bioRxiv preprint doi: https://doi.org/10.1101/2020.02.18.955096; this version posted March 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

"Sulcus Sink": A Compact Binary and Semi-Automated Inverse Dijkstra-based System for Describing Sulcal Trajectories

Rudolph Pienaar[1,3†], Christian Hasselgrove[4], Kiho Im[1,3], David Kennedy[4], P Ellen Grant[1,3], Denise Boriel[2], Lena Tang[5], Nikos Makris[2,3]

[†] Corresponding author: Rudolph Pienaar, rudolph.pienaar@childrens.harvard.edu, +1-781-640-1096
[1] Children's Hospital Boston, Radiology Department, 300 Longwood Ave, Boston, MA, 02115
[2] Center for Morphometric Analysis, Massachusetts General Hospital, Neurology Department, 149 13th Street, Charlestown, MA, 02129
[3] Harvard Medical School, 25 Shattuck Street, Boston, MA, 02115
[4] Child and Adolescent NeuroDevelopment Initiative (CANDI), Department of Psychiatry, University of Massachusetts Medical School Biotech One, Suite 100 365 Plantation Street Worcester, MA 01605
[5] Respiratory Acute Care Unit, Department of Anesthesia, Massachusetts General Hospital, 55 Fruit Street, Boston, MA, 02115

Abstract

We present a description of a system that uses a compact binary representation to describe and trace sulci on a reconstructed human cortical surface, based on a set of human-generated targets. The inputs to the system were manually created on a training set of 20 normal subjects (11 females, 9 males) with ages 22 – 40 years. T1 weighted MPRAGE images were collected on a Siemens 3T Trio scanner, with TR/TE = 2530/3.3, matrix = 256x256, FOV = 256mm, slice thickness 1.33mm. The resultant images were reconstructed with Freesurfer, and 10 sulci on each hemisphere were traced by an expert human operator and independently assessed for accuracy. Presented with these input trajectories in its training phase, the system attempted to determine a compact binary feature vector of each sulcus on each subject using as descriptor a binary parametrized function of several surface-geometry variables (such as mean curvature, sulcal depth, edge length, etc.). This function was optimized in a supervised learning fashion using a Dijkstra-based graph theory formulation, in which the binary weights were used to define graph edge costs. In the setup phase, the system was presented with sulcal trajectories already defined on surfaces, and then adjusted its parametrized weights in a binary fashion to minimize differences between the training input path and its Dijkstra-generated output path. Once the setup phase was complete and sulci had been described in a per-sulcus, per-subject manner, we generalized the per-sulcus description across all the subjects to construct a template binary word for each specific sulcus. The performance of the system for each subject and each sulcus, and for each template sulcus group was measured against the original human reference in both a quantitative and qualitative manner. Individual subjects generally showed very good optimization to their manually traced training samples across all sulci, with 91% average overlap within 4mm of the human target. Generalized group results, as expected, showed less overlap with the original human targets, but still performed on average with 80% overlap. Quantitatively, the group results were nonetheless for the most part quite acceptable to an independent human evaluator. The parametrized binary weight description that drives the Dijkstra path optimization is presented as a mechanism to succinctly and compactly describe individual human sulci and groups of sulci.

Keywords: sulcus-tracing, Dijkstra's algorithm, optimization, sulcus-description

I. Introduction

THIS paper presents a description of a methodology called "sulcus sink" that has been developed at the Massachusetts General Hospital's (MGH) Center for Morphometric Analysis (CMA). Sulcus sink attempts to describe sulci using a compact binary parametrized description. Presented with a human generated reference sulcal trace, the method searches in a binary fashion through a parametrized weight space of graph edge costs and uses Dijkstra's algorithm to optimize the sulcal traces that result from a given weighting. The system attempts to find a weight parametrization that minimizes the error between the human reference and Dijkstra-generated trajectories.

This methodology arose from the very practical need to aid human operators as they manually traced sulcal lines on reconstructed surfaces. The operational use of the system would allow human operators to simply and relatively quickly select sulcal start and terminal points on a reconstructed surface. The sulcus sink system, trained on a cohort of human generated references, would complete the sulcal trace, based on a generalization of earlier optimization analyses.

The purpose of the system is therefore not to automatically extract sulci, but rather to attempt to capture the relative weighting that human operators might place on components of cortical geometry as they generate manual traces, and encode this information in a compact binary word. An important implication of such a weight parametrization description is that it allows us to attempt to propose unbiased sulcal-specific descriptions that have been generalized from an underlying set of per-subject and per-sulcus optimal parametrization. Given such a binary parametrization that weighs graph edges between vertices of the reconstructed cortical mesh, the optimization engine searching through this graph-based representation naturally lends itself to processing by Dijkstra's algorithm [7]. The system is considered "semi-automated" in that it does

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.18.955096; this version posted March 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

require, for a given sulcus, start and terminus points. A complete description, therefore, is the binary parametrization, sulcus start and terminus points, and Dijkstra's algorithm on resultant edge costs.

While the emphasis of the paper is on the technical and operational aspects of the methodology itself, preliminary results on the semi-automated tracing of sets of twenty sulci on twenty subjects is also presented. Based on the characteristics of individually optimized sulcal traces, we also attempt to generalize over all similar sulci in all the subjects.

The automatic (or semi-automatic) tracing of sulci on the cortical surface is useful in a number of neurological applications, from surgical planning to cortical parcellation and morphometric analysis [10, 25, 16, 26, 9, 20], brain registration [13] and exploring brain development during pathologies [39].

A large volume of literature seeks to address the problem of automatically detecting sulci based around geodesics generated by various methods – from dynamic programming [17] using local shape measures (but still requiring the manual selection of sulcal start and end points), through the optimal calculation of geodesics on triangulated mesh surfaces [2] to generate isotropic geodesics. Such geodesics have also been computed by a skeleton extraction approach that solves an Eikonal equation [35] – itself a special type of a Hamilton-Jacobi equation [36]. A refinement to such isotropic geodesics has recently been extended to anisotropic conditions which are more highly adaptive to the anisotropic nature of cortical surface shape and which can lead to sulcal lines lying more accurately within sulcal valleys [34]. Statistical methods based on mean curvature values have also been considered [38] – these require surface registration and statistical training sets.

Geodesic approaches are intuitively appealing since they map to our understanding of the geometry of surfaces. Indeed, a geodesic is by definition a locally length minimizing curve. However, in order to better explore a more abstract parametrized description (including not only curvature, sulcal depth, and distance but also products of the same), we believe a more graph-theoretical approach wherein edge weights can be arbitrarily (and possibly less intuitively obviously) defined affords us better representational power. Such a graph-theoretical approach is ideally suited to Dijkstra's algorithm.

The geodesic and statistical approaches for the most part attempt to solve or model geometric properties using a relatively small subset of available geometric properties – often the mean curvature. Extending these approaches to less intuitive geometric properties or considering the behavior of these systems under such circumstances is often not explored. Our approach here is to explicitly formulate our problem from the perspective of generating a compact binary weight description of multiple geometric properties, and to use this formulation as a descriptor of a given sulcus.

Using Dijkstra's algorithm [7] in this manner as an optimizer for tracking human-traced sulcal paths is a relatively novel approach. In the context of cortical parcellation, Dijkstra's algorithm has previously been used to calculate lengths on the cortical surface [3], and also as a basis for exploring geometric invariants in Hurdal's analysis in classifying cortical sulci [22]. Conceptually, there are some broad similarities between our approach and Hurdal's, however Hurdal attempts to describe sulci using invariant shape descriptors such as the writhe number, ropelength, and moment-based measures of a sulcal trajectory curve. Our work here parametrizes sulci not as complete curves, but rather as a vector of weights that define a cost function to a Dijskstra-based path search. The weight vector uses features calculated from the reconstructed cortical surface, primarily the mean curvature (using methods of discrete geometry [23][37][8]), distance between vertices, direction to end-point, etc.).

In some respects, this work touches on issues relating to the complexity of folded cortical surfaces. Several studies have previously attempted to quantify folding and/or complexity. Fractal dimension analysis has often been used as a technique for measuring "complexity" (14; 18; 15), as well as folding measures derived from cortical thickness analysis (41; 40) and metric distortions that arise from registering surfaces to an average template [42]. Although fractal measures are interesting in the abstract, we find them difficult tools to use as a means of understanding the topology of cortical folding, and hence sulcal paths. Numerous studies based on cortical thickness employ regional shape measures based on functions of mean curvature.

Other means of describing sulci using Mangin's "sulcal roots" have also been proposed. These "sulcal roots" correspond to the first folding locations during antenatal life (4; 5; 27; 28; 29; 33; 1) and are objects derived from mean curvature minima and saddle points. The highly variable pattern of folding noted across adult brains can be recovered from successive scale-space analysis. Folding patterns are decomposed and their core embedded patterns simplified. Conceptually, Mangin's work constructs "tree-like" structures tracing the development of simple folds to complex pattern at different scales. Our work is a complement to this approach. Sulcal roots provide a means for logically constructing how simple folds become complex patterns without focusing on the detailed curvature characteristics of the final patterns. Our work is less concerned with the exact location of the maximal sulcal depth and its branching pattern, and more focused on the surface topology which we analyze using several curvature functions.

Additional studies have attempted to consider the intrinsic geometry of the cortex from a more purely mathematical and computational basis [12], but followup studies in a similar vein using contemporary computing power have not been pursued. Moreover, these studies have focused on adult, not pediatric or developing newborn brain surfaces.

More recently, spherical wavelets have also used to quantify cortical folding [43, 31]. Development of surface folding was modeled through increasing wavelet powers and these wavelet coefficients were fitted to the Gompertz function, a model of self-limited growth. This allows predictions as the when sulcal trajectories might occur during development, and

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.18.955096; this version posted March 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

sulcus name	abbreviation	aspect
calcarine	calc	medial
central	ce	lateral
cingulate	ci	medial
inferior frontal	if	medial
olfactory	ol	lateral
post. horizontal ramus of Sylvian	phr	lateral
parieto-occipital	po	inferior
pre-central	prc	lateral
superior frontal	sf	superior
superior temporal	st	lateral

Table I
A LIST OF SULCI THAT WERE MANUALLY TRACED, AND SUBSEQUENTLY COST-FUNCTION OPTIMIZED.

also provides measures as to the differential folding rate across developing surfaces. The Gompertz function has been used successfully in the past to describe brain volumetric growth (30; 19). A more direct methodology based on surface curvature features and functions of principle curvature has also been proposed to quantify the shape structure of the cortical manifold [32].

From an organizational perspective, Section II focuses on the experimental methods, describing the data set itself, details on the segmentation and the reconstruction process, and the sulci that were traced. Section III describes the Dijkstra based optimization process, and the sulcus sink system itself. The weight vector space, and search methods through this space are presented, and the creation of general templates from a set of optimal weights is considered. Section IV presents and discusses the results of the optimization, with specific examples of sulcal correlation to the training templates, as well as the generalization of the sulcal weights across all subjects. Finally, the paper is concluded in Section V.

II. EXPERIMENTAL METHODS

The Experimental Methods section presents and discusses the main data collection and preparation steps required by the sulcus sink system. Information specific to the data collection and preparation stream are presented, and the software tools that process the data are discussed.

