

Functional Neurodata Graph Service: A One-Click Pipeline and Web Service for Reliable Functional Connectome Estimation

Eric W. Bridgeford^{1,2,*}, Gregory Kiar^{2,3}, Tanay Agarwal¹, Eric Walker^{1,2}, and Joshua T. Vogelstein²

¹Department of Computer Science, Johns Hopkins University

²Department of Biomedical Engineering, Johns Hopkins University

³Department of Biomedical Engineering, McGill University

*ericwb95@gmail.com

1 Abstract

In recent years, functional magnetic resonance imaging (fMRI) analysis has become one of the most widely used techniques for assessing cognitive function. fMRI allows an unparalleled combination of temporal and spatial acuity, providing researchers a glance at the dynamic activity of individual regions in the brain. Unfortunately, computational restrictions of collecting and analyzing an fMRI dataset necessitate technological investments and expertise for interested researchers. Through the Functional NeuroData Graph Service, we develop a robust fMRI processing pipeline for providing automated acquisition of functional connectomes. Furthermore, we provide an open-source scientific container pre-loaded with our software and all dependencies to facilitate the distribution of our methodologies. Finally, we provide utilities and a web-service for deploying our pipeline using Electronic Cloud Computing, enabling researchers to process entire datasets in just a few hours with no external dependencies.

2 Introduction

For the past thirty years, fMRI has presented one of the most intriguing challenges for computational neuroscience. While researchers have long understood that an increase in blood flow was necessary for increased brain function (such as during periods of high brain activity), it was not until a seminal study by [1] that researchers were able to image the flow of blood in the brain. Oxygen-rich blood (high in Hb, the oxygenated form of hemoglobin) is slightly more magnetic than deoxygenated blood (high in dHb, or deoxygenated hemoglobin). As Hb has different magnetic properties than dHb, a high strength magnetic

field such as that of an MRI scanner is able to pick up time-dependent fluctuations in the concentration of Hb. While regions of the brain showing high activity see a brief drop in the concentration of oxygen in vessels supplying the region, the arteries rapidly overcompensate by the hemodynamic response [2], which sees a significant increase in the concentration of oxygenated blood pushed towards the regions with high brain activity. This fluctuation in the Blood Oxygenation Level Dependent (BOLD) signal is what is imaged in an fMRI scanner.

Recently, researchers have been increasingly concerned with developing maps of the brain, or connectomes (TODO: cite connectome). In the functional sense, the connectome represents the strength, or correlation, with which two particular regions of the brain fire together. We record this strength as the "connectivity" of two regions. Repeating over all regions of the brain, we obtain a map, or network, of the brain. We can then apply network theoretical approaches to gain insight into spatial patterns that arise at the region level (TODO: cite something Dani here).

3 Methods

3.1 Pipeline

The question of identifying an optimal processing pipeline is, unfortunately, heavily non-convex in that the parameter settings of components interact such that testing individual processing steps standalone is unlikely to lead to the optimal processing pipeline [3]. To begin to answer this question, we conducted a full parameter sweep of many of the most popular techniques for fMRI analysis in [4]. We then refined each step using a combination of quantitative and qualitative metrics to yield a pipeline which ensures maximal robustness.

3.1.1 Preprocessing

Subject-specific artifacts caused by head motion and limitations of the fMRI sensor are known to be confounding factors in fMRI analyses [5]. The workflow in Figure (1) shows the solution the FNGS pipeline uses to correct for head motion and slice timing issues inherent in the fMRI scanning technique.

Slice Timing To collect an individual 4D EPI sequence, a 3D volume is constructed as a combination of individual 2D slices. The 2D slices are collected incrementally; that is, we collect each 2D slice for approximately 10 milliseconds, and the entire 3D volume is complete in about 30 2D slices. This gives a repetition time, or TR, for the volume on the order of 1 to 3 seconds depending on the scanner. In response to a stimulus, we expect a typical brain response on the order of about 16 seconds. This means that during the course of a single volume being collected, we may have a different amount of BOLD present in the first volume than when the last slice is actually collected [7]. While intuitively we might want our observations to represent a "snapshot" of the data at a given time point, instead we have a "sliding snapshot"; that is, our observations are not all at a fixed point in time.

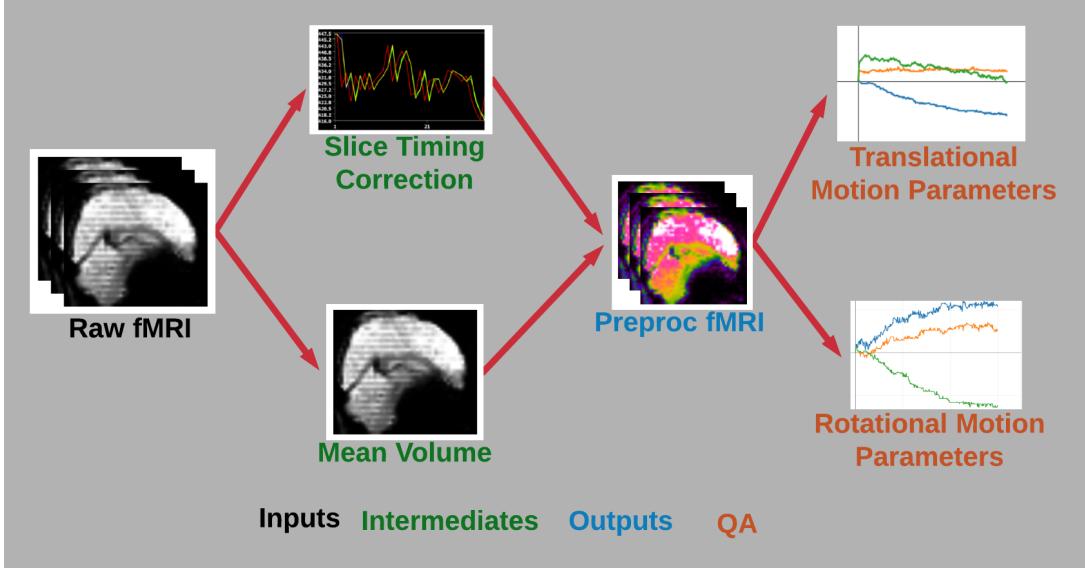


Figure 1: The preprocessing workflow for the FNNGS pipeline.

