Supplementary Material

Functional Data Analysis of Tree Data Objects

A Data Description

Our data set is from a study of Magnetic Resonance Angiography (MRA) brain images (Dumoulin and Hart, 1986) of a set of 98 human subjects. One slice of a 3 dimen-

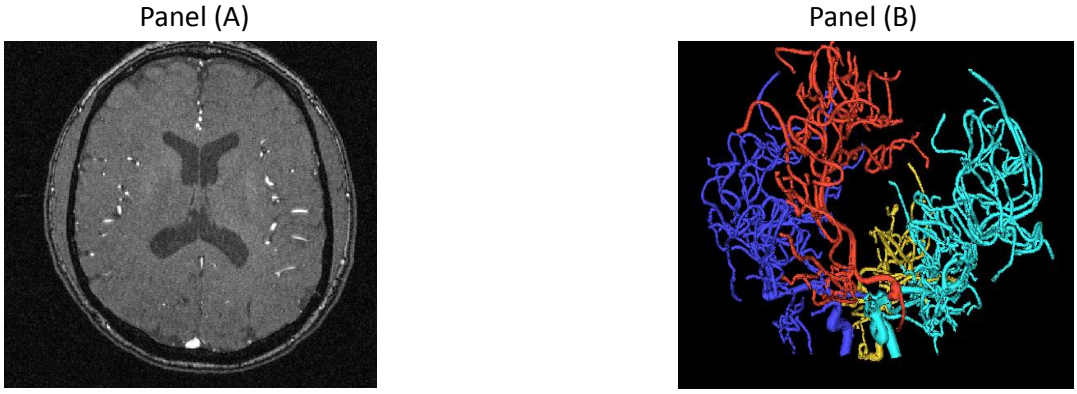

Figure A: Panel (A) is a single slice MRA image for one person with white regions indicating blood flow. The white regions in the 3 dimensional collection of MRA slices generate brain artery trees as shown in Panel (B). The colors in Panel (B) indicate the regions of the brain: Front (red), Back (gold), Left (cyan), Right (blue).

sional MRA image is shown in Panel (A) of Figure A, and it highlights regions of blood flow as white. These white regions are aggregated in 3 dimensions by a tube tracking algorithm, and then combined into brain artery trees, as described in Aylward and Bullitt (2002). The result is shown in Panel (B) of Figure A. These trees are colored according to the region of the brain (using color indicated in the caption of Figure A), and are studied separately. The information stored in these color trees is very rich. Every color tree is a union of branches. We define a *branch* to be a vessel segment between consecutive split points. Each branch is composed of a sequence of spheres, as calculated in Aylward and Bullitt (2002). Each sphere has a center with x, y, z coordinates, indicating the location of the center of the brain artery and its local radius, indicating the arterial thickness at that point. The union of many such spheres gives the vessels shown in the visual rendering in Panel (B) of Figure A.

B Full Size Branch Length Curves

It is difficult to distinguish individual branch length curves in the left panel of Figure 7 in Section 3.3 of the paper, because each branch length curve has many $y = 0$ flat parts due to missing branches. Thus the full size version of this plot is shown in Figure B.

