Connectivity-Informed fMRI Activation Detection

Bernard Ng[1], Rafeef Abugharbieh[1], Gael Varoquaux[2],
Jean Baptiste Poline[2], and Bertrand Thirion[2]

[1] Biomedical Signal and Image Computing Lab, UBC, Canada
[2] Parietal Team, Neurospin, INRIA Saclay-Ile-de-France, France
`bernardyng@gmail.com`

Abstract. A growing interest has emerged in studying the correlation structure of spontaneous and task-induced brain activity to elucidate the functional architecture of the brain. In particular, functional networks estimated from resting state (RS) data were shown to exhibit high resemblance to those evoked by stimuli. Motivated by these findings, we propose a novel generative model that integrates RS-connectivity and stimulus-evoked responses under a unified analytical framework. Our model permits exact closed-form solutions for both the posterior activation effect estimates and the model evidence. To learn RS networks, graphical LASSO and the oracle approximating shrinkage technique are deployed. On a cohort of 65 subjects, we demonstrate increased sensitivity in fMRI activation detection using our connectivity-informed model over the standard univariate approach. Our results thus provide further evidence for the presence of an intrinsic relationship between brain activity during rest and task, the exploitation of which enables higher detection power in task-driven studies.

Keywords: activation detection, connectivity prior, fMRI, resting-state.

1 Introduction

The standard approach for analyzing functional magnetic resonance imaging (fMRI) data involves comparing each brain voxel independently against an expected response to estimate the likelihood of activation [1]. However, accumulating evidence suggests that brain function is also mediated through the interactions between brain regions in what is referred to as functional connectivity [2]. Although incorporation of functional connectivity may provide activation models that better reflect the nature of brain activity, few existing methods have been designed for such purposes [3]. Instead, research efforts have largely focused on modeling the spatial structure of fMRI data through spatial regularization [4-7]. These methods indirectly account for local voxel interactions, but long-range interactions are completely neglected.

The discovery of structure in ongoing brain activity in the absence of external stimulus has ignited enormous research interest [8-10]. Many detected resting-state (RS) networks were found to exhibit high resemblance to those engaged during task performance [9]. These findings suggest potential relationships between brain activity during task and rest, which may serve as an additional source of information for detecting stimulus-evoked activation. In particular, task-based fMRI data typically

display rather low signal-to-noise ratio (SNR), especially in patients due to difficulties in performing certain tasks [10]. Since acquiring RS data requires minimal task demands and RS data are less susceptible to behavioural confounds [10], extracting connectivity priors from RS data to inform activation detection may enhance detection sensitivity, which is especially beneficial for studying diseased populations.

In this paper, we propose a novel generative model for integrating connectivity and task-evoked responses under a unified analytical framework. Assuming brain regions displaying functional correlations at rest are more likely to co-activate during task, incorporating RS-connectivity information should improve task activation detection. To learn the connectivity structure from RS data, we employ and compare graphical LASSO (GL) [11] and the oracle approximating shrinkage (OAS) technique [12]. The resulting connectivity information is then used as a prior on the task activation effects. Unlike most existing generative models [4-7] that employ either approximate inference or sampling methods for parameter estimation, our model has the distinct advantage of permitting exact closed-form solutions for both the posterior activation effect estimates and the model evidence. We apply our model on data from a cohort of 65 subjects undergoing a variety of experimental conditions and show increased sensitivity in detecting group activation over the standard univariate method.

2   Connectivity-Informed Activation Model

Motivated by recent studies that showed high resemblance between functional networks detected during rest and task [9], we propose to integrate RS-connectivity and task-evoked responses under a unified generative model. Specifically, let Y be an $d \times n$ matrix containing the intensity time courses of d voxels or d brain regions of a subject. Our proposed model can be summarized as follows:

$$Y \sim N(AX, V_1) = \frac{1}{\left|2\pi V_1\right|^{\frac{n}{2}}} \exp\left(-\frac{1}{2} tr\left((Y - AX)^T V_1^{-1} (Y - AX)\right)\right), \qquad (1)$$

