

# Evaluating Brain Anatomical Correlations via Canonical Correlation Analysis of Sulcal Lines

Pierre Fillard\*      Xavier Pennec\*      Paul M. Thompson†  
Nicholas Ayache\*

December 19, 2007

Corresponding author:

Pierre Fillard (Pierre.Fillard@sophia.inria.fr)

INRIA, Asclepios Project,

2004 route des Lucioles - BP 93,

06902 Sophia Antipolis Cedex, FRANCE.

tel: +33 4 92 38 76 22

fax: +33 4 92 38 76 69

---

\*INRIA Sophia Antipolis - Asclepios Project - France.

†Laboratory of Neuro Imaging, Dept. of Neurology, UCLA School of Medicine.

## Abstract

Modeling and understanding the degree of correlations between measures of brain structures is a challenging problem in neuroscience. Correlated anatomic measures may arise from common genetic and trophic influences across brain regions, and may be overlooked if structures are modeled independently. Here, we propose a new method to analyze structural brain correlations based on a large set of cortical sulcal landmarks (72 per brain) delineated in 98 healthy subjects (age:  $51.8 \pm 6.2$  years). First, we evaluate the correlation between any pair of sulcal positions via the total covariance matrix, a  $6 \times 6$  symmetric positive-definite matrix. We use Log-Euclidean metrics to extrapolate this sparse field of total covariance matrices to obtain a dense representation. Second, we perform canonical correlation analysis to measure the degree of correlations between any two positions, and derive from it a p-value map for significance testing. We present maps of both local and long-range correlations, including maps of covariation between corresponding structures in opposite hemispheres, which show different degrees of hemispheric specialization. Results show that the central and superior temporal sulci are highly correlated with their symmetric counterparts in the opposite brain hemisphere. Moreover, several functionally unrelated cortical landmarks show a high correlation as well. This structural dependence may be attributable to common genetic programs controlling growth rates, experience-driven plasticity, and coordinated brain growth in the presence of anatomical and functional links along neural pathways.

# 1 Introduction

Understanding correlations between measures from different brain structures is a challenging problem in neuroscience. Most computational anatomic studies of development and disease study deficits or changes by modeling individual brain structures independently, often creating voxel-based maps of group anatomical differences. This reveals factors (e.g., age or disease) that influence each brain structure individually, but may miss important supra-regional correlations, such as brain regions that develop or fail to develop together, or correlations between structures in opposite brain hemispheres. Mechelli et al. [1] discovered (1) spatial correlations in gray matter density between corresponding regions in opposite brain hemispheres, except in the visual cortex, and (2) some negative correlations between functionally distinct regions within the same hemisphere. Neuroscientists have been interested in identifying the reasons for such long-range brain correlations, and what may cause them at a genetic and environmental level. Normally coordinated systems may be disrupted in neuropsychiatric disorders such as schizophrenia, autism, and Williams syndrome, and systems with known fiber or functional connectivity may exert mutual influences on each other's growth, via activation-dependent plasticity and pruning. Structures in different lobes may also influence each other's degeneration, for instance when loss of neuronal input from one brain region is associated with secondary degeneration in another. Despite many hypotheses, few tools have been developed to measure such long-range correlations, making them difficult to study empirically. The work by Lerch et al. [2] describes a study of the potential correlations between brain regions by looking at cortical thickness. Here, we look at correlation between brain regions by looking at anatomical positions (i.e., correlations between cortical positions in space). These two sources of correlations are complementary and could give different results (correlations of cortical thickness are not equivalent to correlations of anatomical positions). The comparison of both, however, will not be developed here.

An efficient, parsimonious model of the complex patterns of brain correlations may help to identify factors that influence them. Inter-hemispheric correlations

(i.e., correlations between point in anatomically homologous structures in both hemispheres) are of interest as they may shed light on how functional regions in contralateral areas become specialized. Improved modeling of correlations of brain structures could make it easier to isolate specific effects of genetic polymorphisms on these normal correlations. Furthermore, information on statistical correlations could reduce the difficulty of automated segmentation and labeling of brain structures. Accessing anatomical correlations also opens up a broad range of studies and comparing groups (e.g., disease versus normal) and a new path to generate hypotheses regarding patterns of brain growth.

We previously built first-order models (mean anatomical templates), and second-order models (variability models) [3] of the sulcal lines to capture the 3D variations of each anatomical point independently around a mean anatomical model of the cortex, after linear registration of multi-subject anatomical images to a common reference space. These variations are often represented by covariance matrices, or variability tensors, as variations may not be of the same magnitude of likelihood in all directions. Here, we go one step further and model the joint variability of all pairs of anatomical points, to see how the displacement of any point in a specific subject with respect to a reference anatomy covaries with the displacement of neighbouring or distant points (e.g., anatomically homologous regions in the opposite hemisphere).

