Minimum Partial Correlation: An Accurate and Parameter-free Measure of Functional Connectivity in fMRI

Lei Nie[1,2], Xian Yang[1], Paul M. Matthews[3], Zhiwei Xu[2], and Yike Guo[1,*]

[1] Department of Computing, Imperial College London, London, United Kingdom
[2] Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China
[3] Department of Medicine, Imperial College London, London, United Kingdom
[*]Correspondence: y.guo@imperial.ac.uk

Abstract. Functional connectivity, a data-driven modelling of spontaneous fluctuations in activity in spatially segregated brain regions, has emerged as a promising approach to generate hypotheses and features for prediction. The most widely used method for inferring functional connectivity is full correlation, but it cannot differentiate direct and indirect effects. This disadvantage is often avoided by fully partial correlation, but this method suffers from Berksons paradox. Some advanced methods, such as regularised inverse covariance and Bayes nets, have been applied. However, the connectivity inferred by these methods usually depends on crucial parameters. This paper suggests minimum partial correlation as a parameter-free measure of functional connectivity in fMRI. An algorithm, called elastic PC-algorithm, is designed to approximately calculate minimum partial correlation. Our experimental results show that the proposed method is more accurate than full correlation, fully partial correlation, ICOV, network deconvolution algorithm and global silencing algorithm in most cases.

Keywords: fMRI, functional connectivity, partial correlation, PC-algorithm

1 Introduction

Functional connectivity, which is defined as measurable temporal dependences among spatially segregated brain areas, does not rest on any generative models about brains [10]. Although functional connectivity does not necessarily reveal mechanisms about how the brain works, it is expected to imply true links (i.e. direct causal effects) between regions of interest as accurately as possible. Such studies are motivated by the belief that more accurate descriptions of functional connectivity may generate more reliable hypotheses of brain organizations [3] and provide a better basis for predictions [6], such as in decoding cognitive states [18], or in evaluating brain disease or its severity [12].

The accuracy of functional connectivity is typically evaluated from three aspects: the skeleton, directionality and strength. This paper focuses on inferring the skeleton, which is to determine whether a link between two ROIs (i.e.

nodes) exists or not. Smith et al. [19] systematically investigated the accuracy of many methods for skeleton inference using 28 synthetic networks and their corresponding BLOD signals generated from the dynamic causal model [11]. Their results show that the top-3 methods are inverse covariance (ICOV), fully partial correlation and Bayes net, which are followed by full correlation.

Among the above four methods, full correlation is the most widely used method. It has been demonstrated that full correlation offers deep insights into brain structures [13] and powerful features for brain decoding [22]. Full correlation is robust and parameter-free, but there is an intrinsic limitation to differentiation of direct and indirect effects. Fig. 1(A) shows an example of differentiating direct and indirect effects in the network where Node 1 $\rightarrow$ Node 2$\rightarrow$ Node 3. The full correlation between Node 1 and Node 3 is as high as 1.0, but there is no real link between the two nodes.

Fully partial correlation [15] often successfully identifies the direct effects in Fig. 1(A) by controlling the intermediate node. The fully partial correlation of two nodes is their correlation when all other nodes in the network are controlled. It is worth noting that "fully partial correlation" is called "partial correlation" in previous studies about functional connectivity. Mathematically, partial correlation is also used to express the correlation when any subset of nodes are controlled. To avoid ambiguity, fully partial correlation and partial correlation are explicitly distinguished in this paper. In [19], fully partial correlation performs excellently. However, this strategy sometimes causes two independent nodes to become conditionally dependent, which is known as Berksons paradox [2]. As illustrated by Fig. 1(B), there is no link between Node 1 and Node 2, but the fully partial correlation between these two nodes is very close to 1.0.

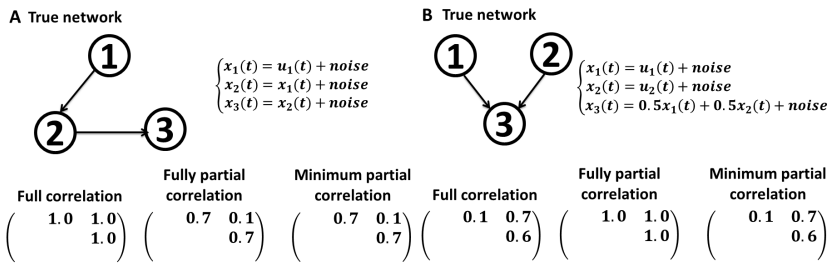

Fig. 1. (A) An example of differentiating direct and indirect effects. (B) An example of the Berkson's paradox.

