Specific aims

There is a dire need for biomarkers to identify, diagnose, and predict treatment outcome for mental disorders (Insel and Cuthbert, 2009). Neuroimaging capabilities have grown and computational tools for processing these data have become more sophisticated, but the statistical tools for probing these data in search of biomarkers have not been very successful. An unmistakeable opportunity has arisen to try and establish biomarkers for mental disorders with *A. access to multi-site, multimodal brain data from hundreds of patients, B. the advent of new computational topology tools intended for finding high-order structure in complex data, and C. the shared interests of cross-disciplinary collaborators with complementary strengths.*

We introduce radically new methods for computing high-order associations in high-dimensional brain data. Specifically, we wish to use *topology* to summarize functional connectivity in the brain activity of *individuals* (as measured by functional magnetic resonance imaging (fMRI) and electroencephalography (EEG)). The second goal of this research is to identify in *samples* of such summaries clinically relevant features of mental disorders to learn more about the pathophysiology of mental disorders, to predict course of illness, and to personalize treatment using these features as biomarkers. Thus, there are two levels of statistical methodology: subject level and sample level. The third goal is to ground the topological biomarkers in anatomical structures and connections of the human brain.

The specific aims of the proposal are the following:

1. **Develop methods for rapidly computing several potential topological biomarkers from multivariate time series data acquired from *individual* patients or human subjects.**
2. **Develop statistical methods to analyze *samples* of such biomarkers to learn more about the pathophysiology of mental disorders, predict course of illness, and personalize treatment.**
3. **Implement the above as easily maintainable, well documented, open source, freely available software tools for general use by the neuroscientific community.**
4. **Apply the above software to clinically relevant data to establish biomarkers for mental illnesses.**
5. **Conduct diffusion tensor imaging (DTI) investigations to relate the biomarkers to network architecture.**

We will demonstrate our methods on resting state functional connectivity and other fMRI data, as well as EEG data, from multiple sources. These sources will include:

- publicly available resources, such as the 1000 Functional Connectomes Project (http://www.nitrc.org/projects/fcon_1000/)
- data acquired through two different grants for which our Co-Investigator Ramin Parsey is a Principal Investigator:
  - "Biological Predictors of Response to Antidepressants" (MH074813)
  - "Biosignature Discovery for Personalized Treatment of Depression" (1U01MH092250-01)

We will develop our methods on data that is publicly available and from the first grant to determine the range of variation of our topological biomarkers, and will validate on the multi-site data from the second grant to try and diagnose individuals with MDD and predict treatment outcome based on remitter/nonremitter data.

We believe this proposal to be significant because it will provide a means of computing high-order associations in many brain regions, something that has eluded researchers until now, and it will attempt to diagnose and predict successful treatment options for individuals with different mental illnesses, such as MDD that affects millions of Americans. This proposal is innovative because it will introduce formidable new methods in computational topology to clinical brain research and will attempt to explain the results of these new methods with respect to more conventional anatomical connections.

1 Significance

In this section, we motivate the need for biomarkers of mental illness by discussing one example in which biomarkers could serve to hasten the identification and treatment of a mental disorder, major depressive disorder (MDD). We then discuss functional connectivity data as a potential data source from which biomarkers could be drawn, expose the problems with using traditional statistical approaches to do so, and finally introduce topological methods that we have evidence to believe will overcome these problems.

1.1 The need for biomarkers of mental illness

Diagnosis of mental disorders suffers from a dearth of reliable biomarkers [17]. The importance of identifying biomarkers for mental disorders is reflected by its inclusion in the National Institute of Mental Health's Strategic Objectives (Strategy 1.3): "Currently, very few biomarkers have been identified for mental disorders due in part to their complexity and an incomplete understanding of the neurobiological basis of mental disorders..." We attribute the elusiveness of biomarkers to the fact that traditional methods used to analyze brain image data do not adequately reflect their complexity.

The most significant contributions we intend to make with the proposed research are to develop alternative methods to overcome this complexity (**Specific Aim 1**) and identify biomarkers of mental disorders (**Specific Aim 2**), determine the range of variation of these biomarkers, and to use them to try and diagnose individuals and predict treatment outcome (**Specific Aim 4**).

1.2 The impact of Major Depressive Disorder (MDD)

The tools we will develop to identify biomarkers are intended to be applied to any mental illness, such as Alzheimer's disease, bipolar disorder, schizophrenia – indeed to analyze differences in brain activity between any two populations or study the association of brain activity with measures of, for example, severity. However, we have targeted attention deficit hyperactivity disorder (ADHD) for our initial experiments (3.4) and major depressive disorder (MDD) to guide further development and conduct validation of our methods. We will process MDD data from two different grants for which our Co-Investigator Ramin Parsey, a leading researcher of depression, is a Principal Investigator:

> "Biological Predictors of Response to Antidepressants" (MH074813)
>
> "Biosignature Discovery for Personalized Treatment of Depression" (1U01MH092250-01)

The latter is a large, multi-site U01 acquiring multimodal brain imaging data from 400 individuals, specifically designed to make such data available to establish biomarkers for MDD (1.2). Arno Klein (P.I.) is also a Co-Investigator on the U01.

In addition to having access to expertise, resources, and data related to MDD, we wish to focus on MDD because of its overwhelming impact on the health of Americans, as noted by the NIMH (http://www.nimh.nih.gov/health/publications/the-numbers-count-mental-disorders-in-america/index.shtml):

> Major Depressive Disorder is the leading cause of disability in the U.S. for ages 15-44
> (http://www.who.int/healthinfo/global_burden_disease/GBD_report_2004update_AnnexA.pdf)
>
> Major Depressive Disorder affects approximately 14.8 million American adults,
> or about 6.7 percent of the U.S. population age 18 and older in a given year
> (http://www.census.gov/popest/national/asrh/) [21].

The largest clinical trial of MDD ever conducted, STAR*D, indicates that two-thirds of patients treated with a first-step antidepressant do not achieve remission of symptoms [31]. Furthermore, successive treatment steps lead to diminishing remission rates [26] and a large number of patients discontinue treatment prematurely due to side effects [31]. The trial and error method currently used in clinical practice often leads to repeated failures before an effective treatment is identified. Given this relative ineffectiveness of treatments for depression and resulting practice of trial and error multiple treatment steps, there is an urgent need to identify factors that can be used to personalize treatment (i.e., markers that maximize effectiveness and minimize the risk for toxicity). The development of biomarker predictors of antidepressant

response languished after multiple candidates, most notably the dexamethasone suppression test (DST), proved to have inadequate prognostic clinical utility [14]. However, the emergence of new technologies in neuroimaging has sparked new interest in developing biomarkers that might predict antidepressant response. Because of limited understanding of the pathophysiology of MDD and the limited range of the mechanism of action of available antidepressants (monoaminergic uptake inhibition or receptor modulators), we are currently unable to match treatments to patients.