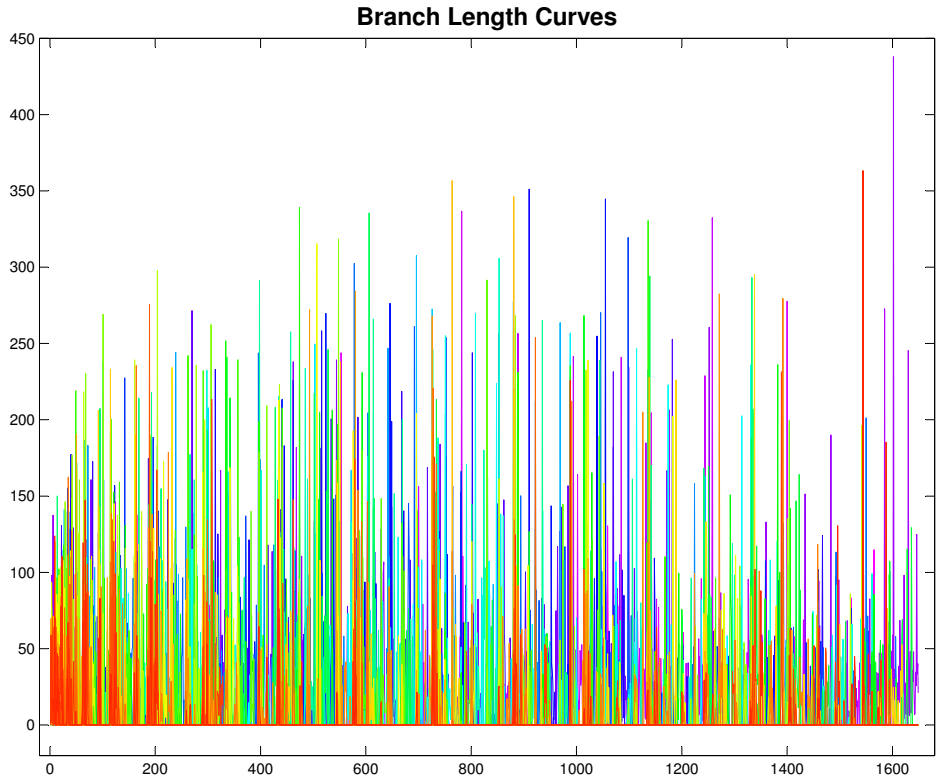

Figure B: The Branch Length Representation of the Full Data Set. The color ranges from magenta (for young) to red (for old).

C PCA of Dyck Path curves

This section provides additional plots for the PCA of Dyck Path curves. Figure C contains a full size version of the mean centered Dyck path curves. In Section 4.1 of the paper, we only studied PC1 projected curves. The PC2 and PC3 projected curves are now provided in Panels (C) and (D) of Figure D. Panel (C) shows that the main PC2 variation is from left to middle, appearing in regions that are disjoint from the PC1 main variation region. The PC3 projection curve plot in Panel (D) shows that the main variation region

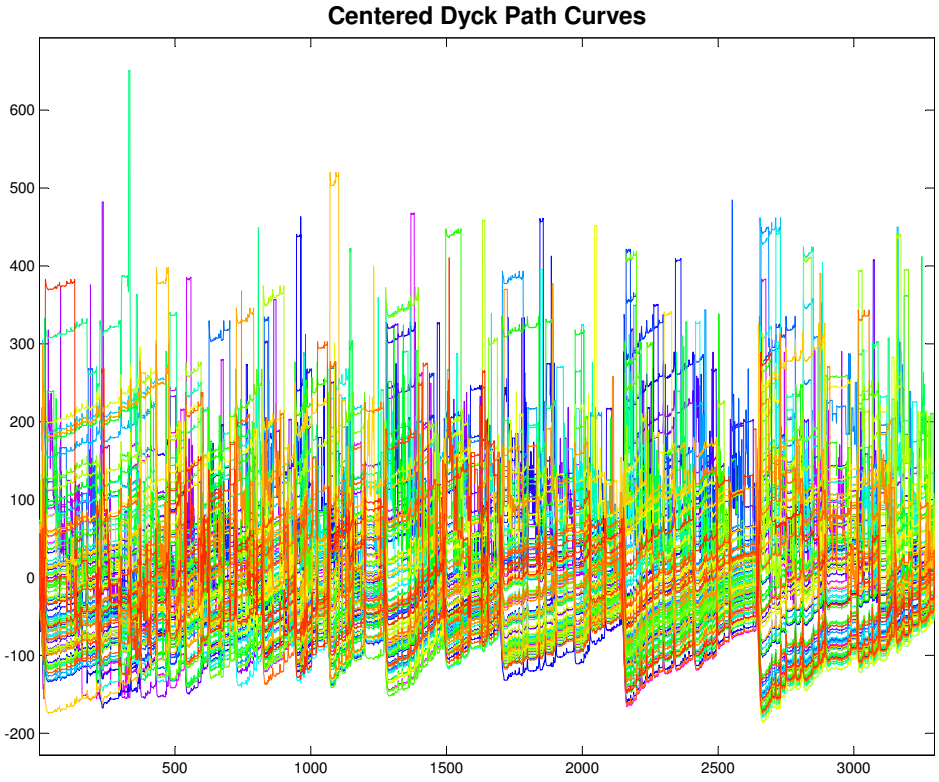

Figure C: The Mean Centered Dyck path Curves. The color again ranges from magenta (for young) to red (for old).