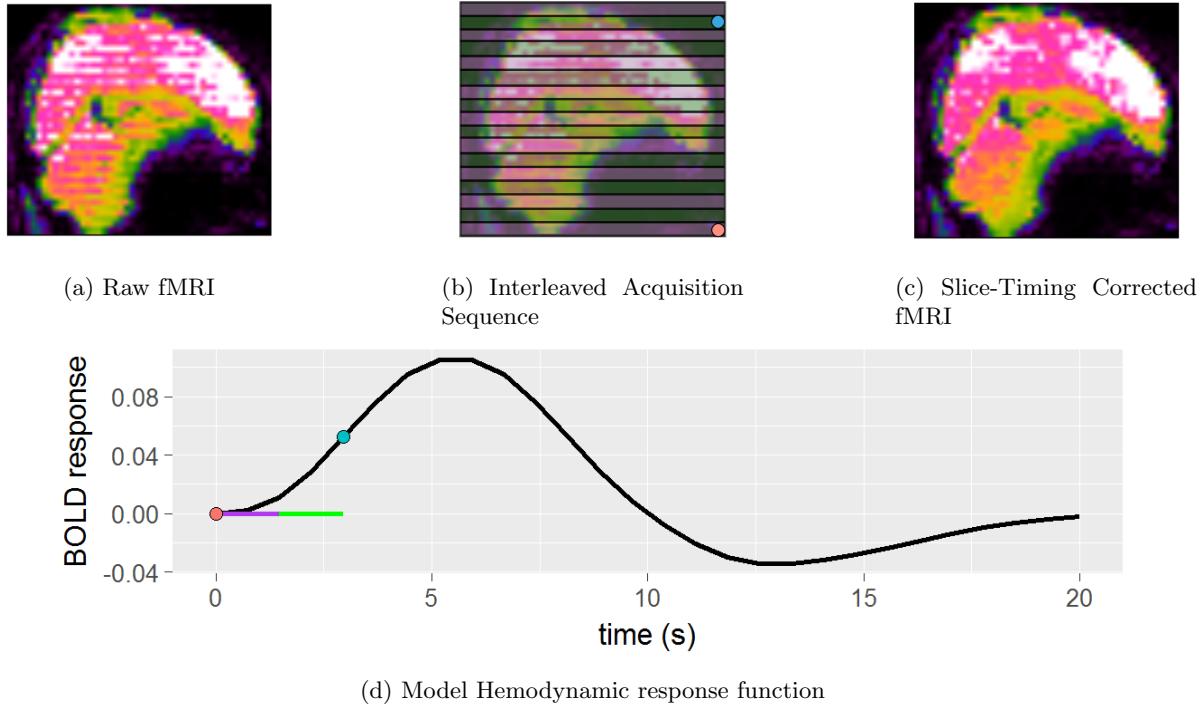


Figure 2: In the raw fMRI shown in (a), we can see the slice-timing effect. Notice the horizontal banding, characteristic of the interleaved-acquisition pattern. In (b), we can see how the interleaved acquisition sequence works. The 3D volume is acquired from bottom to top; first the purple 2D slices, and next the green 2D slices. In (d), we can see an example hemodynamic response, where it is clear that the first slice taken (the red dot) is at a noticeably different time point than the last slice taken (the blue dot) for the first TR (the purple corresponds to the time period when the purple 2D slices are taken, and vice-versa for the green). In (c), we can see the fMRI after slice-timing correction. Data shown is from the HNU1 dataset [6].

To begin, we accept the user-provided acquisition sequence of a single 3D volume. Given the acquisition sequence of the 2D slices, we can compute the TR shift of each 2D slice given the TR information from the header of the brain image. A slice that occurs first in a TR will have a shift of 0, while a slice that occurs at the end of a TR will have a shift of 1. A slice that occurs exactly in the middle of a TR has a shift of 0.5. For each voxel in an individual slice, we use interpolation to re-center our observations to all have a TR shift of 0.5. For details, see Figure (2). We accomplish slicetiming using the slicetimer utility provided by FSL [7]. Optionally, users may choose to skip this step, as slice timing correction is only needed when users are focusing on temporal characteristics of an fMRI session.

Motion Correction During an fMRI session, participants sit in a small, cramped scanner often for between 5 and 10 minutes. During the course of a study, it is fairly common for participants to shift and fidget, even if only small amounts. Small shifts will lead to a person’s head being in different spatial positions at each timestep, which will hamper our efforts to standardize the spatial properties of each subject’s brain down the line through registration. This is because registration entails aligning each 3D volume in a 4D sequence using the same linear or non-linear transformation. This means that if each 3D volume is not already the same spatially, we may get minor inconsistencies in our registration that will generally decrease our functional connectome quality [8].

Fortunately, given that the subject’s brain is the same scaling for each 3D volume (the brain shape itself is not changing in time), we can estimate a 6 degree of freedom rigid affine transformation for each 3D volume (we have 1 translational and 1 rotational parameter per x, y, z direction the subject could move his/her head) using the mean fMRI slice as our reference. Motion correction is implemented using the mcflirt utility [9], which is a simplification of FSL’s FLIRT registration tool.

3.1.2 Registration

During an fMRI session, the goal is to analyze the functional properties associated with a given brain. For each subject, however, we have unique structural properties associated with the shape and layout of that particular brain. Using linear and nonlinear realignments, we can reshape the brains of each subject to that of an average human (or template) brain. With all of the brains in a common spatial domain, we can simply identify the voxels associated with a particular region in the template brain, and then subset these same voxels for all subjects in the dataset. In Figure (3), we can see the registration workflow for the FNGS pipeline.

Registration Score Registration is one of the easiest places for automated fMRI processing to fail. A slight misalignment at one step can have dramatically negative consequences on resulting downstream inferences; we occasionally introduce tearing, distortions, or more serious problems into our scans. To define a quality registration, we introduce the registration score r , defined as:

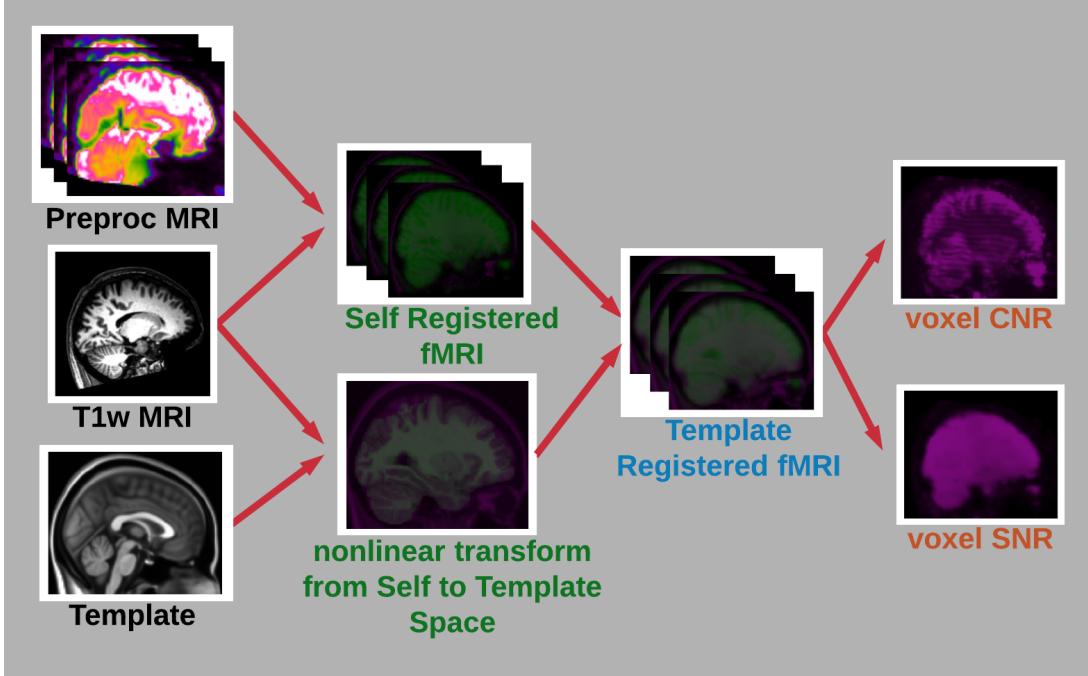


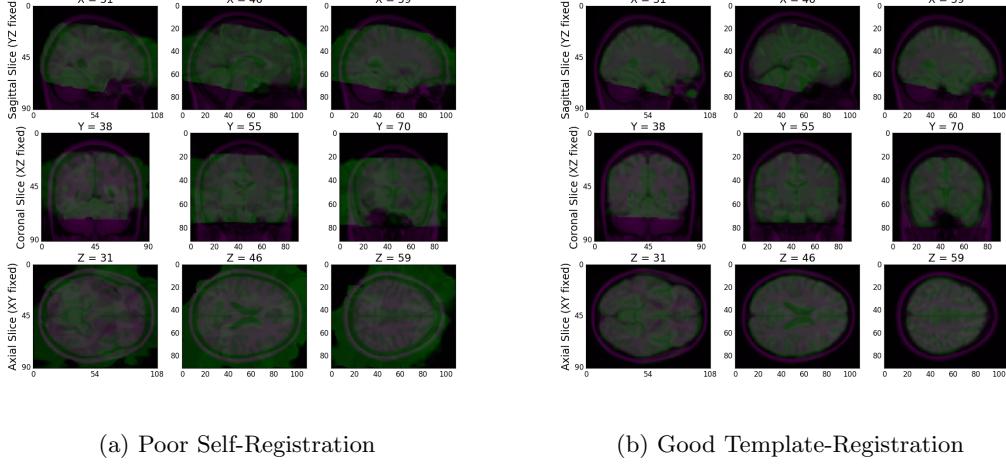
Figure 3: The registration workflow for the FNNGS pipeline. We take as input to the pipeline a preprocessed brain, an anatomical brain, and a template brain. We estimate a linear transformation from the fMRI to the structural image to account for the higher resolution of the anatomical image, and estimate a non-linear transformation from the anatomical image to the template image. Finally, we apply the non-linear transformation to the functional brain in the anatomical brain space.

$$r_\gamma(A, B) = \frac{|A \cap B|}{|A \cup B|}$$

for a brain A being aligned to a reference B using registration procedure γ , where a higher quality registration has a value of $r_\gamma(A, B) > c$ for some cutoff c . This statistic captures the intuition of registration in its most basic form; we expect that properly registered brains will have a spatial overlap that is comparatively sized to their spatial union. Often, techniques that are less robust provide better alignments on high-quality data. Since registration failures are not immediately apparent in downstream derivatives, automatically recognizing when one particular technique fails is of the utmost importance.

Self Registration When collecting an fMRI sequence, researchers also collect a higher resolution anatomical scan, using a procedure known as T1-weighting [10]. Using the T1w imaging scheme, researchers can collect imaging data at resolutions exceeding $1\text{mm} \times 1\text{mm} \times 1\text{mm}$, whereas fMRI sequences will often be of substantially lower resolution (generally around $3\text{mm} \times 3\text{mm} \times 3\text{mm}$). This enhanced resolution makes identifying anatomical landmarks far easier using the T1w than the fMRI sequence, which will greatly improve our template registration down the line.

To register our input fMRI to our reference T1w image, we begin by first estimating a locally-optimized



(a) Poor Self-Registration

(b) Good Template-Registration

Figure 4: Above, we can see the quality control image for one brain analyzed in our pipeline. In (a), we can see a low-quality example of a template-registration error due to FNIRT distorting our image. This brain has $r = 0.46$ with respect to the template after alignment. In (b), we show the brain after it is realigned instead using FLIRT with a 12 degree of freedom linear registration.

affine transformation with FSL’s FLIRT [11]. This locally-optimised transformation from our fMRI space to our T1w space is heavily robust, and has its hyperparameters tuned to focus on local features of the input (fMRI) and reference (T1w) spaces. We use this cost function because often, input fMRI will have narrow fields of view, resolution constraints, or tearing that will perform poorly using a more global alignment.

Next, we compute a second alignment using our locally-aligned fMRI to our structural T1w space using the more advanced epireg function provided by FSL [12]. The epireg generally produces higher quality registrations from functional to structural space, and makes use of the powerful boundary-based registration cost function to allow us to align individual gyri between the functional and structural sequences. To use epireg, however, requires a high quality anatomical scans that are high enough quality to allow for a near-perfect segmentation into white matter, gray matter, and cerebrospinal fluid (csf) probability maps. Additionally, if our fMRI provides a low contrast, we will not be able to fit individual gyri in the brain, which could further distort our resulting alignment. We compute our registration score between our registered fMRI and T1w brains, and if our registration score is below $c = 0.8$, we instead use our more robust locally-optimized FLIRT alignment instead.

Template Registration Now that our fMRI sequence is in T1w space, we can align our brain to a reference template. A template brain represents the anatomical average brain of the sampling of subjects it is collected over. This anatomically average brain theoretically represents the average brain we will find during our external investigations, allowing minimal alignment on average. For FNGS, we assume that users will be using the MNI152 template [13].

We first perform a gentle linear transformation of our high-resolution T1w brain to our reference template brain using a more extensive 12-DOF global FLIRT alignment than for the self-registration case. Given that our template brain will theoretically be less similar than simple translations, rotations and scalings can provide, we use a non-linear registration from our T1w to our template space. We accomplish this using FSL’s FNIRT algorithm [14], with hyper-parameter tuning specific for the MNI152 template. We then apply this non-linear transformation to our fMRI sequence in our T1w space.

Like epireg, FNIRT assumes that the input anatomical image is high enough quality that anatomical features can be extracted and individually aligned to our template. If these features cannot be extracted, our resulting transformation may introduce error to our registration. To account for this, we again compute our registration score, and define a successful registration as having a score greater than $c = 0.8$. In Figure (4), we can see an example of a brain that is poorly aligned using FNIRT, recognized as having poor alignment by the registration score, and then realigned with a more robust FLIRT technique.

3.1.3 Nuisance Correction

Linear Drift Removal Over the course of an fMRI scanning session, many sources of noise arise that must be corrected for in order to make quality data inferences. The scanner inadvertently heats up (producing a high strength magnetic field for sessions lasting up to ten minutes produces an enormous amount of heat). As the scanner heats, the fine electronic equipment inadvertently drifts in the signal it picks up (first demonstrated by [15] when they showed that a heated scanner detected "brain activity" in cadavers). This drift has been shown to be approximately quadratic [16].