$$A \sim MN(M, V_2, K) = \frac{\left|K\right|^{\frac{d}{2}}}{\left|2\pi \alpha^{-1} V_2\right|^{\frac{m}{2}}} \exp\left(-\frac{1}{2} \alpha \cdot tr\left((A - M)^T V_2^{-1} (A - M)K\right)\right), \qquad (2)$$

where X is an $m \times n$ matrix with m regressors along the rows for modeling stimulus-invoked response and confounds [1]. A is the $d \times m$ activation effect matrix to be estimated (Section 2.1). V_1 is the $d \times d$ covariance matrix of Y, V_2 is an $d \times d$ covariance matrix modeling the correlations between the activation effects of the d brain regions, and K is an $m \times m$ covariance matrix modeling the correlations between the experimental conditions. $MN(M, V_2, K)$ denotes the matrix normal distribution, which serves as the conjugate prior of (1) [13] with α controlling the degree of influence of this prior on A (Section 2.3). To ensure that our estimate of A is invariant to affine transformations on Y and X, we set M to $0_{d \times m}$ and K to XX^T [13]. Setting M to $0_{d \times m}$ also ensures that the activation effect estimates will not be biased towards non-zero values, which could induce false detections. V_1 and V_2 are assumed to be known. Compared to the model in [13] where V_1 and V_2 are assumed to be equal, we show that exact

closed-form solutions for the posterior estimate of A and the model evidence can be derived even for the more general case with V_1 and V_2 being distinct. Permitting distinct V_1 and V_2 accounts for how Y and A might have different correlation structures. Also, we hypothesize that brain regions functionally correlated at rest are more likely to co-activate during task performance. We thus set V_2 to the covariance estimates learned from RS data (Section 2.2). We assume $V_1 = I_{d \times d}$ as conventionally done for analytical simplicity, and defer learning V_1 from data for future work.

2.1   Posterior Activation Effects Estimation

To estimate A, we first derive the joint distribution of Y and A by taking the product of (1) and (2) and isolating the terms involving A from those involving Y:

$$
\begin{aligned}
P(Y, A \mid X, V_1, V_2) &\propto \exp\left(-\frac{1}{2} tr\left((Y - AX)^T V_1^{-1}(Y - AX) + \alpha X^T A^T V_2^{-1} AX\right)\right) \\
&= \exp\left(-\frac{1}{2} tr\left((V_1^{-1} + \alpha V_2^{-1})\left(AXX^T A^T - 2(V_1^{-1} + \alpha V_2^{-1})^{-1} V_1^{-1} YX^T A^T\right) + V_1^{-1} YY^T\right)\right)
\end{aligned}
\tag{3}
$$

Since the terms involving A take a quadratic form, the posterior distribution of A is again a matrix normal distribution, as expected from the conjugacy of (1) and (2). By completing the square, the maximum a posteriori mean of A can be derived:

$$
M_A = \left(V_1^{-1} + \alpha V_2^{-1}\right)^{-1} V_1^{-1} YX^T \left(XX^T\right)^{-1} .
\tag{4}
$$

Computing M_A requires inverting V_1 and V_2, which is unstable if V_1 and V_2 are set as sample covariance estimated from data with more brain regions than time points. In this work, we assume $V_1 = I_{d \times d}$, and employ and compare two state-of-the-art techniques, namely GL and OAS, for obtaining a well-conditioned V_2.

2.2   Functional Connectivity Estimation

Graphical LASSO: Given a sample covariance matrix, S, computed from data that are assumed to follow a centered multivariate Gaussian distribution, one can estimate a well-conditioned sparse inverse covariance matrix, $\hat{\Lambda}$, by minimizing the negative log-likelihood of the data distribution over the space of positive definite (p.d.) matrices while imposing an l_1 penalty on $\hat{\Lambda}$ [11]:

$$
\hat{\Lambda} = \underset{\Lambda > 0}{\arg\min}\, tr(\Lambda S) - \log \det(\Lambda) + \lambda \|\Lambda\|_1 ,
\tag{5}
$$

where $\| \cdot \|_1$ is the element-wise l_1 norm and λ controls the level of sparsity. Enforcing sparsity simplifies interpretation, since $\hat{\Lambda}_{ij} = 0$ implies brain regions i and j are not connected. To optimize (5), we employ the Two-Metric Projection method [14].