A. Data Collection and Surface Reconstruction

For this paper, twenty subjects were imaged. This population was age and gender balanced, with 11 males and 9 females, and ages ranging from 22 – 40 years old. All images were acquired on a 3T Siemens Trio system, T1 weighted MPRAGE images with TR/TE = 2530/3.3, matrix = 256x256, FOV = 256mm, slice thickness 1.33mm, and in-plane resolution of 1x1 mm^2.

Once collected, these images were processed using Freesurfer [11, 6] and surfaces reconstructed. The gray/white junction (or simply, “white matter”) surface was used as the primary analysis surface. For each subject, each sulcus of interest was manually traced on an inflated representation of the white matter surface. FreeSurfer surfaces are tessellated wire-frame structures, typically consisting of more than 120,000 vertices for an average adult brain and provides an isotropic resolution of 1 mm^3. In this study, the average vertex separation distance between individual mesh vertices was about 0.79mm. The mesh structure contains vertex-specific data, including: list of neighboring vertex indices; distance (in mm) to each neighboring vertex; Cartesian coordinates of each vertex in an anatomical space; and surface curvature along each link connecting a vertex to its neighbors.

Table I enumerates the sulci that were considered on each hemisphere, and Figure II.1 shows some manual traces on an example subject.

Initially, the system was “trained” in a supervised learning fashion that guided the optimization process. For each sulcus on each subject under consideration, an expert human operator manually traced the path of the sulcus using projections on the reconstructed brain surface. These paths were cross-checked against planar slices in a volumetric space to assure that they lay along anatomically correct parts of the sulcus [24].

III. OPTIMIZATION METHODS

This section describes the actual optimization method used to minimize error between the human specified sulcal path, and the path generated by the “sulcal sink” software. It will describe the binary weight space, the cost function, and its overlap with the human reference from a Dijkstra graph-theory perspective.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.18.955096; this version posted March 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

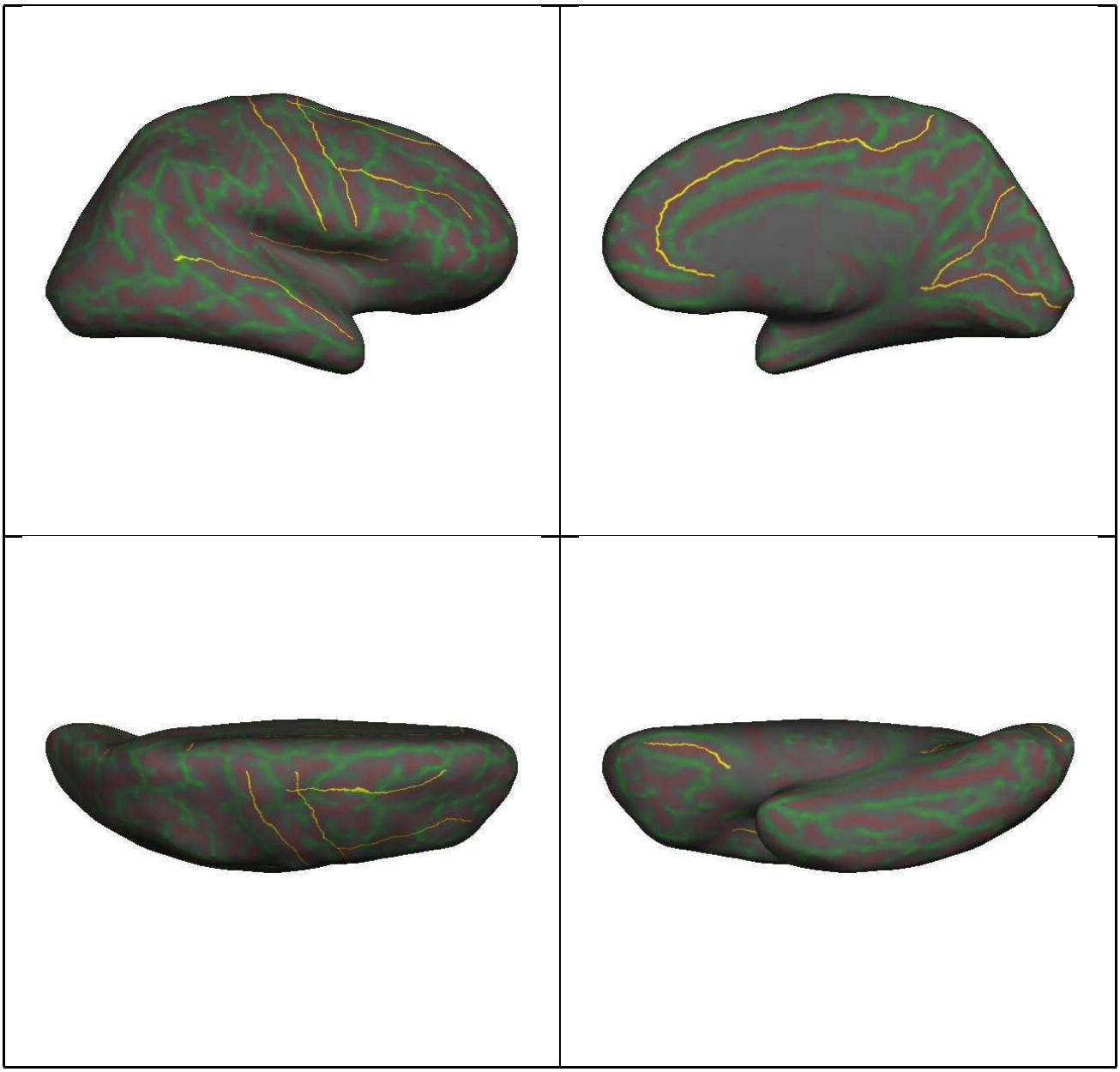

Figure II.1. Sulci as traced by a human operator on a sample reconstructed right hemisphere inflated surface. In the top row, on left, is the lateral view. The sulci shown, from bottom and moving up in a clockwise direction: *sf, ce, prc, sf, if, phr*. On the top right is the medial aspect, dominated by the long *ci*, with the *po* (above) and *calc* (below) at the occipital pole. On the bottom left is the superior aspect showing the *ce* and *prc* (right to left) and the *sf* and *if* (top to bottom). The bottom right shows the *ol* on left. The same sulci were also traced on the contra lateral hemisphere. The background coloration of red and green denote curvature values of the white matter surface projected onto the inflated reconstruction. Red curvatures are positive denoting "inward" depression into the surface; green curvatures are negative, denoting "outward" bulges from the surface.

A. Dijkstra development and weight vector

The reconstructed surface structure generated by Freesurfer is a tessellated mesh that lends itself both visually and conceptually to a graph-theory based approach and in turn to processing by Dijkstra's algorithm [7]. Dijkstra's algorithm is a greedy algorithm that solves the single-source shortest path problem for a directed graph with non-negative edge weights. Essential components are (1) the fact that it is a "locally greedy" algorithm, i.e. operates under the assumption that a global optimum can be found by at each step in an iterative algorithm by choosing the local optimum, and (2) that connections (or weights) between nodes in the graph are non-negative values of some underlying cost function. Dijkstra's algorithm will optimize for the lowest cost path in this connected graph.

Figure III.1 shows a part of the wire-frame mesh for a typical Freesurfer surface reconstruction, and a path running along this mesh. Given an *a priori* path connecting two points in a directed graph as shown, we attempt to find an edge

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.18.955096; this version posted March 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

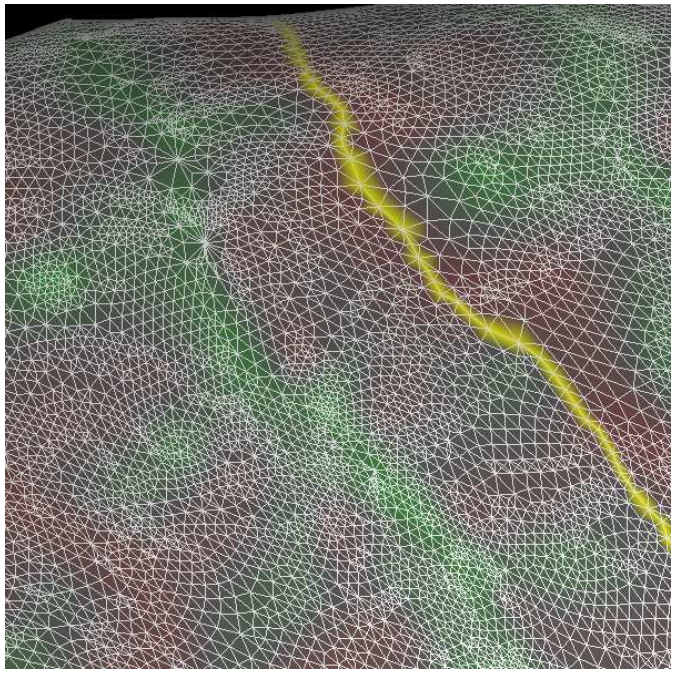

Figure III.1. A path connecting two vertices on a Freesurfer surface mesh. The green and red colored faces denote "outward" and "inward" mean curvature respectively.

weighting that maximizes the overlap between the given path and a Dijkstra optimization,

In the general case, we can create a cost function that is a multi-variable function that contains the vertex edge, a weight vector, and a conditional penalty matrix:

$$C_{ij} = \mathcal{N}\left\{\mathbf{P}_{ij}, \mathcal{E}_{ij}, w_l\right\} \forall i, j \in S \tag{III.1}$$

where $\mathbf{P}_{ij}$ is a penalty weight matrix, $\mathcal{E}_{ij}$ the edge itself between each vertex point $\mathcal{V}_i$ and its neighbor $\mathcal{V}_j$, w_l a weight vector, and surface S . The cost function should have low (positive) values for graph edges that are deep within the sulcus, and higher cost values elsewhere. The connecting edge structure $\mathcal{E}_{ij}$ contains several important fields, viz:

$$\begin{aligned}
\mathcal{E}_{ij} : c_{ij} &= \text{curvature between nodes } \mathcal{V}_\text{i} \text{ and } \mathcal{V}_\text{j} \\
\mathcal{E}_{ij} : h_{ij} &= \text{sulcal height of neighboring node } \mathcal{V}_\text{j} \\
\mathcal{E}_{ij} : d_{ij} &= \text{distance between nodes } \mathcal{V}_\text{i} \text{ and } \mathcal{V}_\text{j} \\
\mathcal{E}_{ij} : c_{max} &= \text{maximum curvature of whole mesh} \\
\mathcal{E}_{ij} : h_{max} &= \text{maximum sulcal height of whole mesh}
\end{aligned} \tag{III.2}$$

Within Freesurfer, both curvature and sulcal height are signed quantities. Since Dijkstra's algorithm depends on non-negative cost values between nodes, we re-cast Freesurfer's signed curvature and sulcal height values to be meaningful in the optimization context we are attempting to solve. This simply entailed shifting the sulcal displacement and surface curvature values upwards by the their respective maximum values for the whole surface:

$$\begin{aligned}
C_{ij} &= \mathcal{E}_{ij} : \{c_{max} - c_{ij}\} \\
H_{ij} &= \mathcal{E}_{ij} : \{h_{max} - h_{ij}\}
\end{aligned} \tag{III.3}$$

which casts high positive curvatures and sulcal heights close to zero and assigns highly negative curvatures and sulcal heights high positive values.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.18.955096; this version posted March 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

1) Fundamental cost function parameters: For the purposes of this paper, we use three fundamental edge parameters, the edge length, d_{ij}, the edge curvature c_{ij}, and the edge sulcal height, h_{ij}. Within a typical Freesurfer reconstruction, these parameters are all of the same magnitude – hence normalization of these values to some arbitrary reference was not necessary. Derived parameters are all the multiplicative combinations of these three fundamentals.

