

NETWORK INFERENCE AND CHANGE POINT DETECTION FOR PIECEWISE-STATIONARY TIME SERIES

Hang Yu, Chenyang Li, and Justin Dauwels

School of Electrical and Electronics Engineering, Nanyang Technological University, Singapore

ABSTRACT

Graphical models are powerful tools to describe complex systems. Especially sparse graphical models are currently en vogue, as they allow us to infer network structure from multiple time series (e.g., functional brain networks from multi-channel electroencephalograms). So far, most of the literature deals with stationary time series, whereas real-life time series often exhibit non-stationarity. In this paper, techniques are proposed to infer graphical models from piecewise stationary time series; first change point are detected in the time series, and then graphical models are inferred for each stationary segment. Specifically, a low-complexity algorithm based on Pruned Exact Linear Time method is proposed to identify change points. Copula Gaussian graphical models (with and without hidden variables) are then generated for each stationary segment. The crux of the proposed approach is that it determines the number and location of the change points as well as the graphical models in a fully automated manner. Results for both synthetic data and scalp electroencephalograms of epileptic seizure patients are provided to validate the model.

Index Terms— change point detection, PELT, Gaussian copula, graphical model, functional network, seizure

1. INTRODUCTION

Graphical models have rich modeling capabilities and great computational efficiency [1, 2]. By means of graphical models, the statistical relations between numerous stochastic variables can be captured as a sparse network. For instance, functional brain networks can be inferred from electroencephalograms using graphical models. Most of the literature on Graphical models deals with Gaussian variables [1, 2]. However, recently developed Gaussian copula graphical models enable us to tie any kind of marginal distributions (both Gaussian and non-Gaussian) together to form a joint distribution [3, 4, 5, 6, 7, 8]; the precision matrix (inverse covariance) in the latent Gaussian layer describes the interaction between the variables in a compact manner, yielding efficient inference algorithms. Gaussian copula graphical models have been applied in such diverse areas as computational biology [3, 4] and neuroscience [5], geophysics [6] and extreme events analysis [7], and sociology [8].

So far, most literature on network inference via graphical models focuses on stationary data. However, real data are

often non-stationary, and statistical models designed for stationary data may not yield accurate results. For example, during epileptic seizures, functional brain networks are shown to evolve through a sequence of distinct topologies [9]. Inferring such evolving networks in the framework of graphical models has received little attention until now. A reasonable approach is to detect change points, and then infer graphical models in the stationary segments between the change points. Xuan *et al* [10] employed the Bayesian change point detection approaches: they adopt a geometric prior on the time segment lengths, and then iterate between MAP segmentation and graphical model inference. The main restriction, however, is that the graph for all segments must be decomposable. On the other hand, a greedy binary segmentation scheme is proposed in [11]. A change point is inserted such that the BIC of the two graphical models of the data before and after the change point is minimized; this procedure is repeated until no further splits reduce the BIC score. Unfortunately, besides the high computational complexity, the binary segmentation can be misleading and overestimate the number of change points, as pointed out in [12]. To address this concern, dynamic programming is applied in [13], resulting in joint estimation of all the change points. Unfortunately, the method has computational complexity of order $\mathcal{O}(n^3)$ in the number of fixed points n , which is impractical for most real-life time series with tens or hundreds of change points. Another limitation of the aforementioned methods is the assumption of Gaussian distributed data, which is not always fulfilled in practice.

In this paper, we aim to establish Gaussian copula graphical models for non-stationary, in particular, piecewise stationary time series. Since those graphical models rely on copulas, they are applicable to non-Gaussian data (cf. [3, 4, 5, 6, 7, 8]). In order to reduce the computational complexity, we disentangle the process of change point detection and graphical model inference. Specifically, we first detect the change points by minimizing a cost function defined on covariance matrix using low-complexity Pruned Exact Linear Time (PELT) method [14], and next learn the graphical model based on the covariance of each stationary time segment. The procedure also infers the number of change points in an automated fashion. Numerical results for both synthetic and real data show that the proposed method provides an effective and efficient tool to identify change points and infer networks.

