

# Unified Statistical Approach to Cortical Thickness Analysis

Moo K. Chung<sup>1</sup>, Steve Robbins<sup>2</sup>, and Alan C. Evans<sup>2</sup>

<sup>1</sup>Department of Statistics, University of Wisconsin-Madison

<sup>2</sup>Montreal Neurological Institute, McGill University, Canada  
mchung@stat.wisc.edu, {steve, alan}@bic.mni.mcgill.ca

**Abstract.** This paper presents a unified image processing and analysis framework for cortical thickness in characterizing a clinical population. The emphasis is placed on the development of data smoothing and analysis framework. The human brain cortex is a highly convoluted surface. Due to the convoluted non-Euclidean surface geometry, data smoothing and analysis on the cortex are inherently difficult. When measurements lie on a curved surface, it is natural to assign kernel smoothing weights based on the geodesic distance along the surface rather than the Euclidean distance. We present a new data smoothing framework that address this problem implicitly without actually computing the geodesic distance and present its statistical properties. Afterwards, the statistical inference is based on the random field theory based multiple comparison correction. As an illustration, we have applied the method in detecting the regions of abnormal cortical thickness in 16 high functioning autistic children.

## 1 Introduction

The human cerebral cortex has the topology of a 2D highly convoluted grey matter shell of average thickness of 3mm. The thickness of the grey matter shell is usually referred as the *cortical thickness* and can be obtained from magnetic resonance images (MRI). The cortical thickness can be used as an anatomical index for quantifying cortical shape variations. The thickness measures are obtained after a sequence of image processing steps which are described briefly here. The first step is to classify each voxel into three different tissue types: cerebrospinal fluid (CSF), grey matter, and white matter. The CSF/grey matter interface is called the *outer cortical surface* while the grey/white matter interface is called the *inner cortical surface*. These two surfaces bound the gray matter. The mainstream approach in representing the cortical surface has been to use a fine triangular mesh that is constructed from deformable surface algorithms [10] [14]. Cortical thickness is estimated by computing the distance between the two triangular meshes [11] [14]. In our study, we have used the method presented in [14]. In order to compare cortical thickness measures across subjects, it is necessary to align the cortical surfaces via surface registration algorithms [16] [20].

For cross-comparison between subjects, surfaces are registered into the *template surface* which serves as reference coordinates.

The image segmentation, thickness computation and surface registration procedures are expected to introduce noise in the thickness measure. In order to increase the signal-to-noise ratio (SNR) and smoothness of data for the random field theory, some type of data smoothing is necessary. For 3D whole brain MRIs, Gaussian kernel smoothing is widely used to smooth data, in part, due to its simplicity in numerical implementation. The Gaussian kernel weights an observation according to its Euclidean distance. However, data residing on the convoluted brain surface fails to be isotropic in the Euclidean sense. On the curved surface, a straight line between two points is not the shortest distance so one may incorrectly assign less weights to closer observations. So when the observations lie on the cortical surface, it is more natural to assign the weights based on the geodesic distance along the surface. Previously *diffusion smoothing* has been developed for smoothing data along the cortex before the random field based multiple comparison correction [1] [6] [7]. By solving a diffusion equation on a manifold, Gaussian kernel smoothing can be indirectly generalized. Although diffusion smoothing has been used widely in image analysis starting with [15], most of previous work is about surface fairing [19]. There is a very few publications that smooth out observations defined on surface for data analysis [1] [3] [6] [7]. The drawback of the previous diffusion smoothing approach is the need for setting of up a finite element method (FEM) to solve the diffusion equation numerically and making the algorithm converges [6]. To address this problem, we have developed a simpler and more efficient method based on the heat kernel convolution on a manifold.

As an illustration, the method was applied to groups of autistic and normal subjects, and we were able to detect the regions of statistically significant cortical thickness difference between the groups.