ICOV (short for Inverse COVariance) is a regularised version of fully partial correlation, which tends to reduce the number of non-zero elements of the precision matrix [9, 25]. Hinne et al. [14] extended ICOV to incorporate both functional and diffusion imaging data. ICOV may correct Berksons paradox if the regularisation parameter is set to an appropriate value. The main drawback of ICOV is that its accuracy highly depends on the regularisation parameter.

One possible way to overcome this drawback is to use cross-validation for automatically adjusting the regularisation parameter. For real datasets, the cross-validation scores were usually calculated based on the accuracies of predicted BOLD signals rather than inferred links [25], because the real links are unknown. However, the validity of tuning the regularisation parameter in this way is still in doubtful for two reasons. One reason is that a value of the parameter that results in accurate prediction (e.g. BOLD signals) is not usually same as the one that is likely to recover the true model (e.g. real links) [16]. The other reason is that functional connectivity is data summarization rather than generative models [10], so it is not reasonable to use functional connectivity to produce BOLD signals for cross-validation.

This paper suggests minimum partial correlation as a measure of functional connectivity in fMRI. The minimum partial correlation between two nodes is the minimum of all absolute values of partial correlations by controlling on all possible subsets of other nodes. Fully partial correlation, which controls all other nodes rather than a proper subset of other nodes, can be regarded as an approximation of minimum partial correlation. Under the faithfulness and Gaussian assumptions, the minimum partial correlation between two nodes is zero if and only if there is no link between the two nodes in the corresponding causal Bayesian network [20]. It is time-consuming to calculate minimum partial correlation, since the number of all possible controlling subsets grows exponentially with the number of nodes. We modify the PC-algorithm to approximate minimum partial correlation, which called elastic PC-algorithm. Unlike the PC-algorithm and its other variations, the elastic PC-algorithm is not based on a specific significance threshold. Within a given time budget, the elastic PC-algorithm can approach the minimum partial correlation as possible. Furthermore, we evaluate and illustrate the elastic PC-algorithm using the NetSim dataset [19] and resting-state fMRI data from the Human Connectome Project [24].

2 Methods

2.1 Preliminaries

Given a directed acyclic graph (DAG) $\mathcal{G} = (V, E)$, which consists of a node set $V = \{r_1, r_2, \cdots, r_N\}$ and an edge set E. Each node denotes a random variable and each directed edge represents a causal effect from one random variable to another one. A causal model represented by $\mathcal{G}$ generates a joint probability distribution over the nodes, which is denoted as $P(V)$. The DAG $\mathcal{G}$ and the distribution P satisfy the causal Markov condition if and only if for every node $r \in V$, the node r is independent of $V \backslash (Descendants(r) \bigcup Parents(r))$ given $Parents(r)$ [20].

When the causal Markov condition is satisfied, the conditional independence relations in the generated distribution P are entailed in the DAG $\mathcal{G}$ using the concept called d-separation. For a 3-node path $r_1 \rightarrow r_2 \leftarrow r_3$, the middle node r_3 is called collider. A path is said to be d-separated by a node set Z if and only if at least one non-collider is in Z or at least one collider and all its descendants

are not in Z; two nodes are said to be d-separated by a node set Z if and only if all path between the two nodes are d-separated by Z [20].

If two nodes r_i and r_j in the DAG $\mathcal{G}$ are d-separated by a node set Z, then r_i and r_j are conditionally independent in distribution P given Z. Here we only emphasis that there is no edge between nodes r_i and r_j if and only if they can be d-separated by some node set. This means that the nodes r_i and r_j are conditionally independent in the distribution P given some node set, if there is no edge in DAG $\mathcal{G}$ between the two nodes. However, the reverse statement is not true. For some DAG $\mathcal{G}$ and distribution P that satisfy the causal Markov condition, the conditional independence between two nodes r_i and r_j does not necessarily imply the absence of an edge between r_i and r_j. In order to exclude these cases, two equivalent concepts, faithfulness [20] and stability [17] , are introduced. A distribution P is faithful to a DAG $\mathcal{G}$ if all and only the conditional independence relations in P are entailed by $\mathcal{G}$ [20]. In other words, there is an edge between two nodes r_i and r_j in $\mathcal{G}$ if and only if r_i and r_j are conditionally dependent in P given any node set.