1.3 Functional connectivity

Advances in neuroimaging brain activity have opened up tremendous stores of rich data from which biomarkers may be drawn. An important aspect of brain activity is the interaction among brain regions. (A "region" may be an anatomical region (3.5.1), a functionally defined region (3.5), or even just a single voxel (3.5.3).) This interaction is reflected in "functional connectivity" in neuroimaging time series, *"the observed temporal correlation between two electro/neurophysiological measurements from different parts of the brain."* [11]

For concreteness we discuss functional connectivity in the context of blood oxygenated level-dependent (BOLD) functional magnetic resonance imaging (fMRI) with the understanding that our discussion applies to general multivariate time series. Moreover, for now we consider the problem of describing or summarizing the functional connectivity as manifested in the fMRI *of an individual person*. This is the level of statistics with which **Specific Aim 1** is concerned. In our initial experiment (3.4) and in future experiments we will consider using these descriptions to compare clinical groups to help get insight into the pathophysiology of mental illness (**Specific Aim 2**). An advantage of summarizing the fMRI of an individual patient is that the summary can be used for prediction of disease course and treatment outcome for that patient and to personalize his/her treatment accordingly.

1.3.1 Taxonomy of functional connectivity methods

Methods for functional connectivity analysis can be arrayed along two dimensions: number of regions and "order." Order refers to the maximum size of groups of regions whose functional connectivity is to be assessed. For example, one may take BOLD fMRI data from 50 regions and look for association between every pair in the 50 (so order $= 2$). There are three levels of numbers of regions: "few regions," meaning 10 or fewer regions; "many," at most a few hundred; and "very many," possibly tens of thousands. (We use "many" to mean moderately or very many.) We also differentiate two levels of *order*: "low," 2 regions at a time; and "high," more than 2.

It makes sense to look for high-order functional connectivity. It is difficult to believe that the brain does its work by coordinating the activity of at most two regions at a time (low order). True, for jointly normal data all the connectivity information is already contained in the 2nd-order associations (reflected in correlations or coefficients in linear regressions). Real data are not jointly normal, however, so one might miss important connections by just restricting attention to 2nd- (i.e., low-) order associations. It makes sense to look for high-order associations directly instead of trusting possible joint normality.

Looking at a few regions makes sense for a focused, hypothesis-driven, model-based analysis. But there is a place for looking at many or very many regions, e.g., in exploratory analyses looking for functional connectivity among regions widely distributed throughout the brain (or even the whole brain). Functional connectivity requires regions to connect. The more regions one considers, the more possibilities there are for connectivity.

Another reason for looking at many regions is to search for functional connectivity at high resolution. Standard anatomical regions often contain subregions which may exhibit independent functional connectivity. So functional connectivity may be more profitably examined at the anatomical subregion level. But higher level of resolution means more regions. We describe at some length and demonstrate on preliminary data a "concurrence-based" method 3.2 that is high-order, moderately many-region. We also propose an alternative method that is high-order, very many-region 3.6.

1.4 Statistical considerations

In the "concurrence-based" approach outlined in this proposal, our focus is on contemporaneous associations, i.e., activity in multiple regions at the *same* time point. Moreover, we are only interested in associations among regional activity for a given task. Thus, the data we use in this proposal are acquired during a single task, for example "rest."

We consider an even-handed or "agnostic" analysis that puts all regions on an equal footing. (This is in contrast to a "seed-based" analysis of the interaction between a specific small group of regions and the remaining regions.) Finally, our emphasis is on fMRI data on moderately many regions; the example presented below (3.4) is based on data with more than 100 regions.

We seek descriptions of functional connectivity that are independent of time, that characterize connectivity over the entire fMRI run. In order for this to be useful, we require that the BOLD signal from any single task be "stationary," i.e., that the statistical character of the multivariate time series does not change over time. This can be assessed to some extent from the data, and some departures from stationarity can be corrected. It is our practice to remove in each region linear time trends in the data as well as remove time trends in the spread in the data (heteroscedasticity).

Describing functional connectivity means tracking the joint activity in multiple regions across time. This might be done measuring the association among groups of regions using the "joint cumulant" [2] of their activity. (The second-order joint cumulant, i.e., the joint cumulant of two variables, is the "covariance," on which Pearson correlation is based.) Another possibility is regression or path analysis, including interaction terms in multiple regression in order to capture higher-order associations.

Functional connectivity will obviously depend in general on the regions involved. In an even-handed approach, the collective activity of *any* group of regions might be of interest. That creates a difficult problem because when there are large numbers of regions in the analysis the numbers of groups of regions can be *extremely* large. One can easily end up with more cumulants or regression coefficients than are are numbers in the original BOLD data!

How can one cope with such information overload? One can simplify correlation or covariance matrices by extracting principal components. There is, however, no such thing as principal components analysis for 3- and higher-order cumulants. One might employ "parallel factor analysis" and "Tucker3" [3] but even they produce fairly large, difficult-to-interpret data summaries. Unlike principal component analysis, which looks for *uncorrelated* sources of variation, independent component analysis (ICA) [16], finds approximately *independent* sources of variation. That makes ICA a high-order analysis method, but only *implicitly*. While ICA can be used with very many regions, "A possible disadvantage of ICA methods might include that the independent components are often perceived as more difficult to understand than traditional seed-dependent fcMaps, as they contain a more complex representation of the data, which could complicate the translation of between-group results to clinical relevance." [34] We propose methods that are radically different. Our approach produces summaries of functional connectivity that are of manageable size and structured in an interpretable way.

1.5 Topological methods

The difficulty with functional connectivity data is that it can be computationally very demanding to analyze, and it is extremely difficult to conceptualize and visualize the connectivity in the data. Topology, a major area of mathematics, is *the* field of mathematics to describe the connectedness among structures. (Graph theoretic methods [34] are inherently 2nd-order, whereas topology can describe connectedness at any order.)

Topology has principally been used in neuroscience for characterizing anatomy, to help guide or constrain registration and cortical surface reconstruction [23, 10, 25, 18], as well as segmentation of anatomical structures [23, 1]. Topology has also been used to characterize some basic properties of the functional organization of the visual cortex (e.g., orientation and spatial frequency maps) [30, 29]. A new method in computational algebraic topology is persistent homology, or *"persistence"* (3.3.2). Persistence has been applied to 2-D cortical thickness data and demonstrates its feasibility for discriminating populations (autistic vs. controls) [6].