OAS: Assume the data for estimating the ground truth covariance, Σ, is generated from a multivariate Gaussian distribution. The most well-conditioned covariance estimate of Σ is $F = tr(S)/d \cdot I_{d \times d}$ [12]. The idea of OAS is to shrink the ill-conditioned sample covariance, S, towards F so that a well-conditioned covariance estimate, $\hat{\Sigma}$, can be obtained. Specifically, OAS optimizes the following cost function [12]:

$$\hat{\rho} = \underset{\rho}{\arg\min}\ E\left[\left\|\hat{\Sigma} - \Sigma\right\|_F^2\right],\ s.t.\ \hat{\Sigma} = (1-\rho)S + \rho F\ , \tag{6}$$

where ρ controls the amount of shrinkage with the optimal value given by [12]:

$$\hat{\rho} = \frac{\left(1 - \frac{2}{d}\right)tr\left(S^2\right) + tr^2(S)}{\left(n + 1 - \frac{2}{d}\right)\left(tr\left(S^2\right) - \frac{tr^2(S)}{d}\right)}. \tag{7}$$

Thus, no parameter selection is required and the inverse covariance matrix can be obtained by inverting $\hat{\Sigma}$ with stable inversion guaranteed.

2.3   Hyper-Parameters Estimation

A critical hyper-parameter in our model is α, which controls the degree of influence of the connectivity prior on the activation effect estimates. To set α, we first derive the model evidence by integrating $P(Y,A|\alpha,V_1,V_2)$ over A:

$$P(Y\,|\,\alpha,V_1,V_2) = \frac{\left|\left(V_1^{-1} + \alpha V_2^{-1}\right)^{-1}\right|^{\frac{m}{2}}}{\left|2\pi V_1\right|^{\frac{n}{2}}\left|\alpha^{-1}V_2\right|^{\frac{m}{2}}} \exp\left(-\frac{1}{2}tr\left(V_1^{-1}\left(YY^T\,\ldots\right.\right.\right.$$
$$\left.\left.\left. - YX^T\left(XX^T\right)^{-1}XY^TV_1^{-1}\left(V_1^{-1} + \alpha V_2^{-1}\right)^{-1}\right)\right)\right) \tag{8}$$

Since V_2^{-1} estimated from GL or OAS is p.d., V_2^{-1} can be decomposed into $Q\Gamma Q^T$ where Q are the eigenvectors of V_2^{-1} and Γ contains the eigenvalues of V_2^{-1} along the diagonal. By exploiting this property of p.d. matrix and that V_1 is assumed to be $I_{d\times d}$ hence $V_1^{-1} = I_{d\times d} = QI_{d\times d}Q^T$, the log of (8) can be simplified into:

$$\ln P(Y\,|\,\alpha,V_1,V_2) = -\frac{m}{2}\sum_{i=1}^{d}\left(\ln(1 + \alpha\gamma_i) - \ln(\alpha\gamma_i) - \frac{1}{m}\frac{B_{ii}}{1 + \alpha\gamma_i}\right) + C\ , \tag{9}$$

$$B = Q^T YX^T\left(XX^T\right)^{-1}XYQ\ , \tag{10}$$

where γ_i is the ith eigenvalue of V_2^{-1}, and terms that do not depend on α, V_1, or V_2 are grouped into C. Since (9) is a single variable function of α, the optimal α at which (9) is maximized can be efficiently determined using any generic optimization routines.

To optimize the choice of λ for the case where V_2^{-1} is estimated using GL, we define a grid of λ values at which the percentage of non-zero elements in V_2^{-1} roughly ranges from 10% to 90%. For each V_2^{-1} associated with a given λ, we determine the optimal α at which (9) is maximized and compute the log model evidence. The λ associated with the largest log model evidence is then taken as the optimal λ.