## 2 Heat Kernel Smoothing

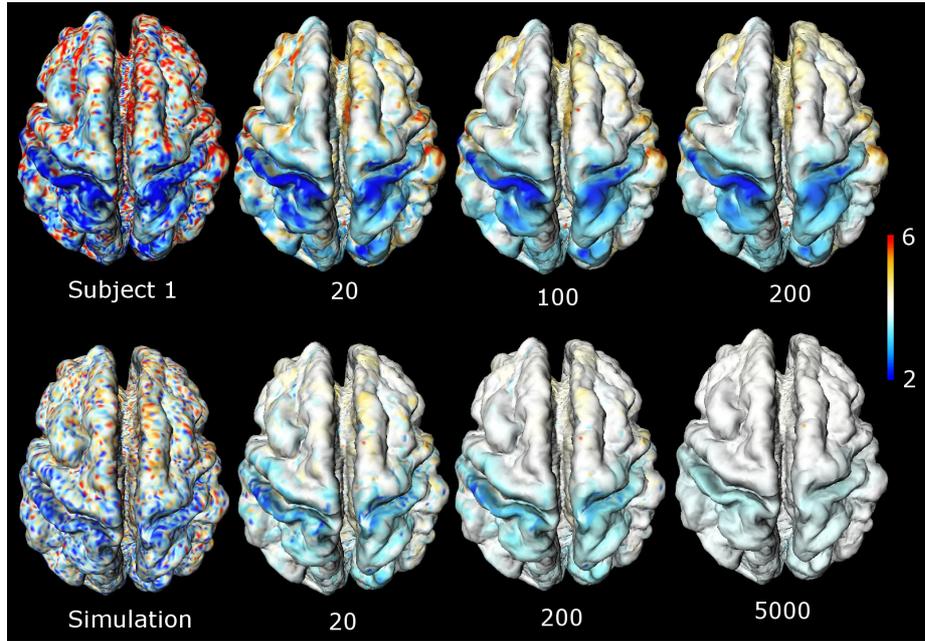
The cortical surface  $\partial\Omega$  can be assumed to be a  $C^2$  Riemannian manifold [12]. Let  $p = X(u^1, u^2) \in \partial\Omega$  be the parametric representation of  $\partial\Omega$ . We assume the following model on thickness measure  $Y$ :

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

where  $\theta(p)$  is a mean thickness function and  $\epsilon(p)$  is a zero-mean random field, possibly a Gaussian white noise process, with covariance function  $R_\epsilon(p, q)$ . The Laplace-Beltrami operator  $\Delta$  corresponding to the surface parameterization  $p = X(u^1, u^2) \in \partial\Omega$  is given by

$$\Delta = \frac{1}{\det g^{1/2}} \sum_{i,j=1}^2 \frac{\partial}{\partial u^i} \left( \det g^{1/2} g^{ij} \frac{\partial}{\partial u^j} \right),$$

where  $g = (g_{ij})$  is the Riemannian metric tensor. Solving equation  $\Delta\psi = \lambda\psi$ , we order eigenvalues  $0 = \lambda_0 \leq \lambda_1 \leq \lambda_2 \leq \dots$  and corresponding eigenfunc-



**Fig. 1.** Top: Heat kernel smoothing of cortical thickness with  $\sigma = 1$  and  $k = 20, 100, 200$  iterations. Bottom: Heat kernel smoothing on simulated data with  $\sigma = 1$  and  $k = 20, 200, 5000$  iterations. The mean thickness and the variance are estimated from 12 normal subject data and Gaussian white noise is added to the mean function

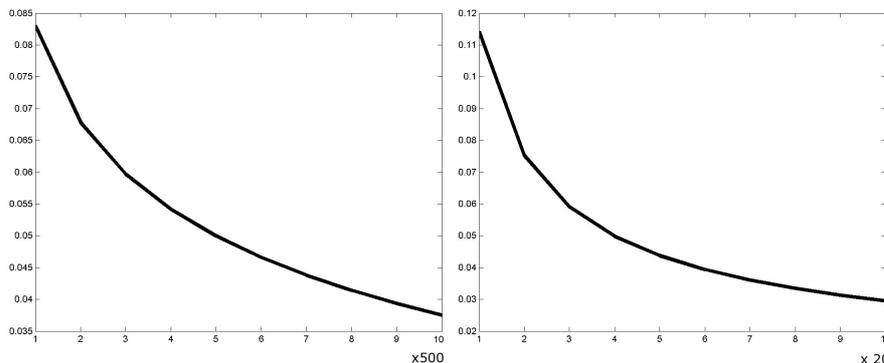
tions  $\psi_0, \psi_1, \dots$ . The eigenfunctions  $\psi_j$  form orthonormal basis of  $L^2(\partial\Omega)$ , the  $L^2$  space of functions defined on  $\partial\Omega$ . On the unit sphere, the eigenvalues are  $m(m+n-1)$  and the corresponding eigenfunctions are spherical harmonics  $Y_{lm}$  ( $|m| \leq l, 0 \leq l$ ) [21]. On an arbitrary surface, the explicit representation of eigenvalues and eigenfunction are only obtained through numerical methods. Based on orthonormal basis, the *heat kernel*  $K_\sigma(p, q)$  is analytically given as

$$K_\sigma(p, q) = \sum_{j=0}^{\infty} e^{-\lambda_j \sigma} \psi_j(p) \psi_j(q), \quad (1)$$

where  $\sigma$  is the bandwidth of the kernel [2] [17]. When  $g_{ij} = \delta_{ij}$ , the heat kernel becomes the Gaussian kernel, which is the probability density of  $N(0, \sigma^2)$ . natural extension of the Gaussian kernel. This can be interpreted as the transition probability density for an isotropic diffusion process with respect to the surface area element [22]. The kernel is symmetric, i.e.  $K_\sigma(p, q) = K_\sigma(q, p)$  and isotropic with respect to the geodesic distance  $d(p, q)$ .