2.2 Minimum partial correlation

In this paper, we assume the mechanism that generates BOLD signals can be represent by a DAG $\mathcal{G} = (V, E)$, where $V = \{r_1, r_2, \cdots, r_N\}$. The node r_i denotes the random variable of the ith region of interest (ROI). The random variable r_i generates T samples of the BOLD signals of the ith ROI, which are denoted as $X_i = \{x_i^1, x_i^2, \cdots, x_i^T\}$. All random variables generate the same number of samples. The Pearson's correlation coefficient $\rho_{i,j}$ between r_i and r_j can be estimated from samples as follows:

$$\hat{\rho}_{i,j} = \frac{\sum_{t=1}^{T}(x_i^t - \bar{x}_i)(x_j^t - \bar{x}_j)}{\sqrt{\sum_{t=1}^{T}(x_i^t - \bar{x}_i)^2 \sum_{t=1}^{T}(x_j^t - \bar{x}_j)^2}} \tag{1}$$

where $\bar{x}_i$ and $\bar{x}_j$ are sample means of random variables r_i and r_j. The Pearson's correlation coefficient is also called full correlation.

Full correlation count both direct and indirect effects between two random variables. To remove indirect effects, we can calculate the second order correlation between r_i and r_j by removing the effects of a node set $Z \subseteq (V - \{r_i, r_j\})$. This second order correlation is called the partial correlation $\rho_{i,j \cdot Z}$ controlled the node set Z. Similarly, partial correlation can be estimated from samples:

$$\hat{\rho}_{i,j \cdot Z} = \frac{\sum_{t=1}^{T}(\varepsilon_{i \cdot Z}^t - \bar{\varepsilon}_{i \cdot Z})(\varepsilon_{j \cdot Z}^t - \bar{\varepsilon}_{j \cdot Z})}{\sqrt{\sum_{t=1}^{T}(\varepsilon_{i \cdot Z}^t - \bar{\varepsilon}_{i \cdot Z})^2 \sum_{t=1}^{T}(\varepsilon_{j \cdot Z}^t - \bar{\varepsilon}_{j \cdot Z})^2}} \tag{2}$$

where $\{\varepsilon_{i \cdot Z}^t\}$ is a set of the residuals of the linear regression with r_i as the response variable and Z as the predictor variable set. So is $\{\varepsilon_{j \cdot Z}^t\}$. The sample partial correlation is the sample full correlation on the residuals when some controlling factors are regressed out.

The partial correlation between two nodes varies with the controlling set. For a node set V, the minimum partial correlation (MPC) between two nodes is the minimum of all absolute values of partial correlations by controlling all possible subsets of other nodes, which is formally defined as follows:

Definition 1. *Given two nodes r_i and r_j in a node set V, the minimum partial correlation $\omega_{i,j}$ between the two nodes is:*

$$\omega_{i,j} = \min\{|\rho_{i,j\cdot Z}||Z \in 2^{(V-\{r_i,r_j\})}\} \tag{3}$$

where $\rho_{i,j\cdot Z}$ is the partial correlation between r_i and r_j by controlling on the node set Z; $2^{(V-\{r_i,r_j\})}$ represents all subsets of the set $V-\{r_i,r_j\}$. A controlling set Z that satisfies $\rho_{i,j\cdot Z} = \omega_{i,j}$ is called a minimum controlling set for nodes r_i and r_j.

Next, we will discuss an important property of minimum partial correlation that implies that it is good measure of existence of functional connectivity.

Proposition 1. *Assuming the distribution P of the nodes V in a DAG $\mathcal{G}$ is multivariate Gaussian and faithful to $\mathcal{G}$, there is an edge between the two nodes r_i and r_j in $\mathcal{G}$ if and only if the minimum partial correlation $\omega_{i,j} \neq 0$.*

According to Proposition 1, minimum partial correlation is a measure of existence of functional connectivity (i.e. the links in functional networks). Because minimum partial correlation is symmetrical, it cannot infer the directions of functional connectivity. According to Definition 1, minimum partial correlation can be calculated only based on partial correlation. Thus, minimum partial correlation can be estimated from samples, which is denoted as $\hat{\omega}_{i,j}$.