2) Locally Greedy Component: Ignoring for the moment the non-locally greedy penalty matrix of the cost function in Equation III.1, we can express the locally greedy component as a product between the weight vector w_l and a set of edge parameters:

$$ {}^{l}C_{ij} = \begin{bmatrix} w_1 & w_2 & \cdots & w_n \end{bmatrix} \begin{bmatrix} \mathcal{E}_{ij} : p_1 \\ \mathcal{E}_{ij} : p_2 \\ \vdots \\ \mathcal{E}_{ij} : p_n \end{bmatrix} \tag{III.4} $$

These edge parameters, dropping for notational simplicity the $\mathcal{E}_{ij}$: suffix, are described by the seven element set of all combinations of curvature, sulcal height, and distance, as well as an additional parameter, p_{ie}, which is the alignment of the current edge direction i with the vector linking the current vertex node to the end vertex node e. The edge parameters are therefore (with reference to Equation III.2):

$$ \mathbf{p}_{ij,ie} = \begin{bmatrix} \begin{bmatrix} d & c & h & dc & dh & ch & dch \end{bmatrix}_{ij} & p_{ie} \end{bmatrix} \tag{III.5} $$

where compound parameters imply the product of each underlying parameter, i.e. dc implies distance times curvature. Note that the p_{ie} parameter is in fact $1 - P_{ie}$ where P_{ie} is the actual projection of the current edge the linking vector (for complete co-linear projection, $P_{ie} = 1$, which we would weigh as zero in a Dijkstra context, while a tangential direction would have zero projection, and hence a high p_{ie}). Our locally greedy cost function is therefore:

$$ \begin{aligned} {}^{l}C_{ij} &= \begin{bmatrix} w_d & w_c & w_h & w_{dc} & w_{dh} & w_{ch} & w_{dch} & w_p \end{bmatrix} \mathbf{p}_{ij,ie}^T \\ &= \mathbf{w}_l \mathbf{p}_{ij,ie}^T \end{aligned} \tag{III.6} $$

with the ij subscript denoting the parameters for the edge connecting vertex i and j, and the ie subscript denoting the direction from the current vertex i to the end vertex e. The first component of Equation III.6, the locally greedy edge weight vector $\mathbf{w}_l$, is the component that we seek to optimize.

3) Penalty Matrix: The non-locally greedy component of the cost function is a penalty matrix is triggered by curvature transitions along the trajectory path. These arise when strict adherence to the underlying Dijkstra "greedy" behavior is in fact a bad strategy. This might arise if "bumps" or interruptions are found along a sulcal path - particularly when a gyrus "interrupts" or crosses over a sulcus. A purely locally "greedy" algorithm might route *around* this interruption instead of continuing *over* it to reach the remainder of the current sulcus. The non-linear component consists of a penalty matrix that toggles between unity and a penalty value based on any changes in curvature sign that are encountered along a path,

$$ \begin{aligned} \mathbf{P}_{ij} &= \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \forall c_{ij} < 0 \\ & \begin{bmatrix} \Delta w_1 & 0 & \cdots & 0 \\ 0 & \Delta w_2 & \cdots & 0 \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & \Delta w_n \end{bmatrix} \forall c_{ij} \geq 0 \end{aligned} \tag{III.7} $$

where the Δw_k penalties weigh the original weight vector with an additional per-edge weight to counter the locally greedy nature of Dijkstra's algorithm. The final cost function follows from Equations III.1, III.6, and III.7:

$$ C_{ij} = \mathbf{w}_l \mathbf{P}_{ij} \mathbf{p}_{ij,ie}^T \tag{III.8} $$

and has as diagonal penalty vector

$$ \begin{aligned} \text{diagonal} &= \begin{bmatrix} \Delta w_d & \Delta w_c & \Delta w_h & \Delta w_{dc} & \Delta w_{dh} & \Delta w_{ch} & \Delta w_{dch} & \Delta w_p \end{bmatrix} \\ &= \begin{bmatrix} 1 & 10 & 10 & 10 & 10 & 10 & 1 & 1 \end{bmatrix} \end{aligned} \tag{III.9} $$

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.18.955096; this version posted March 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

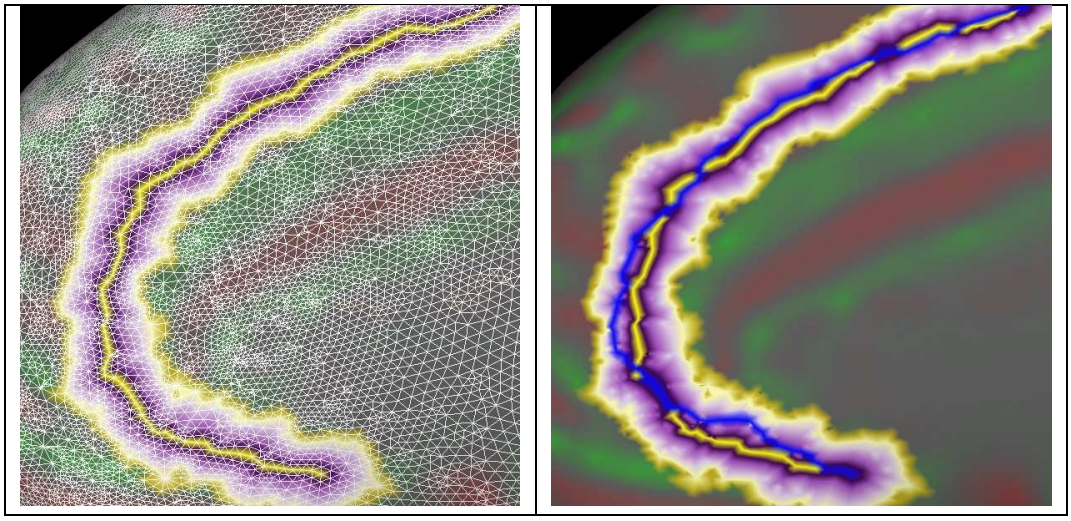

Figure III.2. A series of ISO-regions about each target path were used to evaluate the utility of a candidate weight vector. On the left, the original path surrounded by its ISO Regions, shown projected on a wire-frame. On the right, a candidate path in blue that attempts to track the original path. Notice how different parts of the candidate path intersect different ISO Regions. The "fitness" of a candidate path is determined by how many vertex points along its length are outside successive ISO regions. Thus, an exact match would have 0 points outside all the ISO regions.

Since the penalty diagonal matrix was kept constant throughout, and only the weight vector was variable across a whole Dijkstra path, we can simplify Equation III.8 as

$$C_{ij} \quad = \quad \mathbf{w}_l \mathbf{P} \mathbf{p}_{ij,ie}^T \tag{III.10}$$

and the total cost for a given Dijkstra path $\mathcal{P}$ as the sum of all individual edge costs for a given route:

$$C = \sum_{ij \in \mathcal{P}} C_{ij}$$

4) Goodness of Fit: Given the cost function described in Equation III.8, we now need to develop a "goodness of fit" metric relative to the problem at hand. Referring back to Figure III.1, we are trying to find a weight vector, $\mathbf{w}_l$, that when used by Dijkstra's algorithm with an edge cost defined by Equation III.8 will yield a path that is co-linear (or very-almost co-linear) with the original reference path specified.

Consider now Figure III.2 which shows part of a target path about which are successive colored regions. Each region is one vertex removed from its immediate inner region – hence they are referred *ISO*-regions. The target path occupies the inner-most region. About it, and at a distance of one vertex removed, is ISO Region 1. ISO Region 1 therefore defines a strip that is three vertices wide – the inner target path itself, and a single vertex wide strip on either side of this path. In a similar manner, the single-vertex region about ISO Region 1 defines ISO Region 2, and so forth until ISO Region 6, which is a ribbon patch 13 vertices wide with the original path at its very center.

Conceptually, the most straightforward "goodness of fit" measure is simply the error distance between our reference signal and the Dijkstra generated signal. We can computationally simplify this measurement without needing to calculate distances, but consider instead an ISO region weighted function of the number of mesh vertices of the Dijkstra trajectory that lie in each ISO region. For a given ISO region r, we can express a percentage overlap of the Dijkstra path in this region as

$$f_r = \frac{1}{d} \sum_{k=1}^{N} \mathcal{V}_k \;\; \text{where } \mathcal{V}_k = \begin{cases} 1 & \forall \mathcal{V}_k \in r \\ 0 & \forall \mathcal{V}_k \notin r \end{cases} \tag{III.11}$$

where d is the length of the original reference path, and $\mathcal{V}_k$ is a straightforward membership function of vertex k in the Dijkstra generated path for ISO region r about the reference path. Given this overlap fraction for ISO Region r, we can express the combined weighted overlap across all regions as

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.18.955096; this version posted March 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

$$F = \frac{1}{R}\sum_{r=1}^{R}\left(1 - f_r\left(1 + \delta\left(r-1\right)\right)\right) \tag{III.12}$$

where δ is a weight penalty that penalizes overlap in higher ISO regions. This simpler calculation provides the same benefit to using a more complex distance metric: a Dijkstra path that meanders further away from the human reference has a lower overlap F. An example Dijkstra path is shown in blue on the right of Figure III.2, projected on the original path and with the ISO Regions shown. Notice how the Dijkstra path travels through several different ISO Regions.

B. Weight vector optimization

The underlying nature of C_{ij} is highly non-linear in its greedy weight vector, $\mathbf{w}_l$ – in other words small $\delta\mathbf{w}_l$ perturbations have no or little practical effect on C_{ij}. This negated the use of gradient-based search methods. Moreover, the stated purpose of this paper is to describe sulci as a binary weight vector. This design decision considerably simplified our optimization regime – specifically for our nine-dimensional weight vector, a binary search needs only consider $2^9 = 512$ possible weight factors, which lay comfortably within the realm of exhaustive exploration. For each of the 512 possible weight vectors for a given sulcus on a given subject, we generated a shortest path using Dijkstra's algorithm off the weight vector, and calculated the percentage overlap between the generated path and our reference using the weighted ISO method of Equation III.12. The weight vector with the highest overlap was chosen as the optimal for the particular subject and particular sulcus.

C. Generalizing optimal weight vectors to templates

Each sulcus of each subject of each hemisphere can potentially have its own weight vector (or a set of weight vectors that map to the same "best" fitness value). One of the goals of this project was to determine if there is any "similarity" in terms of weight vectors for a given sulcus. In other words, is it possible to generalize features of a particular sulcus's weight vector across all similar sulci from all the subjects?