To fit the quadratic drift, we use a General Linear Model (GLM). For our n voxels, the t timestep BOLD signal, we can decompose $T_{raw} \in \mathbb{R}^{t \times n}$ as:

$$\begin{aligned}
T_{raw} &= \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \begin{bmatrix} c_{c,1} & c_{c,2} & \dots & c_{c,n} \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ \vdots \\ t \end{bmatrix} \begin{bmatrix} c_{l,1} & c_{l,2} & \dots & c_{l,n} \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ \vdots \\ t^2 \end{bmatrix} \begin{bmatrix} c_{q,1} & c_{q,2} & \dots & c_{q,n} \end{bmatrix} + \epsilon \\
&= \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ \vdots & \vdots & \vdots \\ 1 & t & t^2 \end{bmatrix} \begin{bmatrix} c_{c,1} & c_{c,2} & \dots & c_{c,n} \\ c_{l,1} & c_{l,2} & \dots & c_{l,n} \\ c_{q,1} & c_{q,2} & \dots & c_{q,n} \end{bmatrix} + \epsilon \\
&= RC + \epsilon
\end{aligned}$$

where ϵ is the timeseries corrected without the quadratic RC . Since we know that brain dynamics

behave non-quadratically, we know that minimizing the squared-error loss of T with respect to RC will provide a best-estimate of the coefficients C of our quadratic regressors R , where our desired, quadratic-removed timeseries is just ϵ since we know that ϵ is the components of our raw signal that cannot be fit by a quadratic [16]. We can solve using the least-squares solution to regression problems:

$$C = (R^T R)^{-1} R^T T$$

Using our estimate C for our regressors R , we can then find our quadratic-corrected timeseries to be:

$$\epsilon = T_{raw} - RC$$

Low Frequency Drift Removal Using our quadratic-corrected timeseries, we can then remove low-frequency drift that may still be present in our data. We know that any physiological response due to a stimulus will have a period of around 16 seconds (or a frequency of 0.2 Hz), as it has been shown in [17]. Using this information, we can select to remove sinusoidal components with frequencies far exceeding those of our predicted responses. Conservatively, we set a threshold of 0.01 Hz for highpass-filtering out low-frequency noise.

T1 Effect Removal During the fMRI session, the first few volumes may appear to have brighter intensities as the T1 effects are not fully saturated (TODO: cite this effect). External attempts that remove the T1 effect include component-correction and global mean normalization of each slice, both of which have been shown to potentially remove brain signal [18]. To account for this, we simply discard the first 2 volumes from our fMRI sequence, which tends to account for the majority of contrast-dependent issues (TODO: include an image of a timeseries with the T1 effect, and corrected with slices removed).

3.1.4 Connectivity Estimation

Voxel Timeseries In fMRI processing, we often want to think about each voxel in the brain itself as a function of the time spent in the scanner. This allows us the ability to perform further analyses leveraging the temporal and spatial characteristics of each voxel. To accomplish this, we take a mask of known brain voxels associated with the registration template, and we extract the voxels of known brain tissue.

ROI Timeseries Similarly, it is often of neurological significance to think of the brain as segmented into regions of neurons performing a similar task based on their spatial layout. To accomplish this,

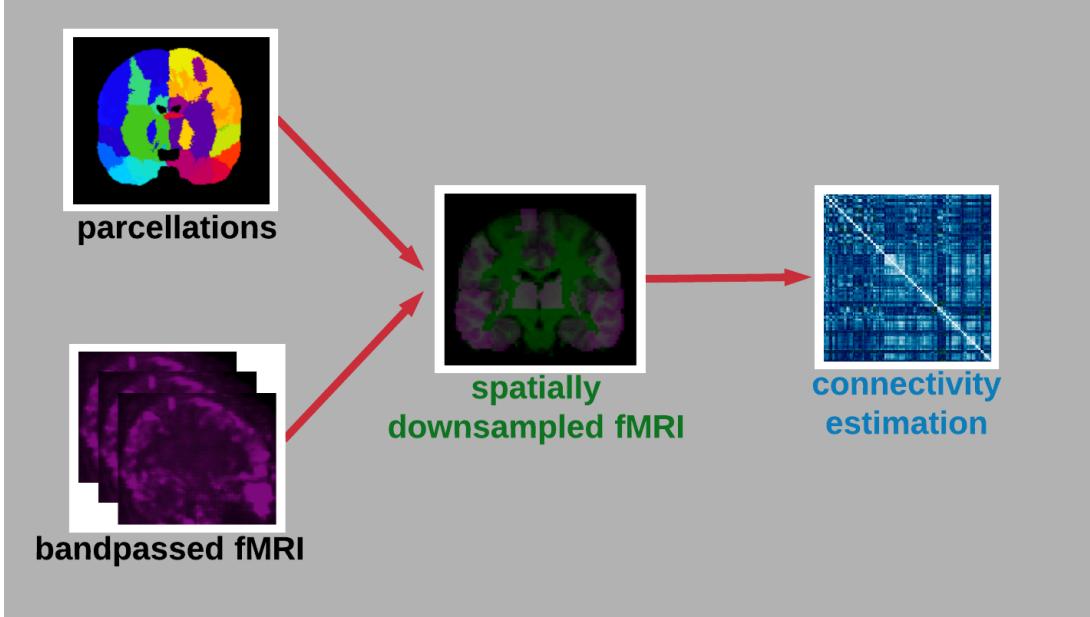


Figure 5: The connectivity matrix workflow for the FNGS pipeline.

neuroscientists parcellate the registered brains into parcellation atlases, whereby they characterize individual voxels in the registration template into individual functional or structural groups. For each region in the parcellation atlas, we subset all of the voxels associated with the given region of interest (ROI) and average them spatially. This gives us a significantly downsampled representation of the brain, often yielding in excess of 100 fold compression and also reducing voxel-by-voxel noise that may still be present. The downsampled timeseries can then be analyzed using traditional statistical methods with greater robustness and computational efficiency [19].

Finally, we can use our ROI timeseries to estimate a connectome. For our connectome, we want each edge to represent a relationship between two regions of the brain. Intuitively, the functional timeseries gives a functional relationship, whereby the simplest comparison is in terms of the firing simultaneity of two regions. For this, we estimate the correlation between each pair of regions from our ROI timeseries for each region in the brain. This yields us a simplified network representation of our original brain.