This paper is structured as follows. In Section 2, we briefly introduce Gaussian copula graphical models (with hidden variables) for stationary data. In Section 3 we extend those models to piecewise stationary data. In Section 4 we present results for synthetic and real data, validating the proposed model. We offer concluding remarks in Section 5.

2. COPULA GAUSSIAN GRAPHICAL MODELS

2.1. Standard Copula Gaussian Graphical Models

Let the observed non-Gaussian variables and hidden Gaussian variables be Y_1, \dots, Y_P and Z_1, \dots, Z_P respectively. A Gaussian copula graphical model is defined as [3, 15]:

$$Z \sim \mathcal{N}(0, K^{-1}) \quad Y_k = F_k^{-1}(\Phi(Z_k)), \quad (1)$$

where K is the precision matrix whose inverse (the covariance matrix) has normalized diagonal, Φ is the CDF (cumulative distribution function) of the standard Gaussian distribution, and F_k is the CDF of Y_k . It is straightforward to prove that the graphical model structure is characterized by K , i.e., $K(i, j) = 0$ if and only if the edge (i, j) is absent [3]. Thus, the graphical model can be inferred using the method called graphical lasso (glasso) [16] by solving the convex problem:

$$\hat{K} := \operatorname{argmin}_{K \succ 0} \operatorname{tr}(SK) - \log \det K + \lambda \|K\|_1, \quad (2)$$

where S is the empirical covariance of Z . Note that the regularization parameter λ needs to be selected carefully in order to recover the true precision matrix. As in [4], we apply the BINCO method [17] to select the parameter λ .

2.2. Copula Gaussian Graphical Models with Hidden Variables

The aforementioned model is useful only when all the relevant variables are observed, whereas in many applications some of them may be hidden or unobserved. Consequently, in earlier work [4] we proposed Gaussian copula graphical models with hidden variables (HVGGM), which we will briefly explain in the following. Suppose in the Gaussian latent layer, each Z_{o_k} is associated with an observed non-Gaussian variable Y_k and there exist several hidden variables Z_h that are not associated with observed variables. The joint precision matrix $K_{(o\ h)}$ of both Z_o and Z_h is given by:

$$K_{(o\ h)} = \begin{bmatrix} K_o & K_{o,h} \\ K_{h,o} & K_h \end{bmatrix}. \quad (3)$$

According to Schur complement, the marginalized precision matrix \tilde{K}_o of Z_o can be written as:

$$\tilde{K}_o = K_o - K_{o,h}K_h^{-1}K_{h,o} = K_o - L, \quad (4)$$

with product matrix $L = K_{o,h}K_h^{-1}K_{h,o}$. Those two components have the following properties [18]: K_o is the supposedly sparse conditional precision matrix of Z_o , conditioned on Z_h ; the product matrix L summarizes the effect of marginalization over the hidden variables. The rank of that matrix equals the number of hidden variables Z_h . Resulting from the subtraction, \tilde{K}_o is dense. The graphical lasso [16] and standard copula Gaussian graphical model [3] both would yield dense graphs since they infer \tilde{K}_o . Instead, K_o and L may be

recovered by solving the convex relaxation [18]:

$$\begin{aligned} (\hat{K}_o, \hat{L}) = \operatorname{argmin}_{K_o, L} & \operatorname{tr}((K_o - L)S_o) - \log \det(K_o - L) \\ & + \lambda(\gamma \|K_o\|_1 + \operatorname{tr}(L)), \end{aligned} \quad (5)$$

where S_o is the empirical marginal covariance of Z_o . To recover the correct matrices K_o and L , the parameters λ and γ need to be chosen appropriately, which can be settled via stability selection [19]. We refer to [4] for more details.

3. CHANGE POINT DETECTION

In this section, we address the problem of detecting change points in piecewise stationary multivariate time series. Specifically, we aim to detect changes in the statistical dependence (a.k.a. connectivity or network structure) among the time series. The data between the change points is assumed to be stationary, and will be modeled by copula Gaussian graphical models. Let us assume that we have an ordered sequence of n samples for each of the p variables, $y_k^{(i)}$, where $i = 1, \dots, n$ and $k = 1, \dots, p$. We wish to infer an unknown number m of change points $\tau_{1:m} = (\tau_1, \dots, \tau_m)$. Each change point is an integer between ℓ and $n - \ell$, where ℓ is minimum length of one segment. We further define $\tau_0 = 0$ and $\tau_{m+1} = n$, and thus the m change points will split the data into $m + 1$ segments, where the k th segment is given by $y_{1:p}^{(\tau_{k-1}+1:\tau_k)}$.

