

# The Power Analysis of Sparse Representation based on Laplace-Beltrami eigenfunctions

Seung-Goo Kim

Department of Brain and Cognitive Sciences,  
Seoul National University, Korea.

Term project for *Statistical Methods in NeuroImage Analysis*  
Instructor: Moo K. Chung

December 14, 2011

## 1 Introduction

In localizing a group difference or a covariate of a factor of interest on structure model in *in-vivo* human brains using magnetic resonance imaging (MRI), various approaches have been used. For approaches that analyze the local differences voxel-wisely, voxel-based morphometry (VBM), deformation-based morphometry (DBM) and tensor-based morphometry (TBM) have been proposed [1, 7, 2]. In these methods, the local measure of brain tissue for a voxel is quantified. In VBM, the probability of certain tissue such as gray matter is used as a gray matter density that quantifies local integrity of gray matter of a corresponding voxel. In TBM, the local volume is quantified as Jacobian determinant of deformation vector field that deforms the template volume to an individual volume. Then the difference of local measures at each voxel usually is tested by a general linear model (GLM) while the family-wise error rate is controlled by statistical techniques such as false discovery rate (FDR) or random field theory (RFT).

The methods have shown usefulness and sensitivity in detecting local difference in brain structures during past decade. However, due to limitations of the current MRI acquisition technique such as low resolution and contrasts, it is quite challenging to quantify subcortical structures and subfields automatically. To overcome this shortage, full or partial use of manual segmentation has

been utilized in analyzing subcortical structures.

There have been many approaches in quantifying local feature of subcortical structures in *in-vivo* human brains using magnetic resonance imaging (MRI). In Thompson *et al.*, in 2004 [20], hippocampus and temporal horn were manually delineated then the surfaces are parametrized. The distance between medial axis and boundary of the surface is measured as a local feature of the subcortical structure in order to characterize local expansion and contraction. In Qui *et al.* in 2008 [16], hippocampus first were segmented automatically using FreeSurfer then corrected by injecting a subcortical template based on manual segmentations [15]. Then the atrophy due to normal aging is quantified by normal surface momentum while the group difference was inferred on the Laplace-Beltrami (LB) eigenfunctions. In Chung *et al.*, in 2010 [5], the displacement vector field between the mean surface and the individual surfaces is used as a multivariate response variable. The amygdalar surfaces are modeled using spherical harmonics of manual segmentations.

We previously proposed sparse representation of displacement length on subcortical structures based on LB-eigenfunction as a way of increasing the signal to noise ratio in measurements [12]. For the term project for *Statistical Methods in Neuroimage Analysis*, Fall, 2011, I performed power analysis on the sparse representation method [12]. In this report, I will review the statistical power analysis on random field and compare the power of sparse representation method to that of a conventional least square method.

## 2 Images and preprocessing

The images and processing steps are explained in [12]. Here I briefly summarize the preprocessing in Figure 1. We have T1-weighted MRIs and manual segmentations of amygdala and hippocampus from 52 middle-age and elderly healthy adults (Figure 1a). The ages range 37 to 74 years (mean age =  $55.5 \pm 10.4$  years). There are 16 men and 36 women. We construct template surfaces by averaging normalized masks (Figure 1b). For normalization, advanced normalization tools (ANTS) is used [3]. During the non-linear normalization, we obtain displacement vector field that deforms template back to the individual image. A displacement vector field of a single subject is shown (Figure 1c). We interpolate this field onto vertices of the template surface (Figure 1d). We use the length ( $l_2$ -norm) of the vector field as a local anatomical measure in this study.

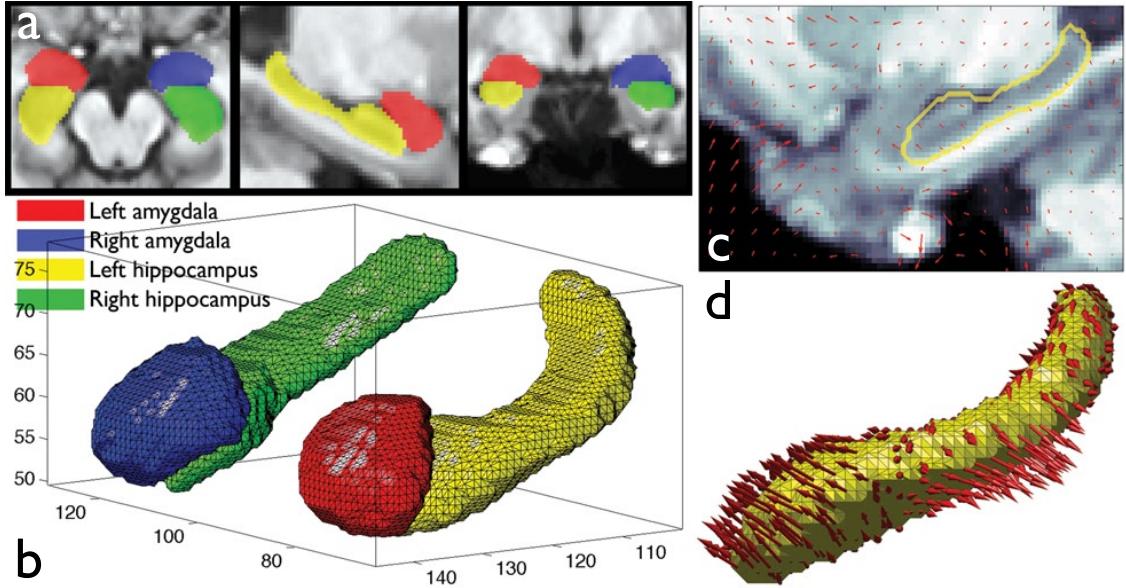


Figure 1: Processing of subcortical structures. Manual segmentations of amygdala and hippocampus (a) were used to construct template surfaces (b). Displacement vector field obtained during non-linear normalization (c) is interpolated onto the template surface (d). The length of displacement is used as an anatomical measure. Please see Kim *et al.* [12] for further explanation.

### 3 Sparse representation of anatomical measurements based on Laplace-Beltrami eigenfunctions

#### 3.1 Measurement parameterization based on Laplace-Beltrami eigenfunctions

Consider a real-valued functional measurement  $Y(p)$  on a manifold  $\mathcal{M} \subset \mathbb{R}^3$ . We assume the following additive model:

$$Y(p) = \theta(p) + \epsilon(p), \quad (1)$$

where  $\theta(p)$  is the unknown mean signal to be estimated and  $\epsilon(p)$  is a zero-mean Gaussian random field.