It has only been very recently that sophisticated tools in computational topology have been developed to analyze high-order associations in high-dimensional functional brain data. So recent, in fact, that we are aware of only two relevant studies. Both analyzed patterns of simultaneously firing neurons, one of simulated rat hippocampal place cells [7] and the other of multielectrode data in anesthetized macaque monkey primary visual cortices [27]. The first study demonstrated a relationship between the topology of an internal representation (patterns of neural activity) and the topology of external stimuli (a maze). We have modeled our concurrence-based approach (3.2) after this work, substituting fMRI for (simulated) multielectrode data. The second study analyzed the persistence of spike train data corresponding to spontaneous activity and image stimulation. We have modeled our second, persistence approach, after

this work (3.6). Both of our approaches are the focus of **Specific Aims 1**, **2**, **3**, to develop and apply new computational topology tools and statistical methods for use by the neuroscience community.

2 Innovation

2.1 Pursue alternatives to traditional statistical approaches

Traditional statistical approaches to functional connectivity work fine for small numbers of regions. But as we have seen in section 1.4, if one looks for high-order connectivity in many regions, one's options seem limited to ICA, which is only *implicitly* high-order.

- *The method we propose is the first explicit high-order, many-region analysis method.*

In this proposal we describe a radically different way to study high-order (i.e., 3 or higher) functional connectivity in fMRI in many regions (1.3.1). Our method sidesteps the difficulty described in (1.4). Instead of producing output consisting of tens or hundreds of thousands or millions or even more numbers, the proposed method generates only dozens of numbers. Moreover, the output of the proposed method is structured in an interpretable way. This makes reduction of the output natural.

- *Our topological approach is radically different from methods currently being employed and, indeed, radically different from standard statistical methods in general.*

We have implemented this procedure and describe its application in the Approaches section (3). We propose to further develop and streamline the proposed method and apply it to a wide variety of brain imaging data sets. Our preliminary data example (3.4) indicates that the proposed method can capture clinically important functional connectivity structure in fMRI data. Moreover, we are able to use computational methods to summarize fMRI data from individual patients, which is a vital component to personalized medicine. In addition to developing alternatives to traditional statistical approaches for analyzing fMRI data, Steven Ellis (P.I.) will lead development of new statistical methods to interpret the data we get as a result of processing our data in this novel manner (**Specific Aim 2**).

2.2 Develop new tools in computational topology

We will not only be applying the newest computational topology tools in our research; we will be developing these tools further with our consultant Dmitriy Morozov, one of the world's experts in the field and the developer of the Dionysus persistent homology software (http://www.mrzv.org/software/dionysus) (**Specific Aims 1** and **3**). will also work with Konstantin Mischaikow of Rutgers University, one of the chief developers of CHomP, the Computational Homology Project (http://chomp.rutgers.edu/). Each of the members of this collaboration has a strong interest in cross-disciplinary research, and we anticipate that this will help to foster exciting methodological innovations.

In addition to the concurrence-based (3.2) strategy we employed in our initial experiments, we have already begun to outline below promising alternative strategies to define regions (3.5) and compute persistence (3.6).

2.3 Make the topology interpretable

The proposed research promises numerous methodological advancements in computational topology and their application to clinical brain research (**Specific Aim 4**). However, identifying topological biomarkers alone will not reveal the relationship between the topological structure of functional connectivity data and the underlying network structure or functions of the brain. We propose to make the topological biomarkers more interpretable by relating them in innovative ways to other imaging data that is better undestood, for example by grounding them in anatomical structures and connections of the human brain by DTI methods (3.7) (**Specific Aim 5**).

3 Approach

3.1 Data and preprocessing

3.1.1 Data

For the duration of the grant period, we intend to make use of several rich sources of data. We will be demonstrating our methods on resting state and other fMRI data as well as EEG data from sources that will include:

- publicly available resources, such as the 1000 Functional Connectomes Project (http://www.nitrc.org/projects/fcon_1000/)
- data acquired through two different grants for which our Co-Investigator Ramin Parsey is a Principal Investigator:
  - "Biological Predictors of Response to Antidepressants" (MH074813)
  - "Biosignature Discovery for Personalized Treatment of Depression" (1U01MH092250-01)

We will develop our methods on data that is publicly available and from the first grant to determine the range of variation of our topological biomarkers, and will validate on the multi-site data from the second (U01) grant to try and diagnose individuals with MDD and predict treatment outcome based on remitter/nonremitter data. Arno Klein is also a Co-Investigator on the U01 study. Over the course of this proposal's R01 period, 400 subjects will participate in the U01. We will make use of all of the data acquired by the U01, which includes structural, functional and diffusion tensor 3.0T MRI data as well as EEG (and behavioral) data:

- Structural 3D axial MPRAGE images (TE: 3.29 ms, TR: 2200 ms, Flip angle 9°, Field of view: 256x192 mm, Slice thickness: 1 mm, Matrix: 256x256, 192 continuous slices, 7:02 min)
- 4 fMRI tasks (emotional conflict, reward processing, PASL, and resting-state connectivity acquisition): 39 axial slices (3.1mm thick, TR/TE=2000/28msec, FOV=205x205cm, matrix=64x64; Flip angle=90°)
- DTI using echo planar imaging (voxel size: 2x2x2mm, 61 and 25 non-colinear directions; b-value=1000s/mm2)
- EEG: resting state (four 2-minute periods), loudness dependency (1KHz, 40ms duration, 10ms rise and decay time), and binaural tones (60–100 dB SPL) with interstimulus intervals ranging from 1600–2100ms (500 trials)

3.1.2 Preprocessing

The following steps serve as an example of how we have prepared properly formatted data from fMRI data:

1. Run motion correction with FMRIB's MCFLIRT [19] and artifact detection with RapidART (within Nipype: http://nipy/nipype), both called from within the Nipype software pipeline tool. Motion artifacts exceed 1mm at the center of at least one of the faces of the brains' bounding boxes for at least one of the time points. Intensity outliers (>3 SD from the mean) are removed from the subjects.
2. Label the gross anatomy with, for example, FreeSurfer (http://surfer.nmr.mgh.harvard.edu/fswiki) [8]. The labeled regions can serve as the regions we refer to throughout the proposal. (This step will be replaced with the new Mindboggle software: http://www.mindboggle.info, P.I.: Arno Klein.)
3. Construct a table for each subject, with each row representing a region and each column a time point. Each table cell contains the average BOLD activity for the corresponding region and time point.
4. Remove linear trends in the rows (for each region over time) by, for example, a robust regression to measure the trend, and then a robust regression to even out the spread in the data across time.

3.1.3 Software

In addition to the above preprocessing software (3.1.2), we will be using multiple computational homology and DTI-related software packages. We will do most of our programming in Python (http://www.python.org) as we develop in Dionysus (created by our consultant Dmitriy Morozov) and use Python libraries for processing image data.

