Beyond the voxel: comparing spatially extended fMRI models

William D. Penny, Guillaume Flandin, Rik Henson and Karl J. Friston

Wellcome Department of Imaging Neuroscience, UCL, London UK.

Corresponding author:

William D. Penny,

Wellcome Department of Imaging Neuroscience,

12 Queen Square, London WC1N 3BG, UK.

Key Words: Variational Bayes, fMRI, spatial priors, effect size, general linear model, autoregressive model, Laplacian, smoothing, spatial regularisation

ABSTRACT

Model comparison of design matrices for modelling HRFs on slices and regions – not just voxels eg. FIR versus Canonical Set. Model comparison of AR error models – slice specific, voxel specific, smoothly varying. Results are shown on simulated data and on data from an event-related fMRI experiment.

1. INTRODUCTION

1.2 Notation

A multivariate normal density over $\mathbf{x}$ is written as $N(\mathbf{x};\mathbf{m},\mathbf{\Sigma})$ where $\mathbf{m}$ denotes the mean and $\mathbf{\Sigma}$ the covariance. A Gamma distribution over x is written $Ga(x;a,b)$ where a and b define the density as shown in the appendix of (Penny et al., 2003h). We will denote matrices and vectors with bold upper case and bold lower case letters respectively, and all vectors are column vectors. Subscripts are used to name different vectors; thus $\mathbf{x}_n$ and $\mathbf{x}_k$ refer to different vectors. The operator $diag(\mathbf{x})$ turns a vector into a diagonal matrix, $\mathbf{x}(k)$ denotes the kth entry in a vector, $\mathbf{X}(j,k)$ the scalar entry in the jth row and kth column, $\otimes$ is the Kronecker product and $\overline{x}$ is used to denote the mean of x.

2.THEORY

We write an fMRI data set consisting of T time points at N voxels as the T-*by*-N matrix $\mathbf{Y}$. In mass-univariate models this data is explained in terms of a T-*by*-K design matrix $\mathbf{X}$, containing the values of K regressors at T time points, and a K-*by*-N matrix of regression coefficients $\mathbf{W}$, containing K regression coefficients at each of N voxels. The model is written

$$\mathbf{Y} = \mathbf{X}\mathbf{W} + \mathbf{E} \tag{1}$$

where $\mathbf{E}$ is a T-*by*-N error matrix. The vector $\mathbf{w}_n$, the nth column of $\mathbf{W}$, therefore contains the K regression coefficients at the nth voxel and the vector $\mathbf{w}_k^T$, the kth row of $\mathbf{W}$, contains an image (after appropriate reshaping) of the kth regression

coefficients. We also make use of the *KN-by-1* vector $\mathbf{w}_v$ which contains all the elements of **W** ordered by voxel. Similarly we define the *NK-by-1* vector $\mathbf{w}_r$ which also contains all elements of W but ordered by regressor. These can both be defined using the *vec* operator which stacks columns of a matrix into one long vector

$$\begin{aligned} \mathbf{w}_v &= vec(\mathbf{W}) \\ \mathbf{w}_r &= vec(\mathbf{W}^T) \\ \mathbf{w}_v &= \mathbf{H}\mathbf{w}_r \end{aligned} \tag{2}$$

where **H** is a *KN-by-KN* permutation matrix. It is useful to define these high-dimensional vectors as the model can then be instantiated using sparse matrix operations.

In this paper the errors are modeled as an autoregressive process. The overall GLM-AR model can be written

$$\begin{aligned} \mathbf{y}_n &= \mathbf{X}\mathbf{w}_n + \mathbf{e}_n \\ \mathbf{e}_n &= \tilde{\mathbf{E}}_n\mathbf{a}_n + \mathbf{z}_n \end{aligned} \tag{3}$$

where, at the *n*th voxel, $\mathbf{a}_n$ is a *P-by-1* vector of regression coefficients, $\mathbf{z}_n$ is a vector of zero mean Gaussian random variables each having precision λ_n and $\tilde{\mathbf{E}}_n$ is a *T-by-P* matrix of lagged prediction errors for the *n*th voxel as defined in section 2 of (Penny et al., 2003g).