Solving

$$c\psi_j = \lambda_j\psi_j, \quad (2)$$

on  $\mathcal{M}$ , we find the eigenvalues  $\lambda_j$  and eigenfunctions  $\psi_j$ . The eigenfunctions  $\psi_j$  form an orthonormal basis in  $L^2(\mathcal{M})$ , the space of square integrable functions on  $\mathcal{M}$  [13]. We may order eigenvalues

ascendingly as

$$0 = \lambda_0 < \lambda_1 \leq \lambda_2 \cdots$$

and corresponding eigenfunctions as  $\psi_0, \psi_1, \psi_2, \dots$ .

Since the closed form expression for the eigenfunctions of the LB-operator on an arbitrary curved surface is unknown, the eigenfunctions are numerically calculated by discretizing the LB-operator. Using the Cotan discretization [4, 14], (2) is linearized as the generalized eigenvalue problem:

$$\mathbf{C}\boldsymbol{\psi} = \lambda \mathbf{A}\boldsymbol{\psi}, \quad (3)$$

where  $\mathbf{C}$  is the stiffness matrix,  $\mathbf{A}$  is the mass matrix and  $\boldsymbol{\psi} = (\psi(p_1), \dots, \psi(p_n))'$  is the unknown eigenfunction evaluated at  $n$  mesh vertices [18].

Once we obtained the eigenfunctions  $\psi_j$ , we can parametrically estimate the unknown mean signal  $\theta(p)$  as the Fourier expansion as

$$\theta(p) = \sum_{j=0}^k \beta_j \psi_j,$$

where  $\beta_j$  is the Fourier coefficients to be estimated. The coefficients are usually truncated at a low degree  $k$  and the high degree terms are ignored, assuming the most of signal is well representable by low degree eigenfunctions and the most of noise is in high degree eigenfunctions. In Styer et al., 12 and 15 degree spherical harmonics expansions were used for hippocampus and caudate respectively [19]. Simply we choose  $k$  as the maximal degree - 1. Once  $k$  is determined, the Fourier coefficients can be obtained by the usual least squares estimation (LSE) by solving

$$\mathbf{Y} = \boldsymbol{\psi}\boldsymbol{\beta}, \quad (4)$$

where  $\mathbf{Y} = (Y(p_1), \dots, Y(p_n))'$ ,  $\boldsymbol{\beta} = (\beta_1, \dots, \beta_k)'$  and  $\boldsymbol{\psi} = (\psi_i(p_j))$  is an  $n \times k$  matrix of eigenfunctions evaluated at mesh vertices [23]. Solving the normal equation (4), the LSE is given as

$$\hat{\boldsymbol{\beta}}_{LSE} = (\boldsymbol{\psi}'\boldsymbol{\psi})^{-1}\boldsymbol{\psi}'\mathbf{Y}. \quad (5)$$

### 3.2 Sparse representation using $l_1$ -minimization

However, the LSE may include a low degree coefficient that does not contribute significantly. Therefore, instead of using LSE, we introduce the  $l_1$ -norm penalty to sparsely filter out insignificant coefficients. The estimation is given [11] as:

$$\hat{\boldsymbol{\beta}} = \arg \min_{\boldsymbol{\beta}} \|\mathbf{Y} - \boldsymbol{\psi}\boldsymbol{\beta}\|_2^2 + \lambda \|\boldsymbol{\beta}\|_1, \quad (6)$$

where the parameter  $\lambda > 0$  controls the amount of sparsity. In the previous study [12], we chose  $\lambda = 1$ .

After we compute a sparse solution for the coefficients estimation using the  $l_1$ -minimization algorithm [11], we thresholded insignificant coefficients to be exact zero, using backward model selection scheme. A similar scheme is used for automatic selection of optimal degree of Fourier expansion in [6, 21]. Here we did not truncate higher degrees. Rather, we choose an optimal threshold for coefficients as in [17]. First we sort the coefficients in descending order. Then from the smallest coefficients (the  $k$ -th largest coefficient) to larger coefficients ( $i < k$ ), we perform a step-wise  $F$ -test to determine whether the exclusion of a certain coefficient significantly increases the error of model fitting.  $F$ -statistic for the  $i$ -th largest coefficient is given as

$$F = \frac{(SSE_{i-1} - SSE_i)/1}{SSE_{i-1}/(n - (i - 1) - 2)} \sim F(1, n - (i - 1) - 2), \quad (7)$$

where  $n$  is total number of measurement.

$$SSE_i = \sum \left( \mathbf{Y} - \sum_{k=1}^i \hat{\beta}_k \psi_k \right)^2$$

where  $\hat{\beta}_k$  is the  $k$ -th largest sparse coefficient estimated from (6) and  $\psi_k$  is the corresponding basis eigenfunction. If the exclusion of coefficient is not significant at  $\alpha$ -level of 0.01, that is, the  $i$ -th coefficient does not significantly contribute to the fitting, thus the process moves on to a larger coefficient ( $i - 1$ ). Otherwise, the procedure terminates and threshold all coefficients smaller than  $i$ -th largest coefficient to be zero. The coefficients estimated by  $l_1$ -minimization and zero-forced coefficients for a single subject's displacement length on a left hippocampus are given in Figure 2.

### 3.3 Gaußianness change due to sparse representation

To quantify how much sparse representation improves Gaußianness of measurements, we computed Kolmogorov-Smirnov statistic (KS-stat), which measures maximal difference between the empirical cumulative distribution function (CDF) of 'whitened' measures  $S(x)$  and the standard normal CDF  $F_z(x)$  as

$$KS = \max |S(x) - F_z(x)|.$$

The relative improved Gaußianness due to  $l_1$ -penalty in relation to LSE can be computed as