As a first step, we transform the non-Gaussian observed variables Y_k into Gaussian latent variables Z_k (associated with the observed variables Y_k), i.e., $Z_k = \Phi^{-1}(\hat{F}_k(Y_k))$.

We next solve the problem of identifying all the change points together in the Gaussian latent layer. Concretely, we minimize a cost function with a penalty on the number of change points, as suggested in the literature [14, 20]:

$$\sum_{k=1}^{m+1} L_k \left(z_{1:p}^{(\tau_{k-1}+1:\tau_k)} \right) + \beta m, \quad (6)$$

where β is the cost associated with each change point, in order to limit overfitting. The negative log-likelihood L_k is defined as $L_k(z_{1:p}^{(\tau_{k-1}+1:\tau_k)}) = (\tau_k - \tau_{k-1})/n \log \det(S_k)$, where S_k is the empirical covariance of segment k . We apply the Pruned Exact Linear Time (PELT) method [14] to efficiently find the global minimum of (6). More specifically, let $F_m(t)$ denote the global minimum of (6) for data $z_{1:p}^{(1:t)}$ and let $\mathcal{T}_t = \{\tau : 0 = \tau_0 < \tau_1 < \dots < \tau_m < \tau_{m+1} = t\}$ be the set of candidate change points. It therefore follows that:

$$\begin{aligned} F_m(t) &= \min_{\tau \in \mathcal{T}_t} \sum_{k=1}^{m+1} \left[L \left(z_{1:p}^{(\tau_{k-1}+1:\tau_k)} \right) + \beta \right] \\ &= \min_s \left\{ \min_{\tau \in \mathcal{T}_s} \sum_{k=1}^m \left[L \left(z_{1:p}^{(\tau_{k-1}+1:\tau_k)} \right) + \beta \right] + L(z_{1:p}^{(s+1:t)}) + \beta \right\} \\ &= \min_s \left[F_{m-1}(s) + L(z_{1:p}^{(s+1:t)}) + \beta \right]. \end{aligned} \quad (7)$$

The expression (7) offers a recursion expressing the minimum for data $z_{1:p}^{(1:t)}$ in terms of the minimum for $z_{1:p}^{(1:s)}$ for $s < t$. As a result, (7) can be solved in turn for $t = \ell + 1, \dots, n$ by

finding the most recent change point s before t . In summary, PELT is a dynamic programming procedure:

1. Initialize $F(0) = -\beta$. The set $cp(t)$ of previous change points at $t \leq \ell$ is initialized as $cp(t) = \emptyset$, and the set $R_{\ell+1}$ of candidate change points at $t = \ell + 1$ is initialized as $R_{\ell+1} = \{0\}$.
2. Compute $F(t) = \min_{\tau \in R_t} [F(\tau) + L(z_{1:p}^{(\tau+1:t)}) + \beta]$ and let $\tau^* = \operatorname{argmin}_{\tau \in R_t} F(t)$
3. Update the set $cp(t)$ of previous change points at time t : $cp(t) = cp(\tau^*) \cup \tau^*$.
4. Prune the set R_t by removing $\{\tau \in R_t : F(\tau) + L(z_{1:p}^{(\tau+1:t)}) > F(t)\}$. If $t \geq 2\ell - 1$, then update R_{t+1} of candidate change points at the next time position $t + 1$: $R_{t+1} = R_t \cup \{t + 1 - \ell\}$.
5. Return to Step 2 if $t < n$ and increase t by 1.

The resulting $cp(n)$ is the optimal set of change points. Note that Step 4 shrinks the set of candidate change points by discarding those τ that can never be minima in the future iterations as proven in [14]; such procedure successfully removes computations that are irrelevant to obtaining the final set of change points and accelerates the algorithm. Under certain conditions, the computational complexity is linear in n [14].