One possible way to generate a template weight vector from a set of subject's weights is to simply find the median value of all the relevant individual weight vectors, based on the assumption that this is the most representative sampling of the solution space. In order to evaluate the utility of this template with reference to the entire space of "best" vectors, we also perform a clustering based analysis to determine that the median vector lies within the larger "cloud" of solution vectors.

Such a clustering methodology is well established in statistical theory, and in fact is available as a set of tools and function calls in MatLAB. For this project, we considered each solution vector as a point in a nine dimensional space, and using a Hamming distance metric, generated a hierarchical cluster tree. We then partitioned the cluster tree using a cutoff value of 10, i.e. partitioned the space into a maximum of 10 clusters. The largest cluster was then determined, and a the overlap between the median template vector and this cluster was calculated, i.e. we determined how many bits were in common between the template and the largest cluster space of all the weight vector data. This overlap was used as a measure of confidence for the template vector's validity.

IV. Results and Discussion

For each of the twenty subjects in this study, ten sulci on each hemisphere were optimized. This resulted in $20 \times 10 \times 2 = 400$ sulcal trajectories. We considered three conditions: (1) the trajectory fit for each sulcus as optimized to each subject; (2) the trajectory fit for each sulcus using a single template weight generalized from the optimal weights of each subject; and (3) the trajectory fit using a NULL (or zero) weight vector. If a given subject and given sulcus as parametrized by the weight vector exactly "fits" over the human generated example, the trajectory fit is 100%. An overview of all the results for each of theses conditions is show in Figure IV.1. The percentage overlap is shown per subject and per sulcus, with a blue-red colormap (bluer colors represent worse overlap, redder colors better overlap). Sulcal indices are as defined in the Figure.

Figure IV.1 shows that the per-subject optimals are close to overlap with the human examples; the generalized sulcal templates had less good overlap; and the NULL weight baseline showed generally poor behavior. Note that we immediately see in the NULL template the poor performance of the cingulate (indices 3 and 13).

The results and discussion are organized into three sub-sections. First, we consider the system behavior on a per-sulcus manner evaluated across all patients. This is followed by a perspective taken from each subject considering all sulci for that subject. Finally, we present a blinded human expert evaluation of all sulcal traces.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.18.955096; this version posted March 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

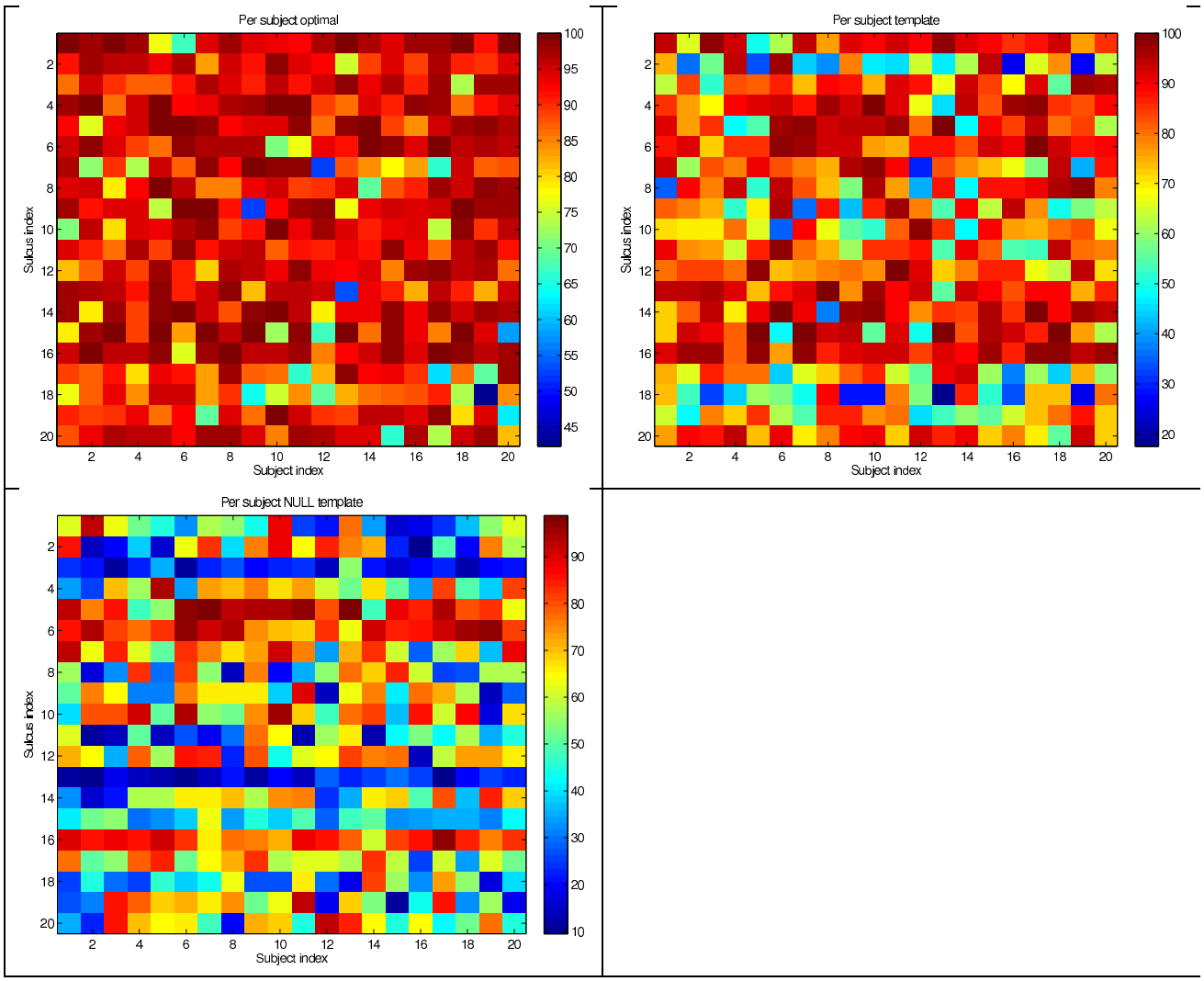

Figure IV.1. Percentage overlap of generated sulcal trajectories per subject and per sulcus. On top left, the overlap of the per subject, per sulcus optimals; on top right the overlap of the per sulcal group binary template weights; on bottom left the baseline NULL weight vector results. Subject indices are from left-to-right, and sulcal indices from top-to-bottom in order: 1: lh_calc; 2: lh_ce; 3: lh_ci; 4: lh_if; 5: lh_ol; 6: lh_phr; 7: lh_po; 8: lh_prc; 9: lh_sf; 10: lh_st; 11: rh_calc; 12: rh_ce; 13: rh_ci; 14: rh_if; 15: rh_ol; 16: rh_phr; 17: rh_po; 18: rh_prc; 19: rh_sf; 20: rh_st.

A. Overlap results from a sulcus perspective

Table II explores Figure IV.1 from the perspective of each sulcus. For each group (optimal, template, and NULL), three result columns are shown. In the "optimal" group, the weight values for a a single example subject and the percentage fit for this subject are presented, followed by the percentage fit averaged over all subjects. Thus, we see that for the "lh_calc" sulcus, one example subject had an optimal binary weight descriptor of "010000011" which resulted in a 99.68% weighted overlap. By comparison, taken across all subjects, the average for the optimal "lh_calc" sulcus weight resulted in a 94.20% overlap over the human reference (and each subject had a uniquely optimized sulcal weight). Also notable are some comparative observations evident between our chosen example and the whole subject group. For this example, the "lh_st" sulcus only optimized to a 70.50% overlap, and was the worst performing sulcus in this case. Taken across all subjects, however, this sulcus averaged 91.56% overlap. Clearly this example performed worse than the group average. The lowest overlap across all subjects was for the "rh_prc" at 81.48%, which also scored low for the subject shown. These observations might indicate that certain specific sulci optimize with more difficulty than others and might do so in a consistent manner.

In the middle part of Table II, the template binary weight for each sulcus is shown, followed by the percentage sulcal trajectory overlap that this template resulted in for the same example subject as the "optimal" part of the table. The "% fit all" column shows the overlap of each template weight sulcus averaged over all subjects. Each sulcus now had the same

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.18.955096; this version posted March 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

	OPTIMAL (O)			TEMPLATE (T)			NULL (N)		
sulcus	$\hat{\mathbf{w}}$ example	% fit example	% fit all	$\bar{\mathbf{w}}$ template	% fit example	% fit all	$\bar{\mathbf{w}}_0$ template	% fit example	% fit all
lh_calc	010000011	99.68	94.20	010101001	94.12	85.07	000000000	61.23	47.97
lh_ce	000010011	91.76	92.27	010000000	74.75	58.53	000000000	85.57	52.66
lh_ci	000000110	86.22	92.33	100001010	62.94	80.24	000000000	23.94	22.37
lh_if	011010000	97.32	94.21	010010000	85.05	85.59	000000000	33.33	59.26
lh_ol	000000000	92.01	93.97	100000000	92.01	82.40	000000000	92.01	82.40
lh_phr	010010000	95.31	93.33	111010000	88.01	88.15	000000000	85.42	84.32
lh_po	010001001	96.61	86.07	010010001	93.45	76.81	000000000	92.42	67.31
lh_prc	010010101	93.71	91.23	010000000	34.31	77.07	000000000	55.97	50.83
lh_sf	011001010	97.73	91.57	111010000	80.71	71.45	000000000	50.22	54.81
lh_st	010000000	70.50	91.56	010000000	70.50	71.75	000000000	39.20	67.07
rh_calc	110010001	93.78	92.73	111000001	90.60	77.80	000000000	62.08	39.81
rh_ce	111001000	81.82	92.35	110010000	79.55	81.06	000000000	71.46	63.58
rh_ci	110011000	97.64	90.85	100000110	94.57	87.14	000000000	11.76	18.51
rh_if	010010000	97.86	94.18	000001010	72.22	86.37	000000000	31.87	55.46
rh_ol	001000010	78.48	89.90	101010110	72.67	83.13	000000000	40.35	40.58
rh_phr	001010011	94.60	95.29	010010000	91.78	91.92	000000000	89.91	81.54
rh_po	000110100	88.89	86.11	011000001	74.87	68.93	000000000	76.46	61.45
rh_prc	011010000	77.32	81.48	010000001	73.75	54.85	000000000	25.53	41.99
rh_sf	010001001	90.37	89.20	010000000	65.00	69.32	000000000	27.13	55.11
rh_st	010000011	87.94	91.46	011010000	76.67	81.21	000000000	35.74	58.24
mean		90.48	91.22		78.38	77.94		54.58	55.26
std		7.98	3.36		14.40	9.69		26.12	17.52