$$G = \frac{KS_{LSE} - KS_{l_1}}{KS_{LSE}}$$

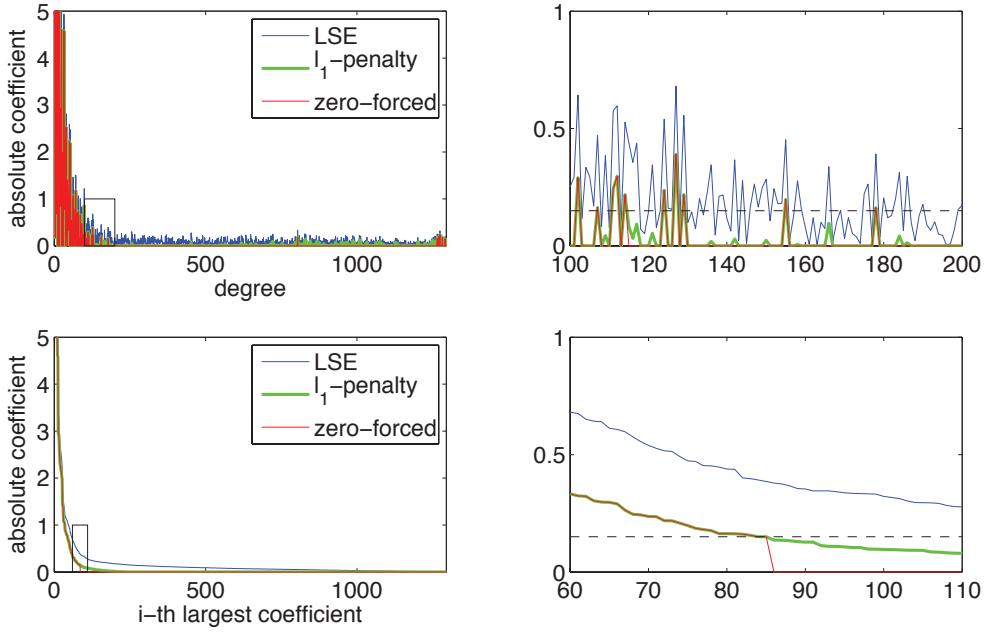


Figure 2: Absolute value of coefficients are compared. Top row shows the coefficients in the order of degree and bottom top shows coefficients in descending order. Dotted line in the magnified boxes in the right column indicate the thresholded value for zero-forcing.

The result is shown in Figure 3. In contrast to presumption, sparse representation of LSE does not improve Gaussianness of the measurement (overall mean = -0.0629 %pt  $\pm$  12.42%pt). Actually Kolmogorov-Smirnov test shows that the two emperical CDF of LSE and sparse representation are significantly different at  $\alpha = 0.05$  for only 15 cases out of 208 cases (52 subjects  $\times$  2 regions  $\times$  2 sides).

The effect of sparse representation on the statistic map in the previous study [12] presumably due to the decrease of sample standard deviation (overall mean relative improvement = 6.37 %pt  $\pm$  2.74%pt; figure not shown), that lead to the improvement of signal-to-noise ratio and the increase of test statistics.

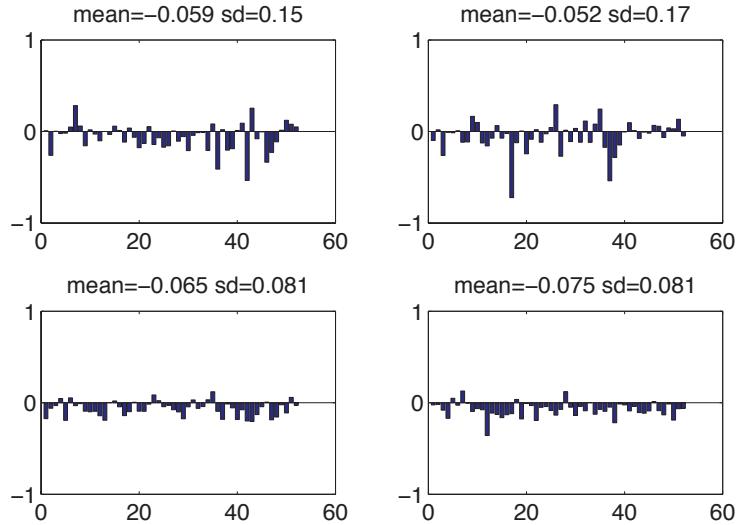


Figure 3: Relative improved Gaussianness due to sparse representation is shown over 52 subjects along the x-axis. The plots in the top row are for the left and right amygdala, and those in the bottom row are for the left and right hippocampus.

## 4 Power analysis

### 4.1 Computing power for a random variable

The power of a statistical test given as the probability of rejecting the null hypothesis  $H_0$  when it is actually not true [9]. Consider a model on an observation  $Y$  as

$$Y = \mu + \epsilon$$

where  $\mu$  is unknown signal and  $\epsilon$  is a zero mean Gaussian noise with unit variance. The null and alternative hypotheses can be given as

$$H_0 : \mu = 0 \text{ vs. } H_1 : \mu \neq 0.$$

If we have  $n$  samples, we can test the hypotheses using one sample  $t$ -test as

$$T = \frac{\bar{Y}}{s/\sqrt{n}}$$

where  $\bar{Y}$  is sample mean and  $s$  is sample standard deviation. We reject the null hypothesis at a given significance level  $\alpha$  if  $T > h(\alpha)$  where

$$h(\alpha) = \Phi_{n-1}^{-1}(1 - \alpha)$$

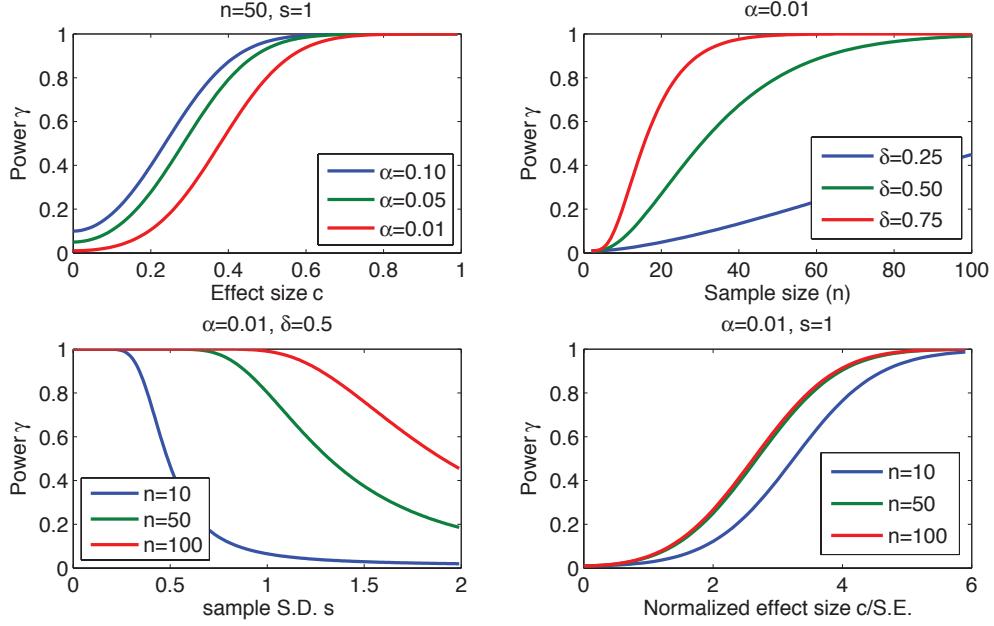


Figure 4: Power  $\gamma$  over effect size  $c$ , sample size  $n$ , sample standard deviation  $s$  and the effect size normalized by standard error  $c/(s/\sqrt{n})$ , respectively.

for a one-tailed test.  $\Phi_\nu$  is cumulative distribution function for the student's  $t$ -distribution with the degree of freedom  $\nu$  and  $(\cdot)^{-1}$  denotes the inverse of function.