In Section 2, we introduce the main tool of our analysis: the total covariance matrix (TCM) between two vector variates, and we recall how to extract from it several matrix and scalar measures to test if these two variables are correlated. In Section 3, we experiment with this framework using TCMs defined from anatomical landmarks (sulcal curves). We start by studying the TCMs between 6 sulcal positions and locations in the rest of the brain, which eventually leads us to analyze the TCMs of all sulcal positions in one hemisphere with their homologous counterparts in the opposite hemisphere.

## 2 The Total Covariance Matrix

### 2.1 Definition

Let  $X = \{X_i\}_{i=1..N}$  and  $Y = \{Y_i\}_{i=1..N}$  be the sets of  $N$  measures of two random vectors whose dimensionality is  $d$ . Computing the correlation between  $X$  and  $Y$  requires us to know not only the variability of each vector (i.e., its covariance matrix), but also their cross-covariance. We therefore define the TCM of  $X$  and  $Y$  that contains this information as  $\Lambda(X, Y)$ :

$$\Lambda(X, Y) = \frac{1}{N-1} \sum_{i=1}^N \begin{pmatrix} X_i - \bar{X} \\ Y_i - \bar{Y} \end{pmatrix} \begin{pmatrix} X_i - \bar{X} \\ Y_i - \bar{Y} \end{pmatrix}^\top, \quad (1)$$

where  $\bar{X} = (\sum_1^N X_i)/N$  and  $\bar{Y} = (\sum_1^N Y_i)/N$ . We denote by  $\Sigma_{XX}$  (resp.  $\Sigma_{YY}$ ) the covariance matrix of  $X$  (resp.  $Y$ ):  $\Sigma_{XX} = E[(X - \bar{X})(X - \bar{X})^\top]$ . The cross-covariance of  $X$  and  $Y$  is given by:  $\Sigma_{XY} = E[(X - \bar{X})(Y - \bar{Y})^\top]$ . By further developing Eq. 1, one can write  $\Lambda(X, Y)$  in a simpler way:

$$\Lambda(X, Y) = \begin{pmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_{YY} \end{pmatrix}. \quad (2)$$

$\Lambda$  is a  $2d \times 2d$  matrix. It has the same properties as a classical covariance matrix: it is symmetric and positive definite. Therefore, we may also call  $\Lambda$  a tensor. In the 3D case,  $\Lambda$  is a  $6 \times 6$  tensor.

### 2.2 Analysis of Total Covariance Matrices

In its current form, it is difficult to appreciate the meaning of the TCM and it cannot be easily visualized (it is an ellipsoid in 6D). However, several matrix, vector and scalar measures may be derived from it. Here, we focus on quantifying the correlation of  $X$  and  $Y$  through Canonical Correlation Analysis.

**Canonical Correlation Analysis (CCA):** CCA [4, 5] refers to the method of

finding vector bases that maximize the correlation between two vector variates, and is the generalization of the correlation coefficient to multivariate data. In the scalar case, we define the correlation coefficient between  $x$  and  $y$  as:  $\rho = \sigma_{xy} / \sqrt{\sigma_{xx} \cdot \sigma_{yy}}$ , where  $\sigma_{xx}$  (resp.  $\sigma_{yy}$ ) is the variance of  $x$  (resp.  $y$ ), and  $\sigma_{xy}$  is the cross-variance of  $x$  and  $y$ . Similarly, the correlation matrix  $\Gamma$  in the multivariate case is defined as:

$$\Gamma(X, Y) = \Sigma_{XX}^{-1/2} \Sigma_{XY} \Sigma_{YY}^{-1/2}. \quad (3)$$

We have the property that  $\Gamma(Y, X) = \Gamma(X, Y)^\top$ . Taking the mean trace of the  $\Gamma$  gives us an average correlation coefficient  $\bar{\rho}$ . The range of  $\bar{\rho}$  lies between 0 (absence of correlation) and 1 (correlation). In multivariate statistics, however, a low average correlation coefficient may hide a potentially high correlation. This is the case, for instance, when only one component of  $X$  is correlated with one component of  $Y$ . Taking the average correlation coefficient may discard this relationship. To distinguish between correlations along potentially different axes, one needs to run a canonical correlation analysis, which is nothing other than decomposing  $\Gamma$  in terms of its singular values:  $\Gamma = U \cdot S \cdot V^\top$ , where  $U$  and  $V$  are orthogonal matrices of correlation vectors, and  $S$  is a diagonal matrix of correlation coefficients  $\rho_i$ .

**Significance Testing:** To test the statistical significance of the obtained correlations, [6] proposed to test the dimensionality of the correlation matrix. If its rank is zero, then there is no correlation ( $\rho_i = 0, \forall i$ ). If we reject this hypothesis, then the rank is at least one, which means that at least two directions in space are correlated. We use the Bartlett-Lawley test [6] with the null hypothesis:  $H_0 : \text{rank}(\Gamma) = 0$ :  $L(\Gamma) = -\left(N - d + \frac{1}{2}\right) \sum_{j=1}^d \log(1 - \rho_j^2)$ .  $L$ 's distribution is chi-squared under the Gaussian assumption on  $X$  and  $Y$  with  $d^2$  degrees of freedom. We can consequently derive a p-value for testing the significance of correlations.