Table II

THE OPTIMAL VS TEMPLATE VS NULL BINARY WEIGHT DESCRIPTOR BEHAVIOR FOR EACH SULCUS. IN THE OPTIMAL CASE (COLUMNS ON LEFT), THE BINARY SULCUS DESCRIPTOR FOR A SPECIFIC SULCUS IS SHOWN IN $\hat{\mathbf{w}}$ FOLLOWED BY THE PERCENTAGE OVERLAP THIS SULCUS DESCRIPTOR GENERATED COMPARED TO THE HUMAN REFERENCE IN THE "% FIT EXAMPLE" COLUMN. THE "% FIT ALL" COLUMN SHOWS THE AVERAGE OVERLAP OF EACH OPTIMAL SULCUS FOR ALL THE SUBJECTS (EACH SUBJECT HAD ITS OWN DESCRIPTOR). IN THE TEMPLATE COLUMNS (CENTER), THE TEMPLATE DESCRIPTOR FOR EACH SULCUS IS SHOWN IN $\bar{\mathbf{w}}$. THIS SAME DESCRIPTOR WAS USED FOR EACH SULCUS ACROSS ALL SUBJECTS. THE "% FIT EXAMPLE" SHOWS THE OVERLAP THAT THE TEMPLATE GENERATED WHEN USED FOR AN EXAMPLE CASE, FOLLOWED BY THE AVERAGE "% FIT ALL" THAT THIS TEMPLATE RESULTED IN ACROSS ALL SUBJECTS. IN A SIMILAR FASHION, THE NULL WEIGHT VECTOR IS SHOWN ON RIGHT FOR A SPECIFIC EXAMPLE AND ACROSS ALL SUBJECTS.

template weight for all subjects, for example the "lh_calc" had a binary template of "010101001" which was applied to all "lh_calc" sulci on all subjects. Overall, the relative overlap percentages for the templated weights scored lower, which given that we have generalized away from a set of optimals matches our expectations. The deviation in overlap for the template weights, however, was lower than the subject-specific optimals, suggesting that generalization was preserving certain underlying properties of sulci across all subjects. Certain sulci, such as the "rh_prc" performed uniformly badly. The template binary resulted in an average 54.85% overlap across all subjects. It also scored comparatively low on the per-subject optimal shown in the first part of the table at 77.32% overlap, with also the lowest group optimal at 81.48%.

Figures IV.2 provides a comparative "real world" references for the overlap percentages of Table II, with human traced sulcal references compared to the optimally weighted sulci on left, and also the human reference sulci compared with the templated binary sulci descriptions on right for the same subject.

The final component of the Table shows the resultant behavior using the NULL template. In this case, all sulcus weights were set to zero, and the algorithm essentially showed no bias to any cost weight, and merely traveled along each mesh edge until it happened to connect the start and terminus vertices. The NULL templates are used to define a baseline reference against which the optimal and templated weights can be compared. The mean overlap in this case for all sulci across all subjects was 54.58%, with a 26.12% standard deviation. Note though how sensitive some sulci can be to weights. The NULL "lh_ce" sulcus had for the example shown in the table, a surprising 85.57% overlap with the human target. The template weight for this sulcus differed in only one bit, the w_c bit. This template weight for the example subject shown, dropped the overlap to 74.75% (and in fact, for this subject, the optimal "lh_ce" weighting was from a Hamming distance measure quite distal).

B. Overlap results from a subject perspective

Using a similar organization as in Section IV-A, Table III presents overlap results organized per subject and shown across all sulci for that subject. Some additional information is also shown. Instead of weight vectors, we present the fractional degradation in performance: $\frac{T}{O}$ is the ratio of the Template overlap to the Optimal overlap, and $\frac{N}{T}$ the ratio of the NULL overlap to the Template overlap. The product of these represents the overall degradation in overlap between the Optimal and the NULL weights.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.18.955096; this version posted March 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

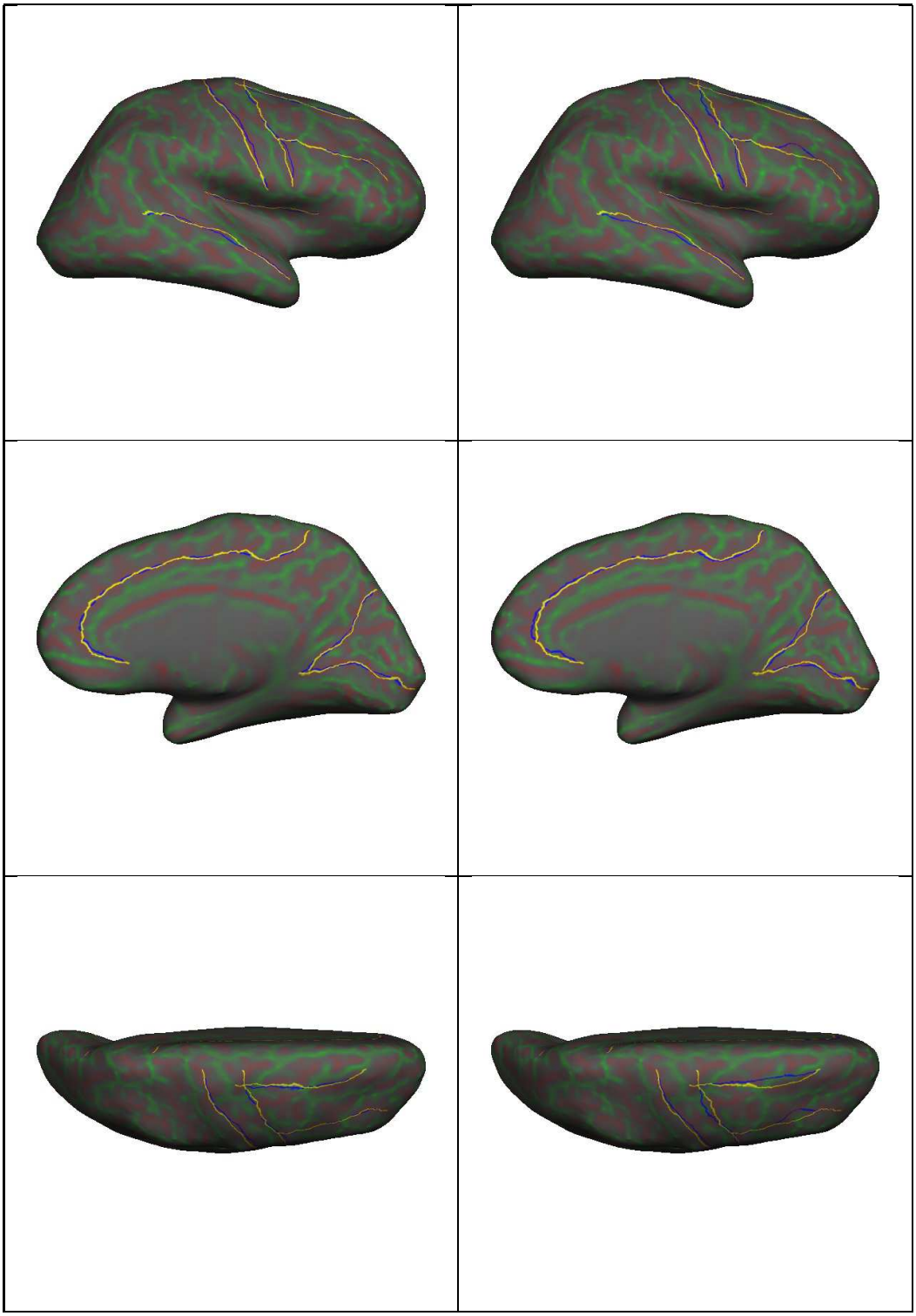

Figure IV.2. Human traced sulci (yellow) vs. binary weighted Dijkstra paths (blue) for the subject of Table II, and shown on the right hemisphere inflated surface. On left, lateral, medial, and superior aspects comparing the human references to the optimal binary descriptions. On right, human reference traces compared with the binary template description. The background coloration of red and green denote curvature values of the white matter surface projected onto the inflated reconstruction.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.18.955096; this version posted March 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

		OPTIMAL (O)	TEMPLATE (T)		NULL (N)		
subject	$\bar{v}_{sep}$	%fit	%fit	$\frac{T}{O}$	%fit	$\frac{N}{T}$	$\frac{T}{O}\frac{N}{T}$
2787_5	0.7959	90.48	78.38	0.87	54.58	0.70	0.60
2859_4	0.7950	91.23	75.53	0.83	47.40	0.63	0.52
2884_5	0.7927	92.09	78.07	0.85	54.04	0.69	0.59
2979_5	0.7965	90.80	77.67	0.86	56.53	0.73	0.62
2980_5	0.7882	93.00	75.92	0.82	50.97	0.67	0.55
2981_5	0.7949	91.76	78.79	0.86	60.30	0.77	0.66
2982_5	0.7910	91.84	77.86	0.85	60.61	0.78	0.66
2997_5	0.7959	94.24	79.26	0.84	53.09	0.67	0.56
2998_5	0.7905	90.09	78.80	0.87	62.67	0.80	0.70
3042_5	0.7990	91.41	80.05	0.88	60.07	0.75	0.66
3156_5	0.7975	92.76	83.97	0.91	59.64	0.71	0.64
3157_5	0.7847	86.90	72.93	0.84	49.34	0.68	0.57
3161_6	0.7946	89.56	73.07	0.82	63.58	0.87	0.71
3181_5	0.7905	91.52	78.88	0.86	60.68	0.77	0.66
3279_5	0.7819	91.38	81.40	0.89	50.03	0.61	0.55
3280_5	0.7968	91.96	73.97	0.80	46.62	0.63	0.51
3305_5	0.7946	90.36	80.46	0.89	58.02	0.72	0.64
3320_5	0.7981	92.00	79.54	0.86	50.88	0.64	0.55
3321_5	0.7963	90.42	76.47	0.85	51.62	0.68	0.57
3322_5	0.7937	90.51	77.76	0.86	54.59	0.70	0.60
mean		91.22	77.94	0.85	55.26	0.71	0.61
std		1.50	2.74	0.026	5.21	0.064	0.058

Table III

THE OPTIMAL VS TEMPLATE VS NULL BINARY WEIGHT DESCRIPTOR BEHAVIOR FOR EACH SUBJECT. THE $\bar{v}_{sep}$ COLUMN INDICATES THE AVERAGE VERTEX SEPARATION FOR A PARTICULAR SUBJECT. THE OPTIMAL % FIT COLUMN PRESENTS THE RESULTS OF ALL THE OPTIMAL OVERLAPS FOR A GIVEN SUBJECT AVERAGED OVER ALL SULCI FOR THAT SUBJECT. SIMILARLY FOR THE TEMPLATE %FIT AND NULL %FIT. THE $\frac{T}{O}$ COLUMN SHOWS THE TEMPLATE FIT AS A PERCENTAGE OF THE OPTIMAL FIT. SIMILARLY FOR THE $\frac{N}{T}$ COLUMN. THE FINAL COLUMN PRESENTS THE OVERLAP OF THE NULL WEIGHT DESCRIPTOR AS A PERCENTAGE OF THE OPTIMAL BEHAVIOR.

As before, we note that the per-subject optimal weight vectors scored the highest overlap with the human reference sulci, with expected decreases when using sulcal template weights and a low scoring NULL reference. Taken over all subjects and sulci, the Template weights performed 85% as good as the optimal weights. The NULL weights performed 61% as well.