Then the power  $\gamma$  at a significance level  $\alpha$  is given as

$$\gamma(\alpha) = 1 - \Pr(Y \leq h(\alpha) | H_1 \text{ is true}) \quad (8)$$

In order to compute power we need to determine a specific alternative hypothesis rather than the negate of the null hypothesis. For instance, if an alternative hypothesis  $H_1 : \mu = c > 0$  is true for a effect size  $c$ , then the power is computed as

$$\gamma(\alpha) = 1 - \Phi_{n-1} (h(\alpha) - c/(s/\sqrt{n})) .$$

The effects of those parameters are given in Figure 4. Note that the power not only subject to the effect size  $c$  but also to the significance level  $\alpha$ . As the effect size  $c$  is normalized by standard error  $s/\sqrt{n}$ , sample standard deviation  $s$  and the sample size  $n$  also affect the power. Additionally, the sample size  $n$  also affects the cumulative distribution function  $\Phi_\nu$ , but it is trivial for a large number of samples ( $n \geq 50$ ).

In neuroimaging studies, the significance level  $\alpha$  is conventionally given such as 0.05 or 0.01 and

effect size  $c$  is unknown but assumed to be constant. The sample size  $n$  can be determined using the effect size based on previous literature when one designs a study. But once the data collection is done, the sample size is given at the image processing stage. What we can change at this stage is only sample standard deviation  $s$  of measurements. We can do this using smoothing techniques.

## 4.2 Computing power for a random field

Now we consider a random field  $Y$  on a manifold  $\mathcal{M}$ . Then we model the observation at position  $x$  as

$$Y(x) = \mu(x) + \epsilon,$$

where  $\mu(x)$  is unknown signal and  $\epsilon$  is a zero mean Gaussian random field with unit variance. Then the collective null and alternate hypotheses are given as

$$\mathcal{H}_0 : \bigcap_{x \in \mathcal{M}} \mu(x) = 0 \text{ vs. } \mathcal{H}_1 : \bigcup_{x \in \mathcal{M}} \mu(x) \neq 0.$$

That is, now the power is the probability to detect a signal at least from one position.

When we specify the collective alternate hypothesis, it might be given as

$$\mathcal{H}_1 : \bigcap_{x \in \mathcal{M}} \mu(x) = c > 0. \quad (9)$$

However it would be less realistic in neuroimaging studies. For instance, in the shape modeling studies on subcortical structures [16, 20, 5], the effect is not uniform over all region, but localized on some subfield.

Thus it might be better to specify the alternative hypothesis as

$$\mathcal{H}_1 : \begin{cases} \bigcap_{x \in \mathcal{M}_1} \mu(x) = c > 0 \\ \bigcap_{x \in \mathcal{M}_0} \mu(x) = 0 \end{cases} \quad (10)$$

where  $\mathcal{M}_1$  is a subset of the manifold  $\mathcal{M}$  that shows effects of  $c$  and  $\mathcal{M}_0 = \mathcal{M} - \mathcal{M}_1$

In Hayasaka *et al.* [10], the power computation for a random field is formulated using a non-central random field. A non-central  $T$ -random field  $S$  with degree of freedom  $\nu$  and non-centrality scalar parameter  $\delta > 0$  is given [10] as

$$S = \frac{\sqrt{\nu}(Z + \delta)}{\sqrt{V}} \quad (11)$$

where  $Z$  is a Gaussian random field with a certain FWHM and  $V$  is a chi-square random field with degree of freedom  $\nu$ .

When the corrected  $p$ -value while controlling family-wise error rate (FWER) is given as

$$p = \Pr \left( \sup_{x \in \mathcal{M}} T(x) > h(\alpha) \right) \approx \sum_{D=0}^k R_D(\mathcal{M}) \rho_D(h(\alpha)) \quad (12)$$

where  $T(x)$  is a central  $T$ -random field given as  $T(x) = \bar{Y}(x)/(s(x)/\sqrt{n})$ ,  $k$  is the number of dimension of data,  $R_D$  is *resel* (resolution element) count in the  $D$ -th dimension and  $\rho_D$  is Euler Characteristic (EC) density in the  $D$ -th dimension [22].

Using (12), first we compute a threshold that corresponds to a given significance level  $h(\alpha)$ . Then the power  $\gamma$  is given as the probability that the maximum of  $t$ -random field exceeds the given threshold  $h(\alpha)$  within  $\mathcal{M}_1$  as

$$\gamma(\alpha) = \Pr \left( \sup_{x \in \mathcal{M}_1} T(x) > h(\alpha) \right) \approx \sum_{D=0}^k R_D(\mathcal{M}_1) \rho'_D(h(\alpha); \delta) \quad (13)$$

where  $\rho'_D(h(\alpha); \delta)$  is the  $D$ -dimensional EC density function for non-central  $T$ -random field  $S$  given by the threshold  $h(\alpha)$  and the non-centrality  $\delta$ . The non-centrality parameter for the non-central  $T$ -field is estimated using Cohen's  $d$  [8] in Hayasaka *et al.* [10] as  $\delta = d\sqrt{\nu}$  where  $\nu$  is the degree of freedom. Since  $d$  is given as an effect size normalized by sample standard deviation [8],  $\delta$  can be rewritten using the notation in this report as

$$\delta = c\sqrt{n-1}/s, \quad (14)$$

where  $c$  is effect size in its measured unit,  $n$  is the number of samples and  $s$  is sample standard deviation.

Note that it is sufficient to consider the positives (random field exceeds a certain threshold) within  $\mathcal{M}_1$  and ignore any positives within  $\mathcal{M}_0$  which is false positives. Because the power is the probability of true positives.

