Graduate College Dissertations and Theses

Dissertations and Theses

2015

Reverse Engineering the Human Brain: An Evolutionary Computation Approach to the Analysis of fMRI

Nicholas Allgaier

University of Vermont, nallgaier@gmail.com

Follow this and additional works at: http://scholarworks.uvm.edu/graddis

Part of the Computer Sciences Commons, Mathematics Commons, and the Neurosciences Commons

Reverse Engineering the Human Brain: An Evolutionary Computation Approach to the Analysis of fMRI

A Dissertation Presented

by

Nicholas Allgaier

to

The Faculty of the Graduate College

of

The University of Vermont

In Partial Fullfillment of the Requirements
for the Degree of Doctor of Philosophy
Specializing in Mathematical Sciences

January, 2015

Defense Date: November 10, 2014
Dissertation Examination Committee:

Chris Danforth, Ph.D., Advisor
Peter Sheridan Dodds, Ph.D., Advisor
Josh Bongard, Ph.D., Chairperson
Hugh Garavan, Ph.D.
Cynthia J. Forehand, Ph.D., Dean of the Graduate College

Abstract

The field of neuroimaging has truly become data rich, and as such, novel analytical methods capable of gleaning meaningful information from large stores of imaging data are in high demand. Those methods that might also be applicable on the level of individual subjects, and thus potentially useful clinically, are of special interest. In this dissertation we introduce just such a method, called nonlinear functional mapping (NFM), and demonstrate its application in the analysis of resting state fMRI (functional Magnetic Resonance Imaging) from a 242-subject subset of the IMAGEN project, a European study of risk-taking behavior in adolescents that includes longitudinal phenotypic, behavioral, genetic, and neuroimaging data.

Functional mapping employs a computational technique inspired by biological evolution to discover and mathematically characterize interactions among ROI (regions of interest), without making linear or univariate assumptions. Statistics of the resulting interaction relationships comport with recent independent work, constituting a preliminary cross-validation. Furthermore, nonlinear terms are ubiquitous in the models generated by NFM, suggesting that some of the interactions characterized here are not discoverable by standard linear methods of analysis. One such nonlinear interaction is discussed in the context of a direct comparison with a procedure involving pairwise correlation, designed to be an analogous linear version of functional mapping. Another such interaction suggests a novel distinction in brain function between drinking and non-drinking adolescents: a tighter coupling of ROI associated with emotion, reward, and interceptive processes such as thirst, among drinkers. Finally, we outline many improvements and extensions of the methodology to reduce computational expense, complement other analytical tools like graph-theoretic analysis, and possibly allow for voxel level functional mapping to eliminate the necessity of ROI selection.

CITATIONS

Material from this dissertation will be submitted for publication in *Neuroimage* in the following form:

Allgaier, N.A., Whelan, R., Banaschewski, T., Barker, G.J., Bokde, A.L.W., Bongard, J.C., Bromberg, U, Büchel, C., Conrod, P., Danforth, C.M., Dodds, P.S., Flor, H., Frouin, V., Gowland, P., Heinz, A., Ittermann, B., Mann, K., Martinot, J-L., Nees, F., Paus, T., Pausova, Z., Rietschel, M., Robbins, T., Smolka, M.N., Ströhle, A., Schumann, G., Garavan, H., IMAGEN consortium. (2014). Nonlinear Functional Mapping of the Human Brain. *Neuroimage.*

Dedicated to

Deana, without whom I would surely languish, and our beautiful daughter Ruby Mae.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF FIGURES

List of Tables

Chapter 1

Introduction and Literature Review

In this chapter we introduce the topic of this dissertation, a novel methodology for inferring nonlinear models of brain function from observational data, and frame it within the context of neuroimaging research. In particular, we discuss the development of mathematical and computational tools for the analysis of functional magnetic imaging (fMRI) data, focusing primarily on the last decade of advances, and also describe some recent applications of those tools, emphasizing the potential contribution of our methodology along the way. In Section 1.1, we provide motivation and general background information. We discuss the relevant neuroimaging literature in Section 1.2, and a present a brief review of the literature describing a central component of the proposed technique, an evolutionary algorithm called *genetic programming* (GP), in Section 1.3. We conclude the introduction in Section 1.4 with a general outline of our procedure, and summary of its potential contributions to neuroimaging research.

1.1 Introduction

Neurological research aims to improve our understanding of the relationship between brain function and physical, emotional, and cognitive processes, and develop diagnostic and treat-

ment regimes for neurological disorder. The search for neural correlates of certain cognitive processes (e.g., language), behavioral proclivities (e.g., substance misuse), and neurological disfunction (e.g., autism spectrum disorders) require the investigation of a functioning human brain, and fMRI provides that capability.

The human brain can be observed, in vivo and non-invasively, by fMRI. The primary benefit of fMRI over other modes of observation, such as electroencephalography, is spatial resolution, typically on the order of a few millimeters. Fine spatial resolution, however, comes at the cost of temporal resolution. Though the blood oxygen level dependent (BOLD) signal measured by fMRI is only a proxy for neuronal activity, research has shown that with proper treatment it can be a reliable one, for example in Mukamel et al. (2005); Birn et al. (2009); Murphy et al. (2011).

Many advances in our understanding of brain function have been achieved through analysis of fMRI data, and indeed, primarily linear and univariate approaches have yielded impressive results for the past 20 years. Some early accomplishments include assessment of the risk of postoperative motor defect in patients with tumors in Mueller et al. (1996), localization of regions responsible for particular tasks, such as episodic memory in Nolde et al. (1998), and human face recognition in Kanwisher et al. (1999). We discuss a number of recent, exciting achievements in more detail in Section 1.2.

Large stores of imaging data provide an exciting opportunity to expand upon these achievements. Desirable characteristics of methods for deriving insight from "Big Data" include the capability to incorporate information from multiple domains (e.g., neuroimaging and genetic), interpretability of explanatory or predictive models, and minimal assumptions about the nature of relationships among observed quantities. The methodology we present in this dissertation has all of these characteristics, and in particular, avoids the assumption of linearity.

In the following literature review, we discuss many of the mathematical tools used to derive insight from imaging data, focusing on the last decade of research. We assess which of the desirable characteristics are possessed by each methodology, and describe how our procedure might complement various approaches and contribute to the body of neuroimaging research along the way.

1.2 MATHEMATICAL TOOLS FOR NEUROIMAGING

The majority of fMRI studies make use of the GLM to determine neural correlates for various tasks and stimulus responses, as in the previously mentioned studies Mueller et al. (1996); Nolde et al. (1998); Kanwisher et al. (1999). Though this approach has been used since the advent of fMRI 20 years ago, creative applications continue to produce results today. Here we'll review some recent work reported in Ide et al. (2013), in which the GLM was used to determine neural correlates of the components in a hypothesized computational model of inhibitory control. Ten years earlier, in the review article O'Doherty (2004), reward learning components of a hypothesized model were already being attributed to particular regions of the brain. In the later review Gläscher and O'Doherty (2010), the steps were outlined for the general procedure of combining fMRI, computational models of cognitive processes, and the GLM for the purposes of attributing model-component encodings. In Ide et al. (2013) we see the culmination of this approach.

Ide et al., 2013. Bayesian Prediction and Evaluation in the Anterior Cingulate Cortex. The Journal of Neuroscience.

In this work, a Bayesian ideal observer model is posited to be the driver of inhibitory control as executed in the stop-signal task, in which subjects performing a repetitive go task (i.e., press a button at a "go" signal) are periodically signaled to inhibit by the

3

presentation of a "stop" signal at some delay after a "go" signal. The delay is adjusted so that each subject succeeds at inhibiting in approximately half of the stop trials. The study includes 66 healthy subjects, teenagers to 60 year-olds split evenly along gender lines, and the main goal is to determine quantitatively what components of the posited model are encoded in a particular region of the brain, the dorsal anterior cingulate cortex (dACC), which has been implicated in a number of related cognitive control functions.

Behavioral data and fMRI were recorded for each subject while performing the task. The first analysis performed was a successful verification that the behavioral data indicated the use of Bayesian predictions by the participants, based on predicted variation of reaction times. To accomplish the main goal of attribution of model components to the dACC, regression models were generated for the various components of several different versions of the Bayesian model, corresponding to different possible categorizations of trial outcomes and different monitored quantities (e.g., signed and unsigned prediction error). The resulting five GLM partitioned the model components in such a way as to allow the attribution of particular (even sub-) components to the dACC, while separating other possibly dependent components.

The results of the study strongly suggest the encoding of Bayesian "surprise of stop" (unsigned prediction error of a stop trial), and signed prediction error of an error trial, in the dACC. Further, once these effects are removed (as was determined by the creative use of multiple GLM), it appears that the dACC does not encode probabilistic expectations of either error or stimulus type overall, nor does it have any residual correlation with conflict or error likelihood as had been previously hypothesized.

Applications of this methodology are not limited to Bayesian models and the stop-signal task, of course, and in Wiecki and Frank (2013) similar strategies are employed using a neural circuit model involving several regions of the brain that include a frontal executive control component, and an action selection component in the basal ganglia to execute

inhibitory control in a variety of tasks. Despite these creative uses of the GLM, however, it is necessarily blind to nonlinear relationships of any kind, and can only *confirm* hypothesized computational models of function, not *discover* them. The methodology presented in this dissertation provides the capability to uncover nonlinear relationships, and also has the potential to reveal computational models of cognitive function, as we'll discuss in Chapter 2.

Group-level inferences from fMRI have also been performed using linear ICA (independent component analysis), as described in Calhoun et al. (2001). Though ICA and the GLM can be used in conjunction, for example in Liu et al. (2010) to investigate the neural effects of stimulation of a particular acupoint, ICA is particularly useful in circumstances that preclude the use of the GLM, such as the analysis of resting-state data, for which there is no task or stimulus regressor. Covarying networks have been suggested by ICA of resting-state fMRI in Smith et al. (2009), and functional, hierarchical classification of these networks has been automated through HCA (hierarchical cluster analysis) of aggregated experimental metadata in Laird et al. (2011), work which we review in detail here.

Laird et al., 2011. Behavioral interpretations of intrinsic connectivity networks.. The Journal of Cognitive Neuroscience.

In Smith et al. (2009), ICA was carried out on thousands of activation maps from the BrainMap database, and, independently, on 36 subjects at rest, and the resulting low-dimensional decompositions were compared. The main result, based on the similarity of the decompositions, was that networks of brain regions that cooperate during the performance of tasks or stimulus response are *also* active and covarying during rest. Surprisingly, it was found that the amplitude of variation is not diminished during rest, as might be expected.

The main goal of Laird et al. (2011) was to automatically generate a functional classification of these identified intrinsic connectivity networks (ICN) by hierarchical cluster analysis (HCA) of the metadata in the BrainMap database. This classification was accomplished in four stages: (1) activation images were reconstructed from peak coordinates from the 8637 published experiments catalogued in the BrainMap database; (2) ICA was performed on these activation images to generate 20 spatial maps interpreted as ICN; (3) each experimental activation image was expressed as the linear mixing of the ICN, allowing the assignment of weights to each ICN for each experiment, *and* its associated metadata (14 fields, with specific classes of allowable entries), which was then aggregated over experiments; and (4) HCA was performed on the aggregated ICN-metadata weightings to cluster metadata values into behavioral domains, and then ICN into functional classifications.

The results are presented as a resource of functional interpretations for networks of brain regions, and indeed, they are used precisely this way in the body of this dissertation, presented in Chapter 2. Functional groupings include visual, motor/visuospatial, and emotional/interoceptive clusters, and networks that were not clustered were interpreted as divergent cognitive, and were often associated with audition and speech. An interesting example is the default mode network (DMN), the suppression of which has been associated with improved performance in a number of tasks. For example, in the review Anticevic et al. (2012), evidence of DMN suppression across disciplines is discussed in the context of health and disease, and in particular, associations between suppression *deficits* and severe mental illness are presented.

These studies provide strong evidence for the existence of a hierarchical, *functional* interactome of regions within the human brain, which are not necessarily directly connected *physically*. The relationship between physical and functional structure is an open question in neuroscience, and Haimovici et al. (2013) present a compelling theory that connects critical dynamics on the structural network of the brain with a number of experimental

functional findings, including the emergence of resting state networks like those described in Smith et al. (2009); Laird et al. (2011). The network perspective on brain function is a theme of the modern literature, and any achievable relaxation of assumptions in the inference of network architecture is desirable. For example, it was determined early on, as in McKeown and Sejnowski (1998), that nonlinear interactions within the brain need to be addressed in order to properly determine functional architecture.

Although ICA algorithms that employ nonlinear mixing functions exist, severe restrictions on those functions are required to avoid non-uniqueness of solutions, as explained in Hyvärinen and Pajunen (1999). Due to this failing, other methodologies have been employed in the attempt to account for nonlinearity. Examples include various forms of nonlinear regression, as in Kruggel et al. (2000), and dynamic causal modelling, as described in Friston et al. (2003). In the former, a particular nonlinear form must be posited *a priori*. In the latter, though nonlinear effects can be incorporated into the model, the model form is such that those effects are not *interpretable*, and thus the capability to *discover* previously unknown nonlinear interactions within the brain is diminished. Nonlinear functional mapping is free-form, and provides directly interpretable mathematical characterizations of function that can be used to infer both small-scale inter-region interaction, and larger scale hierarchical network functional architecture, as we'll demonstrate in Chapter 2.

Structure of the human brain functional interactome, as well as that of the physical connectome, has been shown to posses a number of topological properties of complex networks (e.g., small-world architecture, modularity, hubs) as described in Bullmore and Sporns (2009). As a result, over the past 5 years a number of graph-theoretic analytical tools, in particular a collection of complex network summary measures that provide a quantitative description of both local and global connectivity properties, have been brought to bear in the domain of neuroimaging. A comprehensive description of these measures, and various examples of their applicability, are described in Rubinov and Sporns (2010). A freely avail-

able MATLAB toolbox for computing these measures accompanies the article, and can be accessed at http://www.brain-connectivity-toolbox.net.

Here we review in more detail Bullmore and Sporns (2012), which describes a likely mechanism for the development of complex network architecture, both structural and functional, within the brain. Indeed, the proposed economically motivated mechanism suggests a plausible and direct connection between the structural and functional network organizations. The crux of the argument presented in this work is that brain structure, human and otherwise, is *not* governed exclusively by material and metabolic cost minimization, but instead by an economic balance of that minimization with the allowance for potentially costly, but adaptively valuable, topological patterns of both physical and functional connectivity.

Bullmore and Sporns, 2012. The economy of brain network organization. Nature Reviews Neuroscience.

The basis for cost minimization, derived from the spatial embedding of the network, is revealed in the form of various empirically established allometric scaling laws. White-matter vs. grey-matter scaling supports wiring (material) cost minimization, for example, while the scaling of glucose metabolism with brain volume suggests metabolic cost minimization. Additionally, empirical evidence shows fractal scaling of connections within the human brain, (i.e., a power law governs the scaling of the number of connections to a region of grey-matter with the neuronal density of that region), which is considered indicative of cost-efficient spatial embedding. However, further empirical evidence is presented showing that brain networks are not *strictly* cost-minimal, in either the material or metabolic sense, and indeed the presence of complex network architecture (e.g., small-world), which often comes at a cost premium, demonstrates cost sub-optimality.

Such complex network properties provide enormous benefits, however, including simultaneous segregated (specialized) and integrated (distributed) information processing, and robustness to insult through fast, large-scale reconfiguration in response to perturbation. To substantiate the claimed economic trade-off between cost-minimization and this desirable topological complexity, hypothetical computational rewirings of animal (monkey, worm) brain networks that minimize cost are shown to reduce network efficiency. Similar experiments for the human brain are presented, based on fMRI and magnetoencephalography, though they are necessarily less precise. This experimental evidence supports the theorized, effectively multi-objective, optimization of cost-*efficiency*.

Additionally, more dynamic measurements of functional connectivity within single subjects (e.g., during task switching), and comparisons of connectivity between age demographics (i.e., children vs. adolescents, young vs. elderly adults), suggest that this cost-efficiency optimization occurs on multiple time scales, allowing for task-demand, developmental, *and* evolutionary-scale adaptation. In other words, the mechanism driving complex architecture of brain networks is a cost-constrained spatial embedding subject to efficiency demand from selection (evolutionary), aging (developmental), and task-induced pressures. Finally, several neurological disorders are cast in the framework of brain network economy, and indeed associated topological deficiencies (e.g., reduced long-distance connections, greater path-length, and clustering for Alzheimer's disease) are shown to be consistent with this interpretation.

This work provides compelling motivation for approaching the investigation of brain function from a complex networks viewpoint, and though much has been learned in this vein, network models do face certain limitations and challenges (e.g., various biases in network measures, challenges in network definition, and limited insight into neural computation mechanisms). A concise presentation of recent accomplishments and limitations is provided in Sporns (2014). We emphasize that NFM addresses some of these limitations directly.

Specifically, functional network definition by NFM, while vastly more computationally expensive than the typical pairwise correlation, would provide a fuller, more accurate picture of network dynamics by capturing nonlinear interaction. Also, crucially, NFM may provide direct and interpretable insight into neural computation mechanisms.

