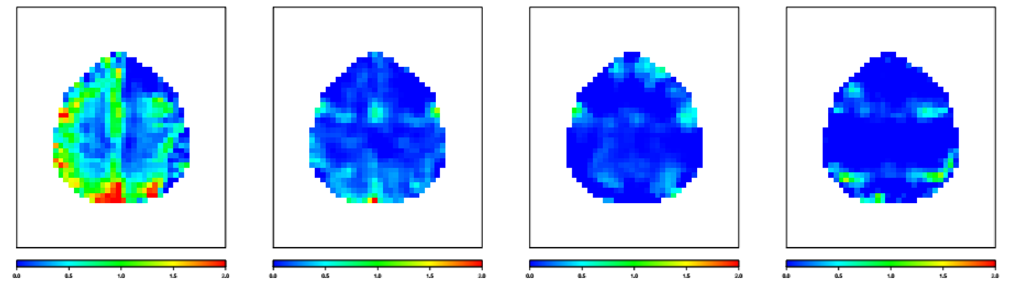
# Regularized Aggregation of Statistical Parametric Maps


Jongik Chung[1], Li-Yu Wang[1], Cheolwoo Park[1], Hosik Choi[2], Amanda L. Rodrigue[1], Jordan E. Pierce[1], Brett A. Clementz[1], and Jennifer E. McDowell[1]

The University of Georgia[1] and Kyonggi University[2]


## Statistical Parametric Maps

In a **Statistical Parametric Map (SPM)**, each voxel is indexing how closely associated the BOLD signal is with the task of interest.



(a) Subject 10 (b) Subject 7 (c) Subject 14 (d) Subject 30

Group analyses in fMRI studies involve combining individual SPMs. Our main question is **how to combine** these individual maps to produce accurate and robust group maps.

## Regularized Unsupervised Learning

To aggregate the estimated regression coefficients at voxel $v$ for $N$ participants is:

$$\hat{\beta}_v = \sum_{i=1}^{N} w_i \hat{\beta}_{vi}$$

where $\hat{\beta}_{vi}$ is the estimated regression coefficient at voxel $v$ for the $i$ th subject.

To determine $\boldsymbol{w}$, we consider the following objective function:

$$\min_{\mathbf{y},\mathbf{w}} L(\mathbf{y},\mathbf{w}) = \min_{\mathbf{y},\mathbf{w}} \left\{ \|\mathbf{y}-\mathbf{X}\mathbf{w}\|_2^2 + \lambda \|\mathbf{D}\mathbf{w}\|_\gamma^\gamma \right\},$$

$$\text{s.t} \sum_{i=1}^{N} w_i = 1, w_i \geq 0, i = 1, \ldots, N$$

We focus on the ridge-type penalty with $\gamma = 2$ and $\boldsymbol{D} = \boldsymbol{I}$.

## Bootstrap-based weighted $t$ test

We develop a **bootstrap-based weighted $t$ test** that utilizes the estimated weights. The weights are determined by neighborhood of the voxel being tested. We use parametric bootstrap sampling to construct the null distribution.

$$T_v^w = \frac{\hat{\beta}_v^w}{S_v^w/\sqrt{N}} \quad \hat{\beta}_v^w = \sum_{i=1}^{N} w_i \hat{\beta}_i \quad S_v^w = \sqrt{\sum_{i=1}^{N} w_i (\hat{\beta}_{vi} - \hat{\beta}_v^w)^2}$$

$$2 \times \min \left\{ \frac{1}{B} \sum_{l=1}^{B} I_{(t_0^{(l)} \geq t_v^w)}, \frac{1}{B} \sum_{i=1}^{B} I_{(t_0^{(i)} \leq t_v^w)} \right\}$$

## Simulation Study

Based on 20 subjects we repeat each simulation 100 times. Also we compare the results from the proposed approach (RA) with those from the random effect (RE) model.

In the **homogeneous case**, the RA and RE successfully produces weights close to the simple average weight (0.05), while RA estimates the weights more stably.

| | RA | RE | | RA | RE | | RA | RE | | RA | RE |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $w_1$ | 0.050 | 0.050 | $w_6$ | 0.050 | 0.050 | $w_{11}$ | 0.050 | 0.051 | $w_{16}$ | 0.050 | 0.049 |
| | (0.004) | (0.017) | | (0.005) | (0.015) | | (0.005) | (0.015) | | (0.005) | (0.013) |
| $w_2$ | 0.050 | 0.047 | $w_7$ | 0.050 | 0.050 | $w_{12}$ | 0.050 | 0.051 | $w_{17}$ | 0.051 | 0.053 |
| | (0.004) | (0.014) | | (0.004) | (0.014) | | (0.004) | (0.015) | | (0.004) | (0.017) |
| $w_3$ | 0.050 | 0.050 | $w_8$ | 0.050 | 0.050 | $w_{13}$ | 0.050 | 0.047 | $w_{18}$ | 0.050 | 0.048 |
| | (0.004) | (0.014) | | (0.004) | (0.015) | | (0.004) | (0.014) | | (0.004) | (0.014) |
| $w_4$ | 0.050 | 0.049 | $w_9$ | 0.050 | 0.048 | $w_{14}$ | 0.050 | 0.049 | $w_{19}$ | 0.049 | 0.051 |
| | (0.003) | (0.014) | | (0.004) | (0.014) | | (0.004) | (0.016) | | (0.004) | (0.017) |
| $w_5$ | 0.050 | 0.052 | $w_{10}$ | 0.050 | 0.054 | $w_{15}$ | 0.050 | 0.052 | $w_{20}$ | 0.050 | 0.050 |
| | (0.005) | (0.015) | | (0.003) | (0.015) | | (0.003) | (0.015) | | (0.005) | (0.015) |