The EC density functions for non-central  $T$ -random field with degree of freedom  $\nu$  and a certain non-centrality  $\delta$  are given in Hayasaka *et al.* [10] as

$$\begin{aligned} \rho'_0(h; \delta) &= \int_h^\infty f_s(y) dy \\ \rho'_1(h; \delta) &= \left( \frac{4 \log 2}{2\pi} \right)^{1/2} \sqrt{\nu} \left( 1 + \frac{h^2}{\nu} \right) E \left[ W^{-1/2} \right] f_s(h) \\ \rho'_2(h; \delta) &= \frac{4 \log 2}{2\pi} \sqrt{\nu} \left( 1 + \frac{h^2}{\nu} \right) \left\{ (\nu - 1) \left( \frac{h^2}{\nu} \right)^{1/2} E \left[ W^{-1} \right] - \left( 1 + \frac{h^2}{\nu} \right)^{-1/2} E \left[ W^{1/2} \right] \delta \right\} f_s(h) \end{aligned}$$

where  $f_s$  is the probability density function of a non-central  $T$ -random variable with degree of freedom  $\nu$  and a certain non-centrality  $\delta$ , and  $W$  is a non-central chi-square random variable with degree of freedom  $\nu + 1$  and a certain non-centrality  $\delta^2$ . The probability density function of a non-central  $T$ -random variable is given as

$$f_s(h) = \frac{1}{2^{\frac{\nu-1}{2}} \Gamma(\nu/2) \sqrt{\nu\pi}} \int_0^\infty y^{\frac{\nu-1}{2}} \exp(-y/2) \exp\left(-\frac{1}{2} \left\{ h \sqrt{\frac{y}{\nu}} - \delta \right\}^2\right) dy,$$

where  $\Gamma$  is gamma function.

The expectation of  $W^b$  is given as

$$E[W^b] = 2^b \Gamma\left(b + \frac{\nu}{2}\right) \sum_{j=0}^{\infty} \frac{\Gamma(b+1)}{\Gamma(j+1)\Gamma(b-j+1)} \left(\frac{\delta}{2}\right)^j \frac{1}{\Gamma + \frac{\nu}{2}}.$$

The non-centrality  $\delta$  can be done using a local mean of  $T$ -field. Suppose a local subset is from non-central  $T$ -field  $S$ , then the expectation of  $S$  is

$$\mathbb{E}S = \mathbb{E}\left(\frac{Z + \delta}{\sqrt{V/\nu}}\right) = \frac{\mathbb{E}Z + \delta}{\sqrt{\mathbb{E}V/\nu}} = \delta \quad (15)$$

where  $Z$  is Gaussian Random field and  $V$  is  $\chi^2$  random field with the degree of freedom  $\nu$ .

### 4.3 Implementation using Monte-Carlo approach

The one of the advantage of using random field theory is its efficiency in actual computation. However, implementation of EC density function for the non-central  $T$ -filed is not yet publicized by Hayasaka [10]. Due to time limitation, I failed to implement infinite integrals in the formulation for this project.

However, the main goal of this project is not to implement power estimation using non-central  $T$ -field, but to compare the proposed  $l_1$ -minimization method to a conventional method such as LSE. Thus we numerically compute the power as the ratio of detection over all simulations. As the power for a certain degree of freedom  $\nu$  and  $\alpha$  level is given as

$$\gamma(\nu, \alpha) = \Pr\left(\sup_{x \in \mathcal{M}_1} T(x) > h(\alpha)\right),$$

we simply counted the number of instances of the maximum exceeding the given threshold at  $\alpha = 0.05$ , and divided by the total number of simulations.

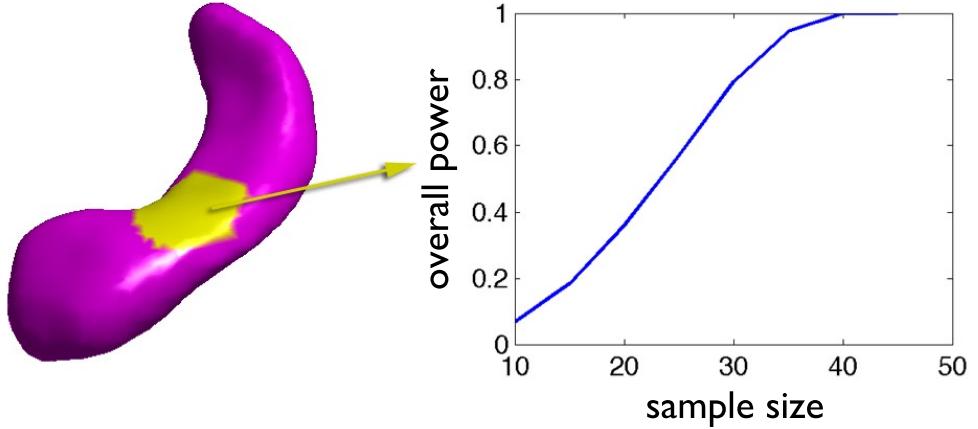


Figure 5: The overall power from an artificial dataset. The signal region (yellow) in the left panel is 6-ring neighbors of the central vertex. Right panel shows corresponding overall power for varying sample sizes.

**Simulation** For a synthetic dataset, we generated 50 Gaussian random fields ( $\sim N(0, 1)$ ) on the left hippocampus surface and added a signal with an effect size of 1 sample standard deviation for 25 cases. Then  $l_1$ -penalty is applied as explained in the previous section (3.2). Then the power is obtained for subsamples  $n = 10, 15, 20, \dots, 50$  over 1000 resamplings. The obtained overall power curve is shown in Figure 5. The power exceeded 80% with more than 30 cases.

**Real data** For a real dataset, we assume the result in the previous study with  $n=52$  subjects [12] as a ground truth. For the sake of simplicity, we look for the effect of gender: men *vs.* women. Subsamples of  $n = 10, 15, 20, \dots, 50$  were resampled from the original dataset for 1000 iterations each. The overall power curves for LSE and  $l_1$ -minimization are given in Figure 6. In relation to LSE,  $l_1$ -minimization shows higher power for a small number (<40) of samples, demonstrating the advantage of sparse regression over LSE.

#### 4.4 Power map

In Hayasaka *et al.* [10], the procedure of power map generation is explained. In a real data, we do not know the exact size and location of the real signal region  $\mathcal{M}_1$ . Instead, we can compute local power curves as if a certain small subregion is the real  $\mathcal{M}_1$ . If we compute local power curves for different locations, then we obtain different kind of power curves as in Figure 7.