We denote the cardinality of the controlling node set Z as $|Z|$. According to [8], the distribution of the estimated partial correlation $\hat{\rho}_{i,j\cdot Z}$ is the same as the full correlation estimated from $T-|Z|$ samples. Under the hypothesis that $\rho_{i,j\cdot Z} = 0$, the z-score $\tilde{\rho}_{i,j\cdot Z}$ of the estimated partial correlation $\hat{\rho}_{i,j\cdot Z}$ can be calculated as follows:

$$\tilde{\rho}_{i,j\cdot Z} = \frac{1}{2}\log(\frac{1+\hat{\rho}_{i,j\cdot Z}}{1-\hat{\rho}_{i,j\cdot Z}})\sqrt{T-|Z|-3}. \tag{4}$$

However, the distribution of estimated minimum partial correlation is unknown. In this paper, we approximate the absolute value of z-score $\tilde{\omega}_{i,j\cdot Z}$ of the estimate minimum partial correlation $\hat{\omega}_{i,j\cdot Z}$ as follows:

$$\tilde{\omega}_{i,j\cdot Z} = \min\{|\tilde{\rho}_{i,j\cdot Z}||Z \in 2^{(V-\{r_i,r_j\})}\}. \tag{5}$$

Given the BOLD signal samples from all ROIs of a brain, we can calculate the estimated value or the absolute value of its z-score of the minimum partial correlation for each pair of ROIs. These estimated values or absolute values of z-scores form a symmetric matrix Ω. The matrix Ω represents the skeleton of functional connectivity in a fractional way. Given a directed graph $\mathcal{G}$, its skeleton is an undirected graph by stripping away all arrowheads from its links and

getting rid of redundant undirected links. A pair of ROIs with higher estimated minimum partial correlation value or absolute value of the z-score is more likely to have a link between them. The matrix Ω is therefore a parameter-free measure of functional connectivity, which can be used for decoding cognitive states, evaluating brain disease, and etc.

2.3 Elastic PC-algorithm

It is in polynomial time to estimate partial correlation by controlling a given node set. However, we need to enumerate all possible controlling sets in order to estimate minimum partial correlation; the number of all possible controlling sets is exponential. Several algorithms for causal inference can be easily modified to calculate minimum partial correlation, such as SGS [20], PC [20], TPDA [4], MMHC [21] and TPMB [26]. This paper proposes an algorithm, called elastic PC-algorithm (EPC), to approximately calculate the absolute value of the z-score of minimum partial correlation.

Algorithm 1 Elastic PC-algorithm (EPC)

Input: signal samples $\{x_i^t \| t \in [1:T], i \in [1:N]\}$, significance threshold α, previous significance threshold β, previous s-cube $\mathcal{D}$ for β
Output: s-cube $\mathcal{C}$ for α

for each ordered node pair (x_i, x_j) **do**
  $\mathcal{C}(i, j, 0 : (N-2)) \leftarrow |\tilde{\rho}_{i,j}|$
end for
for $k = 1 : (N-2)$ **do**
  initialize the referenced skeletons $\mathcal{S}_\alpha$ and $\mathcal{S}_\beta$ as complete undirected graphs
  for each unordered node pair (x_i, x_j) **do**
    if $\mathcal{C}(i, j, k-1) \leq N^{-1}(1 - \frac{\alpha}{2})$ **then**
      delete the undirected edge (x_i, x_j) from the referenced skeleton $\mathcal{S}_\alpha$
    end if
    if $\mathcal{D}(i, j, k-1) \leq N^{-1}(1 - \frac{\beta}{2})$ **then**
      delete the undirected edge (x_i, x_j) from the referenced skeleton $\mathcal{S}_\beta$
    end if
  end for
  $\mathcal{C}(:, :, k) \leftarrow \mathcal{D}(:, :, k)$
  for each ordered node pair (x_i, x_j) that satisfies $|adj(\mathcal{S}_\alpha, x_i) \backslash \{x_j\}| \geq k$ **do**
    for each controlling node set $Z \subseteq adj(\mathcal{S}_\alpha, x_i) \backslash \{x_j\}$ that satisfies $|Z| = k$ and $Z \bigcup \{x_j\} \not\subseteq adj(\mathcal{S}_\beta, x_i)$ **do**
      if $\mathcal{C}(i, j, k) > |\tilde{\rho}_{i,j \cdot Z}|$ **then**
        $\mathcal{C}(i, j, k : (N-2)) \leftarrow |\tilde{\rho}_{i,j \cdot Z}|$
        $\mathcal{C}(j, i, k : (N-2)) \leftarrow |\tilde{\rho}_{i,j \cdot Z}|$
      end if
    end for
  end for
end for

