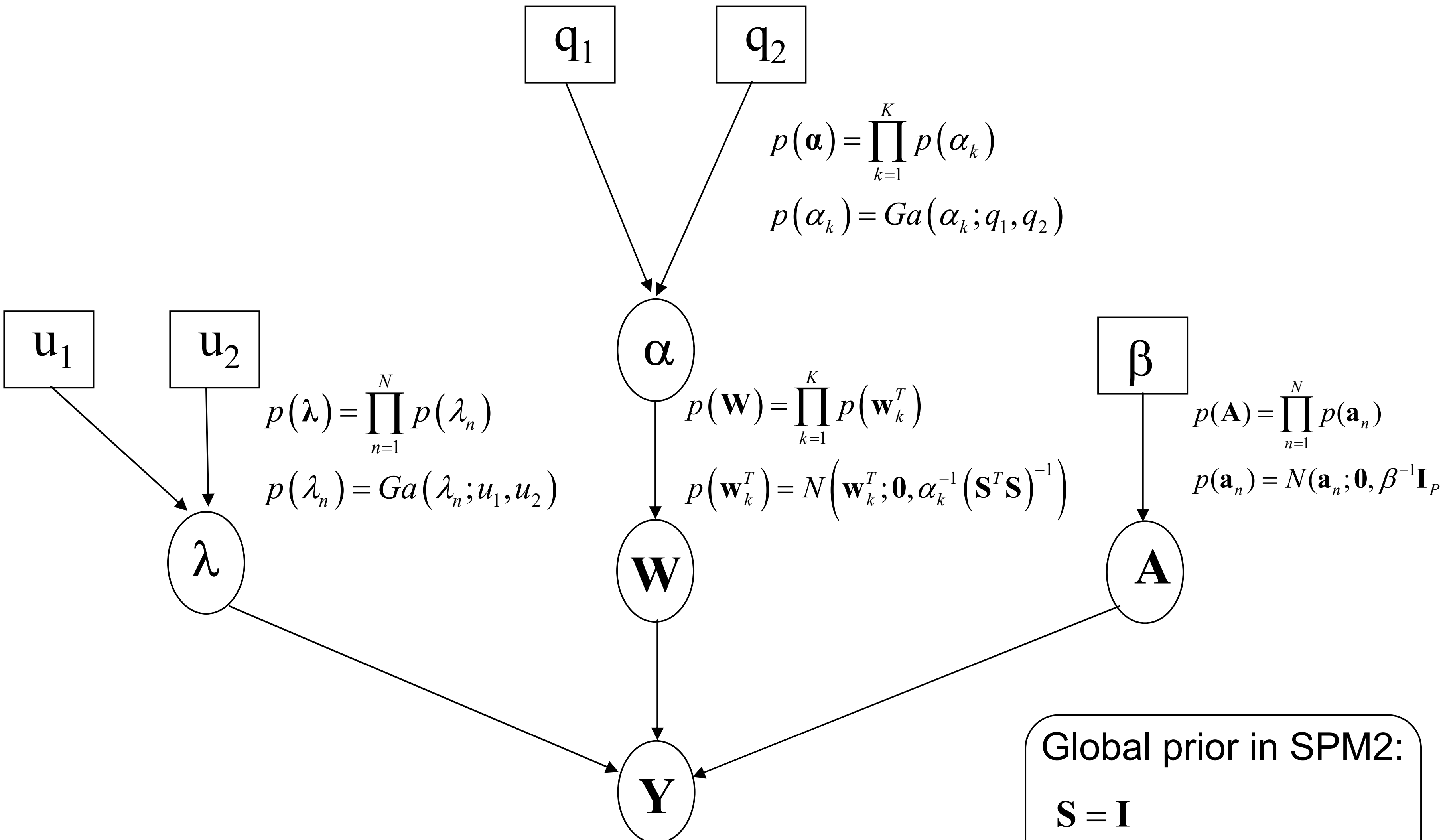
PPMs with Spatial Priors

Will,Nelson,Karl

GLM-AR(p) models (+Stefan) with spatial priors on regression/AR coefficients for single-subject fMRI



Y=XW+E

[TxN] [TxK] [KxN] [TxN]

Global prior in SPM2:

$\mathbf{S} = \mathbf{I}$

Laplacian prior:

$\mathbf{S} = \mathbf{L}$

Variational Bayes

In prior, **W** factorises over k:

$$p(\mathbf{W},\mathbf{A},\boldsymbol{\lambda},\boldsymbol{\alpha}) = \left(\prod_k p(\mathbf{w}_k \mid \alpha_k)\, p(\alpha_k \mid q_1, q_2)\right)\left(\prod_n p(\mathbf{a}_n \mid \beta)\, p(\lambda_n \mid u_1, u_2)\right)$$

In *chosen* approximate posterior, **W** factorises over n:

$$q(\mathbf{W},\mathbf{A},\boldsymbol{\lambda},\boldsymbol{\alpha} \mid \mathbf{Y}) = \left(\prod_k q(\alpha_k \mid \mathbf{Y})\right)\left(\prod_n q(\mathbf{w}_n \mid \mathbf{Y})\, q(\mathbf{a}_n \mid \mathbf{Y})\, q(\lambda_n \mid \mathbf{Y})\right)$$