Another issue with solving (6) is the selection of regularization parameter β , which determines the final number of change points. However, there is no uniform rule to compute β for all kinds of data. An adaptive method to determine the number of change points is introduced in [20], which keeps increasing the number of change points until the negative log-likelihood of the entire time series ceases to decrease significantly. In the example of Fig. 1(a), the optimal number of change points is three, as the negative log-likelihood starts decreasing more slowly for larger number of change points. We modify the approach to choose β in the same spirit.

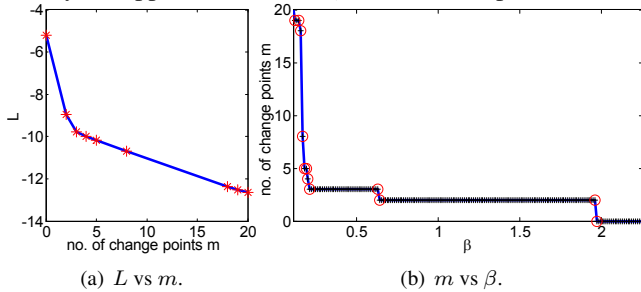


Fig. 1. Adaptive regularization selection.

Let $L = \sum_{k=1}^{m+1} L_k$ and define a general penalty $f(m)$, which equals m in our case. It is apparent that the estimated number of change points $\hat{m}(\beta)$ is a piecewise constant function of β . As such, if $\hat{m}(\beta) = a$, $L(a) + \beta f(a) < \min_{b \neq a} (L(b) + \beta f(b))$. Therefore, β satisfies the following condition [20]:

$$\max_{b>a} \frac{L(a) - L(b)}{f(b) - f(a)} < \beta < \min_{b<a} \frac{L(b) - L(a)}{f(a) - f(b)}. \quad (8)$$

In the example of Fig. 1, for β such that $\hat{m}(\beta) = 3$, $L(3) - L(4) < \beta < L(2) - L(3)$. There hence are ordered sequences

$1 = m_1 < m_2 < \dots$ and $\infty = \beta_0 > \beta_1 \dots$ defined as [20]:

$$\beta_i = \frac{L(m_i) - L(m_{i+1})}{f(m_{i+1}) - f(m_i)}, \quad i \geq 1, \quad (9)$$

such that $\hat{m}(\beta) = m_i, \forall \beta \in [\beta_i, \beta_{i-1})$ (as marked the red circles in Fig. 1(b)). To find the m for which L ceases to decrease significantly, we need to look for a break in the slope of the function $L(f)$, which is defined by β according to (9). Therefore, the change of the slope is determined by the length ℓ_{m_i} of the interval $[\beta_i, \beta_{i-1})$. As a result, the regularization selection procedure can be executed as follows:

1. Select an arithmetic sequence of $\beta = (\beta_1, \beta_2, \dots)$ (black plus signs in Fig. 1(b)). In practice, we set the interval between two β values to be $c \log(n)p(p+1)/4n$, where $c = 0.02$ is a user-defined constant and $\log(n)p(p+1)/4n$ is the regularization parameter when Bayesian Information Criteria (BIC) is employed for regularization selection.
2. For each β_i , estimate the number of change points $\hat{m}(\beta_i)$ by solving (6) using PELT method.
3. Compute the length ℓ_{m_j} of the interval by counting the number of β_i that generates m_j change points, which is represented by the length between two circles with the same number of change points in Fig. 1(b).
4. Choose the smallest value of β such that $\ell_{m_j} \gg \ell_{m_k}$, for $k > j$. In Fig. 1(b), as an example, β corresponding to $\ell_{\hat{m}(\beta)=3}$ is optimal since $\ell_{\hat{m}(\beta)=3} \gg \ell_{\hat{m}(\beta)=4}$.

Once we obtain the optimal segmentation of the time series, we can learn graphical models based on the empirical covariance S_k for each time segment. Here we consider copula Gaussian graphical models without and with hidden variables, which are inferred by solving (2) and (5) respectively.

4. NUMERICAL RESULTS

In this section, we validate the proposed model on both synthetic and real data.

4.1. Synthetic Data

We generate 100 25-dimensional data sets with 2220 non-Gaussian samples for each variable. The common true values of change points are (661, 1561, 1861). In order to assess both the copula Gaussian graphical models with and without hidden variables, the first half of 100 data sets does not contain hidden variables while the other half does.

