

# Predictive Connectome Subnetwork Extraction with Anatomical and Connectivity Priors

Colin J Brown<sup>a</sup>, Steven P Miller<sup>b</sup>, Brian G Booth<sup>a</sup>, Jill G Zwicker<sup>c,d</sup>, Ruth E Grunau<sup>c</sup>, Anne R Synnes<sup>c</sup>, Vann Chau<sup>b</sup>, Ghassan Hamarneh<sup>a</sup>

<sup>a</sup>Medical Image Analysis Lab, Simon Fraser University, Burnaby, BC, Canada

<sup>b</sup>Department of Paediatrics, The Hospital for Sick Children and the University of Toronto, Toronto, ON, Canada

<sup>c</sup>BC Children's Hospital Research Institute and Department of Pediatrics, University of British Columbia, Vancouver, BC, Canada

<sup>d</sup>Department of Occupational Science & Occupational Therapy, University of British Columbia, Vancouver, BC, Canada

---

## Abstract

We present a new method to identify anatomical subnetworks of the human connectome that are optimally predictive of targeted clinical variables, developmental outcomes or disease states. Given a training set of structural or functional brain networks, derived from diffusion MRI (dMRI) or functional MRI (fMRI) scans respectively, our sparse linear regression model extracts a weighted subnetwork. By enforcing novel backbone network and connectivity based priors along with a non-negativity constraint, the discovered subnetworks are simultaneously anatomically plausible, well connected, positively weighted and reasonably sparse. We apply our method to 1) predicting the cognitive and neuromotor developmental outcomes of a dataset of 168 structural connectomes of preterm neonates, and 2) predicting the autism spectrum category of a dataset of 1013 resting-state functional connectomes from the Autism Brain Imaging Data Exchange (ABIDE) database. We find that the addition of each of our novel priors improves prediction accuracy and together outperform other state-of-the-art prediction techniques. We then examine the structure of the learned subnetworks in terms of topological features and with respect to established function and physiology of different regions of the brain.

---

## 1. Introduction

### 1.1. Motivation

Increasingly, diffusion MR and functional MR images (dMRI and fMRI, respectively) of the human brain are being represented as networks (i.e., connectomes) in order to naturally convey white matter connections and correlated blood oxygen level dependent (BOLD) activity between distinct, sometimes spatially disparate regions of the brain. Connectome analysis has been used to study a wide variety of aspects of brain structure, function and development along with many neurological disorders and diseases (Bassett & Bullmore (2009); Castellanos *et al.* (2013)). Furthermore, it has been shown that machine learning models trained on connectome data can be successful in identifying and predicting many of these same diseases and disorders (Brown & Hamarneh (2016)).

Representing an MRI of the brain as a network allows a reduction of dimensionality from millions of (time series or tensor-valued) voxels down to thousands of connections (edges); a reduction informed by the anatomy of the brain as opposed to the representation of the image. However, for the purposes of prediction, thousands of features may still be too many and cause over-fitting when limited numbers of scans are available (Munsell *et al.* (2015)).

Furthermore, certain parts of the brain may be more informative than others with respect to a particular outcome or state. For instance, region of interest (ROI) based studies suggest that structural abnormalities related to poor neurodevelopmental outcomes are not spread evenly across the entire preterm brain but instead localized to particular brain anatomy (Chau

*et al.* (2013)). Similarly, only particular functional subnetworks of the brain are known to be altered by autism spectrum disorders (ASD) (Assaf *et al.* (2010)). Thus, there is motivation to discover which particular subnetworks (group of connections or edges) in the brain network most affect different brain functions or are most impacted by different disorders.

In this paper, we present a method to identify anatomical subnetworks of the human connectome that are optimally predictive of targeted clinical diagnoses or developmental outcomes. We define a subnetwork as meaning any connected or disconnected subset of edges in a brain network (not necessarily only those with a known functional or anatomical purpose). Similar to Munsell *et al.* (2015), our method is based on a regularized linear regression on the output label of choice. Here, however, we leverage two novel informed priors designed to find predictive edges that are well integrated into a connected subnetwork and are unlikely to be dominated by image noise. Furthermore, we impose a constraint that ensures the non-negativity of subnetwork edge weights. We apply our method to 1) predicting the cognitive and neuromotor developmental outcomes of a dataset of 168 structural connectomes of preterm infants acquired shortly after birth, and 2) the autism spectrum category of a dataset of 1013 functional connectomes from the Autism Brain Imaging Data Exchange (ABIDE). We demonstrate that these priors effectuate the desired effect on the learned subnetworks and that, consequently, our model achieves higher accuracies on these challenging prediction tasks. Finally, we discuss the structure of the learned subnetworks in the context of the underlying neuroanatomy.

Building on our prior work (Brown *et al.* (2016)), here we introduce a new bounded connectivity prior which improves on the original connectivity prior, validate our approach on the larger ABIDE dataset and provide expanded discussions of related works, the implementation and runtime complexity of our novel priors and our experimental results. Finally, we also examine the relationship between subnetworks predictive of cognitive neurodevelopmental outcomes versus those predictive of ASD.

## 1.2. Related Works

A variety of works, especially in past decade, have extracted brain subnetworks as a form of feature selection for subsequent analysis or prediction tasks (Brown & Hamarneh (2016)). Typically, these approaches either 1) perform some test of discriminance or signal strength at each edge independently, 2) use a generative model to find a basis set of subnetworks that are used in further analysis or 3) use a predictive model to find edge weights that predict a specific outcome.

The first approach, filtering edges independently, is not ideal in the setting of connectome analysis since it is assumed that there are relationships between different connections in the brain, especially those local to one another. For this reason, subnetwork selection approaches that operate on each edge individually are typically only used as preprocessing steps to find candidate edges. For instance, Zhu *et al.* (2014) used t-tests at each edge in a dataset of functional connectomes for group discriminance followed by correlation-based feature selection (CFS) (Hall & Smith (1999)) to find subnetworks that were maximally predictive of Schizophrenia. In contrast, in our method we perform feature selection in a single step which is preferable as it allows all model parameters to be optimized simultaneously.

The second approach, is to find a set of basis subnetworks that represent localized variation across connectomes in the dataset (An *et al.* (2010); Calhoun *et al.* (2008); Ghanbari *et al.* (2014)). One of the early works of this kind was by Calhoun *et al.* (2008) who used independent component analysis (ICP) to find basis subnetworks for functional brain networks. Similarly, Ghanbari *et al.* (2014) used non-negative matrix factorization to find a sparse set of non-negative basis subnetworks in structural connectomes derived from dMRIs. Along with Anderson *et al.* (2014), they argued that non-negative subnetwork edge weights are more anatomically interpretable, especially in the case of structural connectomes which have only non-negative edge feature values.

One drawback of these generative methods is that they do not explicitly expose the factors that cause the variation in each subnetwork. Further analysis must be performed to check if particular subnetworks co-vary with other variables of interest. To address this issue, Ghanbari *et al.* (2014) introduced age-regressive, group-discriminative and reconstructive regularization terms on groups of subnetworks, encouraging each group to covary with a known factor. While this encourages subnetworks to be related to specific outcomes, it does not guarantee that the networks will be optimally predictive of those outcomes, since the degree to which the regularization is enforced

can be weighted by any amount. Furthermore, the number of subnetworks to extract is also a free parameter and it is not clear how this value should be set.

The third common approach is to frame subnetwork extraction as a sparse feature selection problem over connectome edges for the task of regression or classification. This class of methods, to which our proposed method belongs, has the benefit of being able to find a subnetwork of edges that are optimally predictive of an outcome of choice. Generally, given a data set  $\{(\mathbf{x}_i, y_i)\}$  for  $i \in [1, \dots, N]$ , of observed input variables  $X = [\mathbf{x}_1, \dots, \mathbf{x}_N]^T \in \mathbb{R}^{N \times M}$ , and a vector of outcomes (output variables),  $\mathbf{y} = [y_1, \dots, y_N]^T \in \mathbb{R}^{N \times 1}$ , training of these models requires optimizing an objective function of the form,

$$F(\mathbf{w}|X, \mathbf{y}) = \mathcal{L}(\mathbf{w}|X, \mathbf{y}) + \lambda_{L1}\|\mathbf{w}\|_1 + \lambda_{L2}\|\mathbf{w}\|_2^2 + \mathcal{R}(\mathbf{w}|X), \quad (1)$$

where  $\mathcal{L}$  is the loss function, which fits the model parameters,  $\mathbf{w}$ , to the data,  $\lambda_{L1}$  and  $\lambda_{L2}$  govern the strength of  $L1$  and  $L2$  regularization and  $\mathcal{R}$  is an aggregate of any additional (possibly data dependent) regularization terms. The  $L2$  term penalizes larger parameters more strongly whereas the  $L1$  term penalizes parameters independent of magnitude, leading small parameters to go to zero and thus encouraging sparsity. When  $\mathcal{R} = 0$ , the regularization on  $F$  is called *Elastic Net* regularization and when also  $\lambda_{L2} = 0$ , it is known as *LASSO* regularization (Zou & Hastie (2005); Tibshirani (1994)). In the case that only  $\lambda_{L2}$  is non-zero and  $\mathcal{L}$  is a linear regression,  $F$  is the widely used *Ridge Regression* (Hoerl & Kennard (1970)).

Many recent works on extracting subnetworks have employed some form of Eqn. 1. For instance, Casanova *et al.* (2012) used a LASSO regularized random forest classifier to find a sparse set of edges from fMRI connectomes that distinguished between adult male and female brains. Similarly, Munsell *et al.* (2015) used an Elastic Net based subnetwork selection in a 2-stage prediction framework for predicting the presence of temporal lobe epilepsy (TLE) and the success of corrective surgery in those with TLE. The use of the  $L1$  regularizer in these methods encourages sparse selection of stable features, useful for identifying those edges most important for prediction (Grosenick *et al.* (2013)). These methods, however, fail to leverage the underlying structure of the brain networks that might inform the importance or relationships between edges.

Watanabe *et al.* (2014) extended these approaches and exploited the relative position of edges by using a spatially regularized support-vector machine (SVM) to find a structural subnetwork, predictive of schizophrenia. However, their approach required parcellation of the brain into a regular grid, which is not ideal since functional boundaries of the brain are ignored. In contrast, Li *et al.* (2012) employed a Laplacian-based regularizer which encourage subnetwork weights to smoothly vary between neighbouring edges. Note that the combination of an  $L1$  and a Laplacian-based regularization term is called GraphNet (Grosenick *et al.* (2013); Ng *et al.* (2012)). The smoothing encouraged by a Laplacian, however, may reduce sparsity by promoting many small weights and blur discontinuities between the weights of neighbouring edges that should be pre-

served. An ideal regularizer would encourage a well connected subnetwork while preserving sparsity and discontinuities.

A number of recent papers have used neural network based models for performing regression and classification (He *et al.* (2018); Li *et al.* (2017); Kawahara *et al.* (2016)). While extremely powerful as learning models, the use of neural networks for brain network data remains challenging due to the large number of training instances typically required to prevent over-fitting.