For computational homology, we will continue to use:
Sage (http://www.sagemath.org/),
CHomP (http://chomp.rutgers.edu/), and
Dionysus (http://www.mrzv.org/software/dionysus).

For DTI work, we will use primarily dipy (http://nipy.sourceforge.net/dipy/),
FSL (http://www.fmrib.ox.ac.uk/fsl/),
Camino (http://web4.cs.ucl.ac.uk/research/medic/camino/pmwiki/pmwiki.php), and
ANTS (http://www.picsl.upenn.edu/ANTS/) packages.

3.2 A concurrence-based approach to functional connectivity

3.2.1 Concurrences and their topology

We outline an approach to the analysis of functional connectivity among moderately many regions based on "concurrences" (defined presently). In the *concurrence-based approach*, the first stage in processing the tabular data from section 3.1.2 is to dichotomize the BOLD signal.

At each time point, on the basis of BOLD values, each region is classified as either "active" or "inactive." We do this by thresholding the BOLD signal at a threshold that depends on the region. (In (3.6) we describe an alternative method that does not involve dichotomizing.)

A danger with dichotomizing is that tiny fluctuations in BOLD activity that should be ignored instead get magnified. Suppose a region has a steady level of BOLD activity throughout the run. This might be neurobiologically interesting, however, functional connectivity is concerned with *variation*. So regions whose BOLD activity hardly varies should be ignored. However, tiny variations in BOLD activity might carry it back and forth across a region's threshold and thus can result in large variations in dichotomized BOLD in that region. Such spurious variations in dichotomized BOLD activity can seriously degrade the connectivity information in the dichotomized BOLD data. To prevent this, before dichotomizing, we drop the least variable regions (based on a robust version of coefficient of variation). (For the preliminary data analysis 3.4.1 we dropped the 20% least variable regions in each subject's data.)

As mentioned in (1.3), we limit ourselves to studying contemporaneous functional connectivity, i.e., joint *simultaneous* behavior of brain regions. This means that once BOLD has been dichotomized, analysis of connectivity reduces to studying "concurrences." A "concurrence" is merely the simultaneous achievement of the active state in several regions (or, abusing nomenclature, even in just one region), and is described by merely listing the regions that are active at that time point. The following is important: A concurrence is a list of regions active at the same time point *but the time at which they are all active is ignored*. Thus, there is no temporal information in the collection of concurrences. Here again there is potential for losing important information, but we will review time-shift-invariant alternatives to this approach as well (3.5.2).

The proposed method is adapted from an algorithm described in [7], which presents a simulation performed to study a problem in theoretical neuroscience. The outline of their algorithm (call it the "recipe") is as follows.

1. Dichotomize the continuous measure of regional brain activity. This gives rise to a collection of concurrences.
2. Interpret the collection of concurrences as a geometric shape.
3. Describe the shape.

Very briefly, the idea in step 2 is as follows. One represents each concurrence by a "simplex" that is "spanned" by points ("vertices"), each of which corresponds to a brain region in the concurrence. The points corresponding to the regions can be any collection of points in "general position" in a (necessarily) high-dimensional space. The location of the point corresponding to a brain region *has nothing to do with the region's anatomical location in the brain*. Our

interest is functional, not structural, i.e., not anatomical. A concurrence consisting of just one region is interpreted as a point. A concurrence of two regions becomes a line segment joining the two vertices corresponding to its constituent regions. Three regions becomes a triangle. Four becomes a "tetrahedron," and so on. The collection of all the simplices is a "simplicial complex" (or "complex" for short). A complex forms a shape.

3.2.2 Filtering

For the purposes of describing functional connectivity, the recipe is not quite adequate. It only registers the presence or absence of a concurrence. (If a simplex is already present in the complex, there is no room to add another copy.) But to get a full picture of the regional association of dichotomized BOLD activity, not just the presence or absence of a concurrence, but the *frequency* of its occurrence is needed. To incorporate frequency information we use the concurrence data to create not just one simplicial complex, but a "filtered" complex that consists of a series of complexes nested within each other like Russian dolls.

One creates a filtered complex as follows. First, create a complex using all the concurrences that occur at least once. This is obviously all the concurrences and is the same as step 2 in the recipe. Label that complex as "frequency level 1." Next, exclude all the concurrences that occur just once and form a complex from those that occur at least twice ("frequency level 2"). That complex will be smaller than that at frequency level 1. The complex at frequency level 3 is still smaller, etc.

3.3 "Describing the shape"

Step 3 in the recipe ("describe the shape") means describing the topology of the shape, i.e., describe its gross qualitative features. In this section we describe some methods for doing this.

3.3.1 Betti numbers

The basic strategy is to count the gaps, tunnels, holes, voids, and cavities in the complex (or rather in the shape formed by the complex). The gaps between the Hawaiian islands, the hole in a donut, and the void inside a beach ball represent qualitatively different forms of absence. Therefore, the gaps, etc., are not well represented by just one number. Instead, a series of non-negative whole numbers, the "Betti numbers", $\beta_0, \beta_1, \beta_2, \ldots$ are used for this purpose in topology ([9, 20, 24]).

The subscript d in β_d is, for technical reasons, called the "dimension" of the Betti number and β_d provides a (partial) summary of what is called the "d-dimensional homology" of the complex. (Do not confuse this usage of "homology" with its usage in biology.) If one finds $\beta_d > 0$, then one has found structure in the functional connectivity involving at least $d + 2$ regions. ($\beta_d = 0$ does not signal absence of structure, only ambiguity.) Thus, $\beta_1 > 0$ indicates structure involving at least 3 regions. Structure involving 3 regions reaches the threshold of what we call "high-order" functional connectivity (1.3.1). We find that in real data β_1 is often not just positive but large (1a). In other words, *"interesting structure in high-order functional connectivity is common in fMRI data.*

A description of the topology of the filtered concurrence complex is provided by the Betti numbers of the complex at each frequency level. Thus, the Betti numbers have a two-way structure: frequency and dimension. This structure of the collection of Betti numbers is easy to interpret. As explained in the last paragraph, *dimension* is related to number of regions (i.e., to the "order" of functional connectivity).

Strength of association in the joint activity of regions can be measured by the *frequency* of occurrence of concurrences. Consider, for simplicity, three regions a, b, and c. Suppose the concurrences (a, b), (a, c), and (b, c) occur at least 3 times, but the larger concurrence (a, b, c) occurs fewer than 3 times. This would give rise to a "hole in the complex" at frequency level 3 and be counted by β_1 at frequency level 3. Thus, Betti numbers detect contrasts among different levels of positive association ("excitation") in the same groups of regions. The pattern just described suggests a looseness or incompleteness in the connectivity among activity in a, b, and c. (We return to this example in (3.3.2).)

