Data Generation

The dataset contains 2000 simulated images, with 1000 images each in groups A and B. Each image has 256 pixels arranged in a 16×16 grid. The β matrix, set to 8 in an 8×8 central region and 0 elsewhere, modifies the pixel values in group A. The `group_ind` vector differentiates the groups, assigning 1 to group A and 0 to group B. Noise ϵ_i, generated from a multivariate normal distribution with zero mean and a covariance matrix from an exponential correlation function with rate 1, is added to the pixel values.

To adapt the data for the input specifications of the `myresf_vc` function, the dataset was transformed into a long format. In this format, the columns `x` and `y` denote the coordinates, while `pixel_value` corresponds to the simulated y values.

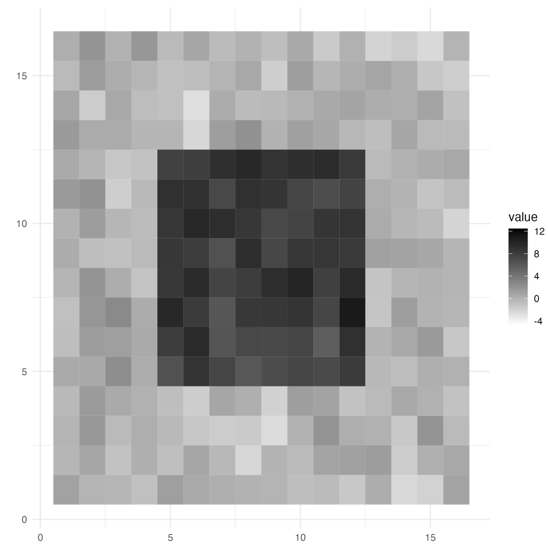

Figure 1: Example image from group A

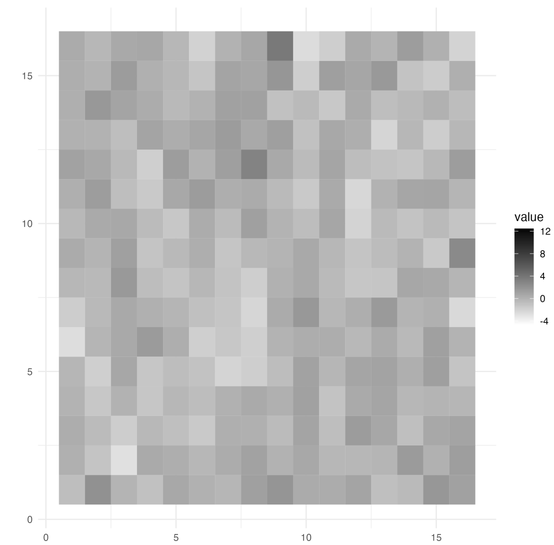

Figure 2: Example image from group B

VBM

In the VBM analysis, a Generalized Linear Model (GLM) was applied pixel-wise to assess group effects on pixel intensities across 100 iterations. For each iteration, the model generated effect size estimates and p-values for each pixel. These p-values were then corrected for multiple comparisons using the Bonferroni method. Figure **??** depicts the frequency of significant p-values in across pixels, with pixels showing significant β in all 100 iterations appearing in black, and those never showing significance in white.

