A Bayesian General Linear Modeling Approach to Cortical Surface fMRI Data Analysis: **Appendix**

Appendix A: Proof of Equation (10)

$$
\begin{aligned}
\pi(\boldsymbol{y} \mid \boldsymbol{\theta}) &= \int \pi(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_m \mid \boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_m, \boldsymbol{\theta}) \pi(\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_m \mid \boldsymbol{\theta}) d\boldsymbol{\beta}_1 \cdots \boldsymbol{\beta}_m \\
&= \int \prod_{m=1}^{M} \pi(\boldsymbol{y}_m \mid \boldsymbol{\beta}_m, \boldsymbol{\theta}) \pi(\boldsymbol{\beta}_m \mid \boldsymbol{\theta}) d\boldsymbol{\beta}_1 \cdots \boldsymbol{\beta}_m \\
&= \prod_{m=1}^{M} \int \pi(\boldsymbol{y}_m \mid \boldsymbol{\beta}_m, \boldsymbol{\theta}) \pi(\boldsymbol{\beta}_m \mid \boldsymbol{\theta}) d\boldsymbol{\beta}_m \\
&= \prod_{m=1}^{M} \pi(\boldsymbol{y}_m \mid \boldsymbol{\theta}) \\
&\propto \prod_{m=1}^{M} \pi(\boldsymbol{\theta} \mid \boldsymbol{y}_m) \pi(\boldsymbol{\theta})^{-1} \\
&= \pi(\boldsymbol{\theta})^{-M} \prod_{m=1}^{M} \pi(\boldsymbol{\theta} \mid \boldsymbol{y}_m)
\end{aligned}
$$

Appendix B: Simulation Figures

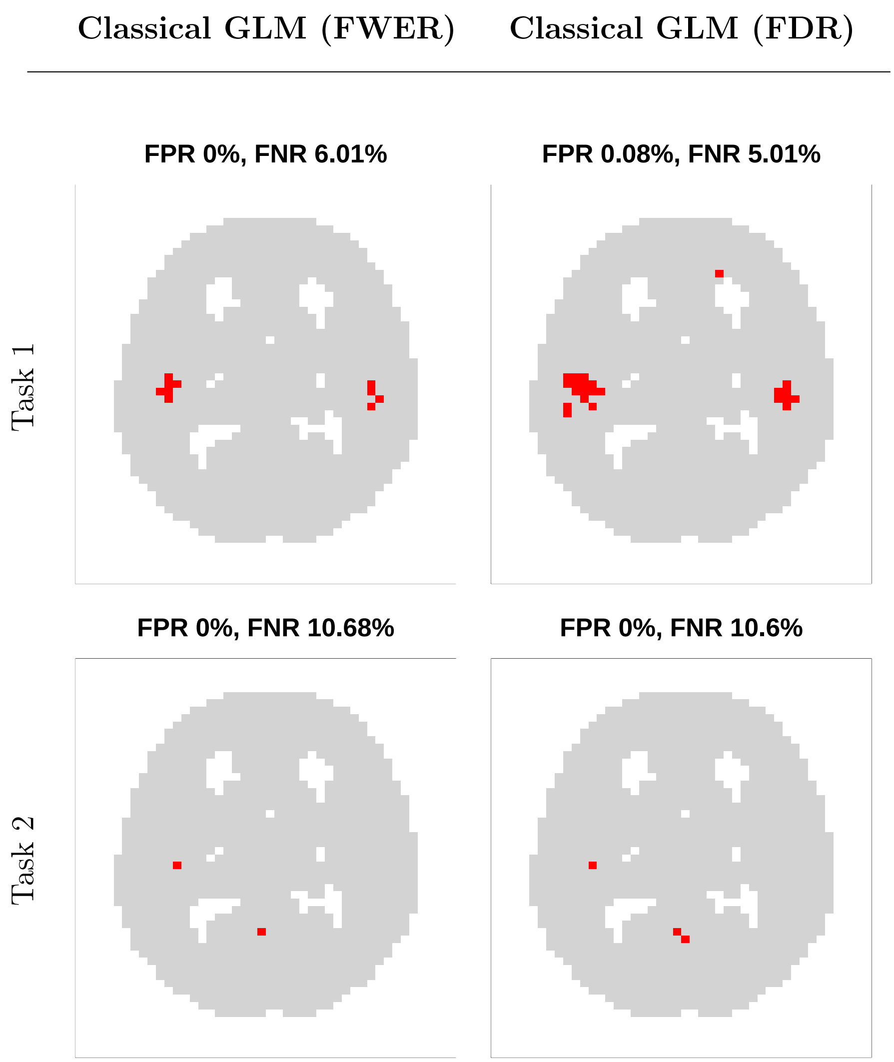

Figure B1: Estimated regions of activation in the simulation study using the classical GLM, based on unsmoothed data. The false positive rate (FPR) and false negative rate (FNR) are reported for each method and activation. Regions of activation are estimated by performing a hypothesis test on the task coefficient at each location, correcting for multiple comparisons through FDR control ($q = 0.01$) and FWER control ($\alpha = 0.01$).

