Published in IEEE 12th International Symposium on Biomedical Imaging (ISBI)
Year: 2015
Pages: 30 - 33, DOI: 10.1109/ISBI.2015.7163809

LARS NETWORK FILTRATION IN THE STUDY OF EEG BRAIN CONNECTIVITY

Yuan Wang[⋆] *Moo K. Chung*[⋆] *David R.W. Bachhuber*[⋆]
Stacey M. Schaefer[⋆] Carien M. van Reekum[†] Richard J. Davidson[⋆]

[⋆] University of Wisconsin-Madison, U.S.A., [†]University of Reading, U.K.

ABSTRACT

In a brain network, weak and nonsignificant edge weights between nodes signal spurious connections and are often thresholded out of the network. The traditional practice of thresholding edge weights at an arbitrary value can be problematic. Network filtration provides an alternative by summarizing the changes in the network topology with respect to a broad range of thresholds. A well established network filtration approach depends on the graphical-LASSO (least absolute shrinkage and selection operator) model, where a sequence of binary networks are obtained based on non-zero sparse inverse covariance (IC) estimates of partial correlations at a range of sparsity parameters. The limitation of the graphical-LASSO network model is that it relies on the structural information rather than actual entries of the sparse IC matrices and therefore can only yield approximate dynamic topological changes in the network. In the current study, we propose a new network filtration approach based on least angle regression (LARS) that yields exact filtration values at which network topology changes, and apply it to study brain connectivity in response to emotional stimuli across different age groups via electroencephalographic (EEG) data.

Index Terms— LARS, EEG, network filtration, emotion, brain connectivity

1. INTRODUCTION

Correlation-based network analysis has been widely used to measure the strength of brain connectivity [1, 2]. In a dense network, correlation matrices are often thresholded to reveal strong and significant connections between nodes and regions, followed by graph theoretic measures to quantify the brain connectivity differences [3]. Nevertheless, thresholding the correlation matrix at a single value provides only a snapshot of the network connectivity, thus losing a dynamic sense of connectivity changes in the network. Most importantly, the optimal choice of a single threshold can vary arbitrarily across studies and affect consistency of interpretation [4].

Network filtration provides an alternate approach of thresholding edge weights at a broad range of values and summarizing changes in the network topology as the threshold varies. An early development in this direction was the sparse inverse covariance estimation (SICE) model proposed in [5]. The model utilizes the popular graphical-LASSO method to calculate LASSO-penalized maximum likelihood estimates (MLE) of the inverse covariance (IC) matrix over a range of sparsity parameters λ. However, the graphical-LASSO method is not reliable in estimating the magnitude of the non-zeros entries of the IC matrix due to its shrinkage property. So only the structural information of zero and non-zero entries is utilized in defining a binary network. At a given λ, two nodes i and j are *directly connected* by an edge if the (i, j)-entry in the λ-penalized IC matrix is non-zero. As λ increases, nonzero entries thin out and a sequence of binary networks corresponding to subgraphs of the complete network is created via the disappearance of edges. Since the magnitude estimation is not reliable, the model can only yield approximate λ values where the topology of node clusters changes. Also, topological information embedded in the filtration of subgraphs was not fully exploited in [5]. Only edge numbers within and between brain regions were used to quantify changes in connectedness of node clusters and to subsequently compare brain networks in Alzheimer's patients and normal controls.

In this paper, we propose a new network filtration method named *persistent LARS (pLARS)* that accommodates a direct and exact thresholding scheme through the LARS algorithm [6]. We also utilize it to investigate whether age affects EEG network connectivity in response to viewing pictures of positive, negative or neutral emotional content [7]. Our key contributions are: (a) establishing a persistent data structure in the LARS algorithm, and exploiting this structure to construct a network filtration; (b) applying the pLARS model to a 128-channel EEG dataset to determine possible aging effect on EEG brain connectivity with respect to different types of emotional stimuli. To the best of our knowledge, this is the first study exploring a LARS filtration for a task-related high-density EEG network.

2. METHOD

The LARS algorithm successively builds up a linear estimate for a response vector over normalized vector covariates [6]. Suppose data vectors $\mathbf{x}_i = (x_{i1}, \ldots, x_{in})', i = 1, \ldots, p$, are centered and normalized EEG signals recorded at the p

978-1-4799-2374-8/15/$31.00 ©2015 IEEE

nodes for n time points, i.e. $\mathbf{x}_i'\mathbf{x}_i = 1$ and $\sum_{i'=1}^{n} x_{ii'} = 0$ for all $i = 1, \ldots, p$. We assume that the EEG measurement at any node i is linearly correlated with measurements at all the other nodes:

$$\mathbf{x}_i = \sum_{j \neq i} \beta_{ij}\mathbf{x}_j + \boldsymbol{\varepsilon}_i, \tag{1}$$

with $var(\boldsymbol{\varepsilon}_i) = 1/\pi_{ii}$ and $cov(\boldsymbol{\varepsilon}_i, \boldsymbol{\varepsilon}_j) = \pi_{ij}/(\pi_{ii}\pi_{jj})$.