Broadly, our approach extends the LASSO regression framework, adding two further regularization terms, tailored to the task of predictive subnetwork extraction human brain networks: an anatomical prior term that discourages the subnetwork from including anatomically implausible edges and a connectivity prior that encourages well connected subnetworks (as opposed to subnetworks comprising scattered, isolated edges). Additionally, we include a non-negativity constraint for more interpretable weights in the context of non-negative brain network connectivity values.

Qiao *et al.* recently presented an approach for estimating functional brain networks that incorporated a prior term to encourage modular structure of the inferred connectome (Qiao *et al.* (2016)). For the same task, Yu *et al.* introduced a weighted sparsity term and a term designed to leverage group structure (Yu *et al.* (2017)). While there are parallels between their proposed terms and ours (e.g., in the inspiration to encode connectivity informed sparsity and to impose some form of group structure as a prior), their terms were embedded in generative models to infer functional connectivity rather than in an end-to-end discriminative model as is proposed here.

We describe our new objective function in detail in Section 2 and then compare its performance against other similar objective functions in Section 3.

## 2. Method and Materials

Here we present our novel subnetwork extraction method (Section 2.1), details about each of the novel prior terms (Sections 2.2, 2.3 and 2.4) and the details of the two datasets used for validation (Sections 2.5 and 2.6).

### 2.1. Subnetwork Extraction

In this section, we outline our proposed method for extracting a subnetwork of connectome edges that is:

1. predictive (i.e., contains edges that accurately predict a neurodevelopmental outcome),
2. anatomically plausible (i.e., edges correspond to functional or structural connections),
3. well connected (i.e., high network integration, as described in Brown *et al.* (2014)),
4. reasonably sparse and
5. non-negative.

Each connectome is represented as an undirected graph  $G(V, E)$  comprising a set of vertices,  $V$ , and  $M$  edges,  $E$ . Note that for both datasets used in this work,  $|V| = 90$  and so

$M = |E| = 90 \times 89/2 = 4005$ . The structural or functional connectivities (i.e., tract counts or correlations, respectively) associated with the edges are represented as a single feature vector  $\mathbf{x} \in \mathbb{R}^{1 \times M}$  and the entire training set of  $N$  subjects is represented as  $X \in \mathbb{R}^{N \times M}$  with labels (e.g., outcome scores or ASD categories)  $\mathbf{y} \in \mathbb{R}^{N \times 1}$ . To find a subnetwork that fits the above criteria, we optimize an objective function over a vector of subnetwork edge weights,  $\mathbf{w} \in \mathbb{R}^{M \times 1}$ :

$$\mathbf{w}^* = \underset{\mathbf{w}}{\operatorname{argmin}} \|\mathbf{y} - X\mathbf{w}\|^2 + \lambda_{L1}\|\mathbf{w}\|_1 + \mathcal{R}(\mathbf{w}) \quad (2)$$

$$\text{such that } \mathbf{w} \geq 0, \quad (3)$$

$$\text{and } \mathcal{R}(\mathbf{w}) = \lambda_B(\mathbf{w}^T B \mathbf{w}) + \lambda_C(\mathbf{w}^T C \mathbf{w}), \quad (4)$$

where  $\|\mathbf{w}\|_1$  is a sparsity regularization term,  $B$  is the network backbone prior matrix (see Section 2.3), and  $C$  is the connectivity prior matrix (see Section 2.4). Hyper-parameters,  $\lambda_B, \lambda_C$  and  $\lambda_{L1}$  are used to weight each of the regularization terms. Given a set of learned weights,  $\mathbf{w}^*$ , the label of a novel, unseen connectome,  $\mathbf{x}_{new}$  can be predicted as  $y_{pred} = \mathbf{x}_{new} \mathbf{w}^*$ .

To perform this optimization we used the method (and software) of Schmidt (2010).

### 2.2. Network Backbone Prior

Given a set of regions in the brain, many possible connectome edges between those regions may be anatomically unlikely (i.e., between regions not connected by white matter fibers or regions that do not co-activate) but may have non-zero connectivity in certain scans due to imaging noise and accumulated pipeline error (i.e. due to atlas registration, parcellation and for structural connectomes, tractography) (Cheng *et al.* (2012)). Especially in the case of high dimensional data (i.e., thousands of edges) and relatively few training samples, some edges may appear discriminative by pure chance, when in fact they are just noise. Therefore, we propose a network backbone prior term that encodes a penalty discouraging the subnetwork from including edges with a low signal-to-noise ratio (SNR) in the training data. The SNR of the  $j$ -th edge can be computed as the ratio  $\operatorname{MEAN}(X_{:,j})/\operatorname{SD}(X_{:,j})$ . However, in the case that the dataset contains multiple classes, this ratio may falsely declare an edge as noisy when the variability (c.f. denominator) in the edge value is actually due to the edges' values changing in a manner that correlates with the class label of the subject. To counteract this problem, we compute the SNR separately for each class.

Assume that there are two classes,  $H$  and  $U$  (e.g., adverse neurodevelopmental outcomes versus normal neurodevelopmental outcomes, or ASD versus controls). Let  $X_\Omega$  represent a matrix with a subset of the rows in  $X$  where  $\Omega \in \{U, H\}$ . The SNR for each edge,  $j$ , in each class,  $\Omega$ , is computed as  $\operatorname{SNR}(X_{\Omega,j}) = \frac{\operatorname{MEAN}(X_{\Omega,j})}{\operatorname{SD}(X_{\Omega,j})}$ . In order not to favour the strongest brain connections over weak yet important connections, we threshold the SNR at each edge conservatively, to exclude only the least anatomically likely edges. An edge,  $j$ , is only penalized if both  $\operatorname{SNR}(X_{U,j})$  and  $\operatorname{SNR}(X_{H,j})$  are less than or equal to

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

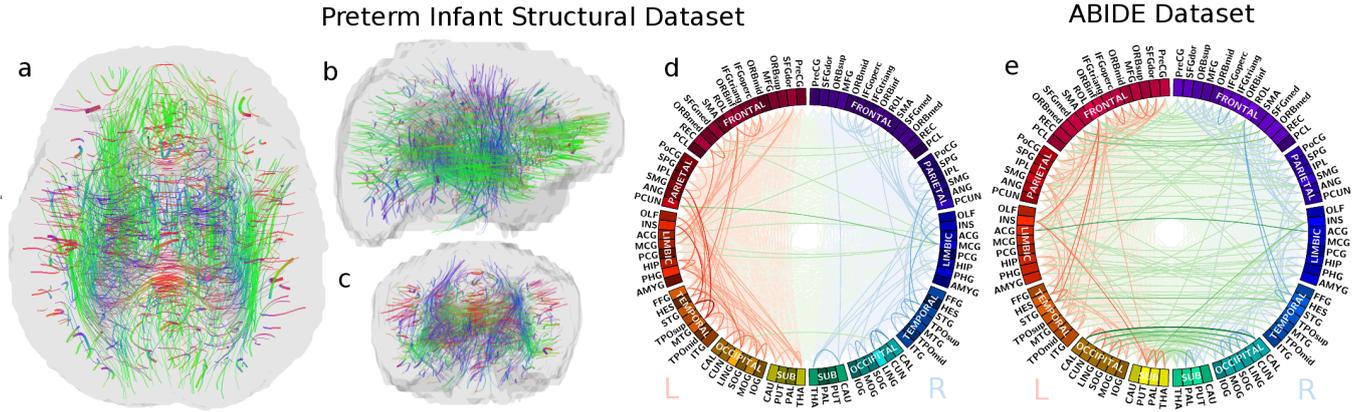


Figure 1: Sample backbone prior networks for preterm infant data (left) and ABIDE data (right). **a-c)** A backbone prior network (i.e.,  $B_{i,j} = 0$ ) for the preterm infant structural connectome data rendered as curves representing the mean shape of all tracts between those connected regions (from one representative subject’s scan). Colours represent the direction of the curve at each location. Axial, sagittal and coronal views in **b)** **c)** and **d)** respectively. **d)** The same network mapped on to a Circos ideogram where green links represent inter-hemispherical connections and red and blue links represent intra-hemispherical connections in the left and right hemispheres respectively. Opacity of each link represents the product of SNR values of the two classes (i.e.,  $SNR(H)_i \times SNR(U)_i$ ). **e)** A backbone prior network for the ABIDE functional connectome data also mapped on to a Circos ideogram.

1 (i.e., signal is weaker than noise in both classes). In particular,  $B$  is an  $M \times M$  diagonal matrix, such that,

$$B_{j,j} = \begin{cases} 1, & \text{if } SNR(X_{H,j}) \leq 1 \text{ and } SNR(X_{U,j}) \leq 1 \\ 0, & \text{otherwise.} \end{cases} \quad (5)$$

So  $\mathbf{w}^T B \mathbf{w}$  only penalizes edges that do not pass the SNR threshold among either instances in class  $H$  or in class  $U$ , and thus are likely noisy. If a dataset does not contain multiple classes, then all connectomes can be considered as part of a single class and the first case in Eqn. 5 collapses to a single inequality. Fig. 1 shows examples of  $B$  for the two datasets used in this paper. Note that in the case of structural connectomes, even edges with high SNR may not represent white matter fibers but instead high FA from other causes, especially for infant connectomes (Brown *et al.* (2014)). Similarly, for functional connectomes, edges with high SNR may not indicate an excitatory (or inhibitory) interaction between two regions, but instead regions that are co-activating due to some tertiary cause (Fallani *et al.* (2014)). Nevertheless, such high-SNR edges are not likely due to noise but instead to some real effect and thus may aid prediction.

While it may also be reasonable to use a statistical threshold (e.g., a t-test) to select entries in  $B$ , the SNR intrinsically provides a reasonable and interpretable threshold (i.e., signal greater than noise). Furthermore, the SNR values are not guaranteed to be normally distributed, which a t-test requires.

Naïvely computing  $\mathbf{w}^T B \mathbf{w}$  directly would require  $O(M^2)$  space and time since  $B \in \mathbb{R}^{M \times M}$ . However, because only the diagonal elements are non-zero,  $B$  does not need to be constructed directly and  $\mathbf{w}^T B \mathbf{w} = \sum_j^M \mathbf{w}_j^2 B_{j,j}$  can be computed in  $O(M)$  space and time.