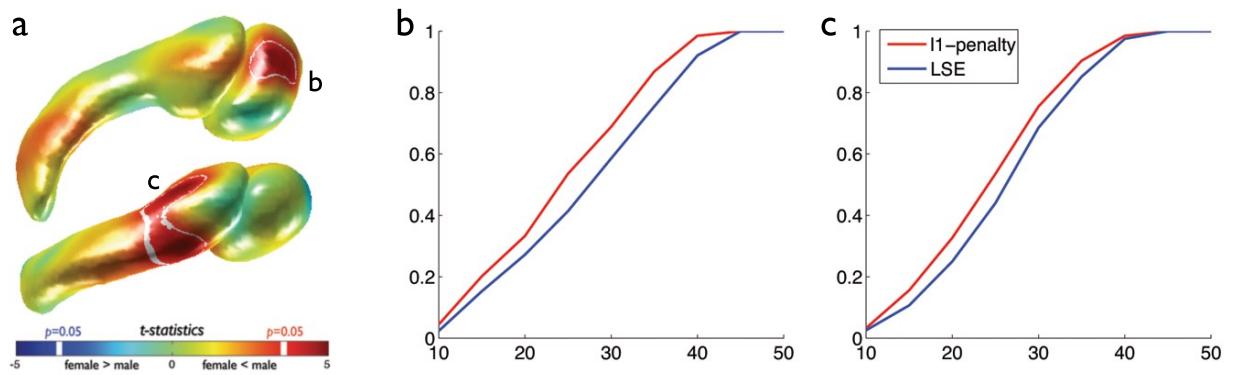


Figure 6: The overall power using a real dataset. The significant gender effect with whole 52 subjects (a), the overall power curves for  $l_1$ -minimization (red) and LSE (blue) are given for the right amygdala (b) and the left hippocampus (c).

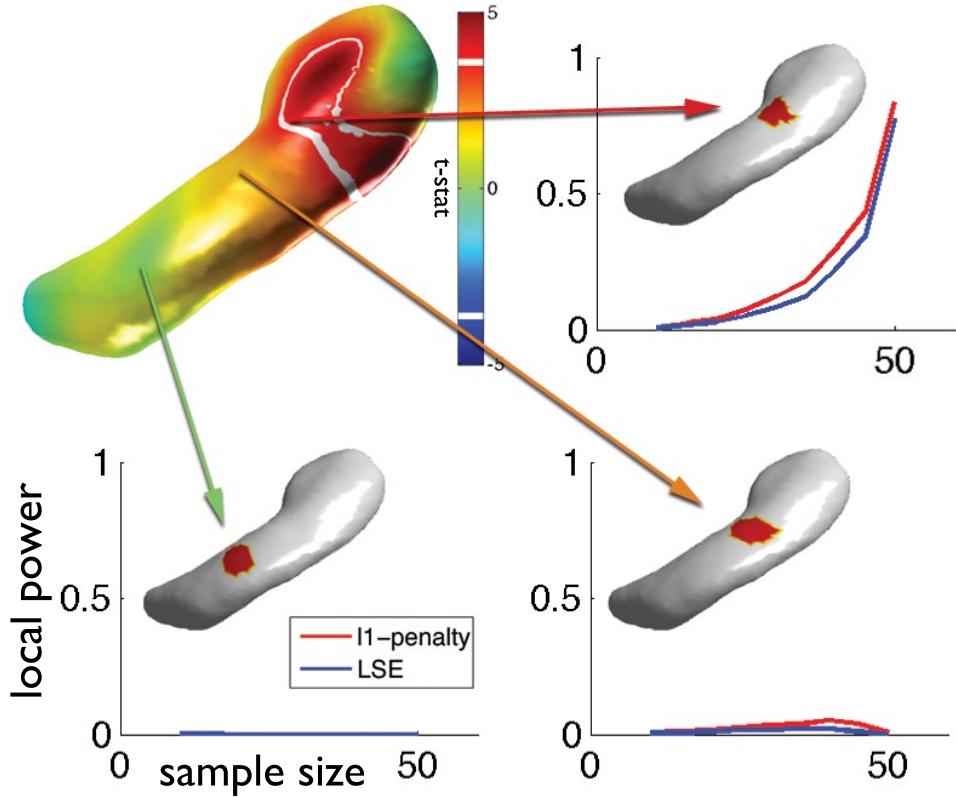


Figure 7: The local power using a real dataset. The size of local region (3-ring neighbors) is indicated as orange area in each cases. The local power curves for  $l_1$ -minimization (red) and LSE (blue) are given over sample size along the x-axis.

Suppose we only have a small number of subjects, and do not know about the actual extend of signal. If we estimate the power using random field theory, then we can obtain a power map that can be used to predict the number of samples to detect a signal to a given extend. In our case, we only use Monte-Carlo method, thus it is not applicable to a real pilot study. Even though, we can see the advantage of power map as in the following scenario:

1. You start a pilot study to see the effect of gender on subcortical structures with  $n=10$  subjects.  
You cannot detect any signals (Figure 8a).
2. You compute power map from the data using random field theory. It would be similar to the result from the simulation (Figure 8b,c).
3. You determine the optimal sample number as  $n=30$ . With the number of subjects, you detect the signals (Figure 8d).

In figure 9, the power maps and sample size maps obtained from LSE and  $l_1$ -minimization are compared. As seen in the comparison of overall power curves,  $l_1$ -minimization shows higher power with a larger extend for a given sample size than LSE. And also the sparse representation needs smaller sample size to achieve a given power level. The power maps demonstrates that sparse representation has advantage not only in the overall power, but also in the discoverable extend of signal.

## 5 Conclusion

In this report, we compared sparse representation based on Laplace-Beltrami eigenfunctions to a conventional least square estimation in terms of statistical power. Even though the implementation of power estimation using random field theory was not done during the preparation of this report, using Monte-Carlo approach, the power estimation was made. We observed the benefit of sparse representation over LSE in terms of overall power and using the power map, and checked that we are more likely to find (due to higher statistical power) a larger extend signal using  $l_1$ -minimization.

## References

- [1] J. Ashburner and K. Friston. Voxel-based morphometry - the methods. *NeuroImage*, 11:805–821, 2000.

