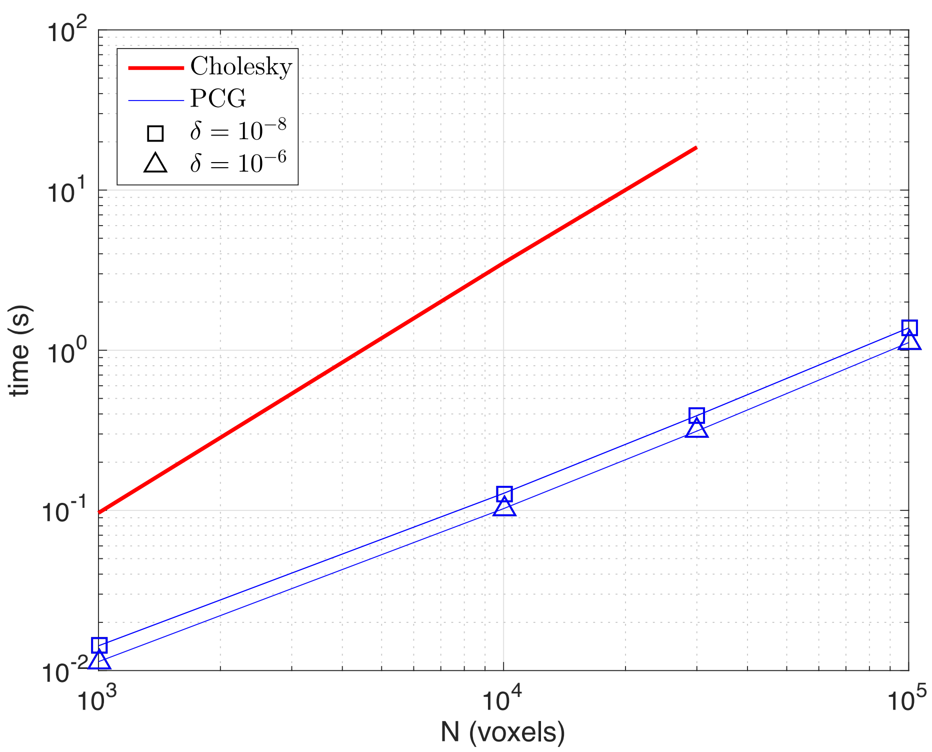
- Estimation with MCMC and Spatial VB
- Matrix reorderings for faster sparse matrix algebra
- Preconditioned conjugate gradient (PCG) methods

Replace $\text{solve } \mathbf{Qx} = \mathbf{b}$

with $\min_{\mathbf{x}} ||\mathbf{Qx} - \mathbf{b}||_2$

- Efficient posterior covariance approximations





Bayesian inference

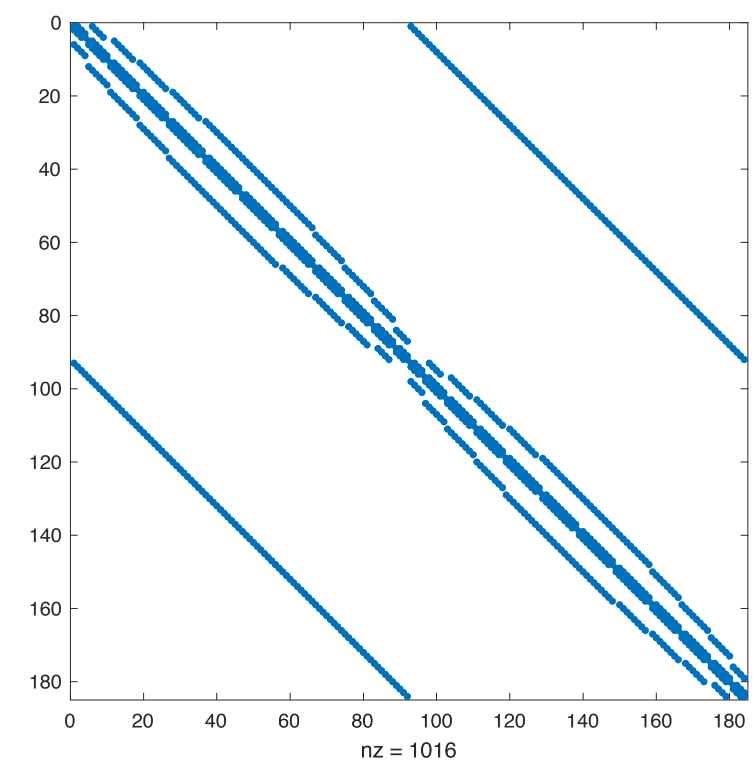
- Conjugate priors —> conditional posteriors also GMRF/Gamma.