To conclude our review of the literature on mathematical tools of neuroimaging, we discuss a recent, exciting application of machine learning to generate predictive models of both current and future binge drinking in adolescents (current = 14 years old, future = 16 years old) participating in the IMAGEN project, a longitudinal collection of multi-domain data including fMRI, genetic data, personality and history survey, and self-report behavioral questionnaires. Previous work with this same dataset reported in Whelan et al. (2012) demonstrated that reduced activity in a specific orbitofrontal subcortical network during the stop-signal task was associated with a higher likelihood for initiation of drug use. In Whelan et al. (2014) the results of a large scale, multi-domain predictor/classifier of alcohol misuse are presented.

Whelan et al., 2014. Neuropsychosocial profiles of current and future adolescent alcohol misusers. Nature.

The predictive model presented in this study consists of a logistic regression, with elastic net regularization and feature selection, across seven different data domains: brain, personality, history, cognition, demographics, genetics and site. In addition to the model incorporating all domains, the regression was also performed for each domain separately to asses individual predictive capabilities, and excluding each individual domain to determine its orthogonal contribution to the full model. In each case, tenfold cross validation was implemented, with three levels of nested cross-validation in order to: (1) optimize imaging

thresholds for the fMRI data; (2) optimize the two regularization parameters of the elastic net; and (3) validate the models.

Models were generated using this procedure to classify current, and predict future binge drinkers, and the area-under-the-curve (AUC) receiver-operator characteristic (ROC) was used to measure effectiveness. The classifier of current alcohol misuse had an AUC of 0.9 after removal of smoking from the model (which is highly co-occurent with alcohol use). The most robust brain classifiers for current alcohol misuse were the ventromedial prefrontal cortex, and the left inferior frontal gyrus, and the history domain provided the most unique contribution to the model (i.e., the worst 6-domain model neglected history).

The predictive model resulted in an AUC of 0.75, with a 66% rate of correct binge drinking prediction at the optimal point on the AUC curve. Interestingly, the predictive power of the brain and personality domains were on par with history (individual AUC of 0.63, 0.67, and 0.68, respectively). External validation was performed for the predictive model, which performed about as well on the fully independent data as it did within the tenfold cross-validation. The most predictive brain regions were the right middle and precentral gyri, and bilateral superior frontal gyrus.

As might be expected, a number of characteristics were effective classifiers and predictors, for example the novelty seeking trait from the personality domain, and many features were effective in only one of the model classes. Crucially, no single feature provides significant predictive value, highlighting the benefit of high-dimensional, multi-domain analyses. Risk profiles generated by such analyses are unquestionably useful clinically, allowing in this case, for example, potential targeted interventions to prevent the progression from binge drinking to alcoholism.

The implementation of such high-dimensional, multi-domain analyses often requires machine learning algorithms like the elastic net used in Whelan et al. (2014). Extensions of that technique are described in Grosenick et al. (2013), where another important benefit of

this algorithm is discussed: interpretability. In the effort to better understand brain function, predictive power is not the only concern; models that convey insight about functional interaction are required. However, elastic nets, like the majority of the mathematical tools in the neuroimaging literature, come with the assumption of linearity built in. The central component of NFM is symbolic regression performed by an evolutionary algorithm called genetic programming. The algorithm generates directly interpretable models, and does not require an assumption of linearity. In the next section we review a demonstration of the technique showing inference of natural physical laws (most of them nonlinear) directly (and only) from data.

1.3 Genetic programming

In this section we review an example application of symbolic regression by the evolutionary algorithm called genetic programming, and briefly discuss several modifications proposed in the literature that may be of particular interest in tailoring the algorithm for use in the neuroimaging domain. None of these modifications has yet been incorporated into the NFM procedure we demonstrate in Chapter 2, and thus they form a natural next phase in the development of the technique.

Schmidt and Lipson, 2009a. Distilling free-form natural laws from experimental data. Science.

In this work, a computational technique, consisting of symbolic regression by genetic programming, is demonstrated to automatically determine analytical expressions for physical processes, and natural laws by which they are governed, directly from observational data. In particular, equations of motion, Lagrangians, and Hamiltonian invariants are recovered for single and double air-track oscillators and pendula. An important benefit of

the technique is the customizability of the fitness metric used to evaluate models throughout evolution, which is illustrated in this work by the use of a partial-derivative-pairings metric designed to avoid trivial invariants (of which there are infinitely many for any physical system).

Another virtue of the algorithm is the production of a collection of models that balance explanatory power (as determined by user-specified fitness metric) with parsimony. A "winning" model can be chosen from this collection, as in this work, by selecting the simplest model with greatest increase in predictive ability. For each of the physical systems examined in the study, this model was an exact theoretical law (e.g., the double pendulum's Hamiltonian). Alternatively, the entire collection of models can be analyzed broadly (especially for the analysis of a large number of related model searches) to probabilistically infer underlying relationships among the observed variables. This is the approach employed in NFM.

Finally, a bootstrapping approach involving the seeding of model-search for more complicated systems (e.g., double pendulum) with results for related, simpler systems (single pendulum), is shown to significantly reduce required search time (by an order of magnitude for the pendulum example). This strategy will likely be an effective tool for the application of the algorithm to high-dimensional data from more complex phenomena.

Automated condensation of observational data into meaningful analytical expressions does not replace the role of the scientist. Rather, this capability facilities the rapid progression from observation to mathematical characterization, and only the interpretation and application of that characterization by scientists constitutes knowledge. Furthermore, though the algorithm is capable of fully automating analytical expression of processes from the data they generate, cooperation between the scientist and the algorithm will likely outperform either in isolation for this endeavor. This was already suggested by the efficacy of

bootstrapping the evolutionary search for the dynamical laws of the double pendulum in Schmidt and Lipson (2009a).

A thorough examination of this human-computer cooperative methodology is presented in Schmidt and Lipson (2009b), in which the seeding of model search with expert knowledge is investigated. An interesting aspect of the results is that seeding the search with simpler expressions of expert knowledge tended to outperform more complex seedings. Incorporation of expert knowledge is one example extension of the algorithm that will likely prove beneficial in the application of GP to the study of brain function. A vast number of other modifications and extensions are explored in the literature. Here we briefly describe two more that might also prove effective in the neuroimaging domain: a method for handling noisy data, and data reduction.

An effective extension of GP for model-search in the presence of noise is the inclusion of stochastic elements with user-defined distributions (e.g., Gaussian or uniform) as potential explanatory "variables", as described in Schmidt and Lipson (2007). These generators can themselves end up inside complex functions within the models, providing those models the capability of reproducing realistic noise distributions more likely to be at play than the typical Gaussian. In that study, this modified version of GP was shown to effectively identify exact underlying analytical models in the presence of nonlinear, non-Gaussian and nonuniform noise.

Data dimensionality (i.e., the number of potential explanatory variables) has a large effect on GP search time. To effectively apply the algorithm to high-dimensional data, a combination approach involving some sort of feature selection and data reduction is likely necessary. One such method, described in Icke et al. (2014), is a hybrid method of symbolic regression employing a machine learning algorithm called FFX (Fast Function Extraction) described in McConaghy (2011) as a first pass, and then GP search on the resulting reduced set of potential explanatory variables (which may include nonlinear terms expressing inter-

action among the original variables). In Icke et al. (2014), this hybrid approach is shown to outperform GP alone on an fMRI dataset, and thus it will likely prove beneficial in the wider application of symbolic regression in the neuroimaging domain.

1.4 CONCLUSION

Though considerable progress has been made in our understanding of brain function, using a cadre of mathematical and computational strategies spanning the gamut of sophistication, a detailed mathematical characterization of human brain function remains on the frontier of modern science. In this dissertation we introduce a methodology called nonlinear functional mapping (NFM) designed to contribute to such a characterization, and provide insight at the group, subject, and ROI levels. NFM provides the capability to incorporate *multi-domain* data in the free-form generation of *nonlinear* and *directly interpretable* mathematical models. No other method possess all three of these capabilities.

Our procedure consists of ROI selection, inter-ROI *symbolic regression* (a model-free form of nonlinear regression), accomplished by an evolutionary algorithm called genetic programming (a form of stochastic optimization), and statistical analysis of the resulting models. With some modification, analysis of higher dimensional data is likely attainable, allowing for eventual application at the voxel scale and eliminating the necessity of ROI selection. In Chapter 2, we demonstrate an application of NFM to the analysis of resting state fMRI from the IMAGEN project, a European study of adolescents that includes longitudinal phenotypic, behavioral, genetic, and neuroimaging data.

Chapter 2

Nonlinear Functional Mapping of the Human Brain

The field of neuroimaging has truly become data rich, and novel analytical methods capable of gleaning meaningful information from large stores of imaging data are in high demand. Those methods that might also be applicable on the level of individual subjects, and thus potentially useful clinically, are of special interest. In the present study, we introduce just such a method, called *nonlinear functional mapping* (NFM), and demonstrate its application in the analysis of resting state fMRI (functional Magnetic Resonance Imaging) from a 242-subject subset of the IMAGEN project, a European study of adolescents that includes longitudinal phenotypic, behavioral, genetic, and neuroimaging data. NFM employs a computational technique inspired by biological evolution to discover and mathematically characterize interactions among ROI (regions of interest), without making linear or univariate assumptions. We show that statistics of the resulting interaction relationships comport with recent independent work, constituting a preliminary cross-validation. Furthermore, nonlinear terms are ubiquitous in the models generated by NFM, suggesting that some of the interactions characterized here are not discoverable by standard linear methods of analysis. We discuss one such nonlinear interaction in the context of a direct comparison

with a procedure involving pairwise correlation, designed to be an analogous linear version of functional mapping. We find another such interaction that suggests a novel distinction in brain function between drinking and non-drinking adolescents: a tighter coupling of ROI associated with emotion, reward, and interceptive processes such as thirst, among drinkers. Finally, we outline many improvements and extensions of the methodology to reduce computational expense, complement other analytical tools like graph-theoretic analysis, and allow for voxel level NFM to eliminate the necessity of ROI selection.

2.1 INTRODUCTION

Many advances in our understanding of brain function have been achieved through analysis of fMRI data. Though the BOLD (blood oxygen level dependent) signal obtained from fMRI is a proxy, physiological confounds such as breathing and heart rate are separable from neuronal-induced signal, as demonstrated in Birn et al. (2009). Inter-subject differences in cerebral blood flow and volume can be modeled as shown in Murphy et al. (2011), and BOLD has been directly shown to provide a reliable measure of neuronal activity in specific circumstances, as in Mukamel et al. (2005). The many years of successful research before and since support that assessment. Accomplishments include localization of regions responsible for particular tasks, such as episodic memory in Nolde et al. (1998) and human face recognition in Kanwisher et al. (1999), assessment of the risk of postoperative motor defect in patients with tumors in Mueller et al. (1996), analysis of the effects of acupuncture in Hui et al. (2000), and recently, identification of neural markers for both current *and future* alcohol use among adolescents in Whelan et al. (2012) and Whelan et al. (2014).

These examples, and indeed the majority of fMRI studies, make use of the GLM (general linear model) to determine neural correlates for various tasks and stimulus responses. Additionally, recent work reported in Ide et al. (2013) investigated a hypothesized computational model of inhibitory control, and used the GLM to determine neural correlates of

that *cognitive* function. Though typical analyses have been performed at the group level with a univariate approach, other recent work reported in Rio et al. (2013) has extended the capabilities of the GLM to analyze multivariate signal in the Fourier domain to reduce confounds from time-correlated noise, thus improving the suitability of the GLM for subject level analysis. Despite these advances, however, the GLM is necessarily blind to nonlinear relationships of any kind, i.e., can only *confirm* hypothesized nonlinear models of function, not *discover* them.

Group-level inferences from fMRI have also been performed using linear ICA (independent component analysis), as described in Calhoun et al. (2001). Though ICA and the GLM can be used in conjunction, for example in Liu et al. (2010) to investigate the neural effects of stimulation of a particular acupoint, ICA is particularly useful in circumstances that preclude the use of the GLM, such as the analysis of resting-state data, for which there is no task or stimulus regressor. Covarying networks have been suggested by ICA of resting-state fMRI in Smith et al. (2009), and functional, hierarchical classification of these networks has been automated through HCA (hierarchical cluster analysis) of aggregated experimental metadata in Laird et al. (2011). However, it was determined early on, for example in McKeown and Sejnowski (1998), that nonlinear interactions within the brain need to be addressed in order to properly determine functional architecture.

Although ICA algorithms that employ nonlinear mixing functions exist, severe restrictions on those functions are required to avoid non-uniqueness of solutions, as explained in Hyvärinen and Pajunen (1999). Due to this failing, other methodologies have been employed in the attempt to account for nonlinearity. Examples include various forms of nonlinear regression, as in Kruggel et al. (2000), and dynamic causal modelling, as described in Friston et al. (2003). In the former, a particular nonlinear form must be posited *a priori*. In the latter, though nonlinear effects can be incorporated into the model, the model form is such that those effects are not *interpretable*, and thus the capability to *discover* previously un-

known nonlinear interactions within the brain is diminished. As a result, a fuller picture of the *nature* of intra- and inter-network functional connectivity within the brain is missing from the literature.

Here we introduce a methodology designed to accomplish such a mathematical characterization, provide insight at the group, subject, and ROI levels, and through which linear and univariate assumptions are avoided. With some modification, analysis of higher dimensional data is likely attainable, allowing for eventual application at the voxel scale and eliminating the necessity of ROI selection. After standard preprocessing (slice-timing and motion correction, normalization, smoothing, etc.), our procedure consists of ROI selection, inter-ROI *symbolic regression* (a model-free form of nonlinear regression), accomplished by an evolutionary algorithm called *genetic programming* (GP; a form of stochastic optimization), and statistical analysis of the resulting models. We demonstrate our technique on a 242-subject collection of resting-state data from the IMAGEN project, though analysis of task or stimulus experiments can be accomplished with little or no modification.

We organize the paper as follows. In Section 2.2, we discuss the data and selection of ROI, provide some background on GP, and describe the procedural details of NFM by symbolic regression. In Section 2.3, we report results of applying the technique to the IMAGEN data, including statistical and hierarchical visualizations, comparison with previous results for cross-validation, effects of nonlinearity, and an example of group-level variation. We discuss the results and potential applications of the technique in Section 2.4, and conclude the paper in Section 2.5.

2.2 MATERIALS AND METHODS

In this section, we first briefly describe the source of the data for our study, and then provide the details of ROI selection that allow for comparison with recent work. We then provide some background on the GP algorithm in general and the specific implementation employed

here, along with the method by which it is applied to BOLD signal time series extracted from the selected ROI. Finally, we describe the statistical technique used to interpret the roughly quarter of a million mathematical models that result from the application of GP to all 52 ROI time series extracted from each of the 242 subjects.

2.2.1 DATA

The data investigated here are a subset of the fMRI scans from the IMAGEN study, a European research project with the goal of better understanding teenage risk-taking behavior. The project is longitudinal, and utilizes several forms of high and low-tech experimental protocols including self-report questionnaires, behavioral assessment, interviews, neuroimaging, and blood sampling for genetic analyses. Each of the 2000 participating adolescents was 14 when entering the study, which itself commenced in late 2007, and data collection continues today.

More specifically, the data for the present study are 6-minute resting-state fMRI time series of 242 of the adolescent subjects who were asked to keep their eyes open while in the scanner, but were presented with no other task or stimulus. To allow for comparison with previous work, locations of the ROI were chosen based on results from Laird et al. (2011), in which statistical analysis across thousands of previous imaging studies (both stimulus/task-based and resting-state) was used to identify networks of brain regions that tend to activate together, termed ICN (intrinsic connectivity networks). The ICN were determined by ICA, from which z-statistic maps were derived. To select ROI for this study, a z-statistic threshold was set for each ICN to determine the number of regions in the network, and ROI were defined as rough spheres with radii of 3 voxels (9mm) and centered at the location of peak z-statistic in each region.