### 3 Experiments

We used a dataset of sulcal landmark curves manually delineated in 98 subjects by expert anatomists, according to a precise protocol with established reliability within and across raters [3]. The dataset consists of sulcal landmarks from 46 men and 52 women (age:  $51.8 \pm 6.2$  years), all normal controls. The lines were traced in 3D on the cortical surface. We included the maximal subset of all curves that consistently appear in all normal subjects, 72 in total (36 per hemisphere). By repeated training, the maximum allowed inter- and intra-rater error (reliability) was ensured to be no greater than 2mm everywhere. MR images used for delineations were first linearly aligned to the ICBM stereotactic space [7].

We used the methodology outlined in [3] to determine the mean curve for each sulcal line by modeling samples as deformations of a single average curve. Mean curve computation involves filtering each sample by B-spline parameterization, minimization of total variance, and sulcal matching by dynamic programming.

In the following, we investigate the potential correlations between locations on different sulci. First, we study the correlation between particular sulcal lines and other cortical points not belonging to the same structure: we call this study *sulcal correlation*. Second, we assess *inter-hemispheric correlations* between corresponding anatomical points in the two hemispheres.

#### 3.1 Sulcal Correlation for 6 Specific Positions

**Methodology:** Obviously, this study is a combinatorial challenge. We sampled the 72 mean sulci with approximately 500 points (average of 7 points per sulcus), which gives a total of about 250000 pairs of points to process. To limit the number of pairs investigated, we focused on two major sulcal lines: the Central Sulcus (CS) and the Superior Temporal Sulcus (STS). The CS is a relatively stable landmark, dividing the frontal and parietal lobes, and separates primary sensory and motor cortices. The STS is an important landmark in studies of visual motion processing, and its posterior limit has been regarded as a sulcal landmark with a relatively tight association with the visual motion area V5/MT in the lateral temporal cortex [8, 9].

These sulci lie in different lobes, develop at different times during cortical gyrogenesis (the CS developing earlier) and are distant in terms of fiber and functional connectivity, so they are good candidates for assessing inter-structure correlation, as little correlation is expected *a priori*. For each of these lines, three reference positions were picked: the beginning, middle, and end point. First, for each of the three reference positions, we extract the set of corresponding sulcal positions in each of the 98 subjects. Second, we computed the TCM of Eq. 2 with each of the remaining 499 average sulcal positions. We end up with a sparse field of TCMs. However, we would be more useful to obtain a dense field of TCMs, as we could map those onto an average cortical model to make it easier to interpret the results visually. We use Log-Euclidean (LE) metrics [10] and the methodology developed in [3] (combination of a radial basis function interpolation with an harmonic partial differential equation) to extrapolate TCMs on a mean cortical surface. This type of interpolation was shown to preserve all the features of a covariance matrix, and has desirable properties like absence of the so-called “swelling” effect, and a smooth interpolation of the eigenvectors. Moreover, leave-one-out tests showed that this interpolation was able to predict missing data in regions locally correlated and is consequently well adapted for TCMs. The correlation matrix and the p-value derived from the CCA can now be computed at any point of the mean cortex. Notice that even if we only focused here on the p-value defined in Sec. 2.2, other measures are interesting, such as the principal vectors of correlation which are currently under investigation. This is why we need to extrapolate the full TCMs and not just the p-value.

The main problem for the curve matching procedure proposed in [3] is the *aperture problem*: correspondences in the direction tangent to the curve are almost impossible to retrieve without additional constraints based on expert knowledge. To keep our results independent of this, we need to cancel the contribution in this tangential direction. The method proposed in this paper is as follows: we define at each position of the mean sulci the Frenet frame, which gives us the plane orthogonal to the curve. Then, we project the sulcal positions onto this plane, which zeroes out the tangential component (see Fig. 1). Note that we lose one degree of freedom in



the dimensionality of the data: vectors no longer have three degrees of freedom but two. This must be accounted for in the statistical tests of Sec. 2.2.

**Results:** A significance map showing the p-values derived from CCA is shown in Figure 2. The significance level was set to 0.0001 to correct for multiple comparisons (Bonferroni correction applied with a threshold at 5% with 500 observations:  $0.05/500 = 0.0001$ ). Notice that as our statistics are irregularly sampled in the 3D space, it is difficult to use standard methods based on the theory of Gaussian random field [11] to correct for multiple comparisons. To obtain non arguable results, we chose the most conservative approach (Bonferroni’s correction) even if this considerably lowers the powerfulness of the test.