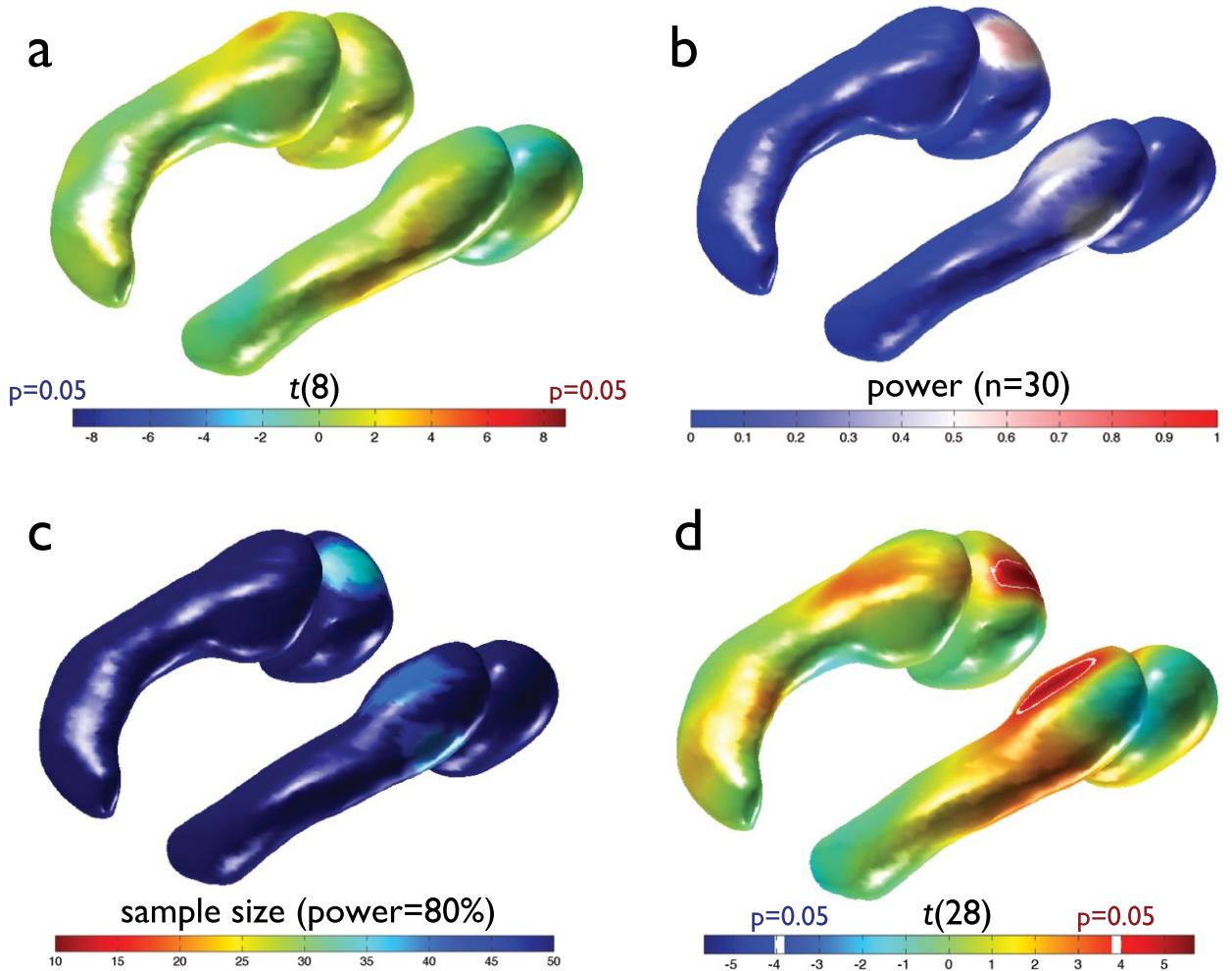


Figure 8: Scenario of pilot study. Non significant results with pilot data  $n=10$  (a). Power maps with  $n=30$  (b) and sample size map needed for power=80% (c) can be derived from the pilot data using random field theory. In this report, they are generated from Monte-Carlo method. Suppose the number of subjects are determined as  $n=30$ , as predicted, significant signals are detected (d).