The elastic PC-algorithm is based on PC-stable algorithm [5]. Different from the PC-algorithm, PC-stable and other variants, the elastic PC-algorithm is not based on a given significant threshold, but automatically increases the significant threshold within a given time budget in order to approach the minimum partial correlation as possible. Moreover, the proposed algorithm can avoid repeatedly calculating same partial correlation. The key of the elastic PC-algorithm is a data structure that called skeleton cube or s-cube for short. A s-cube is a 3D array. Its dimension is $N \times N \times (N-1)$, where N is the number of nodes. The referenced skeleton in the kth stage can be calculated by using the kth slice of the s-cube. Referenced skeletons are used to reduce the number of possible controlling sets.

The elastic PC-algorithm is presented in Algorithm 1. The symbol $N^{-1}(\cdot)$ represents the inverse of the standard normal CDF. The z-score of full correlation $\tilde{\rho}_{i,j}$ and the z-score of partial correlation $\tilde{\rho}_{i,j \cdot Z}$ can be calculated according to Equation 4. To compute the absolute value of the z-scores of minimum partial correlation for a DAG, we start from a small significance threshold and use the EPC to calculate its corresponding s-cube. Based on the s-cube for previous significant threshold, we use the EPC again to calculate a new s-cube for a larger significant threshold. According to Line 16 in Algorithm 1, the EPC does not calculate the z-scores of partial correlation that have been calculated in previous significant threshold. The EPC is repeatedly executed within the given time budget. Due to the space limitation, the details and discussions of the elastic PC-algorithm will be presented in the full version of this paper.

3 Results

The NetSim dataset [19] was used to evaluate the proposed method. There were 28 simulated brain networks with different properties in the NetSim dataset. The signals of brain networks were generated using dynamic causal models with nonlinear balloon models for the vascular dynamics [11]. For each network, 50 simulated subjects were generated by slightly changing the values of parameters. Following [19], we used c-sensitivity of skeleton inference to measure the accuracy. For the details of the NetSim dataset and the definition of c-sensitivity, please refer to [19]. We compared the elastic PC-algorithm with full correlation, fully partial correlation, ICOV [9], network deconvolution (ND) algorithm [7] and global silencing (GS) algorithm [1]. For the elastic PC-algorithm, the initial significant threshold was 0.05 and the step size was 0.005. Ten steps were executed. For ICOV, the regularisation parameter was set to 5 and 100 [19].

Table 1 shows the results of the 28 simulations in the NetSim dataset. The figures highlighted by yellow in Table 1 are the highest c-sensitivity scores in their corresponding simulations. Generally speaking, the elastic PC-algorithm (EPC) performed best. The EPC gained highest c-sensitivity in 24 simulations, full correlation (full) in 1 simulation, fully partial correlation (FP) in 2 simulations, ICOV-5 in 3 simulations, ICOV-100 in 1 simulation, network deconvolution (ND)

in 3 simulations, and global silencing (GS) in 4 simulations. The c-sensitivity values of the EPC were above 0.85 for 19 simulations.

Table 1. Summary of c-sensitivity results over simulations and methods.