Appendix C: Additional Results, Motor Task Study

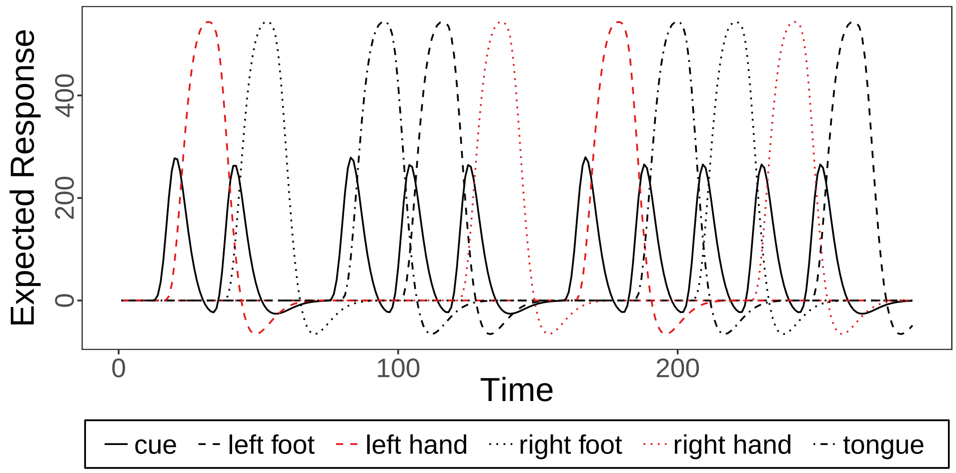

Figure C1: Activation profiles for the visual cue and five motor tasks in the motor study.

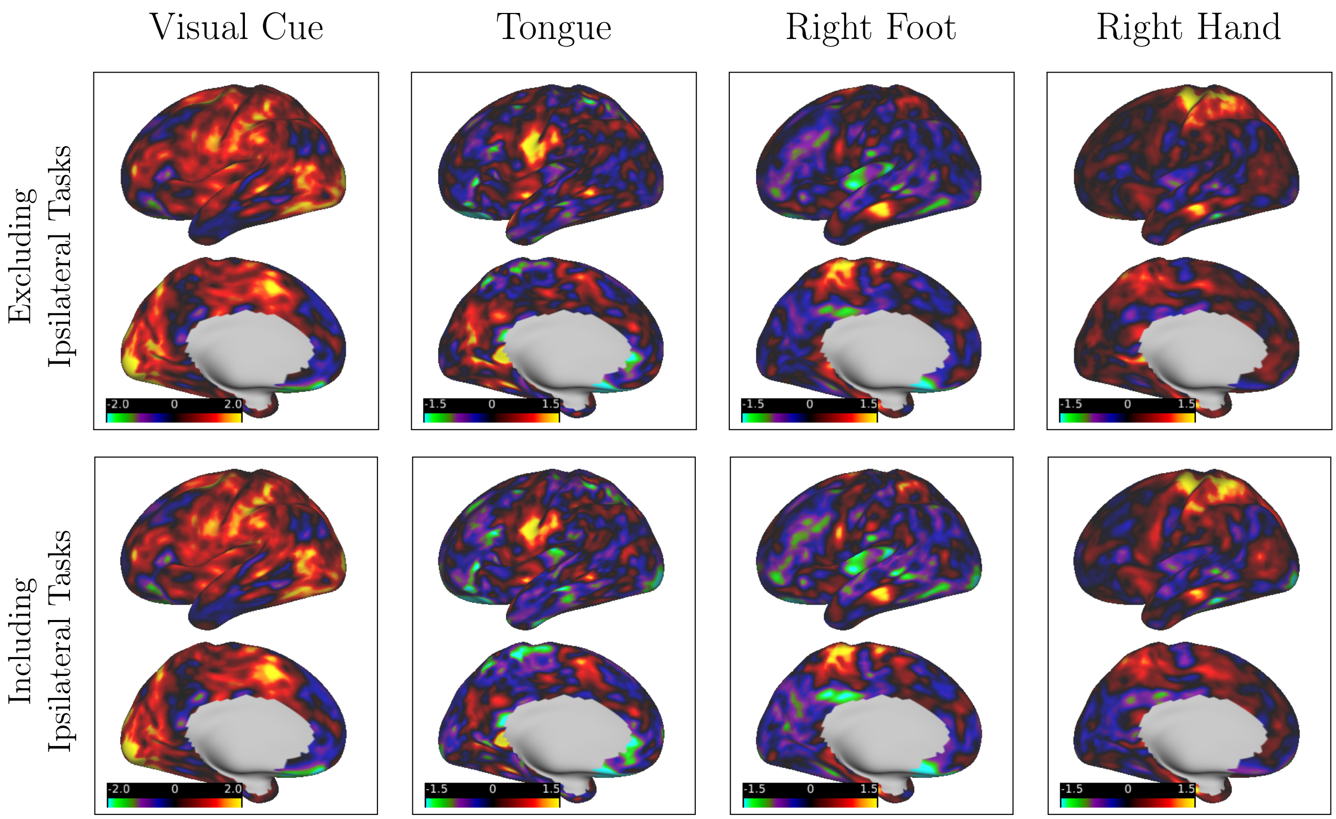

Figure C2: Subject-level estimates of activation amplitude for the left hemisphere Bayesian model, with and without ipsilateral tasks (left finger tapping, left toe wiggling) included in model. For the results of the motor study presented in the paper, we excluded ipsilateral tasks for the model within each hemisphere since activation due to lateral tasks is primarily expected on the contralateral side of the brain. Excluding these two tasks dramatically reduces the computational burden of the model for each hemisphere (see Computation Times section in main text) and, as shown here, produces very similar activation estimates compared with the "full" model.