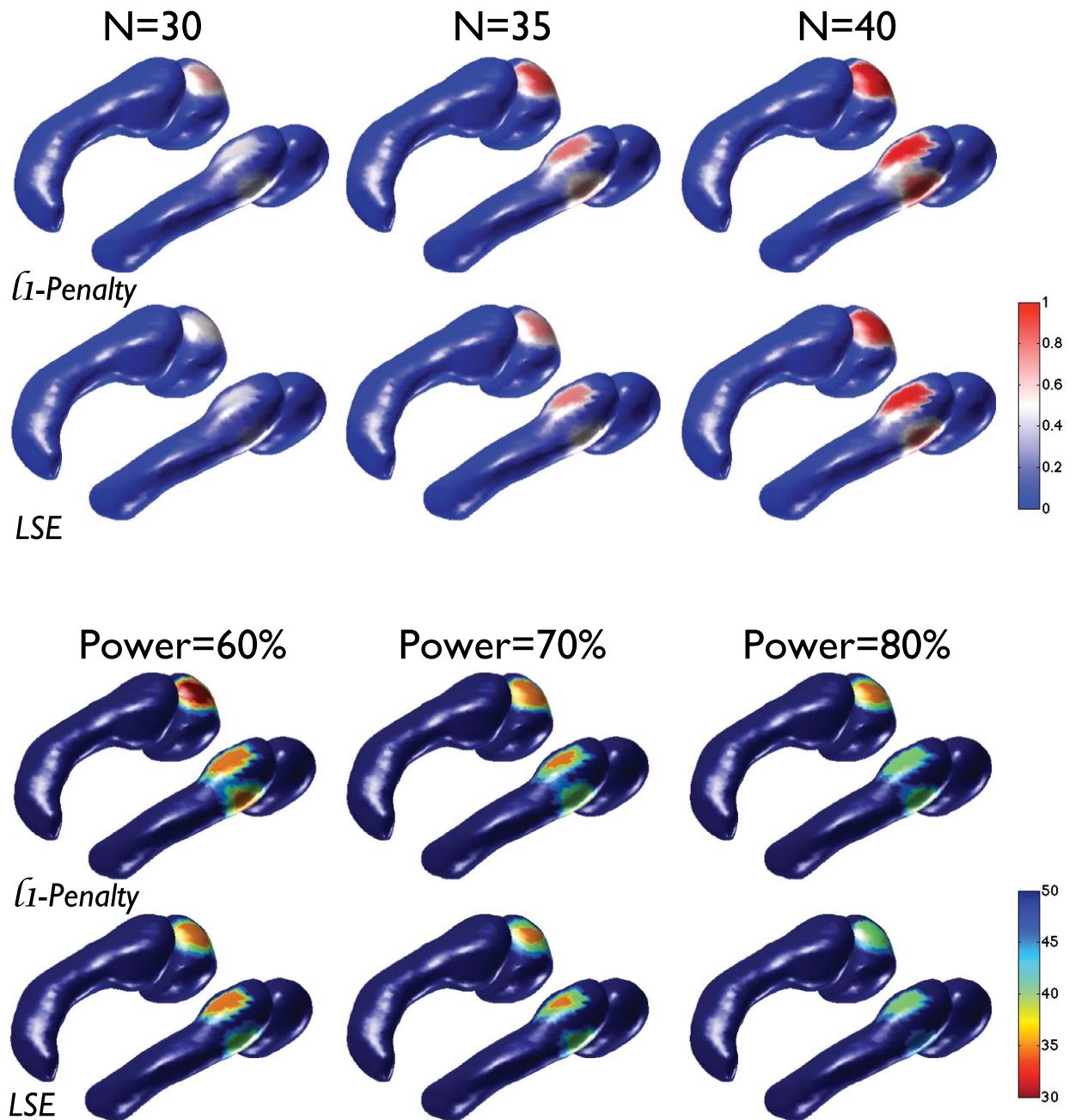


Figure 9: Power maps (first and second rows) and sample size maps (third and forth rows) are compared between methods for varying powers and sample sizes. Note that sparse representation shows higher power with a given sample size than LSE and requires less sample size to achieve a given power level.

- [2] J. Ashburner, C. Good, and K. Friston. Tensor based morphometry. *NeuroImage*, 11S:465, 2000.
- [3] B. Avants, C. Epstein, M. Grossman, and J. Gee. Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, 12(1):26–41, 2008.
- [4] M. Chung and J. Taylor. Diffusion smoothing on brain surface via finite element method. In *Biomedical Imaging: Nano to Macro, 2004. IEEE International Symposium on*, pages 432–435. IEEE, 2004.
- [5] M. Chung, K. Worsley, B. Nacewicz, K. Dalton, and R. Davidson. General multivariate linear modeling of surface shapes using surfstat. *NeuroImage*, 2010.
- [6] M. K. Chung, K. J. Worsley, B. M. Nacewicz, K. M. Dalton, and R. J. Davidson. General multivariate linear modeling of surface shapes using surfstat. *NeuroImage*, 53(2):491 – 505, 2010.
- [7] M. K. Chung, K. J. Worsley, T. Paus, D. L. Cherif, C. Collins, J. Giedd, J. L. Rapoport, , and A. C. Evans. A unified statistical approach to deformation-based morphometry. *NeuroImage*, 14:595–606, 2001.
- [8] J. Cohen. *Statistical power analysis for the behavioral sciences : Jacob Cohen.* Lawrence Erlbaum, 2 edition, 1988.
- [9] J. Cohen. Statistical power analysis. *Current Directions in Psychological Science*, 1(3):pp. 98–101, 1992.
- [10] S. Hayasaka, A. Peiffer, C. Hugenschmidt, and P. Laurienti. Power and sample size calculation for neuroimaging studies by non-central random field theory. *NeuroImage*, 37(3):721–730, 2007.
- [11] S. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky. An interior-point method for large-scale l1-regularized least squares. *Selected Topics in Signal Processing, IEEE Journal of*, 1(4):606–617, 2007.
- [12] S.-G. Kim, S. Schaefer, M. K. Chung, C. Van Reekum, and R. Davidson. Sparse shape representation using the laplace-beltrami eigenfunctions and its application to modeling subcortical structures. In *IEEE workshop on Mathematical Methods in Biomedical Image Analysis (MM-BIA)*, 2012.

- [13] B. Lévy. Laplace-beltrami eigenfunctions towards an algorithm that understands geometry. In *Shape Modeling and Applications, 2006. SMI 2006. IEEE International Conference on*, pages 13–13. IEEE, 2006.
- [14] A. Qiu, D. Bitouk, and M. Miller. Smooth functional and structural maps on the neocortex via orthonormal bases of the laplace-beltrami operator. *Medical Imaging, IEEE Transactions on*, 25(10):1296–1306, 2006.
- [15] A. Qiu, T. Brown, B. Fischl, J. Ma, and M. Miller. Atlas generation for subcortical and ventricular structures with its applications in shape analysis. *Image Processing, IEEE Transactions on*, 19(6):1539 –1547, June 2010.
- [16] A. Qiu and M. Miller. Multi-structure network shape analysis via normal surface momentum maps. *NeuroImage*, 42(4):1430–1438, 2008.
- [17] S. Seo, M. Chung, K. Dalton, and R. Davidson. Multivariate cortical shape modeling based on sparse representation. In *17th Annual Meeting of the Organization for Human Brain Mapping (OHBM)*, 2011.
- [18] S. Seo, M. K. Chung, and H. K. Vorperian. Heat kernel smoothing using Laplace-Beltrami eigenfunctions. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2010*, volume 6363 of *Lecture Notes in Computer Science*, pages 505–512. Springer, 2010.
- [19] M. Styner, I. Oguz, S. Xu, C. Brechbuhler, D. Pantazis, J. Levitt, M. Shenton, and G. Gerig. Framework for the statistical shape analysis of brain structures using spharm-pdm. In *Insight Journal, Special Edition on the Open Science Workshop at MICCAI*, 2006. <http://hdl.handle.net/1926/215>.
- [20] P. Thompson, H. K.M., de Zubicaray G., J. A.L., R. S.E., S. J., H. M.S., D. Herman, D. Gravano, D. Doddrell, and T. A.W. Mapping hippocampal and ventricular change in alzheimer disease. *NeuroImage*, 22:1754–1766, 2004.
- [21] S. Wang. *Weighted fourier image analysis and modeling*. PhD thesis, Univeristy of Wisconsin-Madison, 2008.
- [22] K. Worsley, S. Marrett, P. Neelin, A. Vandal, K. Friston, A. Evans, et al. A unified statistical approach for determining significant signals in images of cerebral activation. *Human brain mapping*, 4(1):58–73, 1996.

- [23] H. Zhang, O. van Kaick, and R. Dyer. Spectral methods for mesh processing and analysis. In *Eurographics 2007 – State of the Art Reports*, pages 1–22, 2007.