Beyond linear dynamic functional connectivity: a vine copula change point model

Xin Xiong
Department of Biostatistics
Harvard T. H. Chan School of Public Health
and
Ivor Cribben *
Department of Accounting and Business Analytics
Alberta School of Business

August 7, 2021

Abstract

To estimate dynamic functional connectivity for functional magnetic resonance imaging (fMRI) data, two approaches have dominated: sliding window and change point methods. While computationally feasible, the sliding window approach has several limitations. In addition, the existing change point methods assume a Gaussian distribution for and linear dependencies between the fMRI time series. In this work, we introduce a new methodology called Vine Copula Change Point (VCCP) to estimate change points in the functional connectivity network structure between brain regions. It uses vine copulas, various state-of-the-art segmentation methods to identify multiple change points, and a likelihood ratio test or the stationary bootstrap for inference. The vine copulas allow for various forms of dependence between brain regions including tail, symmetric and asymmetric dependence, which has not been explored before in the dynamic analysis of neuroimaging data. We apply VCCP to various simulation data sets and to two fMRI data sets: a reading task and an anxiety inducing experiment. In particular, for the former data set, we illustrate the complexity of textual changes during the reading of Chapter 9 in *Harry Potter and the Sorcerer's Stone* and find that change points across subjects are related to changes in more than one type of textual attributes. Further, the graphs created by the vine copulas indicate the importance of working beyond Gaussianity and linear dependence. Finally, the R package **vccp** implementing the methodology from the paper is available from CRAN.

Keywords: Time-varying networks; graphical models; computational statistics; networks: fMRI

*To whom correspondence should be addressed.

1 Introduction

Functional magnetic resonance imaging (fMRI) experiments yield high dimensional data sets that contain complex spatial correlations, often referred to as functional connectivity (FC) networks (see for example, Cribben and Fiecas, 2016, for a review). These FC networks have been previously studied to expose important characteristics of brain function and individual variations in cognition and behavior. In particular, Greicius et al. (2004); Menon (2011); Bakhtiari et al. (2017) and Hart et al. (2018) showed that neurological disorders disrupt the FC structure of the brain and Hrybouski et al. (2021) studied the changes in FC in healthy aging subjects.

Recently, there has been a surge in the development of several new statistical methods for investigating the extent to which FC networks change over time. These changes are commonly referred to as time-varying, or dynamic FC in the field of neuroimaging. This reconstruction of dynamic FC has a major impact on the understanding of the functional organization of the brain. Similar to other biological networks, understanding the complex, dynamic organization and characterizations of the brain can lead to profound clinical implications (Bullmore and Sporns, 2009).

To estimate dynamic FC for fMRI data, two approaches have dominated: sliding window and change point methods. For the former, a sliding window of pre-specified length is defined and the correlations between distinct regions of the brain throughout the duration of the window are estimated. While the sliding window approach is computationally feasible, it also has limitations (Hutchison et al., 2013). For example, the choice of window size is crucial and sensitive, as different window sizes can lead to quite different FC patterns. Another disadvantage is that equal weight is given to all k neighbouring time points and 0 weight to all the others. Hence, researchers considered change point methods that partition the time series into optimal windows. Determining change points may also reveal properties of brain networks as they relate to experimental stimuli and disease processes. There exists an extensive literature and a long history on change point detection. The most widely discussed problems have been concerned with finding multiple change points in univariate time series (Inclan and Tiao, 1994; Chen and Gupta, 1997). Recently, the multiple change

point detection problem in multivariate time series has received some attention especially in non-stationary practical problems. To detect changes in the covariance matrix of a multivariate time series, Aue et al. (2009) introduced a method using a nonparametric CUSUM type test, Dette and Wied (2016) proposed a test where the dimension of the data is fixed while Kao et al. (2018) considered the case where the dimension of the data increases with the sample size. Sundararajan and Pourahmadi (2018) proposed a new method for detecting multiple change points in the covariance structure of a multivariate piecewise-stationary process, while Dette and Gösmann (2020) developed a likelihood ratio approach to detect change points for a general class of parameters including changes in correlations.

In other work, Barnett and Onnela (2016) considered a method for detecting change points in correlation networks. Gibberd and Nelson (2014) identifies both change points and the graphical dependency structure in multivariate time series. Li et al. (2019) considered multiple structural breaks in large contemporaneous covariance matrices of high dimensional time series satisfying an approximate factor model. Cho and Fryzlewicz (2015) segmented the multivariate time series into partitions based on the second-order structure.

For neuroscience applications, Cribben et al. (2012, 2013) first introduced the idea of detecting FC change points by introducing Dynamic Connectivity Regression for detecting multiple change points in the precision matrices (undirected graphs) from a multivariate time series. Schröder and Ombao (2019); Kirch et al. (2015); Barnett and Onnela (2016); Cribben and Yu (2017); Kundu et al. (2018); Dai et al. (2019); Ofori-Boateng et al. (2021); Anastasiou et al. (2020); Ondrus et al. (2021) among others have since introduced new network change point methods. However, all of these methods have limitations. The most obvious is that they all consider a Gaussian distribution for and linear dependencies between the fMRI time series.

In this paper, we introduce a new methodology, called *Vine Copula Change Point* (VCCP), to estimate multiple change points in the FC structure between brain regions. The new method combines vine copulas, various state-of-the-art segmentation methods to identify multiple change points, and a likelihood ratio test or the stationary bootstrap for

inference. The proposed VCCP method has the following unique and significant attributes. First, it is the first statistical method that has applied vine copulas to neuroimaging data. (Fontaine et al., 2020, apply copula models to local field potential data of rats and focus on static dependence). Second, it is the first statistical method that considers change points in vine copulas. Vine copulas split the multivariate distribution into marginal distributions and a dependence measure without a need to assume that the data follows a parametric distribution (e.g., Gaussian). Hence, VCCP is very flexible. Third, VCCP is the first method that allows us to describe various forms of dependence structures such as tail, symmetric and asymmetric dependence between fMRI time series in a dynamic fashion, which has not been explored before in the analysis of neuroimaging data. Therefore, VCCP is capable of modeling a wider range of dependence patterns and hence allows for the detection of more change points. Fourth, the layers of a vine (the number of trees) can be cut, which leads to a simplified and sparse dependence matrix which reduces computation time when the dimension of the data expands. Fifth, VCCP allows for the exploration of network dynamics during a reading fMRI experiment (Chapter 9 in *Harry Potter and the Sorcerer's Stone*) and an anxiety inducing fMRI experiment. In the former, it detects change points across subjects that coincide to more than one type of textual attributes. Further, the graphs created by the vine copulas indicate the importance of working beyond Gaussianity and linear dependence. Sixth, as VCCP uses various state-of-the-art segmentation methods, the paper can be viewed as an extensive comparison of these methods. Seventh, while motivated by fMRI data, VCCP could also be applicable to electroencephalography (EEG), magnetoencephalography (MEG) and electrocorticography (ECoG) data, and other time series applications where the network structure is changing. Finally, the R package **vccp** implementing the VCCP methodology is available from CRAN (Xiong and Cribben, 2021).

This paper is organized as follows. In Section 2 we provide a background on copulas, vine copulas and explain the setup of our proposed new methodology, Vine Copula Change Point (VCCP). In Section 3, we describe the simulated data sets with known change points locations and the two fMRI data sets. We present the performance of VCCP in Section 4, have a discussion in Section 5 before concluding in Section 6.

2 Methods

2.1 Copulas

Table 1 provides a summary of the notation used in the paper. We now introduce copulas

Table 1: Notation for the Vine Copula Change Point (VCCP) model.

Index	Meaning
p	Data dimension
$X = (x_1, x_2, ..., x_p)$	A p-dimensional random variable
$C(\cdot, \cdots, \cdot)$	A copula function
$c(\cdot, \cdots, \cdot)$	A copula density function
n	Sample size
T	The length of a time series
δ	The minimum distance between two adjacent candidate change points

and vine copulas. A copula "couples" marginal distributions into a joint distribution. Hence, copulas allow for the independent construction of joint distributions and marginal distributions. This is convenient as marginal distributions in many cases can be adequately estimated from data, whereas dependence information involves summary indicators and judgment. More specifically, let $\mathbf{X} = (X_1, ..., X_p)$ be a p-dimensional random variable or a multivariate time series from p regions of interest (ROIs) with cumulative distribution function (cdf) $F(x_1, ..., x_p)$. A copula, which is also a multivariate cdf, serves as the link that connects the marginal distributions of $\mathbf{X}$ to its multivariate cdf, $F(x_1, ..., x_p)$. A formal definition of a copula is given by:

Definition 1 (**Copula**) *A p-dimensional copula, C, is a multivariate cdf defined on $[0,1]^p$, $C : [0,1]^p \to [0,1]$, and its univariate margins have a uniform distribution.*

According to Sklar (1959), assuming the marginals of $\mathbf{X}$ are continuous, every continuous cdf F has a unique copula C that satisfies

$$F(x_1, ..., x_p) = C(F_1(x_1), ..., F_p(x_p)). \tag{1}$$

The availability of high-dimensional copula models is limited, but there are several parametric

bivariate copula models such as Gauss, t, Clayton, Gumbel, and Frank copula (see the Supplementary Materials for definitions) to name but a few (the accompanying R package **vccp** allows for many more choices of copulae). These have led to the development of hierarchical models, constructed from cascades of bivariate copula models, called pair-copula constructions (PCCs) or vine copulas.

2.2 Vine copulas

PCCs is a method for building a multivariate distribution by firstly decomposing a joint density function into a sequence of conditional densities, such as

$$f(x_1, ..., x_p) = f(x_1) \cdot f(x_2|x_1) \cdot ... \cdot f(x_p|x_1, ..., x_{p-1}), \tag{2}$$

making the construction of a possible complex dependence structure both flexible and tractable. Then using the definition of conditional densities and the derivative of (1), Joe (1996) proved that $f(x_1, ..., x_p)$ can be further decomposed to:

$$f(x_1, ..., x_p) = \prod_{i=1}^{p} f_i(x_i) \cdot \prod_{i=2}^{p} \prod_{j=1}^{i-1} c_{ij|(j+1)...(i-1)}(F(x_i|x_{j+1}, ..., x_{i-1}), F(x_j|x_{j+1}, ..., x_{i-1})), \tag{3}$$

where f_i denotes the marginal density of x_i, $F(x_i|x_{j+1}, ..., x_{i-1})$ is its univariate conditional distribution function, and $c_{ij|(j+1)...(i-1)}$ is the density of the conditional pair copula associated with the bivariate conditional distribution of x_i and x_j given the subset, $(x_{j+1}, ..., x_{i-1})$. There are many other decompositions of the joint density based on the assigned conditional sets in (2). In fact, Morales Napoles (2016) proved that there are $\frac{p!}{2} \cdot 2^{\binom{p-2}{2}}$ possible decompositions for a p-dimensional distribution. Bedford and Cooke (2002) organized the PCCs decomposition as a one-to-one graphical model, a sequence of $p-1$ nested trees (where edges correspond to the bivariate copulas), called a regular vine, or R-vine. For tree m ($m = 1, ..., p-1$), define $T_m = (V_m, E_m)$, where V_m and E_m represent the node set and edge set, respectively, then the tree sequence (called the structure of the PCC) is an R-vine if it satisfies a set of conditions guaranteeing that the decomposition is a valid joint density.

To specify a p-dimensional distribution on an R-Vine structure, $\mathcal{V}$, with a node set $\mathcal{N} := N_1, ..., N_{p-1}$ and an edge set $\mathcal{E} := E_1, ..., E_{p-1}$, we need to connect each edge $e = j(e), k(e)|D(e)$ in E_i to a conditional bivariate copula density $c_{j(e),k(e)|D(e)}$, where $j(e)$ and $k(e)$ are the <u>conditioned set</u>, while $D(e)$ is the <u>conditioning set</u>. Finally, the joint copula density can be written as the product of all pair-copula densities $c = \prod_{m=1}^{p-1} \prod_{e \in E_m} c_{j(e),k(e)|D(e)}$. To estimate vine copulas, it is common to follow a sequential approach from higher trees (more nodes) to lower trees (less nodes), since there are $\frac{p!}{2} \cdot 2^{\binom{p-2}{2}}$ possible R-vine structures for a p-dimensional cdf. If the vine structure is known, the pair-copulas of the first tree, T_1, can be estimated directly from the data. This is not possible for the other trees in the sequence, since data from the densities $c_{j(e),k(e)|D(e)}$ are not observed. However, it is straightforward to create "pseudo-observations" using appropriate data transformations, resulting in the following estimation procedure: estimate tree T_1 directly from the data, for each edge in the tree, estimate all pairs, construct pseudo-observations for the next tree and then iterate. As the tree sequence $T_1, T_2, ..., T_{p-1}$ is a regular vine, it guarantees that at each step, all required pseudo-observations are available. In this work, in order to specify every tree's connecting network, we utilize Algorithm 1 (Dißmann, 2010).

Two special cases of an R-vine are the canonical-vine (C-vine) and the drawable-vine (D-vine). An R-vine is called a *C-vine* if there is one unique node (center node) of degree $p - i$ in each tree T_i $(i = 1, ..., p - 1)$ (in every tree, one node serves as the center that links all the other nodes). An R-vine is called a *D-vine* if the maximal degree of freedom of nodes in each tree is 2 (no node in any tree can be linked with more than 2 other nodes). Figure 1 shows examples of a D-vine and a C-vine that are based on 5 variables (or 5 ROIs). The colored point on each tree of the C-vine represents the center node, which is also indicated by the dashed line on the previous tree.

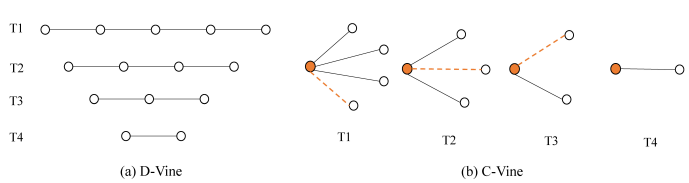

Figure 1: A connected graph for a D-vine (left) and a C-vine (right) that is based on 5 variables (or 5 ROIs).