### 2.3. Connectivity Prior

We also want to encourage the subnetwork to be highly integrated as opposed to being a set of scattered, disconnected

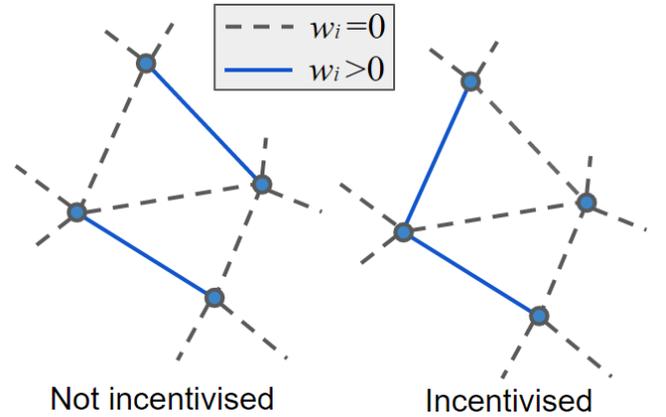


Figure 2: Edge connectivity prior: By including this prior, the model objective function encourages strong weights on pairs of edges that share a node (right) but not pairs that do not share a node (left).

edges. This is motivated by the fact that important brain networks tend to be defined by inter-connectivity as opposed to disparate connections across distinct parts of the brain. For instance, functional brain network activity is generally constrained to white matter structure (Honey *et al.* (2009a)) and white matter structure is organized into well connected link communities (de Reus *et al.* (2014)). One example is the default mode network, which is understood to facilitate certain functions (e.g., internal thought at rest) and is both well functionally connected and well structurally connected (Fair *et al.* (2009)). Furthermore, brain diseases, like amyotrophic lateral sclerosis (ALS) have been shown to impact well-connected networks of the brain (Schmidt *et al.* (2016)).

Thus, we do not expect there to be a large number of disconnected sub-parts of the brain that are all highly responsible for any particular neurodevelopmental outcome type. To embed this prior, we incentivize pairs of edges in the target subnetwork to share common nodes. For edge  $e_{i,j}$ , between nodes  $i$  and  $j$ ,

and edge  $e_{p,q}$  between nodes  $p$  and  $q$ , we construct the matrix,

$$C(e_{i,j}, e_{p,q}) = \begin{cases} -1, & \text{if } i = p \text{ or } i = q \text{ or } j = p \text{ or } j = q \\ 0, & \text{otherwise,} \end{cases} \quad (6)$$

such that the term  $\mathbf{w}^T C \mathbf{w}$  becomes smaller (i.e., more optimal) for each pair of non-zero weighted subnetwork edges sharing a node. This term places a priority on retaining edges in the subnetwork that are connected to hub nodes (Fig. 2). This is desirable since subnetwork hub nodes indicate regions that join many connections (i.e., edges) predictive of outcome. In contrast to a Laplacian based regularizer, which would encourage subnetwork weights to become locally similar, reducing sparsity, our proposed term simply rewards subnetworks with stronger hubs.

Similarly to the backbone network prior,  $\mathbf{w}^T C \mathbf{w}$  can be computed efficiently, without requiring matrix  $C \in \mathbb{R}^{M \times M}$  to be explicitly constructed. Note that the  $j$ th row of  $C$  has a non-zero entry for each edge that neighbours on (i.e., shares a node with) edge  $j$ . With this fact, we can more efficiently select the neighbouring edges of a given edge. Let  $\mathbf{w}(e_{i,j})$  represent the weight of the edge between nodes  $i$  and  $j$  and similarly, let  $(C\mathbf{w})(e_{i,j})$  represent the  $e_{i,j}$ th element of the product  $C\mathbf{w}$ . Then each element  $(C\mathbf{w})(e_{i,j})$  can be computed as,

$$(C\mathbf{w})(e_{i,j}) = \left( \sum_{\{k \in V | k \neq i\}} \mathbf{w}(e_{k,j}) + \sum_{\{k \in V | k \neq j\}} \mathbf{w}(e_{i,k}) \right). \quad (7)$$

Each sum requires  $|V| = \sqrt{M}$  steps and only  $O(M)$  memory. Eqn. 7 must be computed  $M$  times (i.e., once for each edge), and then  $\mathbf{w}^T C \mathbf{w}$  can be computed as,

$$\mathbf{w}^T C \mathbf{w} = \sum_{e_{i,j} \in E} \mathbf{w}(e_{i,j})(C\mathbf{w})(e_{i,j}), \quad (8)$$

with only an additional  $M$  steps (and again,  $O(M)$  memory). Thus, the entire computation requires only  $O(M \sqrt{M})$  time and  $O(M)$  space.

#### 2.4. Bounded Connectivity Prior

One drawback of the connectivity prior,  $\mathbf{w}^T C \mathbf{w}$ , as presented is that it is not bounded and can take on arbitrarily large negative values by increasing the weights of a pair of neighbouring edges, possibly causing degenerate solutions by dominating the objective function. In practice, by initializing the weights reasonably (i.e., all values near zero) we did not find this to occur (due to the influence of the  $L1$  prior which encourages edge weights to decrease) (Brown *et al.* (2016)). Nevertheless, for general applications, this property may pose a problem.

One possible solution is to try to reformulate the connectivity prior as a penalty, rather than as an incentive, by penalizing strong edge pairs that are not neighbours (i.e., do not share a node). However, there may reasonably exist pairs of strong edges in different parts of the subnetwork that are connected but do not share a node and we do not want to discourage this structure. In our preliminary tests, we found that this penalty

Group	# Infants	Birth GA	# Scans	Scan PMA
Normal	100	27.92	146	35.5
Low Motor	21	27.25	29	35.6
Low Cog.	11	27.46	13	35.4
All	115	27.85	168	35.56

Table 1: Demographics of preterm infant dataset from BC Childrens Hospital. Ages are averaged and listed in weeks.

encouraged weights of edges around a single node to become strong and all other weights to go to zero, as expected. Instead, our solution is to maintain the original form of the matrix  $C$  and to place an upper bound on the contribution of each weight. Let

$$g(w_i) = \frac{|w_i|}{|w_i| + 1}. \quad (9)$$

Note that  $g(0) = 0$  and  $\lim_{w_i \rightarrow \infty} g(w_i) = 1$ . Let  $\mathbf{g}(\mathbf{w}) = [g(w_1), \dots, g(w_M)]$ . Then the new bounded connectivity prior term,  $\mathbf{g}(\mathbf{w})^T C \mathbf{g}(\mathbf{w})$ , approaches its largest negative value as all elements of  $\mathbf{w}$  go to infinity and the elements of  $\mathbf{g}(\mathbf{w})$  approach 1. In the limit, each pair of neighbouring edges contributes  $-1$  to the objective function, so  $\mathbf{g}(\mathbf{w})^T C \mathbf{g}(\mathbf{w}) > -2M(|V| - 2)$ , since there are  $M$  edges and each edge has  $2(|V| - 2)$  neighbouring edges. With this modification, the connectivity prior is bounded, preventing degenerate solutions independently of weight initializations. Fig. 3 shows the incentive contribution from two neighbouring edges, across a range of weight values for both the original and the bounded connectivity prior. We test this bounded connectivity prior in Section 3.2.

#### 2.5. Preterm Infant Data

The preterm neonate dataset contains 168 scans taken between 27 and 45 weeks post-menstrual age (PMA) from a cohort of infants, born between 24 and 32 weeks gestational age (GA). Scans were acquired at the British Columbia (BC) Children's Hospital on a Siemens (Berlin, Germany) 1.5T Avanto using VB 13A software. Each scan was a 3D axial volumetric diffusion tensor image set (TR 4900 ms; TE 104 ms; FOV 160 mm; slice thickness, 3 mm; no gap) with 3 averages of 12 non-colinear gradient directions over 2 diffusion weightings of 600 and 700 s/mm<sup>2</sup> (b-value), with an in-plane resolution of 1.3mm. Each diffusion weighted image set was preprocessed using the FSL Diffusion Toolbox (FDT) pipeline<sup>1</sup> and tensors were fit using RESTORE (Chang *et al.* (2005)). Nearly half of the subjects were scanned twice for a total of 168 diffusion tensor images (DTI). Note that the cerebellum was omitted from this analysis as it was not fully captured in many of the scans.

Connectomes were generated for each scan by aligning each DTI to a neonatal template and atlas from the University of North Carolina (UNC) School of Medicine at Chapel Hill (Shi *et al.* (2011)). The UNC neonatal atlas comprises  $|V| = 90$  anatomical regions as defined in the automated anatomical labeling (AAL) atlas (Tzourio-Mazoyer *et al.* (2002)). Full-brain streamline tractography was then performed in order to

<sup>1</sup><http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FDT>

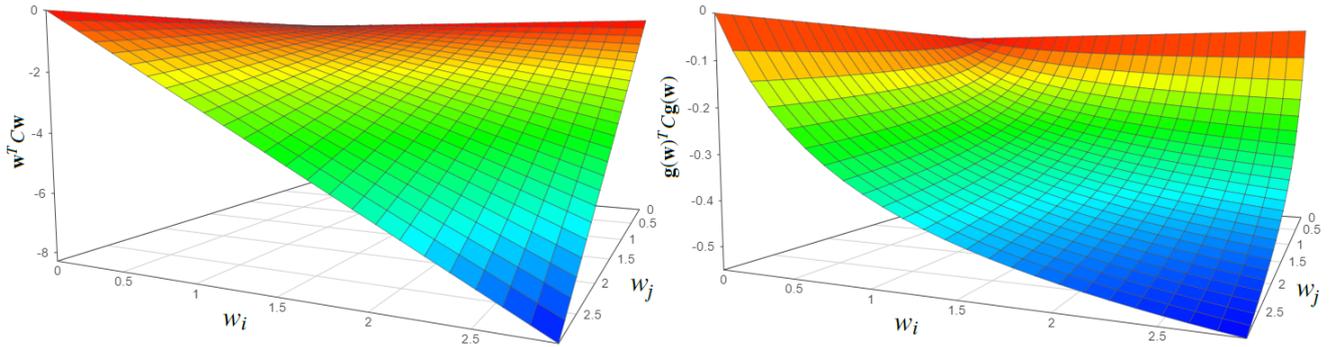


Figure 3: Plots of original and bounded connectivity priors assuming only two neighbouring edges,  $i$  and  $j$ . Connectivity term incentives are plotted across a range of values for the weights of the two edges,  $w_i$  and  $w_j$ . In both cases, the prior gives increased incentive (i.e., negative contribution) when both of the neighbouring edges have stronger weights. However, the original connectivity prior (left) provides unbounded incentive whereas the modified connectivity prior (right) is bounded from below at -1.

count the number of tracts connecting each pair of regions. See [Brown et al. \(2015\)](#) for a more detailed description of the connectome construction process.

Cognitive and neuromotor function of each infant was assessed at 18 months of age, corrected for prematurity, using the Bayley Scales of Infant and Toddler Development, Third Edition (Bayley-III) ([Bayley \(2006\)](#)). The scores are normalized with mean of 100 and standard deviation of 15; we considered adverse motor or cognitive outcomes as scores at or below 85 (i.e.,  $\leq 1$  std.). Table 1 details the demographics of each outcome group.