Homogeneous Case

In the **three outlying case**, both RA and RE detect the difference among outlying subjects and produce small weights (RA=0.015-0.016 and RE=0.001) for such subjects.

| | Three | | | Three | |
|---|---|---|---|---|---|
| | RA | RE | | RA | RE |
| $w_1$ | 0.057 | 0.058 | $w_{11}$ | 0.058 | 0.060 |
| | (0.008) | (0.020) | | (0.009) | (0.017) |
| $w_2$ | 0.056 | 0.056 | $w_{12}$ | 0.057 | 0.060 |
| | (0.009) | (0.016) | | (0.009) | (0.017) |
| $w_3$ | 0.054 | 0.059 | $w_{13}$ | 0.058 | 0.055 |
| | (0.008) | (0.016) | | (0.008) | (0.016) |
| $w_4$ | 0.057 | 0.058 | $w_{14}$ | 0.055 | 0.058 |
| | (0.009) | (0.017) | | (0.007) | (0.018) |
| $w_5$ | 0.055 | 0.060 | $w_{15}$ | 0.055 | 0.061 |
| | (0.008) | (0.018) | | (0.008) | (0.017) |
| $w_6$ | 0.057 | 0.059 | $w_{16}$ | 0.056 | 0.058 |
| | (0.008) | (0.018) | | (0.008) | (0.016) |
| $w_7$ | 0.055 | 0.058 | $w_{17}$ | 0.056 | 0.062 |
| | (0.008) | (0.016) | | (0.008) | (0.020) |
| $w_8$ | 0.056 | 0.058 | $w_{18}$ | 0.016 | 0.001 |
| | (0.008) | (0.018) | | (0.007) | (0.000) |
| $w_9$ | 0.056 | 0.056 | $w_{19}$ | 0.015 | 0.001 |
| | (0.009) | (0.016) | | (0.007) | (0.000) |
| $w_{10}$ | 0.056 | 0.063 | $w_{20}$ | 0.016 | 0.001 |
| | (0.010) | (0.017) | | (0.006) | (0.000) |

Three outlying Case

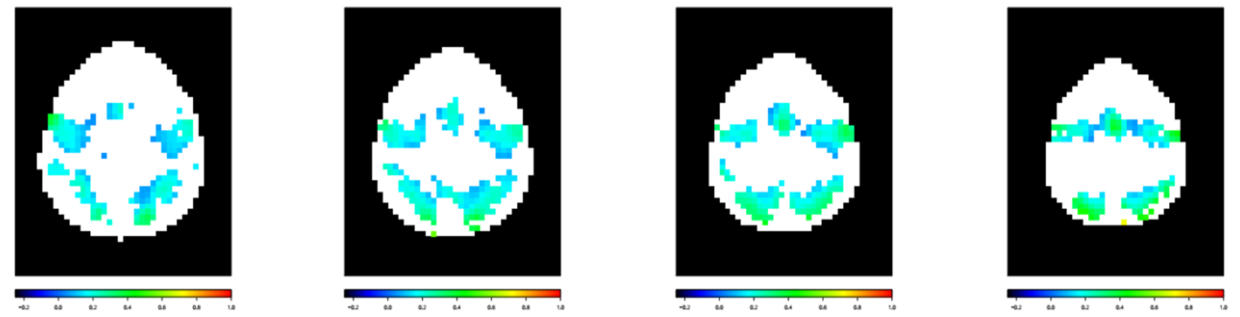
In the **multiple clusters,** 20 subjects are divided into three clusters with 10, 7, 3 subjects. The RA assigns different weights on different clusters, while the RE mainly assigns the weights only to the first cluster (0.094–0.107). This results in producing a map that only reflects the majority cluster.