Algorithm 1: A sequential method to select a vine copula model

```
Input: An n × p matrix X = (x_1, ..., x_p).
for i ← 1 : (p − 1) do
    Let (y_1, ..., y_{p−i+1}) = (x_1, ..., x_{p−i+1});
    Calculate Kendall's τ between y_1, ..., y_{p−i+1};
    Build T_i by selecting the tree that maximizes the sum of Kendall's absolute τs
      without violating the R-vine definition (MST method; Cormen et al., 2009);
    m = 1;
    for s ← 2 : (p − i + 1) do
        for t ← 1 : (j − 1) do
            if y_s and y_t are connected in T_i then
                Select the best pairwise copula c for y_s and y_t from the candidate
                  copulas;
                Estimate parameter(s) in the copula function;
                Calculate the pseudo observation x_m based on y_s, y_t and c;
                m = m + 1;
            end
        end
    end
end
Output: A vine copula model, (T_1, ..., T_{p−1}), with each edge specified by the best
        fitting copula.
```

2.3 Segmentation Methods

In this work, we are not only interested in estimating vine copulas but in estimating change points in the vine copula structure, where we assume that those change points are discrete and at some distance from each other. In order to find the change points, VCCP uses the Bayesian Information Criterion (BIC: Schwarz et al., 1978) metric. BIC is equal to $-2\log\hat{L} + k * \log(n)$, where $\hat{L}$ is the maximized value of the likelihood function of the vine copula model, n is the number of observations and k is the number of parameters estimated in the vine copula. However, a time axis of length T and r change points can be divided into $\frac{(T-1)!}{(r-1)!*(T-r)!}$ possible sub-partitions (with unknown r). It is clear that as T increases, the number of sub-partitions explodes. Hence, VCCP uses four segmentation methods to simplify the process, which we now detail.

Binary segmentation (BS) is a generic segmentation method. Assuming an unknown number of change points, BS can be used to find multiple change points with low compu-

tational complexity. BS begins by searching the entire time course for one change point. We set an auxiliary time point i that shifts from $(\delta + 1)$ to $(T - \delta - 1)$, where i indicates a candidate change point and δ is a parameter that represents the minimal distance between two candidate change points. δ needs to be sufficiently large enough to estimate a stable R-vine structure, but also small enough in order to not miss candidate change points. Thus, three R-vine copula specifications are estimated on the data: between time points $(1 : i - 1)$, $(i : T)$ and $(1 : T)$. The reduced BIC as a result of partitioning the time course into two intervals is calculated as i moves across time. We choose time t_1, which corresponds to the largest BIC reduction as the first candidate change point. After inference, if it is deemed a change point, the time course is split into two sub-intervals (where the binary name originates from) and a similar search is repeated on each of them. The recursion does not stop until BIC does not decrease by partitioning the time course, or the candidate point corresponding to the maximum BIC reduction fails the inference test. An example of BS is shown in Figure 2. As 'arguably the most widely used change point search method', the BS method is conceptually simple and is very easy to use. Despite this, as it is a 'greedy' procedure, each step except the first depends on the previous steps, which are never re-visited. This sequential method leads to inconsistent estimates when the minimum spacing between two adjacent change points is of order less than $T^{3/4}$. We propose an

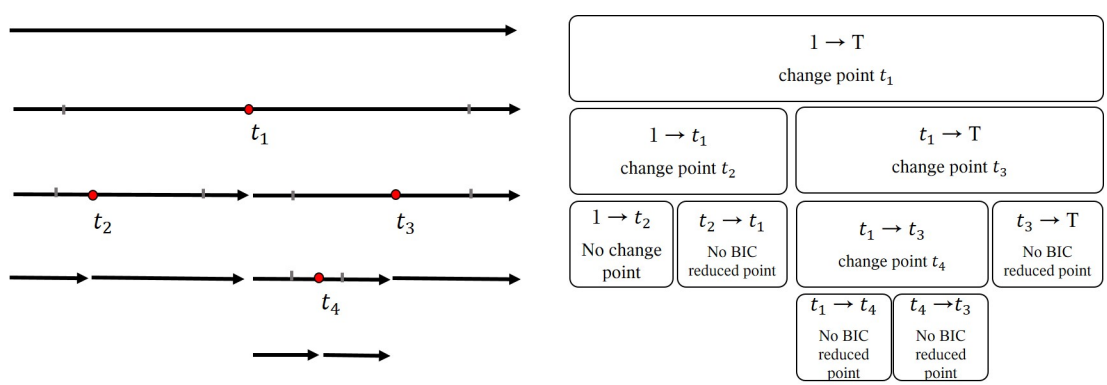

Figure 2: An example of Binary Segmentation (BS). Here, t_1 corresponds to the first candidate change point. It is deemed a change point, hence the data is then split into two sub-intervals. On each sub-segment, we restart a similar search to further segment the data.

adapted BS method (NBS), which is similar to the old classic BS (OBS), but does not carry out inference on each change point until the segmentation has been exhausted (the stopping

criteria for NBS is whether the maximal BIC reduction is larger than 0). Once no further candidate change points reduce the BIC score, NBS orders the candidate change points in sequence and recalculates the reduced BIC:

$$\Delta\text{BIC}_{t_i} = \text{BIC}_{t_{i-1}:(t_{i+1}-1)} - [\text{BIC}_{t_{i-1}:(t_i-1)} + \text{BIC}_{t_i:(t_{i+1}-1)}] \tag{4}$$

where $1 = r_0 < r_1 < ... < t_r < t_{r+1} = T$ and t_1, ..., t_r are the candidate change points detected by NBS. Then, inference is performed on each non-zero BIC reduction in partitions $[t_{i-1}, t_{i+1}]$.

VCCP also includes more recent segmentation methods, Wild Binary Segmentation (WBS: Fryzlewicz, 2014) and moving sum (MOSUM: Eichinger and Kirch, 2018). Briefly, WBS starts by randomly drawing S sub-samples $Y_s = (X_{b_s}, X_{b_s+1}, \cdots, X_{e_s})$, where b_s, e_s and s are integers such that $1 \leqslant b_s < e_s \leqslant T$, $s = 1, ..., S$. Then it estimates candidate change point t_s^c on each sub-sample Y_s based on a single binary search (see t_1 in Figure 2) and a corresponding statistic, $W_s = \Delta\text{BIC}_{t_s^c}$. After, WBS selects the candidate change point as

$$t_i = t_s^c, s = \begin{cases} \underset{m\in\{1,...,S\}}{\arg\max}\ \{W_s\} \text{ if } i = 1; \\ \underset{s\in\{1,...,S\}}{\arg\max}\ \left\{W_s : \{t_1, ..., t_{i-1}\} \bigcap [b_s, e_s] = \varnothing\right\} \text{ if } i > 1. \end{cases} \tag{5}$$

This segmentation method exhausts when t_i is not a change point or each random interval ($[b_s, e_s]$) covers at least one detected change point. To simplify computation, we remove sub-intervals that contain less than 2δ time points before computing the BIC reduction. By 'randomly localizing' the range of BIC reductions, we overcome the issue of greedy searching for certain configurations of multiple change points. Also, since intervals are drawn randomly, we avoid issues such as choosing the appropriate window size prevalent in other segmentation methods.

MOSUM (Eichinger and Kirch, 2018) detects change points in the mean and variance based on the moving sum statistics for a univariate time series. For multivariate time series with a changing network structure, it is not possible to directly apply MOSUM.

Hence, we adapt MOSUM for VCCP as follows. First, we calculate the BIC reduction series within a local range $[i-\delta+1,\, i+\delta]$. Second, we apply MOSUM to this newly created time series, resulting in several boundary points $(a_1, ..., a_r)$. Third, we separate the entire time axis into intervals $[m_i, m_{i+1}], i = 0, ..., r, m_0 = 1, m_{r+1} = T$ and record time points $t_1, ..., t_r$ corresponding to the local maximum BIC reduction in the different intervals. These represent the candidate change points.

2.4 Inference

To perform inference on the candidate change points, VCCP uses the Vuong test (Vuong, 1989) or the stationary bootstrap (SB: Politis and Romano, 1994). The Vuong test is used to compare non-nested models. In our R-vine setting, non-nested models can refer to different copulas for edges of the same tree structure, different vine structures, or a combination of the two. For a random variable X, the Vuong test uses the Kullback-Leibler information criterion to compare two models $F_\beta = f(X|\beta)$ and $G_\gamma = g(X|\gamma)$. Suppose β_* and γ_* are the true parameters of the two candidate models and $\hat{\beta}_n$ and $\hat{\gamma}_n$ are their corresponding maximum likelihood estimates. Then the null hypothesis is

$$H_0 : E_0\left[\log\left(\frac{f(X|\beta_*)}{g(X|\gamma_*)}\right)\right] = 0 \tag{6}$$

where E_0 is the expectation under H_0, and the likelihood ratio test statistic is

$$LR_n(\hat{\beta}_n, \hat{\gamma}_n) = \log\left[\frac{\prod_{i=1}^n f(x_i|\hat{\beta}_n)}{\prod_{i=1}^n g(x_i|\hat{\gamma}_n)}\right] = \sum_{i=1}^n \log\left(\frac{f(x_i|\hat{\beta}_n)}{g(x_i|\hat{\gamma}_n)}\right). \tag{7}$$

Vuong (1989) proved that this test statistic is asymptotically normal, which leads to the asymptotic test, $\hat{V}_n := \frac{LR_n(\hat{\beta}_n, \hat{\gamma}_n)}{\sqrt{n}\hat{\omega}_n} \xrightarrow{D} N(0,1)$ as $n \to \infty$ and

$$\hat{\omega}_n^2 = \frac{1}{n}\sum_{i=1}^n \left[\log\left(\frac{f(x_i|\hat{\beta}_n)}{g(x_i|\hat{\gamma}_n)}\right)\right]^2 - \left[\frac{1}{n}\sum_{i=1}^n \log\left(\frac{f(x_i|\hat{\beta}_n)}{g(x_i|\hat{\gamma}_n)}\right)\right]^2.$$

If we do not reject H_0 under significance level α, the conclusion is that both models fit the data equally well. On the contrary, if we were to reject H_0 and the statistic $\hat{V}_N$ is

positive, the test suggests that model F_β on the numerator of the likelihood ratio statistic is superior, while a negative value supports model G_γ on the denominator. Similar to BIC, Vuong (1989) included the following a penalty term in the likelihood ratio statistic, $LR_n^{Schwarz}(\hat{\beta}_n, \hat{\gamma}_n) = LR_n(\hat{\beta}_n, \hat{\gamma}_n) - \frac{\log(n)}{2}(p+q)$, where p and q are the number of parameters in models F_β and G_γ, respectively.

The SB is an adaptation of the block bootstrap, which allows us to deal with the auto-correlation in time series by retaining the dependence structure of the data. Unlike resampling blocks of data with a fixed block size in the block bootstrap, the SB randomly specifies each block size. For any strictly stationary time series $\{X_t, t = 1, ..., T\}$, the SB procedure consists of generating pseudo-samples $X_1^*, ..., X_T^*$ from the sample $X_1, ..., X_T$ by taking the first T elements from $X_{H_1}, ..., X_{H_1+L_1-1}, ..., X_{H_N}, ..., X_{H_N+L_N-1}$ where H_i is an independent and identically distributed sequence of random variables uniformly distributed on {1,...,T} and L_i is an i.i.d. sequence of geometrically distributed random variables with $P(L_i = h) = q(1-q)^{h-1}, h = 1, 2, ...,$ for some $q \in$ (0,1). Note that the mean block size is $\eta = \frac{1}{q}$, where the choice of η in the SB is similar to the choice of the block length in the block bootstrap.

3 Simulated and fMRI data descriptions

In order to show the performance of our VCCP methodology we use several simulated data sets, and two task-based fMRI data sets. To carry out inference on these data sets, we have to adjust the Vuong test (Section 2.4) when we combine it with the segmentation methods. For NBS and MOSUM, the first step is to detect all candidate change points $1 < t_1 < ... < t_k < T$. For interval $I_i = [t_{i-1}, t_{i+1}]$, three vine copulas are estimated based on data in the intervals $[t_{i-1}, t_{i+1}]$, $[t_{i-1}, t_i - 1]$, $[t_i, t_{i+1}]$, which we denote by VC_0, VC_l and VC_r. In each interval, the data are assumed to be i.i.d and that there are no other change points between time points t_{i-1} and t_{i+1} except t_i. If H_0: t_i is not a change point in $[t_{i-1}, t_{i+1}]$ is rejected, then constructing two VC models in the two partitions separately provides a superior fit to the data than a VC model using the entire data set. Hence, we perform a Vuong test on VC_0 and VC_l based on data in $[t_{i-1}, t_i - 1]$ and a Vuong test on

VC_0 and VC_r using data in the interval $[t_i, t_{i+1}]$. If both tests show that VC_0 is worse than either VC_l or VC_r, we conclude that t_i is a change point. For OBS, the only difference in the above process is that the left and right end of the intervals are the endpoints of the partitioned data in each BS search, since OBS runs detection and inference sequentially. For WBS, the intervals are based on the endpoints of randomly chosen sub-samples. The rest of their procedures are the same as NBS and MOSUM.

The only difference between SB and the Vuong test is that instead of conducting likelihood-based inference, SB creates SB pseudo-samples on $[t_b, t_e]$ and compares them to the original BIC reduction value, ΔBIC_{t_i}, where t_i is the candidate change point, $[t_b, t_e] = [t_{i-1}, t_{i+1}]$ for NBS and MOSUM, $[t_b, t_e] = [1_{b_{s_i}}, T_{b_{s_i}}]$ for OBS and $[t_b, t_e] = [t_{b_m}, t_{e_m}]$ for WBS.

In the simulations, we also considered the 5-copula family but we obtained similar results but with slower speed, hence we only show the results from using the Gauss copula.