3.2 Discriminability

Throughout our investigation, we assume that our data is structured as:

$$T_{n,t} = g_{\psi,t}(f_{\varphi,t}(v_n))$$

Thus, we take an explicit observation $T_{n,t}$ of some latent signal v_n for subject n . However, this signal is distorted, first by measurement distortion $f_{\varphi,t}$ and second by the processing options we have chosen $g_{\psi,t}$. Here, φ represents error introduced by scanner parameters, and ψ represents error introduced

by our processing options themselves at a particular measurement session t . As the measurement and processing-dependent distortions are random and unknown, we intuitively want a robust pipeline to be one in which $T_{n,t}$ is as close as possible to $T_{n,t'}$, or the measurement of subject n taken at a time t' . To accomplish this, we measure the discriminability $r(\psi, \varphi)$ [4], which provides a statistic that is near 1 when the observations $T_{n,t}$ and $T_{n,t'}$ are similar, and near 0.5 when $T_{n,t}$ and $T_{n',t'}$ for a different subject n' are indistinguishable. Then we seek the pipeline that maximizes:

$$\psi_{best}, \varphi_{best} = \operatorname{argmax}_{\psi, \varphi} r(\psi, \varphi)$$

During a resting-state MRI session, researchers do not control the stimuli leading to brain responses in time. Theoretically, the idea is to measure the resting-state fluctuations unique to a particular person, which may or may not be the exact same from one session to the next. Then comparing timeseries to timeseries does not make experimental sense; we have no idea what stimuli are impacting the observable dynamics of the brain at a given time, so we cannot assume that they are the same from one session to the next. To overcome this obstacle, we instead assume that, over the course of a scanning session, we get enough variation of a person’s brain activity to be able to make inferences about which regions in the brain function together. We tune FNNGS with respect to the maximal discriminability with the scan id as the label. This provides an upper-bound on downstream inference errors, by ensuring maximum similarity between observations of the same subject that are thought to be maximally similar [4].

To estimate this functional connectivity, we try several strategies:

Correlational Approach The simplest and most intuitive functional connectome is the connectivity matrix as we have previously described and provide natively as the connectome estimation in the FNNGS pipeline. Here, we consider the correlation over time between each pair of edges in our ROI timeseries [20]. With this definition of a spatially-motivated functional connectome, we assume that over time, regions that are more connected will tend to fire simultaneously in a similar way from session to session.

Fourier Approach Moreover, it might also be valid to think of the timeseries in terms of its fourier properties. While temporal properties themselves may not be consistent from session to session, instead it may be the case that the frequency with which individual ROI timeseries operate is more characteristic of a given person’s brain activity [21]. Following Chen et. al, we first consider the fourier transform for each ROI, and highpass filter any activity below the 0.01 Hz range, as this activity can be thought of as low-frequency drift. We normalize the remaining fourier components to form a probability distribution. Next, we consider the Kullback-Leibler divergence between each pair of regions to obtain a spatial connectome

for each subject that instead considers the frequency-domain similarities between regions.

Ranking In addition, for both the correlational and fourier approaches above, we consider the ranked approximation of each connectome. To compute the ranked connectome, We rank each edge in the connectome from lowest to highest value with a scalar ranking. Essentially our hypothesis is that this allows relative robustness, is affine-transformation invariant (ie, a global scaling of the edge weights will not be reflected after ranking), and allows us to recover monotonic nonlinear fluctuations that may be present better than a linear-transformation such as z-scoring of the correlations [4].

The distance matrix is required for computation of the discriminability score from a connectome.

Frobenius Norm of Difference For our correlational estimates, we consider the distance matrix D to be:

$$D(A, B) = \|A - B\|_F$$

or the frobenius norm of the difference each pair of connectomes A, B obtained in our study, where A, B are correlation-derived connectomes.

Hellinger Distance For our frequency-derived connectomes, we consider the distance matrix H to be:

$$H(C, D) = \frac{1}{\sqrt{2}} \|\sqrt{C} - \sqrt{D}\|_2$$

or the hellinger distance between each pair of connectomes C, D obtained in our study, where C, D are fourier-derived connectomes.

3.3 Software

3.3.1 Docker Container

Through FNNGS, we have developed a reliable functional connectome estimation solution. To maximize the deployability of our software, we have developed a publicly-accessible docker container preloaded with all of our software. Using docker ensures compatibility with any modern (Windows, Linux, OSx) computing platform, as all dependencies are folded into the container directly.

3.3.2 Deployment

Leveraging our docker container, we are able to offer multiple levels of access for the FNNGS pipeline. The pipeline can be directly on single-scans using any input format they choose with a local entrypoint in the docker container. For higher levels of analysis on large numbers of scans at once, we require

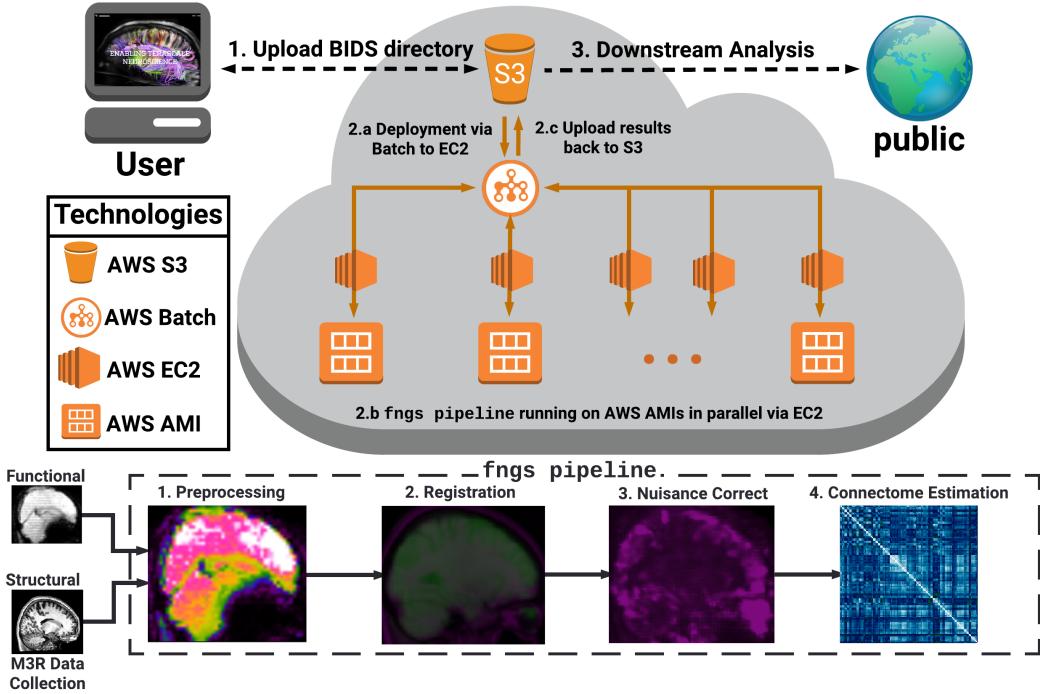


Figure 6: The deployment workflow for the FNNGS pipeline. Users provide a directory according to the BIDs spec to the FNNGS cloud controller via the FNNGS web-service or the docker container at (1). The data is uploaded directly to AWS S3 cloud drives. The controller then initiates the Batch deployment procedure at (2.a), which interfaces between S3 and EC2 cloud computers to provide the MRI scans to EC2 instances pre-loaded with the FNNGS pipeline for analysis. After the scans are finished being analyzed on the EC2 instances at (2.b), the results are then re-uploaded back to the S3 cloud drive at (2.c). The user can then navigate to the S3 cloud drive for downstream analyses.

users to format their data according to the BIDs spec [22], a popular structure for formatting fMRI, diffusion MRI, and various other modalities of neuroimaging data. Users can also optionally call the FNNGS pipeline on their local data using another endpoint in the docker container. If users do not have a high performance computer or a cluster, we leverage Amazon Web Services [23] to allow users to batch-analyze large amounts of data for relatively low cost. We provide a command-line endpoint and automation scripts that allows users to scale their analyses in parallel on the cloud with virtually no setup required other than appropriate configuration of their Amazon Web Services account. The procedure can be seen in Figure (6) and is detailed in our provided tutorials below.