In this setting, with an abuse of notation, the LARS algorithm builds up an estimate $\widehat{\mathbf{x}}_i$ for each response $\mathbf{x}_i$ on the left-hand side in (1) in exactly $p-1$ steps. Geometrically, the algorithm follows a piecewise linear path with $p-1$ turning points. It begins at the origin and sets off in the direction of a node most correlated with $\mathbf{x}_i$. When a second node has as much correlation with the difference between $\mathbf{x}_i$ and its current estimate as node one, the path turns to a direction equiangular between the two nodes and continues until a third node has as much correlation with the difference between $\mathbf{x}_i$ and its current estimate as the first two. The path turns and continues in such a manner until a final estimate for $\mathbf{x}_i$ is attained. The algorithm can be implemented as follows.

---

(1) Let $\widehat{\mathbf{x}}_i^0 = 0$.
(2) Let $k = 1$.
(3) Compute the *current residual* $\mathbf{r}_i = \mathbf{x}_i - \widehat{\mathbf{x}}_i^{k-1}$.
(4) Compute the correlation c_{ij}^k between the current residual $\mathbf{r}_i$ ($j \neq i$) and the nodes $\mathbf{x}_j$: $c_{ij}^k = \mathbf{x}_j'\mathbf{r}_i, j \neq i$, which are the *current correlations* at step k.
(5) Define the *maximal absolute current correlation* (MACC) as the maximum along all current correlations at node i with all other nodes:

$$C_i^k = \max\{|c_{ij}^k| : j \neq i\}. \tag{2}$$

(6) Compute the *equiangular vector* $\mathbf{u}_k = X_k A_k (X_k' X_k)^{-1}\mathbf{1}_k$, where

$$X_k = (sign(c_{ij_1}^k)\mathbf{x}_1, \ldots, sign(c_{ip}^k)\mathbf{x}_p)$$

and

$$A_k = (\mathbf{1}_k (X_k' X_k)^{-1}\mathbf{1}_k)^{-1/2}.$$

(7) Compute the projection vector $\mathbf{a}^k = (a_1^k, \cdots, a_p^k)$, where $a_j^k = \mathbf{x}_j'\mathbf{u}_k, j \neq i$.
(8) The estimate for $\mathbf{x}_i$ is updated to

$$\widehat{\mathbf{x}}_i^k = \widehat{\mathbf{x}}_i^{k-1} + \widehat{\gamma_k}\mathbf{u}_k,$$

where $\widehat{\gamma_k} = \min_{j'}^{+}\left\{\frac{C_i^k - c_{ij}^k}{A_k - a_j^k}, \frac{C_i^k + c_{ij}^k}{A_k + a_j^k}\right\}$ with $\min^+$ taking minimum over positive components within $\{\cdot\}$ for all $j' \neq j_1, \ldots, j_k$. The algorithm then proceeds to the next step with the updated $\widehat{\mathbf{x}}_i^k$.
(9) If $k \leq p-1$, let $k+1 \to k$ and go to step **(2)**; otherwise terminate.

Defining a LARS network. We first define a directed network: the weight for the edge in the direction $\mathbf{x}_{i'} \to \mathbf{x}_i, i' \neq i$, is defined to be the MACC C_i^ℓ at the step ℓ when the covariate node $\mathbf{x}_{i'}$ enters the LARS path for the response node $\mathbf{x}_i$; it is treated as the contribution of node i' to node i relative to all the covariate nodes. Similarly, the weight for the edge in the direction $\mathbf{x}_i \to \mathbf{x}'_i$ is defined as the MACC $C_{i'}^{\ell'}$ at the step ℓ' when the covariate node $\mathbf{x}_i$ enters the LARS path for the response node $\mathbf{x}_{i'}$. Diagonal edge weights are set to 1. We then symmetrize the directed network by defining the *edge weight* $w_{ii'}$ of a node pair (i, i') as the minimum of the bilateral edges, i.e.