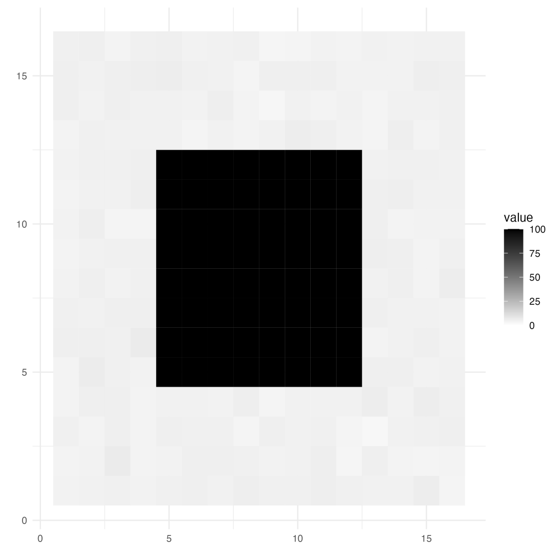

Figure 3: % of significant p-values across pixels in VBM analysis

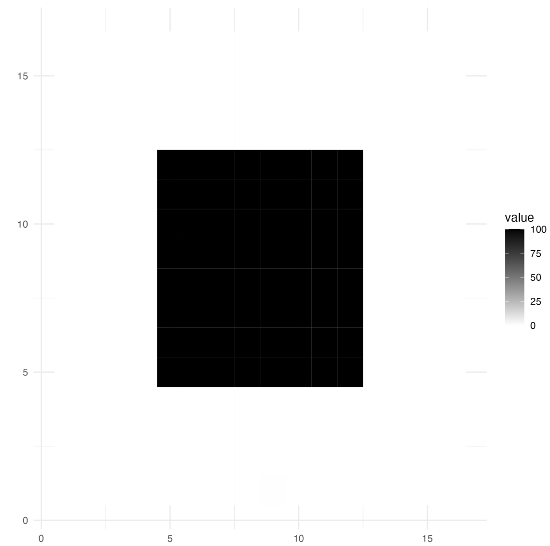

Figure 4: % of significant p-values after correction across pixels in VBM analysis

spVBM

The spVBM model is:

$$
\begin{aligned}
&y_s^i = \sum_{k=1}^{K} x_{s,k}^i \beta_{s,k}^{SVC} + \mathbf{Z^i b^i} + \varepsilon_s^i \\
&\beta_{s,k}^{SVC} = \beta_k + [\mathbf{E}\Gamma]_{s,k} \\
&\mathbf{b^i} \sim \mathcal{N}(\mathbf{0}, G), \quad \varepsilon_i \sim \mathcal{N}\left(0, \sigma^2\right), \quad \Gamma_{,k} \sim \mathcal{N}\left(\mathbf{0}, \sigma_k^2 \mathbf{\Lambda}\left(\alpha_k\right)\right)
\end{aligned}
$$

y_s^i denote the spatial outcome for subject i voxel s. $\mathbf{Z}$ denote non-spatial subject-level covariates for non-spatial random effects.

In our simulated data, this could be simplified to $y_s^i = x_s^i \beta_s^{SVC} + (?)$. My question is, which term captures the exponential correlation structure in the model? G or Γ?

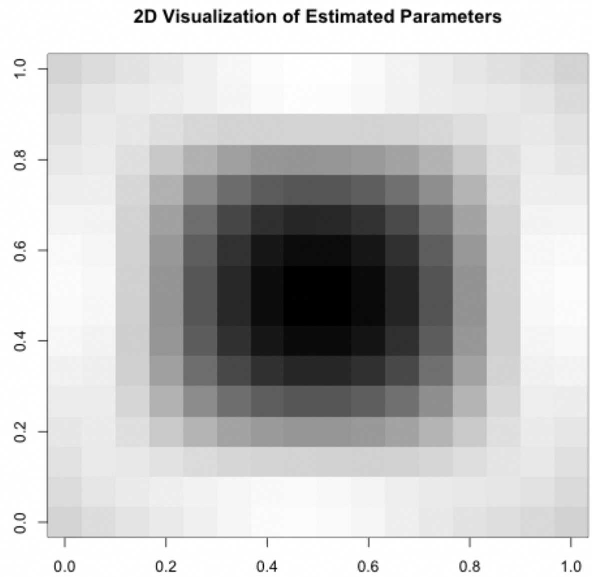

Figure 5: Estimated coefficients

LASSO

Predict `group_ind` from imaging using LASSO regresion via the `hdi` package in R. In each iteration, randomly partition the data, fit LASSO to the first subset to identify non-zero coefficients, and then apply OLS using these coefficients on the second subset to estimate p-values. Repeat this process to obtain average p-values for the coefficients.

Frequency

An empirical correlation matrix was computed from the 2000×256 pixel matrix, followed by the extraction of eigenvectors and eigenvalues. The dataset X was then transformed using the eigenvectors corresponding to positive eigenvalues. A Lasso regression model was fitted on the transformed data to predict `group_ind`. The significance of the model coefficients was assessed using p-values obtained from 1000 permutation tests.