## 2.6. ABIDE Dataset

We also validated our method on a dataset of resting-state functional brain networks from 1112 subjects in the ABIDE database, aggregated from 16 imaging sites and made freely available ([Craddock et al. \(2013\)](#)). Data was accessed through the Preprocessed Connectomes Project website<sup>2</sup>, which offers ABIDE data, preprocessed using a number of different pipelines. The data used in this study was preprocessed using the connectome computation system ([Xu et al. \(2015\)](#)) with band-pass filtering and global signal regression. fMRIs were registered to the MNI152 template space and the time series were averaged within 90 regions defined in the AAL atlas (i.e., the same regions defined in the UNC infant atlas used for the preterm infant dataset). Given an averaged time series for each atlas region, Pearson’s correlation was computed between each pair of regions to construct the functional connectome. Following [Yoldemir et al. \(2015\)](#), negative correlation values were set to zero. Ninety nine subjects were found to have missing or erroneous data for one or more atlas regions and were discarded from our dataset for a total of 1013 functional connectomes.

Each participant in the ABIDE cohort was designated as a control, or diagnosed with one of 4 autism spectrum categories: 1) Asperger’s; 2) Asperger’s or pervasive developmental disorder not otherwise specified (PDD-NOS); 3) PDD-NOS; 4) autism. The distribution of number of participants in each category is reported in Table 2. Individuals in the dataset had mean

Autism Category	# Scans	Age
Control	539	16.6 $\pm$ 7.2
Asperger’s	87	19.5 $\pm$ 11.3
Asperger’s or PDD-NOS	4	36.2 $\pm$ 14.3
PDD-NOS	35	18.1 $\pm$ 11.9
Autism	348	16.0 $\pm$ 6.8
All	1013	16.8 $\pm$ 7.88

Table 2: Demographics of the used subset of the ABIDE dataset organized by autism spectrum categories. Mean ages (and standard deviations) listed in years.

age 16.8  $\pm$  7.88 years old, though with a wide range of 6.47 to 64 years old.

## 2.7. Validation Measures

To evaluate our proposed method and novel regularization terms and compare our approach to others, we used three primary validation measures including Pearson’s correlation ( $r$ ) between ground truth and predicted scores and the area over the regression error characteristic curve (AOC), which provides an estimate of regression error ([Bi & Bennett \(2003\)](#)). Some previous studies have focused on predicting a binary abnormality label instead of predicting actual scalar outcome scores ([Brown et al. \(2015\)](#); [Ziv et al. \(2013\)](#)). Thus, to compare more directly to these works, we also evaluate the accuracy (acc) of our models as a binary classifier. For the preterm infants, abnormality is defined as any scores below 85 (i.e., one standard deviation below the mean of 100) and for the ABIDE data, abnormality is defined as any class that is not control. Similar to [Brown et al.](#), an SVM was used to classify normal from abnormal instances as it was found to perform better than thresholding the predicted scores or autism spectrum categories. SVM learns a max-margin threshold for the predicted output (i.e., one input feature), optimal for classification over the training set.

## 3. Results

We followed a similar procedure for training models and predicting labels on both preterm neonate and ABIDE datasets.

<sup>2</sup><http://preprocessed-connectomes-project.org/>

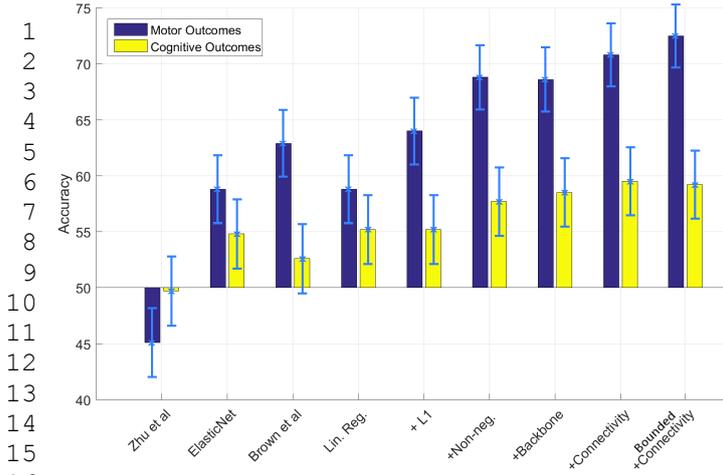


Figure 4: Accuracies of different prediction methods reported in Table 3 along with 95% confidence intervals.

For each method (both proposed and competing), coarse grid searches were performed in powers of two over the method’s hyper-parameters to find the best performance (e.g., for the proposed method, this search was over  $\lambda_{L1}, \lambda_C, \lambda_B \in \{2^0, \dots, 2^9\}$ ). A finer grid search was not performed to avoid over-fitting to the dataset.

### 3.1. Prediction of Neurodevelopmental Outcomes from Preterm Connectomes

On the preterm neonate dataset, our goal was to predict both cognitive and motor neurodevelopmental scores assessed at 18 months of age, corrected for prematurity. For each setting of the model hyper-parameters, a leave-2-out, 1000-round cross validation test was performed. If two scans were of the same infant, those scans were not split between test and training sets.

The preterm neonate dataset is imbalanced, containing fewer scans of infants with high and low outcome scores. In order to flatten this distribution, the number of connectomes in each training set was doubled by synthesizing instances with high and low outcome scores, using the synthetic minority over-sampling technique [Chawla & Bowyer \(2002\)](#).

Table 3 shows a performance comparison of the different methods tested on the preterm neonate connectomes for prediction of motor and cognitive outcome scores. Accuracies are plotted with 95% confidence intervals in Figure 4. Our proposed method with backbone and (bounded or original) connectivity priors achieved the highest correlations, lowest AOCs and best 2-class classification accuracies for both motor and cognitive scores. For prediction of motor and cognitive outcomes, respectively, the best parameter settings,  $[\lambda_{L1}, \lambda_C, \lambda_B]$ , using the original connectivity prior were  $[2^2, 2^1, 2^6]$  and  $[2^5, 2^2, 2^5]$  and using the bounded connectivity prior were  $[2^4, 2^1, 2^5]$  and  $[2^3, 2^2, 2^1]$ . Note that, beginning with standard linear regression, the correlation values of our proposed model improved as each regularization term was added. All tested methods achieved statistically significant ( $p < 0.05$ ) ground truth/predicted correlation values since, for  $1000 \times 2 = 2000$

total predictions, the threshold for 95% significance is  $r \geq 0.0439$ .

For 2-class classification in particular, the proposed method outperformed our previous method, of training SVM on global measures of network topology and metadata [\(Brown et al. \(2015\)\)](#), by 7.4% on average. Note that the reported accuracies for this network-measure-based method are lower here than what was previously reported because in the current study, we include infants with Bayley-III scores of 85 in the abnormal class, whereas in the previous work, 85 was treated as normal. The proposed regularization terms also outperformed GraphNet [\(Grosenick et al. \(2013\)\)](#) by 7% and 2%, Elastic-Net [\(Munsell et al. \(2015\)\)](#) by 12% and 5%, and the feature extraction method of [Zhu et al. \(2014\)](#) by 26% and 10% higher accuracy on motor and cognitive outcomes, respectively. Other than the performance of the GraphNet-regularized model predicting cognitive outcome scores, we found all of these performance differences to be statistically significant (using a two-proportion z-test,  $p < 0.05$ ).

We also compared our results to those from our previously proposed BrainNetCNN, a convolutional neural network (CNN) model for brain network data [\(Kawahara et al. \(2016\)\)](#). While BrainNetCNN was tested on the same preterm neonate connectome data, only 3-fold cross-validation was used (versus 1000 rounds of leave-2-out) as BrainNetCNN requires much longer to train than the proposed regularized linear model. Thus, while our proposed method outperformed BrainNetCNN in terms of correlation, some of the difference in the results may be due to the discrepancy between validation schemes.

Figs. 5 and 6 display the predictive subnetworks learned by our best performing proposed method (averaged over all rounds of cross validation) as Circos ideograms and rendered spatially as a set of tracts in the brain, respectively. It is reasonable to visualize the structure of the subnetworks averaged over different rounds of cross-validation since they remained stable across rounds: 93.6% of all edges were consistently in or out of the subnetwork 95% of the time (on average for cognitive and motor subnetwork). Notable anatomical features of these subnetworks are discussed in Section 4.

We examined the structure of the learned subnetworks to analyse the effect of the proposed regularization terms. By including the  $L1$  regularization term, the learned subnetworks were very sparse, having an average of 71.6% and 98.2% of edge weights set to zero for motor and cognitive scores, respectively, up from only 6.7% (for either score) without the  $L1$  term. Adding the backbone network prior reduced the number of low SNR edges (i.e.,  $B_{j,j} = 1$ ) by 18.6% percent for motor score prediction and 11.2% for cognitive score prediction. Adding the connectivity prior improved subnetwork efficiencies (a measure of network integration described in [Brown et al. \(2014\)](#)) by a factor of 6.8 (from 0.0059 to 0.0403) and 2.2 (from 0.2807 to 0.6215) for subnetworks predictive of motor and cognitive scores, respectively. In Section 4 we discuss notable features of these subnetworks.

As was hypothesized in Section 1.2 the Laplacian term of the GraphNet regularized model discouraged sparsity. None of the edge weights in either the predictive motor outcome or cogni-

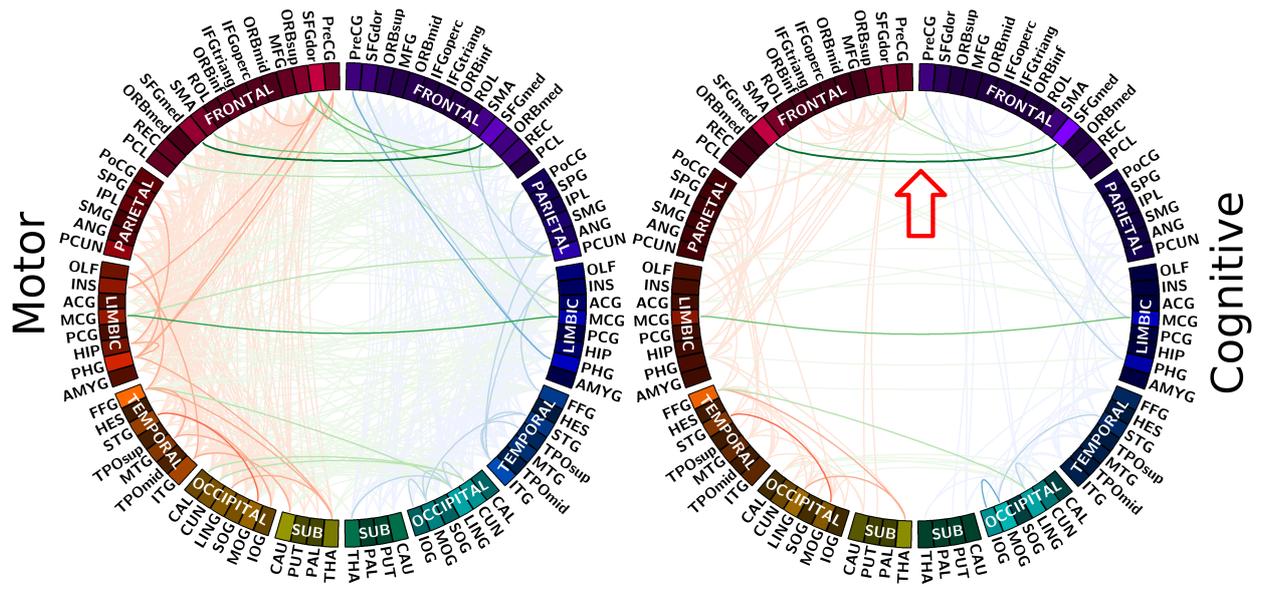


Figure 5: Circos ideograms for the preterm neonate structural connectome subnetworks best predictive of motor (left) and cognitive (right) neurodevelopmental outcomes, averaged over 1000 rounds of cross-validation. Thickness of a curve represents relative strength of that connection in the subnetwork. Colour denotes whether a curve is inter-hemispheric (green) or intra-hemispheric on the left (red) or right (blue) side of the brain. The brightness of each region indicates the relative connectedness. The red arrow highlights the especially strong connection between the left and right medial superior frontal gyri in the cognitive subnetwork.