3   Materials

Synthetic Data: We generated 300 synthetic datasets based on our proposed model. Each dataset consisted of 10 subjects. Each subject's data comprised 100 regions with the first 20 regions set to be mutually correlated and activated across all subjects. The next 20 regions were set to be mutually correlated but not activated to test if our model will falsely declare correlated regions as activated. The remaining regions were not correlated nor activated. To simulate this scenario, we generated 100 signal time courses, where the first 20 time courses were sine functions with Gaussian noise added. The next 20 time courses were cosine functions with Gaussian noise added. The remaining 60 time courses were simply Gaussian noise. The empirical covariance of these signal time courses was taken as the group covariance of the activation effects and the RS networks, Ω^g. A random p.d. matrix was added to Ω^g to introduce inter-subject variability into each subject's covariance matrix, Ω^i. The degree of variability was controlled by restricting the Kullback-Leibler divergence of $N(0,\Omega^g)$ and $N(0,\Omega^i)$ to be ~1.5. For each subject i, 100 RS time courses of length N_t were drawn from $N(0,\Omega^i)$. We set N_t to 25 to emulate the typical situation where the number of regions exceeds N_t. To generate task data, samples of A were drawn from (2) with V_2 set to Ω^i. To simulate activation, the means of the first 20 regions in (2) were set to a small positive number, δ, that depended on the SNR, δ^2/σ^2. Gaussian noise was added to δ to further introduce inter-subject variability. The resulting A were then used to draw samples of Y from (1) with V_1 set to $\sigma^2 I$ and X being boxcar functions convolved with the hemodynamic response function [1]. Three SNRs were tested: 0.25, 0.5, and 0.75, with 100 synthetic datasets generated at each SNR level.

Real Data: fMRI data were collected from 65 healthy subjects at multiple imaging centers. Each subject performed 10 language, computation, and sensorimotor tasks over a period of ~5 min (140 brain volumes) similar to those in [15]. RS data of ~7 min duration (187 brain volumes) were also collected. 3T scanners from multiple manufacturers were used for acquiring the data with TR = 2200 ms, TE = 30 ms, and flip angle = 75°. Standard preprocessing, including slice timing correction, motion correction, temporal detrending, and spatial normalization, were performed on the task-based data using the SPM8 software. Similar preprocessing was performed on the RS data except a band-pass filter with cutoff frequencies at 0.01 to 0.1 Hz was applied to isolate the signal of interest [8]. Signals from cerebrospinal fluid and white-matter voxels were regressed out from the gray-matter voxels.

To ensure stable sparse inverse covariance estimation using GL [11], we reduced the dimension of the data by grouping the voxels into 1000 parcels. Specifically, we concatenated the RS voxel time courses across subjects and applied the parcellation technique of [16] to generate a group parcellation map. Each subject's brain images (in normalized space) were then parcellated using the group parcel labels. The mean voxel time courses of each parcel from rest and task were taken as the input to our model. To account for scanner variability across imaging centers, we normalized the parcel time courses by subtracting the mean and dividing by the standard deviation.

4   Results and Discussion

***Validation*:** For validation, we compare the sensitivity of our model with connectivity prior estimated from OAS and GL against the ridge regression model (i.e. V_2 set to I) and the standard univariate model [1] in detecting group activation. We denote these models as OAS-CM, GL-CM, R-UM, and S-UM. Since the ground truth activated brain regions are unknown for real data, we employ the max-*t* permutation test [17] to enforce strict control on false positive rate (FPR) so that we can safely base our validation on the number of parcels detected. We note that for each permutation, the *entire* activation effect map of each subject is multiplied by 1 or -1 chosen at random. Hence, the spatial covariance structure of the activation effect maps is preserved. Also, we emphasize that our validation criterion is independent of the criterion used for optimizing the model parameters, which mitigates bias from being introduced.