$$\mathbf{w}|\mathbf{Y},\boldsymbol{\alpha},\boldsymbol{\lambda} \sim \mathcal{N}\left(\tilde{\boldsymbol{\mu}},\tilde{\mathbf{Q}}^{-1}\right)$$

- Allows for Gibbs sampling (MCMC)
- Alternatively, use spatial variational Bayes (SVB)

$$p\left(\mathbf{W},\boldsymbol{\alpha},\boldsymbol{\lambda}|\mathbf{Y}\right) \approx q\left(\mathbf{W}\right)q\left(\boldsymbol{\alpha}\right)q\left(\boldsymbol{\lambda}\right).$$

$\tilde{\mathbf{Q}}$

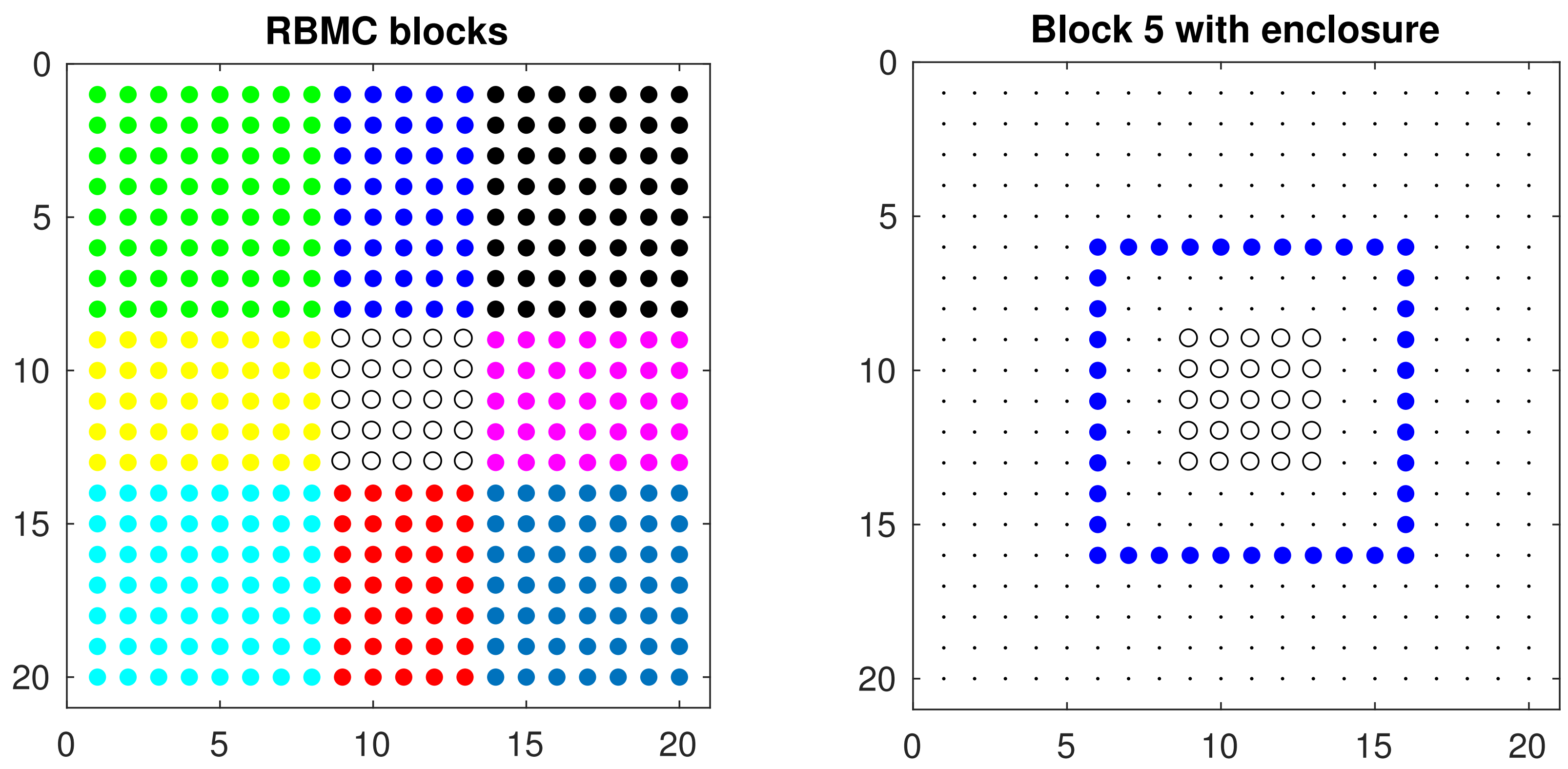


Efficient covariance computations using Rao-Blackwellization

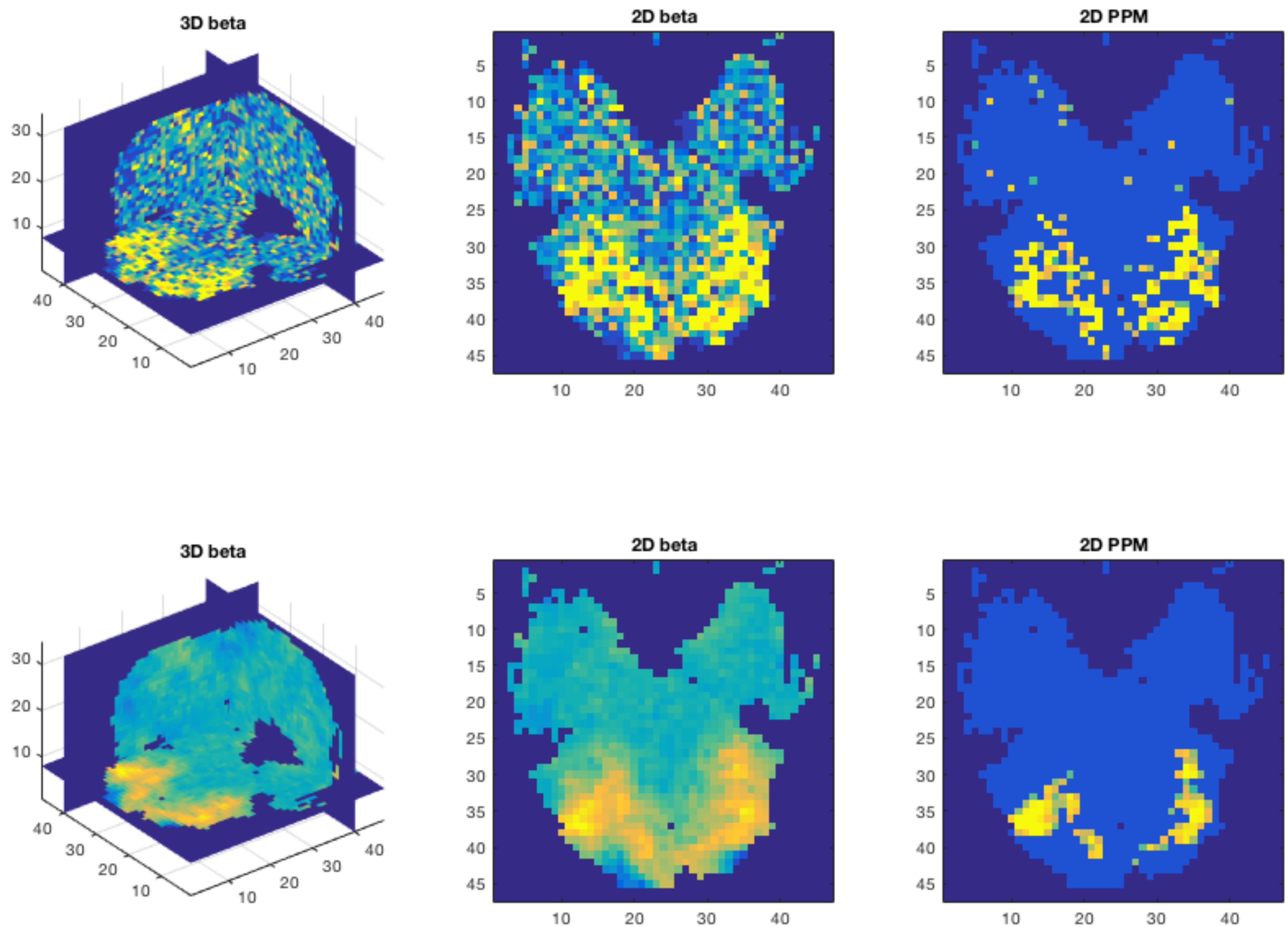
- Problem: Compute diagonal of $\mathbf{\Sigma} = \tilde{\mathbf{Q}}^{-1}$

$$Var\left(x_i\right) = E\left[Var\left(x_i|\mathbf{x}_{-i}\right)\right] + Var\left[E\left(x_i|\mathbf{x}_{-i}\right)\right]$$

$$\approx Q_{i,i}^{-1} + \frac{1}{N_s}\sum_{j=1}^{N_s}\left(Q_{i,i}^{-1}\mathbf{Q}_{i,-i}\mathbf{x}_{-i}^{(j)}\right)^2$$