3.1 Simulated data

Table 2 provides a summary of the simulations, while Figure 3 provides a visual display. We include two data types: multivariate normal (MVN) data and data from a vector autoregression (VAR) model. A VAR model is a generalization of a univariate AR process. Given a p-dimensional multivariate time series $\mathbf{X}$, the lag-1 vector autoregression model, VAR(1), is defined as $\mathbf{X}_t = \Pi_1 \mathbf{X}_{t-1} + \epsilon_t,\ t = 2, ..., T$, where Π_1 is an $(p \times p)$ coefficient matrix and ϵ_t is an $(p \times 1)$ unobservable mean white noise vector process with time invariant covariance matrix. The VAR model is used to reconstruct the linear inter-dependency element prevalent among multivariate time series applications such as fMRI data. In the simulations overall, we attempt to emulate the properties in fMRI data. In particular, we consider no change points (null data), multiple change points, general changes in the correlation/network structure, changes in the strength of the connectivity, pre-whitened (MVN) and un-whitened (VAR) data, and inspired by the experimental fMRI data, an off-on-off pattern. We run each simulation using 100 iterations.

Unless stated otherwise, we have the following settings for the simulations. The minimum

Table 2: A summary of the simulated data sets. MVN and VAR denote multivariate normal and vector autoregression data, respectively. Weak means that only the edge weights are adjusted between change points with the non-zero edges themselves remaining constant over the entire time course, while small means that only small changes in the networks are present between change points.

Sim	Data type	Change type	p	Locations of change points	ROI connectivity
1	MVN	None	15	NA	No change points
2	MVN	Normal	15	61;121;181;241	1-60s: 1-3-5-8-10 & 2-6-9; 61-120s: 1-4-5-7-9; 121-180s: 2-3-4-6-8-10 & 1-7; 181-240s: 2-7-9; 241-300s: 1-5-6 & 7-8-10.
3	MVN	Weak	15	61;121;181;241	1-300s: 1-4-8-14 & 3-6-9; (Changes in connecting strength)
4	MVN	Small	15	76;151;226;301;376;451;526	1-75s & 301-375s & 526-600s: 1-5-8-10-15 & 3-6-13; 76-150s: 3-6-13; 151-225s: 1-5-8-10 & 3-6-13; 226-300s: 1-5-8-10-15; 376-450s: 1-5-8-10-15 & 3-6-13; 451-525s: 1-5 & 3-6-13.
5	VAR	Normal	15	61;121;181;241	1-60s: 1-3-5-8-14 & 4-15; 61-120s: 2-6-10 & 3-9-13; 121-180s: 4-6-11-14 & 5-12; 181-240s: 1-4-7; 241-300s: 8-10-13-15.
6	VAR	Weak	15	61;121;181;241	1-300s: 1-5-10-15 & 2-9-14; (Changes in connecting strength)
7	VAR	Small	10	76;151;226;301;376;451;526	1-75s & 301-375s & 526-600s: 1-5-8-10 & 2-7-9; 76-150s: 2-7-9; 151-225s: 1-5-8 & 2-7-9; 226-300s: 2-7-9; 376-450s: 1-5-8-10 & 2-7-9; 451-525s: 5-8 & 2-7-9.

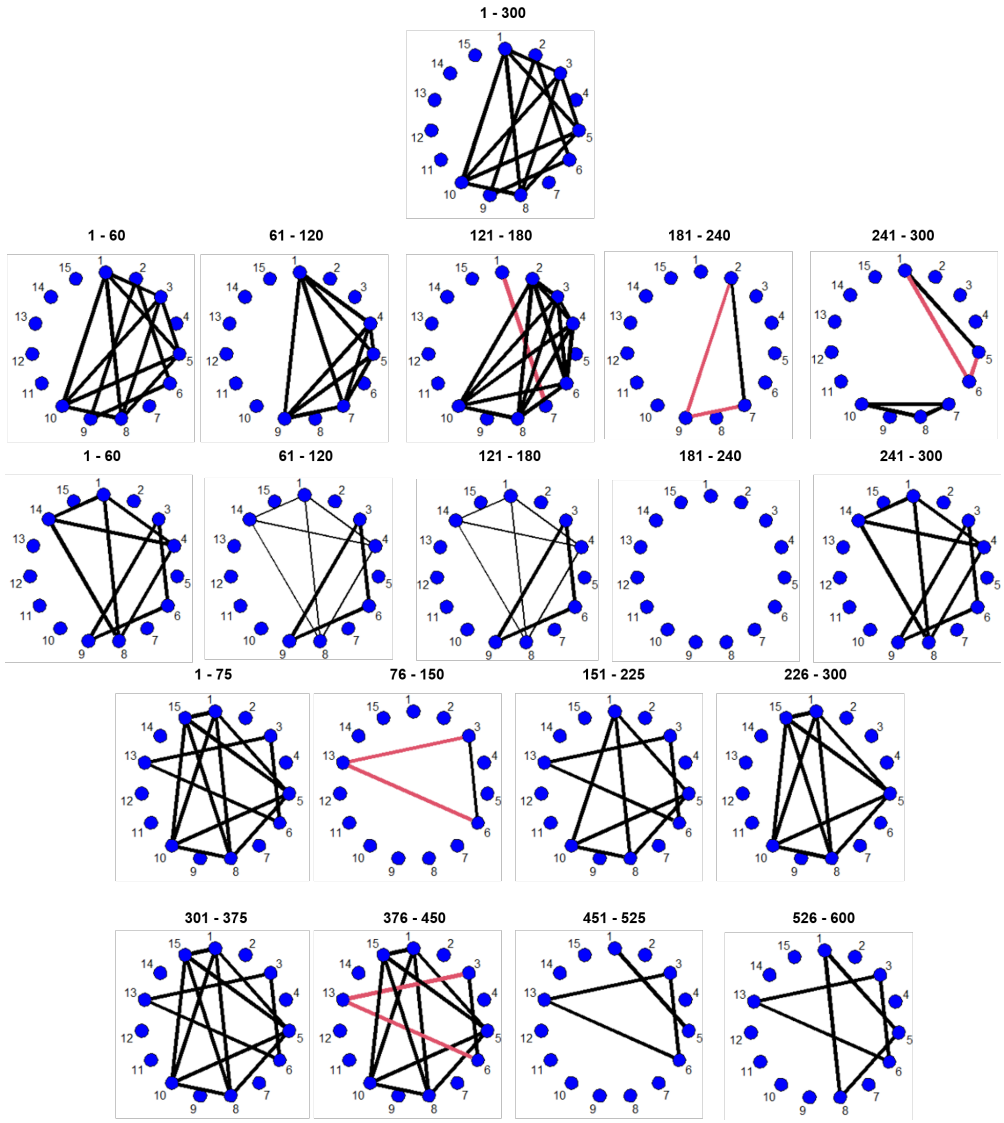

Figure 3: The network structures for the simulations: (first row) i.i.d data with no change points (Simulation 1), (second row) VAR data set with four change points (Simulation 5), (third row) VAR data set with four change points of small magnitude (Simulation 6), and (fourth and five rows) VAR data set with seven change points (Simulation 7). The time series are represented by nodes, black (red) edges infer positive (negative) connectivity, and the strength of connection between the regions is directly related to the thickness of the edges, that is, the thicker the edge the stronger the connection.

distance between change points is set to $\delta = 30$. We fix the tree structure to be D-vine. The significance level of the Vuong test is set to $\alpha = 0.1$, as two tests (left and right partitions) have to be rejected simultaneously before identifying a change point. The average block size and the number of resamples in the SB is set to be $1/0.3$ (or $p = 0.3$) and 100, respectively.

For MOSUM, we set the bandwidth to be $G = 0.1 * T$ and the number of random intervals to be $M = [9 * \log(T)]$ in WBS, where $[x]$ denotes the maximal integer smaller than x.

We compare the True Positive rate (TP: the ability to detect the true change points), False Positive rate (FP: detecting incorrect change points) and False Negative rate (FN: failing to detect the true change points). As a measure of the accuracy of the detected locations in time compared to the location of the true change points, we also provide the scaled Hausdorff distance,

$$d_H = n_s^{-1} \max \left\{ \max_j \min_r \left| t_j - \hat{t}_r \right|, \max_r \min_j \left| t_j - \hat{t}_r \right| \right\},$$

where n_s is the length of the largest segment, $\hat{t}_r$ are the estimated change points and t_j are the true change points. The optimal model obtains a minimum scaled Hausdorff distance. The bias standard we use in the tables is 10 time points, hence detected change points located no more than 10 units away from the true change points are considered a correct detection.

Many of the methods mentioned in the introduction assume Gaussianity and linear dependence for data between the change points, which is equivalent to using a Gauss copula in our simulations and fMRI data sets. However, we also compare VCCP to 5 non-parametric methods, e.divisive (Matteson and James, 2014), e.cp3o, e.cp3o_delta, ks.cp3o and ks.cp30_delta (Zhang et al., 2017), implemented in the ecp package (James and Matteson, 2015). We ignore the e.agglo method since it requires an initial segmentation of the data. The five methods were built for multivariate time series data and multiple change points detection. For e.divisive, we keep *min.size* (the minimum number of observations between change points) parameter to be the default value 30, which is equivalent to $\delta = 30$ in our VCCP setting. The moment index α used for determining the distance between and within segments can be 1 or 2, and the number of change point locations to estimate is based on inference test in each turn. The e.cp3o and kp.cp3o methods use the E-statistic and the Kolmogorov-Smirnov statistic for the goodness-of-fit measure, respectively. Both of these methods requires users to input *Kmax*, the maximum number of possible change points and *min.size*. E.cp3o_delta and ks.cp3o_delta narrow down the window size used

to calculate the complete portion of the approximate test statistic, significantly reducing the computation time. Likewise, it requires $Kmax$, δ (window size) and α. In simulation 2, 3, 5 and 6, $Kmax$ can be 4 or 8 while in simulation 4 and 7, $Kmax$ is set either as 7 or 14.

3.2 fMRI data

We also apply our new VCCP methodology to two task-based fMRI data sets. The first fMRI data set was acquired from 9 right-handed native English speakers (5 females and 4 males) aged 18–40 years, while they read Chapter 9 of *Harry Potter and the Sorcerer's Stone* (one subject's data was discarded due to artifacts). All subjects had read the Harry Potter book series, or seen the movie series prior to participating in the experiment. All the subjects therefore were familiar with the characters and the events of the book, and were reminded of the events leading up to Chapter 9 before the experiment. This chapter was chosen because it involves many characters and spans multiple locations and scenes. The chapter was read using Rapid Serial Visual Presentation (RSVP): the words of the chapter were presented one by one in the center of the screen, for 0.5 s each. The sampling rate of fMRI acquisition was 2 seconds per observation, hence, four consecutive words were read during the time it took to scan the whole brain once. The 45-minute experiment was divided into four runs, each starting with 20 seconds (= 10 TRs) of rest, and ending with 10 seconds (= 5 TRs) of rest. After discarding data collected during the resting periods, we obtained 1290 scans, during which approximately 5200 words were presented. Figure 4 shows the main plotline within each run. Data was preprocessed as in Wehbe et al. (2014). Fourteen ROIs defined using the Automated Anatomical Atlas (AAL: Tzourio-Mazoyer et al., 2002) were extracted from the data set (Table 3). These regions contain a variety of voxels that can distinguish between the literary content of two novel text passages based on neural activity while these passages are being read and are important for understanding reading processes (Wehbe et al., 2014). In this work, we focus on exploring whether the FC network of the 14 ROIs changes as the subjects read. We focus on the varying textual features about story characters (e.g., emotion, motion and dialog) that the dynamic networks encode.

A description of the second fMRI data set (SET) can be found in the Supplementary

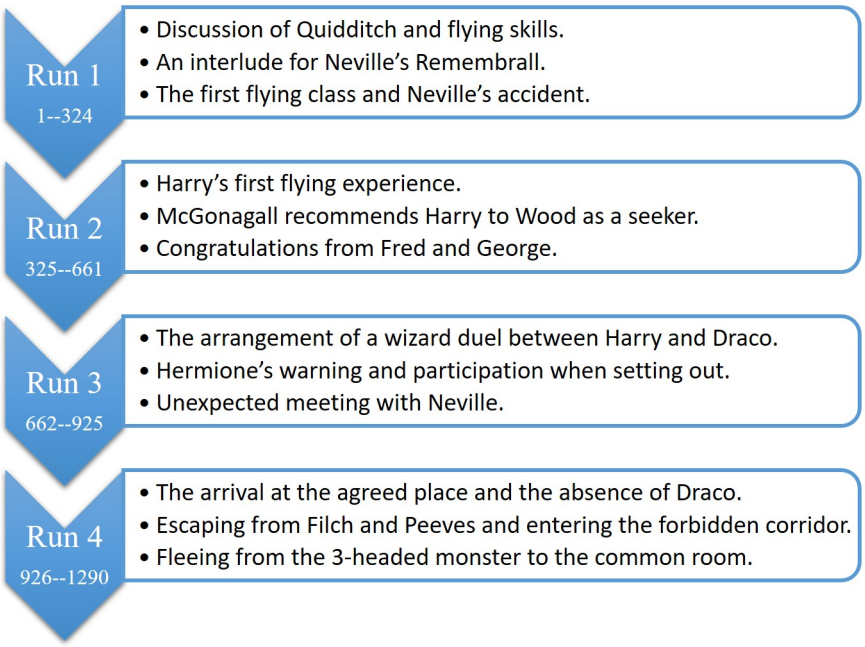

Figure 4: The main plot lines within each run of Chapter 9 of *Harry Potter and the Sorcerer's Stone*.

Table 3: Information on the 14 ROIs from the AAL atlas extracted from subjects in the Harry Potter task-based fMRI data set.