***Synthetic Data*:** The mean receiver operating characteristics curves averaged over the synthetic datasets are shown in Fig. 1. At all FPR and SNR levels, OAS-CM and GL-CM achieved higher true positive rates (TPR) than R-UM and S-UM, thus confirming that given the activation effects are inherently correlated, our model can exploit this information to improve group activation detection. We note that there is less need for imposing a prior (which introduces a bias) at higher SNR since more signals are available to estimate A, hence the decrease in sensitivity for OAS-CM and GL-CM. Also, the higher sensitivity achieved by OAS-CM compared to GL-CM might be due to how GL tends to produce unstable covariance estimates for large-scale problems.
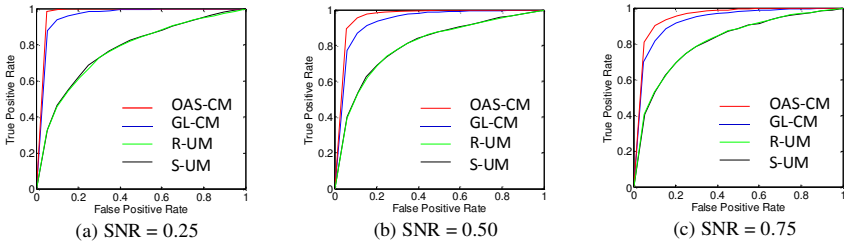

(a) SNR = 0.25   (b) SNR = 0.50   (c) SNR = 0.75

Fig. 1. Synthetic data results. OAS-CM (red) and GL-CM (blue) achieved higher TPR for all FPR and SNR levels than R-UM (green) and S-UM (black).

***Real Data*:** To test the generality of our model, we examined 21 contrasts between the 10 experimental conditions. Contrasts included computation vs. sentence processing task, auditory vs. visual task among others. Shown in Fig. 2(a) is the percentage of parcels detected with significant activation differences averaged over contrasts for *p*-value thresholds ranging from 0 to 0.5. Incorporating a connectivity prior using OAS-CM and GL-CM outperformed R-UM and S-UM even under the simplifying yet common assumption that $V_1 = I_{d \times d}$. We note that adding a shrinkage prior to control overfitting, as in the case of R-UM, only improved detection mildly. Thus, our results suggest an intrinsic relationship between the correlation structure of activation effects and RS-connectivity, and this relationship is consistent across subjects, hence the

improved group activation detection. Qualitatively, incorporating a RS-connectivity prior resulted in more detections of bilateral activation in brain regions implicated for the contrasts examined, examples of which are shown in Fig. 2(b) and (c).
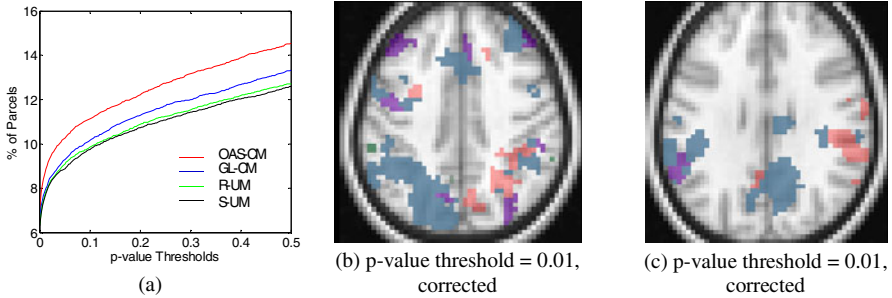

(a)

(b) p-value threshold = 0.01, corrected

(c) p-value threshold = 0.01, corrected

Fig. 2. Real data results. (a) % of parcels with significant activation differences averaged across contrasts vs. *p*-value thresholds. (b) Parcels detected by contrasting computation against sentence processing task, and (c) auditory against visual task. Red = detected by only OAS-CM. Purple = detected by both OAS-CM and GL-CM. Blue = detected by all methods.

5   Conclusions

We proposed a novel generative model for integrating connectivity and stimulus-induced response under a single analytical framework. Our model permits exact closed-form solutions for the posterior activation effect estimates and the model evidence without resorting to approximate inference or computationally-expensive sampling methods. On real data, we demonstrated that integrating a RS-connectivity prior improves sensitivity in detecting group activation. Our results thus support that the correlation structure of task activation effects is pertinent to RS connectivity, and this relationship is common across subjects. The flexibility of our model permits other V_1 and V_2 to be easily examined. In particular, integrating an anatomical connectivity prior estimated from diffusion MRI data would be a promising direction to explore. We expect that our model can similarly improve intra-subject activation detection, which enables more refined subject-specific analysis for studying patient populations.