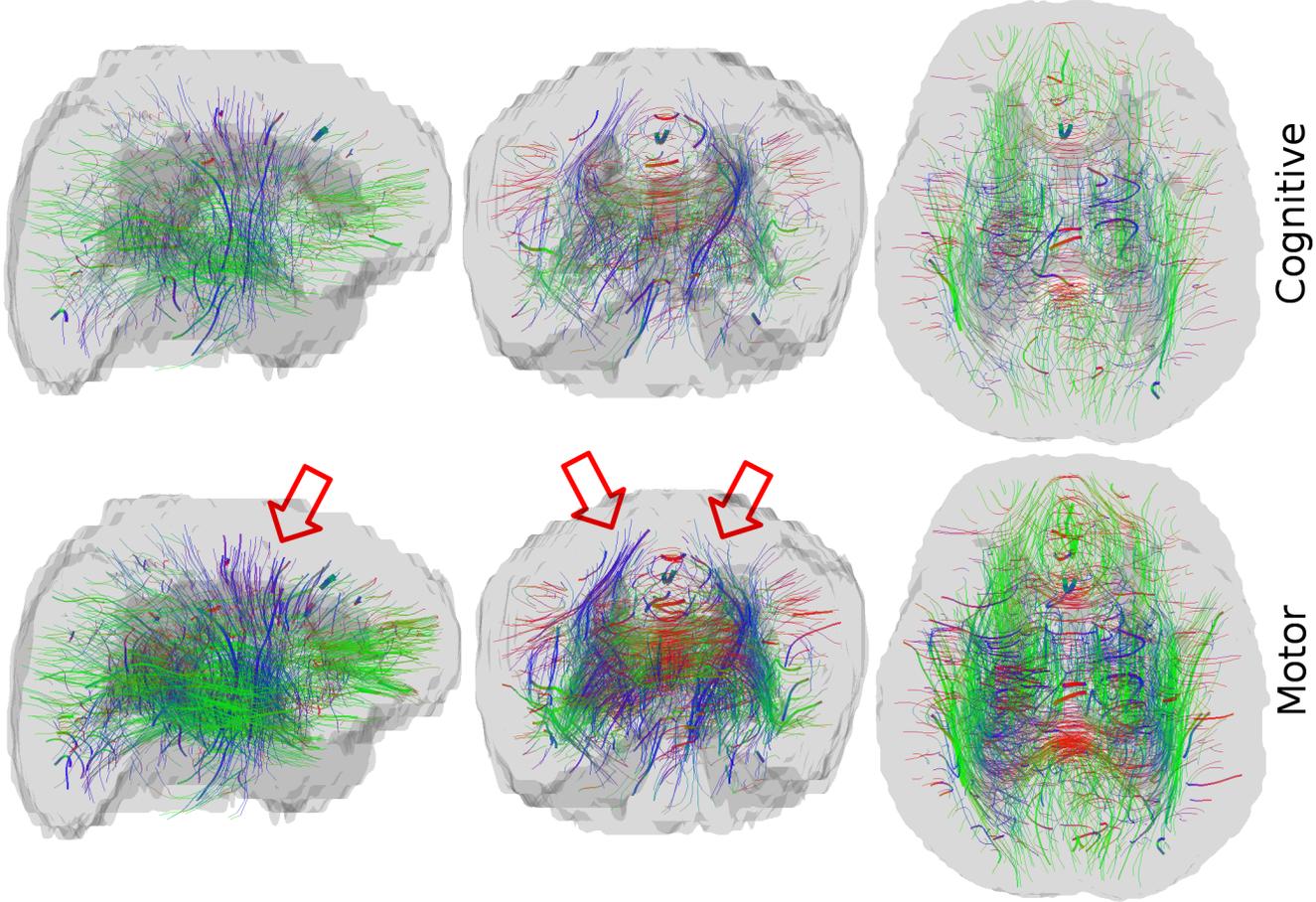


Figure 6: Preterm infant structural connectome subnetworks best predictive of cognitive (top row) and motor (bottom row) Bayley-III scores. Subnetwork edges are mapped on to the mean shape of all tracts between those connected regions (from one representative subject's scan). The thickness of each curve represents the relative edge weight. The red arrows locate the superior terminations of the corticospinal tract, which are relatively strong in the predictive motor subnetwork.

Method	Motor			Cognitive		
	r	AOC	acc.	r	AOC	acc.
T-test + CFS (Zhu <i>et al.</i> (2014))	0.159	27.39	45.1	0.021	28.05	49.7
Elastic Net (Munsell <i>et al.</i> (2015))	0.270	24.58	58.8	0.207	24.83	54.8
GraphNet (Grosenick <i>et al.</i> (2013))	0.326	23.30	63.5	0.256	25.72	57.5
SVM + Network Measures (Brown <i>et al.</i> (2015))	-	-	62.9	-	-	52.6
BrainNetCNN* (Kawahara <i>et al.</i> (2016))	0.310	-	-	0.188	-	-
Linear Regression	0.270	24.78	58.8	0.245	24.72	55.2
+ L1 regularization	0.314	18.55	64.0	0.244	24.75	55.2
+ Non-neg. Constraint	0.433	14.53	68.8	0.317	17.73	57.7
+ Backbone Prior	0.436	14.47	68.6	0.327	17.82	58.5
+ Connectivity Prior	<b>0.442</b>	14.25	70.8	0.343	17.38	<b>59.5</b>
or Bounded Connectivity	0.441	<b>14.01</b>	<b>72.5</b>	<b>0.442</b>	<b>15.37</b>	59.0

Table 3: Prediction of 18-month Bayley-III neurodevelopmental outcomes from structural connectomes of preterm neonates. Correlation ( $r$ ) between ground-truth and predicted scores, area over REC curve (AOC) values and classification accuracy of scores at or below 85 (acc.) for each model, assessed via 1000 rounds of leave-2-out cross validation (\* except BrainNetCNN which was validated with 3-fold cross validation). Note that our previous method (Brown *et al.* (2015)) performed binary classification only. BrainNetCNN results are from our previous paper in which the model was only validated as a regressor (i.e., no two-class accuracy) and not measured in terms of AOC (Kawahara *et al.* (2016)).

Method	r	AOC	acc.
T-test + CFS (Zhu <i>et al.</i> (2014))	0.213	1.952	58.2
Elastic Net (Munsell <i>et al.</i> (2015))	0.421	1.490	64.3
GraphNet (Grosenick <i>et al.</i> (2013))	0.405	1.542	64.8
Linear Regression	0.328	1.640	59.8
+ L1 regularization	0.414	<b>1.464</b>	63.7
+ Backbone Prior	0.414	<b>1.464</b>	64.0
+ Connectivity Prior	0.419	1.493	<b>66.7</b>
or Bounded Connectivity	<b>0.423</b>	1.484	64.6
+ Non-neg. Constraint	0.294	1.728	63.5

Table 4: Comparison of different models for the prediction of autism spectrum categories from functional connectomes in the ABDIE dataset. Pearson’s correlation ( $r$ ) between ground-truth and predicted categories, area over REC curve (AOC) values and binary classification accuracy of controls versus ASD are reported for each model, assessed via 10-fold cross validation.

tive outcome subnetworks learned by the GraphNet regularized model were set to zero. In contrast, an average of 28% of the edge weights in the subnetworks learned with the Elastic Net regularizer were set to zero.

### 3.2. Prediction of autism spectrum category from ABIDE connectomes

For the ABIDE dataset, we posed the task of predicting an autism category from a functional connectome as a regression problem, attempting to predict a continuous, approximate category value from 0 (control) to 4 (full autism). Because the larger dataset (1013 scans versus 168 for preterm neonate dataset) allows for more training data per round, we used a 10-fold cross validation scheme.

Table 4 shows that the best correlation between predicted and ground-truth autism spectrum categories improves as the L1, backbone network and connectivity regularization terms are added to the training objective function. Replacing the connectivity prior term with the bounded connectivity prior

term improves this measure even further, providing the best predicted/ground-truth correlation value of  $r = 0.423$ . However, in contrast to the results on the preterm neonate structural connectomes, we found that the non-negativity constraint generally reduced predictive performance on this dataset (and thus, in Table 4, we report results with the non-negativity constraint enforced last).

Binary prediction of controls versus ASD also exhibited progressively increasing accuracy with the additions of the L1, backbone network and connectivity regularization terms, achieving a top accuracy of 66.7% when all the priors were included. The bounded connectivity term, however, does not outperform the original connectivity term in terms of this measure. Compared to the results on the infant data, Elastic Net and GraphNet regularized models performed relatively well, nearly matching our top performing proposed models with respect to correlation and 2-class accuracy. In contrast, using a t-test and then CFS to define the subnetworks performs relatively poorly, as it did for the infant data.

Similarly to the subnetwork extracted by our method for the preterm neonate data, functional connectome subnetworks predictive of autism categories, learned by our best performing proposed method (i.e., with the bounded connectivity term), were stable across folds: 80.4% of edges were consistently zero or non-zero in all 10 folds. These subnetworks were also sparse: 75.6% of edges in the mean subnetwork, averaged across the 10 rounds of cross validation, were set to zero. Further, only 10% of the 2474 edges penalized by the backbone network prior were non-zero in this subnetwork. This subnetwork is visualized as a Circos ideogram in Fig. 7 and spatially as edges connecting regions in the brain in Fig. 8. Note that we could not render Fig. 8 as fiber bundle curves (as was done in Fig. 6) since the ABIDE dataset did not contain the dMRI data to reconstruct the white matter fibers.