We provide a cut-out illustrating ROI selection for the default mode network (ICN 13) in Figure 2.1. Figure 2.2 contains axial cross sections showing many of the ROI derived
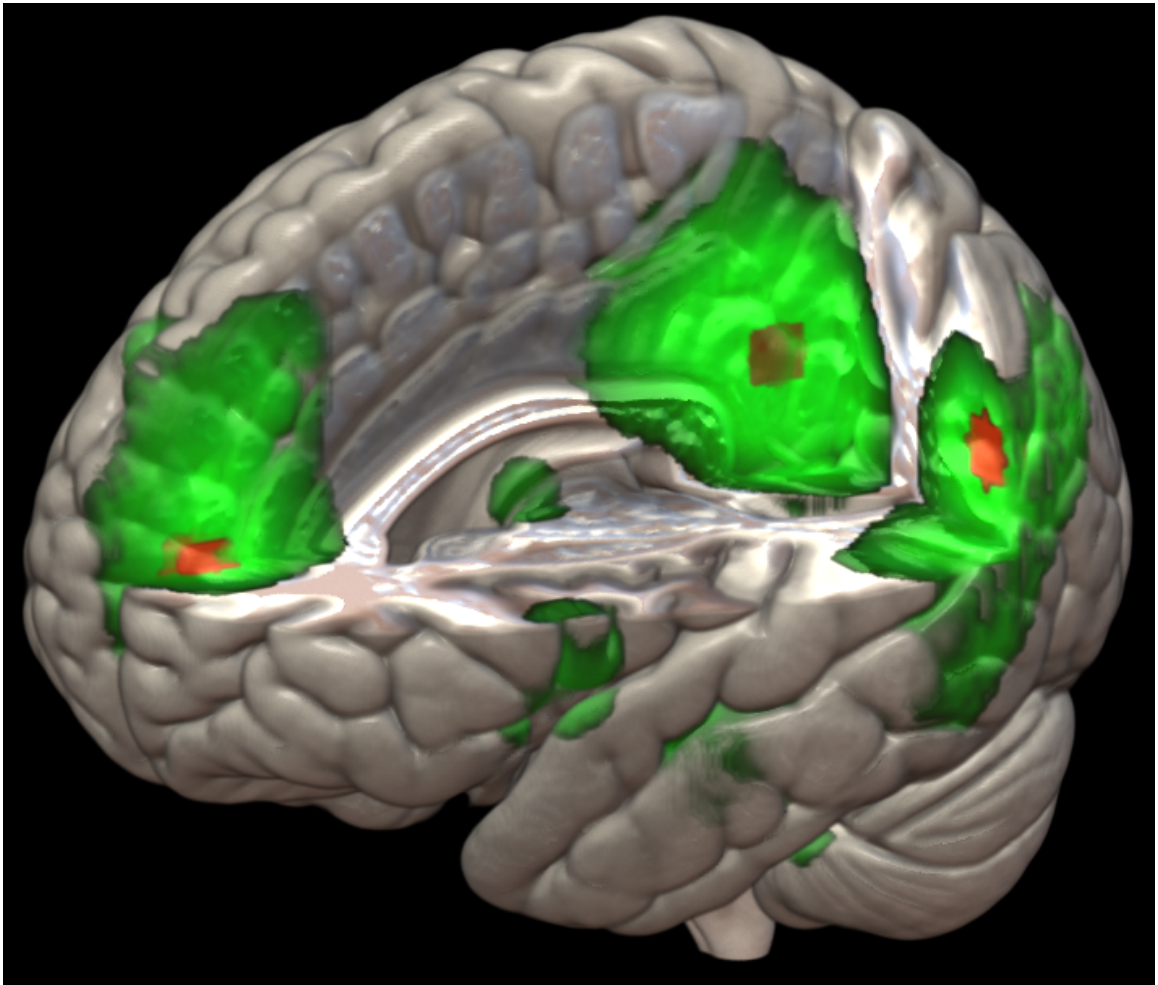
Figure 2.1: ROI Selection. (Red) ROI from within the default mode network, with radii of 3 voxels and centers corresponding to the highest z-statistics (green) in each region as determined in Laird et al. (2011).

from the 18 non-artifactual ICN in Laird et al. (2011). Subsequent to ROI definition, a gray matter mask was applied to assure that only appropriate voxels were contained within each ROI. In some cases this resulted in a considerable reduction of ROI voxels, but the majority maintained the full complement of about 100 voxels. For each of the 242 subjects, time series were extracted from each of the 52 resulting ROI by averaging the BOLD signal over all voxels within the ROI. These time series then form the input to the GP algorithm.
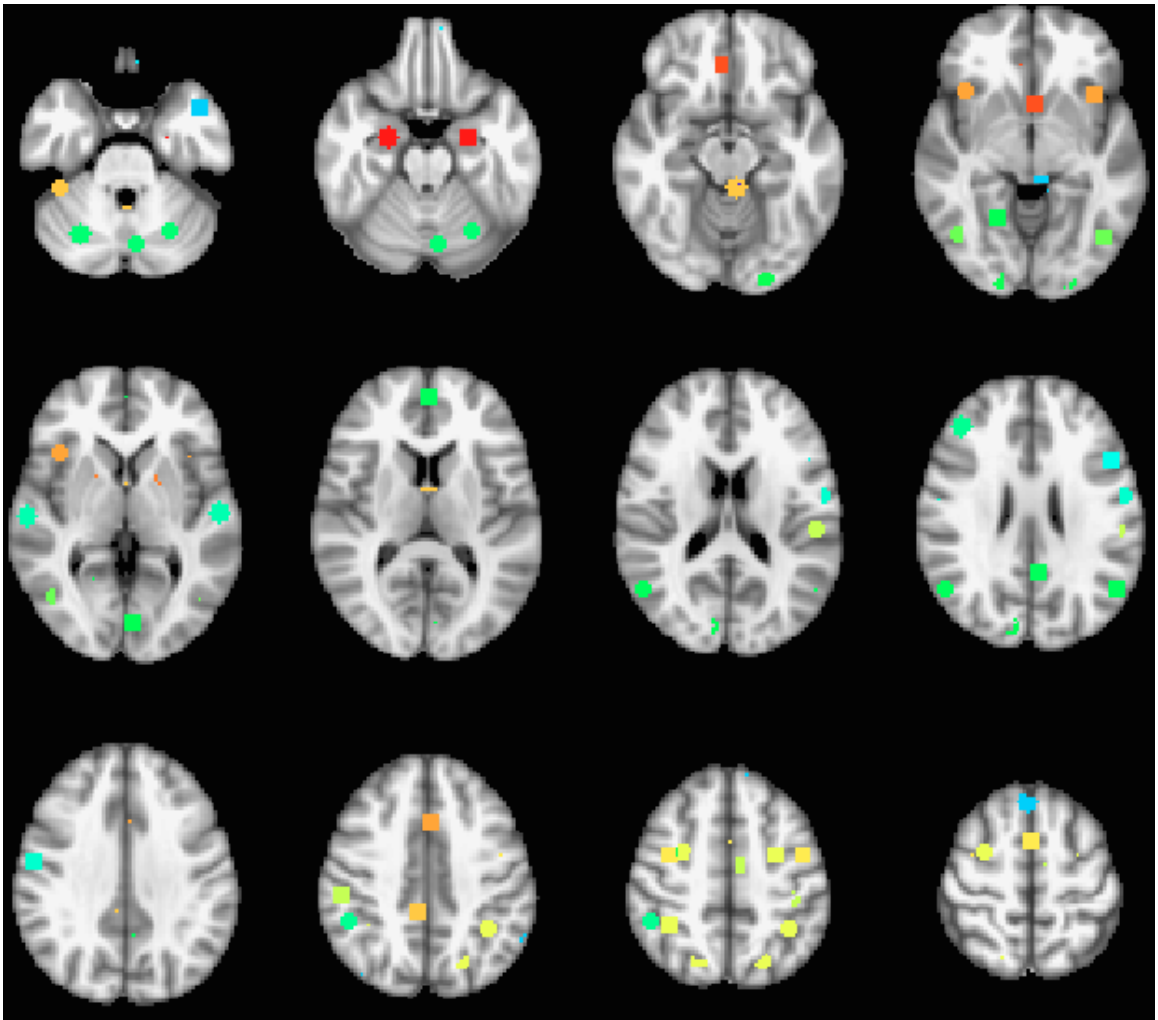

Figure 2.2: Visualization of ROI. Axial cross sections showing many of the ROI derived from the ICN in Laird et al. (2011).
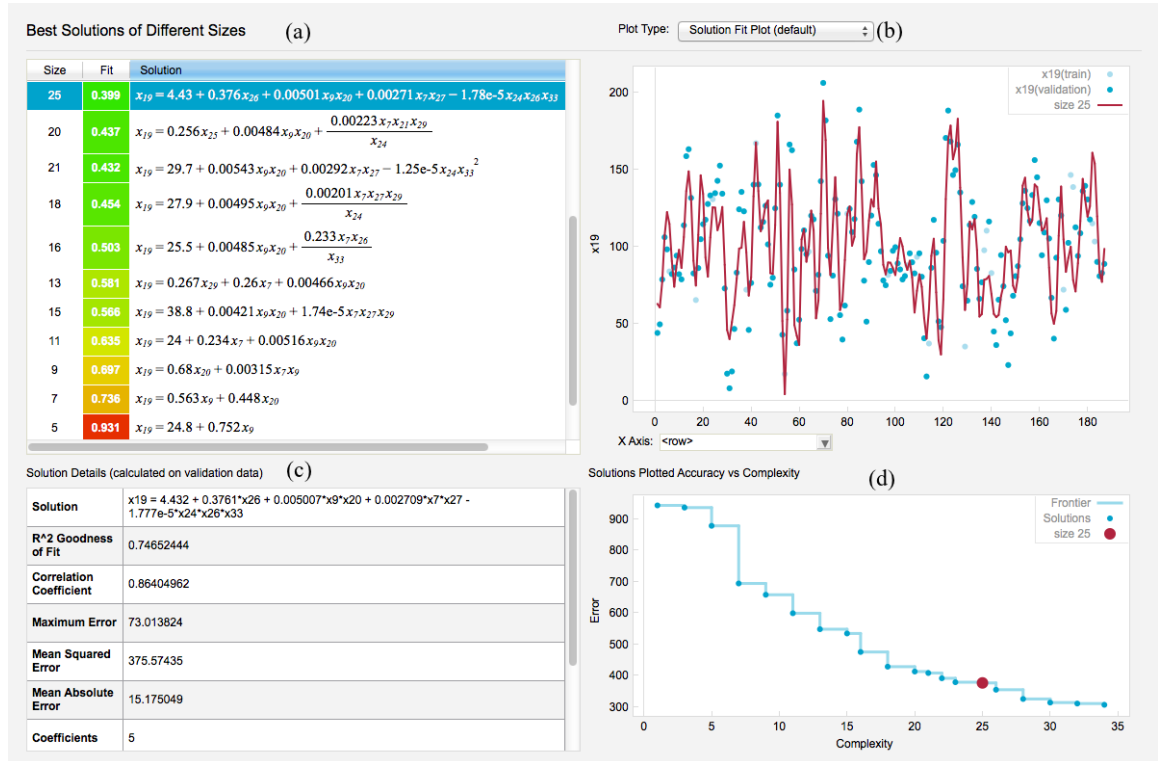
2.2.2 GENETIC PROGRAMMING

Figure 2.3: Screen shot of the GP package Eureqa during a search for models of the activity in ROI 19 in a single subject, as a function of activity in the other 51 regions. (a) The current set of models along the Pareto front of accuracy vs. parsimony, shown in (d) where each point represents a model and the red point represents the highlighted model. (b) Data from ROI 19 (points) over the 6-minute time series for this subject (x-axis in scans, roughly 2 seconds each). The highlighted model is shown in red, and statistics for this model's fit appear in (c).

GP is a biologically inspired, population-based machine learning algorithm. It is most commonly employed for symbolic regression: the algorithm searches for models explaining some quantity of interest (e.g., average BOLD signal from an ROI in the brain) as a function of some other possibly related observable quantities, statistics, or summary data (e.g., BOLD signals from other ROI). The algorithm proceeds by evolving the functional forms of a population of potential models, which are initially constructed at random from user-specified mathematical building blocks (available variables, arithmetic functions, parameter

constants, etc.). In brief, the models that better explain the data produce more offspring, leading to a gradual reduction of error within the population. We show a representative set of models produced by this approach in Figure 2.3(a). A key advantage of the technique is that no assumption (e.g., linearity) is imposed on the form of solutions, other than the choice of building blocks from which they can be made (we use arithmetic operations in the present work).

Typically, some measure of error (e.g., root mean square error) constitutes a model's explanatory fitness, and some measure of its size (e.g., number of operators, constants, and variables in the equation) represents its parsimony. The next generation of potential models is obtained by mutation (e.g., a single change of variable or operator) and recombination (i.e., swapping of function components between models) of the current set of *non-dominated* solutions: those models for which no simpler model in the population has less error. This set of non-dominated models is said to approach the ideal *Pareto front* of fitness versus parsimony as the population evolves. An important aspect of GP is that the result of a single search is this entire set of potential models, providing a trove of information for statistical analysis. Figure 2.3 is a screenshot of the off-the-shelf GP package Eureqa from Schmidt and Lipson (2009) performing a search (Eureqa version 0.97 Beta was used to generate the results reported in this study).

To apply GP to the fMRI data, for each of the 242 subjects we extract a single BOLD signal time series from each of the 52 selected ROI by averaging over the voxels within that ROI. Then the GP algorithm is run 52 times, one for each ROI, using all other ROI as potential explanatory variables. Note that the algorithm has no knowledge of the hypothesized networks from which these regions were chosen.

We describe the computational expense of the algorithm in terms of core-hours, i.e., the number of hours required for a single processor core to perform the necessary computation. Specifically, twelve core-hours of search were performed for each region, amounting to 624

core-hours per subject, and over 17 total core-years of computation were required for the population of 242 subjects. This substantial investment of resources yielded roughly 12 thousand Pareto fronts comprised of a quarter million models for statistical analysis.

The results of this analysis characterize the entire population of 242 subjects. Alternatively, results can be aggregated over phenotypic groups to produce group-level characterizations, or many GP searches can be run for a single individual to produce a subject-level characterization. We report results of population- and group-level analyses in Section 2.3, and discuss an example subject-level analysis in Appendix A.

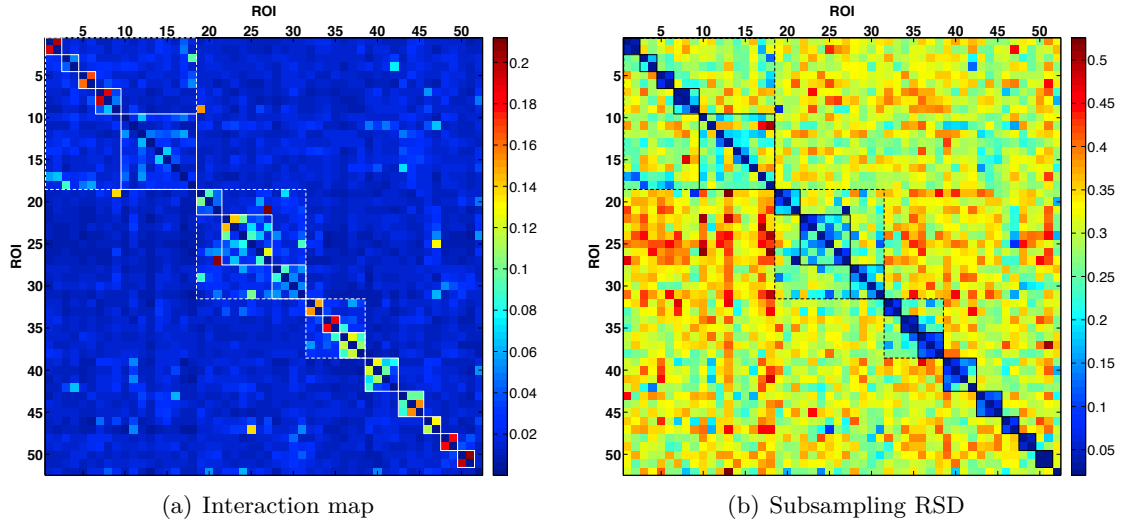(a) Interaction map   (b) Subsampling RSD

Figure 2.4: Functional interaction map. (a) Interaction map across all 242 subjects, and (b) map of RSD (relative standard deviation) of the interaction rates over 100 subsamples with 100 randomly selected subjects each. Solid outlines indicate ICN and dashed outlines indicate functional groupings of ICN from Laird et al. (2011).

2.2.3 ANALYSIS

The output of the GP algorithm poses a challenge for interpretation. Here we present a coarse statistical analysis of this rich mathematical characterization. For each ROI, we count the number of models for that ROI, across the Pareto fronts for all 242 subjects, that

have a particular (other) region on the right-hand side of the equation. We compute this count for each of the other 51 regions.

For example, consider the GP search for models of ROI 19 within a single subject illustrated in Figure 2.3. Upon completion, all 20 of the models along the Pareto front for this subject had at least one term containing ROI 9, and 17 models had terms containing ROI 20. In the subject pool as a whole, the total counts are 2990 and 1984, respectively. Specifically, of the roughly 5000 models for ROI 19 across all subjects, about 60% have terms containing ROI 9, and about 40% have terms containing ROI 20. Note that these frequencies are not properly normalized, because most models contain several ROI. Thus we normalize by the sum of the counts for all ROI. In the case of ROI 19, this sum is 22016.

The result is a vector for each ROI that describes, in a statistical sense, its relative dependence on each of the other regions. We interpret this vector as a distribution of likely interaction, and define the computed values to be relative *interaction rates* (IR). Note that both linear and nonlinear interactions, as well as weakly and strongly weighted basis functions, are counted equally. We form an interaction map by stacking these IR row vectors to visualize interaction across all 52 ROI, shown in Figure 2.4(a). The value in row 19 column 9, for example, is $2990/22016 \approx 0.1358$.

Note that the IR map is not symmetric by construction (though it appears nearly so), and indeed the value in row 9, column 19 is $0.148 \neq 0.1358$. We interpret a row of the IR map as a distribution of relative *dependence* of the corresponding ROI on each of the other regions. We interpret a column, on the other hand, as a measure of the *influence* of the corresponding ROI on each of the other regions. Taking the average of the IR map and its transpose produces a symmetric, *overall* IR map (not shown) that can be used in hierarchical analysis. We examine the interaction map, and provide results of hierarchical analysis, in the next section.

2.3 Results

Figure 2.4(a) shows the interaction map generated by the normalized frequency analysis of the NFM procedure, summarizing ROI interaction across all 242 subjects. To test the robustness of the computed interaction map, we form 100 random subsamples (with replacement) from the pool of 242 subjects, each with 100 subjects. For each sample, we perform the same counting procedure to produce the interaction map corresponding to that sample. A heat map of relative standard deviation (RSD) of IR over the 100 subsamples is shown in Figure 2.4(b).