is between the middle and the right, and seems disjoint with the PC1 and PC2 main variation regions.

In addition, the bottom left panel of Figure E is the scree plot that shows the individual and cumulative variability explained by the first 10th PCs. For each of the first three PC directions, one of the diagonal panels plots the corresponding PC scores (in the horizontal axis) against random generately jittered vertical heights, where the points (corresponding to the trees) are colored according to the age, from magenta (for young) to red (for old). A kernel density estimate for the PC scores is superimposed to illustrate the distribution of the scores. Note the bimodal distribution among the PC1 scores. The plots above the diagonal are the pairwise scatter plots of the PC scores on the first three PC directions, where the points are again colored according to age.

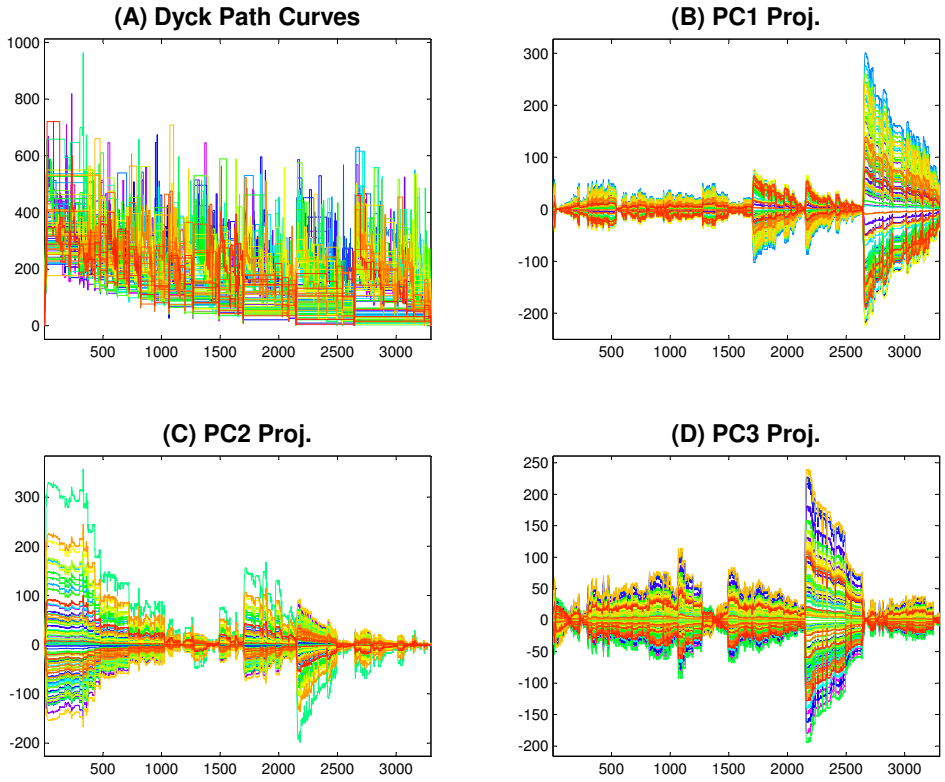

Figure D: PCA of the Dyck Path Curves. The Dyck path curves are first centered and then projected onto the first 3 PC direction to explore the modes of the variation of the Dyck path curves.

References

Aylward, S. and E. Bullitt (2002). Initialization, noise, singularities, and scale in height ridge traversal for tubular object centerline extraction. *IEEE Transactions on Medical Imaging 21*(2), 61–75.