| | RA | RE | | RA | RE | | RA | RE | | RA | RE |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $w_1$ | 0.071 | 0.099 | $w_6$ | 0.074 | 0.099 | $w_{11}$ | 0.036 | 0.001 | $w_{16}$ | 0.035 | 0.001 |
| | (0.009) | (0.033) | | (0.011) | (0.029) | | (0.016) | (0.000) | | (0.014) | (0.000) |
| $w_2$ | 0.070 | 0.094 | $w_7$ | 0.072 | 0.099 | $w_{12}$ | 0.037 | 0.001 | $w_{17}$ | 0.036 | 0.001 |
| | (0.012) | (0.027) | | (0.012) | (0.026) | | (0.016) | (0.000) | | (0.016) | (0.000) |
| $w_3$ | 0.070 | 0.099 | $w_8$ | 0.073 | 0.099 | $w_{13}$ | 0.036 | 0.001 | $w_{18}$ | 0.008 | 0.000 |
| | (0.010) | (0.027) | | (0.013) | (0.029) | | (0.015) | (0.000) | | (0.009) | (0.000) |
| $w_4$ | 0.070 | 0.098 | $w_9$ | 0.075 | 0.095 | $w_{14}$ | 0.036 | 0.001 | $w_{19}$ | 0.009 | 0.000 |
| | (0.011) | (0.028) | | (0.012) | (0.028) | | (0.016) | (0.000) | | (0.010) | (0.000) |
| $w_5$ | 0.071 | 0.102 | $w_{10}$ | 0.072 | 0.107 | $w_{15}$ | 0.039 | 0.001 | $w_{20}$ | 0.009 | 0.000 |
| | (0.012) | (0.029) | | (0.011) | (0.029) | | (0.016) | (0.000) | | (0.009) | (0.000) |

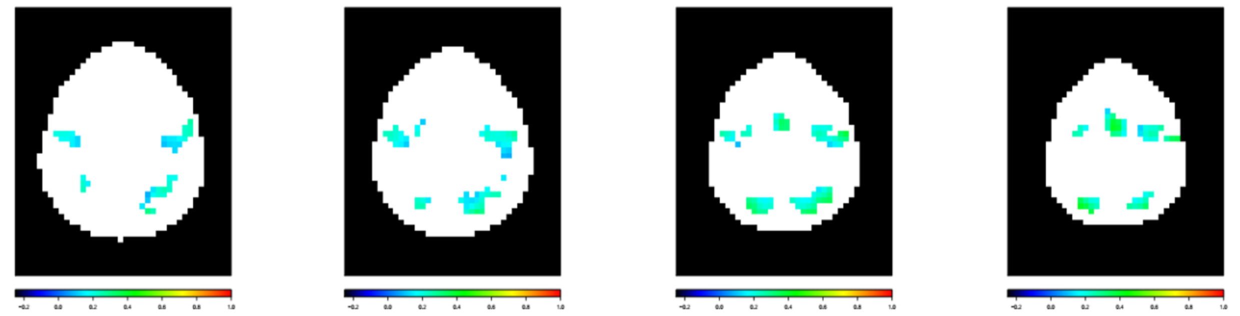
Multiple Clusters

## Real Data Analysis

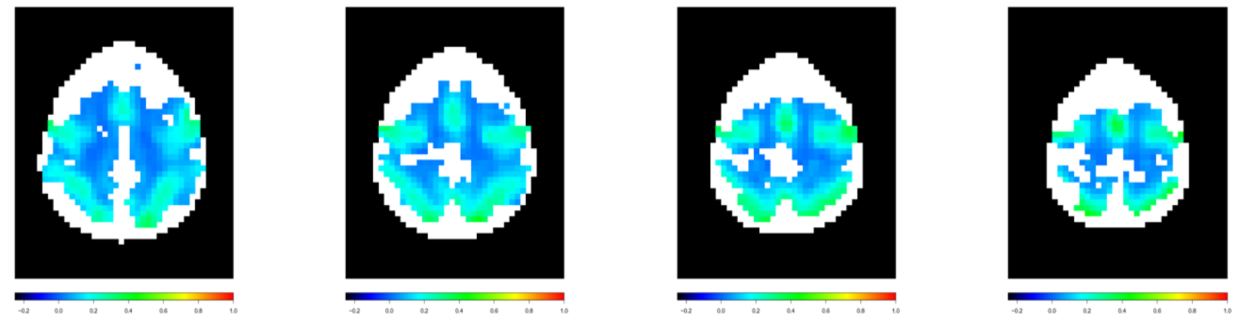
Real fMRI data were collected from healthy participants. Based on an antisaccade task, two groups ("high / low cognitive control group") were generated. We create a resampled group that consists of 17 subjects from the high cognitive control group and 3 from the low cognitive control group.



(a) RA (40) (b) RA (44) (c) RA (48) (d) RA (52)



(e) Classical t (40) (f) Classical t (44) (g) Classical t (48) (h) Classical t (52)



(i) RE (40) (j) RE (44) (k) RE (48) (l) RE (52)

The RA provides larger and well-formed clusters than the classical t test. On the contrary, the **RE** model shows too many activated regions.


## Acknowledgement

- National Science Foundation (NSF-1021RR193164)
- Basic Science Research Program (2017R1D1A1B05028565)
- National Institute of Mental Health (MH094172)