The strong block-diagonal structure of the interaction map corresponds directly to the grouping of ROI into ICN. For example, regions 39-42, which form a partial block in the figure, are the four ROI that make up the default mode network (ICN 13) in Laird et al. (2011). Robustness (across subjects) of intra-network interaction is supported by the matching block-diagonal structure of low subsampling RSD (mean intra-network RSD < 20%), for all but ICN 2 (ROI 3-4) and ICN 5 (ROI 10-18). In addition to the strong primary block-diagonal structure, there is a secondary structure of lighter blocks that group ICN together. For example, regions 32-38 are composed of the strong blocks 32-33, 34-35, and 36-38 (corresponding to ICN 10, 11 and 12 respectively). There is a lighter block structure that suggests interaction among these three ICN which are, in fact, *together* responsible for visual processing. The secondary structure visible for regions 19-31 is comprised of ICN 6-8, which perform motor and visuospatial tasks. Each of these examples shows a matching secondary structure of moderate subsampling RSD (mean inter-network RSD < 30%), indicating fairly robust inter-network interaction as well.

2.3.1 HIERARCHICAL ANALYSIS

To further illustrate and clarify the hierarchical organization suggested by the interaction map, we generate the dendrogram in the top of Figure 2.5 by HCA (hierarchical cluster analysis, implemented in MATLAB with the nearest distance algorithm), using the reciprocal of the overall IR between each pair of ROI as the distance between them. For example, ROI 1 and 2 have an approximate overall IR of 0.2, and thus the distance between them is 5. We emphasize that the organization of ROI into networks, and clustering of those networks into functional groups described in Laird et al. (2011), are both captured by NFM. Some examples:

- The red group forms the visual cluster. ROI 32 and 33, the middle temporal gyri, form one network (ICN 10), while ROI 34 and 35, the lateral, and ROI 36-38, the medial posterior occipital cortices, (including V1, V2, and V3), form two other networks (ICN 11 and 12) from within the visual cluster.

- Regions 39-42 (the orange group) form ICN 13, the default mode network, and interact with ROI 4, the orbitofrontal cortex, from ICN 2.

- The green group to the far left includes all but one of the ROI from the motor and visuospatial complex. Interaction of this complex with the dorsal anterior cingulate cortex (dACC, ROI 9) and the network composed of ROI 46 and 47, thought to be responsible for multiple cognitive processes such as attention and inhibition, is indicated as well, suggesting that this interaction was common among many of the subjects.

- ICN 1 (ROI 1,2), 3 (ROI 5,6), the first two regions from ICN 4 (ROI 7-9), ICN 14 (ROI 43-45), 16 (ROI 48,49), and 17 (ROI 50,51) are also indicated.
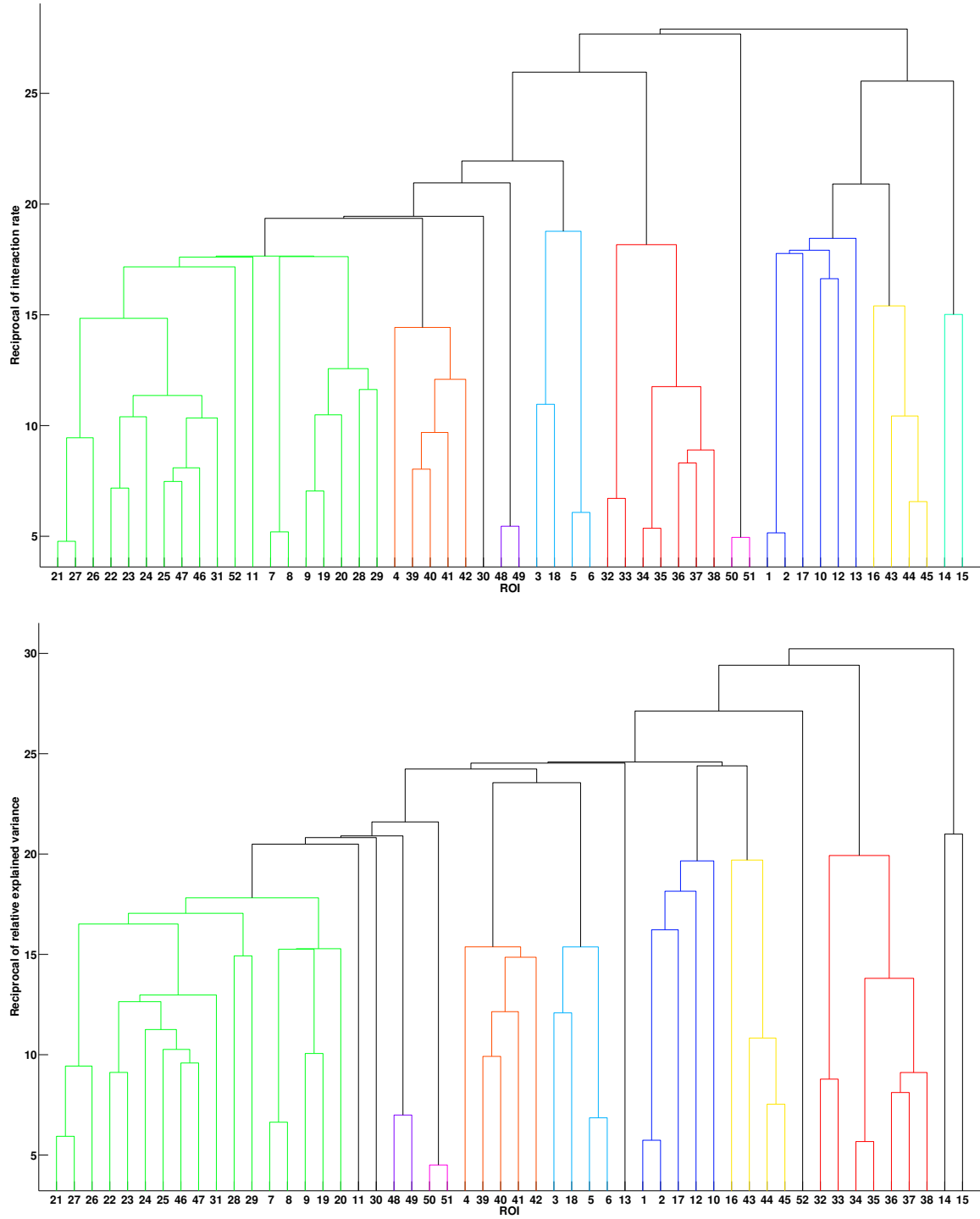
Figure 2.5: Hierarchical cluster analysis (HCA) of interaction among ROI, generated with NFM (top) and correlation analysis (bottom).

- Many of the regions from ICN 5 (ROI 10-18) interact with ICN 1 (ROI 1-2), and also form a loose interaction group with ICN 14 (ROI 43-45), the cerebellum, the most robust connection of which appears to be between ROI 16 and 43.

The robustness of each of the interactions discussed in this list is supported by low interaction rate subsampling RSD, shown in Figure 2.4(b).

2.3.2 IMPACT OF NONLINEARITY

In this section we demonstrate that the NFM procedure both captures the hierarchical structure of ROI interaction indicated by linear analyses, *and* reveals nonlinear interactions not discoverable by such methods. To accomplish this, we compare the population-level hierarchy generated by NFM with the results of an analogous linear procedure involving pairwise correlation analysis. Furthermore, we validate nonlinear relationships suggested by NFM in a setpwise multiple regression, the results of which we describe at the end of this section.

Comparison with correlation analysis

For each of the 242 subjects, we compute the correlation matrix for the 52 ROI time series. Squaring the elements of the correlation matrix and normalizing each row (after setting the diagonal to zero) provides the relative explained variance (relative R^2) of the ROI corresponding to that row by each of the other 51 ROI. The average of the 242 normalized subject matrices is interpreted as the linear version of the population-level IR map generated by NFM. As with IR, the reciprocal of relative explained variance can be considered a distance between ROI (higher relative R^2 means closer). The resulting hierarchy generated by HCA is shown in the bottom of Figure 2.5.

As expected, much of the large-scale structure revealed by NFM is also indicated by the linear correlation analysis. The similarity of the generated hierarchies supports the validity

of the models discovered by GP (i.e., the algorithm is not excessively overfitting the data), and the subtle differences between them suggest potentially interesting interactions that are missed if linearity is assumed. In the following we investigate one of these differences.

Interaction of the dACC (ROI 9) with the motor visuospatial complex is evident in both hierarchies. However, in the linear analysis it appears more closely connected with its own ICN (ROI 7,8), and only with the supplementary motor area (SMA, ROI 19) from ICN 6. In contrast, NFM reveals that activity in the dACC is related to more components of the motor system. The nonlinear models generated by GP show a strong connection between the dACC, the SMA, and the central sulcus (CS) of the primary motor cortex (ROI 29 from ICN 8) shown in red, green, and blue, respectively, in Figure 2.6. Specifically, about 20% of all of the models generated for the activity in the SMA, across all subjects and levels of complexity, contain *both* the dACC and the CS as explanatory variables.

For many of these models, the dACC and CS only show up as linear terms, so it is reasonable to wonder why the correlation analysis did not pick up this interaction. The vast majority of models containing dACC and CS in only linear terms *also* contain nonlinear terms in other ROI. It is the SMA *along with* these nonlinear terms that is correlated with the dACC and CS. Thus the interaction is hidden from linear analyses. Furthermore, many of the models *do* contain nonlinear terms involving the dACC and CS. In fact, the product of the activity in these two regions shows up in 78 models for the SMA across 21 different subjects, and it is always additive. This term is involved in models across the spectrum of complexity, including instances where it is the *only* term.

Validation of nonlinear terms by stepwise regression

To validate first order nonlinearity (pairwise product and quotient terms, as well as recipro-cals) suggested by NFM, we first randomly assign 100 subjects to a training group, and 100 different subjects to a testing group. NFM results are aggregated over the training group
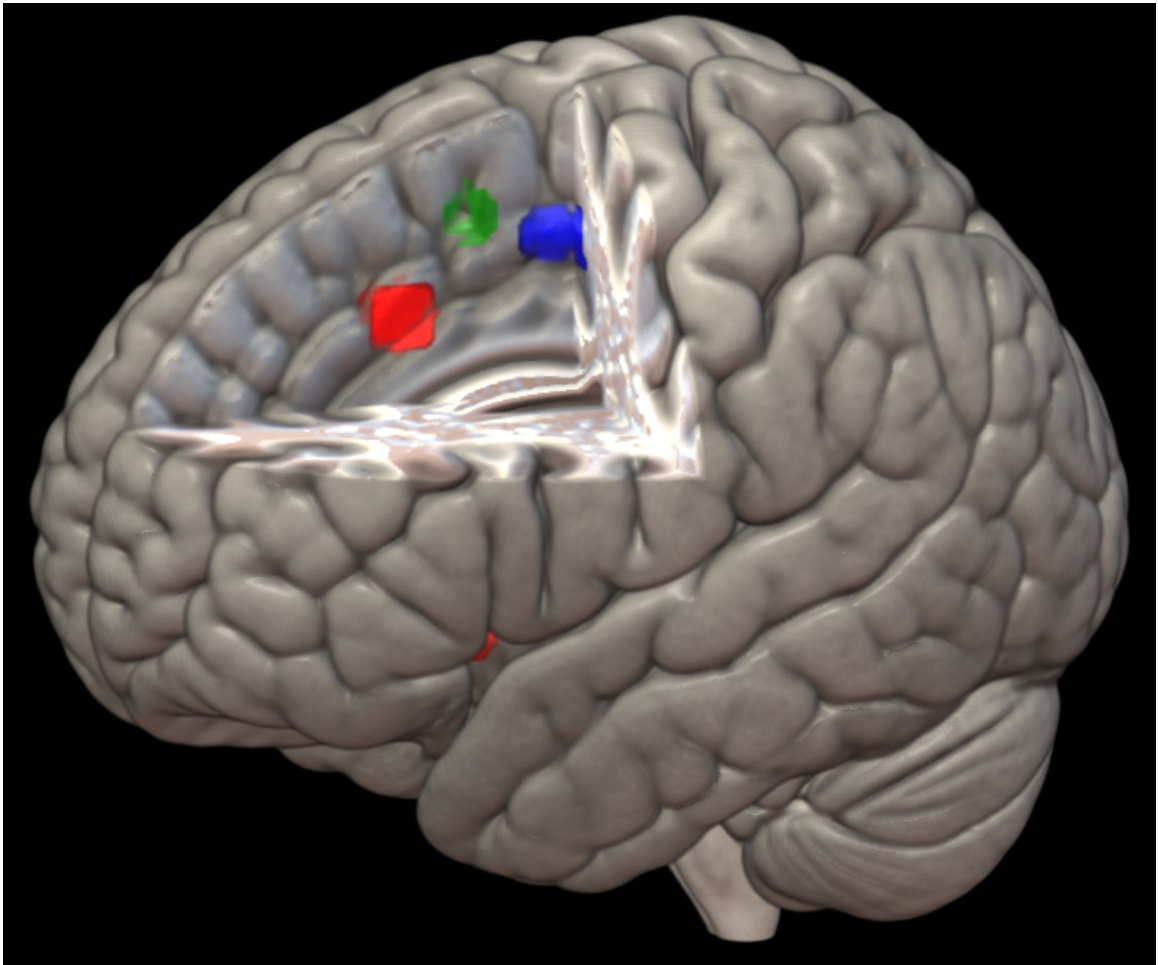
Figure 2.6:   NFM reveals a nonlinear interaction among these three ROI: dorsal anterior cingulate cortex (red), supplementary motor area (green), and central sulcus in the primary motor cortex (blue).

to produce an IR map and hierarchy (not shown) summarizing ROI interaction within the training group as a whole. The roughly 2000 specific models generated by NFM for each region (approximately 20 models per training subject) are then used to inform the modeling of ROI activity in that region within the testing group, by stepwise regression.

For each ROI and each testing subject, we first perform a standard stepwise linear regression using the other 51 ROI as regressors. We then perform a stepwise *nonlinear* regression by including all first order nonlinear terms suggested by NFM over the training group in addition to the 51 linear regressors. Statistics of the linear and nonlinear models are compared to determine the effect of including these first order terms. To illustrate, we describe results of the validation procedure for the SMA (ROI 19) here.

We show a histogram of increase in the percentage of explained variance for the nonlinear versus linear regression models for the SMA in Figure 2.7. The inclusion of first order nonlinear terms suggested by NFM over the training group increases the percentage of explained variance for *every* test subject, with a mean increase of 12.5% and maximum increase of 42%. The nonlinear models contain more terms (mean 46, compared with mean 19 for linear models), so a potential concern is that the increase in R^2 might simply be a result of the additional degrees of freedom. However, for each test subject the nonlinear model F-statistic is also greater than that of the linear model (mean increase of 83.5, maximum increase of 1000), and comparisons of adjusted R^2, which account for differences in degrees of freedom, show only slightly smaller increases for all test subjects. This suggests that the increase in explained variance is due to explanatory power of the nonlinear terms, and not simply the additional degrees of freedom in the nonlinear models.

2.3.3 GROUP-LEVEL VARIATION

Variation among individuals (illustrated in Appendix A) suggests that statistics of interaction rates among ROI may differ between phenotypic groups. The hierarchical organization
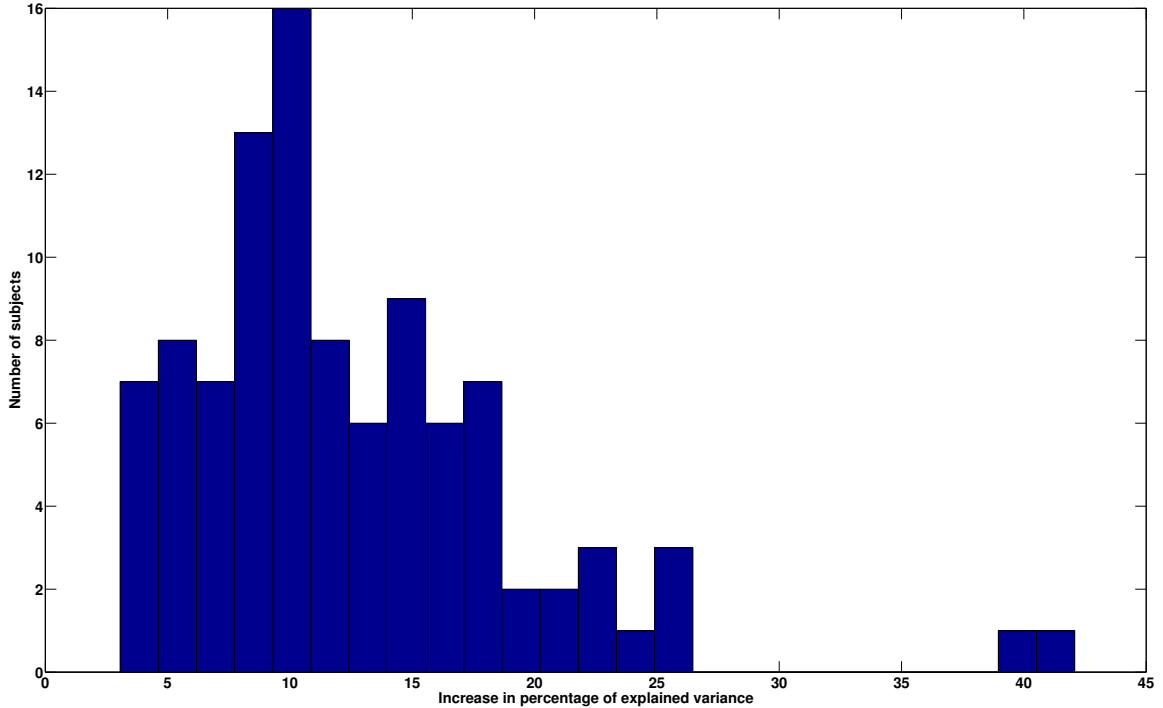
Figure 2.7: Histogram of increase in explained variance. The inclusion of first order nonlinear terms suggested by NFM over the training group, in a stepwise regression analysis for the SMA in the testing group, increases the percentage of explained variance for every test subject, with a mean increase of 12.5% and maximum increase of 42%.

of ROI induced by IR might illuminate, in such cases, variation in functional dynamics associated with demographic, behavioral, or genetic characteristics. An example illustrating this potential is provided by the contrast between drinking (D) and non-drinking (ND) adolescents from the IMAGEN dataset. In Figure 2.8, we show hierarchies for the top and bottom 100 subjects in terms of lifetime drinking score, determined by self-report questionnaire, corresponding to those who have had 2 or more lifetime drinks, and those who have had 1 or fewer, respectively. The two hierarchies are similar to one another (and comparable to the population level hierarchy), but subtle differences between them suggest group-differentiating factors.