A large area around the reference points shows high p-values: as expected, points that are anatomically close to the reference do have a correlated distribution among individuals. More interestingly, regions with high p-values most often include the structures’ opposite hemisphere counterparts, but not always: the first (upper) and middle CS positions are highly correlated in the other hemisphere (Fig. 2 top panel, 1st and 2nd row), while the most inferior position is not, most likely because its variability across subjects is extremely low (Fig. 2 top panel, 3rd row). In right-handed subjects, we know that some measures of motor skill (maximum finger tapping rate) correlate with gray matter volume positively in the left CS, but negatively in the right CS [12], which is plausible given the left hemispheric specialization for fast repetitive movements in right-handers. Logically, such functional specializations may promote correlations between the two hemispheres in these regions of the CS, where tissue volumes are known to directly correlate with skill levels. The posterior part of the STS shows lowest correlation with its opposite hemisphere counterpart. Unlike the bottom of the CS, the posterior tip of the STS is highly variable and asymmetric in structure and function - it is specialized, to some extent, for understanding the semantics of language in the left hemisphere, but for understanding prosodic aspects of language in the right hemisphere (see, e.g. [13]) This may suggest partially independent developmental programs for these functionally specialized structures. The long-range correlation between the posterior ramus of the STS and the left and

right intra-parietal sulci is of interest, as the planum temporale and planum parietale are two distinct areas that have been extensively examined in structural studies of hemispheric asymmetry.

Nevertheless, it is intriguing that 5 of the 6 sulcal positions studied show a correla-

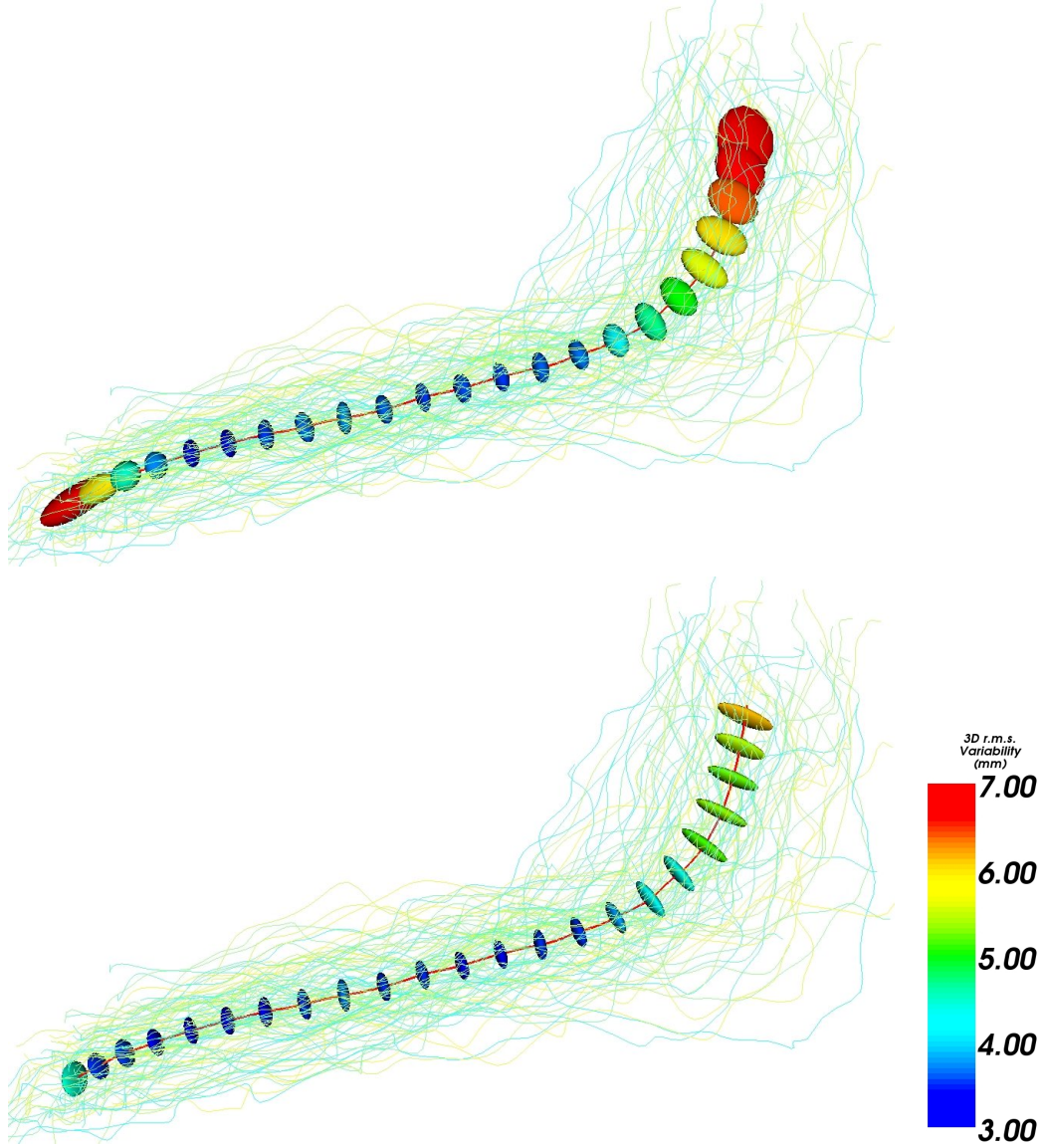


Figure 1: **Canceling the tangential component.** **Left:** Covariance matrices extracted along the mean Sylvian Fissure as in [3]. **Right:** The same tensors after canceling the tangential component. A small tangential value is kept for visualization purposes (tensors should be flat).