We first test the accuracy of change point detection. Specifically, we benchmark the proposed copula PELT method with adaptive regularization selection (denoted as ‘‘CPELT-A’’) against three other approaches: copula PELT method with the regularization parameter selected via BIC (denoted as ‘‘CPELT-B’’), the original PELT method with adaptive regularization selection (‘‘PELT-A’’), and the original PELT method with BIC (‘‘PELT-B’’). The results are listed in Table 1, where we show the distribution of the detected number of change points. We also report the mean absolute error (MAE) between the estimated position of change points and the ground truth, for the cases where the correct number of

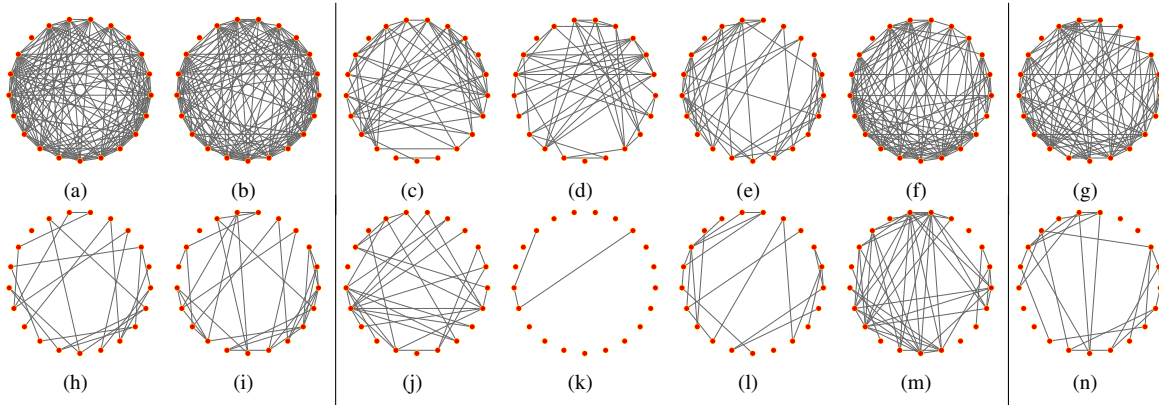


Fig. 2. Results of functional networks resulting from graphical models without hidden variables (a)-(g) and with hidden variables (h)-(n) (with the vertical lines denoting the onset and end of the seizure)

Table 1. Comparison of different methods for accuracy of change point (cp) detection

| Methods | Distribution of detected cp no. | | | | MAE (cp no. = 3) |
|---------|---------------------------------|-----|----|------|------------------|
| | 0 | 1 | 2 | 3 | |
| CPELT-A | 0% | 0% | 0% | 100% | 1.4 |
| CPELT-B | 100% | 0% | 0% | 0% | |
| PELT-A | 0% | 31% | 1% | 68% | |
| PELT-B | 91% | 9% | 0% | 0% | |
| | | | | | |

change points is inferred. Clearly, the proposed CPELT-A method greatly outperforms other methods, reliably identifying the number and the position of change points. On the other hand, CPELT-B and PELT-B seriously underestimate the number of change points; BIC does not seem to be suitable here for regularization selection. The PELT-A method can often detect the correct number of change points (in 68% of the cases), but the inferred position of change points is not as accurate compared to CPELT-A. We also validated the graphical models inferred from the stationary segments between change points obtained from CPELT-A. From our results (not shown here for the sake of brevity), we observe that the proposed methods can reliably recover the true graph (precision matrices generated in Step 1), as reported in [4, 8].

4.2. Real Data

We now apply the proposed model to scalp electroencephalograms (EEG) recorded during epilepsy seizures [21]. The EEG time series were collected from 23 pediatric patients (5 males, ages 3-22; and 17 females, ages 1.5-19) with intractable seizures, using the international 10-20 system at a sampling frequency of 256 Hz. During the recordings, the patients experienced 173 events that were judged to be clinical seizures by experts. Here, we randomly select five events and extract EEG data starting 60s before seizure onset and ending 30s after seizure termination. The EEGs are band-pass filtered with a digital 3rd order Butterworth filter between 3 and 40Hz. We then down sample the signal to retain one fourth of all the samples, to limit the computational complexity.