- The ROI pair 3,18, the subgenual ACC and anterior nuclei of the thalamus, respectively, are coupled in both the D and ND groups. However, their arrangement in the
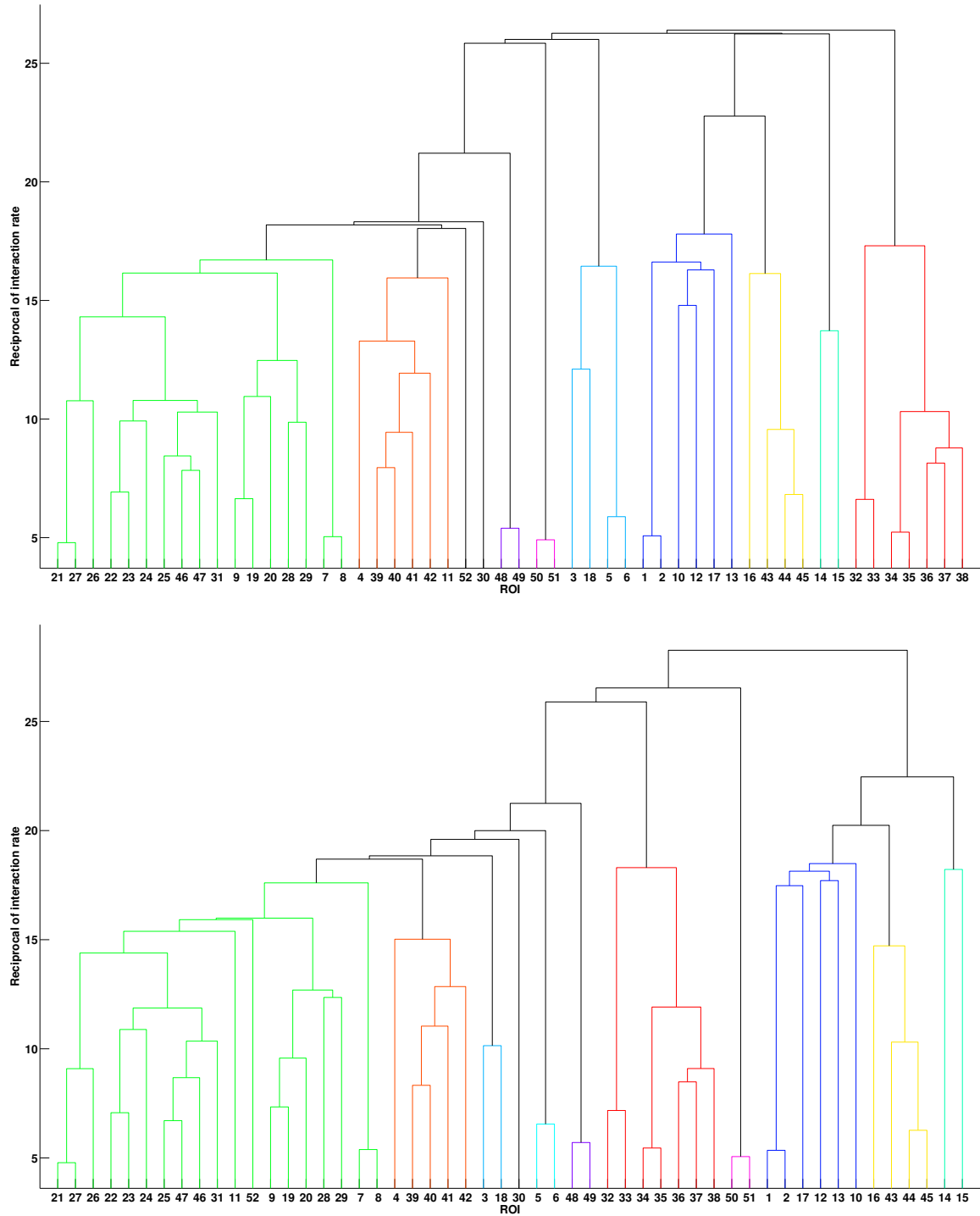
Figure 2.8: Hierarchies for groups with high (top) and low (bottom) alcohol consumption rates, defined by two or more lifetime drinks and one or fewer lifetime drinks, respectively.

hierarchies is different, as we'll describe in a moment, resulting from the following two distinguishing interaction rates.

- For the ND group, there is a 22% lower IR between ROI 6, the left putamen, and the anterior nuclei of the thalamus, ROI 18. We note that this reduced interaction is completely missed by pairwise correlation analysis, (which indicates a slightly reduced interaction among *drinkers*, see Appendix B), and thus appears to be an entirely nonlinear effect.

- In contrast, there is a 33% higher intra-network IR within ICN 2, comprised of ROI 3-4, the subgenual ACC and the orbitofrontal cortex, respectively, among non-drinkers. Though this difference is also indicated by correlation analysis, only about half of the effect is captured (a 16% elevation).

- These two differences in interaction cooperate to shuffle the hierarchical arrangement of ROI in the D versus ND group. The subgenual ACC and anterior nuclei of the thalamus are most closely associated with the default mode network in non-drinkers, through the orbitofrontal cortex. Among drinkers, in contrast, they are grouped directly with the bilateral basal ganglia of ICN 3 (ROI 5-6). In other words, in drinkers there is a tighter coupling among ROI most strongly linked to reward and thirst tasks as reported in Laird et al. (2011). The relevant ROI are shown in Figure 2.9.

- The largest single difference between the D and ND groups is a 74% elevated IR between the right angular gyrus (ROI 41 in the default mode network) and ROI 11, the dorsal posterior cingulate cortex, among drinkers. These ROI are shown in red and green, respectively, in Figure 2.10. About half of this effect is captured by correlation analysis.
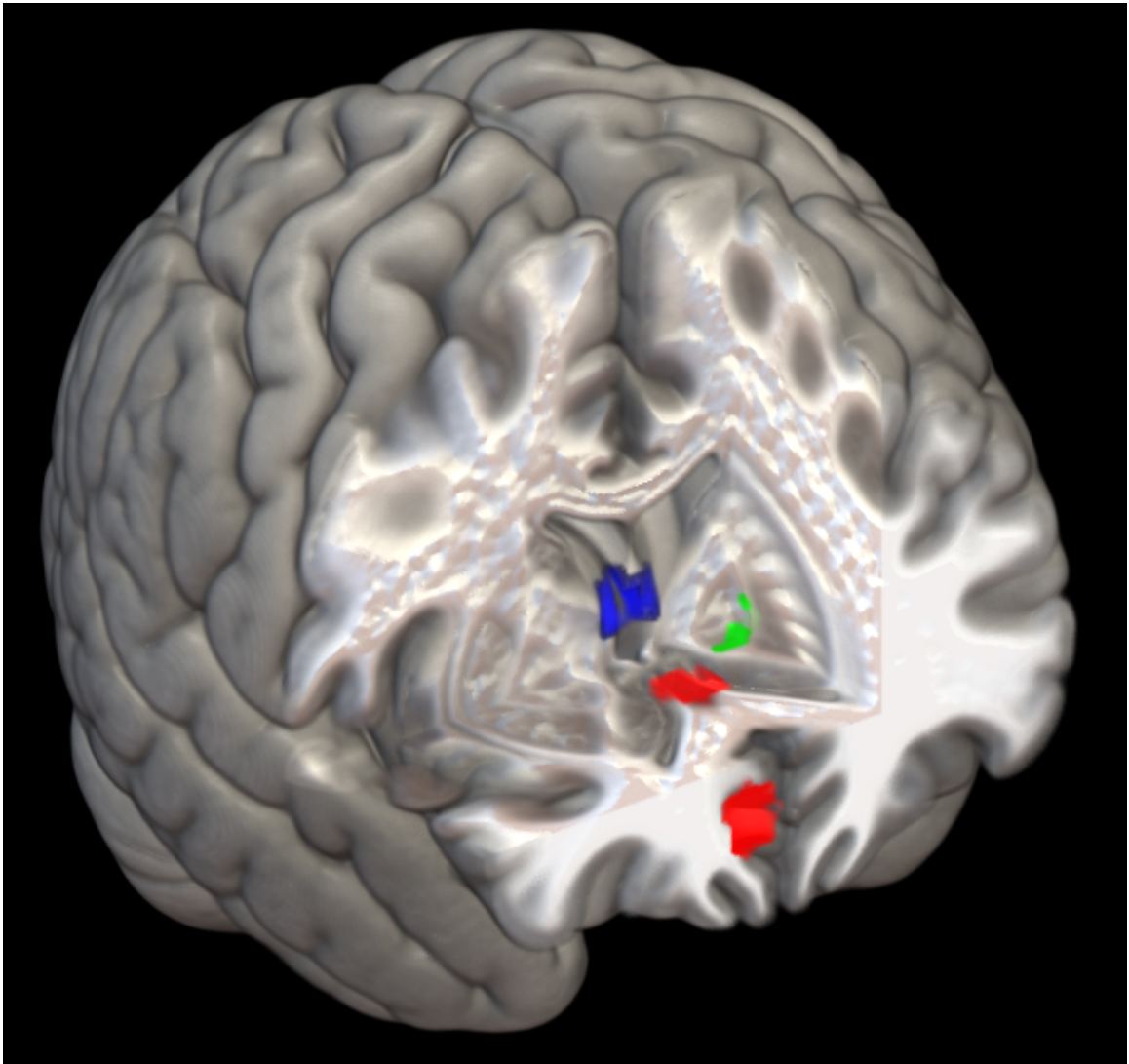
Figure 2.9:   *Interaction between the subgenual ACC (top red) and orbitofrontal cortex (bottom red) is lower among drinkers, who also show elevated interaction between the left putamen (green) and anterior nuclei of the thalamus (blue), an apparently nonlinear effect.*
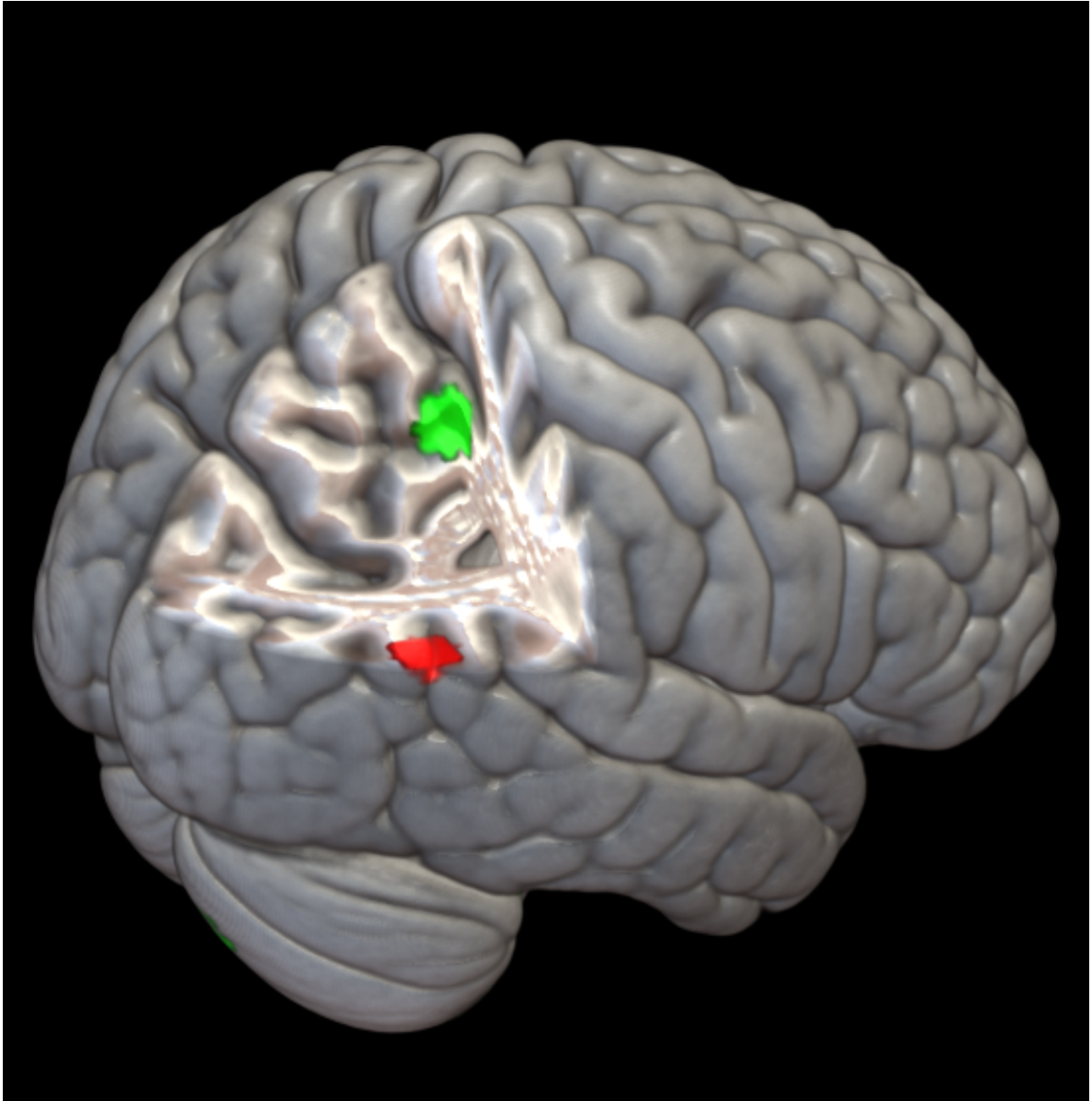
Figure 2.10: Interaction between the right angular gyrus from the default mode network (red), and dorsal posterior cingulate cortex (green) is 74% higher among drinkers.

2.4 DISCUSSION

First we note that results presented in this work are intended to demonstrate the applicability of NFM. Validation over independent data has not yet been completed, and therefore the results themselves should be considered preliminary. Nonetheless, the reproduction of ICN, and their hierarchical organization into functional groups, using an entirely different approach than that described in Laird et al. (2011), provides strong evidence for the analytical potential of the technique.

Furthermore, NFM reveals nonlinear interactions that are not discoverable with standard linear techniques, or without prior hypotheses. Such relationships could provide a new window into brain function, and results we present here suggest that that window may provide a better view of the connection between brain function and behavior. This highlights the potential of the methodology as a *hypothesis generator*. Of course proper care must be taken (with regard to independence of observations, etc.) in the ensuing investigations of such data-driven hypotheses. Nonetheless, hypothesis generation is a powerful tool for scientific exploration, and has been used recently to inform biomedical research, such as in Abedi et al. (2012) and Spangler et al. (2014).

In addition to providing insight on its own, the NFM procedure complements other modes of analysis. A potentially promising extension, especially for a hybrid version capable of voxel level analysis (discussed in Appendix C), would be to use it in conjunction with graph-theoretic analyses such as those described in Bassett and Bullmore (2006), Stam and Reijneveld (2007), and van den Heuvel et al. (2008). The general technique, as detailed in Bullmore and Sporns (2009) and Rubinov and Sporns (2010), is to compute pairwise correlations among all voxels, set a threshold above which two voxels are considered connected, and calculate various network summary measures (e.g., degree distribution, assortativity,

diameter, etc.). By simply replacing correlations in these networks with interaction rates determined by NFM, the assumption of linearity is left behind.

Finally, it is important to note that the specific forms of the models in the output of the GP algorithm have been analyzed simplistically in the present work. A major potential benefit of NFM is the insight that might be gained from precisely analyzing these mathematical descriptions of the relationships among ROI in the brain. Of course, ascribing meaning to any particular one of these models would have to be done cautiously. However, given the results we describe here, obtained by a coarse treatment, the *collection* of models determined by GP may offer a number of as yet undiscovered insights. This seems a potentially fruitful avenue for future theoretical research.

2.5 Conclusions

Results produced in our study suggest that there is potential analytical power in the use of NFM, or some modification thereof, in the neuroimaging domain. The procedure we investigated here utilizes commercially available, out-of-the-box GP software, and preliminary statistical analysis of its output. Many improvements and extensions are possible, only some of which we have suggested in this work. Reproduction of recent results constitutes a measure of cross-validation, and the preliminary results presented demonstrate the unique capability of NFM to discover nonlinear relationships among regions of the brain that hold promise for illuminating differences in brain function between subject groups. Further, the mathematical characterizations we have achieved, which are not limited by linear or univariate assumptions, are ripe for future investigation.