$$w_{ii'} = \min\{C_i^\ell, C_{i'}^{\ell'}\}, \tag{3}$$

where we take the $\min(\cdot)$ norm to ensure that we only consider direct inter-nodal weight as strong when bilateral weights are strong. We shall refer to this weighted network as the *LARS network*. A binary network then follows by defining the adjacency matrix $B(\lambda) = (b_{ij}(\lambda))$:

$$b_{ij}(\lambda) = \begin{cases} 1 & \text{if } w_{ij} > \lambda \\ 0 & \text{otherwise} \end{cases}. \tag{4}$$

on which the connectedness of the LARS network $G(\lambda) = (V, E(\lambda))$ is defined: two nodes i and j are directly connected by an edge if the (i, j)-entry in the adjacency matrix B is 1, or equivalently when $w_{ij} > \lambda$. The vertex set V consists of p nodes and the edge set $E(\lambda)$ consists of all the edges, or arcs, present in $B(\lambda)$.

LARS network filtration. A data structure is said to be *persistent* if its state is preserved when it is modified according to a certain criterion. A persistent data structure can manifest itself in various forms that facilitate data exploration. In a graph theoretic setting, the graph $G(\lambda)$ for some parameter λ is defined to be persistent if it has the nested subset structure $G(\lambda_1) \supset G(\lambda_2) \supset \cdots$ for $\lambda_1 \leq \lambda_2 \leq \lambda_3 \cdots$ [2]. Here we establish a persistent structure in the LARS network via its graphs $G(\lambda)$. Without proof, we present Theorem 1, which states that the sequence, or filtration, of the subgraphs of the LARS network is persistent. Also, the established persistent graph filtration, which we shall refer to as the *persistent LARS filtration*, is maximal in the sense that no additional levels of subgraphs can be added to the filtration [2].

Theorem 1. *For data $\mathbf{x}_i = (x_{i1}, \ldots, x_{in})', i = 1, \ldots, p$, that are centered and normalized, let $w_{(1)}, w_{(2)}, \ldots, w_{(q)}$ be the order statistics of the unique edge weights w_{ij} with $q = p(p+1)/2$ being the number of distinct entries. Then we have the maximal filtration of subgraphs $G(w_{(1)}) \supset G(w_{(2)}) \supset \cdots \supset G(w_{(q)})$.*

Persistence as a quasi-measure of connection strength. The nested structure in Theorem 1 constitutes the *persistent edge property* of the LARS network filtration: *an edge that disappears at λ_1 does not appear again for any $\lambda_2 > \lambda_1$,*

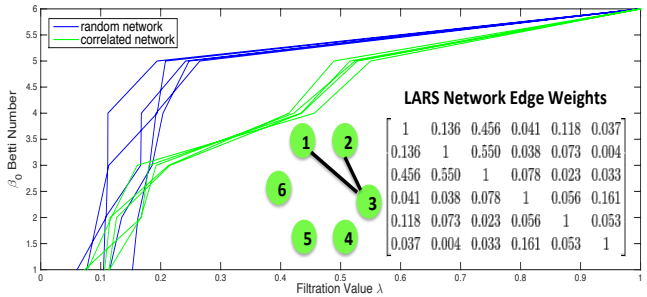

Fig. 1. Barcode representation of pLARS on five simulated random and correlated networks, where the latter have underlying connection between nodes 1, 2 and 3 (middle). The edge weight matrix corresponds to one of the correlated networks.

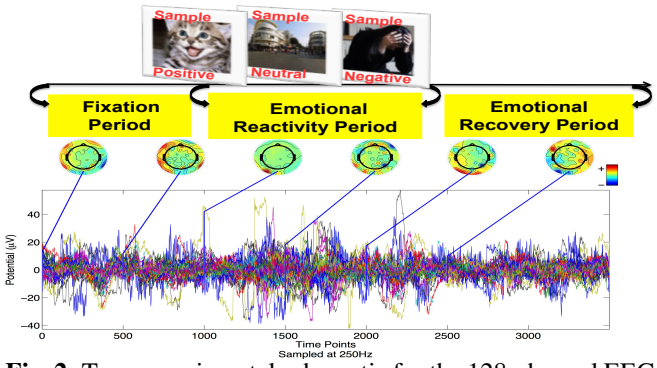

Fig. 2. Top: experimental schematic for the 128-channel EEG data. Main: a sample set of EEG signals from the 128 channels across the whole time span. Data analysis is based on averaging over 30 sets of signals in each picture category.

which is a stronger property than the *monotone property* of the SICE model [5]. As a result, the *persistence* represented by λ qualifies as a quasi-measure of connection strength between nodes. It indirectly measures how long node pairs, hence path-connected components, stays connected as the filtration value λ increases.