The results are summarized in Table 2. We observe that the proposed method detects change points near the start and end of seizures, as one would expect. Typically there is change point within two seconds from the start or end of a

Table 2. Results for seizure data

| No. | True period | | Estimated results: change points and network density | | | | | | | |
|-----|-------------|-----|------------------------------------------------------|-------|--------|--------|--------|--------|--------|----|
| | | | change points | 64.34 | 72.09 | 77.45 | 90.75 | | | |
| 1 | 60 | 76 | no. of edges (glasso) | 142 | 106 | 104 | 101 | 130 | | |
| | | | no. of edges (HVGGM) | 34 | 12 | 40 | 23 | 23 | | |
| | | | change points | 35.23 | 60.78 | 72.69 | 110.44 | | | |
| 2 | 60 | 109 | no. of edges (glasso) | 140 | 131 | 111 | 119 | 126 | | |
| | | | no. of edges (HVGGM) | 18 | 50 | 18 | 33 | 29 | | |
| | | | change points | 60.88 | 69.42 | 98.52 | 119.52 | | | |
| 3 | 60 | 100 | no. of edges (glasso) | 146 | 108 | 149 | 135 | 125 | | |
| | | | no. of edges (HVGGM) | 25 | 26 | 28 | 33 | 33 | | |
| | | | change points | 61.48 | 117.02 | 118.48 | 120.25 | 122.66 | 141.05 | |
| 4 | 60 | 123 | no. of edges (glasso) | 137 | 135 | 59 | 48 | 52 | 106 | 91 |
| | | | no. of edges (HVGGM) | 23 | 34 | 32 | 3 | 22 | 50 | 24 |
| | | | change points | 30.00 | 59.16 | 66.63 | 94.20 | 141.47 | | |
| 5 | 60 | 142 | no. of edges (glasso) | 104 | 90 | 62 | 105 | 123 | 114 | |
| | | | no. of edges (HVGGM) | 16 | 30 | 16 | 19 | 54 | 19 | |
| | | | change points | | | | | | | |

seizure, indicating that the proposed method may be further developed for detecting the onset and ending of seizures.

We also inferred the functional network in each stationary time segment by using the Gaussian copula graphical models with and without hidden variables. As an illustration, we show the networks without and with hidden variables for the 4th data set in Fig. 2. Kramer *et al.* [9] analyzed the dynamics of functional networks through the entire seizure in intracranial electrocorticogram (ECoG) recordings. They found that the networks are dense at seizure onset and termination, but sparse during the middle portion of the seizure. Interestingly, although scalp EEG is more noisy than intracranial ECoG recordings, the proposed method can find the same pattern as observed in [9]. Moreover, the inferred number of stationary segments for all 5 data sets are similar, implying that seizures undergo characteristic progressions that can be stretched or compressed in time, as assumed in [9]. Remarkably, in contrast to [9], the proposed method allows us to infer the number and location of change points in an automated fashion.

5. CONCLUSIONS

An effective method is proposed to infer abruptly changing network structures from piecewise-stationary time series. A low-complexity PELT-based algorithm is introduced to detect change points. Next graphical models (with and without hidden variables) are inferred for each stationary time segment. Numerical results for synthetic and real data demonstrate the utility of the proposed model.