**Definition 1.** *Heat kernel smoothing of cortical thickness  $Y$  is the convolution:*

$$K_\sigma * Y(p) = \int_{\partial\Omega} K_\sigma(p, q) Y(q) dq. \quad (2)$$



**Fig. 2.** Left: Within-subject variance plotted over the number of iterations of heat kernel smoothing with  $\sigma = 1$ . Decreasing variance implies the convergence of the heat kernel smoothing to the mean thickness (Theorem 4). Right: Between-subject variance plotted over the number of iterations illustrating Theorem 5

It can be written in terms of basis function expansion:

$$K_\sigma * Y(p) = \sum_{j=0}^{\infty} \alpha_j \phi_j(p),$$

where  $\alpha_j = e^{-\lambda_j \sigma} \int_{\partial\Omega} \phi_j(q) Y(q) dq$ . We also define the *heat kernel estimator* of unknown signal  $\theta(p)$  to be  $\hat{\theta}_\sigma(p) = K_\sigma * Y(p)$ . As  $\sigma \rightarrow 0$ ,  $K_\sigma(p, q)$  becomes the Dirac delta function  $\delta(p, q)$  so the heat kernel estimator becomes unbiased as  $\sigma \rightarrow 0$ , i.e.  $\lim_{\sigma \rightarrow 0} \mathbb{E} \hat{\theta}_\sigma(p) = \theta(p)$ . As  $\sigma$  gets larger, the bias increases. However the total bias over all cortex is always zero, i.e.  $\int_{\partial\Omega} [\theta(p) - \mathbb{E} \hat{\theta}_\sigma(p)] dp = 0$ . Let us list important nontrivial properties of heat kernel smoothing.

**Theorem 1.**  $K_\sigma * Y$  is the unique solution of the following isotropic diffusion equation at time  $t = \sigma^2/2$ :

$$\frac{\partial f}{\partial t} = \Delta f, f(p, 0) = Y(p), p \in \partial\Omega \tag{3}$$

This is a well known result [17]. This theorem implies that the heat kernel smoothing isotropically assigns weights on  $\partial\Omega$ .

**Theorem 2.**

$$K_\sigma * Y(p) = \arg \min_{\theta(p) \in L^2(\partial\Omega)} \int_{\partial\Omega} K_\sigma(p, q) [Y(q) - \theta(p)]^2 dq.$$

The proof can be found in [5]. This shows that the heat kernel smoothing can be formulated as a regression on a manifold.

**Theorem 3.**

$$\underbrace{K_\sigma * \dots * K_\sigma}_{k \text{ times}} * Y = K_{\sqrt{k}\sigma} * Y.$$

This can be seen as a scale space property of diffusion. From Theorem 1,  $K_\sigma * (K_\sigma * Y)$  can be taken as the diffusion of signal  $K_\sigma * Y$  after time  $\sigma^2/2$  so that  $K_\sigma * (K_\sigma * Y)$  is the diffusion of signal  $Y$  after time  $\sigma^2$ . Hence

$$K_\sigma * K_\sigma * Y = K_{\sqrt{2}\sigma} * Y.$$

Arguing inductively we see that the general statement holds. We will denote the  $k$ -fold iterated kernel as  $K_\sigma^{(k)} = \underbrace{K_\sigma * \dots * K_\sigma}_{k \text{ times}}$ . This is the basis of our iterated

heat kernel smoothing. Heat kernel with a large bandwidth will be performed by iteratively applying heat kernel smoothing with a smaller bandwidth. For instance iterated heat kernel smoothing with  $\sigma = 1$  and  $k = 200$  will generate heat kernel smoothing with the effective bandwidth of  $\sqrt{200} = 14.14\text{mm}$ . Figure 1 shows the process of iterated heat kernel smoothing.

**Theorem 4.**

$$\lim_{\sigma \rightarrow \infty} K_\sigma * Y = \frac{\int_{\partial\Omega} Y(q) dq}{\mu(\partial\Omega)}.$$

Here  $\mu(\partial\Omega)$  is the total surface area of  $\partial\Omega$ . This theorem shows that when we choose large bandwidth, heat kernel smoothing converges to the sample mean of data on  $\partial\Omega$ . Figure 1 (bottom) shows the convergence of heat kernel smoothing to the within-subject mean cortex 4mm as the bandwidth increases. Figure 2 (left) shows the convergence of the within-subject variance indirectly implying  $K_\sigma * Y$  converges to a constant, which is the average thickness over the cortex.