Dumoulin, C. and H. Hart (1986). Magnetic resonance angiography. *Radiology 161*(3), 717–720.

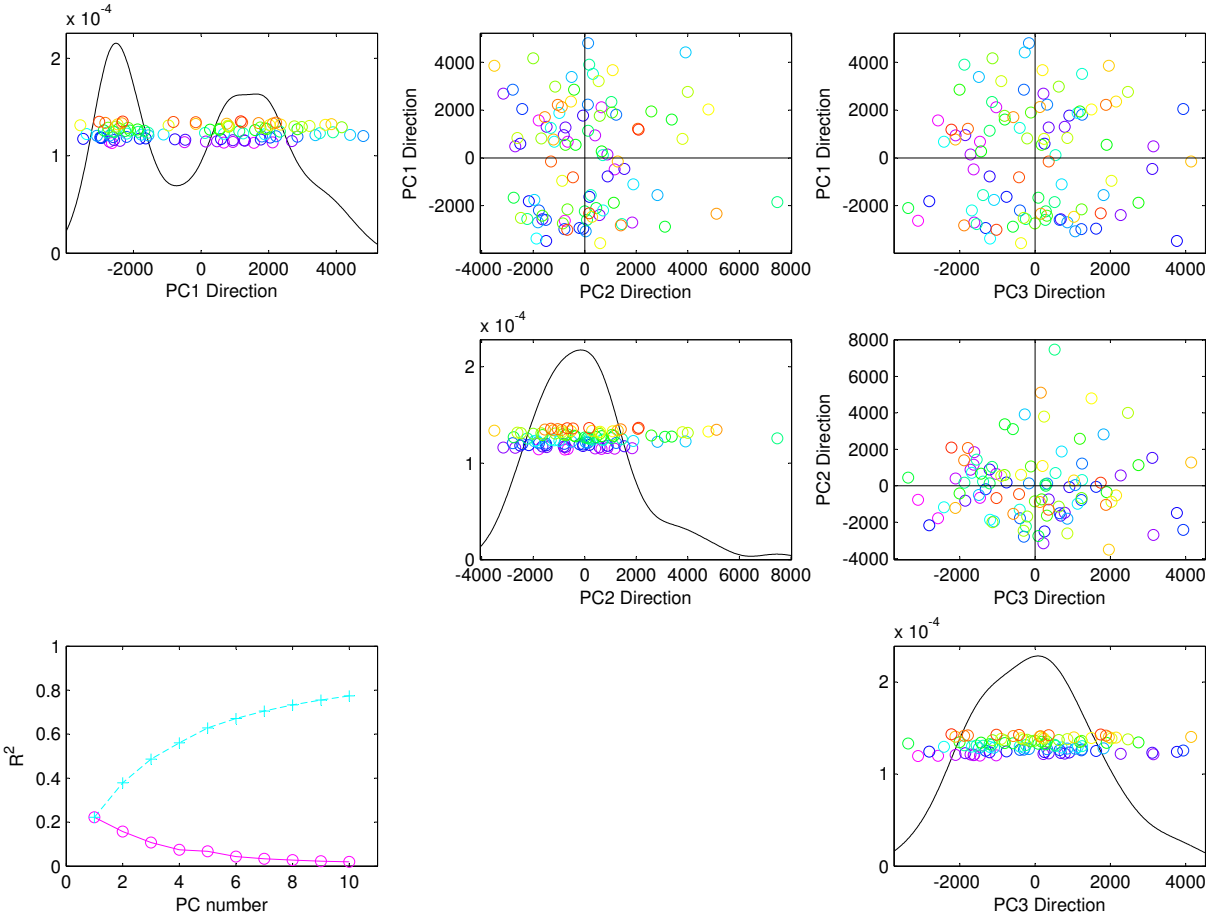

Figure E: Scree Plot and Plots of PC Scores on the First Three PC Directions. The bottom left panel provides the scree plot based on the first ten PCs. The other panels plots PC scores on the first three PC directions.