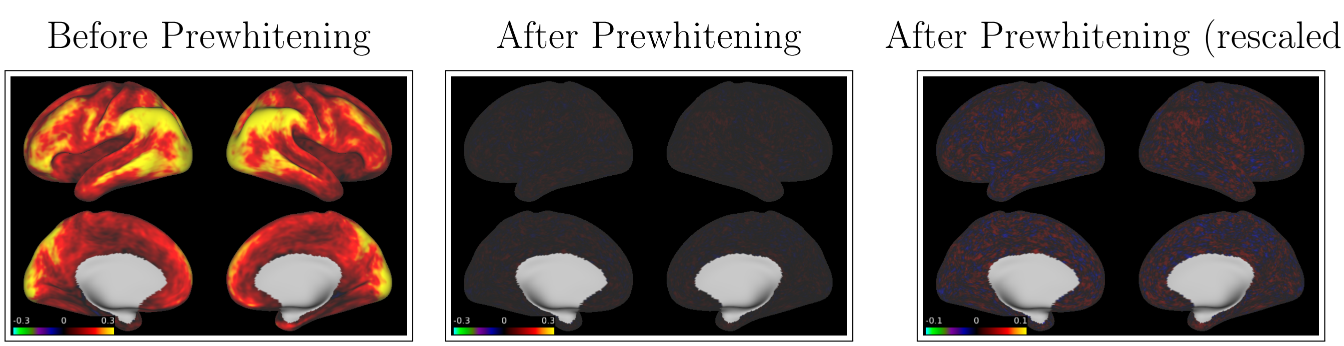

Figure C3: First AR coefficient of the residuals of a classical GLM, before and after prewhitening. Even without multiple iterations, the temporal autocorrelation of the residuals appears to be negligible after prewhitening.

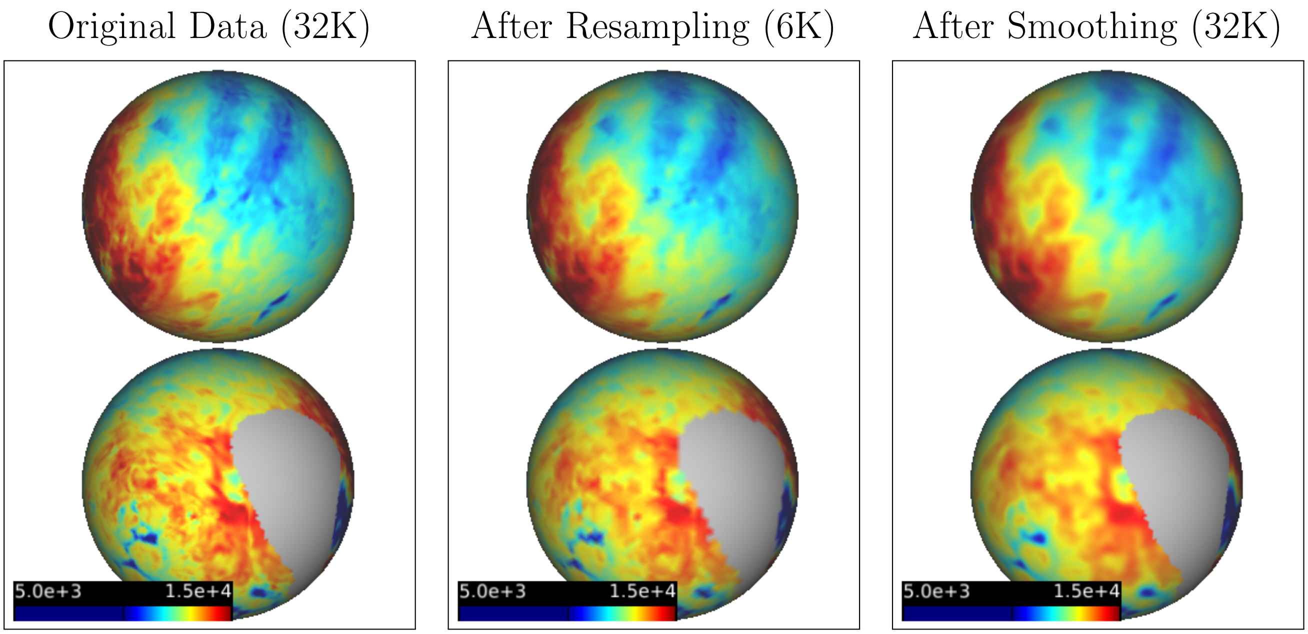

Figure C4: For a single volume of one subject, fMRI BOLD values at the original 32K resolution, after resampling to 6K, and after smoothing the original data with a 6mm FWHM Gaussian kernel. The values are displayed on the spherical surface of the left hemisphere. The data is substantially more blurred due to smoothing versus resampling.