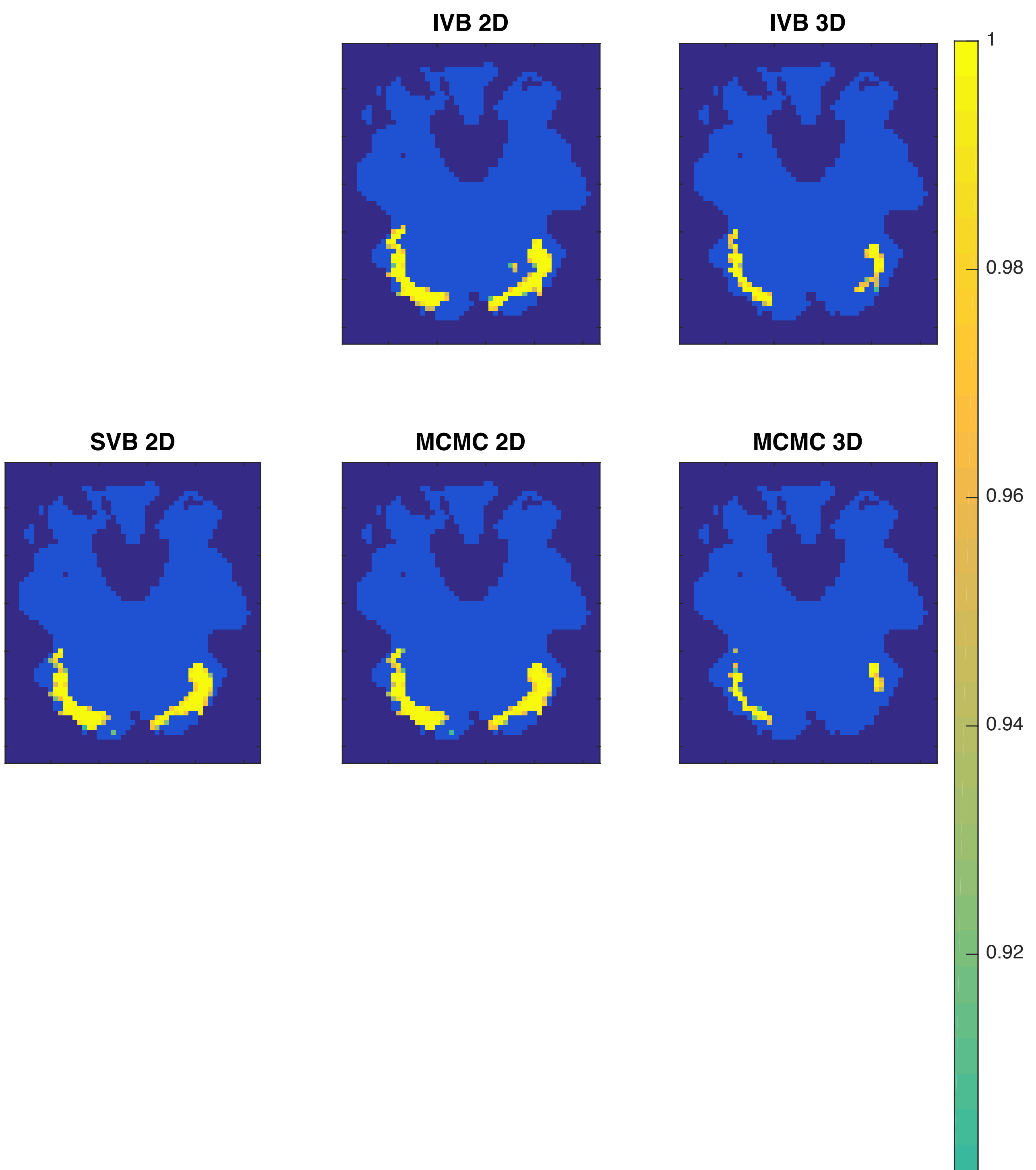


Results

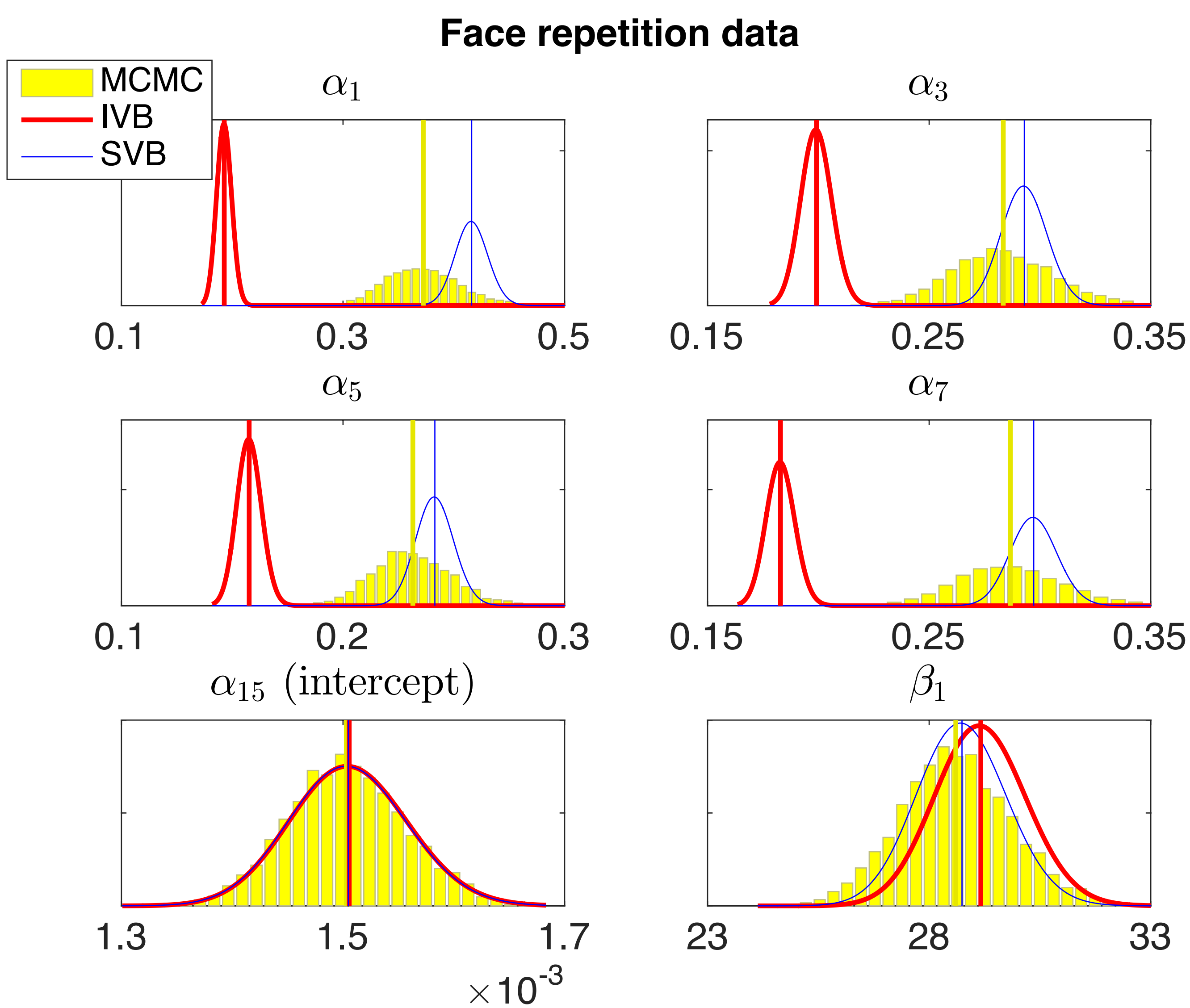


Posterior Probability Maps (PPMs)

$$PPM_n = P\left(\beta_n > \gamma|\mathbf{Y}\right)$$



Posterior of hyperparameters



Future work

- Better spatial priors (e.g. Matérn, non-stationary)
- Model selection/validation
- Group analysis

Thank you!

Mattias Villani

- Papers:
  Fast Bayesian whole-brain fMRI analysis with spatial 3D priors, NeuroImage (2017).

  Efficient Covariance Approximations for Large Sparse Precision Matrices, arXiv:1705.08656 (2017), to appear in Journal of Computational and Graphical Statistics.

David Bolin

- Code:
  http://www.fil.ion.ucl.ac.uk/spm/ext/#BFAST3D
  https://github.com/psiden/CovApprox

Finn Lindgren

- Funding:
  Vetenskapsrådet (2013-5229).

Anders Eklund

LINKÖPING UNIVERSITY