ROI id	**Regions**	**number of voxels**	**Label**
1	Angular gyrus	838	AG
2	Fusiform gyrus	1447	F
3	Inferior temporal gyrus	2013	IT
4	Inferior frontal gyrus, opercular part	747	IFG 1
5	Inferior frontal gyrus, orbital part	1008	IFG 2
6	Inferior frontal gyrus, triangular part	1355	IFG 3
7	Middle temporal gyrus	2795	MT
8	Occipital lobe	6371	O
9	Precental gyrus	2030	PCG
10	Precuneus	2014	PC
11	Supplementary motor area	1322	SM
12	Superior temporal gyrus	1635	ST
13	Temporal pole	1353	TP
14	Supramarginal gyrus	550	SG.R

Materials. For both fMRI data sets, we used the same settings as the simulations (Section 3.1), except for the first fMRI data set, we set $\delta = 40$. In addition, since the experiment lasted 45 minutes, and given the computational burden for BS, we run VCCP on each of the four runs separately.

4 Results

4.1 Simulations

The results for the multivariate normal (MVN) data are displayed in the Supplementary Materials. Table 4 shows the results from applying all variations of VCCP and the two best ecp methods (e.cp3o and ks.cp3o with $Kmax$= 8) to the VAR data with 4 change points (Simulation 5). For conciseness, we only display the best two settings of non-parametric methods. For all other results see the Supplementary Materials. The NBS.D.SB (NBS segmentation, D-vine, and stationary bootstrap for inference) combination has the best performance, it almost perfectly recognizes the four true change points within the bias range for each iteration. Similar to the corresponding MVN simulation, the combinations NBS.D.SB and OBS.D.SB also have the smallest scaled Hausdorff distance, indicating the detected change points are located closest to the true change points. They also have low FN and FP rates. The results for the Vuong test and the SB are very similar in combination with all the segmentation methods. Other combinations such as MOSUM.D present good results, but with slightly higher FP rates. WBS appears to be too conservative with inferior accuracy but smaller FP rates.

Overall, e.cp3o and ks.cp3o have inferior performance to all variations of VCCP with lower TP, higher TN and higher FP. e.cp3o and ks.cp3o show similar results, with ecp being more conservative, leading to lower TP as well as lower FP and FN rates. Both detected more than 4 change points on average. If we reduce the $Kmax$ to 4 (see Supplementary Materials for the results table), all methods in ecp detected fewer than 3 change points.

In Simulation 6, only the edge weights in the network are fluctuating between change points with the non-zero edges themselves remaining fixed over the entire time course. This

Table 4: The results from applying all variations of VCCP and the two best performing non-parametric methods in the **ecp** R package to the VAR data with 4 change points (Simulation 5) over 100 simulated data sequences. NBS, OBS, MOSUM, and WBS denote the adapted binary segmentation, old binary segmentation, moving sum, and the wild binary segmentation methods, respectively. TP, FN, FP, Cand.num, Num.det, and d_H denote the true positive rate, false negative rate, false positive rate, number of candidate change points, number of change points detected, scaled Hausdorff distance, respectively.

	Stationary Bootstrap				Vuong test				e.cp3o	ks.cp3o
	NBS	OBS	MOSUM	WBS	NBS	OBS	MOSUM	WBS	alpha=1 Kmax=8	Kmax=8
TP.61	0.99	0.99	1	0.84	0.99	0.92	1	0.84	0.41	0.57
TP.121	1	1	0.99	0.92	1	0.92	0.99	0.95	0.63	0.51
TP.181	0.98	0.98	0.95	0.94	0.98	0.9	0.95	0.9	0.62	0.5
TP.241	0.99	1	0.94	0.76	0.97	0.92	0.94	0.7	0.18	0.4
FN.61	0	0	0	0.01	0	0.04	0	0	0.42	0.5
FN.121	0	0	0	0	0	0.01	0	0	0.25	0.4
FN.181	0.01	0.01	0	0	0.01	0.01	0	0.03	0.3	0.43
FN.241	0.01	0	0.02	0.07	0.03	0.03	0.02	0.09	0.53	0.5
FP.0-60	0.01	0.01	0.04	0	0.03	0.03	0.12	0.01	0.55	0.7
FP.61-120	0.03	0.03	0.01	0.02	0.06	0.06	0.02	0.02	0.57	0.72
FP.121-180	0.03	0.04	0.03	0.01	0.05	0.05	0.03	0.01	0.82	0.73
FP.181-240	0	0	0.04	0.01	0.05	0.04	0.05	0.02	0.31	0.76
FP.241-300	0	0	0.02	0.01	0.02	0.01	0.04	0.01	0.24	0.47
Cand.num	4.94	4.76	4.37	3.7	4.94	4.49	4.37	3.7	8.9	10.64
Num.det	4.02	4.04	4.03	3.49	4.14	3.84	4.14	3.46	4.37	5.38
d_H	0.09	0.09	0.13	0.51	0.18	0.44	0.16	0.54	1.18	0.81

is a very difficult simulation but NBS in combination with both SB and Vuong perform well, identifying almost all 4 change points (Table 5). Overall, NBS.D.V is the best model, having the smallest scaled Hausdorff distance and the highest TP rate. Ironically, OBS, which had many advantages over other segmentation methods in the previous simulations, has the worst performance, identifying the least number of true change points. Since OBS carries out segmentation and inference sequentially, we attribute its deterioration in performance to this. Once the changes become weak enough, it is likely that the SB and the Vuong test reject the first candidate in OBS, which results in it failing to detect the remaining change points. As small changes are common in resting-state fMRI data, it might be inappropriate to apply OBS to this data type. After NBS.D.V, NBS.D.SB and MOSUM.D.V, WBS has the next best performance in combination with the SB. This may stem from the sub-sampling step in WBS. As stated in Section 2, WBS firstly creates a series of multiple candidates and with their corresponding BIC reductions within each sub-sample. Then an OBS-like inference procedure decides the test order. Compared to the BIC reduction of the original OBS method, the sub-sampling idea in WBS overcomes its drawbacks. In fact, WBS is more similar to the combination of NBS and OBS, which contributes to its unique strength.

In terms of the non-parametric methods, e.cp3o and ks.cp3o_delta perform best. However, they are all inferior to all variations of VCCP. In fact, it becomes hard to pick the "best" nonparametric methods for this simulation, given the variation in TP and FP. When $Kmax = 4$, all but e.divisive have low TP and FP rates. When $Kmax$ increases to 8, the TP rate grows, but the FP rate of e.cp3o_delta soars to 1 in some intervals, and the same applies to other methods requiring $Kmax$ in their detection. This pattern indicates non-robustness of the non-parametric methods. On the contrary, methods depending on the E-statistics are less sensitive with respect to the α value.

We also compared the results in Table 4 (VAR data with 4 change points) with the results in Table 5 (VAR data with 4 weak change points), and find that the results for the latter are inferior due to the less distinctive changing patterns. Similar to the MVN data, the performance degraded from the former data set to the latter, but to a greater extent. The FP rate and the scaled Hausdorff distance increased. Similar to the results in the MVN

data with 4 weak change points, the first change point located at time point $t = 61$ is the most difficult to detect, where the highest TP rate was only 61% from NBS.D.V.

Table 5: The results from applying all variations of VCCP and the two best performing non-parametric methods in the **ecp** R package to the VAR data with 4 weak change points (Simulation 6) over 100 simulated data sequences. NBS, OBS, MOSUM, and WBS denote the adapted binary segmentation, old binary segmentation, moving sum, and the wild binary segmentation methods, respectively. TP, FN, FP, Cand.num, Num.det, and d_H denote the true positive rate, false negative rate, false positive rate, number of candidate change points, number of change points detected, scaled Hausdorff distance, respectively. ks.cp30_d is short for ks.cp3o_delta.

	Stationary Bootstrap				Vuong test				e.cp3o	ks.cp3o_d
	NBS	OBS	MOSUM	WBS	NBS	OBS	MOSUM	WBS	alpha=1 Kmax=8	Kmax=8
TP.61	0.56	0.43	0.35	0.5	0.61	0.19	0.42	0.3	0.7	0.82
TP.121	0.75	0.57	0.67	0.77	0.78	0.35	0.72	0.63	0.62	0.52
TP.181	0.82	0.68	0.73	0.8	0.83	0.34	0.7	0.64	0.08	0.15
TP.241	0.86	0.72	0.89	0.81	0.81	0.4	0.87	0.75	0.38	0.34
FN.61	0.23	0.09	0.35	0.11	0.18	0.21	0.28	0.11	0.18	0.07
FN.121	0.08	0.07	0.11	0.05	0.05	0.05	0.06	0.11	0.31	0.32
FN.181	0.1	0.04	0.06	0.02	0.09	0.03	0.09	0.14	0.44	0.52
FN.241	0.02	0.05	0.01	0.04	0.07	0.18	0.03	0.06	0.51	0.57
FP.0-60	0.07	0.05	0.04	0.04	0.08	0.02	0.03	0.03	0.88	0.82
FP.61-120	0.15	0.14	0.1	0.08	0.25	0.1	0.08	0.08	0.9	0.94
FP.121-180	0.09	0.12	0.16	0.17	0.19	0.1	0.23	0.14	0.41	0.43
FP.181-240	0.1	0.04	0.05	0.04	0.12	0.02	0.05	0.01	0.19	0.28
FP.241-300	0.12	0.09	0.09	0.02	0.15	0.02	0.1	0.03	0.55	0.66
Cand.num	5.3	2.78	4.34	3.17	5.3	2.47	4.34	3.18	8.69	8.6
Num.det	3.44	2.78	3.04	3.17	3.72	1.53	3.18	2.58	4.94	5.03
d_H	0.69	1.31	0.99	0.7	0.64	2.42	0.87	1.26	1.31	0.88

Table 6 shows the results of the VAR model with 7 small change points (Simulation 7). Most of the preponderant methods for the MVN data with 7 small change points (in the Supplementary Materials) are also evident for the VAR data with 7 small change points, except for OBS.D.V. This combination has the largest scaled Hausdorff distance, with TP rates less than 20% for each change point. It detects only one change point on average and has high FN rates at for change points at time points 76, 451 and 526.

To explore the reason behind its unusual performance, we notice that in OBS, the inference result at the beginning of the procedure decides whether to conduct later segmentation. If the first candidate change point is at time point t, the Vuong test compares the goodness

Table 6: The results from applying all variations of VCCP and the two best performing non-parametric methods in the **ecp** R package to the VAR data with 7 small change points (Simulation 7) over 100 simulated data sequences. NBS, OBS, MOSUM, and WBS denote the adapted binary segmentation, old binary segmentation, moving sum, and the wild binary segmentation methods, respectively. TP, FN, FP, Cand.num, Num.det, and d_H denote the true positive rate, false negative rate, false positive rate, number of candidate change points, number of change points detected, scaled Hausdorff distance, respectively.

	Stationary Bootstrap				Vuong test				e.cp3o_delta	
	NBS	OBS	MOSUM	WBS	NBS	OBS	MOSUM	WBS	alpha=1 Kmax=14	alpha=2 Kmax=14
TP_76	1	1	0.69	0.57	0.97	0.16	0.97	0.66	0.45	0.4
TP_151	0.99	0.99	0.62	0.97	0.98	0.14	0.97	0.96	0.22	0.17
TP_226	0.97	0.95	0.53	0.6	0.93	0.11	0.84	0.4	0.18	0.13
TP_301	0.89	0.89	0.63	0.69	0.89	0.12	0.82	0.58	0.26	0.32
TP_376	0.97	0.97	0.75	0.69	0.9	0.13	0.71	0.69	0.4	0.44
TP_451	0.98	0.98	0.89	0.93	1	0.19	0.94	0.87	0.28	0.29
TP_526	0.93	0.91	0.66	0.25	0.88	0.16	0.6	0.19	0.47	0.4
FN_76	0	0	0	0	0	0.29	0	0.02	0.1	0.17
FN_151	0	0	0	0	0	0.02	0.01	0	0.6	0.5
FN_226	0	0.02	0.01	0.1	0.04	0.02	0.07	0.23	0.55	0.52
FN_301	0.03	0	0.05	0.07	0.03	0	0.06	0.07	0.62	0.62
FN_376	0	0	0.01	0	0.06	0.04	0.12	0.02	0.49	0.33
FN_451	0	0	0	0	0	0.3	0	0.02	0.52	0.52
FN_526	0.02	0	0.03	0.18	0.05	0.2	0.16	0.21	0.19	0.26
FP_0-75	0.03	0.02	0.24	0.04	0.08	0	0.07	0.02	0.32	0.25
FP_76-150	0.07	0.07	0	0.01	0.06	0	0	0.03	0.69	0.65
FP_151-225	0.06	0.04	0.08	0.03	0.11	0	0.09	0	0.84	0.82
FP_226-300	0.08	0.1	0.02	0.01	0.08	0	0.04	0.03	0.84	0.81
FP_301-375	0.11	0.13	0.21	0.12	0.14	0.02	0.16	0.14	0.89	0.91
FP_376-450	0.05	0.06	0.07	0.02	0.08	0	0.03	0.01	0.82	0.87
FP_451-525	0.09	0.11	0.07	0.01	0.08	0.03	0.1	0	0.81	0.83
FP_526-600	0.11	0.13	0.21	0.12	0.14	0.02	0.16	0.14	0.89	0.91
Cand.num	9.79	8.24	6.17	5.5	9.6	2.07	8.31	5.31	19.04	18.44
Num.det	7.27	7.29	5.67	4.94	7.24	1.07	6.44	4.53	7.93	7.83
d_H	0.23	0.23	0.9	0.97	0.33	5.81	0.57	1.11	0.53	0.60

of fit between VC_0 and VC_l using the left part of data $(1 : t-1)$, and VC_0 and VC_r using the right part of data $(t : T)$. Only when both tests show VC_0 is inferior do we treat t as a change point. However, if one candidate is close to the beginning (end) of the timeline and the right (left) part of data contains one or more latent change points, segmentation at time point t might not improve the goodness-of-fit of the right (left) part of data. That is to say, neither VC_r (VC_l) nor VC_0 can perfectly describe the distribution of the right (left) part of data with the latent change points undetected. In some cases, the Vuong test even provides a reverse result that VC_0 outperforms the other VC model due to a larger sample size. This happens more frequently when the changes in the structure are less discernible. Therefore, the candidate change point at time point t is not considered a change point, and other latent change points remain undiscovered because the segmentation terminates, as shown in the left panel of Figure 5. Here, for one iteration of Simulation 7, OBS.D.V stops at the second candidate change point ($t = 76$), even though the left part of the test (the dark blue dot) has identified it as a change point. On the contrary, NBS.D.V's performance is superior as it performs the Vuong test on all candidate change points after its exhaustive search.