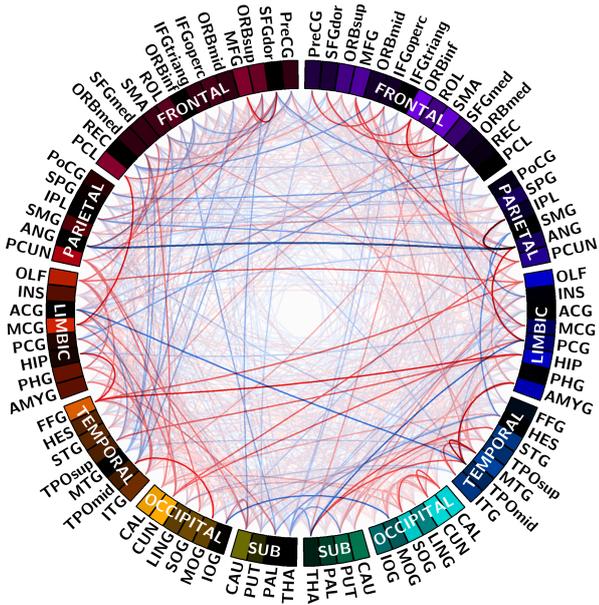


Figure 7: Circos ideogram representing the subnetwork of weights learned by our top performing model on the functional connectome ABIDE data, averaged across 10 folds. Red curves represent positive weights and blue curves represent negative weights. The thickness of a curve represents the relative magnitude of the weight on that connection.

#### 4. Discussion

The experiments on both datasets confirmed our hypothesis that including prior knowledge about the organization of the brain into our learning model, helps to regularize that model and enables it to predict with higher accuracy. Testing on both the preterm neonate structural connectome and the functional ABIDE connectomes, it was found that the structure of the learned subnetworks was stable across rounds of cross-validation. The stability of the subnetworks suggests that the models were not over-fitting to the training data or noise but instead were selecting groups of connections that effect or are affected by ASD or neurodevelopmental outcomes.

In Fig. 1, the backbone networks of the two datasets were visualized, exposing those connections that are most consistent across subjects in the two cohorts. The edges of the preterm neonate backbone network prior (Fig. 1a-d) are distributed similarly to what we would expect in a healthy young infant: The majority of the connections are intra-hemispherical, cortico-cortico connections with only a few strong inter-hemispherical connections and a reasonable degree of bilateral symmetry (Brown et al. (2014)). Fig. 1e shows that the backbone network of the functional connectomes in the ABIDE dataset are also bilaterally symmetric and include many strong cortico-cortico (functional) connections. In contrast to the preterm infant structural connectome, the backbone network for the ABIDE data also includes many strong connections between left and right hemispheric region pairs, which we expect to find in normative functional connectomes (Just et al. (2007); Meindl et al. (2010)).

In terms of the preterm neonate data specifically, the results

in Table 3 suggest that prediction of cognitive outcomes is more difficult than prediction of motor outcomes. This agrees with the findings in our previous work (Kawahara et al. (2016)). As was mentioned in that paper, the greater challenge in predicting cognitive outcome scores may be due to the fact that cognitive development is known to be more affected by the child’s environment than motor development (Grunau et al. (2009)). A longer term follow-up study of this cohort is currently underway and may improve our understanding of these complex factors.

As expected, the predictive motor subnetwork clearly includes the corticospinal tract and pyramidal tracts (red arrow in Fig. 6). These connections were relatively more heavily weighted in the motor subnetwork versus the cognitive subnetwork. The corticospinal tract and the pyramidal tracts in particular are known to be important for motor function (Asanuma (1981)).

The predictive cognitive subnetwork was more sparse and had generally lower weights than the motor subnetwork (as visualized by less dense, more transparent streamlines), due to the larger  $L1$  weight used for best prediction of the cognitive scores. However, the left and right medial superior frontal gyri (SFGmed) and the connection between these two regions that had stronger weights (factor of 2.1) in the cognitive network than in the motor network, (red arrow in Fig. 6). These regions contain the presupplementary motor area which is thought to be responsible for a range of cognitive functions (Zhang et al. (2012)). Thus, the learned cognitive subnetwork, like the learned motor subnetwork, heavily weights connections and regions that are associated with brain functions important for those outcomes.

When validating the proposed method on the ABIDE functional connectome dataset, it was found that the non-negativity constraint was detrimental to prediction results (Table 4). In contrast, enforcing non-negativity of subnetwork weights improved prediction accuracy on the preterm neonate structural connectome dataset. This is likely due to both the differences in the modality from which the connectomes were constructed and in what is being predicted. It has been reported in the literature that functional connectivity in subjects with ASD can be both decreased and increased between different regions of the brain, as compared to normal controls (Mizuno et al. (2006); Noonan et al. (2009); Salmi et al. (2013)). Thus, to predict ASD, the subnetwork model may learn to weight certain connections positively and others negatively (Figure 8). Conversely, causes of adverse neurodevelopmental outcomes that affect structural white matter connectivity (e.g., injury and infection, Back & Miller (2014)) are predominantly associated with reduced connectivity (Chau et al. (2013)), and so a reasonable predictive model can be composed of only non-negative edge weights.

In Figs. 7 and 8 the negative subnetwork weights are visualized as blue edges and are weights that reduce the predicted autism spectrum category (i.e., down towards 0 for controls) for stronger functional connectivity between two regions in a given subject. Thus, these are edges that we expect to typically have reduced functional connectivity in ASD subjects. Two of the strongest negatively weighted edges were found to be be-

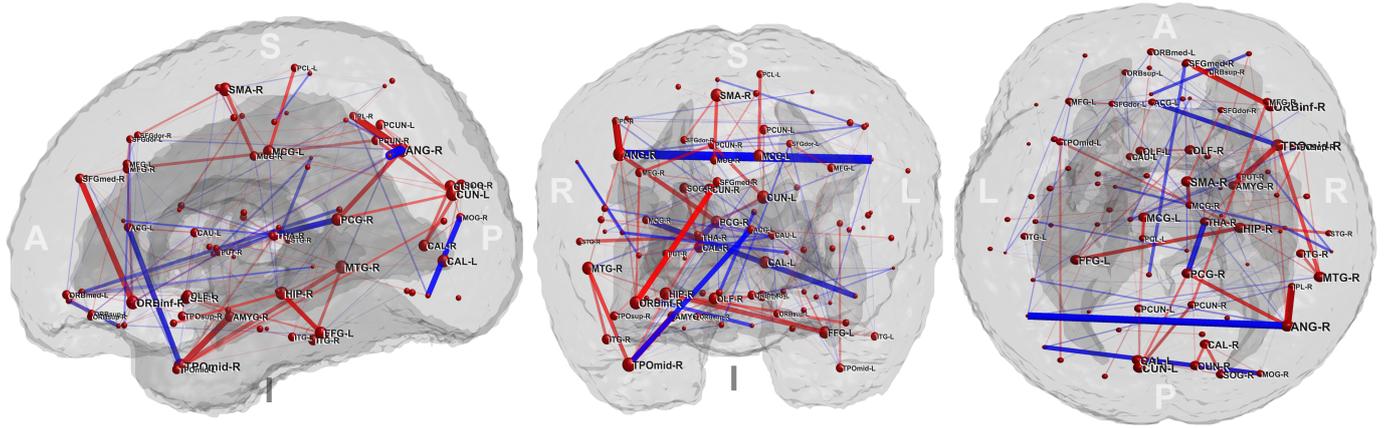


Figure 8: Functional subnetworks best predictive of autism spectrum categories, averaged across 10 folds and rendered spatially on the brain from sagittal (left), coronal (middle) and axial (right) views. Positive and negative subnetwork edge weights are represented by red and blue edges respectively. Edge thickness represents relative weight magnitude. Node size represents relative weighted nodal degree and the names of brain regions associated with each node are listed for those nodes with the highest degrees.

tween the left and right angular gyri (ANG-L and ANG-R) and from the left anterior cingulate gyrus (ACG-L) to the right middle temporal pole (TPO-mid). This aligns with the findings of Salmi *et al.* who found that ASD subjects had reduced functional connectivity from the left angular gyrus and from the left anterior cingulate cortex (among other regions) (Salmi *et al.* (2013)).

Previous studies have investigated prediction of ASD using the ABIDE dataset. For instance, Abraham *et al.* examined binary classification of ASD versus controls using a variety of different processing pipelines and learning models (Abraham *et al.* (2016)). Using a 10-fold cross validation strategy, they found that their top prediction pipelines achieved accuracies of 65.7% to 67.9%. These results are comparable to our best achieved accuracy of 66.7%. They also reported that choices of data subset (e.g., only data from particular sites), parcellation technique and functional covariance measure had significant impacts on prediction accuracy. While the primary purpose of our experiments were to show the relative improvements of including our additional priors, it seems plausible that our best prediction accuracy on this dataset could likely be improved by fine tuning the preprocessing procedure. Also note that Abraham *et al.* also had more rigorous quality control criteria and inspected each scan for visual quality, ultimately selecting only 871 subjects of the original 1112 (versus 1013 used by our study). This suggests that the dataset used by our study represents a broader range of image qualities, which likely makes accurate prediction more difficult.

While the proposed method performed better than competing methods on both prediction tasks, neither task is yet near to being satisfactorily solved (e.g., 95-99% accuracy). One limitation of our method may be that the linearity of the model cannot capture the complexities of the relationships between connectome data and target labels. Thus, a key future area of research is to investigate whether our proposed regularization terms can be applied to more complex, non-linear machine learning models. While BrainNetCNN, the only deep model we tested (Ta-

ble 3), performed relatively poorly, we expect that these results could be greatly improved by regularizing the model with the proposed priors and by reducing the number of learnable parameters per layer, especially on larger datasets like the ABIDE dataset.

Another limitation of our learning model is that it does not leverage any metadata features that may be available at the time of prediction (e.g., subject age, sex, and other clinical variables). Investigation into how to best expand the proposed model to incorporate non edge-wise features may enable further improvements to its predictive performance.

Finally, there is an established link between adverse cognitive outcomes in infants born very preterm and ASD (Johnson *et al.* (2010)). Because the same AAL atlas was used for constructing connectomes from both datasets, there is an opportunity to explore the relationship between structural subnetworks predictive of cognitive outcomes and functional subnetworks predictive of ASD. In particular, we examined the intersection set of region-pairs that were learned by our model to be both i) positively correlated with normal cognitive neurodevelopmental outcomes in structural connectomes (Fig. 5 right) and ii) positively correlated with control (instead of with ASD) labels in functional connectomes (blue connections in Fig. 7, which have negative edge weights). Note that we do not examine edge weights that are positively correlated with ASD labels (i.e., positive edge weights), since there are no equivalent edge weights (predictive of abnormal cognitive outcomes) in the cognitive outcome subnetwork due to the non-negativity constraint enforced in that optimization. The edge weights of this intersection set subnetwork are defined by taking the product of the normalized edge weight magnitudes of i) the subnetwork most predictive of cognitive outcomes on the preterm neonate dataset and ii) the negatively weighted edges of the subnetwork most predictive of ASD on the ABIDE dataset. In both cases, these subnetworks were averaged over all rounds of their respective cross-validation schemes.