2.1 Regression coefficient prior

The prior over regression coefficients is given by

$$p(\mathbf{W}) = \prod_{k=1}^{K} p\left(\mathbf{w}_k^T\right)$$
$$p\left(\mathbf{w}_k^T\right) = N\left(\mathbf{w}_k^T; \mathbf{0}, \alpha_k^{-1}\left(\mathbf{S}^T\mathbf{S}\right)^{-1}\right) \quad (4)$$

where we refer to **S** as an *N-by-N* spatial kernel matrix (to be defined later) and α_k is a spatial precision variable for the *k*th regressor. This equation shows that the prior factorises over regressors. This means that different regression coefficients can have different smoothnesses. For the case of **S** being the Laplacian operator (see below) a sample from the prior is shown in Figure 3. The precision variables α_k are collected together in the *K-by-1* vector $\boldsymbol{\alpha}$. The prior over $\boldsymbol{\alpha}$ is given by

$$p(\boldsymbol{\alpha}) = \prod_{k=1}^{K} p(\alpha_k)$$
$$p(\alpha_k) = Ga(\alpha_k; q_1, q_2) \quad (5)$$

2.2 Autoregressive coefficient prior

In contrast to previous work (Penny et al., 2003f), which has used uninformative priors, in this paper we use a spatial prior given by

$$p(\mathbf{A}) = \prod_{p=1}^{P} p(\mathbf{a}_p)$$
$$p(\mathbf{a}_p) = N(\mathbf{a}_p; \mathbf{0}, \beta_p^{-1}(\mathbf{S}^T\mathbf{S})^{-1}) \quad (6)$$

where ***A*** is a matrix of AR coefficients, $\boldsymbol{a}_p$ is an N-by-1 vector of AR coefficients (the *n*th entry being the *p*th regression coefficient at the *n*th voxel) and β_p is the prior spatial precision of the *p*th regression coefficient. The prior over $\boldsymbol{\beta}$ is given by

$$p(\boldsymbol{\beta}) = \prod_{p=1}^{P} p(\beta_p)$$
$$p(\beta_p) = Ga(\beta_p; r_1, r_2) \quad (7)$$

2.3 Noise precision prior

The observation noise precisions are defined as

$$p(\boldsymbol{\lambda}) = \prod_{n=1}^{N} p(\lambda_n)$$
$$p(\lambda_n) = Ga(\lambda_n; u_1, u_2) \tag{8}$$

The values u_1 and u_2 are set so as to make $p(\lambda_n)$ an uninformative prior as described in (Penny et al., 2003e).

2.3 Approximate Posteriors

Regression coefficients

The posterior over regression coefficients is assumed to factorize over voxels. That is

$$q(\mathbf{W}) = \prod_{n=1}^{N} q(\mathbf{w}_n) \tag{9}$$

This is the key assumption of this paper and is central to the derivation of update rules for the posteriors. These derivations follow the general principles of the VB framework outlined in our previous work and will not be elaborated upon here.

For our model this results in a posterior over regression coefficients at voxel n given by

$$q(\mathbf{w}_n) = N\left(\mathbf{w}_n; \hat{\mathbf{w}}_n, \hat{\boldsymbol{\Sigma}}_n\right) \tag{10}$$

where

$$\begin{aligned}\widehat{\mathbf{w}}_n &= \widehat{\boldsymbol{\Sigma}}_n\left(\overline{\lambda_n}\tilde{\mathbf{b}}_n^{\,T} + \mathbf{r}_n\right)\\ \widehat{\boldsymbol{\Sigma}}_n &= \left(\overline{\lambda_n}\tilde{\mathbf{A}}_n \; + \mathbf{B}_{nn}\right)^{-1}\end{aligned} \tag{11}$$

where the matrix $\tilde{\mathbf{A}}_n$ is related to the data precision at the *n*th voxel and the vector $\tilde{\mathbf{b}}_n$ is related to the data at the *n*th voxel projected onto the design matrix. These quantities are identical to those defined in equations 63 and 64 in (Penny et al., 2003d). The matrix **B** is the *NK-by-NK* spatial precision matrix (with entries ordered by voxel – hence the permutation matrix **H** in the following equation) and is given by

$$\mathbf{B} = \mathbf{H}\left(diag\left[\boldsymbol{\alpha}\right] \otimes \mathbf{S}^T\mathbf{S}\right)\mathbf{H}^T \tag{12}$$