Update SS of approximate posterior to maximise (lower bound on) evidence

$$L = \log p(Y)$$
$$= F + KL[q(\theta \mid Y), p(\theta \mid Y)]$$

Updating approximate posterior

Regression coefficients

$$q(\mathbf{w}_n) = N\left(\mathbf{w}_n; \hat{\mathbf{w}}_n, \hat{\boldsymbol{\Sigma}}_n\right)$$

$$\widehat{\mathbf{w}}_n = \widehat{\boldsymbol{\Sigma}}_n \left(\overline{\lambda_n} \tilde{\mathbf{b}}_n^{\ T} + \mathbf{r}_n\right)$$

$$\widehat{\boldsymbol{\Sigma}}_n = \left(\overline{\lambda_n} \tilde{\mathbf{A}}_n + \mathbf{B}_{nn}\right)^{-1}$$

$$\mathbf{B} = \mathbf{H}\left(diag\left[\boldsymbol{\alpha}\right] \otimes \mathbf{S}^T \mathbf{S}\right) \mathbf{H}^T$$

$$\mathbf{r}_n = -\sum_{i=1, i \neq n}^{N} \mathbf{B}_{ni} \hat{\mathbf{w}}_i$$

AR coefficients

$$q(\mathbf{a}_n) = N(\mathbf{a}_n; \mathbf{m}_n, \mathbf{V}_n)$$

$$\mathbf{V}_n = (\lambda_n \widetilde{\mathbf{C}_n} + \beta \ \mathbf{I}_P)^{-1}$$

$$\mathbf{m}_n = \lambda_n \widetilde{\mathbf{D}_n} \mathbf{V}_n$$

Observation noise

$$q(\lambda_n) = Ga(\lambda_n; b_n, c_n)$$

$$\frac{1}{b_n} = \frac{\widetilde{G_n}}{2} + \frac{1}{u_1}$$

$$c_n = \frac{T}{2} + u_2$$

Spatial precisions

$$q\left(\boldsymbol{\alpha}\right) = \prod_{k=1}^{K} q\left(\alpha_k\right)$$

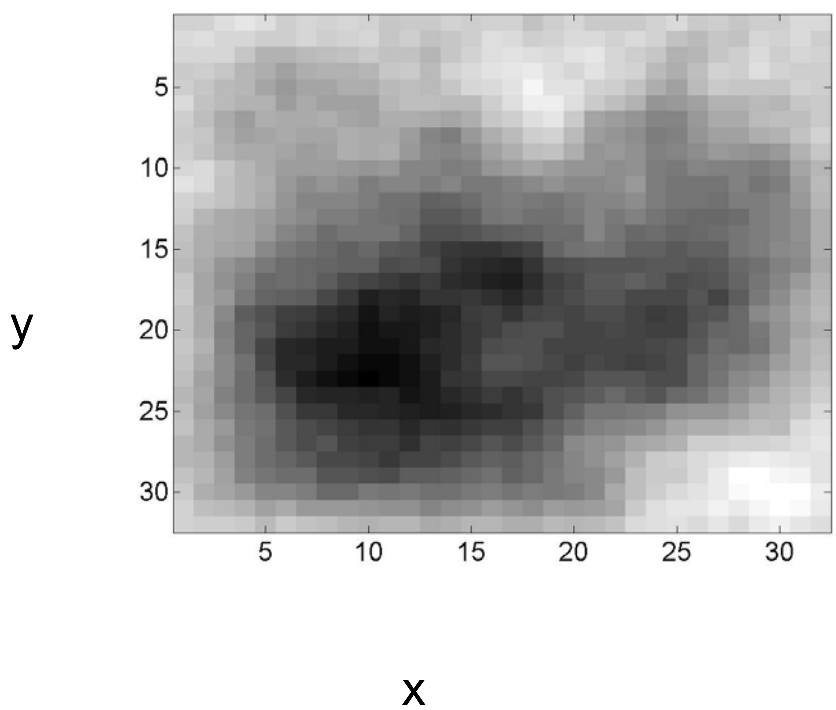
$$q\left(\alpha_k\right) = Ga\left(\alpha_k; g_k, h_k\right)$$

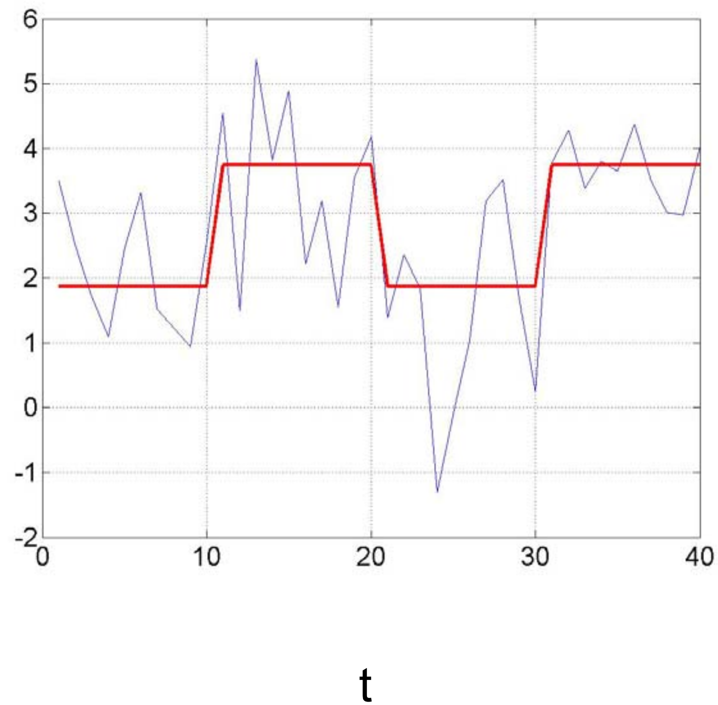
$$\frac{1}{g_k} = \frac{1}{2}\left[Tr\left(\widehat{\boldsymbol{\Sigma}}_k \mathbf{S}^T \mathbf{S}\right) + \widehat{\mathbf{w}}_k^{\ T} \mathbf{S}^T \mathbf{S} \widehat{\mathbf{w}}_k\right] + \frac{1}{q_1}$$