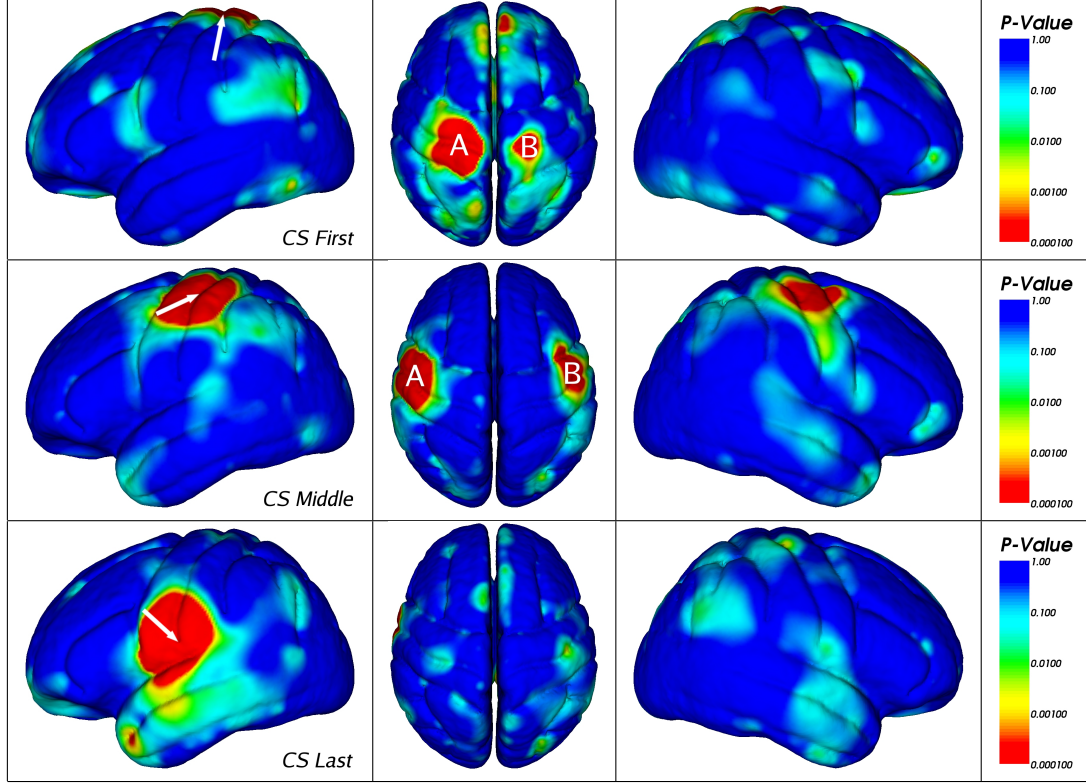


Figure 2: **Correlation Maps between Specific Reference Points and Other Brain Regions.** **Top panel: the Central Sulcus.** The significance level was set to 0.0001 for multiple comparisons correction. A white arrow in each row indicates a reference landmark; correlations with the reference landmark are plotted. Correlations for 3 reference landmarks on the CS are shown: the first (top row), the middle (second row), and the last, i.e. most inferior, position (third row) on the sulcal trace. Corresponding regions in the opposite hemisphere are highly correlated for the top and middle points (marked A and B). The lower end of the sulcus, however, exhibits low correlation with its symmetric contralateral counterpart.

tion with their symmetric counterpart in the opposite hemisphere. In the following, we test whether this observation can be generalized to all sulcal positions.