3.3.3 Resources

Tutorials for usage and demos, and links to the code, and our live version of the website can be found at <https://github.com/neurodatadesign/fngs>. All of our software is open source and coded in python. An exhaustive high-level overview of the pipeline and all derivatives can be found at:

https://neurodatadesign.github.io/fngs/about_fngs/Schematic.html. Note that the image is interactive.

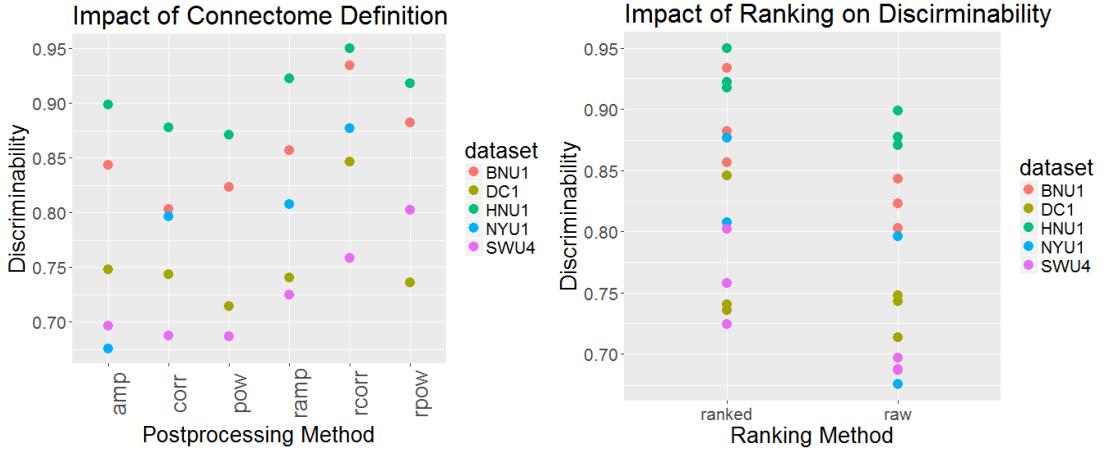
4 Results

4.1 Quality Assurance

The FNNGS pipeline comes pre-packaged with automatic subject-specific quality assurance at each step in the pipeline. Quality assurance is done step-wise to allow researchers assured confidence that each step in the pipeline is performing as-intended on the scans. Downstream connectome acquisition failures are not readily apparent, so visualizations and statistical analyses of the processing procedure are invaluable for aiding researchers in excluding bad or poorly analyzed data. The FNNGS group analysis protocol produces plots of the intra-dataset quantitative properties for all scans in the dataset. This allows users to easily identify potential failed scans more easily than manually looking at the results for each subject so they can then look at the subject-specific quality assurance to identify whether a scan should be excluded from further analysis. Our overview schematic following the link in the resources section above contains an exhaustive look at the pipeline and all derivatives.

4.2 Impact of Connectome Estimation Metric

As we can see in Figure (7), the metric we use to define our connectome plays a substantial role on the within-dataset discriminability results we obtain. Within-dataset discriminability is the discriminability score using the subject ids as labels for single datasets; that is, assessing the closeness of the between-subject distances compared to the out-of-subject distances. Noting in figure (7.a) that the ranking approaches tend to appear to outperform the unranked, or raw edge, approaches, in (7.b) we plot the ranked-edge compared to raw edge approaches. We can see that ranking the edges tends to improve our



(a) Comparison of all 6 permutations of the definitions of a functional connectome (b) comparison of ranking vs. raw (no ranking) of connectome edges

Figure 7: In (7), we can see the discriminability results of each definition of a functional connectome discussed. As we can see, performance tends to be relatively consistent in the unranked case (amp, corr, and pow labels on x axis), whereas we get a slight bump when ranking (ramp, rcorr, and rpow labels on x axis). In (7b), noting that the ranked approaches tend to perform better, we look this time at just ranked approaches compared to unranked approaches. As we can see, ranking the connectome edges tends to perform generally better than using the raw edge weights. Data shown is taken from the CoRR dataset [6], and all results were obtained from the FNGS pipeline using our cloud utilities. For all data points, we show the within-subject discriminability for single-datasets.

discriminability.

4.3 Batch Effects and Multi Dataset Discriminability

It is becoming increasingly common for researchers to examine the cross-dataset robustness of collection techniques [24, 25]. The ability to collect data with similar quantitative properties is very advantageous for researchers; the researchers can achieve a far higher throughput of subjects, and obtain a diverse sampling of the population in the process. In Figure (8), the multi-dataset discriminability is shown for five reference datasets from the CoRR dataset [6]. As we can see, the discriminability of connectomes by subject id varies widely depending on scan-site with connectomes acquired following similar acquisition procedures. Using the collection site as our discriminability label, we obtain a discriminability value of $d = 0.6328$. The random chance discriminability can be computed according to:

$$C = \frac{\sum_{i=1}^K M_i^2}{N^2} \quad (1)$$

shows a theoretical random chance $d = .2461$. A value that is significantly close to $d = .248$ indicates that using the frobenius norm as our distance metric, our connectomes are relatively indistinguishable between connectome site. Our p-value deviating from the theoretical random chance by $d = .6421$ can be shown using the two-sample test permutation test [4] to be significant with probability $p = 0.001$ after 1000 iterations. This indicates that scan-site plays a heavy role in the quantitative

