Let us consider a k-variate Besag-York-Mollié model. We assume disease-specific mixing parameters, hence we rely on the M-models [1] framework.

If we define the matrix-valued mixing parameter$\Phi := \text{diag}(\phi_1, \phi_2, \dots \phi_k)$ and $\bar{\Phi} := \text{diag}(1-\phi_1, 1-\phi_2, \dots 1-\phi_k) = (I_k - \Phi)$, the BYM field $\mathbf{Y}$ is given by:

$$\mathbf{Y} = \mathbf{U}\Phi^{\frac{1}{2}}M + \mathbf{V}\bar{\Phi}^{\frac{1}{2}}M \tag{1}$$

Where

- $\mathbf{U} = (U_1, U_2, \dots U_k)$ is a multivariate and independent ICAR field, such that $U_j \sim \mathcal{N}_n(0, L^+)$ for $j = 1, 2 \dots k$, and L^+ is the pseudoinverse of the scaled graph Laplacian matrix
- $\mathbf{V} = (V_1, V_2, \dots V_k)$ is a collection of *iid* random Normal vectors, such that $V_j \sim \mathcal{N}_n(0, I_n)$ for $j = 1, 2 \dots k$.
- M is a generic $k \times k$ full-rank matrix such that $M^\top M = \Sigma$, being Σ is the scale parameter. For instance, M can be defined as $D^{1/2}E^\top$, where D is the diagonal matrix of the eigenvalues of Σ and E is the corresponding eigenvectors matrix.

And the prior variance of $\mathbf{Y}$ is:

$$\text{Var}\left[\text{vec}(\mathbf{Y}) \mid \Sigma, \Phi\right] = \left(M^\top \otimes I_n\right) \text{diag}\left(S_1, \dots, S_k\right)\left(M \otimes I_n\right)$$

Where each diagonal block S_j is the variance of $U_j\Phi + V_j\bar{\Phi}$, i.e. $S_j = \phi_j L^+ + (1-\phi_j)I_n$.

Extending the approach of [2] to the multivariate case, we know that:

$$\mathbb{E}[\text{vec}(\mathbf{Y}) \mid U, \Phi, \Sigma] = \text{vec}(\mathbf{U}\Phi^{\frac{1}{2}}M) = [(M^\top\Phi^{\frac{1}{2}}) \otimes I_n]\text{vec}(\mathbf{U})$$

and similarly

$$\text{Var}[\text{vec}(\mathbf{Y}) \mid \mathbf{U}, \Phi, \Sigma] = \left[(M^\top\bar{\Phi}^{\frac{1}{2}}) \otimes I_n\right] \mathbb{E}\left[\text{vec}(\mathbf{V})\text{vec}(\mathbf{V})^\top\right] \left[(\bar{\Phi}^{\frac{1}{2}}M) \otimes I_n\right] = \left(M^\top\bar{\Phi}M\right) \otimes I_n$$

The distribution of $\mathbf{Y} \mid \mathbf{U}, \Sigma, \Phi$ then reads:

$$\begin{aligned}
-2\ln\pi\left(\text{vec}(\mathbf{Y}) \mid \mathbf{U}, \Sigma, \phi\right) = C + \text{vec}(\mathbf{Y})^\top &\left[\left(M^{-1}\bar{\Phi}^{-1}M^{-1^\top}\right) \otimes I_n\right] \text{vec}(\mathbf{Y}) \\
-2\text{vec}(\mathbf{Y})^\top \left[\left(M^{-1}\bar{\Phi}^{-1}M^{-1^\top}\right) \otimes I_n\right] &\left[\left(M^\top\Phi^{\frac{1}{2}}\right) \otimes I_n\right] \text{vec}(\mathbf{U}) \\
+\text{vec}(\mathbf{U})^\top \left[\left(\Phi^{\frac{1}{2}}M\right) \otimes I_n\right] \left[\left(M^{-1}\bar{\Phi}^{-1}M^{-1^\top}\right) \otimes I_n\right] &\left[(M^\top\Phi^{\frac{1}{2}}) \otimes I_n\right] \text{vec}(\mathbf{U}) \\
= C + \text{vec}(\mathbf{Y})^\top &\left[\left(M^{-1}\bar{\Phi}^{-1}M^{-1^\top}\right) \otimes I_n\right] \text{vec}(\mathbf{Y}) \\
-2\text{vec}(\mathbf{Y})^\top &\left[\left(M^{-1}\bar{\Phi}^{-1}\Phi^{\frac{1}{2}}M^\top\right) \otimes I_n\right] \text{vec}(\mathbf{U}) \\
+\text{vec}(\mathbf{U})^\top &\left[\left(\Phi\bar{\Phi}^{-1}\right) \otimes I_n\right] \text{vec}(\mathbf{U})
\end{aligned}$$

Now, for brevity let us define the following $k \times k$ matrices:

$$q_{11} := M^{-1}\bar{\Phi}^{-1}{M^{-1}}^{\top}; \quad q_{12} := M^{-1}\bar{\Phi}^{-1}\Phi^{\frac{1}{2}}; \quad q_{22} := \Phi\bar{\Phi}^{-1}$$

Hence

$$\begin{aligned} -2\ln\pi\left(\mathrm{vec}(\mathbf{Y}) \mid \mathbf{U}, \Sigma, \Phi\right) = C + \mathrm{vec}(\mathbf{Y})^{\top}\left(q_{11} \otimes I_n\right)\mathrm{vec}(\mathbf{Y}) \\ -2\mathrm{vec}(\mathbf{Y})^{\top}\left(q_{12} \otimes I_n\right)\,\mathrm{vec}(\mathbf{U}) \\ +\mathrm{vec}(\mathbf{U})^{\top}\left(q_{22} \otimes I_n\right)\mathrm{vec}(\mathbf{U}) \end{aligned}$$

Then, we have

$$\begin{aligned} -2\ln\pi\left(\mathrm{vec}(\mathbf{Y}), \mathrm{vec}(\mathbf{U}) \mid \Sigma, \phi\right) = C + \mathrm{vec}(\mathbf{Y})^{\top}\left(q_{11} \otimes I_n\right)\mathrm{vec}(\mathbf{Y}) \\ -2\mathrm{vec}(\mathbf{Y})^{\top}\left(q_{12} \otimes I_n\right)\,\mathrm{vec}(\mathbf{U}) \\ +\mathrm{vec}(\mathbf{U})^{\top}\left(q_{12} \otimes I_n + I_k \otimes L\right)\mathrm{vec}(\mathbf{U}) \end{aligned}$$

Hence, with some straightforward algebra, it can be concluded that:

$$\begin{pmatrix} \mathrm{vec}(\mathbf{Y}) \\ \mathrm{vec}(\mathbf{U}) \end{pmatrix} \sim N_{2kn}\left(0, \begin{pmatrix} q_{11} \otimes I_n & -q_{12} \otimes I_n \\ -q_{12}^{\top} \otimes I_n & q_{22} \otimes I_n + I_k \otimes L \end{pmatrix}^{-1}\right) \tag{2}$$

Which generalises to the multivariate case the sparse precision derived by [2].

References

[1] P. Botella-Rocamora, M.A. Martinez-Beneito, and S. Banerjee. A unifying modeling framework for highly multivariate disease mapping. Statistics in Medicine, 34(9):1548–1559, 2015

[2] A. Riebler, S. H. Sørbye, D. Simpson, and H. Rue. An intuitive Bayesian spatial model for disease mapping that accounts for scaling. Statistical methods in medical research, 25(4):1145–1165, 2016