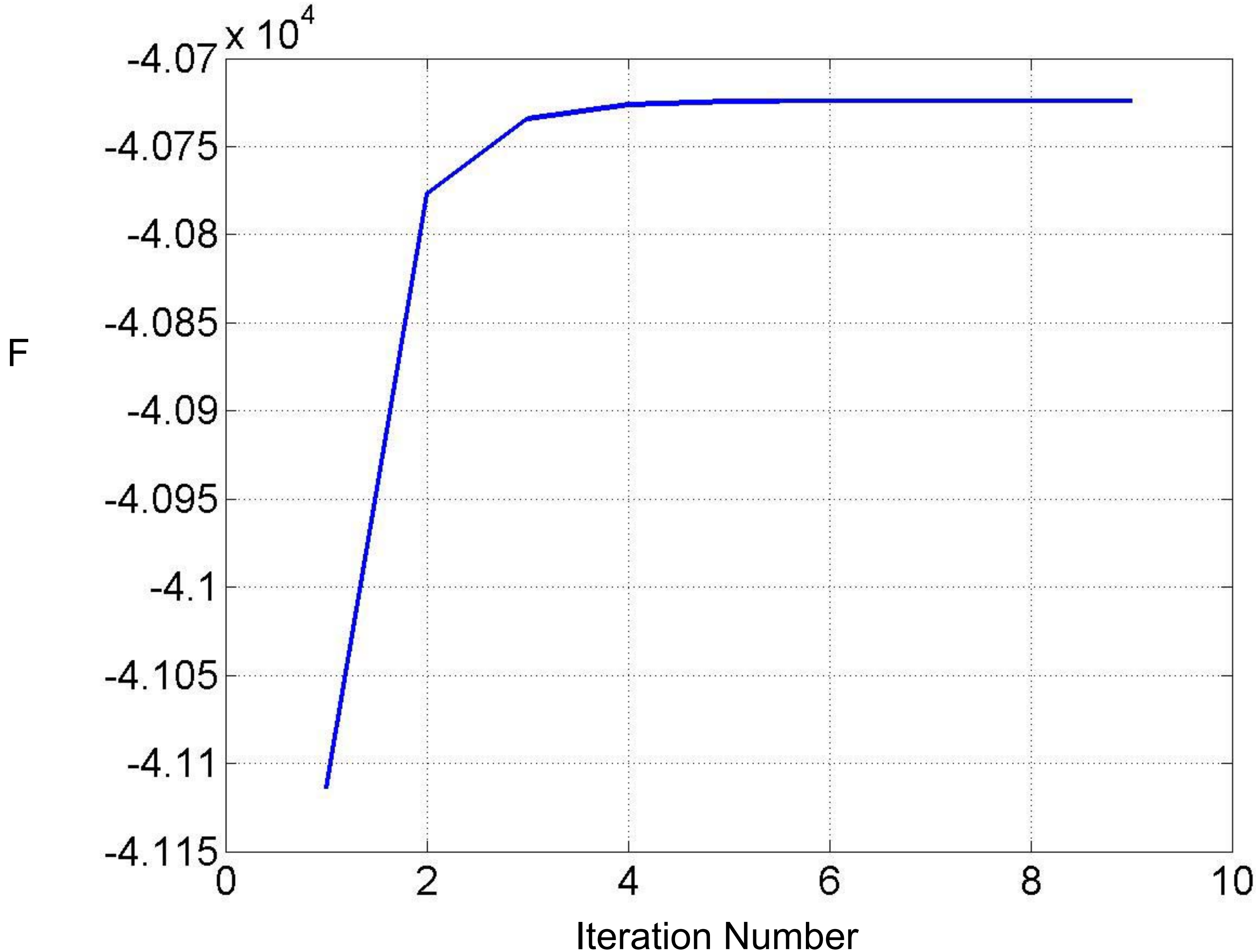
$$h_k = \frac{N}{2} + q_2$$

$$\overline{\alpha}_k = g_k h_k$$

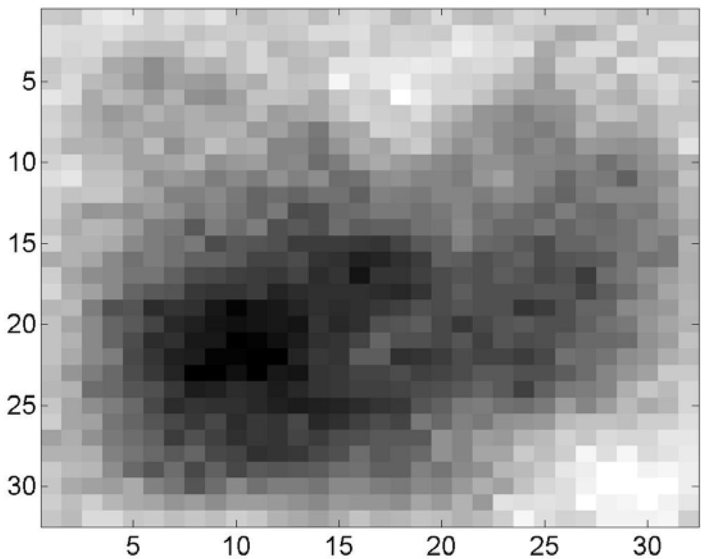
Synthetic Data 1 : from Laplacian Prior






x 10^4
-4.07
-4.075
-4.08
-4.085
-4.09
-4.095
-4.1
-4.105
-4.11
-4.115
F
0
2
4
6
8
10
Iteration Number