### 3.2 Inter-Hemispheric Correlation Analysis

In this study, we specifically target the correlation between all points of one hemisphere and their homologous region in the other hemisphere. To do so, we first map all sulci of the right hemisphere onto the left. Then, we define a global mean, i.e. an average sulcal curve computed from the 98 left and right samples (Fig. 4). Global

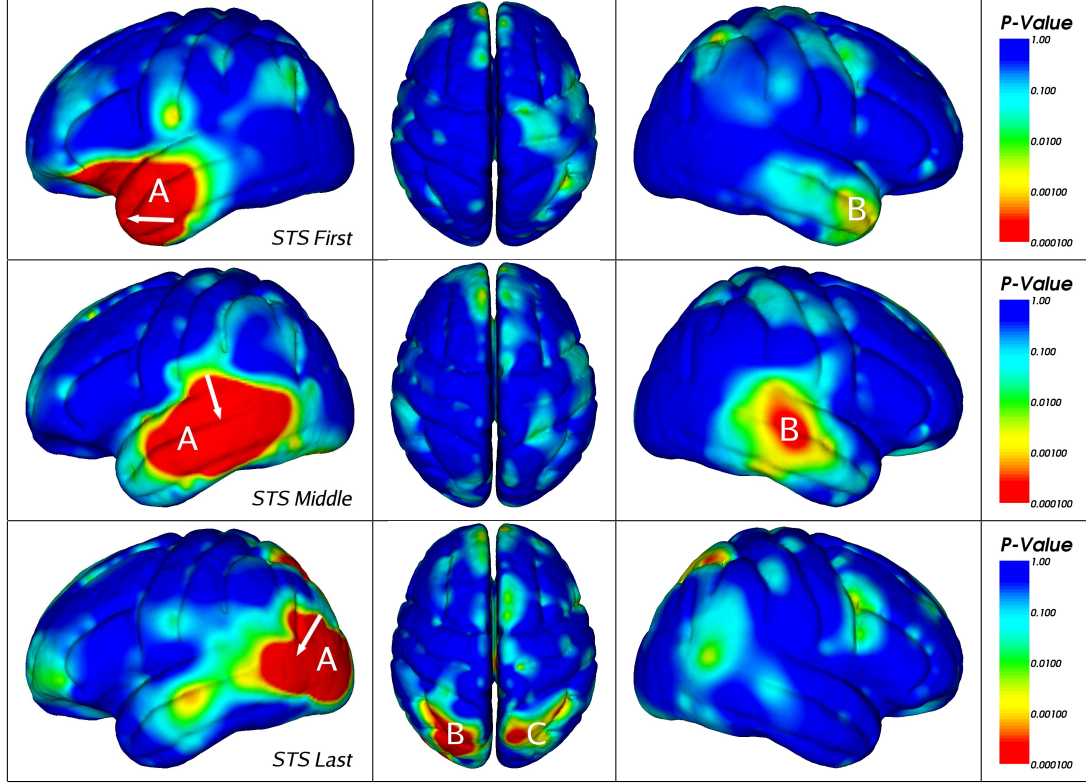


Figure 3: **Maps of Anatomical Correlation between Landmarks in the Perisylvian Language Cortex and other Brain Regions.** The significance level was set to 0.0001 for multiple comparisons correction. The same 3 positions as in Fig. 2 for the Superior Temporal Sulcus are analyzed. The first (top row) and middle (second row) positions are symmetrically correlated (marks A and B). The last position (third row) correlates less with its opposite hemisphere counterpart, than with the intra-parietal sulci (marked B and C).

means provide a common reference curve to compare left and right positions. Correspondences between global means and left and right average curves are computed using the same framework as for the samples. For any given position on the global mean, we obtain corresponding points on left and right average curves, giving in turn correspondences between left and right sulcal positions in the 98 subjects (the choice of the left hemisphere is arbitrary, and we obtained the same results when using the right hemisphere instead). Finally, we compute the TCM of Eq. 2 between left and right positions. As for the sulcal correlation study, we extrapolate this sparse field of TCMs to an average left hemisphere surface, and cancel the tangential component