It is natural to assume the measurements  $Y(p)$  and  $Y(q)$  to have less correlation when  $p$  and  $q$  are away so we assume the covariance function to be  $R_\epsilon(p, q) = \rho(d(p, q))$  for some nondecreasing function  $\rho$ . Then we can show the variance reduction property of heat kernel smoothing.

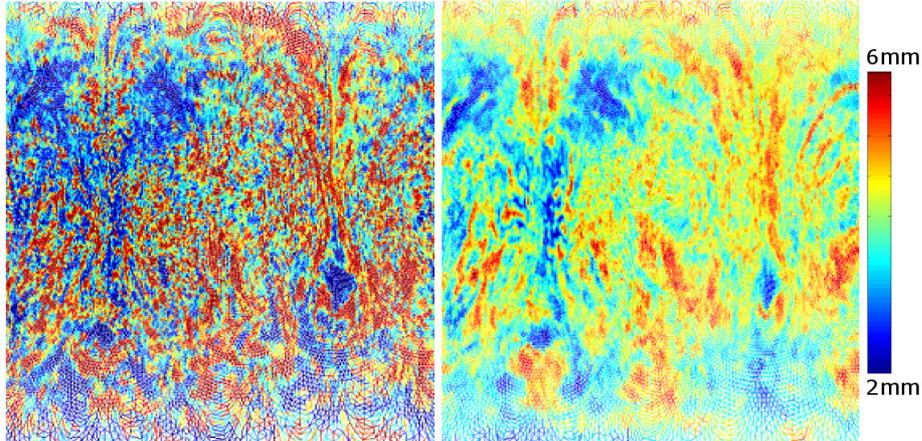
**Theorem 5.**  $\text{Var}[K_\sigma * Y(p)] \leq \text{Var}Y(p)$  for each  $p \in \partial\Omega$ .

Figure 2 (right) shows the between-subject variance decreases as  $\sigma$  increases.

The problem with the heat kernel smoothing on an arbitrary surface is that the explicit analytic form of the heat kernel is unknown. To address this problem we use the *parametrix expansion* of the heat kernel [17] [22]:

$$K_\sigma(p, q) = \frac{1}{(2\pi\sigma)^{1/2}} \exp\left[-\frac{d^2(p, q)}{2\sigma^2}\right] [1 + O(\sigma^2)] \tag{4}$$

for small  $d(p, q)$ . This expansion spells out the exact form of the kernel for small bandwidth. When the metric is flat, the heat kernel becomes a Gaussian kernel, reconfirming that heat convolution is a generalization of Gaussian kernel. The expansion is the basis of our heat kernel smoothing formulation. Heat kernel smoothing with a large bandwidth will be decomposed into iterated kernel



**Fig. 3.** Thickness maps are projected onto a unit square. Left: original noisy thickness map. Right: Heat kernel smoothing with  $\sigma = 1$  and  $k = 200$  iterations

smoothing. We will truncate and normalize the heat kernel using the first order term. For each  $p \in \partial\Omega$ , we define

$$\tilde{K}_\sigma(p, q) = \frac{\exp\left[-\frac{d^2(p, q)}{2\sigma^2}\right] \mathbf{1}_{B_p}(q)}{\int_{B_p} \exp\left[-\frac{d^2(p, q)}{2\sigma^2}\right] dq}, \quad (5)$$

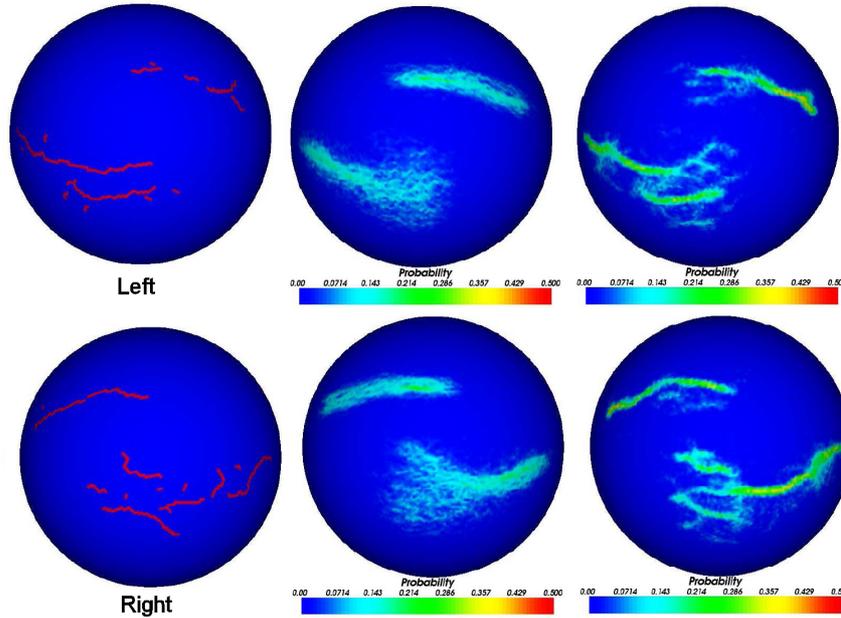
where  $\mathbf{1}_{B_p}$  is an indicator function defined on a small compact domain containing  $B$  such that  $\mathbf{1}_{B_p}(q) = 1$  if  $q \in B_p$  and  $\mathbf{1}_{B_p}(q) = 0$  otherwise. Note that for each fixed  $p$ ,  $\tilde{K}_\sigma(p, q)$  defines a probability distribution in  $B_p$  and it converges to  $K_\sigma(p, q)$  as  $\sigma \rightarrow 0$  in  $B_p$ . This implies