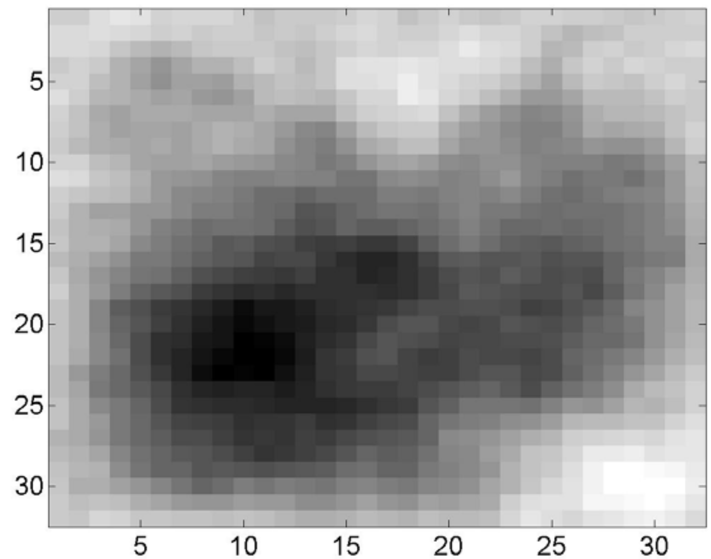
Least Squares



x

Coefficients = 1024

VB – Laplacian Prior



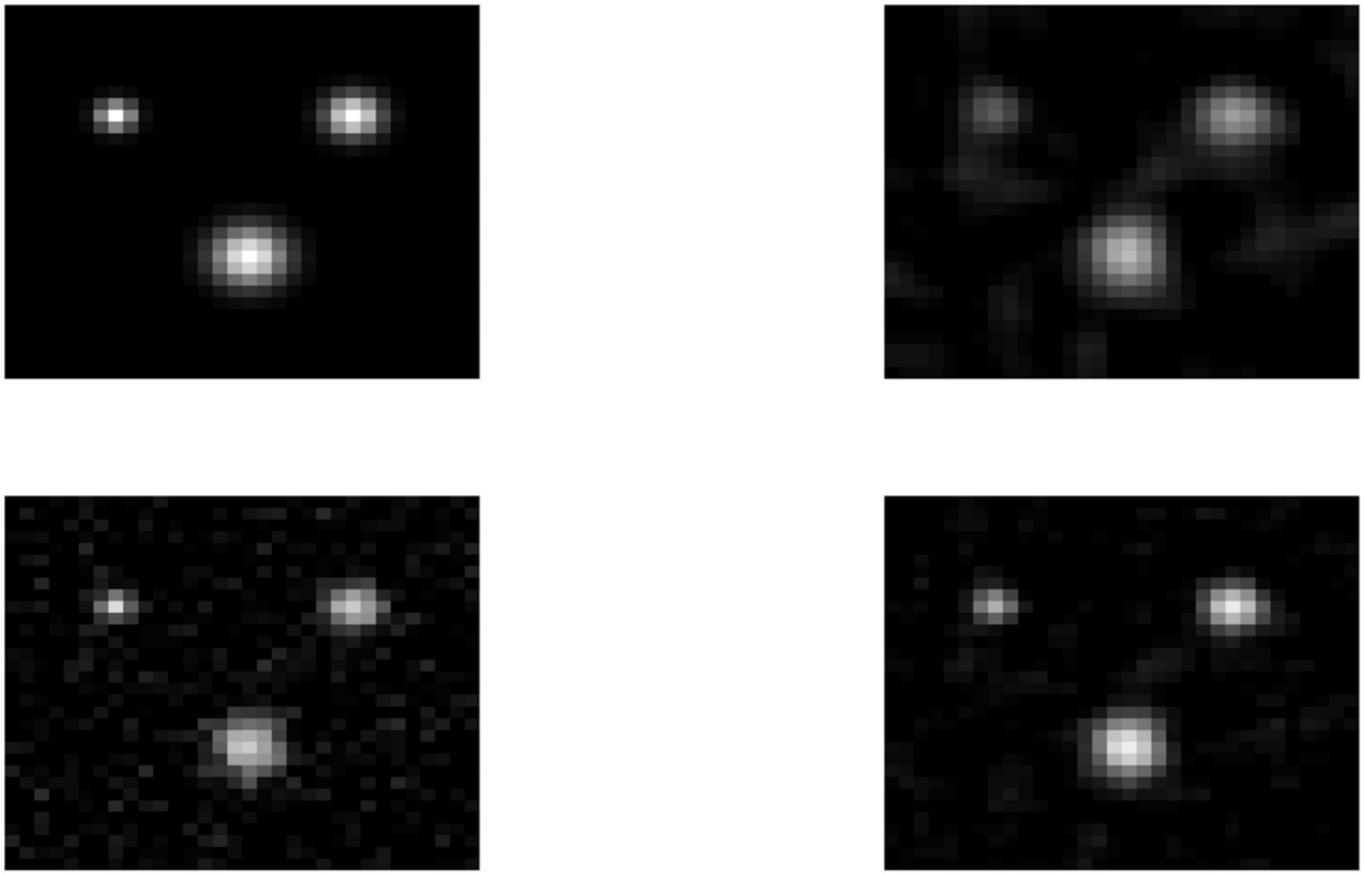
x

`Coefficient RESELS' = 366

Synthetic Data II : blobs

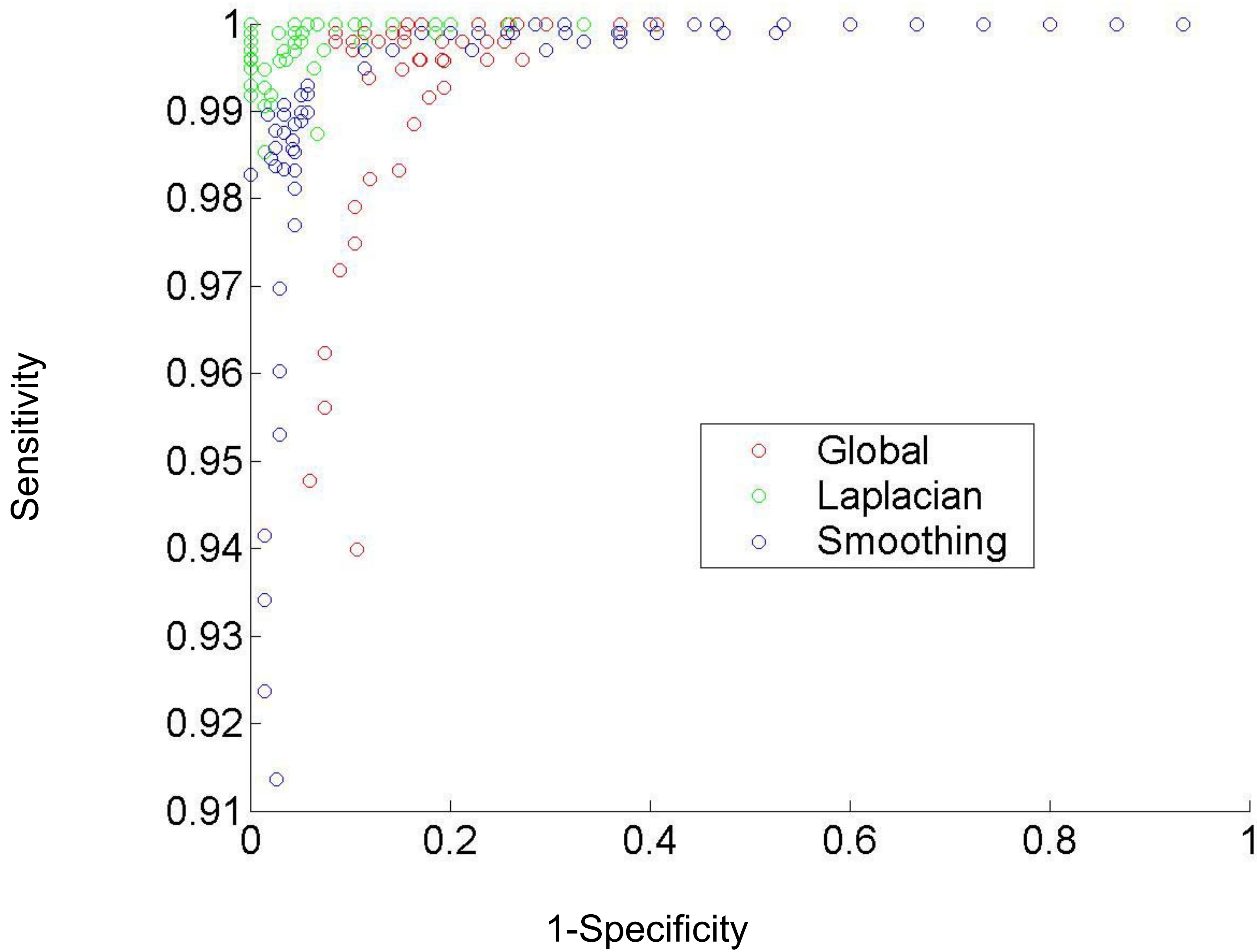