The quantity $\mathbf{r}_n$ is given by

$$\mathbf{r}_n = -\sum_{i=1, i\neq n}^{N} \mathbf{B}_{ni}\hat{\mathbf{w}}_i \tag{13}$$

The subscripts in $\mathbf{B}_{ni}$ denote those entries in **B** pertaining to voxels *n* and *i*. An intuitive description of this posterior is given in section 2.4.

AR coefficients

For the autoregressive coefficients we have

$$\begin{aligned}q(\mathbf{a}_n) &= N(\mathbf{a}_n; \mathbf{m}_n, \mathbf{V}_n)\\ \mathbf{V}_n &= \left(\lambda_n\widetilde{\mathbf{C}}_n + \mathbf{J}_{nn}\right)^{-1}\\ \mathbf{m}_n &= \mathbf{V}_n\left(\lambda_n\tilde{\mathbf{d}}_n + \mathbf{j}_n\right)\\ \mathbf{J} &= \mathbf{H}\left(diag(\bar{\boldsymbol{\beta}}) \otimes \mathbf{S}^T\mathbf{S}\right)\mathbf{H}^T\\ \mathbf{j}_n &= -\sum_{i=1, i\neq n}^{N} \mathbf{J}_{ni}\mathbf{m}_i\end{aligned} \tag{14}$$

where $\widetilde{\mathbf{C}}_n$ and $\tilde{\mathbf{d}}_n$ are quantities related to AR prediction error and are defined in equation 50 in the appendix of (Penny et al., 2003c). The subscripts in $\mathbf{J}_{ni}$ denote those entries in **J** pertaining to voxels *n* and *i*.

Spatial precisions

The posteriors over the spatial precision variables for the regression coefficients are given by

$$
\begin{aligned}
q(\boldsymbol{\alpha}) &= \prod_{k=1}^{K} q(\alpha_k) \\
q(\alpha_k) &= Ga(\alpha_k; g_k, h_k) \\
\frac{1}{g_k} &= \frac{1}{2}\left[Tr\left(\hat{\boldsymbol{\Sigma}}_k \mathbf{S}^T \mathbf{S}\right) + \hat{\mathbf{w}}_k{}^T \mathbf{S}^T \mathbf{S} \hat{\mathbf{w}}_k \right] + \frac{1}{q_1} \\
h_k &= \frac{N}{2} + q_2 \\
\bar{\alpha}_k &= g_k h_k
\end{aligned}
\tag{15}
$$

where $\hat{\boldsymbol{\Sigma}}_k$ is a *N-by-N* diagonal matrix with *n*th entry $\hat{\boldsymbol{\Sigma}}_n(k,k)$.

The posteriors over the spatial precision variables for the AR coefficients are given by

$$
\begin{aligned}
q(\boldsymbol{\beta}) &= \prod_{p=1}^{P} q(\beta_p) \\
q(\beta_p) &= Ga(\beta_p; r_{1p}, r_{2p}) \\
\frac{1}{r_{1p}} &= \frac{1}{2}\left(Tr(\mathbf{V}_p \mathbf{S}^T \mathbf{S}) + \mathbf{m}_p{}^T \mathbf{S}^T \mathbf{S} \mathbf{m}_p \right) + \frac{1}{r_1} \\
r_{2p} &= \frac{P}{2} + r_2 \\
\bar{\beta}_p &= r_{1p} r_{2p}
\end{aligned}
\tag{16}
$$

Observation Noise

The posteriors over the noise precisions are identical to those defined in previous work (Penny et al., 2003b)

$$
\begin{aligned}
q(\lambda_n) &= Ga(\lambda_n; b_n, c_n) \\
\frac{1}{b_n} &= \frac{\widetilde{G_n}}{2} + \frac{1}{u_1} \\
c_n &= \frac{T}{2} + u_2
\end{aligned}
\tag{17}
$$

where $\widetilde{G}_n$ is related to the GLM prediction error and is defined for a single voxel in equation 77 in the appendix of (Penny et al., 2003a).