$$\tilde{K}_\sigma^{(k)} * Y(p) \rightarrow K_\sigma^{(k)} * Y(p) \text{ as } \sigma \rightarrow 0.$$

For a discrete triangular mesh, we can take  $B_p$  to be a set of points containing  $p$  and its neighboring nodes  $q_1, \dots, q_m$ , and take a discrete measure on  $B_p$ , which still make (5) a probability distribution. This can be viewed as a *Gaussian kernel Nadaraya-Watson* type smoothing extended to manifolds [4]. Figure 3 shows a flattened thickness map illustrating how heat kernel smoothing can enhance the thickness pattern by increasing the signal-to-noise ratio.

### 3 Random Field Theory on Cortical Manifold

Here we will describe how to perform multiple comparisons on  $\partial\Omega$  using the random field theory. The random field theory based approach is widely used for correcting multiple comparisons in 3D whole brain volume but rarely used on 2D cortical manifolds [1] [7] [8] [24]. First we combine both the autistic and the



**Fig. 4.** Automatically generated traces of the central and superior temporal sulcal fundi [3]. The first column shows the traces generated for the template surface. The second column shows the probability of sulcal matching based on 149 normal subjects before any surface normalization. The third column shows the probabilities after surface normalization. The first row is the left hemisphere and the second row is the right hemisphere. Note that the distribution is much more spatially concentrated and the matching probabilities are much greater after normalization

control subjects in a single indexing  $j$  and set up a general linear model (GLM) on cortical thickness  $Y_j$  for subject  $j$ :

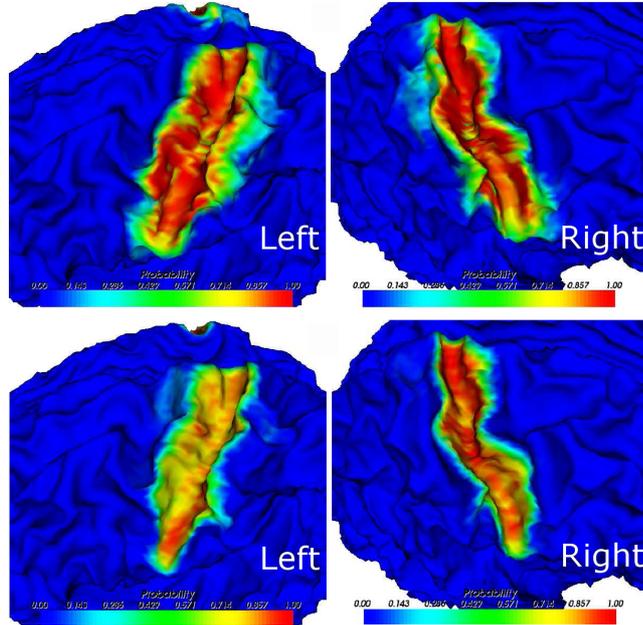
$$K_\sigma * Y_j(p) = \lambda_1(p) + \lambda_2(p) \cdot \text{age}_j + \lambda_3(p) \cdot \text{volume}_j + \beta(p) \cdot \text{group}_j + \epsilon_j \quad (6)$$

is used. Here dummy variable `group` is 1 for the autistic subjects and 0 for the normal subjects. `volume` is the total gray matter volume for subject  $j$ . The total gray matter volume is estimated by computing the volume bounded by the both outer and inner surfaces [8]. The error is modeled as a smooth Gaussian random field which is viewed as the heat kernel convolution with Gaussian white noise, i.e.  $\epsilon_j = K_\sigma * W$ . Then we test the group difference by performing a hypothesis testing:

$$H_0 : \beta(p) = 0 \text{ for all } p \in \partial\Omega$$

v.s.