With respect to the non-parametric methods, we observe similar findings. Apart from OBS combined with Vuong, all variations of VCCP outperform them. In particular, for this data set, the performance of e.cp3o is poor. E.cp3o_delta has the best performance for all α values, yet it still suffers from high FP and FN rates. Additionally, the predominant effect of *Kmax* in Num.det still exists, deteriorating the model's robustness. VCCP, on the contrary, does not have this issue since the number of candidate change points is completely data-driven without any prior knowledge.

4.2 Harry Potter task-based fMRI data

In order to show that other forms of dependence exist beyond linear dependence for fMRI data, we compare the detected change points found by the NBS.D.V combination of VCCP using only the Gauss copula to NBS.D.V using 5 types of copulas (Gauss, t, Clayton, Gumbel, and Frank copula). Figure 6 shows the detected change points.

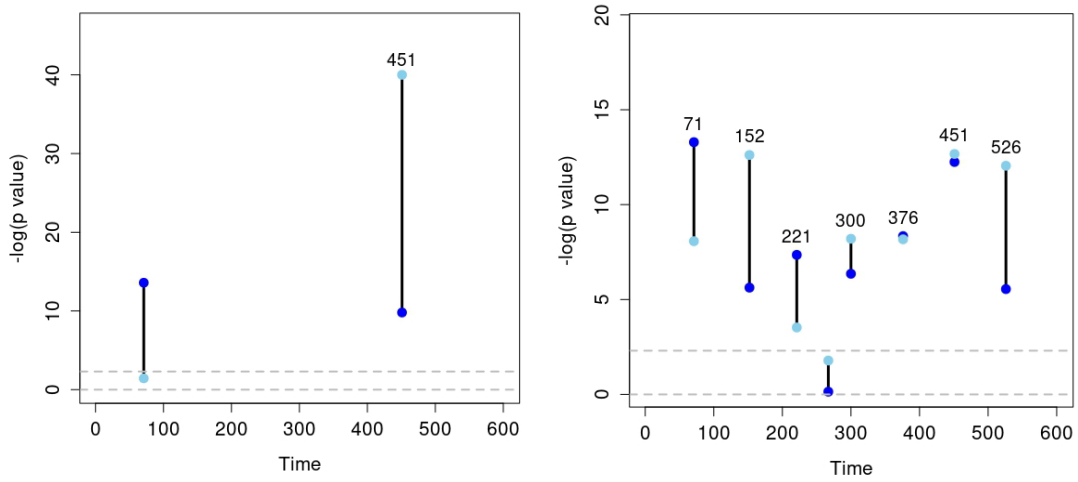

Figure 5: A comparison of the OBS.D.V and NBS.D.V combination methods when applied to one iteration of the VAR data with 7 small change points simulation. The y-axis represents the $-\log(p\text{-value})$ from the Vuong test. The horizontal dashed line represents the significance level, $\alpha = -0.1{*}\log(p\text{-value}) = 2.303$. The dark blue and light blue dots represent the test using the left part and the right part of the data, respectively.

As expected, by considering various forms of dependence and copulas beyond Gaussian, more change points are detected (1/3 more detected). For example, let's consider the detected change points for subject 7. Using NBS.D.V with the Gauss copula, time points $t = 43, 235, 282, 372, 436, 704, 1003, 1099, 1149, 1247$ are identified as change points, while time points $t = 44, 241, 282, 390, 704, 746, 792, 867, 998, 1038, 1096, 1168, 1247$, are identified as change points for NBS.D.V with the 5-copula family. Now let's focus our attention on change points at $t = 998, 1038$ and $t = 1096$. The first and last change points are detected by both methods while the middle change point is missed by NBS.D.V with the Gauss copula. Figure 7 displays the estimated Kendall's τ, the upper tail dependence, the lower tail dependence, and the optimal copula function between pairs of ROIs for data between change points. From segment $867 - 997$ to segment $998 - 1037$, the Kendall's τ network becomes more sparse, while the lower tail network (especially within the IFG sub-network) and the upper tail network become more dense. As a result, the dominant copula type turns out to be the Clayton copula during the second segment $(998 - 1037)$. As the Gauss copula model is unable to model tail and other forms of dependence, the change point at

$t = 998$ is detected only because of the change in Kendall's τ. At time point $t = 1038$, there is not an obvious strong discernible change in Kendall's τ, hence, the Gauss copula fails to find the related change point. Only models free of the Gaussian assumption are capable of discovering this change point. NBS.D.V with the 5-copula family has this flexibility. From segment $1038 - 1095$ to segment $1096 - 1167$, negative edges between nodes IT, IFG1, IFG3 and MT switch to positive, while a new sub-network between F, SM, PCG and IT appears with most of its edges being negative in the network of Kendall's τ. The density of edges also alters. In addition, the lower tail dependence weakens so that only edges between TP and other nodes (TP becomes a hub node) remain, which results in the Clayton copula not being the dominant copula type. Furthermore, there is little change in the upper tail dependence network. Hence, both NBS.D.V with the Gauss copula and the 5-copula family are able to detect the change point at $t = 1096$. To understand the rationale behind the change point at $t = 1038$, the last sentence given to readers before this time is, 'They ripped through a tapestry and found themselves in a hidden passageway, ... which they knew was miles from the trophy room.', and the first sentence after $t = 1038$ is ' "I think we've lost him," Harry panted, leaning against the cold wall and wiping his forehead.' This passage occurs when Harry and the other three characters run away from Filch. The text during this period includes changes in locations, actions and emotions.

For a complete understanding of the changing brain networks when reading, further exploration of the simultaneous changes in the text becomes necessary. Taking advantage of the rich text information available, we can assign attributes to every word in the text. In particular, we focus on whether the word is a character's name (Char), implies a kind of emotion (Emo), refers to a motion (Mo), or contains a specific verb (Verb). Subdivisions of the four main word types, with their frequency among the 5200 words in the text, are shown in Table 7. As the 'Motion' attribute takes place over the course of a sentence, Wehbe et al. (2014) created two features: a punctual feature and a "sticky" feature. The punctual feature coincides when the verb of the motion is mentioned, and the sticky feature is active for the duration of the motion (i.e., the sentence). The same applies to the 'Emotion' attribute, where the punctual features coincide when the emotion is explicitly mentioned, and sticky

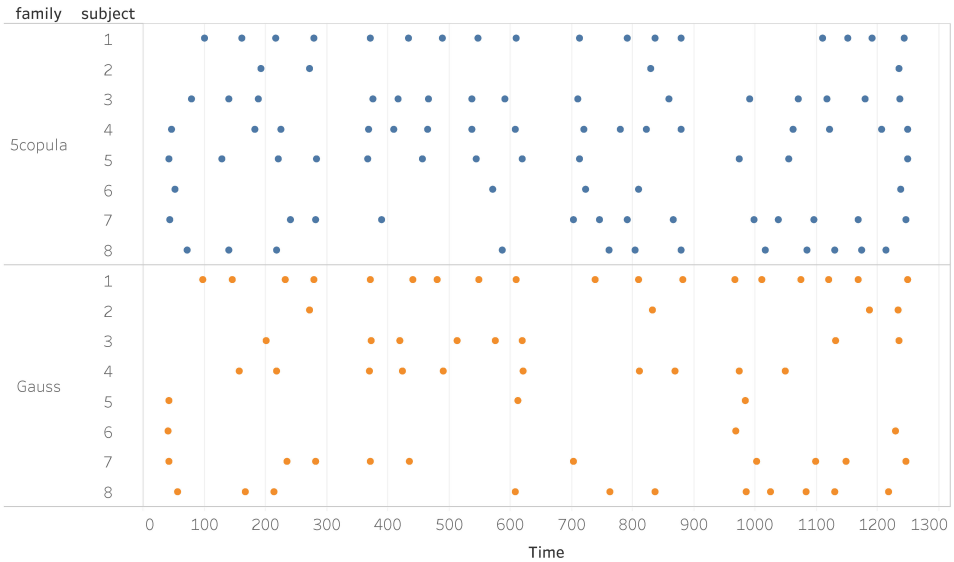

Figure 6: The top and bottom panels represent the change points from NBS.D.V using the 5-copula family and only the Gauss copula, respectively, from the Harry Potter task-based fMRI study.

features align is active for the duration of the emotion felt. Here, we concentrate on the sticky features and changes in them. First, we specify an interval parameter (Int) indicating the range in which text changes take place, for example, $[t - Int, t + Int - 1]$. The range covers two whole sentences to display the complete textual background of the change. Since the upper quantile of the sentence length equals 17 words and 4 words are displayed within 1 TR (2 sec), Int is equal to 4 TR. We define three types of changes in text features: 'off-on', 'on-off', and 'mixed' patterns. If the identifier of a feature remains constant on one side of a point t (e.g., $[t - Int, t - 1]$), but contains a heterogeneous element on the other side (e.g., $[t, t + Int - 1]$), we then call it a 'off-on' ($[0 \rightarrow 0{,}1]$ or $[0{,}1 \rightarrow 1]$) or 'on-off' ($[1 \rightarrow 1{,}0]$ or $[1{,}0 \rightarrow 0]$) change type. If both sides include heterogeneous elements (i.e., $[0{,}1 \rightarrow 0{,}1]$), we call it a mixed type.

To understand the difficulty of our analysis; coinciding changes in brain networks and the variation of a specific text feature, we explore the latent text changes, and in particular we consider the features before and after a particular time point in the text, time point $t = 432$. Figure 8 shows the text changes: three motion features (collide, fly, manipulate) experience 'on-off' changes; one verb (move), one emotion (nervous), and one character

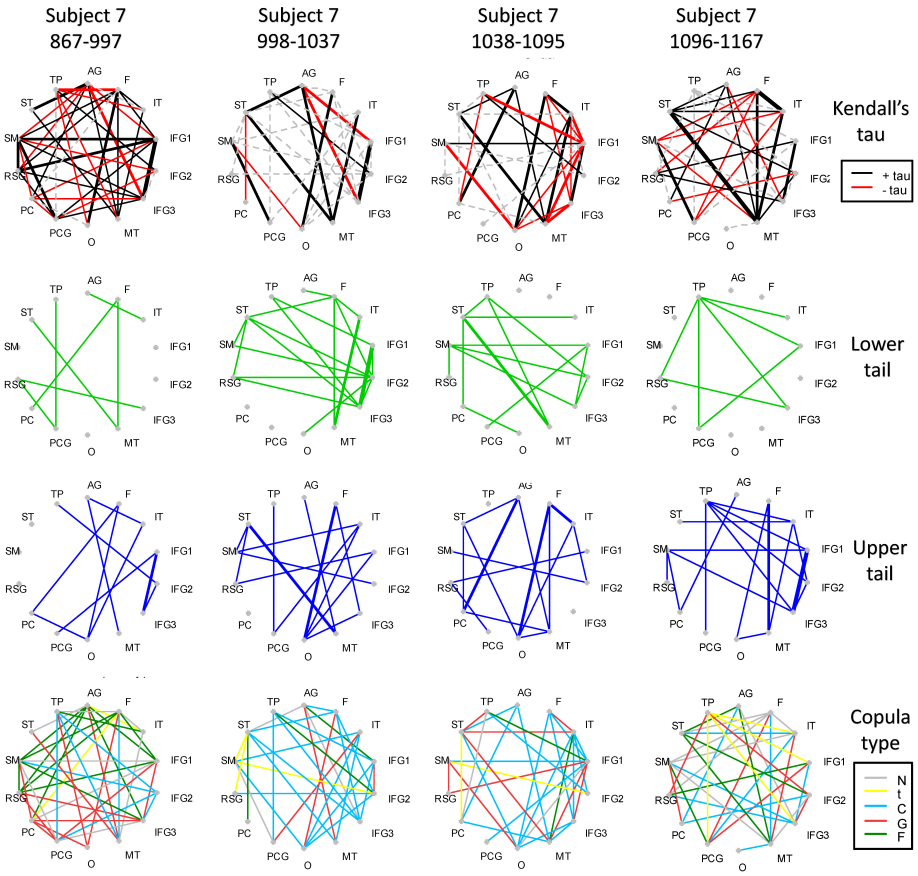

Figure 7: The estimated Kendall's τ, the lower and upper tail dependence, and the optimal copula function, for each edge between pairs of nodes for networks between change points at time points $t = 998, 1003, 1099$ for subject 7 in the Harry Potter task-based fMRI data set using the combination NBS.D.V of VCCP. Black (red) lines in the graph on the first row represent positive (negative) Kendall's τ correlation coefficients. Dashed lines indicate edges between nodes with heavy tail dependence but insignificant Kendall's τ. Green and blue edges in the graphs on the second and the third rows represent the lower and upper tail dependence, respectively. The various colored lines on the bottom row indicate the optimal copula family for edges, with grey, yellow, blue, red and green indicating the Gauss, t, Clayton, Gumbel and Frank copula, respectively.

Table 7: The frequency of words for the following four attributes: a character's name (Char), emotion (Emo), motion (Mo), or a specific verb (Verb) in Chapter 9 of *Harry Potter and the Sorcerer's Stone*.

Motion	move 47.67%				manipulate 19.27%			
	fly 9.78%				collidePhys 3.96%			
Verbs	know 1.00%				see 0.89%			
	hear 0.66%			tell 0.60%			be 0.58%	
Characters	harry 4.19%		draco 1.47%		herm 1.35%		neville 1.18%	
	ron 1.14%		minerva 0.79%		filch 0.54%	peeves 0.48%	wood 0.41%	hooch 0.39%
Emotion	nervous 15.85%		fear 15.03%	annoyed 13.49%	command 11.88%	puzzled 9.89%	wonder 6.45%	dislike 5.72%
	like 4.60%	relief 3.98%	hurtPhys 3.13%	praising 1.76%	hurtMental 1.60%	cynical 0.97%	pride 0.58%	pleading 0.35%

(Minerva) encounter 'off-on' patterns; and a character (Harry) feature has a mixed variation. Hence, for the FC change point at this time point, the corresponding text may involve changes in multiple features, which increases the difficulty of assigning the change points detected to one specific text feature.