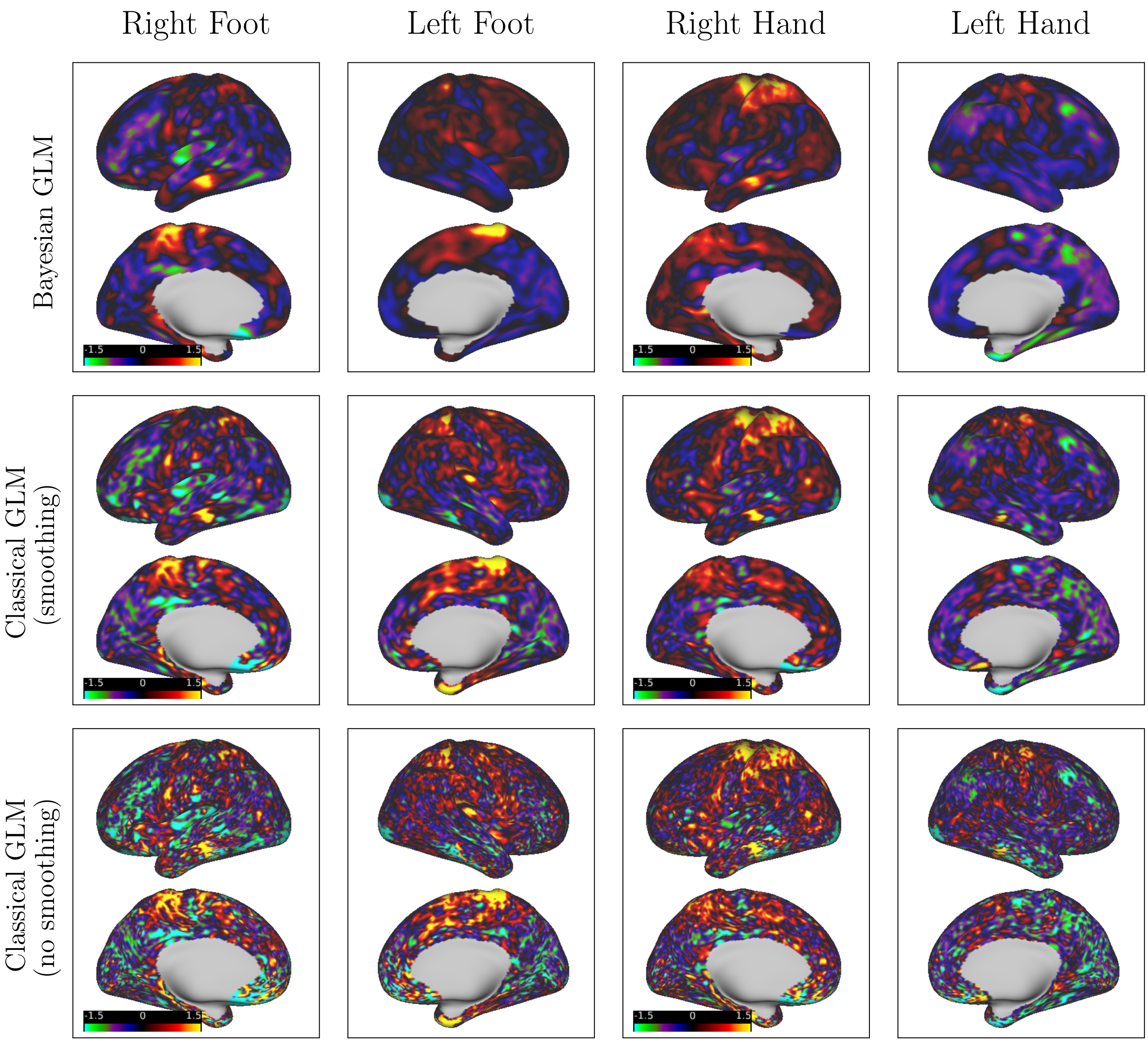

Figure C5: For one randomly selected subject, estimates of activation amplitude for the right foot, left foot, right hand and left hand tasks of the motor study, based on the classical and Bayesian approaches.