True

Smoothing



Global prior

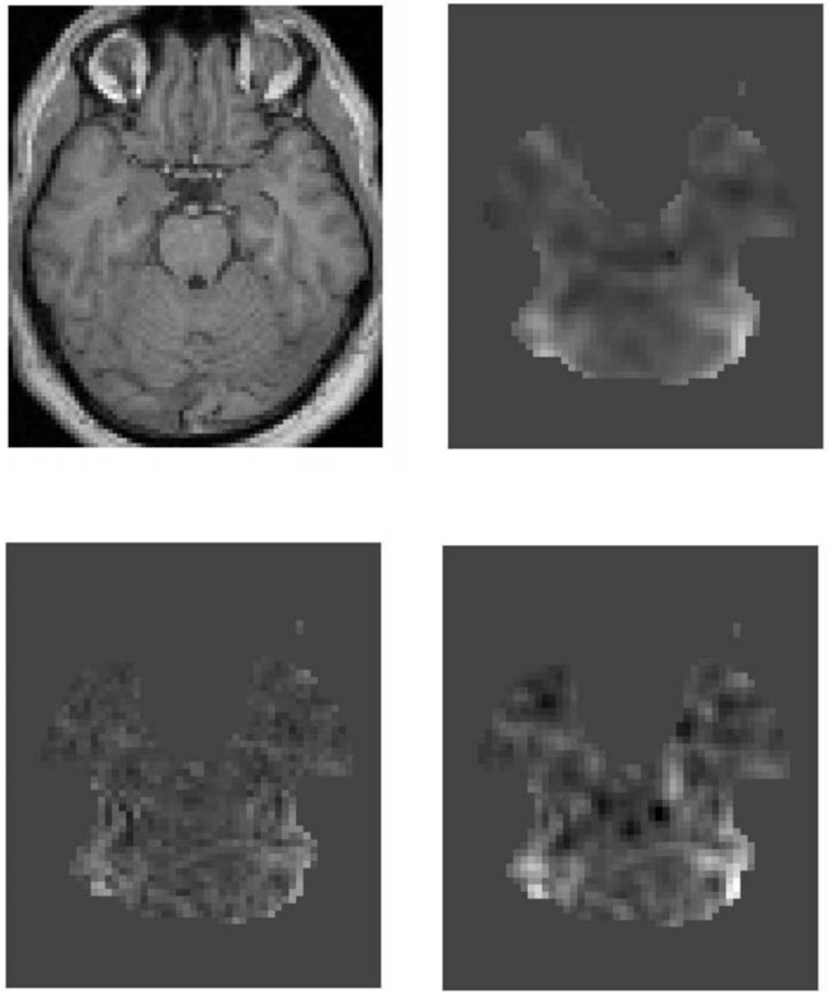
Laplacian prior



Sensitivity
1-Specificity
Global
Laplacian
Smoothing
0.91
0.92
0.93
0.94
0.95
0.96
0.97
0.98
0.99
1
0
0.2
0.4
0.6
0.8
1