2.4 Model Evidence

The objective function for the algorithm is the lower bound on the logarithm of the model evidence which for our model is given by

$$F = L_{av} - \left(KL(\mathbf{W}) + KL(\mathbf{A}) + KL(\boldsymbol{\alpha}) + KL(\boldsymbol{\lambda}) + KL(\boldsymbol{\beta}) \right) \tag{18}$$

where L_{av} is the average log-likelihood and the KL terms are the Kullback-Liebler divergences between the priors and approximate posteriors. These are computed using standard results for KL-divergences for Gamma and Normal distributions given in (Roberts and Penny, 2002). The average log-likelihood is given by

$$L_{av} = \sum_n \frac{T}{2} \left(\varphi(c_n) + \log(b_n) \right) - \frac{\lambda_n}{2} \tilde{G}_n \tag{19}$$

where $\varphi()$ is the digamma function (Press et al., 1992).

3. RESULTS

4. DISCUSSION

Acknowledgements

Will Penny is supported by the Wellcome Trust.

Figure Captions

Figure 1. The figure shows the probabilistic dependencies underlying our generative model for fMRI data. The quantities in square brackets are constants and those in circles are random variables. The spatial regularization coefficients $\boldsymbol{\alpha}$ constrain the regression coefficients **W**, and the spatial regularization coefficients $\boldsymbol{\beta}$ constrain the AR coefficients **A**. The parameters $\boldsymbol{\lambda}$ control the observation noise precision at each voxel. The graph shows that the joint probability of parameters and data can be written

$$p(\mathbf{Y},\mathbf{W},\mathbf{A},\boldsymbol{\lambda},\boldsymbol{\alpha},\boldsymbol{\beta}) = p(\mathbf{Y}\mid\mathbf{W},\mathbf{A},\boldsymbol{\lambda})p(\mathbf{W}\mid\boldsymbol{\alpha})p(\mathbf{A}\mid\boldsymbol{\beta})p(\boldsymbol{\lambda}\mid u_1,u_2)p(\boldsymbol{\alpha}\mid q_1,q_2)p(\boldsymbol{\beta}\mid r_1,r_2)$$

where the first term is the likelihood and the other terms are priors.