TR	word	collide	fly	mani	move	nervous	harry	minerva
427	broom	1	1	1	0	0	0	0
	straight,	1	1	1	0	0	0	0
	and	1	1	1	0	0	0	0
	he	1	1	1	0	0	1	0
428	toppled	1	1	1	0	0	0	0
	gently	1	1	1	0	0	0	0
	onto	1	1	1	0	0	0	0
	the	1	1	1	0	0	0	0
429	grass	1	1	1	0	0	0	0
	with	1	1	1	0	0	0	0
	the	1	1	1	0	0	0	0
	Remembrall	1	1	1	0	0	0	0
430	clutched	1	1	1	0	0	0	0
	safely	1	1	1	0	0	0	0
	in	1	1	1	0	0	0	0
	his	1	1	1	0	0	1	0
431	fist.	1	1	1	0	0	0	0
	+	0	0	0	0	0	0	0
	"HARRY	0	0	0	0	0	1	0
	POTTER!"	0	0	0	0	0	1	0
432	+	0	0	0	0	0	0	0
	His	0	0	0	0	1	1	0
	heart	0	0	0	0	1	0	0
	sank	0	0	0	0	1	0	0
433	faster	0	0	0	0	1	0	0
	than	0	0	0	0	1	0	0
	he"d	0	0	0	0	1	1	0
	just	0	0	0	0	1	0	0
434	dived.	0	0	0	0	1	0	0
	Professor	0	0	0	1	0	0	0
	McGonagall	0	0	0	1	0	0	1
	was	0	0	0	1	0	0	0
435	running	0	0	0	1	0	0	0
	toward	0	0	0	1	0	0	0
	them.	0	0	0	1	0	0	0
	…	…	…	…	…	…	…	…

Interval = 4

Interval = 4

t = 432

change type	off-on	on-off	mixed
collide		✓	
fly		✓	
mani		✓	
move	✓		
nervous	✓		
harry			✓
minerva	✓		

Figure 8: An example of three types of text changes at time point $t = 432$ in Chapter 9 of *Harry Potter and the Sorcerer's Stone*, with the interval parameter fixed at 4 TRs.

In Figure 9 (top panel), we sum the frequency of attributes (Char, Emo, Mo, Verb) that correspond to change points detected by NBS.D.V with the 5-copula family for each subject. For example, the 94% in the first cell implies that amongst all the change points for subject 1, 94% are located at time points that coincide with character switching. Overall, there is a great deal of subject-level heterogeneity. For example, despite there being frequent character switches in the text, the dominant attributes at FC change points for subject 5 is motion. Further, Subjects 3, 4 and 8 react actively to emotion and verb attributes. Considering that pre-specified verbs occur infrequently in Table 7, this large number of

change points (e.g., 13 out of 16 change points for subject 4) strongly implies that some subjects are very sensitive to changing verbs. We also display the locations of the detected change points with the attribute underpinning them Figure 9 (bottom panel). These plots further illustrate the complexity of textual changes. In summary, more than 90% of the change points are related to at least two types of attributes simultaneously. For these attributes, the emergence, rather than the disappearance more frequently leads to a change point, as red segments are more evident than yellow ones.

	1	2	3	4	5	6	7	8	Sum
Char	16 94.12%	3 60.00%	14 93.33%	15 93.75%	10 83.33%	5 100.00%	13 100.00%	9 75.00%	80 89.89%
Emo	10 58.82%	3 60.00%	11 73.33%	12 75.00%	9 75.00%	5 100.00%	10 76.92%	8 66.67%	64 71.91%
Mo	15 88.24%	4 80.00%	13 86.67%	14 87.50%	11 91.67%	4 80.00%	11 84.62%	8 66.67%	76 85.39%
Verb	9 52.94%	2 40.00%	10 66.67%	13 81.25%	7 58.33%	3 60.00%	7 53.85%	8 66.67%	55 61.80%
Sum	17 100.00%	5 100.00%	15 100.00%	16 100.00%	12 100.00%	5 100.00%	13 100.00%	12 100.00%	89 100.00%

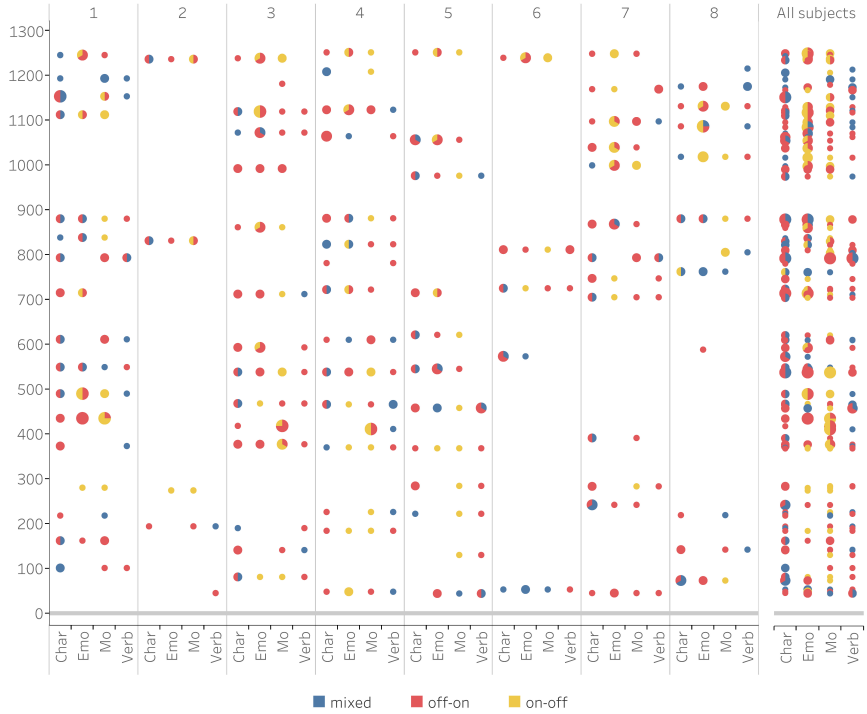

Figure 9: The top panel is the summation of the frequency of attributes (Char, Emo, Mo, Verb) that correspond to change points detected by NBS.D.V with the 5-copula family for each subject. The bottom panel displays the locations of change points corresponding to textual changes for each subject (left) and all subjects (right). Colors denote the three types of patterns.

In Figure 10, we plot the percentage of changing subdivisions (off-on) in their corre-

sponding attributes at the change points detected by NBS.D.V for each subject. For each attribute, the first row (Freq) records the relative frequency of subdivisions in the original text in decreasing order, which we then compare with the frequency of change points for the same subdivision. For example, from the top left graph (Character: off-on), we can conclude that the appearance of characters such as 'Hermione', 'Neville', 'Minerva', and 'Peeves' in the text closely relate to the change points across some of the subjects (match the block size in Freq to the blocks in each of the subjects). The appearance of 'Harry' is not as consistent across subjects. The composition of the emotion graph is even more complex, where subjects are generally susceptible to different emotional changes of characters. Among the fifteen different emotions, 'puzzlement' is the most likely to be associated with changes in the brain networks across the 8 subjects. Even though positive feelings such as 'like', 'relief' and 'pride' seldom appear in the text, subjects 4, 5 and 7 are still sensitive to them. On the contrary, emotions such as 'nervousness' and 'fear' that cover a large portion of the text only stimulate a subset of the subjects (1, 3, 5, 7 and 8), while other subjects seem unaffected by the characters evoking such emotions. The reaction to the 'annoyance' of characters was the most inconsistent. Subjects 2, 3 and 6 experienced more frequent change points in terms of emotional transition to annoyance, while subjects 1, 4 and 7 share a moderate proportion of change points related to annoyance. Subjects 5 and 8, however, did not exhibit any reaction when confronting the presence of annoyance in the text. Such inconsistency could be partly due to the insufficient number of change points detected for certain subjects (e.g., subjects 2 and 6), and perhaps partly to the real heterogeneity of dynamic brain networks towards a certain emotion experienced by characters. The proportion of motion subdivisions in the text is consistent with that of subdivision change points for subjects 3, 4 and 7. The other five subjects did not react to the beginning of the characters flying nor to the colliding in the story. Further research could be carried out by dividing the 'manipulate' and 'move' motions into more detailed subdivisions. The heterogeneity across subjects becomes more evident when it comes to the variation in verbs: subjects 1 and 4 are very sensitive to the verb 'know'; the FC of subject 8 changes intensively when the characters 'hear' a sound; subjects 5, 7 and 8 experience switching patterns in their brain network at the time of

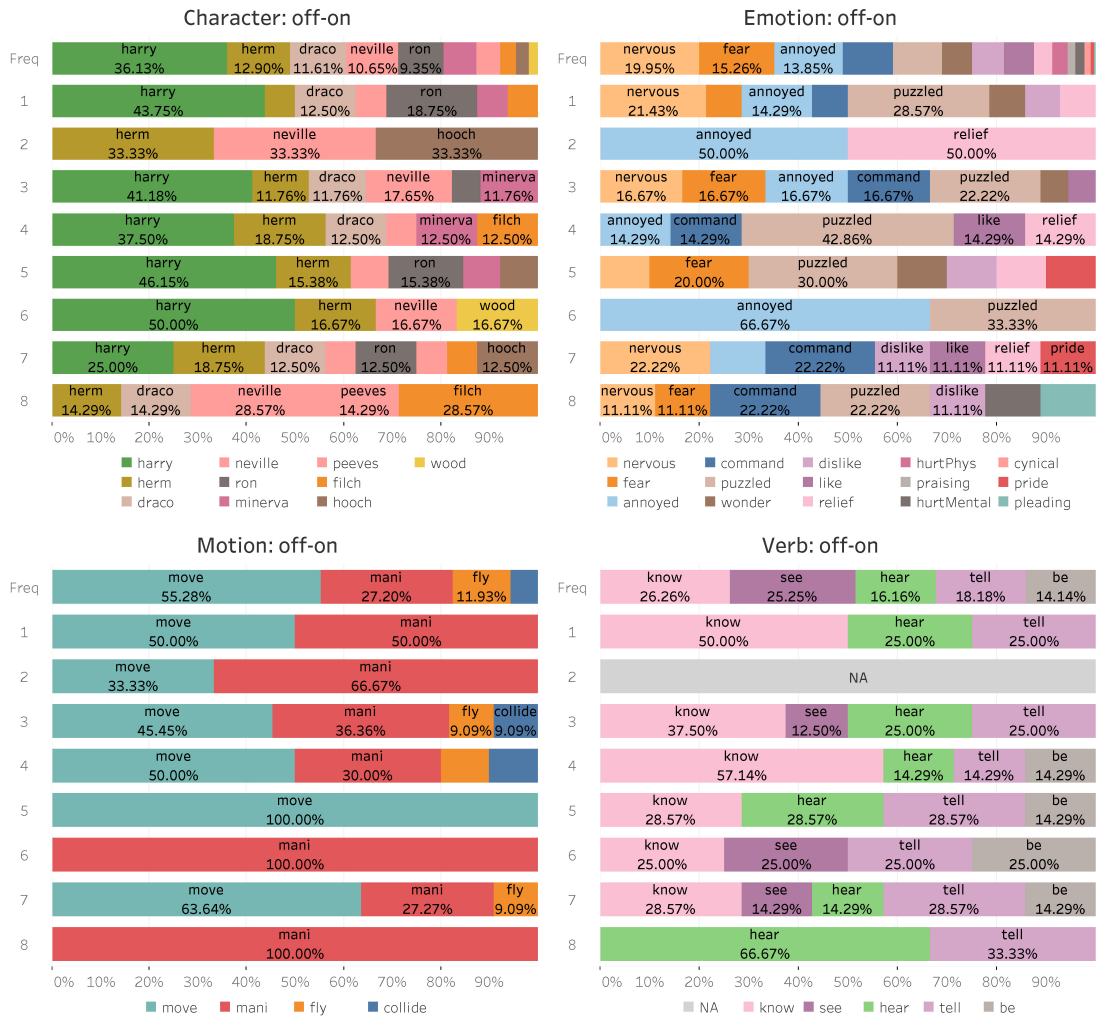

Figure 10: The percentage of changing subdivisions (off-on) in their corresponding attributes at the change points detected by the NBS.D.V combination of VCCP for each subject. For each attribute, the first row (Freq) records the relative frequency of subdivisions in the original text in order (of time).

speaking; subject 2 fails to give any reaction to the appearance of the five types of verbs. Interestingly, brain networks across all 8 subjects remain in a relatively stable status when characters 'tell' something unexpected in Chapter 9. In other words, other than visual description, auditory stimulation to characters can generate more change points in the brain networks for readers.

If we turn to the 'on-off' subdivision pattern in Figure 11, all 8 subjects remain in a brain network longer than the subdivision when confronting the disappearance of a certain verb or the transfer of focus from a certain character. For all attributes, char, Verb, Emo, the 'on-off' subdivision pattern are less likely to coincide with change points, however, for attribute Mo, there is correspondence between the subdivision pattern and the detected change points. Yet, for the attribute Emo, the abnormally higher portion of the 'hurtPhys' element in subjects 1 and 8, the 'like' element in subjects 4 and 5, the 'cynical' element in subject 7 still deserves exploration. For all textual features, the functional network of the subjects' brain varies in accordance with the integrated variation of the text, instead of a single, specific attribute. Considering the complex component of the reading material and the collaborative work of our brain, it is hard to directly relate a connection variation between two nodes to one special type of textual change.

Next, we focus on general change points as a response to plot changes. By selecting a representative set of change points shared by multiple subjects and the contexts read before and after the change, we display the heterogeneity and homogeneity among subjects when confronting the same combined textual variation.