Subjects 2997_5 and 3156_5 scored the highest overlaps for their optimal weighting against their human generated references. In fact, subject 3156_5 scored the highest when using template weights, and (considering standard deviation) also high with the NULL reference weights – indicating that this specific subject's sulci were perhaps geometrically more easily captured by this paper's weight vector formulation and hence better conserved by the generalization process than the structural description of subject 3280_5 which had high optimal overlap, but comparatively low template and NULL weight overlap with its human traced sulcal references.

C. Sulcal-template Optimizations

The template weights are an attempt to generalize the binary weight pattern for a given sulcus across all the subjects. Table IV shows summary results for each sulcus considered by this project. The first column presents the template weight vector, and the second indicates the correlation of this template to within the largest cluster of the space defined by all the individual weight vectors for that sulcus. In general we see that the correlation is 100%, except for 5 sulci that had 88% correlation, meaning that only 1 bit of the weight vector was not within the cluster space.

It is important to note that the binary template description in and of itself is not a complete specification of a sulcus. Note in the table that the "lh_phr" and the "lh_sf" have the same template. While this does suggest that an underlying geometric similarity might exist between these sulci, each of these sulci are only fully specified in conjunction with their start and end vertices. Since Dijsktra's algorithm is locally greedy, different start and end vertices for a sulcus, provided that they do not drift too much, would not change the overall sulcal weight description. We evaluated the effect of locally varying the start and end vertices on a sulcus's weight optimization, and for variations within a range of 10 vertices, no changes in the optimal weights were observed.

The remaining columns of Table IV show the standard deviation of each particular weight bit from the median. For cases where the deviation is very close to 0.50, we can conclude that the underlying variation in that bit's value ranged evenly between 0 and 1. Nonetheless, seen as a whole, some bits over the entire subject space showed less variation that others. In fact, the deviation in the $\sigma_d, \sigma_h, \sigma_{dh}, \sigma_{ch}, \sigma_{dch}, \sigma_p$ were all less than 0.45 while the remaining weights variations, i.e. $\sigma_{\mathbf{P}}, \sigma_c, \sigma_{dc}$ were all higher. We can interpret lower deviation values to imply a relative "bit strength" - the closer to zero deviation, the better the generalization of that particular bit.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.18.955096; this version posted March 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

sulcus	$\bar{\mathbf{w}}$	correlation	$\sigma_{\mathbf{P}}$	σ_d	σ_c	σ_h	σ_{dc}	σ_{dh}	σ_{ch}	σ_{dch}	σ_p
lh_calc	010101001	1.00	0.49	0.47	0.49	0.48	0.50	0.50	0.49	0.49	0.42
lh_ce	010000000	1.00	0.49	0.46	0.44	0.40	0.46	0.47	0.36	0.49	0.46
lh_ci	100001010	1.00	0.46	0.48	0.46	0.35	0.49	0.50	0.50	0.44	0.35
lh_if	010010000	1.00	0.47	0.34	0.36	0.00	0.50	0.49	0.00	0.50	0.44
lh_ol	100000000	1.00	0.50	0.45	0.42	0.46	0.49	0.47	0.40	0.50	0.49
lh_phr	111010000	1.00	0.50	0.49	0.50	0.39	0.50	0.31	0.43	0.40	0.48
lh_po	010010001	1.00	0.49	0.42	0.49	0.46	0.50	0.47	0.32	0.37	0.44
lh_prc	010000000	1.00	0.48	0.24	0.37	0.00	0.49	0.37	0.17	0.30	0.44
lh_sf	111010000	0.88	0.50	0.47	0.50	0.34	0.50	0.40	0.18	0.40	0.44
lh_st	010000000	1.00	0.50	0.47	0.44	0.50	0.49	0.49	0.30	0.22	0.48
rh_calc	111000001	0.88	0.48	0.46	0.50	0.48	0.47	0.29	0.46	0.32	0.46
rh_ce	110010000	1.00	0.49	0.48	0.49	0.38	0.48	0.50	0.35	0.50	0.50
rh_ci	100000110	0.88	0.43	0.39	0.45	0.36	0.43	0.50	0.45	0.48	0.00
rh_if	000001010	0.88	0.47	0.47	0.49	0.47	0.50	0.49	0.49	0.50	0.37
rh_ol	101010110	1.00	0.49	0.49	0.49	0.50	0.49	0.50	0.49	0.49	0.49
rh_phr	010010000	1.00	0.50	0.43	0.47	0.16	0.49	0.40	0.23	0.44	0.41
rh_po	011000001	0.88	0.49	0.49	0.49	0.45	0.49	0.17	0.45	0.28	0.49
rh_prc	010000001	1.00	0.44	0.37	0.43	0.35	0.49	0.15	0.15	0.15	0.50
rh_sf	010000000	1.00	0.49	0.50	0.41	0.40	0.47	0.45	0.47	0.46	0.35
rh_st	011010000	0.88	0.49	0.50	0.50	0.49	0.50	0.47	0.49	0.48	0.48
mean			0.48	0.44	0.46	0.37	0.49	0.42	0.36	0.41	0.42

Table IV

PER-SULCUS TEMPLATE WEIGHTS, CORRELATION, AND DEVIATION. THE CORRELATION COLUMN INDICATES WHAT FRACTION OF THE WEIGHT TEMPLATE VECTOR WAS CONTAINED IN THE LARGEST CLUSTER OF POINTS GENERATED BY ALL THE OPTIMAL WEIGHT VECTORS FOR A PARTICULAR SULCUS. THE REMAINING COLUMNS SHOW THE DEVIATION FOR INDIVIDUAL WEIGHT VECTOR BITS. THE LESS DEVIATION IN A WEIGHT BIT, THE BETTER THE GENERALIZATION OF THAT BIT.

D. Human-expert evaluation

Although we can measure the deviation from the human-generated "gold standard" example templates of this "sulcus sink" system, a basic question remains as to how acceptable such generated sulcal paths are to a human expert. To this end, an independent expert anatomist who was not involved in the generation of the original sulcus paths, was presented with three sets of results for each hemisphere of each subject. Each set showed a complete grouping of projections of ten sulcal traces. One set showed the original traces as created by a human (the "target" set); a second set showed the subject-sulcus specific optimal, i.e. the best weight candidate for each sulcus on the particular subject as generated by the Dijkstra-based optimization (the "optimal" set); and a final set showed, for each sulcus, the trajectory resultant from a generalization of all weight vectors for each sulcus (the "template" set).

Without any knowledge of which set corresponded to which generation process, the evaluator was tasked with rating each sulcus of each set on a four-point (1 – 4) scale:

1) sulcus trajectory was completely unacceptable with significant portions of its path measurably far away from target
2) sulcus trajectory was unacceptable – typically the trajectory followed a neighboring, i.e. non-target, sulcus
3) sulcus trajectory was largely acceptable with portions of its path tracking along sulcal banks and not necessarily along the sulcal trough
4) sulcus trajectory was completely acceptable

Although there is of course a certain evaluator bias in such a rating system, generally speaking sulcal traces of "3" and "4" are acceptable, while those of "1" and "2" are not. Examples of these rating are shown in Figure IV.3 .

In the top row of Figure IV.3, the "1" rated cingulate was the result from using the generic cingulate template generated from all the optimal cingulate sulci of all subjects. The "4" rating on the right was the target trajectory as traced by a human. Interestingly, for the bottom row, the "2" rated cingulate on the left was in fact the trace as originally made by the human operator, and the "4" rated cingulate on the right was the same generic cingulate template that failed so completely for the top row subject.

Consider Table V which summarizes the sulcal trace evaluation on a per-subject basis. For each set of "target", "optimal", and "template", the average rating for all sulci per subject per set was calculated. For the most part, the "target" set had the highest mean across all the subjects, followed by the "optimal", with the "template" performing worst – a general trend that conforms with our expectations. Note however, some interesting deviations, particularly with the right hemisphere rankings. In some subjects (2859_4, 2884_5, 3156_5, and 3161_6) the trend for one or other hemisphere was reversed with the human target rated lowest, and the sulcus trace generated from the generic template rated highest. In these cases, several sulcal traces on these subjects made by the humans were not drawn well. How can the template perform better if some of its underlying domain set descriptors have errors? Examining these cases, we noted that only a small number of

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.18.955096; this version posted March 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

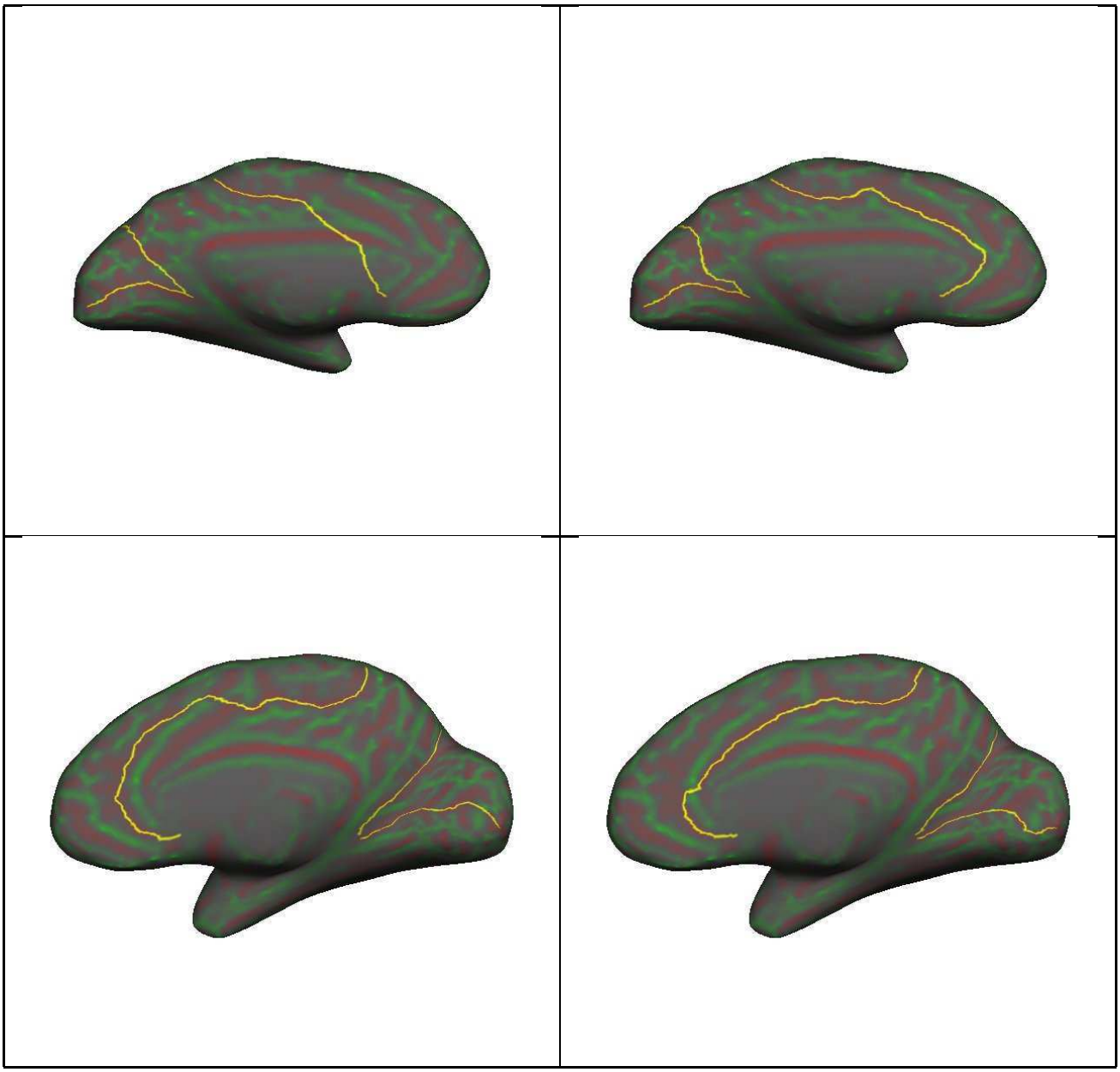

Figure IV.3. Examples of sulcal trace ratings using the cingulate sulcus (the long trace running length-wise along the medial aspect). Along the top row: “1” on left; “4” on right – notice that the “1” trace completely misses the anterior portion of the cingulate. Along the bottom row: “2” on left; “4” on right – in this case the “2” trace runs along the neighboring sulcus for most of its anterior component.

the sulci were traced badly. Our nine-bit binary descriptor had enough descriptive power to capture relevant features from the remaining larger pool of correctly drawn sulci to mitigate the errors in the training set.