Event-related fMRI: Faces versus chequerboard

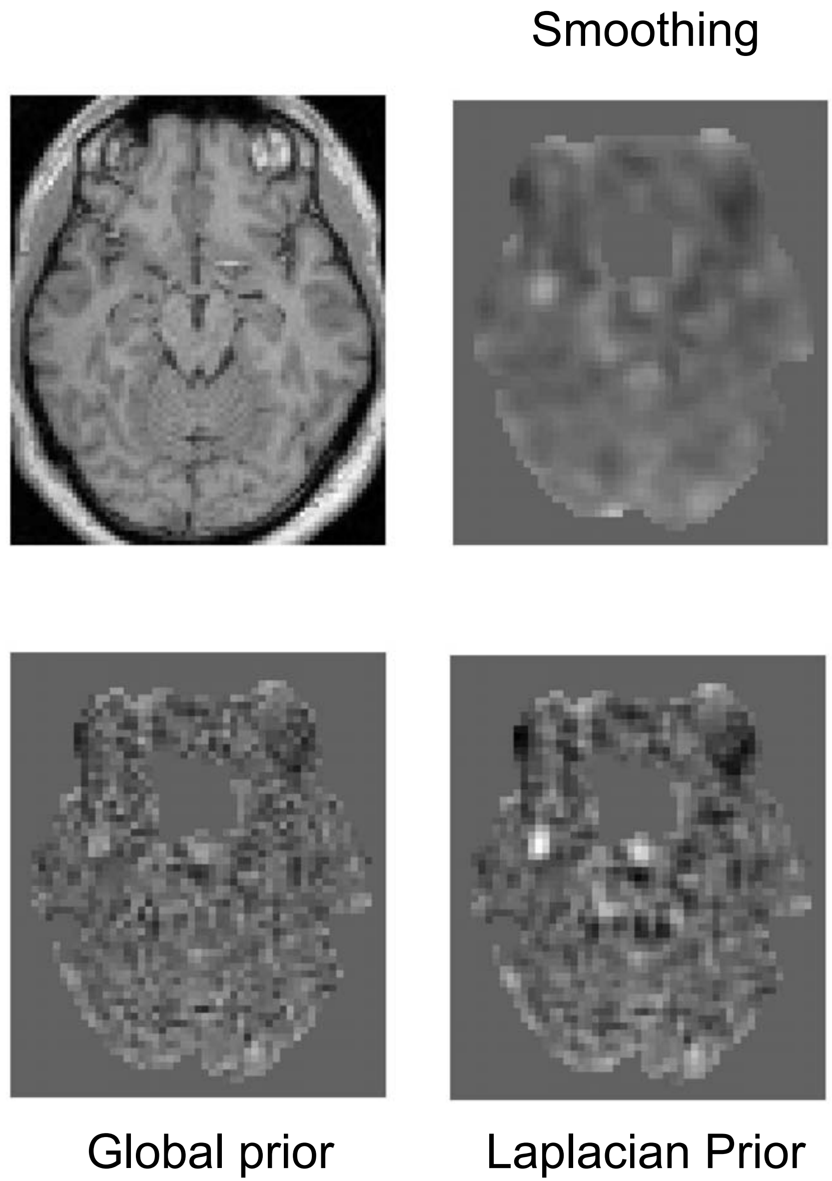
Smoothing



Global prior

Laplacian Prior

Event-related fMRI: Familiar faces versus unfamiliar faces



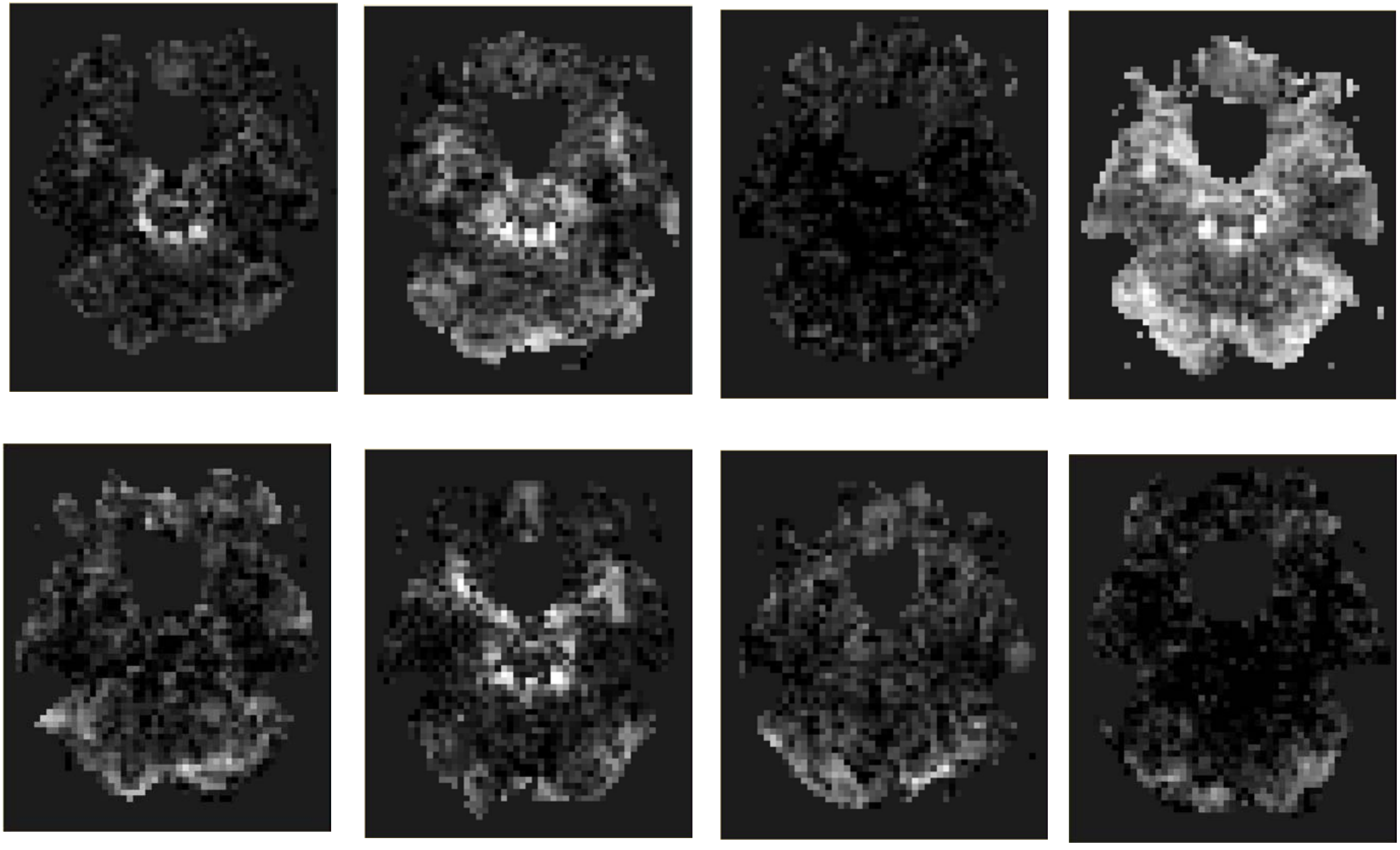
Summary

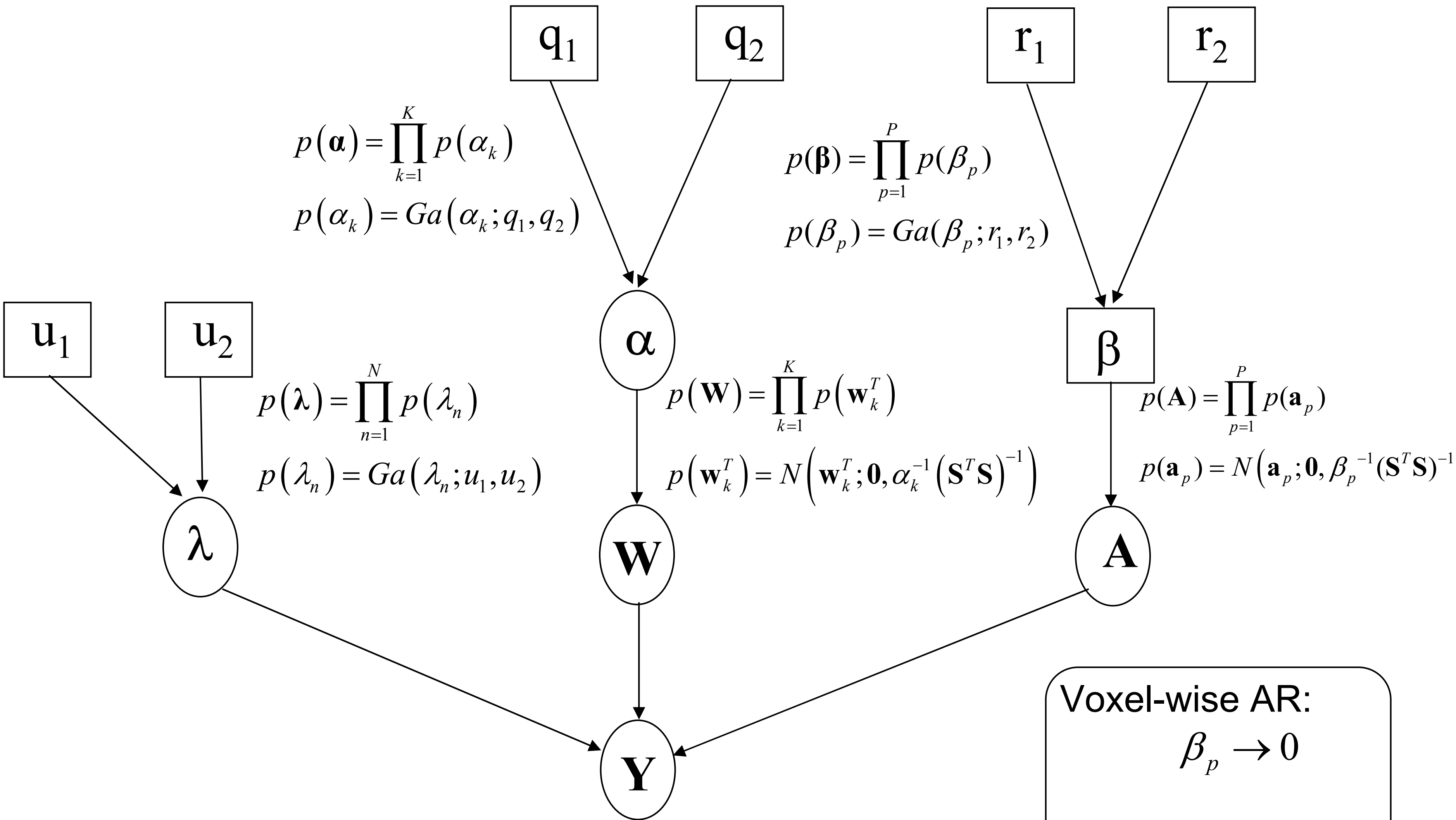
- No need to smooth data at 1st level
- Greater sensitivity
- Implementation in parallel to non-Bayesian methods – `spm/devel`
- 5-10 minutes per slice
- Model comparison of signal and noise models for single/multiple regions/slices

The future

- Laplacian priors for AR coefficients
- Mixture priors
- Wavelet priors
- 2^{nd} level noise models

Maps of AR(1) coefficient





Y=XW+E

[TxN] [TxK] [KxN] [TxN]

Updating approximate posterior