ID	Full	FP	ICOV-5	ICOV-100	ND	GS	EPC
1	84.00	92.40	93.60	89.60	92.80	92.00	95.60
2	80.00	86.73	88.36	88.18	91.45	88.18	93.82
3	80.67	82.56	84.78	88.00	88.00	71.22	89.44
4	91.51	77.02	82.95	91.48	91.34	45.97	90.16
5	97.20	99.60	100.00	99.60	100.00	100.00	100.00
6	94.00	99.64	99.82	99.27	99.82	99.64	100.00
7	98.80	100.00	100.00	100.00	100.00	100.00	100.00
8	47.20	65.20	65.20	58.40	62.40	60.40	67.20
9	57.60	81.20	81.20	69.20	74.80	76.00	80.00
10	80.00	96.80	96.80	95.20	95.20	90.80	97.20
11	14.91	11.45	12.91	13.64	13.82	15.27	12.18
12	76.00	83.64	84.55	86.00	87.64	84.73	88.55
13	61.20	61.20	62.00	61.20	64.00	63.20	65.20
14	81.20	94.00	94.40	90.00	93.20	94.80	94.40
15	59.20	89.20	90.00	84.80	80.00	78.00	95.20
16	69.14	85.71	86.57	80.57	86.86	83.14	86.86
17	87.64	92.18	93.27	96.73	96.73	80.18	97.45
18	81.60	91.60	91.60	89.20	92.40	91.60	94.40
19	86.00	94.00	94.40	94.80	95.20	94.80	96.40
20	87.20	95.60	97.20	95.20	98.00	96.00	98.80
21	82.40	89.60	90.40	87.60	90.80	89.20	92.80
22	61.20	74.00	73.60	68.40	72.80	66.80	76.80
23	46.40	73.20	78.00	63.20	70.40	69.20	80.40
24	32.00	41.20	41.60	24.80	38.40	42.40	45.60
25	65.60	68.00	67.20	69.60	72.00	71.60	73.60
26	51.20	53.60	54.00	57.60	57.60	52.80	60.00
27	65.20	68.00	69.20	71.20	74.00	61.60	75.60
28	74.40	83.20	84.00	84.00	84.40	83.20	87.60

We illustrated the proposed method using the resting-state fMRI data of 10 unrelated subjects from the Human Connectome Project [24]. The 10 subjects were healthy, six females and four males. There were 2 subjects in the 22-25 age range, 3 subjects in the 26-30 age range, and 5 subjects in the 31-35 age range. For each subject, the functional data were acquired in four approximate-15-minute runs, which were carried out in two separate sessions. The AAL atlas [23] was used to parcel a whole brain into to 116 ROIs, which are 90 regions in cerebrum and 26 regions in cerebellum. The time series of a ROI were defined to be the average time series of all voxels in this ROI.

Due to the space limitation, we only summarized the results of this empirical study. Details of this study will be presented in the full version of this paper.

With the increase of the significance threshold, the matrix inferred by the EPC algorithm became sparser. Comparing with full correlation, fully partial correlation, ICOV, network deconvolution algorithm and global silencing algorithm, the results of the EPC algorithm were sparest. The matrices inferred by the EPC algorithm seemed denoised versions of the matrices inferred by fully partial correlation. Many elements of the matrices inferred by the EPC algorithm are much smaller that the corresponding elements of the matrices inferred by fully partial correlation. This is because the EPC algorithm tried to minimize partial correlations and shrinks the absolute value of elements.

We analysed the top 1% links inferred by the EPC algorithm. In these strongest 67 links, there were 26 links that connected two homotypic functional regions in different hemispheres; there were 40 links that connected two heterotypic functional regions in same hemispheres; interestingly, there were only 1 links that connected two heterotypic functional regions in different hemispheres. This asymmetrically inter-hemispheric link connected the the superior occipital gyrus of the left hemisphere with the cuneus in the right hemisphere.

4 Conclusions

We proposed an alternative measure, minimum partial correlation, to infer the skeleton of functional networks using fMRI data. We designed an algorithm, called elastic PC-algorithm, to approximate the minimum partial correlation. The experimental results demonstrated the capability of the proposed method to improve the accuracy of skeleton inference.