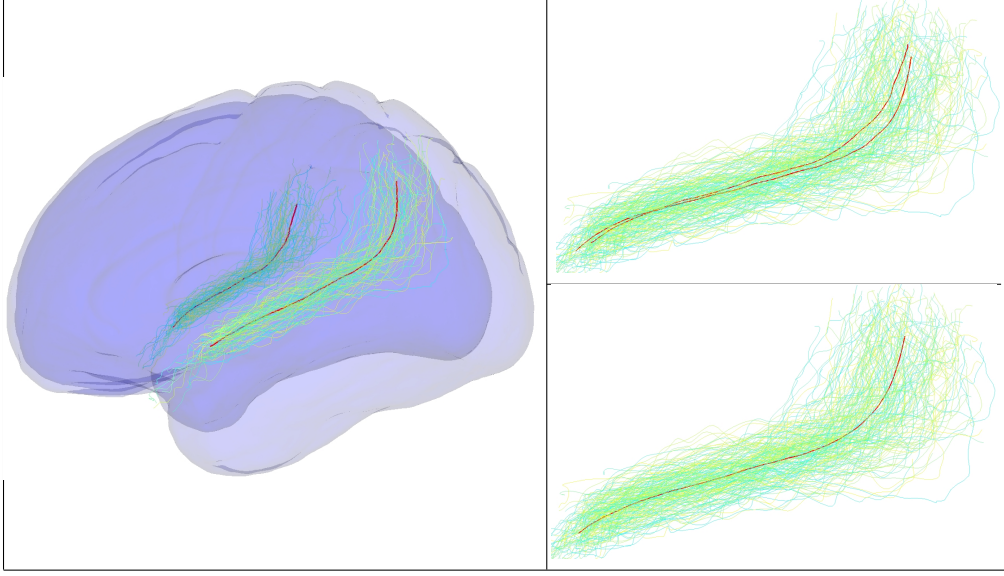


Figure 4: **Finding hemispheric correspondences.** The significance level was set to 0.0001 for multiple comparisons correction. **Left:** Positions of left and right Sylvian fissures (SF), in an oblique view, with a transparent cortical surface overlaid (in blue colors). **Top Right:** All right hemisphere sulci are mapped to the left, including the two mean curves (in red). **Bottom Right:** Same set of sulci, whose global mean was computed and displayed in red. This is now used as a reference to compare left and right sulcal positions.

which is consistently biased towards zero. Then, we extract p-values with the CCA and map those on the surface (Fig. 5). We apply the same Bonferroni correction as for the previous experiments and lower the significance level to 0.0001.

As observed previously on a few specific sulcal positions, most points are correlated with their symmetric counterparts. Regions with lowest correlations include Broca’s and Wernicke’s areas, which have already been shown to exhibit the greatest asymmetries in variability [3]. These cortical regions are specialized for language production and comprehension respectively, but most right-handers show a greater reliance on the left hemisphere for language processing, and the volumes of these regions are highly asymmetric between hemispheres.



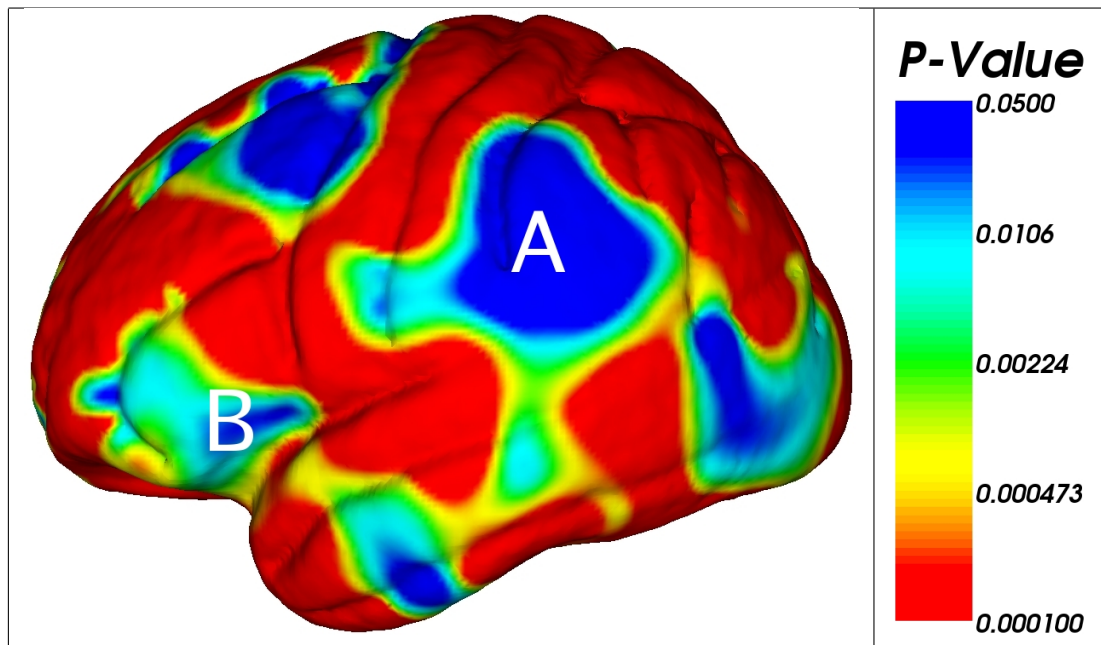


Figure 5: **Hemispheric correlations.** Hot colors (red) correspond to  $p\text{-value} \leq 0.0001$ . Most cortical points show anatomical variations that are correlated with their counterparts in the opposite hemisphere. Un-correlated regions include Wernicke's (marked A) and Broca's areas (marked B), which, in most subjects, are known to be more heavily specialized for language processing in the left hemisphere.

## 4 Conclusions and Perspectives

In this paper, we represent the cross-covariance between one point and any other point of the brain by a total covariance matrix describing not only the variability of the two points, but also their cross-covariance. Canonical correlation analysis allows us to test for the significance of these correlations. As TCMs have the same properties as classical covariance matrices (they are symmetric and positive-definite), we used the Log-Euclidean extrapolation to obtain a dense representation of initially sparsely-defined measures. This extrapolation was indeed previously shown to have good properties [3], contrary to the Euclidean extrapolation of the matrix coefficients.