In order to relate these numbers to brain anatomy, we want to identify the collection ("cycle") of regions surrounding a "hole." Computational homology software can output representative cycles associated with holes. Unfortunately, there may be very many such cycles. (Imagine a tunnel through the complex. There may be many cycles that wrap around

that tunnel.) We will work with our collaborators Morozov and Mischaikow to develop ways for the software to output multiple cycles to increase the chances of generating cycles related to anatomy of scientific interest.

In (3.4.3) we exhibit Betti numbers from some real data. However, at present computing Betti numbers is a difficult, hands-on task. The problem is compounded by the fact that the Betti numbers have to be computed for every subject in a data set, and indeed for every complex at each frequency level. We have written some code that helps, but at present we are unable to automate every computation. We believe we can soon streamline this process.

The complex at frequency level 1 is the largest in a filtered complex and at present we find computing the Betti numbers at frequency level 1 impractical. Fortunately, frequency level 1 is also the least interesting frequency level since it includes events that are not even generalizable within the same subject so are unlikely to be generalizable across subjects. Therefore, not much is lost in ignoring frequency level 1.

An alternative to computing the Betti numbers is to compute the "Euler characteristic" of the complex:

$$\chi = \beta_0 - \beta_1 + \beta_2 - \beta_3 + \cdots . \tag{1}$$

Since χ involves Betti numbers one would think it would be hard to compute, however, surprisingly, χ can usually be computed much more easily than the individual Betti numbers can. We exploit this in (3.3.2).

3.3.2 Persistence

Consider again the example given in (3.3.1). In that example, the difference between the frequencies of (a, b), (a, c), and (b, c) on one hand and (a, b, c) on the other may be as small as 1. So the "hole" associated with the frequency difference may not be so important. But suppose (a, b), (a, c), and (b, c) each occur at least 4 times, while (a, b, c) occurs just once. Then there is a difference of at least 3 in frequency. That would be a striking feature in the data. Such a contrast suggests something more than a mere looseness in the connectivity among those regions. The pattern is consistent with a kind of "inhibition" in which, e.g., activity in a and b "inhibits" activity in c. To see this, suppose both a and b are active at a given time point. Then if c were also active at that time point there would be an instance of the concurrence (a, b, c). But (a, b, c) is far less common than (a, b). In fact, activity in c only occurs at one time point at which a and b are active. Similarly, activity of b and c "inhibits" a, etc. The technique of "persistent homology" ([9]) not only detects holes but also records how "deep" they are in terms of frequency contrast. (Call that the "frequency depth" of the hole.) The holes with the greatest frequency depth are probably the most important.

Another way to think about persistence is as follows. The procedure described in (3.3) examines each of the complexes constituting the filtered complex individually. Persistence describes how they fit together. Persistence analyzes homology at different levels of a stratifying variable, such as frequency or resolution, detecting homology classes (holes) that persist as the stratifying variable is changed.

3.4 Initial experiment

In this section we exemplify the concurrence-based approach described in (3.2) to (3.3.2) using real data.

3.4.1 Data for initial experiment

The data used in our initial experiment are resting state fMRI scans available for download as part of the 1000 Functional Connectomes Project (http://www.nitrc.org/projects/fcon_1000/). Some of these data were used earlier in studies at New York University, where a full description regarding their diagnosis and image acquisition may be obtained (3x3x3 mm voxels; 39 slices; 192 time points; echo planar imaging on a Siemens 3.0T Allegra; TR=2s; echo=25ms; flip=90; matrix=64x64; fov=192mm) [33, 5]. We selected this data set because it contains many subjects performing the same resting task (fixation on the word "Relax" for 6.5 minutes), and it was the only one in the project with controls and a clinical population (ADHD):

22 of the 25 with ADHD: 19M/4F, 20-50 yrs.

39 of the 84 controls: 43M/41F, 7-49 yrs.

3.4.2 Preprocessing results

We found very few motion artifacts (corrected in any case by MCFLIRT): 14 out of 109 subjects, and only 5 subjects with more than two time points with motion artifacts. There were no more than 6 time points in any subject with intensity outliers. FreeSurfer labeled about 113 anatomical regions in each case. While detrending the data we didn't notice any heteroscedasticity.

3.4.3 Experimental results

Once we have computed summaries of homology for a number of subjects we can use those summaries to try to learn something about the populations from which those subjects come. This is the level of statistics intended by **Specific Aim 2**. As an example, we computed the Euler characteristics of the fMRI data for each subject (and separately for each frequency level). (The Euler characteristics were computed using software written by the P.I.s.) To further summarize the Euler characteristics (one Euler characteristic at each frequency level for each subject) we can sum them across several frequency levels. Figure 1a, left panel, shows the Euler characteristics summed across frequency levels 2 through 5. (This is a data reduction that could plausibly be chosen before seeing the data.) Based on this sum we can test for a difference between the ADHD patients and the controls. The p-values are: $t-$test: 0.032, and Wilcoxon rank sum test: 0.045. It is important to note that the statistically significant result was obtained using only the coarsest descriptor of homology (*viz*, Euler characteristic).

Next, we did a similar analysis, this time using Betti numbers. Figure 1b shows the mean Betti numbers of ADHD subjects vs. those of controls broken down by frequency level and dimension. A pair of equal Betti numbers would correspond to a point on the line. The most striking feature of the plot is that the mean Betti numbers of the ADHD subjects in dimension 1 are noticeably higher than the corresponding Betti number means of the controls. Paralleling the analysis of the Euler numbers, we took as our data summary the Betti numbers in dimension 1 in each group over frequency levels 2 through 5 (right panel in Figure 1a). We find that the groups are significantly different (Welch Two Sample t-test: $p = 0.018$; Wilcoxon rank sum test: $p = 0.028$). It is remarkable that we find a significant difference since we made no effort to focus on the regions of the brain believed to be implicated in ADHD. Moreover, as mentioned in section 3.3.1, the fact that group differences are found in β_1 means that the differences lie in 3rd- (and hence high-) order(1.3.1) functional connectivity.

Our interpretation of this finding is as follows. The resting state brain activity is less cohesively connected in ADHD subjects than in controls. This result is in keeping with what the researchers who provided the data concluded based on their own functional connectivity analysis of the same data [33, 5]. They measured functional connectivity among only three spherical regions of interest at what was assumed to be the dorsal anterior cingulate cortex and rostral inferior and middle frontal gyrii. Our method, by contrast, is able to process many regions at high order (2.1).