References

1. Friston, K.J., Holmes, A.P., Worsley, K.J., Poline, J.B., Frith, C.D., Frackowiak, R.S.J.: Statistical Parametric Maps in Functional Imaging: A General Linear Approach. Hum. Brain Mapp. 2, 189–210 (1995)

2. Rogers, B.P., Morgan, V.L., Newton, A.T., Gore, J.C.: Assessing Functional Connectivity in the Human Brain by fMRI. Magn. Reson. Imaging 25, 1347–1357 (2007)

3. Ng, B., Hamarneh, G., Abugharbieh, R.: Detecting Brain Activation in fMRI Using Group Random Walker. In: Jiang, T., Navab, N., Pluim, J.P.W., Viergever, M.A. (eds.) MICCAI 2010. LNCS, vol. 6362, pp. 331–338. Springer, Heidelberg (2010)

4. Descombes, X., Kruggel, F., von Cramon, D.Y.: Spatio-Temporal fMRI Analysis Using Markov Random Fields. IEEE Trans. Med. Imaging 17, 1028–1039 (1998)

5. Penny, W.D., Trujillo-Barreto, N.J., Friston, K.J.: Bayesian fMRI Time Series Analysis with Spatial Priors. NeuroImage 24, 350–362 (2005)

6. Harrison, L.M., Penny, W.D., Asburner, J., Trujillo-Barreto, N.J., Friston, K.J.: Diffusion-based Spatial Priors for Imaging. NeuroImage 38, 677–695 (2007)

7. Woolrich, M.W., Jenkinson, M., Brady, J.M., Smith, S.M.: Fully Bayesian Spatio-temporal Modeling of fMRI Data. IEEE Trans. Med. Imaging 23, 213–231 (2004)

8. Fox, M.D., Raichle, M.E.: Spontaneous Fluctuations in Brain Activity Observed with Functional Magnetic Resonance Imaging. Nat. Rev. Neurosci. 8, 700–711 (2007)

9. Smith, S.M., Fox, P.T., Miller, K.L., Glahn, D.C., Fox, P.M., Mackay, C.E., Filippini, N., Watkins, K.E., Toro, R., Laird, A.R., Beckmann, C.F.: Correspondence of the Brain's Functional Architecture During Activation and Rest. Proc. Natl. Acad. Sci. 106, 13040–13045 (2009)

10. Fox, M.D., Greicius, M.: Clinical Applications of Resting State Functional Connectivity. Front. Syst. Neurosci. 4, 19 (2010)

11. Friedman, J., Hastie, T., Tibshirani, R.: Sparse Inverse Covariance Estimation with the Graphical LASSO. Biostats. 9, 432–441 (2008)

12. Chen, Y., Wiesel, A., Eldar, Y.C., Hero, A.O.: Shrinkage Algorithms for MMSE Covariance Estimation. IEEE Trans. Sig. Proc. 58, 5016–5029 (2010)

13. Minka, T.P.: Bayesian Linear Regression. Technical report, MIT Media Lab (2001)

14. Schmidt, M., Fung, G., Rosales, R.: Fast Optimization Methods for L1 Regularization: A Comparative Study and Two New Approaches. In: Kok, J.N., Koronacki, J., Lopez de Mantaras, R., Matwin, S., Mladenič, D., Skowron, A. (eds.) ECML 2007. LNCS (LNAI), vol. 4701, pp. 286–297. Springer, Heidelberg (2007)

15. Pinel, P., Thirion, B., Meriaux, S., Jober, A., Serres, J., Le Bihan, D., Poline, J.B., Dehaene, S.: Fast Reproducible Identification and Large-scale Databasing of Individual Functional Cognitive Networks. BioMed. Central Neurosci. 8, 91 (2007)

16. Thirion, B., Flandin, G., Pinel, P., Roche, A., Ciuciu, P., Poline, J.B.: Dealing with the Shortcomings of Spatial Normalization: Multi-subject Parcellation of fMRI Datasets. Hum. Brain Mapp. 27, 678–693 (2006)

17. Nichols, T., Hayasaka, S.: Controlling the Familywise Error Rate in Functional Neuroimaging: a Comparative Review. Stat. Methods Med. Research 12, 419–446 (2003)