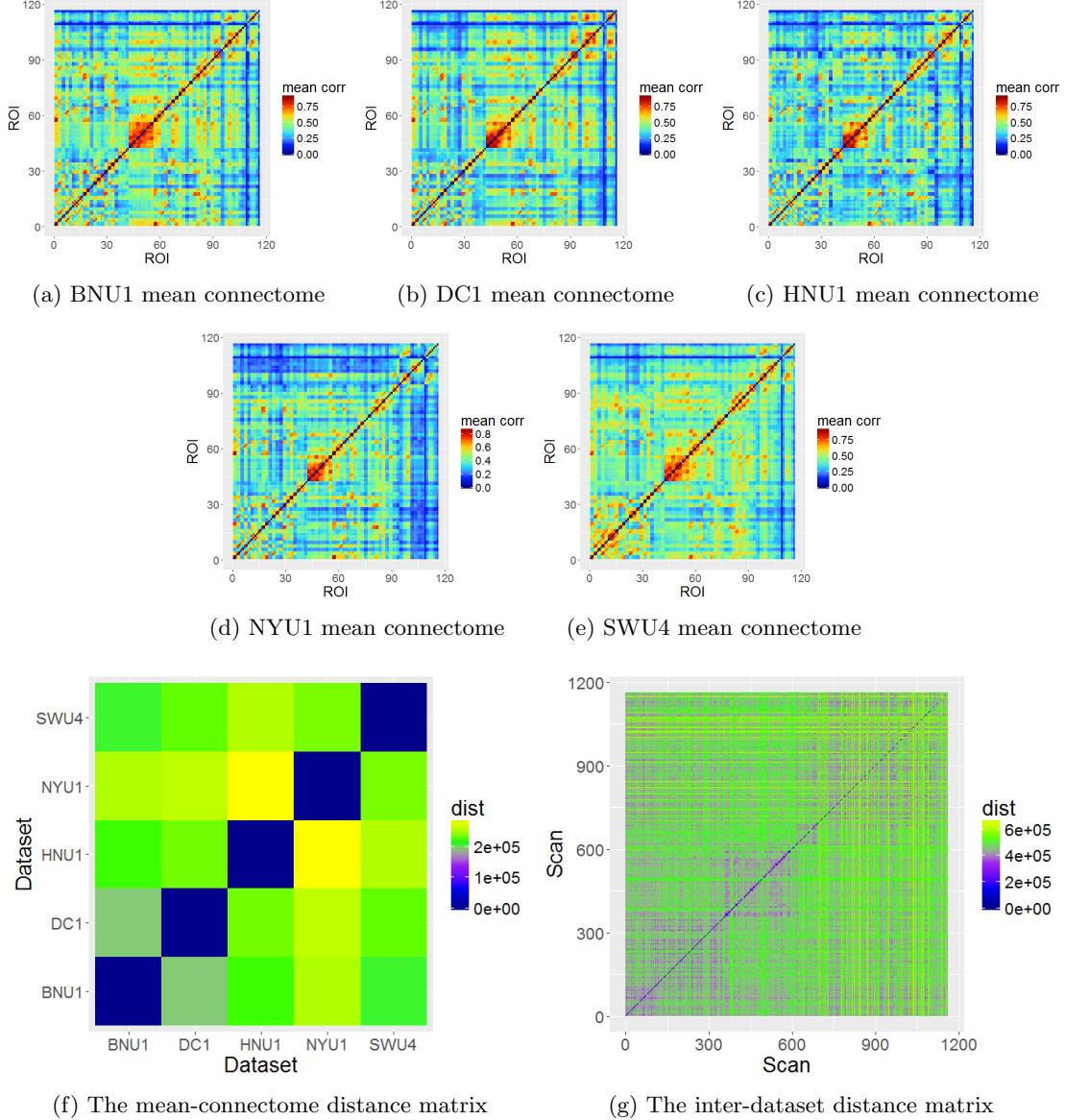


Figure 8: As we can see in (8a) - (8e), the mean connectome for each of our 5 collection sites is relatively qualitatively similar from collection site to collection site. However, in (8f) we can see that the mean connectome matrices are quite quantitatively dissimilar, as evinced by the high inter-dataset mean-connectome distances. In (8g), we compare the scans to the scans of the other datasets. Using the dataset id as our label to compute inter-dataset discriminability, we organize our scans by their dataset, so the block-clusters indicate that datasets tend to have higher distances to the other clusters than within cluster. We see a discriminability score of $d = .6241$, which significantly exceeds the score of $d = .2461$ that would have been observed if the dataset distances were random.

properties implied by the frobenius norm in a connectome, suggesting that there is site-specific signal introduced in functional connectomes. Future work should hinge on investigating the potential sources of the quantitatively dissimilarity of inter-dataset connectomes.

5 Discussion

Functional connectomics is rapidly rising as a method for providing scalable neuroscience analysis [26, 27]. Functional connectomes offer mathematical simplicity and are intuitively relatable due to graphs that can be inferred from the connectomes [28, 29]. Through the FNGS pipeline, we have developed a robust tool tuned for the discriminability for fMRI connectome acquisition. The FNGS pipeline offers state of the art computational performance tuned for maximal discriminability at each step. Our software offers maximal modularity, whereby we can easily customize or modify algorithms as desired into the pipeline as future methods become available. We package our software in docker containers, allowing researchers ease and simplicity when preparing to use our packages, and provide several layers of access points to ease usage of the pipeline. The batch analysis problem suggests that higher numbers of samples may be required to obtain more numerically stable algorithmic results. This suggests that it is vital to provide researchers with the tools to validate their algorithms on large, diverse sets of data to confirm that there is not collection-site bias in the quantitative results obtained by an analysis. We provide three layers of command line access for analyzing individual scans or datasets on local computers, and we provide tools enabling researchers to automatically scale their analyses on the cloud using Amazon Web Services. The FNGS pipeline allows researchers with no infrastructure dependencies to batch-analyze connectomes by the thousands per day, significantly lowering the substantial technical barrier to entry of modern neuroscience.

References

- [1] S Ogawa, T M Lee, A R Kay, and D W Tank. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences of the United States of America*, 87(24):9868–72, dec 1990.
- [2] Richard B. Buxton, Kâmil Uludağ, David J. Dubowitz, and Thomas T. Liu. Modeling the hemodynamic response to brain activation. *NeuroImage*, 23:S220–S233, jan 2004.
- [3] Stephen Strother, Stephen La Conte, Lars Kai Hansen, Jon Anderson, Jin Zhang, Sujit Pulapura, and David Rottenberg. Optimizing the fMRI data-processing pipeline using prediction and reproducibility performance metrics: I. A preliminary group analysis. *NeuroImage*, 23:S196–S207, 2004.
- [4] Shangsi Wang, Zhi Yang, Xi-Nian Zuo, Michael Milham, Cameron Craddock, Gregory Kiar, William Gray Roncal, Eric Bridgeford, Carey E Priebe, and Joshua T Vogelstein. Optimal Decisions

for Discovery Science via Maximizing Discriminability: Applications in Neuroimaging. *request for preprint*.