Fig. 9 shows that 173 shared, non-zero edges were discov-

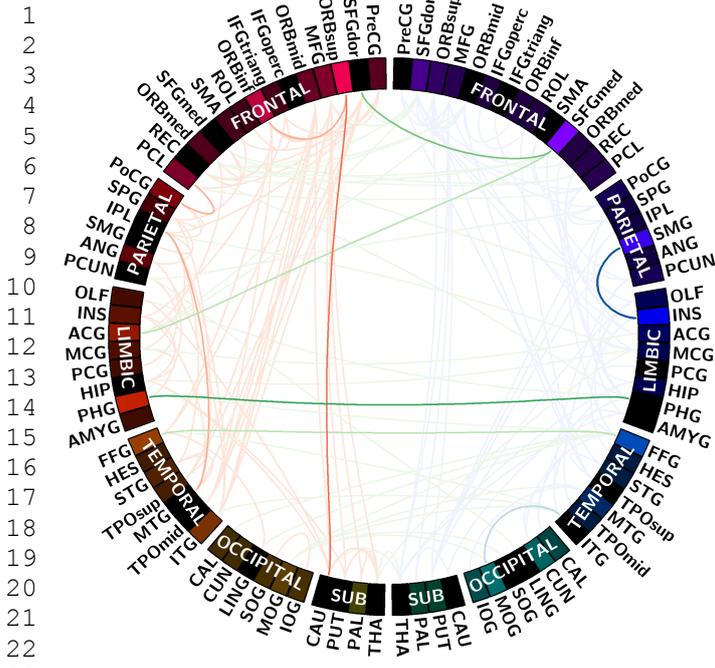


Figure 9: Common connections found in both the structural subnetwork best predictive of cognitive neurodevelopmental outcomes and the negative weights of the functional subnetwork best predictive of ASD. The thickness of a curve represents the product of the normalized weight magnitudes of that edge between the two subnetworks.

ered in the intersection set. The two edges with the highest multiplied weight include a connections between the right insula (INS-R) and right supramarginal gyrus (SMG-R) as well as a connection between the left and right parahippocampal gyri (PHG-L and PHG-R). Notably, PHG-L, PHG-R and SMG-R have all been found to have significantly altered functional connectivity in ASD subjects performing language comprehension tasks and the PHG-R, in particular, has been found to be associated with understanding sarcasm and social context (Just *et al.* (2004); Rankin *et al.* (2009)).

These results, however, must be interpreted carefully. While functional connectivity is well correlated with structural connectivity, functional connectivity fluctuates in time and can be observed between regions with no structural connection (Honey *et al.* (2009b)). Additionally, while the atlas used for both datasets comprises the same AAL defined brain regions, the preterm neonate brains are imaged at an early stage in development when the structural connectivity is changing rapidly (Brown *et al.* (2014)). Thus, it is possible that certain connections observed in the neonate brains will weaken or disappear as they develop towards the age group represented in the ABDIE cohort. More exploration of the relationships between preterm birth and ASD, structural and functional connectomes, and neonate and adult brain structures is required to address these issues.

## 5. Conclusions

We proposed a framework for learning subnetworks of structural or functional connectomes that are predictive of targeted clinical variables. We found that by introducing our novel network backbone prior, the learned subnetworks were robust to noise and included few edges with low SNR weights. By including either of the two proposed connectivity priors, the learned subnetworks were highly integrated, a property we expect for subnetworks pertinent to specific outcomes or disorders. Compared to other methods, our approach achieved the best accuracies for identifying patients with ASD, and predicting cognitive and motor neurodevelopmental outcomes of preterm neonates. Edges and regions learned to be most important for predicting labels in both datasets fit well with established neuroscience on neurodevelopment and ASD. We also discovered a set of connections that predicted both ASD in functional connectomes and adverse cognitive neurodevelopmental outcomes in structural connectomes. Ongoing follow-up of the preterm cohort will allow us to determine the longer-term significance of these findings. In future work, we plan to apply the proposed regularization terms to deep, nonlinear learning models and to leverage a wider variety of features in order to predict neurological disorders and neurodevelopmental outcomes with even higher accuracy.

## 6. Acknowledgements

We thank NSERC, CIHR (MOP-79262: S.P.M. and MOP-86489: R.E.G.), the Canadian Child Health Clinician Scientist Program, Kids Brain Health Network and the Michael Smith Foundation for Health Research for their financial support. S.P.M. is currently the Bloorview Children’s Hospital Chair in Pediatric Neuroscience. R.E.G. is supported by a senior investigator salary award from the BC Children’s Hospital Research Institute. The Titan X Pascal used for this research was donated by the NVIDIA Corporation.

## References

## References

Abraham, Alexandre, Milham, Michael, Di, Adriana, & Craddock, R Cameron. 2016. Deriving reproducible biomarkers from multi-site resting-state data : An Autism-based example. *NeuroImage*, **147**, 736–745.

An, Michael, Ho, Hon Pong, Staib, Lawrence, Pelphrey, Kevin, & Duncan, James. 2010. Multimodal MRI analysis of brain subnetworks in autism using multi-view EM. *Conference Record - Asilomar Conference on Signals, Systems and Computers*, 786–789.

Anderson, Ariana, Douglas, Pamela K., Kerr, Wesley T., Haynes, Virginia S., Yuille, Alan L., Xie, Jianwen, Wu, Ying Nian, Brown, Jesse a., & Cohen, Mark S. 2014. Non-negative matrix factorization of multimodal MRI, fMRI and phenotypic data reveals differential changes in default mode subnetworks in ADHD. *NeuroImage*, **102**, 207–219.

Asanuma, Hiroshi. 1981. The pyramidal tract. *Comprehensive Physiology*.

Assaf, Michal, Jagannathan, Kanchana, Calhoun, Vince D., Miller, Laura, Stevens, Michael C., Sahl, Robert, O’Boyle, Jacqueline G., Schultz, Robert T., & Pearlson, Godfrey D. 2010. Abnormal functional connectivity of default mode sub-networks in autism spectrum disorder patients. *NeuroImage*, **53**(1), 247–256.