Acknowledgements

References

Abedi, V., Zand, R., Yeasin, M., Faisal, F. E., 2012. An automated framework for hypotheses generation using literature. BioData Min 5 (1), 13.

Bassett, D. S., Bullmore, E., Dec 2006. Small-world brain networks. Neuroscientist 12 (6), 512–523.

Birn, R. M., Murphy, K., Handwerker, D. A., Bandettini, P. A., 2009. fmri in the presence of task-correlated breathing variations. NeuroImage 47 (3), 1092 – 1104, brain Body Medicine.

Bullmore, E., Sporns, O., 03 2009. Complex brain networks: graph theoretical analysis of structural and functional systems. Nat Rev Neurosci 10 (3), 186–198.
URL http://dx.doi.org/10.1038/nrn2575

Calhoun, V., Adali, T., Pearlson, G., Pekar, J., 2001. A method for making group inferences from functional mri data using independent component analysis. Human Brain Mapping 14 (3), 140–151.
URL http://dx.doi.org/10.1002/hbm.1048

Friston, K., Harrison, L., Penny, W., 2003. Dynamic causal modelling. NeuroImage 19 (4), 1273 – 1302.

Hui, K. K., Liu, J., Makris, N., Gollub, R. L., Chen, A. J., I. Moore, C., Kennedy, D. N., Rosen, B. R., Kwong, K. K., 2000. Acupuncture modulates the limbic system and subcor-

tical gray structures of the human brain: Evidence from fmri studies in normal subjects. Human Brain Mapping 9 (1), 13–25.

Hyvärinen, A., Pajunen, P., 1999. Nonlinear independent component analysis: Existence and uniqueness results. Neural Networks 12 (3), 429 – 439.

Ide, J. S., Shenoy, P., Yu, A. J., Li, C.-s. R., 2013. Bayesian prediction and evaluation in the anterior cingulate cortex. The Journal of Neuroscience 33 (5), 2039–2047.
URL http://www.jneurosci.org/content/33/5/2039.abstract

Kanwisher, N., Stanley, D., Harris, A., 1999. The fusiform face area is selective for faces not animals. NeuroReport 10 (1).

Kruggel, F., Zysset, S., von Cramon, D., 2000. Nonlinear regression of functional {MRI} data: An item recognition task study. NeuroImage 12 (2), 173 – 183.

Laird, A. R., Fox, P. M., Eickhoff, S. B., Turner, J. A., Ray, K. L., McKay, D. R., Glahn, D. C., Beckmann, C. F., Smith, S. M., Fox, P. T., 2011. Behavioral interpretations of intrinsic connectivity networks. J. Cognitive Neuroscience, 4022–4037.

Liu, P., Zhou, G., Zhang, Y., Dong, M., Qin, W., Yuan, K., Sun, J., Liu, J., Liang, J., von Deneen, K. M., Liu, Y., Tian, J., 2010. The hybrid glmĐica investigation on the neural mechanism of acupoint st36: An fmri study. Neuroscience Letters 479 (3), 267 – 271.

McKeown, M. J., Sejnowski, T. J., 1998. Independent component analysis of fmri data: Examining the assumptions. Human Brain Mapping 6 (5-6), 368–372.

Mueller, W. M., Yetkin, F. Z., Hammeke, T. A., Morris, G. L. I., Swanson, S. J., Reichert, K., Cox, R., Haughton, V. M., 1996. Functional magnetic resonance imaging mapping of the motor cortex in patients with cerebral tumors. Neurosurgery 39 (3).

Mukamel, R., Gelbard, H., Arieli, A., Hasson, U., Fried, I., Malach, R., 2005. Coupling between neuronal firing, field potentials, and fmri in human auditory cortex. Science 309 (5736), 951–954.
URL http://www.sciencemag.org/content/309/5736/951.abstract

Murphy, K., Harris, A. D., Wise, R. G., 2011. Robustly measuring vascular reactivity differences with breath-hold: Normalising stimulus-evoked and resting state {BOLD} fmri data. NeuroImage 54 (1), 369 – 379.

Nolde, S. F., Johnson, M. K., D'Esposito, M., 1998. Left prefrontal activation during episodic remembering: an event?related fmri study. NeuroReport 9 (15).

Rio, D. E., Rawlings, R. R., Woltz, L. A., Gilman, J., Hommer, D. W., 2013. Development of the complex general linear model in the fourier domain: Application to fmri multiple input-output evoked responses for single subjects. Computational and Mathematical Methods in Medicine 2013, 16.
URL http://dx.doi.org/10.1155/2013/645043

Rubinov, M., Sporns, O., 2010. Complex network measures of brain connectivity: Uses and interpretations. NeuroImage 52 (3), 1059 – 1069, computational Models of the Brain.

Schmidt, M., Lipson, H., 2009. Distilling free-form natural laws from experimental data. Science 324 (5923), 81–85.
URL http://www.sciencemag.org/content/324/5923/81.abstract

Smith, S. M., Fox, P. T., Miller, K. L., Glahn, D. C., Fox, P. M., Mackay, C. E., Filippini, N., Watkins, K. E., Toro, R., Laird, A. R., Beckmann, C. F., 2009. Correspondence of the brain's functional architecture during activation and rest. Proceedings of the National Academy of Sciences 106 (31), 13040–13045.
URL http://www.pnas.org/content/106/31/13040.abstract

Spangler, S., Wilkins, A. D., Bachman, B. J., Nagarajan, M., Dayaram, T., Haas, P., Regenbogen, S., Pickering, C. R., Comer, A., Myers, J. N., Stanoi, I., Kato, L., Lelescu, A., Labrie, J. J., Parikh, N., Lisewski, A. M., Donehower, L., Chen, Y., Lichtarge, O., 2014. Automated hypothesis generation based on mining scientific literature. In: Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. KDD '14. ACM, New York, NY, USA, pp. 1877–1886.
URL http://doi.acm.org/10.1145/2623330.2623667

Stam, C. J., Reijneveld, J. C., 2007. Graph theoretical analysis of complex networks in the brain. Nonlinear Biomed Phys 1 (1), 3.

van den Heuvel, M. P., Stam, C. J., Boersma, M., Hulshoff Pol, H. E., Nov 2008. Small-world and scale-free organization of voxel-based resting-state functional connectivity in the human brain. Neuroimage 43 (3), 528–539.

Whelan, R., Conrod, P. J., Poline, J.-B., Lourdusamy, A., Banaschewski, T., Barker, G. J., Bellgrove, M. A., Buchel, C., Byrne, M., Cummins, T. D. R., Fauth-Buhler, M., Flor, H., Gallinat, J., Heinz, A., Ittermann, B., Mann, K., Martinot, J.-L., Lalor, E. C., Lathrop, M., Loth, E., Nees, F., Paus, T., Rietschel, M., Smolka, M. N., Spanagel, R., Stephens, D. N., Struve, M., Thyreau, B., Vollstaedt-Klein, S., Robbins, T. W., Schumann, G., Garavan, H., 06 2012. Adolescent impulsivity phenotypes characterized by distinct brain networks. Nat Neurosci 15 (6), 920–925.
URL http://dx.doi.org/10.1038/nn.3092

Whelan, R., Watts, R., Orr, C. A., Althoff, R. R., Artiges, E., Banaschewski, T., Barker, G. J., Bokde, A. L. W., Buchel, C., Carvalho, F. M., Conrod, P. J., Flor, H., Fauth-Buhler, M., Frouin, V., Gallinat, J., Gan, G., Gowland, P., Heinz, A., Ittermann, B., Lawrence, C., Mann, K., Martinot, J.-L., Nees, F., Ortiz, N., Paillere-Martinot, M.-L., Paus, T., Pausova, Z., Rietschel, M., Robbins, T. W., Smolka, M. N., Strohle, A., Schumann, G., Garavan, H., the IMAGEN Consortium, 08 2014. Neuropsychosocial profiles of current and future adolescent alcohol misusers. Nature 512 (7513), 185–189. URL http://dx.doi.org/10.1038/nature13402

Chapter 3

Conclusion

In this dissertation we introduced a novel methodology, called nonlinear functional mapping (NFM), for deriving a nonlinear mathematical characterization of the human brain functional interactome directly from fMRI data. The models generated by NFM are interpretable, allowing for the extraction of insight at the group, subject, and ROI levels, and proposed extensions would provide the capability to analyze high-dimensional, multi-domain data. No other method possess all of these capabilities. In addition to its stand-alone analytical potential, NFM, or some variant thereof, would complement many other methodologies, granting them the capability of incorporating nonlinearity.

The technique was demonstrated in the analysis of resting-state fMRI recorded from 242 adolescents. The preliminary results are both consistent with recent independent work, and suggest previously unknown nonlinear interactions, which are not discoverable by other methods, that correlate with adolescent alcohol consumption. Discovery of such relationships could provide a new window into brain function, and results presented here suggest that that window may provide a better view of the connection between brain function and behavior. This highlights the potential of the methodology as a *hypothesis generator*. We note that the specific forms of the models generated by NFM have been analyzed simplistically in this demonstration. A major benefit of the technique is the insight that might be

gained from precisely analyzing these mathematical descriptions of the relationships among ROI in the brain, and this is one of many potentially fruitful avenues for future research.

In addition to the precise analysis of the mathematical characterization derived by NFM, there are a number of other directions in which the technique can be immediately extended. The first and most obvious is to investigate associations between the resting state functional interactome and other behavioral (e.g., tobacco, cannabis use), or phenotypic (e.g., gender, verbal/performance IQ) groups. Differences in effective connectivity (functional interaction) may also be associated with neurological disorder, potentially granting insight into the specific nature of particular neural disfunction (e.g., schizophrenia). The capability of NFM to detect nonlinear effects *and* to derive interpretable models of those effects should enhance the potential for uncovering such associations.

A related direction of future research would be to investigate to what extent differences in functional interactome can predict *future* behavior, or onset of neurological disorder. It may be that certain characteristics of resting state effective connectivity are endophenotypic risk factors for various forms of addiction or neural disfunction. There is a tremendous potential benefit for diagnosis and treatment/intervention if such risk factors can be determined. It has already been shown, in Whelan et al. (2014), that there are neural correlates for future alcohol misuse, providing a basis for the notion that brain function can indeed be predictive of behavior. These neural correlates were not, however, determined from resting state fMRI, but from brain activity while subjects were performing certain tasks. The NFM procedure as described in this dissertation is directly applicable to fMRI time series data acquired while subjects are performing tasks as well, and this is another natural extension of the methodology.

In the effort to predict undesirable behavior or neurological disorder, it is almost certainly the case that risk factors exist in many domains (personal/family history, personality, genetics, brain function), and also likely that the influence of these cross-domain factors *are*

not independent, and may, in fact, interact (potentially nonlinearly) in a subject's risk profile. An aspect of the core of NFM, symbolic regression by GP, that was not explored at all in the present work is the capability to directly incorporate data from multiple domains. A promising alternative application of symbolic regression by GP would be to model risk (i.e., predictively classify) based on such multi-domain data.

In addition to determining associations between multi-domain risk factors and behavior or disease, recent work (e.g., Ramsey et al. (2010, 2011)) suggests the possibility of determining *causal* relationships among risk factors that could inform treatment and intervention strategies. Considering recent success using deep brain stimulation for the treatment of various neurological disorders (e.g., Parkinson's and OCD), there is clearly potential for direct intervention upon the discovery of causal mechanisms at play in neurological disfunction. Behavioral and pharmacological therapies might also be informed by such causal mechanisms. Furthermore, the algorithms employed in Ramsey et al. (2010, 2011) to determine causal Bayesian network models might themselves be enhanced by the incorporation of symbolic regression by GP (instead of linear regression).

Finally, in the NFM methodology presented here, interactions among ROI are effectively aggregated over time, constituting a tacit assumption that the structure of functional interaction is static. This is, of course, highly unlikely, even within a single subject. An investigation of the dynamic nature of effective connectivity could be accomplished by a windowing approach, whereby multiple interaction maps are derived by NFM for successive subsets of the fMRI time series. Changes in the resulting hierarchy from window to window might illuminate the ways in which ROI interaction shifts over time, and this could be particularly enlightening in the analysis of task- or stimulus-driven interaction.

Bibliography

Abedi, V., Zand, R., Yeasin, M., Faisal, F. E., 2012. An automated framework for hypotheses generation using literature. BioData Min 5 (1), 13.

Anticevic, A., Cole, M. W., Murray, J. D., Corlett, P. R., Wang, X.-J., Krystal, J. H., 2014/09/16 2012. The role of default network deactivation in cognition and disease. Trends in Cognitive Sciences 16 (12), 584–592.
URL http://www.cell.com/trends/cognitive-sciences/abstract/S1364-6613(12)00244-6

Bassett, D. S., Bullmore, E., Dec 2006. Small-world brain networks. Neuroscientist 12 (6), 512–523.

Birn, R. M., Murphy, K., Handwerker, D. A., Bandettini, P. A., 2009. fmri in the presence of task-correlated breathing variations. NeuroImage 47 (3), 1092 – 1104, brain Body Medicine.

Bullmore, E., Sporns, O., 03 2009. Complex brain networks: graph theoretical analysis of structural and functional systems. Nat Rev Neurosci 10 (3), 186–198.
URL http://dx.doi.org/10.1038/nrn2575

Bullmore, E., Sporns, O., 05 2012. The economy of brain network organization. Nat Rev Neurosci 13 (5), 336–349.
URL http://dx.doi.org/10.1038/nrn3214

Calhoun, V., Adali, T., Pearlson, G., Pekar, J., 2001. A method for making group inferences from functional mri data using independent component analysis. Human Brain Mapping 14 (3), 140–151.
URL http://dx.doi.org/10.1002/hbm.1048

Friston, K., Harrison, L., Penny, W., 2003. Dynamic causal modelling. NeuroImage 19 (4), 1273 – 1302.

Gläscher, J. P., O'Doherty, J. P., 2010. Model-based approaches to neuroimaging: combining reinforcement learning theory with fmri data. Wiley Interdisciplinary Reviews: Cognitive

Science 1 (4), 501–510.
URL http://dx.doi.org/10.1002/wcs.57

Grosenick, L., Klingenberg, B., Katovich, K., Knutson, B., Taylor, J. E., 2013. Interpretable whole-brain prediction analysis with graphnet. NeuroImage 72 (0), 304 – 321.
URL http://www.sciencedirect.com/science/article/pii/S1053811912012487

Haimovici, A., Tagliazucchi, E., Balenzuela, P., Chialvo, D. R., 2013. Brain organization into resting state networks emerges at criticality on a model of the human connectome. Physical review letters 110 (17), 178101.

Harding, S., Banzhaf, W., 2007. Fast genetic programming on gpus. In: Ebner, M., OÕNeill, M., EkǦrt, A., Vanneschi, L., Esparcia-AlcǦzar, A. (Eds.), Genetic Programming. Vol. 4445 of Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 90–101.

Hui, K. K., Liu, J., Makris, N., Gollub, R. L., Chen, A. J., I. Moore, C., Kennedy, D. N., Rosen, B. R., Kwong, K. K., 2000. Acupuncture modulates the limbic system and subcortical gray structures of the human brain: Evidence from fmri studies in normal subjects. Human Brain Mapping 9 (1), 13–25.

Hyvärinen, A., Pajunen, P., 1999. Nonlinear independent component analysis: Existence and uniqueness results. Neural Networks 12 (3), 429 – 439.

Icke, I., Allgaier, N., Danforth, C. M., Whelan, R. A., Garavan, H. P., Bongard, J. C., 2014. A deterministic and symbolic regression hybrid applied to resting-state fmri data. In: Riolo, R., Moore, J., Kotanchek, M. (Eds.), Genetic Programming Theory and Practice XI. Springer, Ch. 9, pp. 155–173.

Ide, J. S., Shenoy, P., Yu, A. J., Li, C.-s. R., 2013. Bayesian prediction and evaluation in the anterior cingulate cortex. The Journal of Neuroscience 33 (5), 2039–2047.
URL http://www.jneurosci.org/content/33/5/2039.abstract

Kanwisher, N., Stanley, D., Harris, A., 1999. The fusiform face area is selective for faces not animals. NeuroReport 10 (1).

Kruggel, F., Zysset, S., von Cramon, D., 2000. Nonlinear regression of functional {MRI} data: An item recognition task study. NeuroImage 12 (2), 173 – 183.

Laird, A. R., Fox, P. M., Eickhoff, S. B., Turner, J. A., Ray, K. L., McKay, D. R., Glahn, D. C., Beckmann, C. F., Smith, S. M., Fox, P. T., 2011. Behavioral interpretations of intrinsic connectivity networks. J. Cognitive Neuroscience, 4022–4037.

BIBLIOGRAPHY

Liu, P., Zhou, G., Zhang, Y., Dong, M., Qin, W., Yuan, K., Sun, J., Liu, J., Liang, J., von
Deneen, K. M., Liu, Y., Tian, J., 2010. The hybrid glmÐica investigation on the neural
mechanism of acupoint st36: An fmri study. Neuroscience Letters 479 (3), 267 – 271.

McConaghy, T., 2011. Ffx: fast, scalable, deterministic symbolic regression technology. In:
Riolo, R., Vladislavleva, E., Moore, J. (Eds.), Genetic Programming Theory and Practice
IX. Springer, Ch. 13.

McKeown, M. J., Sejnowski, T. J., 1998. Independent component analysis of fmri data:
Examining the assumptions. Human Brain Mapping 6 (5-6), 368–372.