$$H_1 : \beta(p) \neq 0 \text{ for some } p \in \partial\Omega.$$



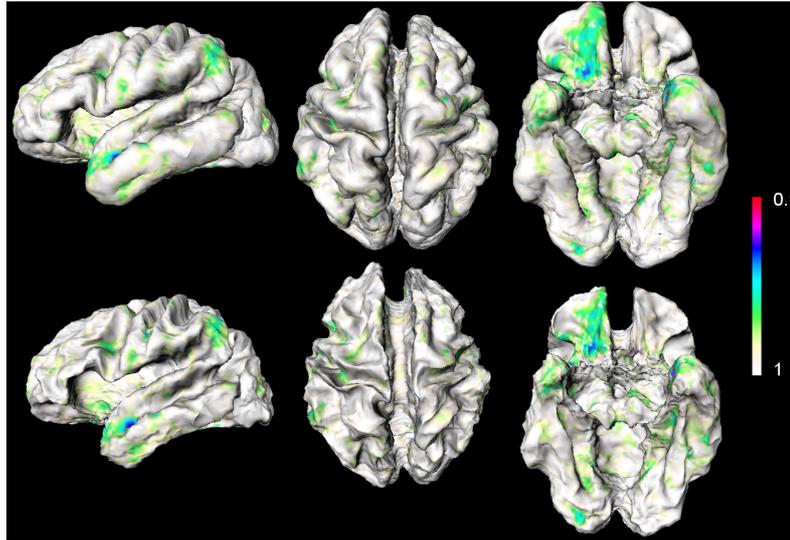
**Fig. 5.** Probability of sulcal matching, after normalization, for 39 manually identified central sulci defined as the surface region surrounded by gyri, not just the fundus. The views are illustrated on a slightly-opened version of the template cortical surface in order to better view inside the sulcus. The warping in 2D localizes the central sulcus nearly completely inside the template central sulcus. Left (right) figure is the left (right) central sulci

The test statistic is the ratio of the sum of the squared residual errors under the null and alternate models. Under  $H_0$ , the test statistic is a  $F$  random field with 1 and  $n = n_1 + n_2 - 4$  degrees of freedom [23]. The null hypothesis is the intersection of collection of hypothesis  $H_0 = \bigcap_{p \in \partial\Omega} H_0(p)$ , where  $H_0(p) : \beta(p) = 0$  for each fixed  $p$ . The type I error for the multiple comparisons is then given by

$$\begin{aligned} \alpha &= P\left(\bigcup_{p \in \partial\Omega} \{F(p) > h\}\right) = 1 - P\left(\bigcap_{p \in \partial\Omega} F(p) \leq h\right) \\ &= 1 - P(\sup_{p \in \partial\Omega} F(p) \leq h) = P(\sup_{p \in \partial\Omega} F(p) > h) \end{aligned}$$

for some  $h$ . The resulting  $p$ -value is usually called the *corrected p-value*. The distribution of  $\sup_{p \in \partial\Omega} F(p)$  is asymptotically given as

$$P(\sup_{p \in \partial\Omega} F(p) > h) \approx \sum_{d=0}^2 \phi_d(\partial\Omega) \rho_d(h) \tag{7}$$



**Fig. 6.** Corrected  $p$  value maps of  $F$ -test removing the effect of age and relative gray matter volume difference projected onto the average outer (top) and inner surfaces (bottom). It shows relatively asymmetric thickness difference between two groups

where  $\phi_d$  are the  $d$ -dimensional Minkowski functionals of  $\partial\Omega$  and  $\rho_d$  are the  $d$ -dimensional Euler characteristic (EC) density of  $F$ -field with  $\alpha = 1$  and  $\beta = n$  degrees of freedom [23]. The Minkowski functionals are  $\phi_0 = 2, \phi_1 = 0, \phi_2 = \text{area}(\partial\Omega)/2 = 49,616\text{mm}^2$ , the half area of the template cortex  $\partial\Omega$ . The EC density is given by

$$\rho_0(h) = \int_h^\infty \frac{\Gamma(\frac{\alpha+\beta}{2})}{\Gamma(\frac{\alpha}{2})\Gamma(\frac{\beta}{2})} \frac{\alpha}{\beta} \left(\frac{\alpha x}{\beta}\right)^{\frac{(\alpha-2)}{2}} \left(1 + \frac{\alpha x}{\beta}\right)^{-\frac{(\alpha+\beta)}{2}} dx,$$