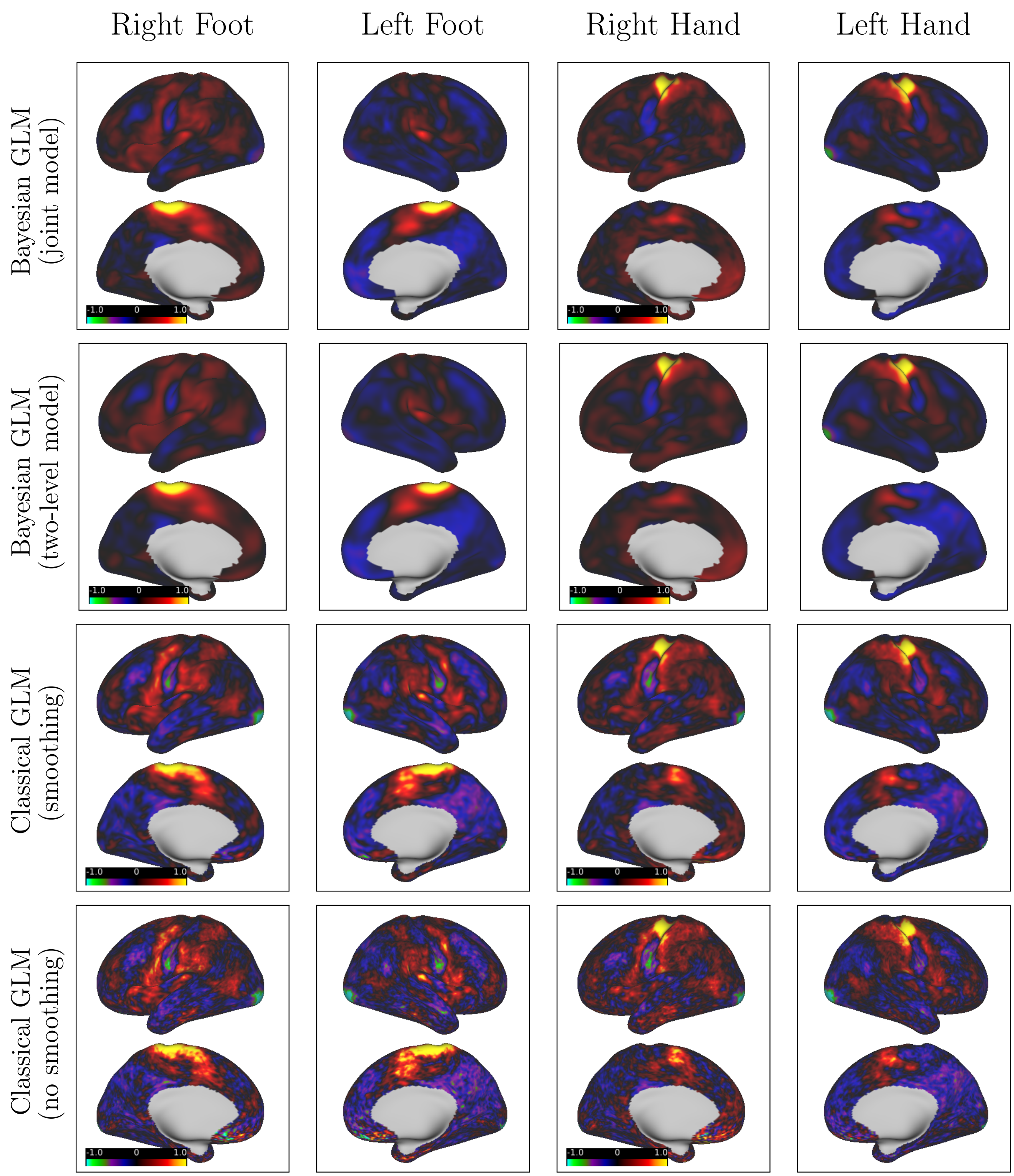

Figure C6: Group-level estimates of activation amplitude for the right foot, left foot, right hand and left hand tasks of the motor study, based on the classical and Bayesian approaches. Results for both the joint and two-level Bayesian multi-subject modeling approaches are displayed. As expected, the two-level approach tends to result in oversmoothed activation estimates.

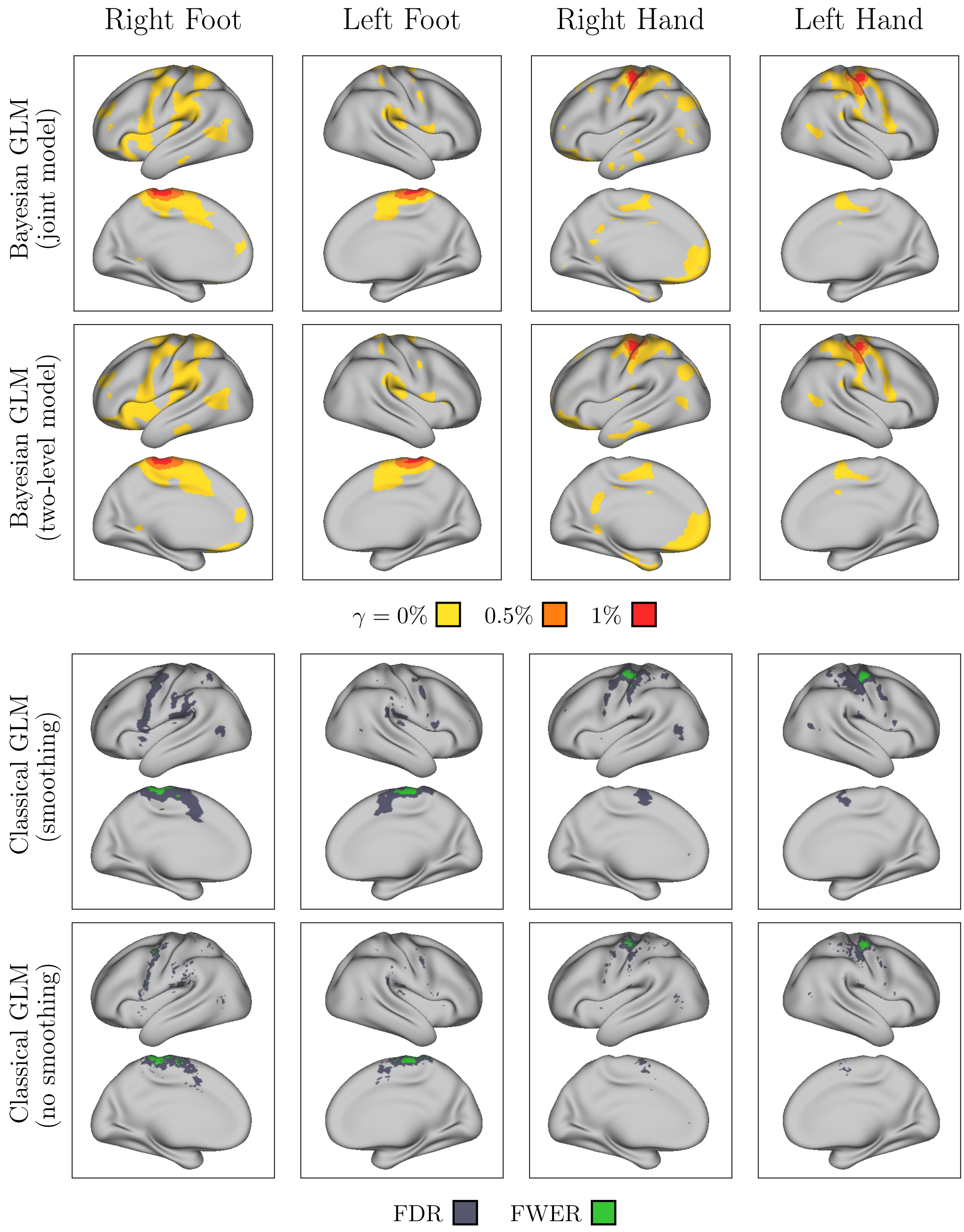

Figure C7: Group-level regions of activation for the right foot, left foot, right hand and left hand tasks of the motor study at significance level 0.01, based on the classical and Bayesian approaches.