We apply this method to study sulcal and hemispheric correlations. We showed that the central sulcus was highly correlated with its symmetric counterpart, except

in its inferior part, which is not highly variable anyway. For the central sulcus, where motor skill is correlated with volume and is also lateralized, a strong hand preference for motor skills is likely to promote negative correlations between hemispheres for volumes in opposite regions. The Superior Temporal Sulcus shows similar intriguing correlations, and its low correlation with its opposite counterpart may reflect their different developmental programs and functions.

Corresponding brain regions in each hemisphere are highly correlated, except for regions including Wernicke's and Broca's areas, which are known to be functionally specialized in one hemisphere. Any longer-range correlation - such as that found between the intra-parietal sulci and superior temporal sulci - is in itself an interesting neuroanatomical finding. The planum parietale and temporale are distinct highly asymmetric systems in each of these regions, and long-range correlations may reflect common factors driving programmed asymmetries for both regions.

Future work will include a more extensive modeling of these correlations and factors that might cause them. We could store, for each point of the brain, a minimal set of correlated positions, such as the local neighborhood and the set of distant most correlated points. This information could be used as a prior to guide inter-subject non-linear registration algorithms. Furthermore, a detailed study of all possible correlations between cortical landmarks could help understand the effects of genes on brain maturation. Validations of long-range correlations could be made using other sources of information, such as functional MRI: Are jointly activated regions, or causal models (achieved through structural equation modeling or dynamic causal modeling) for a given task related to anatomical correlations as well as connectivity (fiber bundles)? Such information on long-range anatomical correlations could contribute to our understanding of the functional organization of the brain. Finally, these results should be compared with those of other methods, such as surface-based versus volumetric registration algorithms; this comparison is currently underway for generating second-order models of brain variability [14].

## References

- [1] Mechelli, A., Friston, K.J., Frackowiak, R.S., Price, C.J.: Structural covariance in the human cortex. *The Journal of Neuroscience* **25**(36) (2005) 8303–8310
- [2] Lerch, J.P., Worsley, K., Shaw, W.P., Greenstein, D.K., Lenroot, R.K., Giebb, J., Evans, A.C.: Mapping anatomical correlations across cerebral cortex (macacc) using cortical thickness from mri. *NeuroImage* **31** (2006) 993–1003
- [3] Fillard, P., Arsigny, V., Pennec, X., Hayashi, K., Thompson, P., Ayache, N.: Measuring brain variability by extrapolating sparse tensor fields measured on sulcal lines. *Neuroimage* **34**(2) (2007) 639–650 Also as INRIA Research Report 5887, April 2006. PMID: 17113311.
- [4] Hotelling, H.: Relations between two sets of variates. *Biometrika* **28** (1936) 231–377
- [5] Rao, A., Babalola, K., Rueckert, D.: Canonical correlation analysis of sub-cortical brain structures using non-rigid registration. In: *Proc. of WBIR*. (2006)
- [6] Fujikoshi, Y., Veitch, L.: Estimation of dimensionality in canonical correlation analysis. *Biometrika* **66**(2) (1979) 345–351.
- [7] Collins, D., Holmes, C., Peters, T., Evans, A.: Automatic 3D model-based neuroanatomical segmentation. *Human Brain Mapping* **3**(3) (1995) 190–208
- [8] Watson, J.D.G., Myers, R., Frackowiak, R.S.J., Hajnal, J.V., Woods, R.P., Mazziotta, J.C., Shipp, S., Zeki, S.: Area v5 of the human brain: Evidence from a combined study using positron emission tomography and magnetic resonance imaging. *Cerebral Cortex* **3** (1993) 19–94
- [9] Dumoulin, S.O., Bittar, R.G., Kabani, N.J., Baker, C.L., Goualher, G.L., Pike, G.B., Evans, A.C.: A new anatomical landmark for reliable identification of human area v5/mt: a quantitative analysis of sulcal patterning. *Cerebral Cortex* **5** (2000) 454–463
- [10] Arsigny, V., Fillard, P., Pennec, X., Ayache, N.: Log-Euclidean metrics for fast and simple calculus on diffusion tensors. *MRM* **56**(2) (2006) 411–421



- [11] Taylor, J., Worsley, K.: Detecting sparse signal in random fields, with an application to brain mapping. *Journal of the American Statistical Association* **102**(479) (2007) 913–928
- [12] Herve, P., Mazoyer, B., Crivello, F., Perchey, G., Tzourio-Mazoyer, N.: Finger tapping, handedness and grey matter amount in the rolando’s genu area. *NeuroImage* **25**(4) (2005) 1133–45
- [13] Cabeza, R., Nyberg, L.: Imaging cognition ii: An empirical review of 275 pet and fmri studies. *J. of Cognitive Neuroscience* **12** (2000) 1–47
- [14] Durrleman, S., Pennec, X., Trouvé, A., Ayache, N.: Measuring brain variability via sulcal lines registration: a diffeomorphic approach. In: *Proc. of MICCAI’07*. (2007) To appear.