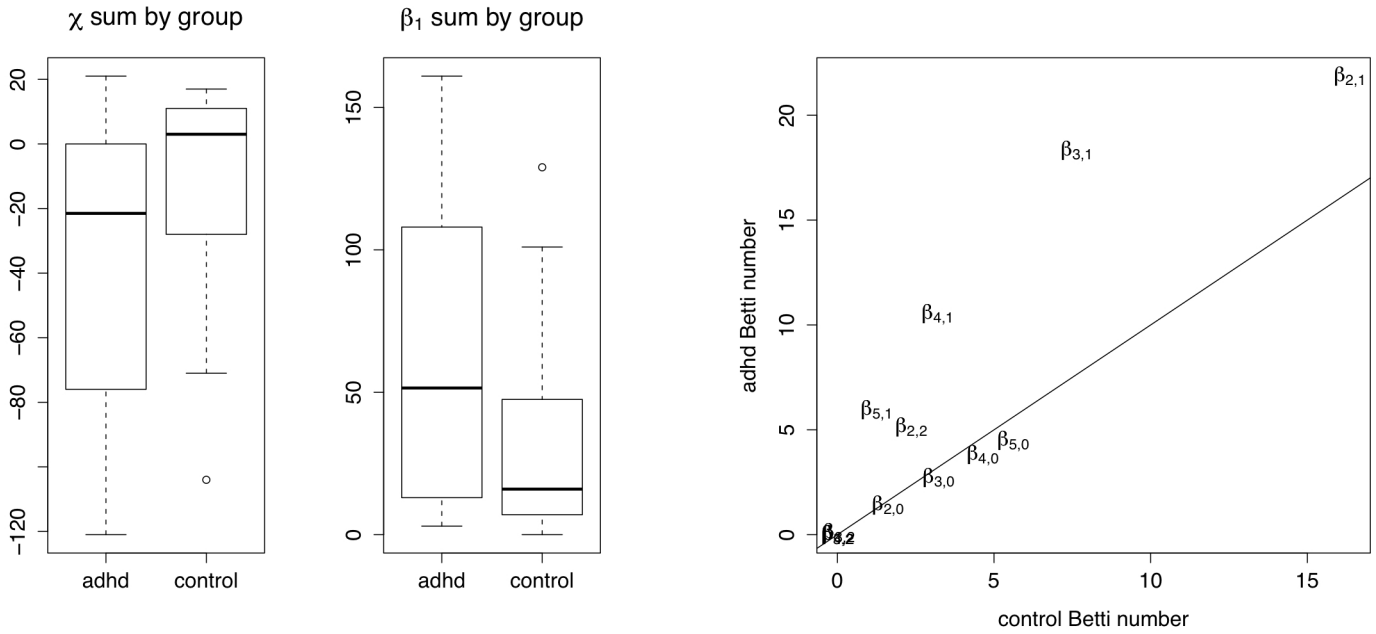
As mentioned in (3.3), it is possible to compute representative "cycles" for the homology described by Betti numbers. Here is an example of *one* cycle (computed using CHomP, http://chomp.rutgers.edu/software/) from another, nonclinical (Function BIRN Phase II Traveling Subject) data set we worked with, using different atlas labels [32]:

$$\begin{array}{cc} Cerebellum-Crus2-R & Cerebellum-3-L \\ Cerebellum-Crus2-L & Cerebellum-3-L \\ Frontal-Mid-R & Fusiform-R \\ Cerebellum-Crus2-R & Cerebellum-6-R \\ Frontal-Mid-R & Cerebellum-Crus2-L \\ Fusiform-R & Cerebellum-6-R \end{array}$$

Each row represents a concurrence of length 2 (that might lie within a larger concurrence). Thus, each row is represented by a line segment in the simplicial complex for the data.

We have only very preliminary data concerning persistence. But for illustration here are some results for one control. In dimension 0 there are "holes" of "frequency depths" 1, 2, 3, 4, 6, and 8. In dimension 1 there are "holes" of "frequency depths" 1, 3, and 4. In dimension 2 there is a "hole" of "frequency depth" 1. And in dimension 3 the "frequency depths" are 1 and 2.

Given that we applied our tools to the only publicly available resting state data of clinical relevance and obtained significant results, we are hopeful that our methods can help to identify, diagnose, and predict treatment outcome for



(a) Left panel: Boxplots of sums of Euler characteristics χ across frequency levels 2 through 5 vs. group. Right panel: Same, but with Betti numbers in dimension 1 (β_1) in place of Euler characteristics. In these data χ is mainly driven by β_1 (Equation 1), so the plots are almost inverses of each other.

(b) Mean Betti numbers of ADHD subjects vs. that of controls by frequency level (1st subscript; ranges from 2 to 5) and dimension (2nd subscript; ranges from 0 to 2). So, e.g., the coordinates of the point labeled "$\beta_{3,1}$" are the β_1 values for controls and ADHD cases (in that order) computed at frequency level 3. Note that except at frequency level 2, all mean Betti numbers in dimension 2 are close to 0. The line is the "adhd β = control β" line.

Figure 1: Initial experimental results of concurrence-based analysis of ADHD 1000 Functional Connectome data

mental disorders.

3.5 Alternative region definitions

As mentioned in 1.3, there are three different types of regions we will experiment with in concurrence-based and non-concurrence-based analysis: anatomical, functional, and individual samples.

3.5.1 Anatomical regions

In our initial experiments (3.4), we use FreeSurfer-defined anatomical labels. However, Arno Klein (P.I.) is involved in the development of a new brain labeling protocol where the label definitions draw from the best of prior protocols and feedback from the neuroscience community, and the application of the protocol to manually label 1,000 brain MRI volumes (http://www.braincolor.org). He is also actively developing a new version of the Mindboggle automated brain labeling software (http://www.mindboggle.info) as part of a 3-year, NIMH-funded grant (MH084029-01) titled "Mindboggling Shape Analysis and Identification." We are of course interested in applying the new protocol and Mindboggle software to supply anatomical labels for this project.

3.5.2 Functional regions

In this proposal we are concerned primarily with functional connectivity, and should not assume that anatomical boundaries are the best way to define our regions. Functional connectivity between regions will be strengthened or weakened depending on how cohesive the data are within a given region, and there is little reason to believe that anatomical regions

act as a homogeneous unit of similar associations with other regions. An alternative way to partition the image data is to cluster voxels according to task-derived functional brain mapping or by the strength of associations among them in resting state data, which is the convention for functional connectivity research. We will use coherence as our measure of the strength of association, because it allows for time lags between similar time series which is more faithful to the definition of functional connectivity (1.3). Voxels with similar time series are clustered together as regions, removing trivial concurrences from concurrence-based analysis.