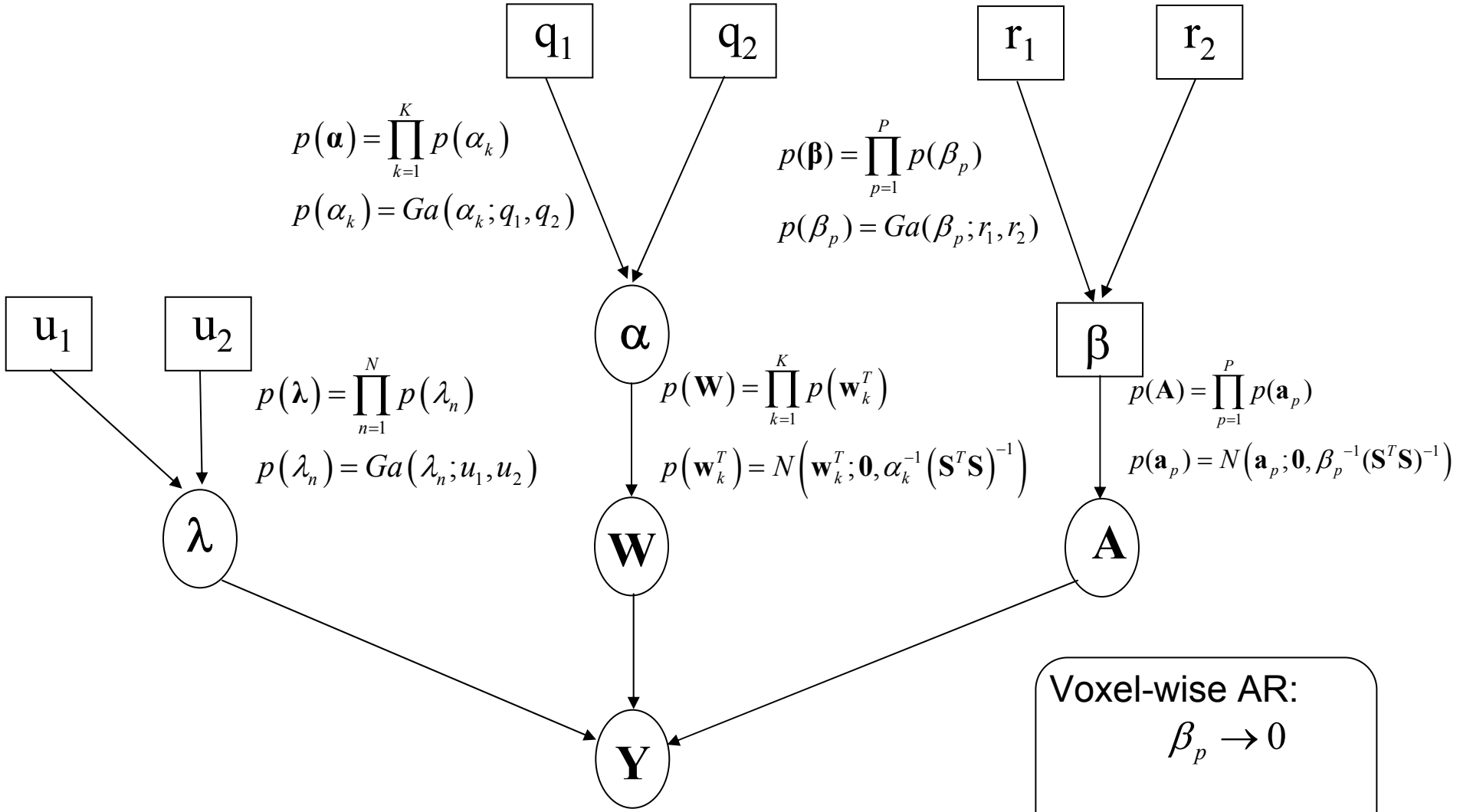
Figure 2. Summary of VB-update equations

Reference List

1. Penny,W., Kiebel,S., and Friston,K. (2003f). Variational Bayesian inference for fMRI time series. Neuroimage. *19*, 727-741.

9. Press,W.H., Teukolsky,S.A., Vetterling,W.T., and Flannery,B.V.P. (1992). Numerical Recipes in C. Cambridge).

10. Roberts,S. and Penny,W.D. (2002). Variational Bayes for generalised autoregressive models. IEEE Trans.Signal Processing *50*, 2245-2257.


q1
q2
r1
r2
$p(\boldsymbol{\alpha}) = \prod_{k=1}^{K} p(\alpha_k)$
$p(\alpha_k) = Ga(\alpha_k; q_1, q_2)$
$p(\boldsymbol{\beta}) = \prod_{p=1}^{P} p(\beta_p)$
$p(\beta_p) = Ga(\beta_p; r_1, r_2)$
u1
u2
α
β
$p(\boldsymbol{\lambda}) = \prod_{n=1}^{N} p(\lambda_n)$
$p(\lambda_n) = Ga(\lambda_n; u_1, u_2)$
$p(\mathbf{W}) = \prod_{k=1}^{K} p(\mathbf{w}_k^T)$
$p(\mathbf{w}_k^T) = N(\mathbf{w}_k^T; \mathbf{0}, \alpha_k^{-1}(\mathbf{S}^T\mathbf{S})^{-1})$
$p(\mathbf{A}) = \prod_{p=1}^{P} p(\mathbf{a}_p)$
$p(\mathbf{a}_p) = N(\mathbf{a}_p; \mathbf{0}, \beta_p^{-1}(\mathbf{S}^T\mathbf{S})^{-1})$
λ
W
A
Y
Y=XW+E
[TxN] [TxK] [KxN] [TxN]
Voxel-wise AR:
$\beta_p \to 0$
Spatial pooled AR:
$\beta_p \to \infty$

Regression coefficients, **W**

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

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

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

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

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

AR coefficients, **A**

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

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

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

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

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

Spatial precisions for **W**

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

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

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

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

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

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

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

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

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