The first flying experience:

> Blood was pounding in his ears. He mounted the broom and kicked hard against the ground and up, up he soared; air rushed through his hair, and his robes whipped out behind him – and in a rush of fierce joy he realized he'd found something he could do without being taught – this was easy, this was wonderful. $|^{s5}$ He pulled his broomstick up a little to make it even higher $|^{s4}$ and heard screams and gasps of girls back on the ground $|^{s1}$ and an admiring whoop from Ron. $|^{s3}$ He turned his broomstick sharply to face Malfoy in midair ...

We detect change points in subjects 1, 3, 4 and 5 that coincide with Harry's first flying

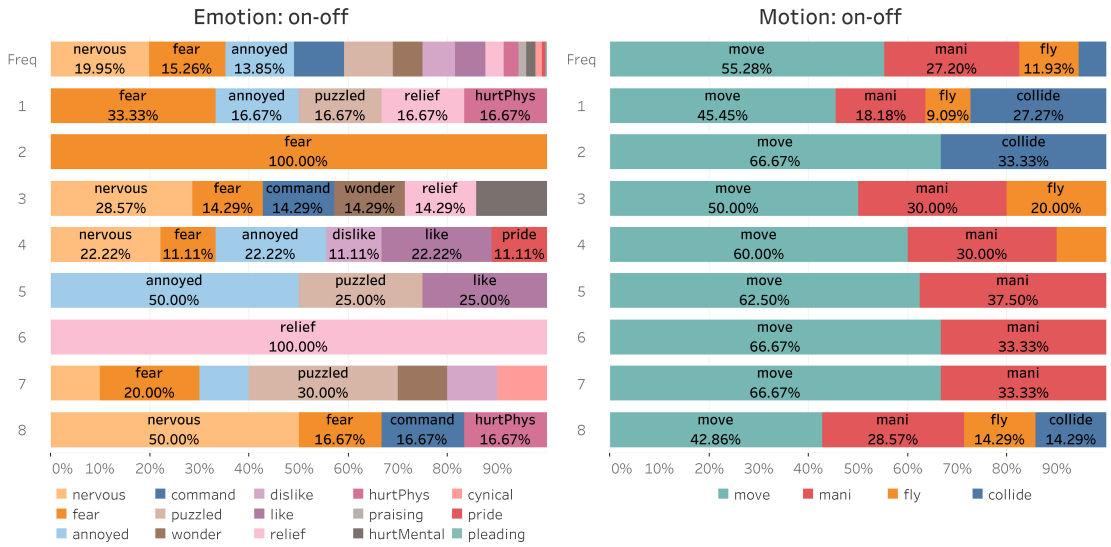

Figure 11: The percentage of changing subdivisions (on-off) in their corresponding textual types at change points for each subject. For each attribute, the first row (Freq) records the relative frequency of subdivisions in the original text in order (of time).

experience. The exact locations of the change points are marked as vertical lines in the quotation above, and the corresponding FC networks before and after the change points are provided in Figure 12 (we only plot the positive correlation between nodes). Several positive emotions such as 'joy' and 'admiring' emerged during this extract, which are rare emotions in Chapter 9. As a reaction to such a 'positive' textual change, the Kendall's τ network of all four subjects become more sparse (Figure 12, top left). For subject 1, the close relationship between ROIs IFG (IFG1, IFG2, IFG3), F and RSG disappears; for subject 3, most of the edges disappear while the connections within the sub-network (ST-AG-MT) become stronger. In subject 4's network, there is disconnection between ST, TP, MT and IFG, with only seven edges remaining. The middle temporal gyrus is sensitive to visual motion (flying), and while traditional language processing areas include the inferior frontal gyrus (Broca's area), superior temporal and middle temporal gyri, supramarginal gyrus and angular gyrus (Wernicke's area), there is evidence that structures in the medial temporal lobe have a role in language processing (Tracy and Boswell, 2008). The FC networks of subject 5 before and after the change point are more closely related, compared to the other three subjects, with decreasing Kendall's τ before and after the change point.

Contrary to the decreasing Kendall's τ, the lower tail networks (Figure 12, bottom left) increase in degree edges while the upper tail dependence networks (bottom right) vary across subjects. Based on the increasing number of grey dashed lines in the networks after the change point, we can ascertain that although some edges disappear, some of the tail dependence edges remain or appear. For example, for subjects 1 and 3, Kendall's τ between ROIs TP and IFG3 disappears but the corresponding lower tail dependence appears or stabilizes, respectively. The temporal pole has been associated with several high-level cognitive processes: visual processing for complex objects and face recognition, naming and word-object labelling, semantic processing in all modalities, and socio-emotional processing (Herlin et al., 2021). Accordingly, changes in the connection (strength) of the tail dependence has less to do with the variation in Kendall's τ. Rich colors in the top right sub-graphs for each subject represent the diversity in the best fitted copula types for edges between each pair of nodes. As Harry is experiencing the joy of his first flying experience, the dominant copulas shift to the Clayton and Gumbel copula after the change point, which is due to the fact that Clayton and Gumbel copulas provide a better fit to the lower and upper tail dependence than the other three available copulas. This is best illustrated by subject 5, where both types of tail dependence networks become denser and the preponderant F copula loses its dominance after time point $t = 367$.

Encountering with the three-headed dog:

> "What?" Harry turned around – and saw, quite clearly, what. For a moment, he was sure he'd walked into a nightmare – this was too much, on top of everything that had happened so far. $|^{s7}$ They weren't in a room, as he had supposed. They were in a corridor. The forbidden corridor on the third floor. $|^{s8}$ And now they knew why it was forbidden. They were looking straight into the eyes of a monstrous dog, $|^{s3}$ a dog that filled the whole space between ceiling and floor. It had three heads. Three pairs of rolling, mad eyes; three noses, twitching and quivering in their direction; three drooling mouths, saliva hanging in slippery ropes from yellowish fangs.

This event is the most thrilling in Chapter 9 of *Harry Potter and the Sorcerer's Stone*. The exact locations of the change points for subjects 3, 7 and 8 are marked as vertical lines in the quotation above, and the corresponding FC networks before and after the change

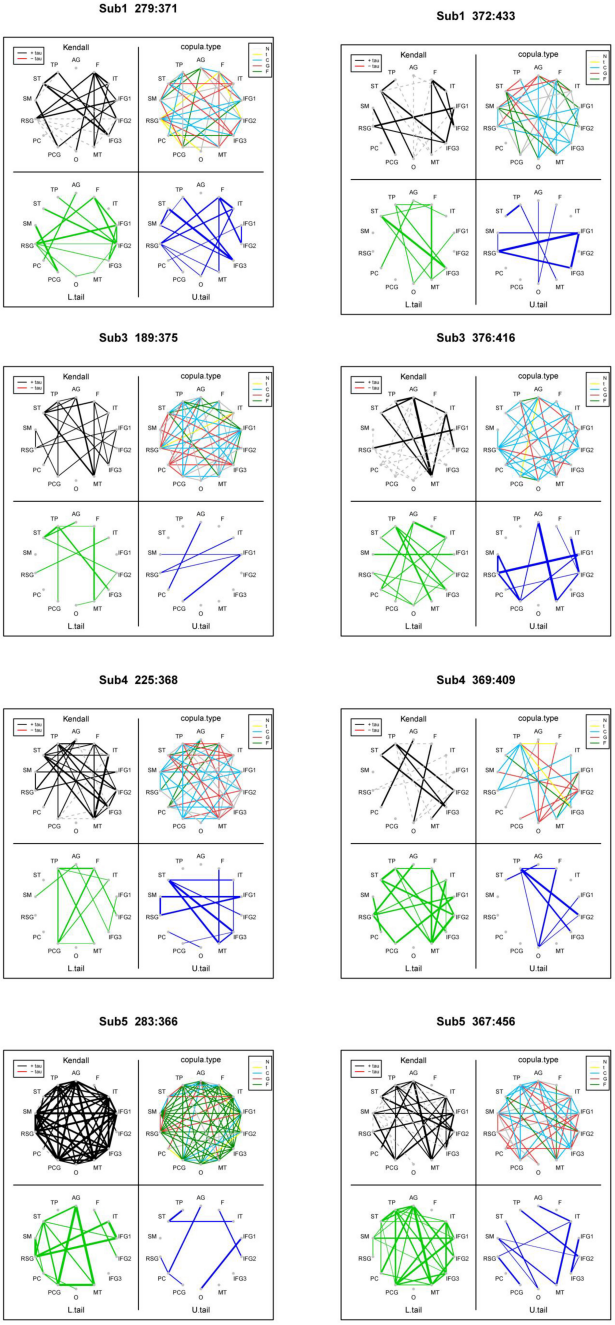

Figure 12: The FC networks estimated using the NBS.D.V combination of VCCP for subjects 1, 3, 4 and 5 at the change point that coincides with Harry's first flying experience. The left plots represent the networks before the change point while the right plots represent the networks after the change point. The top left, top right, bottom left and bottom right represent the Kendall's τ coefficient network, copula type network, lower tail dependence network, upper tail dependence network, respectively.

points are provided in Figure 13. Both consistency and heterogeneity exist in their brain networks. For subject 3, the most obvious change after the change point at time $t = 1179$ is the disappearance of the upper tail network between ROIs RSG, IT, IFG1, IFG2 and IFG3. In addition, strong connections between ROIs TP, AG, ST and F show up in Kendall's τ and the upper tail dependence network after the change point. For subject 7, the Kendall's τ network after the change point resembles that of subject 3, especially for sub-networks MT-ST-AG-F-IT and IFG1-SM-RSG-PCG. The first sub-network hub at MT is also present in the networks of subjects 3 and 7 before their change points, though subject 7's former network had two hub ROIs (MT and ST) and involved more related ROIs (e.g., AG, IFG3). The supramarginal gyrus is essential for visuospatial awareness and it may generate the fictive dream space necessary for the organized hallucinatory experience of dreaming (Pace-Schott and Picchioni, 2017). The sequence of events occurring in the book at this time are dreamlike with the description including the word "nightmare" and the characters moving from a room to a corridor without their own movement. The edge between ROIs SM and IFG1 also shares a reappearing pattern in subjects 7 and 8. Interestingly, subject 8 has a lower tail dependence network hub at ROI ST similar to the Kendall's τ network of subject 7 before the change point. However, the hub node pattern around ROI ST fades and shifts to a stronger Kendall's τ network between SM, RSG, PCG, IFG1 and IFG3 after the change point, which is unique during the post-change period among the three subjects. There are also between-edge, between-subject and between-period variations in the best fitted copula functions, which indicate the heterogeneity in the dependence type.

In general, when subjects share common change points, their dynamic reaction reflected in the FC networks embodies not only uniformity towards a certain combination of textual changes, but also individual heterogeneity. However, to further confirm the specific textual feature leading to the change point, more refined experiments controlling other irrelevant variables are required. Otherwise, we can only attribute the change in brain networks to combined textual factors.

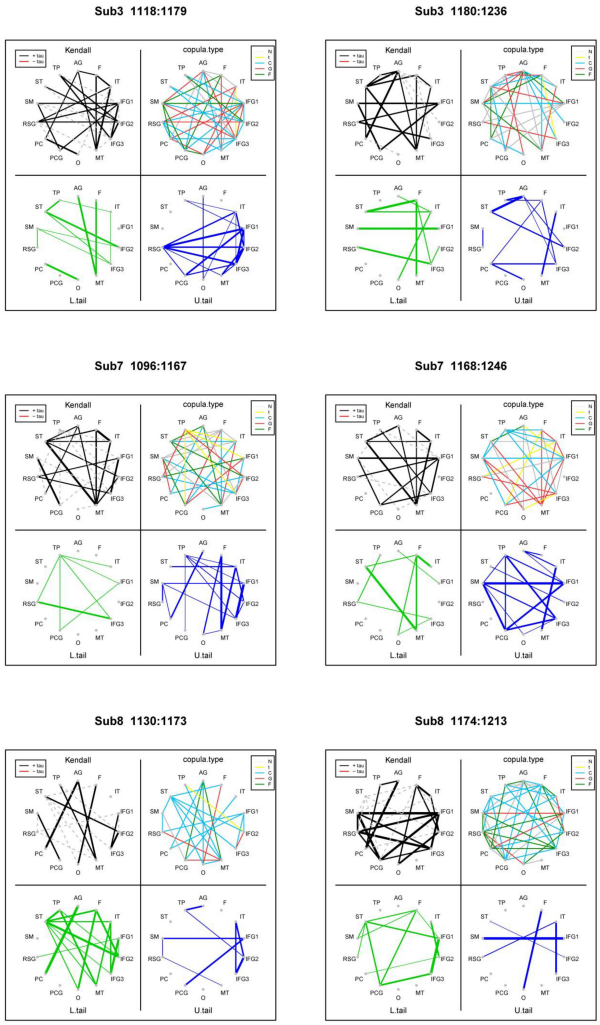

Figure 13: The FC networks estimated using the NBS.D.V combination of VCCP for subject 3, 7, and 8 at the change point that coincides with the four main characters encountering the three-head dog. The left plots represent the networks before the change point while the right plots represent the networks after the change point. The top left, top right, bottom left and bottom right represent the Kendall's τ coefficient network, copula type network, lower tail dependence network, upper tail dependence network, respectively.

4.3 SET fMRI data

The results for the SET data set can be found in the Supplementary Materials. To illustrate the importance of diverse copulae and possible tail dependence between the selected 5 ROIs, we compare the performance of the Gauss copula against the 5-copula family in our VCCP model. We also study the effect of the choice of δ.