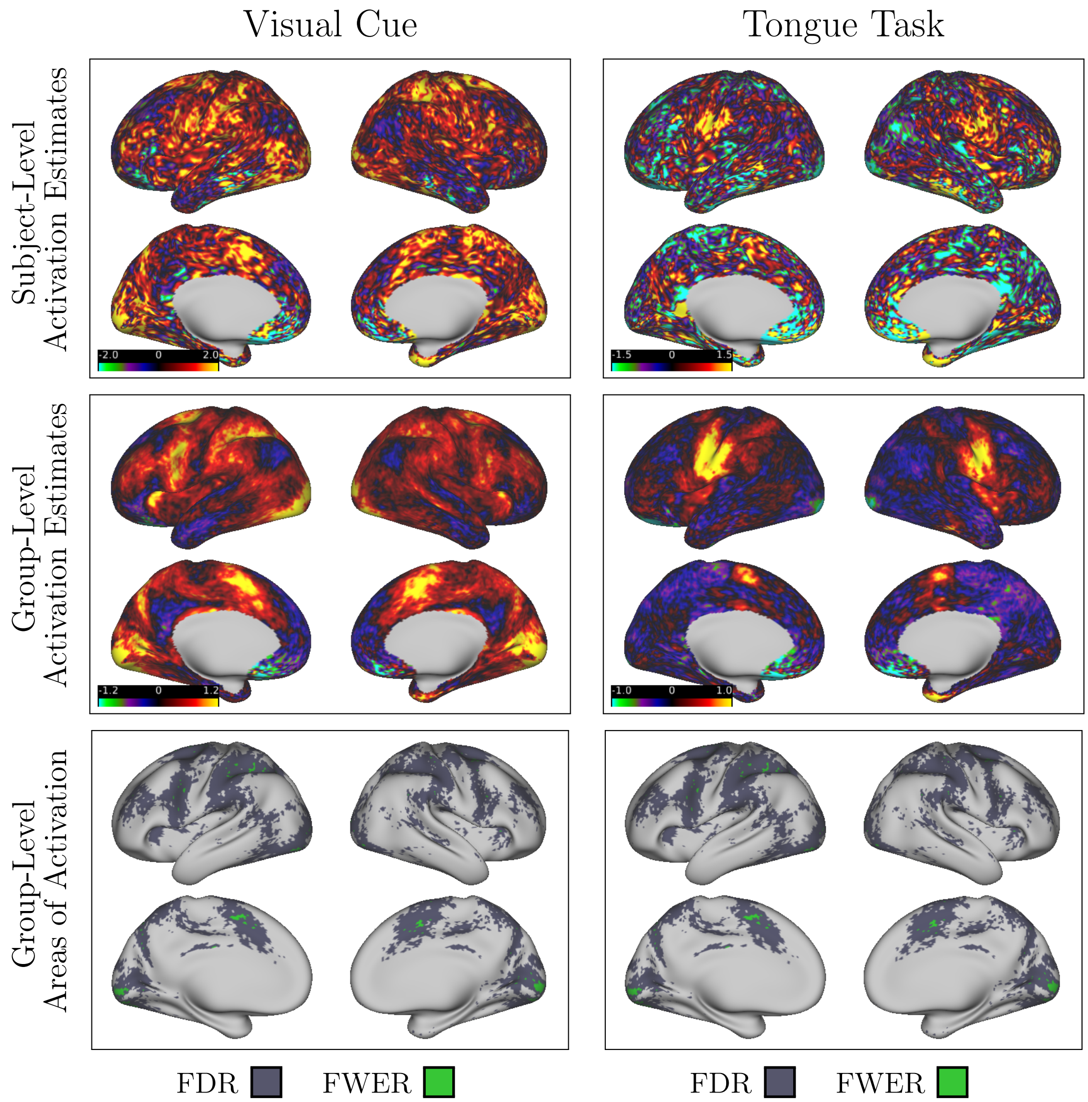

Figure C8: Subject-level estimates of activation amplitude for the visual cue and tongue tasks of the motor study, based on the classical GLM using unsmoothed data.

Appendix D: Results, Gambling Task

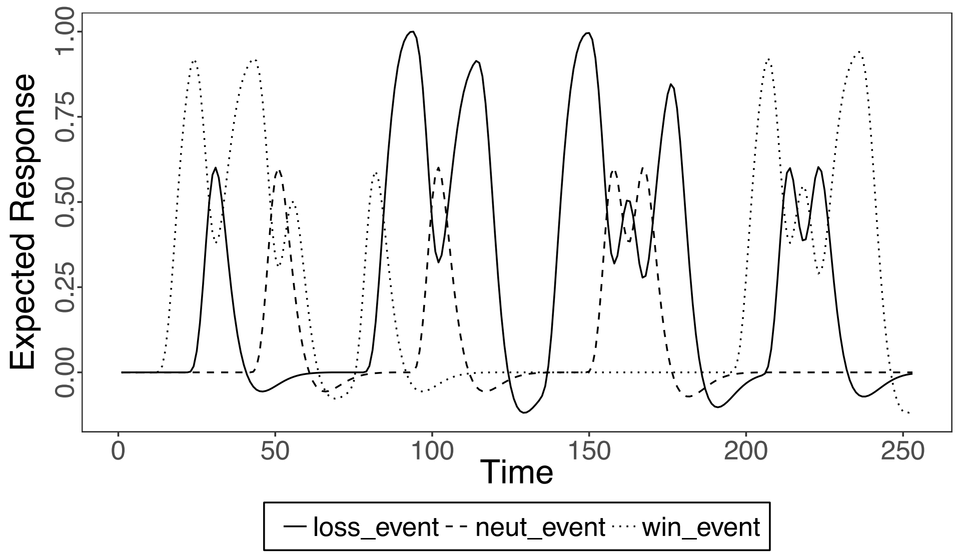

Figure D1: Activation profiles for the gambling task study.