Regression coefficients

$$q(\mathbf{w}_n) = N\left(\mathbf{w}_n; \hat{\mathbf{w}}_n, \hat{\mathbf{\Sigma}}_n\right)$$

$$\widehat{\mathbf{w}}_n = \hat{\mathbf{\Sigma}}_n \left(\overline{\lambda_n} \tilde{\mathbf{b}}_n^{\ T} + \mathbf{r}_n\right)$$

$$\hat{\mathbf{\Sigma}}_n = \left(\overline{\lambda_n} \tilde{\mathbf{A}}_n + \mathbf{B}_{nn}\right)^{-1}$$

$$\mathbf{B} = \mathbf{H}\left(diag[\boldsymbol{\alpha}] \otimes \mathbf{S}^T\mathbf{S}\right)\mathbf{H}^T$$

$$\mathbf{r}_n = -\sum_{i=1, i\neq n}^{N} \mathbf{B}_{ni} \hat{\mathbf{w}}_i$$

AR coefficients

$$q(\mathbf{a}_n) = N(\mathbf{a}_n; \mathbf{m}_n, \mathbf{V}_n)$$

$$\mathbf{V}_n = \left(\lambda_n \widetilde{\mathbf{C}}_n + \mathbf{J}_{nn}\right)^{-1}$$