$$\rho_2(h) = \frac{\lambda}{2\pi} \frac{\Gamma(\frac{\alpha+\beta-2}{2})}{\Gamma(\frac{\alpha}{2})\Gamma(\frac{\beta}{2})} \left(\frac{\alpha h}{\beta}\right)^{\frac{(\alpha-2)}{2}} \left(1 + \frac{\alpha h}{\beta}\right)^{-\frac{(\alpha+\beta-2)}{2}} \times \left[(\beta-1)\frac{\alpha h}{\beta} - (\alpha-1)\right]$$

where  $\lambda$  measures the smoothness of fields  $\epsilon$  and given as  $\lambda = 1/(2\sigma^2)$ . The resulting corrected  $p$ -values maps for  $F$  field is shown in Figure 6. The main use of the corrected  $p$ -value maps are the localization and visualization of thickness difference.

### 4 Application

$T_1$ -weighted MR scans were acquired for 16 autistic and 12 control subjects on a 3-Tesla GE SIGNA scanner. They are all right-handed males. 16 autis-

tic subjects were diagnosed with high functioning autism (HFA). The average age is  $17.1 \pm 2.8$  is for the control subjects and  $16.1 \pm 4.5$  for the autistic subjects. The complete description of the data set, image acquisition parameters, the subsequent image processing routines, and the interpretation of the resulting statistical parametric maps is provided in [5]. Each image underwent several image preprocessing steps. Image intensity nonuniformity was corrected using nonparametric nonuniform intensity normalization method [18]. Then using the automatic image processing pipeline, the image was spatially normalized into the Montreal neurological institute (MNI) stereotaxic space using a global affine transformation. Subsequently, an automatic tissue-segmentation algorithm based on a supervised artificial neural network classifier was used to classify each voxel as cerebrospinal fluid (CSF), gray matter, or white matter [13]. Brain substructures such as the brain stem and the cerebellum were removed automatically. Triangular meshes for inner and outer cortical surfaces were obtained by a deformable surface algorithm [14]. Such a *deformable surface* approach has the advantage that the surface topology can be fixed to be spherical and the deformation process can maintain a non-intersecting surface at all times, obviating the need for topology correction [9]. The mesh starts as an ellipsoid located outside the brain and is shrunk to obtain the inner cortical surface. Then the inner surface is expanded, with constraints, to obtain the outer cortical surface. The triangular meshes are not constrained to lie on voxel boundaries. Instead, the triangular meshes can cut through a voxel, which serves to reduce discretization error and partial volume effect. Thickness is measured using the natural anatomical homology between vertices on the inner and outer cortical surface meshes, since the outer surface is obtained by deforming the inner surface.

Afterwards, thickness measures are smoothed with heat kernel smoothing with parameters  $\sigma = 1$  and  $k = 200$  giving the effective smoothness of  $\sqrt{200} = 14.14$  mm. A surface-to-surface registration to a template surface was performed to facilitate vertex-by-vertex inter-subject thickness comparison. We have formulated it as a registration problem of two functional data on a unit sphere [20]. First a mapping from a cortical surface onto the sphere is established while recording the mapping. Then cortical curvatures are mapped onto the sphere. The two curvature functions on the sphere are aligned by solving a regularization problem that tries to minimize the discrepancy between two functions while maximizing the smoothness of the alignment in such a way that the pattern of gyral ridges are matched smoothly. This alignment is projected back to the original surface using the recorded mapping. This regularization mechanism produces a smooth deformation field, with very little folding. The deformation field is parameterized using a triangulated mesh and the algorithm proceeds in a coarse-to-fine manner, with four levels of mesh resolution. Figure 4 and Figure 5 illustrate the effectiveness of this surface registration algorithm by computing the probability of matching superior temporal sulcal fundi and central sulci.

After smoothing out thickness measurements, statistical analysis is performed following the procedures described in the previous section. The resulting corrected  $p$ -value map ( $< 0.1$ ) for the  $F$  statistic is projected onto the template

surface for visualization. Figure 6 shows statistically significant regions of cortical thickness between two groups. After removing the effect of age and total grey matter volume difference, the statistically significant regions of thickness decreases are highly localized at the right inferior orbital prefrontal cortex, the left superior temporal sulcus and the left occipito-temporal gyrus in autistic subjects.

## 5 Conclusions

This paper has introduced heat kernel smoothing and its statistical properties for data analysis on the cortical manifolds. The technique can be used in smooth out data that is necessary in the random field theory based multiple comparison correction. We have applied the methodology in detecting the regions of abnormal cortical thickness in a group of autistic subjects; however, the approach is not limited to a particular clinical population. The algorithm is implemented in MATLAB and freely available to download on the web <http://www.stat.wisc.edu/~mchung/software/hk>. A sample cortical mesh for a subject and its thickness measures can be also downloaded from the same website for other researchers.