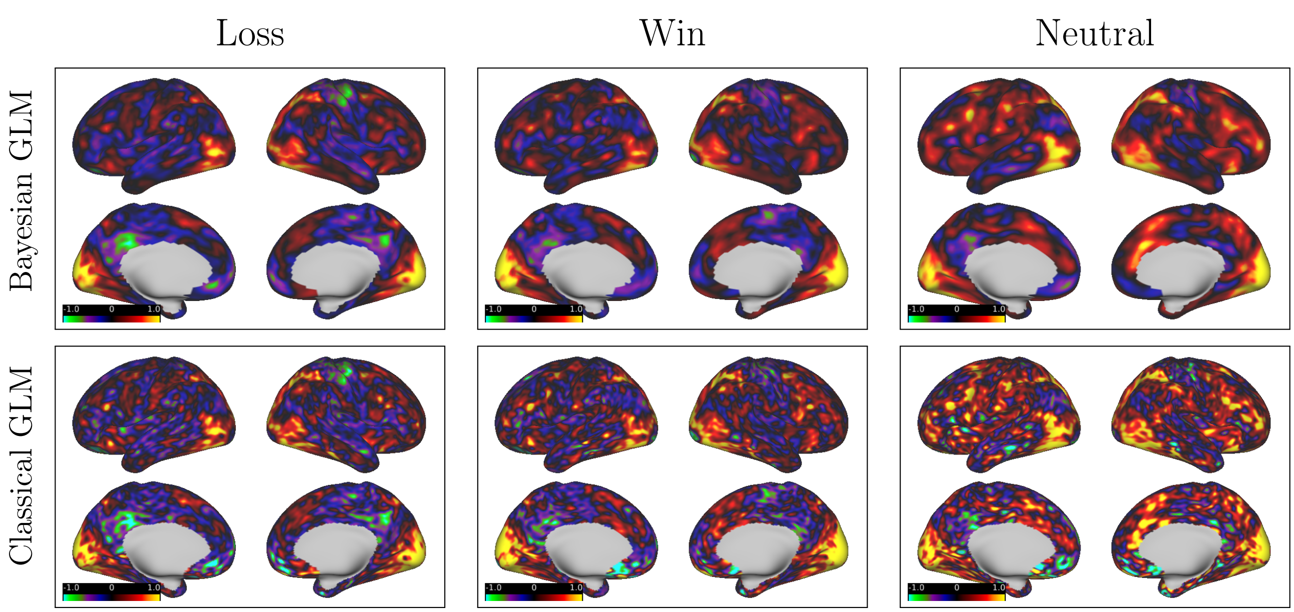

Figure D2: Subject-level estimates of activation amplitude for loss, win and neutral trials in the gambling study using the classical and Bayesian approaches. Classical GLM results are based on smoothed data.

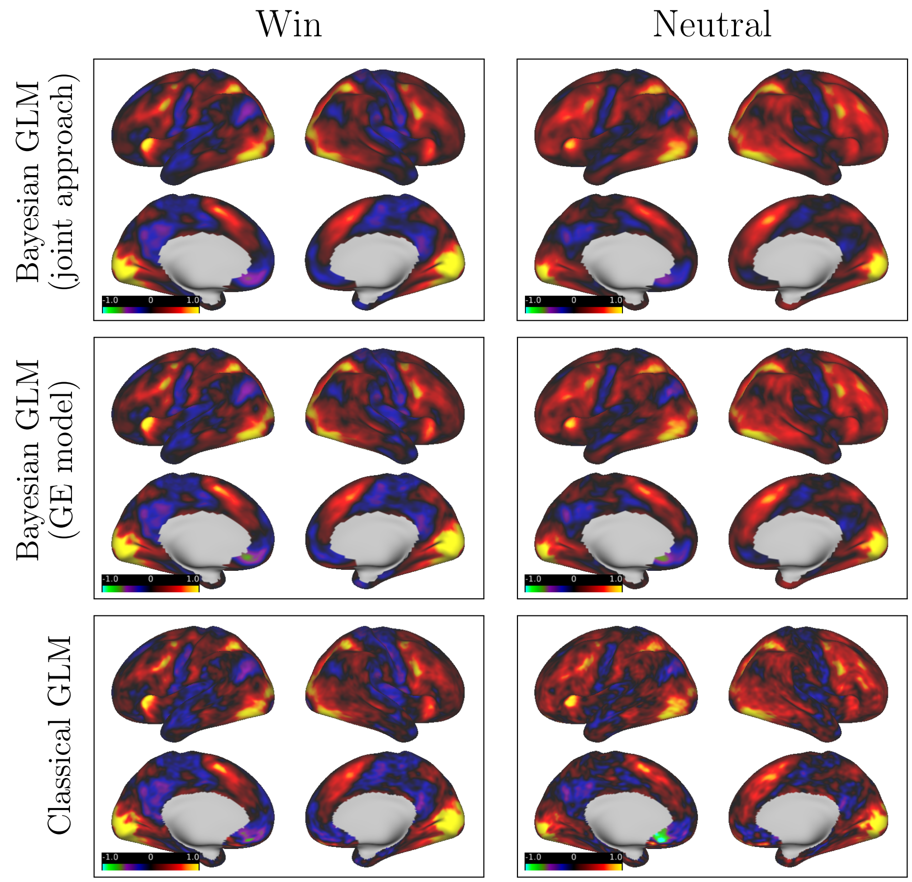

Figure D3: Group-level estimates of activation amplitude for win and neutral trials in the gambling study using the classical and joint Bayesian approaches. All three trial types activate expected areas, including the visual cortex and the insula, which is known to be involved in emotion processing. The win trial has somewhat higher estimates of activation in both areas. Classical GLM results are based on smoothed data.