- [5] Nathan W Churchill, Anita Oder, Hervé Abdi, Fred Tam, Wayne Lee, Christopher Thomas, Jon E Ween, Simon J Graham, and Stephen C Strother. Optimizing Preprocessing and Analysis Pipelines for Single-Subject FMRI. I. Standard Temporal Motion and Physiological Noise Correction Methods. *Human Brain Mapping*, 33:609–627, 2012.
- [6] K J Gorgolewski, N Mendes, D Wilfing, E Wladimirow, C.J. Gauthier, T Bonnen, R Trampel, P L Bazin, and D S Margulies. Measuring variability of human brain activity at rest - a high resolution 7-Tesla test-retest dataset.
- [7] Mark W. Woolrich, Saad Jbabdi, Brian Patenaude, Michael Chappell, Salima Makni, Timothy Behrens, Christian Beckmann, Mark Jenkinson, and Stephen M. Smith. Bayesian analysis of neuroimaging data in FSL. *NeuroImage*, 45(1):S173–S186, 2009.
- [8] Koene R A Van Dijk, Mert R Sabuncu, and Randy L Buckner. The influence of head motion on intrinsic functional connectivity MRI. *NeuroImage*, 59(1):431–8, jan 2012.
- [9] Mark Jenkinson, Peter Bannister, Michael Brady, and Stephen Smith. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, 17(2):825–41, oct 2002.
- [10] Donald G. Mitchell and Mark Cohen. *MRI principles*. Saunders, 2004.
- [11] M Jenkinson and S Smith. A global optimisation method for robust affine registration of brain images. *Medical image analysis*, 5(2):143–56, jun 2001.
- [12] Douglas N. Greve and Bruce Fischl. Accurate and robust brain image alignment using boundary-based registration. *NeuroImage*, 48(1):63–72, oct 2009.
- [13] Günther Grabner, Andrew L. Janke, Marc M. Budge, David Smith, Jens Pruessner, and D. Louis Collins. Symmetric Atlasing and Model Based Segmentation: An Application to the Hippocampus in Older Adults. pages 58–66. Springer, Berlin, Heidelberg, 2006.
- [14] Jesper L R Andersson, Mark Jenkinson, Stephen Smith, and Jesper Andersson. Non-linear registration aka Spatial normalisation. 2007.
- [15] Anne M Smith, Bobbi K Lewis, Urs E Ruttmann, Frank Q Ye, Teresa M Sinnwell, Yihong Yang, Jeff H Duyn, and Joseph A Frank. Investigation of Low Frequency Drift in fMRI Signal. 1999.

- [16] Jody Tanabe, David Miller, Jason Tregellas, Robert Freedman, and Francois G Meyer. Comparison of Detrending Methods for Optimal fMRI Preprocessing.
- [17] Martin A Lindquist, Ji Meng Loh, Lauren Y Atlas, and Tor D Wager. Modeling the hemodynamic response function in fMRI: efficiency, bias and mis-modeling. *NeuroImage*, 45(1 Suppl):S187–98, mar 2009.
- [18] Molly G Bright, Christopher R Tench, and Kevin Murphy. Potential pitfalls when denoising resting state fMRI data using nuisance regression. 2016.
- [19] Russell A. Poldrack. Region of interest analysis for fMRI. *Social Cognitive and Affective Neuroscience*, 2(1):67–70, mar 2007.
- [20] Baxter P Rogers, Victoria L Morgan, Allen T Newton, and John C Gore. Assessing functional connectivity in the human brain by fMRI. *Magnetic resonance imaging*, 25(10):1347–57, dec 2007.
- [21] Cencheng Shen, Carey E Priebe, Mauro Maggioni, and Joshua T Vogelstein. Discovering the Geometry of Relationships Across Disparate Data Modalities. *submitted*.
- [22] Krzysztof J. Gorgolewski, Tibor Auer, Vince D. Calhoun, R. Cameron Craddock, Samir Das, Eugene P. Duff, Guillaume Flandin, Satrajit S. Ghosh, Tristan Glatard, Yaroslav O. Halchenko, Daniel A. Handwerker, Michael Hanke, David Keator, Xiangrui Li, Zachary Michael, Camille Maumet, B. Nolan Nichols, Thomas E. Nichols, John Pellman, Jean-Baptiste Poline, Ariel Rokem, Gunnar Schaefer, Vanessa Sochat, William Triplett, Jessica A. Turner, Gaël Varoquaux, and Russell A. Poldrack. The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Scientific Data*, 3:160044, jun 2016.
- [23] Amazon Inc. *Amazon Elastic Compute Cloud (Amazon EC2)*. Amazon Inc., 2008.
- [24] Jorge Jovicich, Ludovico Minati, Moira Marizzoni, Rocco Marchitelli, Roser Sala-Llonch, David Bartrés-Faz, Jennifer Arnold, Jens Benninghoff, Ute Fiedler, Luca Roccagliata, Agnese Picco, Flavio Nobili, Oliver Blin, Stephanie Bombois, Renaud Lopes, Régis Bordet, Julien Sein, Jean-Philippe Ranjeva, Mira Didic, Hélène Gros-Dagnac, Pierre Payoux, Giada Zoccatelli, Franco Alessandrini, Alberto Beltramello, Núria Bargalló, Antonio Ferretti, Massimo Caulo, Marco Aiello, Carlo Cavalieri, Andrea Soricelli, Lucilla Parnetti, Roberto Tarducci, Piero Floridi, Magda Tsolaki, Manos Constantinidis, Antonios Drevelegas, Paolo Maria Rossini, Camillo Marra, Peter Schönknecht, Tilman Hensch, Karl-Titus Hoffmann, Joost P. Kuijer, Pieter Jelle Visser, Frederik Barkhof, Giovanni B. Frisoni, and PharmaCog Consortium. Longitudinal reproducibility of default-mode network connectivity in healthy elderly participants: A multicentric resting-state fMRI study. *NeuroImage*, 124(Pt A):442–454, jan 2016.

- [25] Stefan J Teipel, Alexandra Wohlert, Coraline Metzger, Timo Grimmer, Christian Sorg, Michael Ewers, Eva Meisenzahl, Stefan Klöppel, Viola Borchardt, Michel J Grothe, Martin Walter, and Martin Dyrba. Multicenter stability of resting state fMRI in the detection of Alzheimer's disease and amnestic MCI. *NeuroImage. Clinical*, 14:183–194, 2017.
- [26] Stephen M. Smith, Diego Vidaurre, Christian F. Beckmann, Matthew F. Glasser, Mark Jenkinson, Karla L. Miller, Thomas E. Nichols, Emma C. Robinson, Gholamreza Salimi-Khorshidi, Mark W. Woolrich, Deanna M. Barch, Kamil Uğurbil, and David C. Van Essen. Functional connectomics from resting-state fMRI. *Trends in Cognitive Sciences*, 17(12):666–682, dec 2013.
- [27] Olaf Sporns. Contributions and challenges for network models in cognitive neuroscience. *Nature Neuroscience*, 17(5):652–660, mar 2014.
- [28] S. M. Smith, P. T. Fox, K. L. Miller, D. C. Glahn, P. M. Fox, C. E. Mackay, N. Filippini, K. E. Watkins, R. Toro, A. R. Laird, and C. F. Beckmann. Correspondence of the brain's functional architecture during activation and rest. *Proceedings of the National Academy of Sciences*, 106(31):13040–13045, aug 2009.
- [29] B. T. Thomas Yeo, F. M. Krienen, J. Sepulcre, M. R. Sabuncu, D. Lashkari, M. Hollinshead, J. L. Roffman, J. W. Smoller, L. Zollei, J. R. Polimeni, B. Fischl, H. Liu, and R. L. Buckner. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, 106(3):1125–1165, sep 2011.