3.5.3 Individual samples

The most agnostic approach to functional connectivity analysis is to simply determine the strength of association between each sample (voxel in an image volume) and every other sample. Though not unreasonable when searching for 2nd-order correlations, this approach would be computationally intractable when searching for high-order associations, such as with concurrence-based methods. However, it would be reasonable to focus attention on a subset of voxels in an fMRI volume, and this is very useful when the researcher has an idea where to look. For example, if one wanted to test the hypothesis that distractors in a visual attention task will fragment otherwise consistent and cohesive patterns of functional activity, then sampling from within "attentional centers" is analogous to how one would choose a waypoint or seed region in DTI tractography.

3.6 Alternative persistence strategies

Concurrence-based analysis is not the only way one could apply topological methods to fMRI data. We will also pursue three alternative persistence strategies that have their own advantages.

3.6.1 Rips complexes

An alternative to filtering binary data would be to filter non-binary data, which would retain more of the original brain activity data. One way would be to treat the unthresholded BOLD activity data as points in a high-dimensional space, with each point representing a region and the number of dimensions equaling the number of time points in the data. Strength of association would be represented by distance in this space. With this formulation, we will treat the data as a Rips complex and conduct persistence by successively filtering the data by degree of proximity between points[27, 9]. In other words, if one expands each point into a sphere of some radius, then as one increases the radii, more and more points become connected, altering the topological structure of the complex. We would then, as before, compare the changes in topological structures across brains to distinguish between two or more populations.

3.6.2 Zigzag persistence

Another alternative to concurrence-based analysis is to localize the topological signal in (physical) time by employing the recently introduced zigzag persistence[4]. The construction of a simplicial complex from signal co-occurrences, described above, does not take the temporal proximity of the signals into account. Zigzag persistence allows one to track homology of simplicial complexes as they grow and shrink with time. In practice, this means that we can slide a fixed-width time window over the data, and track the high-order associations that we find in any given slice across the entire time series.

3.6.3 Filtration in 3-D

A third alternative to concurrence-based persistence would take as its input the result of an fMRI analysis, *viz.* a 3-D statistical map representing the contrast between two experimental conditions, routinely thresholded to give "blobs" of activity. Because this representation collapses the data across time, we are unable to use frequency of coactivation as a filtration method. Instead, we will filter these blobs by changing the threshold of the statistical map, by duration of activity, or by strength of anatomical connectivity derived from DTI tractography.

3.7 Interpretation of topological biomarkers using tractography

Finally, we would like to do more than distinguish between two or more populations of individuals, because diagnosis without insight into the mechanism of a disorder is blind. Resting state functional connectivity research led to studies of the "default mode network" and to excitement about making sense of functional connectivity in the context of anatomical structures and structural connections.[22, 28, 15, 13, 12] This development has been instrumental in helping to ground functional connectivity, which is extremely difficult to conceptualize and visualize, in physical structures and physiology of the brain. We hope to ground topological biomarkers in an analogous manner. Although difficult to predict what the outcome of this effort may be, it is possible that differences in complexity of the topological structures across individuals or after an event may reflect disruptions in tracts or abnormalities in the pattern of distribution of neurotransmitters or receptors in the brain. Or topological cycles (3.3.2) may reflect network architectural elements such as loops (feedback, recurrent, and converging neural pathways).

One method for relating topological biomarkers with anatomical structures and connections, already mentioned in (3.3.1) and exemplified in (3.4.3), is to display the names of regions involved in representative cycles. However, our understanding of the topological biomarkers would be greatly enhanced if we were to relate this anatomical information to independent and complementary data acquired from the same individual, such as by DTI data. We will try to relate these biomarkers and their related anatomical regions to strength of anatomical connections, and explain variation in the biomarkers with variation in the fractional anisotropy and tractography. Any such findings would help tether topological biomarkers to physical and physiological mechanisms, and would lead to a richer understanding of how the brain functions and how it fails. We understand that this component of the proposal is a long-term strategy, but we feel that our collaboration has the right complement of skills and interests to spearhead its development.

3.8 Conclusion and timeline

In the beginning of this proposal, we described the dire need for effective biomarkers of mental illness. We then presented formidable new computational tools that can find structure in complex data – tools that can allow explicitly high-order functional connectivity analyses in many regions, something not possible with current methods. After outlining our methodological innovations, we described our research approach to find biomarkers of, for example, major depressive disorder, and provided preliminary experimental results for ADHD data. We concluded by providing alternative strategies and enhancements, as well as a proposal to use diffusion tensor imaging to help interpret our biomarkers. Our timeline will be as follows:

Year 1: Create software to compute biomarkers within individuals using the concurrence-based framework.
Determine their range of variation. Test their validity on clinical (3.1.1: public and grant 1) data.
Begin development of statistical methods for combining persistence results across individuals.
Design evaluation and DTI comparison protocols.

Year 2: Continue with different (grant 2) data, region definitions, and persistence algorithms.
Compare their range of variation against DTI measures of anisotropy and network connectivity.
Continue development of statistical methods for combining persistence results across individuals.

Year 3: Continue with all work from Year 2.
Test validity of the biomarkers and evaluate their relationship to DTI measures.
Refine, test, and completely document the software for public distribution.

References

[1] Pierre-Louis Bazin and Dzung L. Pham. Homeomorphic brain image segmentation with topological and statistical atlases. *Medical Image Analysis*, 12(5):616–625, October 2008. 1.5

[2] David R. Brillinger. *Time Series: Data Analysis and Theory*. Society for Industrial and Applied Mathematics, Philadelphia, 2001. 1.4

[3] Donald S. Burdick. An introduction to tensor products with applications to multiway data analysis. 28:229–237, 1995. 1.4

[4] Gunnar Carlsson and Vin Silva. Zigzag Persistence. *Foundations of Computational Mathematics*, 10(4):367–405, April 2010. 3.6.2

[5] F Xavier Castellanos, Daniel S Margulies, Clare Kelly, Lucina Q Uddin, Manely Ghaffari, Andrew Kirsch, David Shaw, Zarrar Shehzad, Adriana Di Martino, Bharat Biswal, Edmund J S Sonuga-Barke, John Rotrosen, Lenard a Adler, and Michael P Milham. Cingulate-precuneus interactions: a new locus of dysfunction in adult attention-deficit/hyperactivity disorder. *Biological psychiatry*, 63(3):332–7, February 2008. 3.4.1, 3.4.3

[6] Moo K Chung, Peter Bubenik, and Peter T Kim. Persistence diagrams of cortical surface data. *Information processing in medical imaging : proceedings of the ... conference*, 21(1):386–97, January 2009. 1.5

[7] Carina Curto and Vladimir Itskov. Cell Groups Reveal Structure of Stimulus Space. *PLoS Computational Biology*, 4(10):e1000205, October 2008. 1.5, 3.2.1