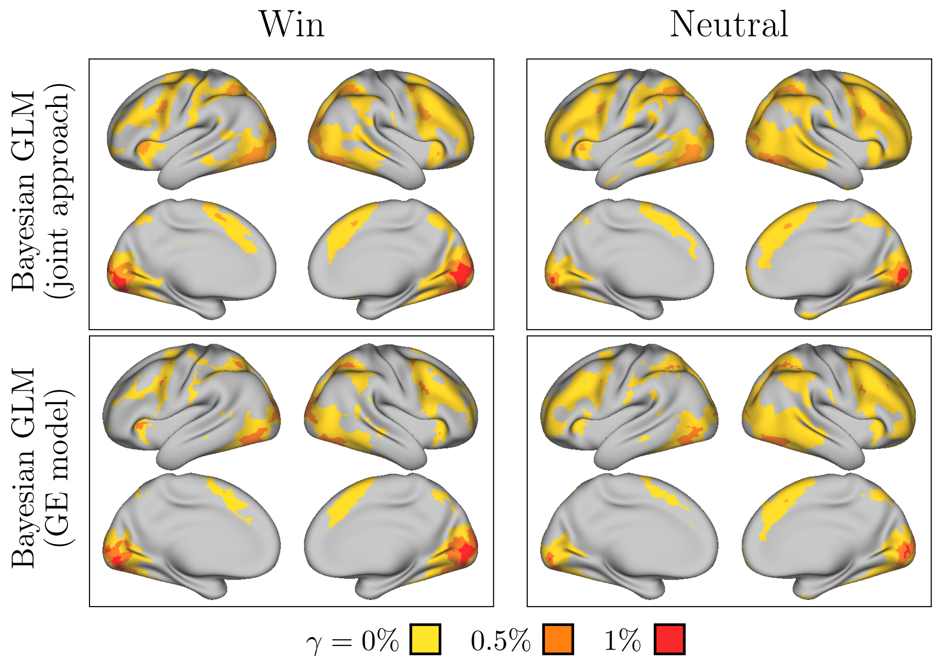

Figure D4: Group-level areas of activation for win and neutral trials in the gambling study using the joint approach and fully Bayesian GE model. Significance level is $\alpha = 0.01$.

Appendix E: Computation Times

The single-subject Bayesian GLM for the motor task study, which included 4 tasks and hence 9 hyperparameters in the model for each hemisphere (two hyperparameters per task, plus one hyperparameter for the residual precision assuming independent errors), required approximately 45 minutes of computation time and 50 gigabytes (GB) of memory per subject and hemisphere. The model including all 6 tasks, resulting in 13 hyperparameters in the model for each hemisphere, required 2 hours of computation time per hemisphere. The gambling task study, which included 3 tasks and therefore 7 hyperparameters, required 30 minutes and 40 GB of memory on average per subject and hemisphere. The difference in computation times for these three models demonstrates the dramatic impact of additional hyperparameters on the computational burden of INLA. Computation time for all models was greatly improved by using the `PARDISO` parallel matrix sparse matrix library. Choosing good starting values for the hyperparameters[1] also substantially improved computation time.

The proposed joint approach for multi-subject analysis, parallelized across samples $\ell = 1, \ldots, L = 50$, required approximately 30 minutes of computation time for the motor task study, including identification of areas of activation for three activation thresholds. Without parallelization, the computation time would have been approximately 6 hours. For the gambling task study, the total computation time using the joint approach was similar at 25 minutes. The two-level group Bayesian approach was more time-consuming due to the need to fit the model for each task in each hemisphere using INLA. Each task required an average of approximately 5 minutes for model fitting and 4-6 minutes for identifying areas of activation at each activation threshold, for a total of approximately 20 minutes per task per hemisphere. The fully Bayesian multi-subject group effect (GE) model for the gambling task study with all 20 subjects required approximately 11 hours of computation time and 270 GB of memory per hemisphere for model fitting, followed by an additional 1 hour for excursion set estimation for each task and activation threshold. The growth in computation time and memory requirements per additional subject was approximately linear for the GE model.

Finally, the single-subject classical GLM took 3h 39m of computation time and required 1.39 GB of memory on average for the motor task study. For the gambling task study, the computation time was 2h 40m on average per subject with similar memory requirements. The relatively high computation time was due to the vertex-wise prewhitening approach, which results in a unique design matrix at each vertex and thus requires looping over all vertices. Refitting the model at each location for each subject using 100 permutations of the prewhitened time points required approximately 1 hour of additional computation time per subject and 1.1 GB of storage space. Performing FWER correction required an additional 3.5 hours of computation time at the group level. Performing FDR correction, by contrast, took less than 1 second at the group level and no additional computation time at the subject level.

[1] We chose -2 for $\log(\tau)$ and 2 for $\log(\kappa)$.