References

1. Barzel, B., Barabási, A.L.: Network link prediction by global silencing of indirect correlations. Nature biotechnology 31(8), 720–725 (2013)
2. Berkson, J.: Limitations of the application of fourfold table analysis to hospital data. Biometrics Bulletin pp. 47–53 (1946)
3. Buckner, R.L., Krienen, F.M., Yeo, B.T.: Opportunities and limitations of intrinsic functional connectivity mri. Nature Neuroscience 16(7), 832–837 (2013)
4. Chenga, J., Greinera, R., Kellya, J., Bellb, D., Liub, W.: Learning bayesian networks from data: An information-theory based approach. Artificial Intelligence 137, 43–90 (2002)
5. Colombo, D., Maathuis, M.H.: Order-independent constraint-based causal structure learning. Journal of Machine Learning Research 15, 3741–3782 (2014)
6. Craddock, R.C., Jbabdi, S., Yan, C.G., Vogelstein, J.T., Castellanos, F.X., Di Martino, A., Kelly, C., Heberlein, K., Colcombe, S., Milham, M.P.: Imaging human connectomes at the macroscale. Nature Methods 10(6), 524–539 (2013)
7. Feizi, S., Marbach, D., Médard, M., Kellis, M.: Network deconvolution as a general method to distinguish direct dependencies in networks. Nature biotechnology 31(8), 726–733 (2013)
8. Fisher, R.A., et al.: The distribution of the partial correlation coefficient. Metron 3, 329–332 (1924)

9. Friedman, J., Hastie, T., Tibshirani, R.: Sparse inverse covariance estimation with the graphical lasso. Biostatistics 9(3), 432–441 (2008)
10. Friston, K.J.: Functional and effective connectivity: a review. Brain Connectivity 1(1), 13–36 (2011)
11. Friston, K.J., Harrison, L., Penny, W.: Dynamic causal modelling. Neuroimage 19(4), 1273–1302 (2003)
12. Hawellek, D.J., Hipp, J.F., Lewis, C.M., Corbetta, M., Engel, A.K.: Increased functional connectivity indicates the severity of cognitive impairment in multiple sclerosis. Proceedings of the Nnational Academy of Sciences 108(47), 19066–19071 (2011)
13. Hermundstad, A.M., Bassett, D.S., Brown, K.S., Aminoff, E.M., Clewett, D., Freeman, S., Frithsen, A., Johnson, A., Tipper, C.M., Miller, M.B., et al.: Structural foundations of resting-state and task-based functional connectivity in the human brain. Proceedings of the National Academy of Sciences 110(15), 6169–6174 (2013)
14. Hinne, M., Ambrogioni, L., Janssen, R.J., Heskes, T., van Gerven, M.A.: Structurally-informed bayesian functional connectivity analysis. Neuroimage 86, 294–305 (2014)
15. Marrelec, G., Krainik, A., Duffau, H., Pélégrini-Issac, M., Lehéricy, S., Doyon, J., Benali, H.: Partial correlation for functional brain interactivity investigation in functional mri. Neuroimage 32(1), 228–237 (2006)
16. Murphy, K.P.: Machine learning: a probabilistic perspective. MIT press (2012)
17. Pearl, J.: Causality: models, reasoning and inference, vol. 29. Cambridge Univ Press (2000)
18. Shirer, W., Ryali, S., Rykhlevskaia, E., Menon, V., Greicius, M.: Decoding subject-driven cognitive states with whole-brain connectivity patterns. Cerebral Cortex 22(1), 158–165 (2012)
19. Smith, S.M., Miller, K.L., Salimi-Khorshidi, G., Webster, M., Beckmann, C.F., Nichols, T.E., Ramsey, J.D., Woolrich, M.W.: Network modelling methods for fmri. Neuroimage 54(2), 875–891 (2011)
20. Spirtes, P., Glymour, C.N., Scheines, R.: Causation, prediction, and search, vol. 81. MIT press (2000)
21. Tsamardinos, I., Brown, L.E., Aliferis, C.F.: The max-min hill-climbing bayesian network structure learning algorithm. Machine learning 65(1), 31–78 (2006)
22. Turk-Browne, N.B.: Functional interactions as big data in the human brain. Science 342(6158), 580–584 (2013)
23. Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M.: Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain. Neuroimage 15(1), 273–289 (2002)
24. Van Essen, D.C., Smith, S.M., Barch, D.M., Behrens, T.E., Yacoub, E., Ugurbil, K., Consortium, W.M.H., et al.: The wu-minn human connectome project: an overview. Neuroimage 80, 62–79 (2013)
25. Varoquaux, G., Gramfort, A., Poline, J.B., Thirion, B.: Brain covariance selection: better individual functional connectivity models using population prior. In: Advances in Neural Information Processing Systems. pp. 2334–2342 (2010)
26. Wang, Z., Chan, L.: Learning bayesian networks from markov random fields: an efficient algorithm for linear models. ACM Transactions on Knowledge Discovery from Data (TKDD) 6(3), 10 (2012)