[8] Rahul S Desikan, Florent Ségonne, Bruce Fischl, Brian T Quinn, Bradford C Dickerson, Deborah Blacker, Randy L Buckner, Anders M Dale, R Paul Maguire, Bradley T Hyman, Marilyn S Albert, and Ronald J Killiany. An automated labeling system for subdividing the human cerebral cortex on {MRI} scans into gyral based regions of interest. *NeuroImage*, 31:968–980, 2006. 2

[9] Herbert Edelsbrunner and John Harer. *Computational topology: an introduction*. 2008. 3.3.1, 3.3.2, 3.6.1

[10] B Fischl, a Liu, and a M Dale. Automated manifold surgery: constructing geometrically accurate and topologically correct models of the human cerebral cortex. *IEEE transactions on medical imaging*, 20(1):70–80, January 2001. 1.5

[11] K.J. Friston, C.D. Frith, P.F. Liddle, and R.S.J. Frackowiak. Functional connectivity: the principle-component analysis of large (PET) data sets. *Journal of Cerebral Blood Flow*, 13(1):5–14, April 1993. 1.3

[12] Michael D Greicius, Ben Krasnow, Allan L Reiss, and Vinod Menon. Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proceedings of the National Academy of Sciences of the United States of America*, 100(1):253–8, January 2003. 3.7

[13] Michael D Greicius, Kaustubh Supekar, Vinod Menon, and Robert F Dougherty. Resting-state functional connectivity reflects structural connectivity in the default mode network. *Cerebral cortex (New York, N.Y. : 1991)*, 19(1):72–8, January 2009. 3.7

[14] F Holsboer and N Barden. Antidepressants and hypothalamic-pituitary-adrenocortical regulation. *Endocrine reviews*, 17(2):187–205, April 1996. 1.2

[15] Christopher J. Honey, Rolf Kotter, Olaf Sporns, and Michael Breakspear. Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *Proceedings of the National Academy of Sciences*, 104(24):10240–10245, June 2007. 3.7

[16] a Hyvärinen and E Oja. Independent component analysis: algorithms and applications. *Neural networks : the official journal of the International Neural Network Society*, 13(4-5):411–30, 2000. 1.4

[17] Thomas R Insel and Bruce N Cuthbert. Endophenotypes: bridging genomic complexity and disorder heterogeneity. *Biological psychiatry*, 66(11):988–9, December 2009. 1.1

[18] Sylvain Jaume, Patrice Rondao, and B. Macq. Open Topology : A Toolkit for Brain Isosurface Correction. *The Insight Journal - 2005 MICCAI Open-Source Workshop*, 2005. 1.5

[19] 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*, 841:825– 841, 2002. 1

[20] Tomasz Kaczynski, Konstantin Mischaikow, and Marian Mrozek. *Computational Homology*. Springer, New York, 2004. 3.3.1

[21] Ronald C Kessler, Wai Tat Chiu, Olga Demler, Kathleen R Merikangas, and Ellen E Walters. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of general psychiatry*, 62(6):617–27, June 2005. 1.2

[22] Martin a Koch, David G Norris, and Margret Hund-Georgiadis. An investigation of functional and anatomical connectivity using magnetic resonance imaging. *NeuroImage*, 16(1):241–50, May 2002. 3.7

[23] Jean-Franï¿œois Mangin, Vincent Frouin, Isabelle Bloch, Jean Rï¿œgis, and Jaime Lï¿œpez-Krahe. From 3D magnetic resonance images to structural representations of the cortex topography using topology preserving deformations. *Journal of Mathematical Imaging and Vision*, 5(4):297–318, December 1995. 1.5

[24] J R Munkres. *Elements of Algebraic Topology*. Benjamin/Cummings, Menlo Park, CA, 1984. 3.3.1

[25] J.L. Prince. Cortical surface reconstruction using a topology preserving geometric deformable model. *Proceedings IEEE Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA 2001)*, pages 213–220. 1.5

[26] a John Rush, Madhukar H Trivedi, Stephen R Wisniewski, Andrew a Nierenberg, Jonathan W Stewart, Diane Warden, George Niederehe, Michael E Thase, Philip W Lavori, Barry D Lebowitz, Patrick J McGrath, Jerrold F Rosenbaum, Harold a Sackeim, David J Kupfer, James Luther, and Maurizio Fava. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. *The American journal of psychiatry*, 163(11):1905–17, November 2006. 1.2

[27] Guillermo Sapiro and Dario L Ringach. Topological analysis of population activity in visual cortex Gurjeet Singh. *Journal of Vision*, 8:1–18, 2008. 1.5, 3.6.1

[28] O Sporns, G Tononi, and G M Edelman. Theoretical neuroanatomy: relating anatomical and functional connectivity in graphs and cortical connection matrices. *Cerebral cortex (New York, N.Y. : 1991)*, 10(2):127–41, February 2000. 3.7

[29] N V Swindale. Visual cortex: looking into a Klein bottle. *Current biology : CB*, 6(7):776–9, July 1996. 1.5

[30] S. Tanaka. Topological Analysis of Point Singularities in Stimulus Preference Maps of the Primary Visual Cortex. *Proceedings of the Royal Society B: Biological Sciences*, 261(1360):81–88, July 1995. 1.5

[31] Madhukar H Trivedi, a John Rush, Stephen R Wisniewski, Andrew a Nierenberg, Diane Warden, Louise Ritz, Grayson Norquist, Robert H Howland, Barry Lebowitz, Patrick J McGrath, Kathy Shores-Wilson, Melanie M Biggs, G K Balasubramani, and Maurizio Fava. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: implications for clinical practice. *The American journal of psychiatry*, 163(1):28–40, January 2006. 1.2

[32] Natalie Tzourio, Laurent Petit, Emmanuel Mellet, Christophe Orssaud, Fabrice Crivello, Karim Benali, Georges Salamon, and Bernard Mazoyer. Use of anatomical parcellation to catalog and study structure-function relationships in the human brain. *Human Brain Mapping*, 5(4):228–232, 1997. 3.4.3

[33] Lucina Q Uddin, a M Clare Kelly, Bharat B Biswal, Daniel S Margulies, Zarrar Shehzad, David Shaw, Manely Ghaffari, John Rotrosen, Lenard a Adler, F Xavier Castellanos, and Michael P Milham. Network homogeneity reveals decreased integrity of default-mode network in ADHD. *Journal of neuroscience methods*, 169(1):249–54, March 2008. 3.4.1, 3.4.3

[34] Martijn P van den Heuvel and Hilleke E Hulshoff Pol. Exploring the brain network: A review on resting-state fMRI functional connectivity. *European neuropsychopharmacology : the journal of the European College of Neuropsychopharmacology*, 20(8):519–534, May 2010. 1.4, 1.5