Table VI organizes the summary rankings on a per-sulcus manner across all subjects. As with Table V, we note that the general trend is with the human “target” ranked highest, followed by the “optimal” track, with the “template” performing the worst. The human “target” also had the lowest score deviation, while the “template” sulcus had a measurably higher standard deviation.

Note however, that in the “rh st” sulcus (right hemisphere superior temporal), and “rh/lh ol” (left and right hemisphere olfactory) cases, the human “target” was ranked lower than the “template” sulcus. In conjunction with Table V, this highlights subjects and sulci where the human operators might have made minor errors in creating the initial “target” traces for the system. For both hemispheres the “ci” (cingulate) sulcus had the lowest “template” rank – attesting to the fact that several “template” cingulate traces suffered from significant errors. This might suggest that the overall sulcus length might

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.18.955096; this version posted March 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

	Left Hemisphere			Right Hemisphere		
subject	Human	Optimal	Template	Human	Optimal	Template
2787_5	3.90	3.80	3.60	3.80	3.90	3.70
2859_4	4.00	3.90	3.90	3.60	3.60	3.80
2884_5	4.00	3.90	3.70	3.90	3.90	4.00
2979_5	3.90	3.70	3.70	4.00	4.00	4.00
2980_5	4.00	3.90	3.90	4.00	4.00	3.70
2981_5	4.00	4.00	4.00	3.80	3.80	3.80
2982_5	4.00	4.00	3.70	4.00	4.00	4.00
2997_5	4.00	4.00	4.00	4.00	4.00	3.90
2998_5	4.00	4.00	4.00	4.00	4.00	3.80
3042_5	3.90	4.00	3.70	4.00	3.90	4.00
3156_5	4.00	4.00	4.00	3.80	3.90	4.00
3157_5	3.90	4.00	4.00	3.80	4.00	4.00
3161_6	4.00	4.00	4.00	3.80	3.80	4.00
3181_5	3.80	3.80	3.80	4.00	4.00	4.00
3279_5	3.90	4.00	4.00	3.80	3.70	3.80
3280_5	3.80	3.80	3.70	4.00	3.90	3.80
3305_5	4.00	4.00	4.00	3.90	3.90	3.90
3320_5	3.80	3.50	3.50	4.00	4.00	4.00
3321_5	4.00	4.00	4.00	4.00	3.90	3.80
3322_5	4.00	4.00	4.00	3.90	3.90	3.70
mean	3.95	3.92	3.86	3.91	3.91	3.89
std	0.0759	0.135	0.167	0.115	0.110	0.118

Table V

SULCUS TRACE EVALUATIONS ON A FOUR POINT SCALE BETWEEN 1 (WORST) AND 4 (BEST) AS RANKED BY AN INDEPENDENT HUMAN EXPERT UNAWARE OF WHICH SULCAL SET IS THE RESULT OF WHICH PROCESS. RESULTS ARE SUMMARIZED ON A PER-SUBJECT BASIS, I.E. THE AVERAGE OF ALL RANKINGS FOR A SUBJECT/HEMISPHERE/SET.

	Left Hemisphere			Right Hemisphere		
sulcus	Human	Optimal	Template	Human	Optimal	Template
calc	3.900	3.90	3.91	3.90	3.90	3.85
ce	4.00	4.00	4.00	4.00	4.00	4.00
ci	3.90	3.75	3.05	3.90	3.95	3.65
if	4.00	3.85	3.91	3.90	3.90	3.90
ol	3.80	3.95	3.95	3.90	4.00	4.00
phr	4.00	4.00	4.00	4.00	3.95	4.00
po	3.90	3.90	3.91	3.85	3.85	3.85
prc	4.00	3.95	3.91	3.95	3.95	3.85
sf	3.95	3.90	3.95	3.95	3.95	3.95
st	4.00	3.95	3.95	3.70	3.60	3.80
mean	3.95	3.92	3.85	3.91	3.91	3.89
std	0.0685	0.0747	0.285	0.0864	0.117	0.11068

Table VI

SULCUS TRACE EVALUATIONS ON A FOUR POINT SCALE BETWEEN 1 (WORST) AND 4 (BEST) AS RANKED BY AN INDEPENDENT HUMAN EXPERT UNAWARE OF WHICH SULCAL SET IS THE RESULT OF WHICH PROCESS. RESULTS ARE SUMMARIZED ON A PER-SULCUS BASIS, I.E. FOR A GIVEN SULCUS, THE AVERAGE ACROSS ALL SUBJECTS FOR A GIVEN HEMISPHERE/SET COMBINATION.

be an important factor in how well this system could generate optimal and hence template paths. The cingulate was by far the longest sulcus, and was consistently the lowest ranked.

V. CONCLUSION

This paper described a methodology for compactly describing and tracing sulcal paths on reconstructed cortical surfaces. Using a sample set of 400 human-traced sulci paths, an optimization process attempted to find a weight vector that minimized the error between a sulcus target trajectory and a sulcus path generated by a Dijkstra-based search using the weight vector to define edge weights in the reconstructed cortical mesh. The system operated thus in a semi-automated supervised fashion since its optimization targets were provided *a priori* – and the optimization process itself was an *inverse* search since we were attempting to solve for weight factors that best mapped to a target path through Dijkstra's algorithm.

The weight vector was a nine-element binary word that simply toggled on or off a set of contribution various features in the three dimensional surface space of a reconstructed brain. As such, the optimal weight vector for a specific sulcus represented along with the sulcus's start and end points, a succinct binary parametrization or description of that sulcus.

Our approach with the methods of this paper is somewhat complimentary or tangential to the tendency to use highly automated methods. Fully automated methods tend to sacrifice a component of robustness while increasing repeatability and dramatically reducing analysis time compared to manual methods. To our knowledge, however, automatic sulcal extraction

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.18.955096; this version posted March 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

suffers from reliably identifying sulci and hence struggle to define starting and ending points of sulci. Moreover, automatic sulcal extraction methods do not only find major sulcal lines but also small and very minor sulcal lines (secondary or tertiary sulci) which can further complicate sulcal labeling and even introduce unwanted noise [21]. Sulcal lines can also be defined in wrong sulcal regions, or not defined in true sulcal regions [34, 35, 38].

During the training phase, optimal weight vectors per subject and per sulcus were found. Subsequently, a generalized binary "template" for each group of sulcal traces was created using a simple cluster median taken across all representative sulci. Thus, for example, all the right hemisphere cingulate weight vectors were analyzed to create a "template" right hemisphere cingulate weight vector. The underlying median and deviation for the group template bits were also analyzed for their relative stability – somewhat interestingly we found that the bit controlling curvature in the sulcus parametrization was less stable (with higher group deviation) than the bits weighing distance between neighboring vertices and sulcal depth. This suggests that sulcal descriptions might not be easily geometrically intuitive.

These template vectors represented thus a group-weighted description for an entire class of sulci (a full sulcus description was contained in the template binary weight and the start/end position of the given sulcus on a reconstructed surface mesh). The performance of these template vectors was measured, and compared with the per-subject optimals performed with an average of 80% overlap with about a 10% standard deviation (the optimal weights, relative to the human reference, had an average 91% overlap with a 3% deviation). We also considered so-called NULL templates with all weights set to zero as an absolute base reference. The NULL templates performed measurably poorly with only 55% average overlap with the human reference.

Finally, we presented a blinded human expert evaluator with sets of three sulci: one set contained the human target sulci for a subject, a second set contained the per-subject optimals as generated by this system, and a third set contained the sulcal paths generated using sulcal-template weights. Without knowledge of which set belonged to which process, the evaluated ranked individual sulci on a four point scale, with "1" being completely unacceptable and "4" completely acceptable. For the most part, the paths created by the template weights, though ranked consistently lower than the human targets or the optimal weights, were still quite acceptable to the evaluator.

Two interesting observations were made in the evaluator experiment. The first observation was that unacceptable paths were sometimes generated by the template weight for the cingulate sulcus, implying that a future automated system based on this work would still require special attention in the case of the cingulate. The second observation was that in a few cases, the human reference contained some errors as far as the evaluator was concerned. Interestingly, the general binary template sulcus weight, when applied to the same subjects containing human errors, were able to outperform the human reference. This indicates that enough information was generalized from the same sulcus that was correctly traced in other subjects.

Taken as a whole, the results presented here demonstrated that per-sulcus optimizations using an inverse Dijkstra-based method showed a very high degree of correlation between human target traces and automated ones. Moreover, generic template weight vectors generated from large sets of subject- and sulcus-hemisphere specific optimal weight vectors resulted, with the exception of the cingulate sulcus in a few cases, in trajectories that were acceptable to an independent human evaluator. We believe that the relative length of the cingulate sulcus compared to other sulci is a major reason for the poor performance in finding a binary description. The underlying geometry over such a long length could conceivably change in such a manner as to not be captured by our nine-element binary vector.

We have presented a method based on Dijkstra's algorithm that can describe human sulcal paths with a nine-element binary weight vector. Per subject, per sulcus optimizations using this nine-bit binary description show on average a 91% overlap with reference human generated targets. Generalizations for groups of sulci based on per-subject optimizations perform well, indicating that such a binary description could be a useful and compact descriptor for sulci when coupled with a Dijkstra path optimizer.

Future work will seek to expand the weight vector with the addition of more parameters, including the Gaussian curvature and functions of the per-vertex principle curvatures. Since we were able to demonstrate good performance with a nine-bit binary description, it is our hope that performance can be measurably improved with straightforward (but still manageable) extensions to the search space size. In addition, more sophisticated generalization methods will be explored, and in the case of problematic cingulate sulcus, we propose to break this single sulcus into multiple joined segments, and optimize each segment independently.