Barcode representation of persistence. We then quantify changes in the connectedness by a barcode summarizing the zeroth Betti number $\beta_0(\lambda)$, which counts the number of path-connected components, across a range of λ [8]. By assuming the edge weights of the LARS network are unique, we can compute the barcode of its filtration via Kruskal's algorithm for finding a minimum spanning tree (MST) [9]. Whenever a new edge is added in ascending order of $(1 - w_{ij})_{i,j}$, the algorithm checks whether it connects two different path-connected components. Two path-connected components are merged when they are connected for the first time. Once we obtain the complete sequence of λ values where components merge, we can subtract them from 1 to obtain the x-coordinates for the β_0 barcode.

Comparing connectivity between groups. To compare the barcodes $\beta_0^1(\lambda)$ and $\beta_0^2(\lambda)$ of two LARS network filtrations, we conduct the test

$$H_0\text{: } \beta_0^1(\lambda) = \beta_0^2(\lambda), \textit{ for all } \lambda \in [0,1];$$

$$H_1\text{: } \beta_0^1(\lambda) \neq \beta_0^2(\lambda), \textit{ for some } \lambda, \in [0,1].$$

where the range $[0, 1]$ of the filtration value λ is determined in Theorem 2, which shows the LARS MACCs, hence the edge weights, are bounded between 0 and 1.

Theorem 2. *Given that* $\mathbf{x}_i = (x_{i1}, \ldots, x_{in})', i = 1, \ldots, p$ *are centered and normalized, the LARS MACCs* C_i^k *as defined by* (2) *satisfy the inequality* $0 \leq C_i^k \leq 1$*, for any* i *and* $k = 1, \ldots, p-1$.

A permutation test is then performed on two groups of barcodes to find the significance of the observed null maximum absolute t-statistic $\max_{\lambda \in [0,1]} |T(\lambda)|$. This will account for multiple comparison over every possible filtration value λ.

3. SIMULATIONS

We evaluate the ability of pLARS to detect actual nodal connectivity in the underlying network in comparison with the graphical-LASSO model [5]. We simulate two sets of ten 6-node networks. In set 1, each network is simulated as a random network with independent series of Gaussian noise $N(0, 0.1)$ of size 100 simulated at each node. In set 2, an independent series Gaussian noise $N(0, 0.1)$ of size 100 is simulated at nodes 2 through 6, and connections are imposed between nodes 1, 2 and 3, i.e. $\mathbf{x}_3 = 0.5\mathbf{x}_2 + 0.5\mathbf{x}_1 + \boldsymbol{\varepsilon}, \boldsymbol{\varepsilon} \sim N(0, 0.1 \times \mathbf{I}_{100})$. In Figure 1, the significance of connections is reflected by the magnitude of the edge weights between the three nodes. In 300 simulations, pLARS identifies 99.7% and 99% of true signals at 5% and 1% significance levels respectively, whereas the graphical-LASSO only identifies 60% and 59.5% at 5% and 1% significance levels respectively. Our results show that pLARS outperforms the graphical-LASSO by a large margin in detecting underlying connections for this group of networks.

4. APPLICATION TO EEG ANALYSIS

Literature suggests age effect on emotional response [7]. We apply the pLARS method to study age effect on EEG connectivity in response to different emotional stimuli.

Data collection and processing. The dataset was acquired as part of the Midlife in US (MIDUS II) project. EEG signals were recorded at a 250Hz rate, referenced at the center node Cz, and band-pass filtered within [0.5Hz,50Hz] for 106 subjects aged between 36 and 84 years old (57 ± 11 years). A total of 90 digital color pictures (30 positive/neutral/negative) selected from the International Affective Picture System (IAPS) were presented to each of the subjects in a randomized sequence [7]. The first second was designated as a fixation period before the onset of a picture. The picture was turned off 4s after onset (Figure 2). We removed bad chan-