Mueller, W. M., Yetkin, F. Z., Hammeke, T. A., Morris, G. L. I., Swanson, S. J., Reichert,
K., Cox, R., Haughton, V. M., 1996. Functional magnetic resonance imaging mapping of
the motor cortex in patients with cerebral tumors. Neurosurgery 39 (3).

Mukamel, R., Gelbard, H., Arieli, A., Hasson, U., Fried, I., Malach, R., 2005. Coupling
between neuronal firing, field potentials, and fmri in human auditory cortex. Science
309 (5736), 951–954.
URL http://www.sciencemag.org/content/309/5736/951.abstract

Murphy, K., Harris, A. D., Wise, R. G., 2011. Robustly measuring vascular reactivity
differences with breath-hold: Normalising stimulus-evoked and resting state {BOLD}
fmri data. NeuroImage 54 (1), 369 – 379.

Nolde, S. F., Johnson, M. K., D'Esposito, M., 1998. Left prefrontal activation during
episodic remembering: an event?related fmri study. NeuroReport 9 (15).

O'Doherty, J. P., Dec 2004. Reward representations and reward-related learning in the
human brain: insights from neuroimaging. Curr Opin Neurobiol 14 (6), 769–776.

Ramsey, J. D., Hanson, S. J., Glymour, C., 2011. Multi-subject search correctly identifies
causal connections and most causal directions in the {DCM} models of the smith et al.
simulation study. NeuroImage 58 (3), 838 – 848.
URL http://www.sciencedirect.com/science/article/pii/S1053811911007142

Ramsey, J. D., Hanson, S. J., Hanson, C., Halchenko, Y. O., Poldrack, R. A., Glymour, C.,
Jan 2010. Six problems for causal inference from fmri. Neuroimage 49 (2), 1545–1558.

Rio, D. E., Rawlings, R. R., Woltz, L. A., Gilman, J., Hommer, D. W., 2013. Development
of the complex general linear model in the fourier domain: Application to fmri multi-
ple input-output evoked responses for single subjects. Computational and Mathematical
Methods in Medicine 2013, 16.
URL http://dx.doi.org/10.1155/2013/645043

BIBLIOGRAPHY

Rubinov, M., Sporns, O., 2010. Complex network measures of brain connectivity: Uses and interpretations. NeuroImage 52 (3), 1059 – 1069, computational Models of the Brain.

Schmidt, M., Lipson, H., 2009a. Distilling free-form natural laws from experimental data. Science 324 (5923), 81–85.
URL http://www.sciencemag.org/content/324/5923/81.abstract

Schmidt, M. D., Lipson, H., 7-11 Jul. 2007. Learning noise. In: Thierens, D., Beyer, H.-G., Bongard, J., Branke, J., Clark, J. A., Cliff, D., Congdon, C. B., Deb, K., Doerr, B., Kovacs, T., Kumar, S., Miller, J. F., Moore, J., Neumann, F., Pelikan, M., Poli, R., Sastry, K., Stanley, K. O., Stutzle, T., Watson, R. A., Wegener, I. (Eds.), GECCO '07: Proceedings of the 9th annual conference on Genetic and evolutionary computation. Vol. 2. ACM Press, London, pp. 1680–1685.

Schmidt, M. D., Lipson, H., 2009b. Incorporating expert knowledge in evolutionary search: a study of seeding methods. In: Proceedings of the 11th Annual conference on Genetic and evolutionary computation. ACM, pp. 1091–1098.

Smith, S. M., Fox, P. T., Miller, K. L., Glahn, D. C., Fox, P. M., Mackay, C. E., Filippini, N., Watkins, K. E., Toro, R., Laird, A. R., Beckmann, C. F., 2009. Correspondence of the brain's functional architecture during activation and rest. Proceedings of the National Academy of Sciences 106 (31), 13040–13045.
URL http://www.pnas.org/content/106/31/13040.abstract

Spangler, S., Wilkins, A. D., Bachman, B. J., Nagarajan, M., Dayaram, T., Haas, P., Regenbogen, S., Pickering, C. R., Comer, A., Myers, J. N., Stanoi, I., Kato, L., Lelescu, A., Labrie, J. J., Parikh, N., Lisewski, A. M., Donehower, L., Chen, Y., Lichtarge, O., 2014. Automated hypothesis generation based on mining scientific literature. In: Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. KDD '14. ACM, New York, NY, USA, pp. 1877–1886.
URL http://doi.acm.org/10.1145/2623330.2623667

Sporns, O., 05 2014. Contributions and challenges for network models in cognitive neuroscience. Nat Neurosci 17 (5), 652–660.
URL http://dx.doi.org/10.1038/nn.3690

Stam, C. J., Reijneveld, J. C., 2007. Graph theoretical analysis of complex networks in the brain. Nonlinear Biomed Phys 1 (1), 3.

van den Heuvel, M. P., Stam, C. J., Boersma, M., Hulshoff Pol, H. E., Nov 2008. Small-world and scale-free organization of voxel-based resting-state functional connectivity in the human brain. Neuroimage 43 (3), 528–539.

Whelan, R., Conrod, P. J., Poline, J.-B., Lourdusamy, A., Banaschewski, T., Barker, G. J., Bellgrove, M. A., Buchel, C., Byrne, M., Cummins, T. D. R., Fauth-Buhler, M., Flor, H., Gallinat, J., Heinz, A., Ittermann, B., Mann, K., Martinot, J.-L., Lalor, E. C., Lathrop, M., Loth, E., Nees, F., Paus, T., Rietschel, M., Smolka, M. N., Spanagel, R., Stephens, D. N., Struve, M., Thyreau, B., Vollstaedt-Klein, S., Robbins, T. W., Schumann, G., Garavan, H., 06 2012. Adolescent impulsivity phenotypes characterized by distinct brain networks. Nat Neurosci 15 (6), 920–925.
URL http://dx.doi.org/10.1038/nn.3092

Whelan, R., Watts, R., Orr, C. A., Althoff, R. R., Artiges, E., Banaschewski, T., Barker, G. J., Bokde, A. L. W., Buchel, C., Carvalho, F. M., Conrod, P. J., Flor, H., Fauth-Buhler, M., Frouin, V., Gallinat, J., Gan, G., Gowland, P., Heinz, A., Ittermann, B., Lawrence, C., Mann, K., Martinot, J.-L., Nees, F., Ortiz, N., Paillere-Martinot, M.-L., Paus, T., Pausova, Z., Rietschel, M., Robbins, T. W., Smolka, M. N., Strohle, A., Schumann, G., Garavan, H., the IMAGEN Consortium, 08 2014. Neuropsychosocial profiles of current and future adolescent alcohol misusers. Nature 512 (7513), 185–189.
URL http://dx.doi.org/10.1038/nature13402

Wiecki, T. V., Frank, M. J., 2013. A computational model of inhibitory control in frontal cortex and basal ganglia. Psychological Review 120 (2), 329–355.

Appendix A

Subject-Level Variation of NFM Hierarchies

Figure A.1 contains individual subject interaction hierarchies for two different (randomly selected) subjects, generated by 100 random restarts of the GP algorithm and subsequent normalized frequency analysis. Though the two hierarchies are quite different from one another, they do show some network organization similar to that illustrated in the population level hierarchy (top of Figure 2.5).

- Portions of the visual cluster (ROI 32-38) are intact in each case.

- Many of the two-region networks remain together, e.g. ICN 1 (ROI 1,2), ICN 16 (ROI 48,49), and though associated with other ROI, also ICN 3 (ROI 5,6) and ICN 17 (ROI 50,51)

- The default mode network (ROI 39-42) is mostly intact in each subject.

The interaction profile of the default mode network illustrates an interesting distinction between the two subjects. For the top subject, the network is fully intact, interacting with ROI 4 (consistent with the population level hierarchy), and also interacting with ROI 26 from the motor visuospatial complex. For the bottom subject, three of the four ROI in the
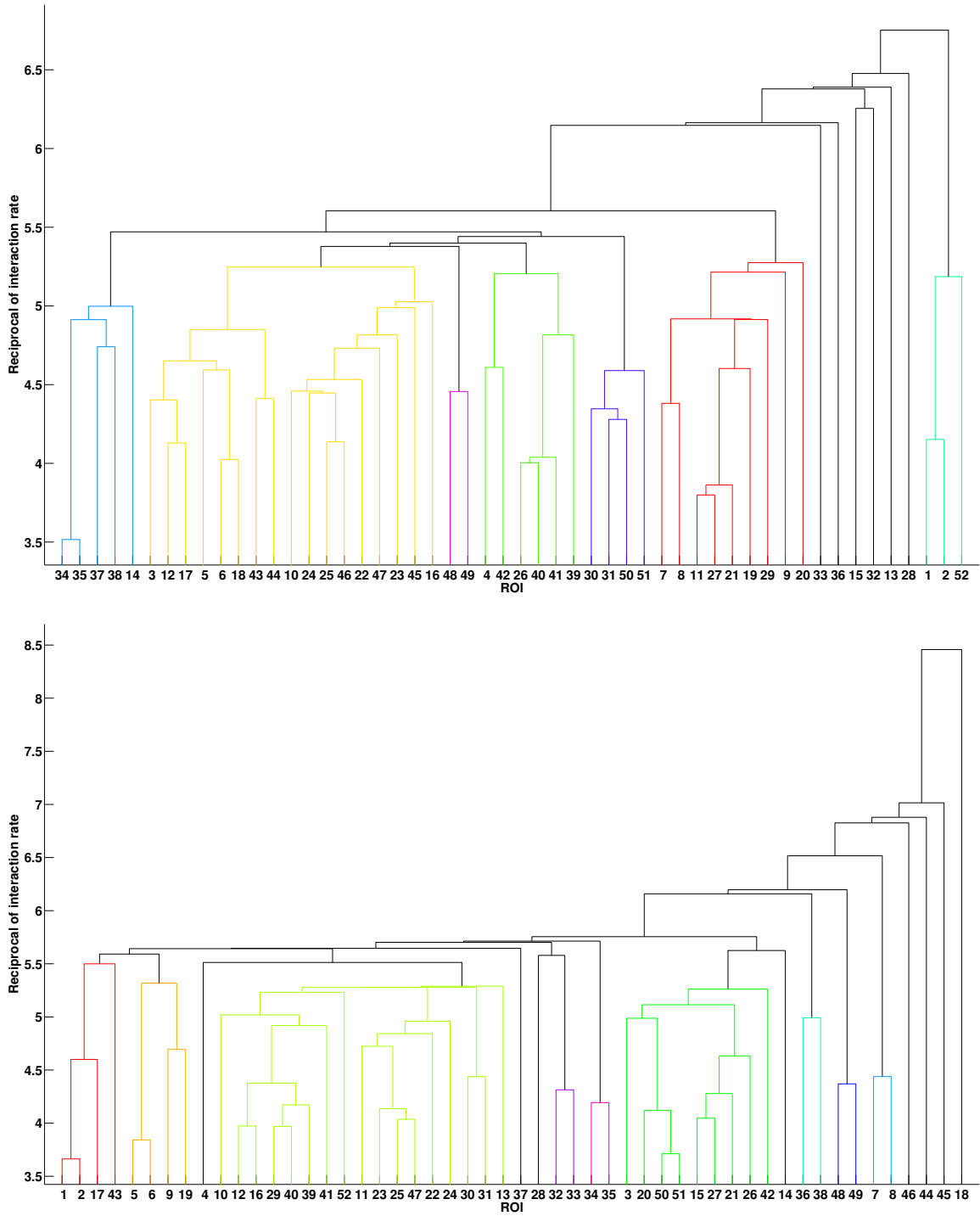
Figure A.1: Example hierarchies for two different individual subjects. Note the large degree of variation between the two.

APPENDIX A. SUBJECT-LEVEL VARIATION OF NFM HIERARCHIES

network remain together, but interact instead with several other ROI from the emotional interoceptive class instead of ROI 4, specifically ROI 10, 12, and 16, and a different ROI from the motor visuospatial complex as well (ROI 29 instead of 26). By themselves, these dendrogram comparisons offer no conclusive evidence regarding connections between cognitive processes. However, an experiment could be designed to test if any inferences can be made from such distinctions. For example, the administration of post-scan surveys might grant some interpretability to the specifics of these single-subject interaction hierarchies.

Appendix B

Linear HCA of Alcohol Consumption

Here we demonstrate that the shuffling of the interaction hierarchy in drinking (D) versus non-drinking (ND) adolescents discovered by NFM is not uncovered by linear correlation analysis. To perform group-level correlation analysis, the normalized relative R^2 matrices for each subject (described in Section 2.3.2) are averaged over the 100 subjects in each group. Recall that these matrices are generated for each subject by computing the correlation matrix for the 52 ROI time series, squaring the elements, and normalizing each row (after setting the diagonal to zero). The reciprocal of relative explained variance can be considered a distance between ROI (higher relative R^2 means closer), and the resulting D and ND hierarchies generated by HCA are shown in the top and bottom, respectively, of Figure B.1.

Comparison with Figure 2.8 suggests that this linear analysis *partially* uncovers a distinguishing difference in interaction between drinking and non-drinking adolescents. Specifically, among non-drinkers, a higher intra-network interaction within ICN 2, comprised of the subgenual ACC and the orbitofrontal cortex (ROI 3 and 4, respectively) is detected here. The result is an indirect coupling, within non-drinkers, of the default mode network (ROI 39-42) and the complex comprised of ROI 3,18,5,6, through the orbitofrontal cortex.

The results of NFM provide further insight in two important ways. First, the elevated interaction within ICN 2 among non-drinkers is detected at twice the strength. Second,

the main interaction responsible for grouping the complex of ROI 3,18,5,6, specifically the interaction between the left putamen (ROI 6) and anterior nuclei of the the thalamus (ROI 18), is lower among non-drinkers. This second effect is entirely missed by correlation analysis, suggesting that it is nonlinear in nature. The result of capturing these effects *together*, as shown in Figure 2.8, is a breakup of the complex in non-drinkers, for whom ROI 3,18 are separated from the bilateral basal ganglia of ICN 3 (ROI 5-6). This breakup is suggestive, as each of these ROI is associated with emotion, reward, and interoceptive processes such as thirst, and experiments reporting activity in ICN 5, including the anterior nuclei of the thalamus (ROI 18), predominantly involved interoceptive stimulation, as reported in Laird et al. (2011).
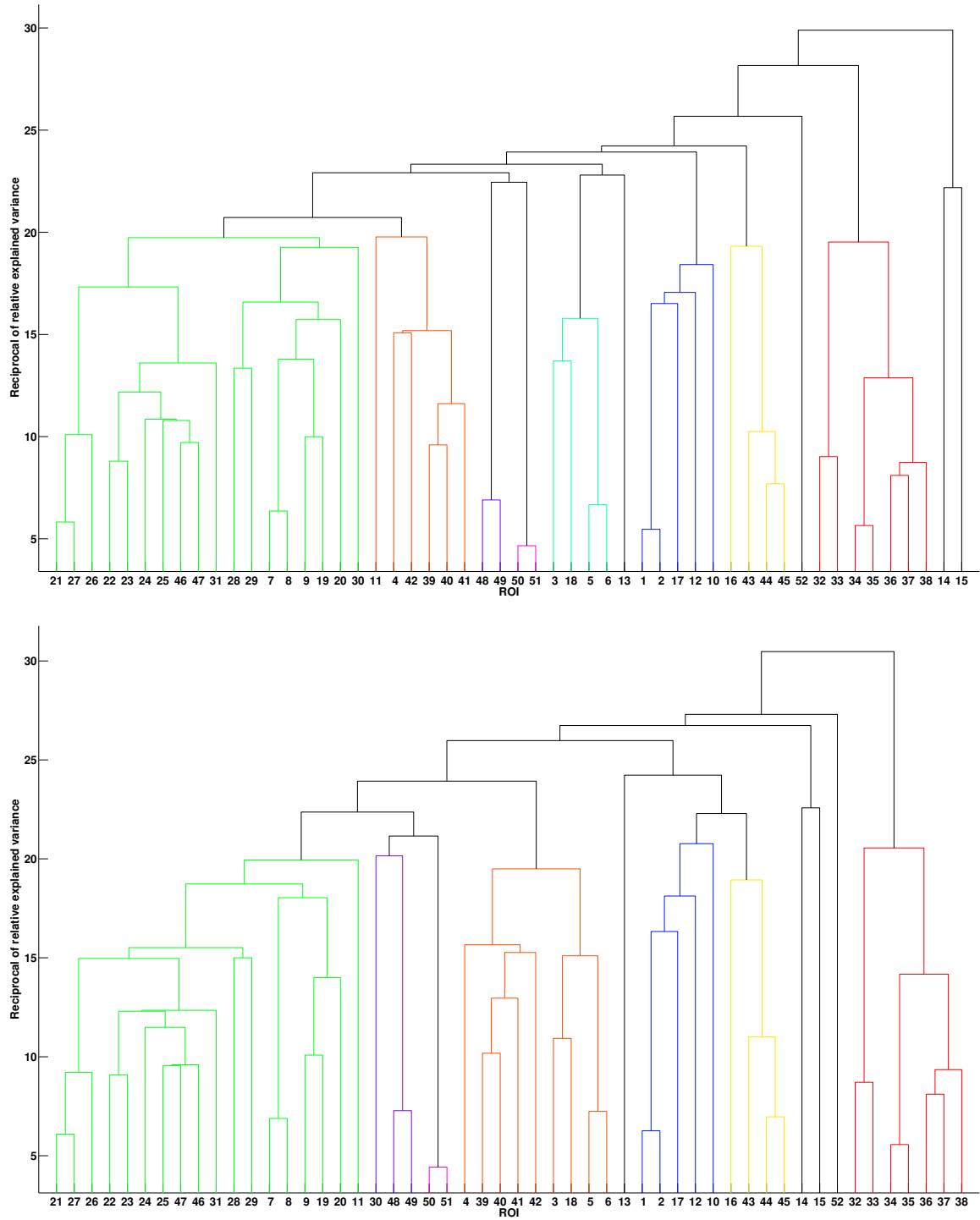
Figure B.1: Linear hierarchies for groups with high (top) and low (bottom) alcohol consumption rates, defined by two or more lifetime drinks and one or fewer lifetime drinks, respectively.