References

[1] J. Andoh, E. Artiges, C. Pallier, D. Riviere, J.F. Mangin, A. Cachia, M. Plaze, M.L. Paillere-Martinot, and J.L. Martinot. Modulation of language areas with functional MR image-guided magnetic stimulation. *Neuroimage*, 29(2):619–27, 2006.

[2] A. Bartesaghi and G. Sapiro. A system for the generation of curves on 3D brain images. *Hum Brain Mapp*, 14(1):1–15, 2001.

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.18.955096; this version posted March 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

[3] Brian A. Wandell and Suelika Chial and Benjamin T. Backus. Visualization and Measurement of the Cortical Surface. *Journal of Cognitive Neuroscience*, 12(5):739–752, 2000.

[4] A. Cachia, J.F. Mangin, D. Riviere, F. Kherif, N. Boddaert, A. Andrade, D. Papadopoulos-Orfanos, J.B. Poline, I. Bloch, M. Zilbovicius, P. Sonigo, F. Brunelle, and J. Regis. A primal sketch of the cortex mean curvature: a morphogenesis based approach to study the variability of the folding patterns. *IEEE Trans Med Imaging*, 22(6):754–65, 2003.

[5] A. Cachia, J.F. Mangin, D. Riviere, D. Papadopoulos-Orfanos, F. Kherif, I. Bloch, and J. Regis. A generic framework for the parcellation of the cortical surface into gyri using geodesic Voronoi diagrams. *Med Image Anal*, 7(4):403–16, 2003.

[6] A M Dale, B Fischl, and M I Sereno. Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage*, 9(2):179–194, Feb 1999.

[7] E. W. Dijkstra. A note on two problems in connection with graphs. *Numerische Mathematik*, 1:269–271, 1959.

[8] Manfredo P. do Carmo. *Differential Geometry of Curves and Surfaces*. Prentice-Hall, Inc., New Jersey, 1976.

[9] Christine Fennema-Notestine, I Burak Ozyurt, Camellia P Clark, Shaunna Morris, Amanda Bischoff-Grethe, Mark W Bondi, Terry L Jernigan, Bruce Fischl, Florent Segonne, David W Shattuck, Richard M Leahy, David E Rex, Arthur W Toga, Kelly H Zou, and Gregory G Brown. Quantitative evaluation of automated skull-stripping methods applied to contemporary and legacy images: Effects of diagnosis, bias correction, and slice location. *Hum Brain Mapp*, 27(2):99–113, Feb 2006.

[10] P A Filipek, M Semrud-Clikeman, R J Steingard, P F Renshaw, D N Kennedy, and J Biederman. Volumetric MRI analysis comparing subjects having attention-deficit hyperactivity disorder with normal controls. *Neurology*, 48(3):589–601, Mar 1997. Clinical Trial.

[11] B Fischl, M I Sereno, and A M Dale. Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. *Neuroimage*, 9(2):195–207, Feb 1999.

[12] L. D. Griffin. The intrinsic geometry of the cerebral cortex. *J Theor Biol*, 166:261–273, 1994.

[13] P. Hellier and C. Barillot. Coupling dense and landmark-based approaches for nonrigid registration. *IEEE Trans Med Imaging*, 22(2):217–27, 2003.

[14] M A Hofman. The fractal geometry of convoluted brains. *J Hirnforsch*, 32(1):103–111, 1991.

[15] K Im, J M Lee, U Yoon, Y W Shin, S B Hong, I Y Kim, J S Kwon, and S I Kim. Fractal dimension in human cortical surface: Multiple regression analysis with cortical thickness, sulcal depth, and folding area. *Hum Brain Mapp*, 27(12):994–1003, Dec 2006.

[16] Lili Ju, Monica K Hurdal, Josh Stern, Kelly Rehm, Kirt Schaper, and David Rottenberg. Quantitative evaluation of three cortical surface flattening methods. *Neuroimage*, 28(4):869–880, Dec 2005.

[17] N. Khaneja, M. I. Miller, and U. Grenander. Dynamic programming generation of curves on brain surfaces, 1998.

[18] V G Kiselev, K R Hahn, and D P Auer. Is the brain cortex a fractal? *Neuroimage*, 20(3):1765–1774, Nov 2003.

[19] M. Koop, G. Rilling, A. Herrmann, and H. J. Kretschmann. Volumetric development of the fetal telencephalon, cerebral cortex, diencephalon, and rhombencephalon including the cerebellum in man. *Bibl Anat*, pages 53–78, 1986.

[20] In Kyoon Lyoo, Young Hoon Sung, Stephen R Dager, Seth D Friedman, Jun-Young Lee, Seog Ju Kim, Namkug Kim, David L Dunner, and Perry F Renshaw. Regional cerebral cortical thinning in bipolar disorder. *Bipolar Disord*, 8(1):65–74, Feb 2006.

[21] I. Lyu, J. K. Seong, S. Y. Shin, K. Im, J. H. Roh, M. J. Kim, G. H. Kim, J. H. Kim, A. C. Evans, D. L. Na, and J. M. Lee. Spectral-based automatic labeling and refining of human cortical sulcal curves using expert-provided examples. *Neuroimage*, 2010.

[22] M. K. Hurdal and J. B. Gutierrez and C. Laing and D. A. Smith. Shape analysis for automated sulcal classification and parcellation of MRI data. *Journal of Combinatorial Optimization*, 15(3):257–275, April 2008.

[23] M Meyer and M Desbrun and P Schröder and A. H. Barr. Discrete Differential-Geometry Operators for Triangulated 2-Manifolds. *Visualization and Mathematics*, 2002.

[24] M. Ono and S. Kubik and C. D. Abernathy. *Atlas of Cerebral Sulci*. Thieme, Stuttgart, 1990.

[25] Nikos Makris, Steven M Hodge, Christian Haselgrove, David N Kennedy, Anders Dale, Bruce Fischl, Bruce R Rosen, Gordon Harris, Verne S Jr Caviness, and Jeremy D Schmahmann. Human cerebellum: surface-assisted cortical parcellation and volumetry with magnetic resonance imaging. *J Cogn Neurosci*, 15(4):584–599, May 2003.

[26] Nikos Makris, John E Schlerf, Steven M Hodge, Christian Haselgrove, Matthew D Albaugh, Larry J Seidman, Scott L Rauch, Gordon Harris, Joseph Biederman, Verne S Jr Caviness, David N Kennedy, and Jeremy D Schmahmann. MRI-based surface-assisted parcellation of human cerebellar cortex: an anatomically specified method with estimate of reliability. *Neuroimage*, 25(4):1146–1160, May 2005.

[27] J.F. Mangin, D. Riviere, A. Cachia, E. Duchesnay, Y. Cointepas, D. Papadopoulos-Orfanos, D.L. Collins, A.C. Evans, and J. Regis. Object-based morphometry of the cerebral cortex. *IEEE Trans Med Imaging*, 23(8):968–82, 2004.

[28] J.F. Mangin, D. Riviere, A. Cachia, E. Duchesnay, Y. Cointepas, D. Papadopoulos-Orfanos, P. Scifo, T. Ochiai,

bioRxiv preprint doi: https://doi.org/10.1101/2020.02.18.955096; this version posted March 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

F. Brunelle, and J. Regis. A framework to study the cortical folding patterns. *Neuroimage*, 23 Suppl 1:S129–38, 2004.

[29] J.F. Mangin, D. Riviere, A. Cachia, D. Papadopoulos-Orfanos, D.L. Collins, A.C. Evans, and J. Regis. Object-based strategy for morphometry of the cerebral cortex. *Inf Process Med Imaging*, 18:160–71, 2003.

[30] J. E. McLennan and F. H. Grilles. *The Developing Human Brain*, chapter A model of growth of the human fetal brain, pages 43–58. John Wright PSG Inc, Boston, MA, 1983.

[31] D Nain, S Haker, A Bobick, and A Tannenbaum. Multiscale 3-d shape representation and segmentation using spherical wavelets. *IEEE Trans Med Imaging*, 26(4):598–618, Apr 2007.

[32] R. Pienaar and B. Fischl and V. Caviness and N. Makris and P. E. Gran. A methodology for analyzing curvature in the developing brain from preterm to adult. *International Journal of Imaging Systems and Technology*, 18:42–68, 2008.

[33] J. Regis, J.F. Mangin, T. Ochiai, V. Frouin, D. Riviere, A. Cachia, M. Tamura, and Y. Samson. "Sulcal root" generic model: a hypothesis to overcome the variability of the human cortex folding patterns. *Neurol Med Chir (Tokyo)*, 45(1):1–17, 2005.

[34] J. K. Seong, K. Im, S. W. Yoo, S. W. Seo, D. L. Na, and J. M. Lee. Automatic extraction of sulcal lines on cortical surfaces based on anisotropic geodesic distance. *Neuroimage*, 49(1):293–302, 2010.

[35] Y. Shi, P. M. Thompson, I. Dinov, and A. W. Toga. Hamilton-Jacobi skeleton on cortical surfaces. *IEEE Trans Med Imaging*, 27(5):664–73, 2008.

[36] Kaleem Siddiqi, Sylvain Bouix, Allen Tannenbaum, and Steven W. Zucker. Hamilton-jacobi skeletons, 1999.

[37] T. Surazhsky and E. Magid and O. Soldea and G. Elber and E. Rivlin. A Comparison of Gaussian and Mean Curvatures Estimation Methods on Triangular Meshes. *{I}{E}{E}{E} International Conference on Robotics and Automation*, 2003.

[38] X. Tao, J. L. Prince, and C. Davatzikos. Using a statistical shape model to extract sulcal curves on the outer cortex of the human brain. *IEEE Trans Med Imaging*, 21(5):513–24, 2002.

[39] P. M. Thompson, K. M. Hayashi, E. R. Sowell, N. Gogtay, J. N. Giedd, J. L. Rapoport, G. I. de Zubicaray, A. L. Janke, S. E. Rose, J. Semple, D. M. Doddrell, Y. Wang, T. G. van Erp, T. D. Cannon, and A. W. Toga. Mapping cortical change in Alzheimer's disease, brain development, and schizophrenia. *Neuroimage*, 23 Suppl 1:S2–18, 2004.

[40] Duygu Tosun, Allan L. Reis, Ursula Bellugi, Albert M. Galaburda, Julie R. Korenburg, Debra L. Mills, Arthur W. Toga, and Paul M. Thompson. Measuring increased cortical complexity in williams syndrome using 3-d cortical morphometry. January 2006.

[41] D C Van Essen, D Dierker, A Z Snyder, M E Raichle, A L Reiss, and J Korenberg. Symmetry of cortical folding abnormalities in williams syndrome revealed by surface-based analyses. *J Neurosci*, 26(20):5470–5483, May 2006.

[42] J J Wisco, G Kuperberg, D Manoach, B T Quinn, E Busa, B Fischl, S Heckers, and A G Sorensen. Abnormal cortical folding patterns within broca's area in schizophrenia: Evidence from structural mri. *Schizophr Res*, May 2007.

[43] P Yu, P E Grant, Y Qi, X Han, F Ségonne, R Pienaar, E Busa, J Pacheco, N Makris, R L Buckner, P Golland, and B Fischl. Cortical surface shape analysis based on spherical wavelets. *IEEE Trans Med Imaging*, 26(4):582–597, Apr 2007.