6. REFERENCES

- [1] A. T. Ihler, S. Krishner, M. Ghil, A. W. Robertson, and P. Smyth, “Graphical Models for Statistical Inference and Data Assimilation,” *Physica D* vol. 230, pp. 72–87, 2007.
- [2] H.-A. Loeliger, J. Dauwels, J. Hu, S. Korl, P. Li, and F. Kschischang, “The factor graph approach to model-based signal processing,” *Proceedings of the IEEE* 95(6), pp. 1295–1322, 2007.
- [3] H. Liu, J. Lafferty, and L. Wasserman, “The Nonparamoral: Semiparametric Estimation of High Dimensional Undirected Graphs,” *Journal of Machine Learning Research*, pp. 2295–2328, 2010.
- [4] H. Yu, J. Dauwels, and X. O. Wang, “Copula Gaussian Graphical Models with Hidden Variables,” *Proceedings of ICASSP 2012*, pp. 2177–2180, 2012.
- [5] J. Dauwels, H. Yu, X. O. Wang, F. Vialatte, C. Latchoumane, J. Jeong, and A. Cichocki, “Inferring Brain Networks through Graphical Models with Hidden Variables”, *Machine Learning and Interpretation in Neuroimaging, Lecture Notes in Computer Science, Springer*, pp. 194–201, 2012.
- [6] H. Yu, J. Dauwels, X. Zhang, S. Y. Xu, and W. I. T. Uy, “Copula Gaussian Multiscale Graphical Models with Application to Geophysical Modeling,” *Proceedings of 15th International Conference on Information Fusion*, pp. 1741–1748, 2012.
- [7] H. Yu, Z. Choo, W. I. T. Uy, J. Dauwels, and P. Jonathan, “Modeling Extreme Events in Spatial Domain by Copula Graphical Models,” *Proceedings of 15th International Conference on Information Fusion*, pp. 1761–1768, 2012.
- [8] J. Dauwels, H. Yu, S. Y. Xu, and X. O. Wang, “Copula Gaussian Graphical Model for Discrete Data”, *Proceedings of ICASSP 2013*, pp. 2177–2180, 2013.
- [9] M. A. Kramer, U. T. Eden, E. D. Kolaczyk, R. Zepeda, E. N. Eskandar, and S. S. Cash, “Coalescence and Fragmentation of Cortical Networks during Focal Seizures”, *The Journal of Neuroscience*, pp. 10076–10085, 2010.
- [10] X. Xuan, and K. Murphy, “Modeling Changing Dependency Structure in Multivariate Time Series”, *Proceedings of the 24th ICML*, 2007.
- [11] I. Cribben, R. Haraldsdottir, L. Y. Atlas, T. D. Wager, and M. A. Lindquist, “Dynamic Connectivity Regression: Determining State-related Changes in Brain Connectivity”, *NeuroImage* 61, pp. 907–920, 2012.
- [12] M. Lavielle, and G. Teyssi re, “Adaptive Detection of Multiple Change-Points in Asset Price Volatility”, in: G. Teyssi re and A. Kirman (Eds), *Long-Memory in Economics*, Springer, pp. 129–156, 2005.
- [13] D. Angelosante, and G. B. Giannakis, “Sparse Graphical Modeling of Piecewise-Stationary Time Series”, *Proceedings of ICASSP 2011*, pp. 1960–1963, 2011.
- [14] R. Killick, P. Fearnhead, I. A. Eckley, “Optimal detection of changepoints with a linear computational cost,” *Journal of the American Statistical Association* vol. 107, no. 500, pp. 1590–1598, 2012.
- [15] A. Dobra and A. Lenkoski, “Copula Gaussian Graphical Models and their Application to Modeling Functional Disability Data,” *Annals of Applied Statistics*, vol. 5, No. 2A, pp. 969–993, 2011.
- [16] J. Friedman, T. Hastie, and R. Tibshirani, “Sparse Inverse Covariance Estimation with the Graphical Lasso,” *Biostatistics* vol. 9, pp. 432–441, 2008.
- [17] S. Li, L. Hsu, J. Peng and P. Wang, “Bootstrap Inference for Network Construction,” *Annals of Applied Statistics*, vol. 7, no. 1, pp. 391–417, 2013.
- [18] V. Chandrasekaran, P. A. Parrilo, and A. S. Willsky, “Latent Variable Graphical Model Selection via Convex Optimization,” *The Annals of Statistics*, vol. 40, no. 4, pp. 1935–1967, 2012.
- [19] N. Meinshausen, P. B uhlmann, “Stability Selection,” *Journal of the Royal Statistical Society*, vol. 72, Series B, pp. 417–473, 2010.
- [20] M. Lavielle, and G. Teyssi re, “Detection of Multiple Change-Points in Multivariate Time Series,” *Lithuanian Mathematical Journal*, vol. 46, pp. 287–306, 2006.
- [21] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C. K. Peng, H. E. Stanley, “PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals,” *Circulation* vol. 101, pp. 215–220, 2000.