## Acknowledgement

Authors wish to thank Kim Dalton and Richard J. Davidson of the Waisman Laboratory for Brain Imaging and Behavior, University of Wisconsin-Madison for providing the data illustrated in this study.

## References

1. A. Andrade, Kherif, J. F., Mangin, K.J. Worsley, A. Paradis, O. Simon, S. Dehaene, D. Le Bihan, and J-B. Poline. Detection of fmri activation using cortical surface mapping. *Human Brain Mapping*, 12:79–93, 2001.
2. N. Berline, E. Getzler, and Vergne M. *Heat kernels and dirac operators*. Springer-Verlag, 1991.
3. A. Cachia, J.-F. Mangin, Rivière D., D. Papadopoulos-Orfanos, F. Kherif, I. Bloch, and J. Régis. A generic framework for parcellation of the cortical surface into gyri using geodesic voronoï diagrams. *Image Analysis*, 7:403–416, 2003.
4. P. Chaudhuri and J. S. Marron. Scale space view of curve estimation. *The Annals of Statistics*, 28:408–428, 2000.
5. M.K. Chung, S. Robbins, Davidson R.J. Alexander A.L. Dalton, K.M., and A.C. Evans. Cortical thickness analysis in autism with heat kernel smoothing. *NeuroImage*, 2005 in press.
6. M.K. Chung and J. Taylor. Diffusion smoothing on brain surface via finite element method. In *Proceedings of IEEE International Symposium on Biomedical Imaging (ISBI)*, 2004.

7. M.K. Chung, K.J. Worsley, S. Robbins, and A.C. Evans. Tensor-based brain surface modeling and analysis. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume I, pages 467–473, 2003.
8. M.K. Chung, K.J. Worsley, S. Robbins, T. Paus, Taylor, J.N. J., Giedd, J.L. Rapoport, and A.C. Evans. Deformation-based surface morphometry applied to gray matter deformation. *NeuroImage*, 18:198–213, 2003.
9. A.M. Dale and B. Fischl. Cortical surface-based analysis i. segmentation and surface reconstruction. *NeuroImage*, 9:179–194, 1999.
10. C. Davatzikos and R.N. Bryan. Using a deformable surface model to obtain a shape representation of the cortex. *Proceedings of the IEEE International Conference on Computer Vision*, 1995.
11. S.E. Jones, B.R. Buchbinder, and I. Aharon. Three-dimensional mapping of cortical thickness using laplace’s equation. *Human Brain Mapping*, 11:12–32, 2000.
12. S.C. Joshi, J. Wang, M.I. Miller, D.C. Van Essen, and U. Grenander. On the differential geometry of the cortical surface. *Vision Geometry IV*, pages 304–311, 1995.
13. K. Kollakian. Performance analysis of automatic techniques for tissue classification in magnetic resonance images of the human brain. Technical Report Master’s thesis, Concordia University, Montreal, Quebec, Canada, 1996.
14. J.D. MacDonald, N. Kabani, D. Avis, and A.C. Evans. Automated 3-d extraction of inner and outer surfaces of cerebral cortex from mri. *NeuroImage*, 12:340–356, 2000.
15. P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 12:629–639, 1990.
16. S.M. Robbins. Anatomical standardization of the human brain in euclidean 3-space and on the cortical 2-manifold. Technical Report PhD thesis, School of Computer Science, McGill University, Montreal, Quebec, Canada, 2003.
17. S. Rosenberg. *The Laplacian on a Riemannian Manifold*. Cambridge University Press, 1997.
18. J.G. Sled, A.P. Zijdenbos, and A.C. Evans. A nonparametric method for automatic correction of intensity nonuniformity in mri data. *IEEE Transactions on Medical Imaging*, 17:87–97, 1988.
19. N. Sochen, R. Kimmel, and R. Malladi. A general framework for low level vision. *IEEE Transactions on Image Processing*, 7:310–318, 1998.
20. P.M. Thompson and A.W. Toga. A surface-based technique for warping 3-dimensional images of the brain. *IEEE Transactions on Medical Imaging*, 15:1–16, 1996.
21. G. Wahba. *Spline models for observational data*. SIAM, 1990.
22. F.-Y. Wang. Sharp explicit lower bounds of heat kernels. *Annals of Probability*, 24:1995–2006, 1997.
23. K.J. Worsley. Local maxima and the expected euler characteristic of excursion sets of  $\chi^2$ ,  $f$  and  $t$  fields. *Advances in Applied Probability*, 26:13–42, 1994.
24. K.J. Worsley, J.E. Taylor, F. Tomaiuolo, and J. Lerch. Unified univariate and multivariate random field theory. *NeuroImage*, 2005.