$$\mathbf{m}_n = \mathbf{V}_n \left(\lambda_n \widetilde{\mathbf{D}}_n + \mathbf{j}_n\right)$$

$$\mathbf{J} = \mathbf{H}\left(diag(\overline{\boldsymbol{\beta}}) \otimes \mathbf{S}^T\mathbf{S}\right)\mathbf{H}^T$$

$$\mathbf{j}_n = -\sum_{i=1, i\neq n}^{N} \mathbf{J}_{ni} \mathbf{m}_i$$

Spatial precisions for **W**

$$q(\boldsymbol{\alpha}) = \prod_{k=1}^{K} q(\alpha_k)$$

$$q(\alpha_k) = Ga(\alpha_k; g_k, h_k)$$

$$\frac{1}{g_k} = \frac{1}{2}\left[Tr\left(\hat{\mathbf{\Sigma}}_k \mathbf{S}^T\mathbf{S}\right) + \widehat{\mathbf{w}}_k^{\ T} \mathbf{S}^T\mathbf{S}\widehat{\mathbf{w}}_k\right] + \frac{1}{q_1}$$

$$h_k = \frac{N}{2} + q_2$$

$$\overline{\alpha}_k = g_k h_k$$

Spatial precisions for **A**

$$q(\boldsymbol{\beta}) = \prod_{p=1}^{P} q(\beta_p)$$

$$q(\beta_p) = Ga(\beta_p; r_{1p}, r_{2p})$$

$$\frac{1}{r_{1p}} = \frac{1}{2}\left(Tr(\mathbf{V}_p\mathbf{S}^T\mathbf{S}) + \mathbf{m}_p^{\ T}\mathbf{S}^T\mathbf{S}\mathbf{m}_p\right) + \frac{1}{r_1}$$

$$r_{2p} = \frac{P}{2} + r_2$$

$$\overline{\beta}_p = r_{1p} r_{2p}$$

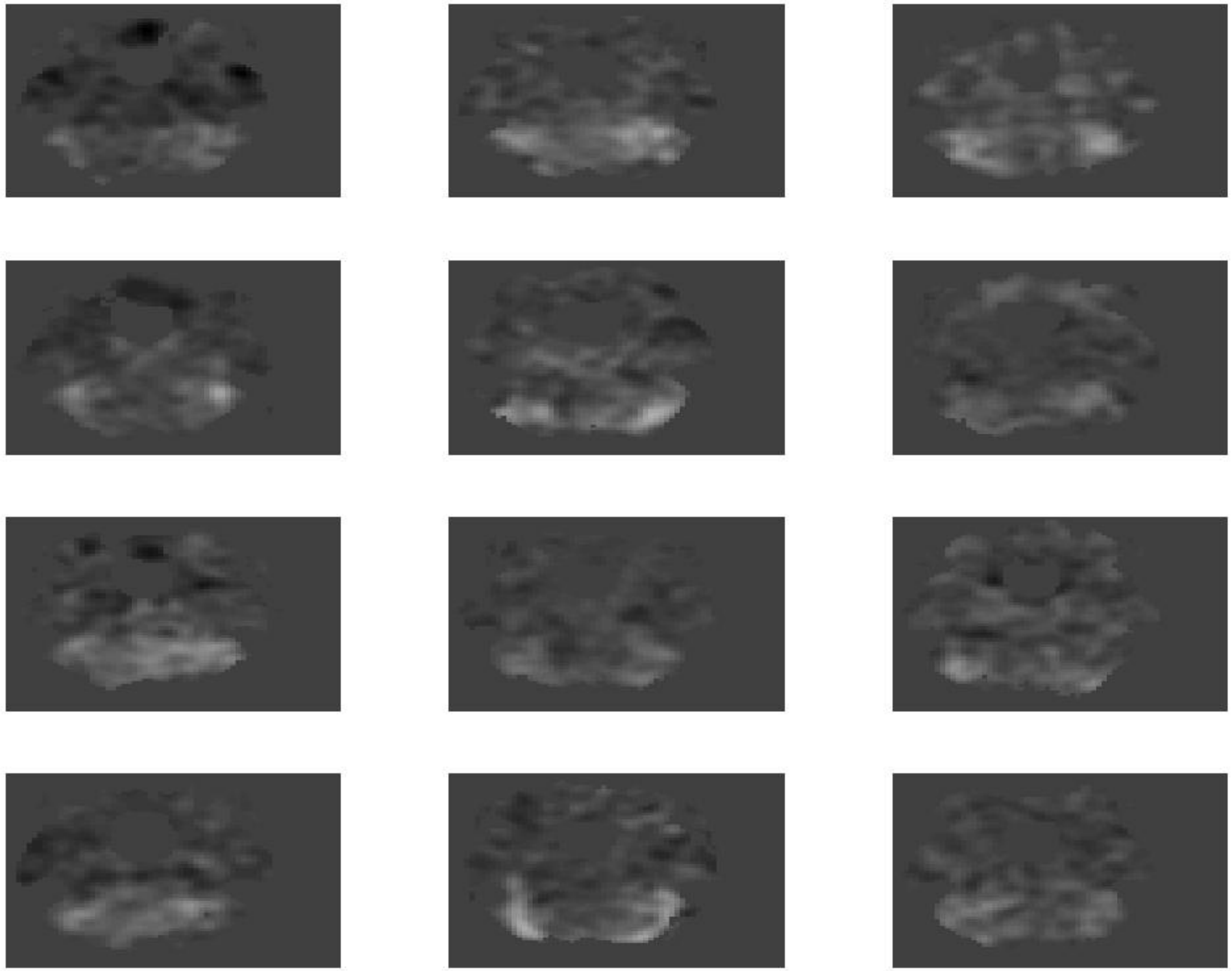
Observation noise

$$q(\lambda_n) = Ga(\lambda_n; b_n, c_n)$$

$$\frac{1}{b_n} = \frac{\widetilde{G}_n}{2} + \frac{1}{u_1}$$

$$c_n = \frac{T}{2} + u_2$$

Gaussian smoothing



Wavelet smoothing