- 1 Back, Stephen a., & Miller, Steven P. 2014. Brain injury in premature neonates:  
2 A primary cerebral dysmaturational disorder? *Annals of Neurology*, **75**(4),  
3 469–486.
- 4 Bassett, Danielle S., & Bullmore, Edward T. 2009. Human brain networks in  
5 health and disease. *Current Opinion in Neurology*, **22**(4), 340–347.
- 6 Bayley, N. 2006. *Bayley Scales of Infant Development - Third Edition*. San  
7 Antonio: Harcourt.
- 8 Bi, Jinbo, & Bennett, Kristin P. 2003. Regression Error Characteristic Curves.  
9 *Proceedings of the Twentieth International Conference on Machine Learning  
10 (ICML-2003)*, 43–50.
- 11 Brown, Colin J., & Hamarneh, Ghassan. 2016. Machine Learning on Human  
12 Connectome Data from MRI. *arXiv preprint arXiv:1611.08699*.
- 13 Brown, Colin J., Miller, Steven P., Booth, Brian G., Andrews, Shawn, Chau,  
14 Vann, Poskitt, Kenneth J., & Hamarneh, Ghassan. 2014. Structural network  
15 analysis of brain development in young preterm neonates. *NeuroImage*, **101**,  
16 667–680.
- 17 Brown, Colin J., Miller, Steven P., Booth, Brian G., Poskitt, Kenneth J., Chau,  
18 Vann, Synnes, Anne R., Zwicker, Jill G., Grunau, Ruth E., & Hamarneh,  
19 Ghassan. 2015. Prediction of Motor Function in Very Preterm Infants Us-  
20 ing Connectome Features and Local Synthetic Instances. *Pages 69–76 of:  
21 MICCAI 2015*.
- 22 Brown, Colin J., Miller, Steven P., Booth, Brian G., Zwicker, Jill G., Grunau,  
23 Ruth E., Synnes, Anne R., Chau, Vann, & Hamarneh, Ghassan. 2016. Predic-  
24 tive Subnetwork Extraction with Structural Priors for Infant Connectomes.  
25 *Pages 175–183 of: MICCAI 2016*.
- 26 Calhoun, Vince D., Kiehl, Kent A., & Pearlson, Godfrey D. 2008. Modulation of  
27 temporally coherent brain networks estimated using ICA at rest and during  
28 cognitive tasks. *Human brain mapping*, **29**(7), 828–838.
- 29 Casanova, R., Whitlow, CT, Wagner, B., Espeland, MA, & Maldjian, JA. 2012.  
30 Combining graph and machine learning methods to analyze differences in  
31 functional connectivity across sex. *The Open Neuroimaging Journal*, **6**(1).
- 32 Castellanos, F. Xavier, Di Martino, Adriana, Craddock, R. Cameron, Mehta,  
33 Ashesh D., & Milham, Michael P. 2013. Clinical applications of the func-  
34 tional connectome. *NeuroImage*, **80**, 527–540.
- 35 Chang, L-C, Jones, D K, & Pierpaoli, C. 2005. RESTORE: Robust Estima-  
36 tion of Tensors by Outlier Rejection. *Magnetic Resonance in Medicine*, **53**,  
37 1088–1095.
- 38 Chau, V, Synnes, A, Grunau, R E, Poskitt, K J, Brant, R, & Miller, S P. 2013.  
39 Abnormal brain maturation in preterm neonates associated with adverse de-  
40 velopmental outcomes. *Neurology*, **81**(24), 2082–2089.
- 41 Chawla, Nv, & Bowyer, Kw. 2002. SMOTE: synthetic minority over-sampling  
42 technique. *arXiv preprint arXiv: . . .*, **16**, 321–357.
- 43 Cheng, Hu, Wang, Yang, Sheng, Jinhua, Kronenberger, William G., Mathews,  
44 Vincent P., Hummer, Tom a., & Saykin, Andrew J. 2012. Characteristics  
45 and variability of structural networks derived from diffusion tensor imaging.  
46 *NeuroImage*, **61**(4), 1153–1164.
- 47 Craddock, C, Benhajali, Y, Chu, C, Chouinard, F, Evans, A, Jakob, A, Khun-  
48 drakpam, B S, Lewis, J D, Li, Q, Milham, M, Yan, C, & Bellec, P. 2013.  
49 The Neuro Bureau Preprocessing Initiative: open sharing of preprocessed  
50 neuroimaging data and derivatives. *Frontiers in Neuroinformatics*.
- 51 de Reus, Marcel a, Saenger, Victor M, Kahn, René S, & van den Heuvel, Mar-  
52 tijn P. 2014. An edge-centric perspective on the human connectome: link  
53 communities in the brain. *Philosophical transactions of the Royal Society  
54 of London. Series B, Biological sciences*, **369**(1653), 20130527–.
- 55 Fair, Damien A, Cohen, Alexander L, Power, Jonathan D, Dosenbach, Nico UF,  
56 Church, Jessica A, Miezin, Francis M, Schlaggar, Bradley L, & Petersen,  
57 Steven E. 2009. Functional brain networks develop from a local to dis-  
58 tributed organization. *PLoS computational biology*, **5**(5), e1000381.
- 59 Fallani, Fabrizio De Vico, Richiardi, Jonas, Chavez, Mario, & Achard, Sophie.  
60 2014. Graph analysis of functional brain networks: practical issues in trans-  
61 lational neuroscience. *Phil. Trans. R. Soc. B*, **369**(1653), 20130521.
- 62 Ghanbari, Yasser, Smith, Alex R., Schultz, Robert T., & Verma, Ragini. 2014.  
63 Identifying group discriminative and age regressive sub-networks from DTI-  
64 based connectivity via a unified framework of non-negative matrix factoriza-  
65 tion and graph embedding. *Medical Image Analysis*, **18**(8), 1337–1348.
- 66 Grosenick, Logan, Klingenberg, Brad, Katovich, Kiefer, Knutson, Brian, &  
67 Taylor, Jonathan E. 2013. Interpretable whole-brain prediction analysis with  
68 GraphNet. *NeuroImage*, **72**(2), 304–321.
- 69 Grunau, R E, Whitfield, M F, Petrie-Thomas, J, Synnes, A R, Cepeda, I L,  
70 Keidar, A, Rogers, M, MacKay, M, Hubber-Richard, P, & Johannesen, D.  
71 2009. Neonatal pain, parenting stress and interaction, in relation to cognitive  
72 and motor development at 8 and 18 months in preterm infants. *Pain*, **143**(1),  
73 138–146.
- 74 Hall, Mark, & Smith, Lloyd a. 1999. Feature Selection for Machine Learning  
75 : Comparing a Correlation-based Filter Approach to the Wrapper CFS :  
76 Correlation-based Feature. *International FLAIRS Conference*, **1999**, 235–  
77 239.
- 78 He, Lili, Li, Hailong, Holland, Scott K, Yuan, Weihong, Altaye, Mekibib,  
79 & Parikh, Nehal A. 2018. Early prediction of cognitive deficits in very  
80 preterm infants using functional connectome data in an artificial neural net-  
81 work framework. *NeuroImage: Clinical*, **18**, 290–297.
- 82 Hoerl, Arthur E, & Kennard, Robert W. 1970. Ridge Regression: Biased Esti-  
83 mation for Nonorthogonal Problems. *Technometrics*, **12**(1), 55–67.
- 84 Honey, C J, Honey, C J, Sporns, O, Sporns, O, Cammoun, L, Cammoun, L,  
85 Gigandet, X, Gigandet, X, Thiran, J P, Thiran, J P, Meuli, R, Meuli, R,  
86 Hagmann, P, & Hagmann, P. 2009a. Predicting human resting-state func-  
87 tional connectivity from structural connectivity. *Proceedings of the National  
88 Academy of Sciences of the United States of America*, **106**(6), 2035–40.
- 89 Honey, CJ, Sporns, O, Cammoun, Leila, Gigandet, Xavier, Thiran, Jean-  
90 Philippe, Meuli, Reto, & Hagmann, Patric. 2009b. Predicting human resting-  
91 state functional connectivity from structural connectivity. *Proceedings of the  
92 National Academy of Sciences*, **106**(6), 2035–2040.
- 93 Johnson, Samantha, Hollis, Chris, Kochhar, Puja, Hennessy, Enid, Wolke, Di-  
94 eter, & Marlow, Neil. 2010. Autism spectrum disorders in extremely preterm  
95 children. *The Journal of pediatrics*, **156**(4), 525–531.
- 96 Just, Marcel Adam, Cherkassky, Vladimir L, Keller, Timothy A, & Minshew,  
97 Nancy J. 2004. Cortical activation and synchronization during sentence  
98 comprehension in high-functioning autism: evidence of underconnectivity.  
99 *Brain*, **127**(8), 1811–1821.
- 100 Just, Marcel Adam, Cherkassky, Vladimir L, Keller, Timothy A, Kana, Ra-  
101 jesh K, & Minshew, Nancy J. 2007. Functional and anatomical cortical  
102 underconnectivity in autism: evidence from an FMRI study of an executive  
103 function task and corpus callosum morphometry. *Cerebral cortex*, **17**(4),  
104 951–961.
- 105 Kawahara, Jeremy, Brown, Colin J., Miller, Steven P., Booth, Brian G., Chau,  
106 Vann, Grunau, Ruth E., Zwicker, Jill G., & Hamarneh, Ghassan. 2016.  
107 BrainNetCNN: Convolutional neural networks for brain networks; towards  
108 predicting neurodevelopment. *NeuroImage*, sep, 1–9.
- 109 Li, Hai, Xue, Zhong, Ellmore, Timothy M., Frye, Richard E., & Wong,  
110 Stephen T. 2012. Identification of faulty DTI-based sub-networks in autism  
111 using network regularized SVM. *Proceedings - International Symposium on  
112 Biomedical Imaging*, **6**, 550–553.
- 113 Li, Xingjuan, Li, Yu, & Li, Xue. 2017. Predicting Clinical Outcomes  
114 of Alzheimers Disease from Complex Brain Networks. *Pages 519–525  
115 of: International Conference on Advanced Data Mining and Applications*.  
116 Springer.
- 117 Meindl, Thomas, Teipel, Stefan, Elmouden, Rachid, Mueller, Sophia, Koch,  
118 Walter, Dietrich, Olaf, Coates, Ute, Reiser, Maximilian, & Glaser, Christian.  
119 2010. Test–retest reproducibility of the default-mode network in healthy  
120 individuals. *Human brain mapping*, **31**(2), 237–246.
- 121 Mizuno, Akiko, Villalobos, Michele E, Davies, Molly M, Dahl, Branelle C,  
122 & Müller, Ralph-Axel. 2006. Partially enhanced thalamocortical functional  
123 connectivity in autism. *Brain research*, **1104**(1), 160–174.
- 124 Munsell, Brent C., Wee, Chong-Yaw, Keller, Simon S., Weber, Bernd, Elger,  
125 Christian, da Silva, Laura Angelica Tomaz, Nesland, Travis, Styner, Martin,  
126 Shen, Dinggang, & Bonilha, Leonardo. 2015. Evaluation of machine learn-  
127 ing algorithms for treatment outcome prediction in patients with epilepsy  
128 based on structural connectome data. *NeuroImage*, **118**, 219–230.
- 129 Ng, Bernard, Silless, Viviana, Varoquaux, Gael, Poline, Jean Baptiste, Thirion,  
130 Bertrand, & Abugharbieh, Rafeef. 2012. Connectivity-informed sparse clas-  
131 sifiers for fMRI brain decoding. *Proceedings - 2012 2nd International Work-  
132 shop on Pattern Recognition in NeuroImaging, PRNI 2012*, 101–104.
- 133 Noonan, Sarah K, Haist, Frank, & Müller, Ralph-Axel. 2009. Aberrant func-  
134 tional connectivity in autism: evidence from low-frequency BOLD signal  
135 fluctuations. *Brain research*, **1262**, 48–63.
- 136 Qiao, Lishan, Zhang, Han, Kim, Minjeong, Teng, Shenghua, Zhang, Limei, &  
137 Shen, Dinggang. 2016. Estimating functional brain networks by incorporat-  
138 ing a modularity prior. *NeuroImage*, **141**, 399–407.
- 139 Rankin, Katherine P, Salazar, Andrea, Gorno-Tempini, Maria Luisa, Sollberger,  
140 Marc, Wilson, Stephen M, Pavlic, Danijela, Stanley, Christine M, Glenn,  
141 Shenly, Weiner, Michael W, & Miller, Bruce L. 2009. Detecting sarcasm  
142 from paralinguistic cues: anatomic and cognitive correlates in neurodegen-

- erative disease. *Neuroimage*, **47**(4), 2005–2015.
- Salmi, J, Roine, U, Glerean, E, Lahnakoski, J, Nieminen-von Wendt, T, Tani, P, Leppämäki, S, Nummenmaa, L, Jääskeläinen, IP, Carlson, S, *et al.* 2013. The brains of high functioning autistic individuals do not synchronize with those of others. *NeuroImage: Clinical*, **3**, 489–497.
- Schmidt, M. 2010. *Graphical model structure learning with l1-regularization*. Ph.D. thesis, University of British Columbia (Vancouver).
- Schmidt, Ruben, de Reus, Marcel A, Scholtens, Lianne H, van den Berg, Leonard H, & van den Heuvel, Martijn P. 2016. Simulating disease propagation across white matter connectome reveals anatomical substrate for neuropathology staging in amyotrophic lateral sclerosis. *Neuroimage*, **124**, 762–769.
- Shi, F, Yap, P-T, Wu, G, Jia, H, Gilmore, J H, Lin, W, & Shen, D. 2011. Infant brain atlases from neonates to 1-and 2-year-olds. *PLoS One*, **6**(4), e18746.
- Tibshirani, Robert. 1994. *Regression Selection and Shrinkage via the Lasso*.
- Tzourio-Mazoyer, Nathalie, Landeau, Brigitte, Papathanassiou, Dimitri, Crivello, Fabrice, Etard, Olivier, Delcroix, Nicolas, Mazoyer, Bernard, & Joliot, Marc. 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*, **15**(1), 273–289.
- Watanabe, Takatori, Kessler, Daniel, Scott, Clayton, Angstadt, Michael, & Sri-pada, Chandra. 2014. Disease prediction based on functional connectomes using a scalable and spatially-informed support vector machine. *NeuroImage*, **96**, 183–202.
- Xu, Ting, Yang, Zhi, Jiang, Lili, Xing, Xiu-Xia, & Zuo, Xi-Nian. 2015. A connectome computation system for discovery science of brain. *Science Bulletin*, **60**(1), 86–95.
- Yoldemir, Burak, Ng, Bernard, Abugharbieh, Rafeef, & Member, Senior. 2015. Stable Overlapping Replicator Dynamics for Brain Community Detection. *IEEE TMI*, **35**(2), 529 – 538.
- Yu, Rensing, Zhang, Han, An, Le, Chen, Xiaobo, Wei, Zhihui, & Shen, Ding-gang. 2017. Connectivity strength-weighted sparse group representation-based brain network construction for MCI classification. *Human brain mapping*, **38**(5), 2370–2383.
- Zhang, Sheng, Ide, Jaime S., & Li, Chiang Shan R. 2012. Resting-state functional connectivity of the medial superior frontal cortex. *Cerebral Cortex*, **22**(1), 99–111.
- Zhu, Dajiang, Shen, Dinggang, Jiang, Xi, Liu, Tianming, & Imaging, Cortical Architecture. 2014. Connectomics Signature For Characterization of Mild Cognitive Impairment and Schizophrenia. *2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI)*, 325–328.
- Ziv, Etay, Tymofiyeva, Olga, Ferriero, Donna M., Barkovich, a. James, Hess, Chris P., & Xu, Duan. 2013. A machine learning approach to automated structural network analysis: Application to neonatal encephalopathy. *PLoS ONE*, **8**(11).
- Zou, H, & Hastie, T. 2005. Regularization and variable selection via the elastic-net. *Journal of the Royal Statistical Society*, **67**, 301–320.