Appendix C

Improvements and Modifications to Functional Mapping

The GP implementation we used for this study is the commercially available package Eureqa from Nutonian, as described in Schmidt and Lipson (2009a). Though much of its behavior can be controlled through the interface or command line, it is proprietary code and thus somewhat of a black box. There are many reasons why a dedicated, open source implementation of GP would be more desirable.

A major challenge for this method of analysis is the computational expense of running a large number of GP searches. Generating the IR map for a single subject requires a large number of random restarts for each ROI. For example, running 100 restarts for each of the 52 ROI in this study, allowing 1 core-hour for each search, requires over 10 hours with access to 500 dedicated processors. The procedure as described here is likely computationally prohibitive for running analyses on large numbers of subjects, or for larger collections of ROI. Intelligent stopping criteria, and many other approaches to the mitigation of computational expense, have been reported at length in the GP literature, an example of which is the use of graphics processors reported in Harding and Banzhaf (2007). It may also be possible to

59

determine an ideal (and smaller) number of restarts that balances computation time with the statistical power of the resulting IR map.

It should also be noted that for collections of ROI much larger than that considered here, in addition to the computational expense resulting from more required searches, each search will take much longer to produce meaningful models due to the larger number of possible explanatory variables. A hybrid method of symbolic regression employing a machine learning algorithm called FFX (Fast Function Extraction) described in McConaghy (2011) as a first pass, and then GP, has great potential for the treatment of higher dimensional data, e.g., large numbers of ROI. A prototype of this method was reported in Icke et al. (2014). FFX is a deterministic algorithm that builds up models with nonlinear terms (e.g., products of ROI signal) in a prescribed fashion and evaluates explanatory power at each stage. By ruling out ROI that are likely not explanatory at each stage, the algorithm reduces the dimensionality of the search. In other words, at the cost of reduced breadth in the search space, the algorithm provides huge reductions in computation time in addition to reducing the number of variables that will eventually be injected into the GP algorithm. Implemented effectively, this hybrid algorithm could eliminate the necessity of ROI selection completely by allowing direct regression over voxel signals.

An ever-present concern in the analysis of fMRI is the level of noise in the data. Particularly in the case of regressing over voxel signals, low signal-to-noise ratio is a major challenge, and indeed GP efficacy is diminished in such circumstances. However, there has been some work on modifying the GP algorithm to better manage noisy data, an example of which is the inclusion of noise generators called *stochastic elements* with user-defined distributions (e.g., Gaussian or uniform) as potential explanatory "variables". These generators can themselves end up inside complex functions within the models, providing those models the capability of reproducing realistic noise distributions more likely to be at play than the typical Gaussian. There is no guarantee that this modification will prove beneficial
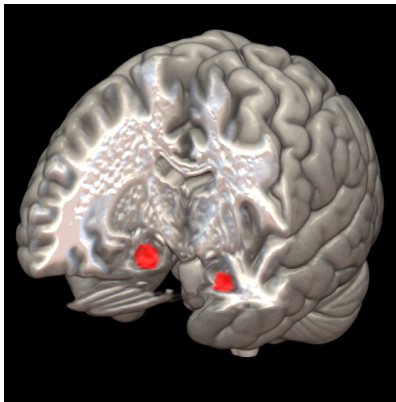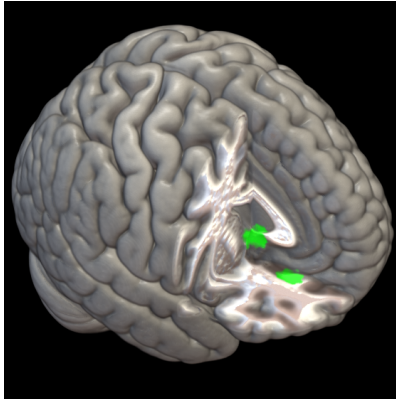
in the case of fMRI, but it has been shown, in Schmidt and Lipson (2007), to effectively identify exact underlying analytical models in the presence of nonlinear, non-Gaussian and nonuniform noise.
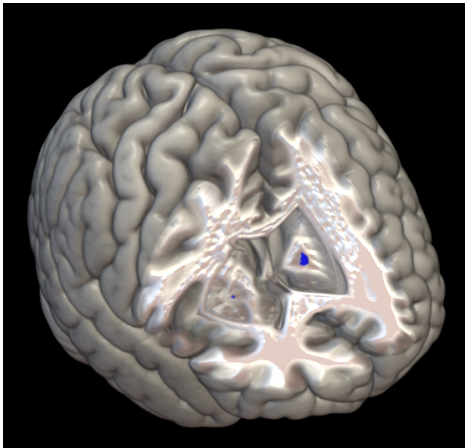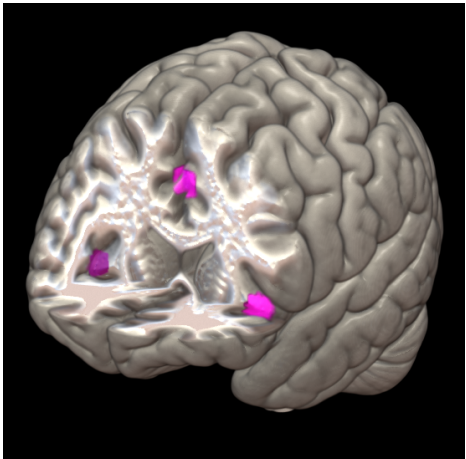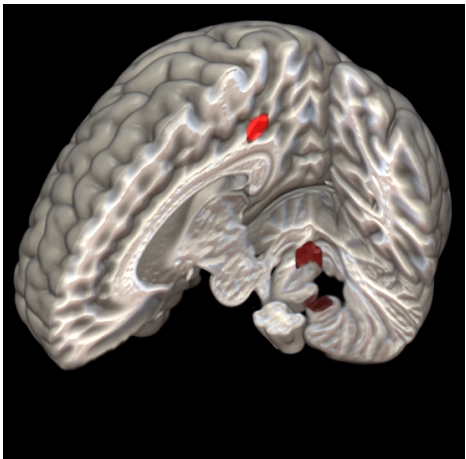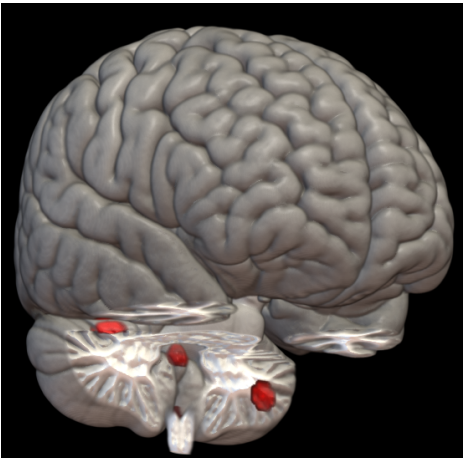
Appendix D

Table of ROI
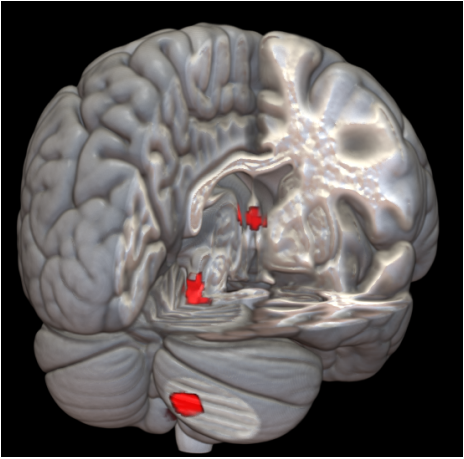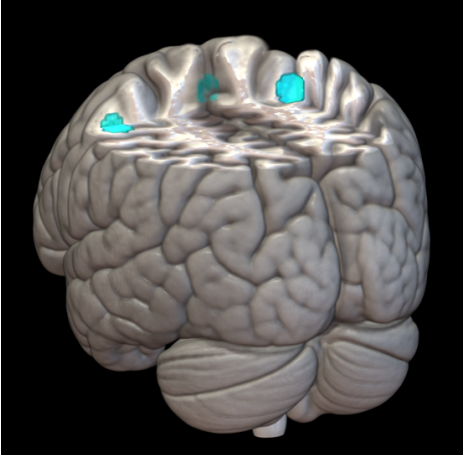
Table D.1: Table of ROI

ROI	ICN	ROI Description	Visualization
1	1	left ventral entorhinal cortex, BA 28	
2	1	right ventral entorhinal cortex, BA 28	
3	2	subgenual anterior cingulate cortex	
4	2	orbitofrontal cortex	
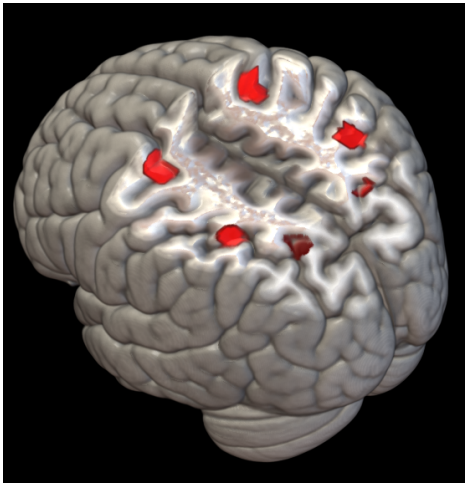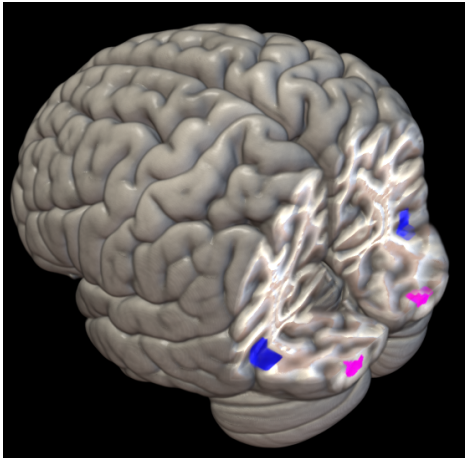
ROI	ICN	ROI Description	Visualization
5	3	right putamen, basal ganglia	
6	3	left putamen, basal ganglia	
7	4	right pars triangularis, BA 45	
8	4	left pars triangular is, BA 45	
9	4	dorsal anterior cingulate cortex, BA 32	
10	5	cerebellum, left tonsil	
11	5	dorsal posterior cingulate cortex, BA 31	
12	5	ventromedial cerebellar cortex	
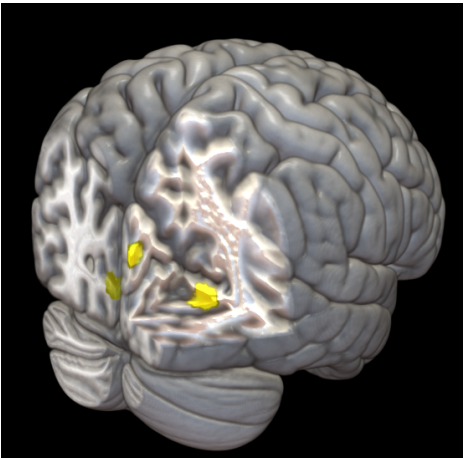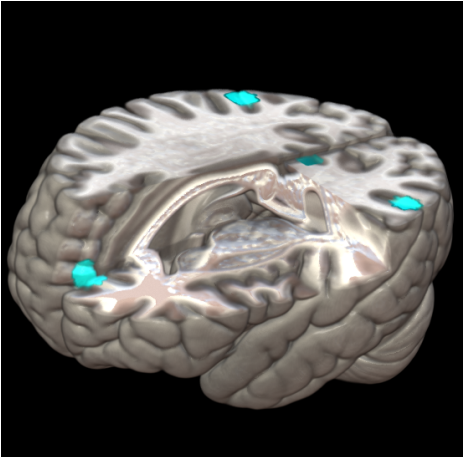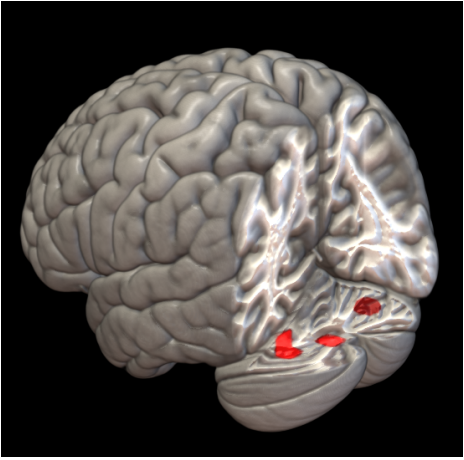
APPENDIX D. TABLE OF ROI

ROI	ICN	ROI Description	Visualization
13	5	midbrain	
14	5	right dorsolateral cerebellar cortex	
15	5	left ventrolateral cerebellar cortex	
16	5	cerebellum, right posterior lobe	
17	5	inferior colliculus, midbrain	
18	5	thalamus, anterior nuclei	
19	6	supplementary motor area, BA 6	
20	6	left premotor cortex, BA 6	
21	6	right premotor cortex, BA 6	
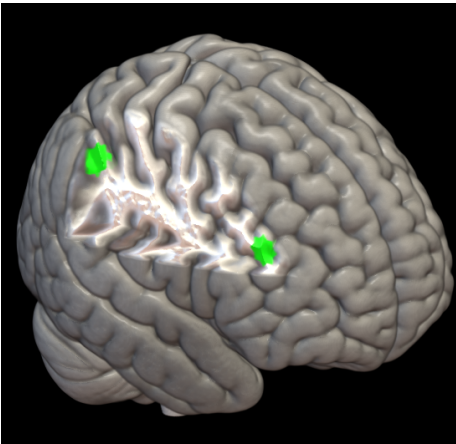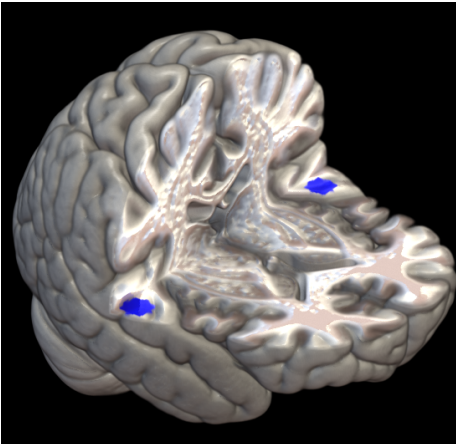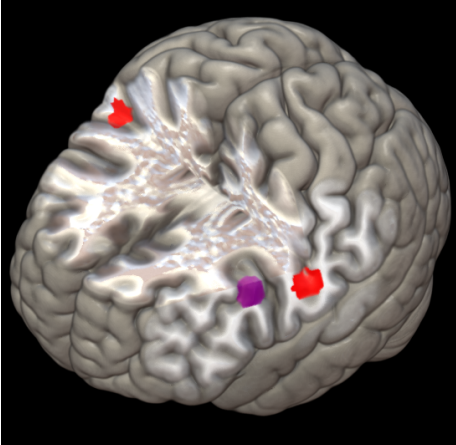
APPENDIX D. TABLE OF ROI

ROI	ICN	ROI Description	Visualization
22	7	left somatosensory cortex, BA 7	
23	7	right somatosensory cortex, BA 7	
24	7	left somatosensory cortex, BA 5	
25	7	right somatosensory cortex, BA 5	
26	7	left primary motor cortex, BA 4	
27	7	right primary motor cortex, BA 4	
28	8	left dorsal primary motor cortex, BA 4	
29	8	central sulcus of the primary motor cortex, BA 4	
30	8	left ventral primary motor cortex, BA 4	
31	8	right primary motor cortex, BA 4	
32	10	right middle temporal gyrus, V5, BA 19	
33	10	left middle temporal gyrus, V5, BA 19	
34	11	left primary visual cortex, V1, BA 17	
35	11	right primary visual cortex, V1, BA 17	

ROI	ICN	ROI Description	Visualization
36	12	medial primary visual cortex, V1, BA 17	
37	12	right medial secondary visual cortex, V2, BA 18	
38	12	right associative visual cortex, V3, BA 19	
39	13	dorsal posterior cingulate cortex, BA 31	
40	13	left angular gyrus, BA 39	
41	13	right angular gyrus, BA 39	
42	13	ventromedial prefrontal cortex, BA 10	
43	14	right cerebellar cortex	
44	14	medial cerebellar cortex	
45	14	left cerebellar cortex	

ROI	ICN	ROI Description	Visualization
46	15	right dorsolateral prefrontal cortex, BA 9	
47	15	right somatosensory association cortex, BA 5	
48	16	left auditory cortex, BA 41	
49	16	right auditory cortex, BA 41	
50	17	right primary somatosensory cortex, BA 3,1,2	
51	17	left primary motor cortex, BA 4	
52	18	left fronto-parietal, Broca's area, BA 44,45	