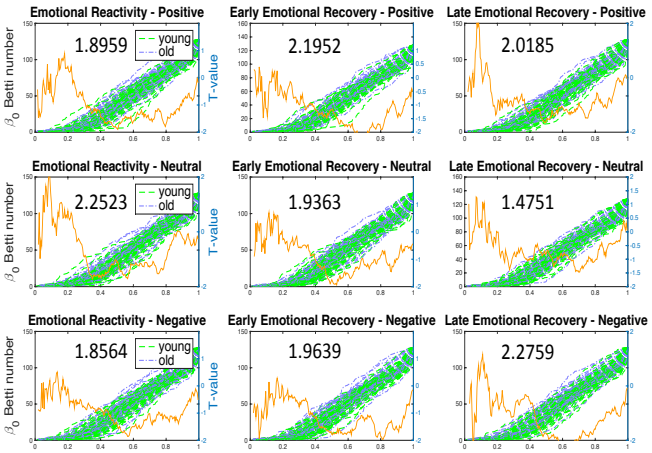

Fig. 3. Barcodes of pLARS networks and their pointwise t-values for three emotional response periods of three picture-viewing tasks; the x-axis represents the filtration value λ, and the left and right y-axis represent the β_0 Betti numbers and t-values of pointwise two-sample t-tests respectively. The maximum absolute p-values are marked in each category.

nels, corrected myogenic artifacts by performing independent component analysis (ICA) on the filtered EEG signals [10], and spherically interpolated missing channels over the skull. We then averaged over signals recorded on each subject in periods viewing 30 pictures in the same category. Signals in the fixation period were treated as baselines and subtracted from signals in subsequent periods second by second.

LARS network filtration. Subjects were divided into two age groups: young (below 60 years) and old (above 60 years), and three emotional periods: emotional reactivity - 4s period after picture onset; early recovery - 2s period after picture offset; late recovery - 2s period after early recovery (Figure 2). The pLARS procedure is performed independently on data from individual subjects for the three emotional periods in the three valence categories. Figure 3 shows barcodes and respective t-values from pointwise two-sample t-tests. The barcodes do not show clear group separation, and there is no highly significant maximum absolute t-value. The permutation test did not yield significant age effect on any cross category of valence and emotional response period for a 5% significance level with Bonferroni correction for 9 tests.

5. DISCUSSION

We conclude that there is no evidence of age effects on EEG study we performed, possibly due to small sample size. Additional study with increased sample size is warranted. Also, we currently require solving LARS in full to threshold by the MACCs. Further exploitation of the LARS structure may yield an alternate filtration value that bypasses a full solution.

Acknowledgements. The research was funded by NIH grants PO1-AG020166, R01-MH043454, UL1-TR000427, P30-HD03352 and the Vilas Associate Award from University of Wisconsin-Madison. We thank Hyekyung Lee at Seoul National University Hospital and Matthew Arnold at the University of Bristol for their helpful discussion on persistent homology.

6. REFERENCES

[1] G. Marrelec, A. Krainik, H. Duffau, M. Peregrini-Issac, S. Lehericy, J. Doyon, and H. Benali, "Partial correlation for functional brain interactivity investigation in functional MRI," *NeuroImage*, vol. 32, pp. 228–237, 2006.

[2] M.K. Chung, J.L. Hanson, J. Ye, R.J. Davidson, and S.D. Pollak, "Persistent homology in sparse regression and its application to brain morphometry," *arXiv 1409.0177*, 2014.

[3] H. Lee, D.S. Lee, H. Kang, B.-N. Kim, and M.K. Chung, "Sparse brain network recovery under compressed sensing," *IEEE Transactions on Medical Imaging*, vol. 30, no. 5, pp. 1154–1165, 2011.

[4] D Meunier, S Achard, A Morcom, and E Bullmore, "Age-related changes in modular organization of human brain functional networks," *Neuroimage*, vol. 44, no. 3, pp. 715–23, 2009.

[5] S. Huang, J. Li, L. Sun, J. Ye, A. Fleisher, T. Wu, K. Chen, and E. Reiman, "Learning brain connectivity of Alzheimer's disease by sparse inverse covariance estimation," *NeuroImage*, vol. 50, no. 3, pp. 935–949, 2010.

[6] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani, "Least angle regression," *Annals of Statistics*, vol. 32, no. 2, pp. 407–499, 2004.

[7] C.M. van Reekum, S.M. Schaefer, R.C. Lapate, C.J. Norris, L.L. Greischar, and R.J. Davidson, "Aging is associated with positive responding to neutral information but reduced recovery from negative information," *Social Cognitive and Affective Neuroscience*, vol. 6, no. 2, pp. 177–85, 2011.

[8] H. Edelsbrunner and J. Harer, *Computational Topology*, American Mathematical Society, 2010.

[9] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*, MIT Press, 2009.

[10] A.J. Shackman, B.W. McMenamin, H.A. Slagter, J.S. Maxwell, L.L. Greischar, and R.J. Davidson, "Electromyogenic artifacts and electroencephalographic inferences," *Brain Topography*, vol. 22, no. 1, pp. 7–12, 2009.