5 Discussion

5.1 Computational speed

Table 8 shows a computational speed comparison across the various combinations of VCCP for the VAR simulation with 4 weak change points (Simulation 6) across 100 iterations. All were completed using a ASUS UX303UA laptop which has an Intel(R) Core(TM) i7-6500U single core processor. Obviously, with access to more cores and introducing parallel computing, VCCP speeds up significantly. Among the four segmentation methods, MOSUM is the most computationally efficient as it reduces the problem of a (possibly) high dimensional time series to a univariate time series problem. WBS is relatively slow as the computational efficiency depends heavily on the number of random draws, M, which should be large enough to guarantee the detection of the candidate change points. NBS and OBS have a similar moderate speed since both apply the idea of BS. With respect to inferential methods, SB depends on the number of resamples: as N increases, SB slows, but the result becomes more reliable. The Vuong test is computationally efficient since it has a known asymptotic distribution. For VCCP, the number of ROIs included has a large influence on its computational speed. As the number of ROIs increases, the MST methods (Cormen et al., 2009) have to consider more edges and hence more time is consumed finding the optimal vine. By specifying the vine structure (D-vine or C-vine), MST only needs to set the linking order or the central node, which can increase computational efficiency. Similarly, if we restrict the candidate family set of copulas (e.g., only the Gauss family) for each edge, we can decrease speed. Finally, as the experimental time course increases, the search space has a wider range, which leads to an increased computing time.

Method	Avg Time (s)	SE
NBS.D.SB	354.6324	47.91403
NBS.D.V	201.568	29.68706
OBS.D.SB	297.7229	94.55046
OBS.D.V	137.1511	49.44729
MOSUM.D.SB	186.8384	23.15093
MOSUM.D.V	47.4841	4.353137
WBS.D.SB	777.8286	141.6771
WBS.D.V	626.5993	90.56946

Table 8: A computational speed comparison across the various combinations of VCCP for the VAR simulation with 4 weak change points (Simulation 6) across 100 iterations. The Avg Time and SE represent the average time in seconds and standard error across 100 iterations, respectively. All are completed using simple sequential computing on a single core processor.

5.2 Extensions

To deal with a large number of ROIs, it is possible to combine VCCP with dimension reduction tools such as Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA) or Feature Annealed Independence Rule (Fan and Fan, 2008). In addition, the distribution of the value of BIC reduction at certain time points could also be explored under the vine copula specification. Here the classic hypothesis test could be applied instead of the SB, in order to improve the computational speed. Furthermore, similar to MOSUM, we could consider other methods that directly explore the local maximums of the BIC reduction time series, especially if we believe there are small and subtle dynamic FC patterns in the data.

VCCP can also be applied to multi-subject data (panel data). The only distinct difference between single-subject and multi-subject data is that the sample size in each interval increases as a result of stacking different subjects' data in the same period. Accordingly, more signal is available. For more details, see Xiong and Cribben (2021).

Applying VCCP to the Harry Potter fMRI data set, we discovered that including using more copula families in the building blocks of VCCP results in more change points detected. To further explore the effect of copula families on the performance of VCCP, we extend our choice to other copulas including Joe, Clayton-Gumbel (BB1), Joe-Gumbel (BB6),

Joe-Clayton (BB7) and the Joe-Frank (BB8) copula, and re applied the change point analysis with $\delta = 40$ on the Harry Potter fMRI data (Figure 14). In comparison to the single Gauss copula (see Figure 6), copulas with heavy tail dependence allow VCCP to possibly detect more change points. This was also evident when we add t, Clayton, Gumbel, Frank (the previous 4 copulas) to the family or add Joe, BB1, BB6, BB7 and BB8 (the new 5 copulas). When we combined all the 9 non-gauss copulas with the Gauss copula (last panel), the number of detected change points exceeds the other cases, although not as evident in moving from Gauss to 5 family copulae. That is, the number of change points detected by VCCP is not strictly linearly correlated with the number of copula families.

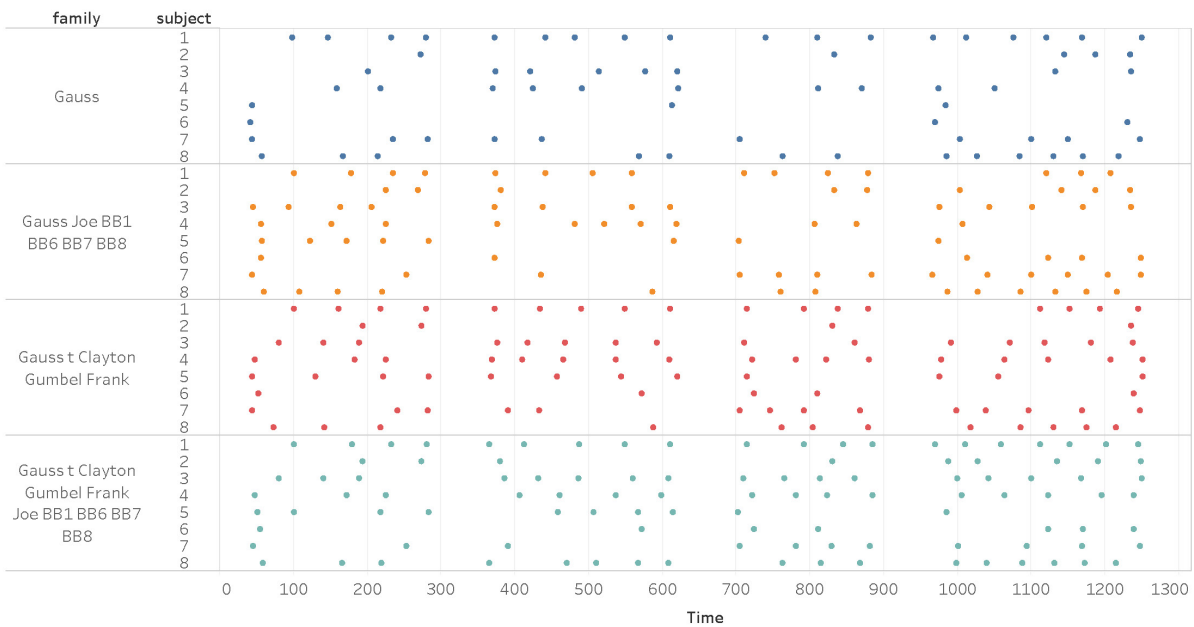

Figure 14: The four panels represent the change points from NBS.D.V using different copula family combinations with $\delta = 40$, from the Harry Potter task-based fMRI study.

5.3 Limitations

While different combinations of segmentation and inference methods in VCCP can be used in different settings, some combinations have shortcomings under certain circumstances. For example, both OBS and NBS segmentation methods have at their core BS, which is limited in that it can only find one change point at a time and then recursively splits the

data. They are also restricted by δ, the minimum between change points and hence the total number of possible change points. It is important that VCCP remains robust to alternative choices to this parameter and the practitioner has the option to obtain more change points by decreasing this value. We studied the robustness of VCCP to changing δ for the SET fMRI set in the Supplementary Materials. Overall, we find that the method finds consistency in the location of the change points but obviously the number of change points alter for different δ. For a theoretical perspective on the minimum between change points, δ, in binary segmentation, see Fryzlewicz (2014).

The MOSUM and WBS methods allow us to bypass specifying this minimum distance between change points. However, MOSUM is excessively conservative when changes in the FC are subtle (change in edge strength, see Simulation 6). This may due to MOSUM's indirect detection in this setting, which reduces the problem to a univariate time series (in our case the BIC reduction series) and then chooses the local maximums as candidate change points. For resting-state fMRI data sets, where subjects smoothly transit between states, MOSUM might confuse the change points with random variations or might amalgamate several moderate peaks into one change point, which ultimately leads to fewer change points. Finally, for the task based fMRI data sets, we obtained conservative results for WBS. This is understandable given WBS stops if the point detected is not a change point or no random intervals remains. If the number of random intervals is not sufficient and the first candidate change point is close to the middle of the time course, the second condition is more likely to occur. Fryzlewicz (2014) proved the lower limit of the number of random draws ($S = \frac{9T^2}{\delta^2}\log(T^2\delta^{-1})$), which we used in this work, but we may need more intervals as additional intervals shorter than 2δ are removed.

6 Conclusion

In this paper, we introduce a new methodology, called Vine Copula Change Point (VCCP), that estimates change points in a vine copula structure. It is motivated by the fact that no other change point method, to the best of our knowledge, considers possible non-linear and non-Gaussian dynamic functional connectivity. Furthermore, the layers of a vine can be

cut, which leads to a simplified and sparse network, and requires less computation time when the dimension of the data set expands.

We also compared different combinations of segmentation methods (NBS, OBS, MOSUM and WBS) and inference methods (Vuong test and the stationary bootstrap) in VCCP. Overall, different combinations serve different needs. NBS with the Vuong test provides the most robust change points for the task-based fMRI data sets, requires fewer inputs, and is more computationally efficient. Although not studied here, perhaps WBS is more suitable for resting-state fMRI data sets as it is sensitive to dense but weak change points in the network. Overall, we found that OBS (used by most previous change point methods in the neuroscience literature) failed to perform well. SB and the Vuong test are both suitable for inference, but Vuong is computationally faster.

We showed using an extensive simulation study that VCCP performs well on multivariate normal data and on Vector autoregression (VAR) data, even in challenging scenarios where the subject alternates between task and rest. The MVN data can be considered whitened data, while the VAR data can be considered un-whitened data. However, the simulation results show the deterioration in performance is not drastic. In addition, the task-based fMRI experimental results clearly indicate the presence of FC change points that are due to non-linear dependence, which has not been explored in a dynamic fashion before in the analysis of neuroimaging data.

While VCCP was applied to task based fMRI data, it could seamlessly be applied to resting-state fMRI data, Electroencephalography (EEG) or Magnetoencephalography (MEG), and electrocorticography (ECoG) data or other time series applications where the network structure is changing. The outputs of VCCP could also be used as an input into a classification model for predicting brain disorders. Finally, the R package **vccp** (Xiong and Cribben, 2021) implementing the methodology from the paper is available from CRAN.

Acknowledgments

X. Xiong was supported by a China Council Scholarship and the work of I. Cribben was supported by the Natural Sciences and Engineering Research Council (Canada) grant

RGPIN-2018-06638 and the Xerox Faculty Fellowship, Alberta School of Business.

References

Anastasiou, A., I. Cribben, and P. Fryzlewicz (2020). Cross-covariance isolate detect: a new change-point method for estimating dynamic functional connectivity. bioRxiv.

Aue, A., S. Hörmann, L. Horváth, M. Reimherr, et al. (2009). Break detection in the covariance structure of multivariate time series models. The Annals of Statistics 37(6B), 4046–4087.

Bakhtiari, R., J. Cummine, A. Reed, C. M. Fox, B. Chouinard, I. Cribben, and C. A. Boliek (2017). Changes in brain activity following intensive voice treatment in children with cerebral palsy. Human brain mapping 38(9), 4413–4429.

Barnett, I. and J.-P. Onnela (2016). Change point detection in correlation networks. Scientific reports 6, 18893.

Bedford, T. and R. M. Cooke (2002). Vines–a new graphical model for dependent random variables. The Annals of Statistics 30(4), 1031–1068.

Bullmore, E. and O. Sporns (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. Nature Reviews Neuroscience 10(3), 186–198.

Chen, J. and A. K. Gupta (1997). Testing and locating variance changepoints with application to stock prices. Journal of the American Statistical association 92(438), 739–747.

Cho, H. and P. Fryzlewicz (2015). Multiple-change-point detection for high dimensional time series via sparsified binary segmentation. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 77(2), 475–507.

Cormen, T. H., C. E. Leiserson, R. L. Rivest, and C. Stein (2009). Introduction to algorithms. MIT press.

Cribben, I. and M. Fiecas (2016). Functional connectivity analyses for fmri data. Handbook of Neuroimaging Data Analysis 369.

Cribben, I., R. Haraldsdottir, L. Atlas, T. Wager, and M. Lindquist (2012, 03). Dynamic connectivity regression: Determining state-related changes in brain connectivity. NeuroImage 61, 907–20.

Cribben, I., T. Wager, and M. Lindquist (2013). Detecting functional connectivity change points for single-subject fmri data. Frontiers in Computational Neuroscience 7, 143.

Cribben, I. and Y. Yu (2017). Estimating whole-brain dynamics by using spectral clustering. Journal of the Royal Statistical Society: Series C (Applied Statistics) 66(3), 607–627.

Dai, M., Z. Zhang, and A. Srivastava (2019). Discovering common change-point patterns in functional connectivity across subjects. Medical image analysis 58, 101532.

Dette, H. and J. Gösmann (2020). A likelihood ratio approach to sequential change point detection for a general class of parameters. Journal of the American Statistical Association 115(531), 1361–1377.

Dette, H. and D. Wied (2016). Detecting relevant changes in time series models. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 78(2), 371–394.

Dißmann, J. F. (2010, May). Statistical inference for regular vines and application. Diplomarbeit, Technische Universität München.

Eichinger, B. and C. Kirch (2018, 02). A mosum procedure for the estimation of multiple random change points. Bernoulli 24, 526–564.

Fan, J. and Y. Fan (2008). High dimensional classification using features annealed independence rules. Annals of statistics 36(6), 2605.

Fontaine, C., R. D. Frostig, and H. Ombao (2020). Modeling dependence via copula of functionals of fourier coefficients. TEST: An Official Journal of the Spanish Society of Statistics and Operations Research 29(4), 1125–1144.

Fryzlewicz, P. (2014). Wild binary segmentation for multiple change-point detection. The Annals of Statistics 42(6), 2243–2281.

Gibberd, A. J. and J. D. Nelson (2014). High dimensional changepoint detection with a dynamic graphical lasso. In ICASSP, pp. 2684–2688.

Greicius, M. D., G. Srivastava, A. L. Reiss, and V. Menon (2004). Default-mode network activity distinguishes alzheimer's disease from healthy aging: evidence from functional mri. Proceedings of the National Academy of Sciences 101(13), 4637–4642.

Hart, B., I. Cribben, M. Fiecas, A. D. N. Initiative, et al. (2018). A longitudinal model for functional connectivity networks using resting-state fmri. NeuroImage 178, 687–701.

Herlin, B., V. Navarro, and S. Dupont (2021). The temporal pole: from anatomy to function-a literature appraisal. Journal of Chemical Neuroanatomy, 101925.

Hrybouski, S., I. Cribben, J. McGonigle, F. Olsen, R. Carter, P. Seres, C. R. Madan, and N. V. Malykhin (2021). Investigating the effects of healthy cognitive aging on brain functional connectivity using 4.7 t resting-state functional magnetic resonance imaging. Brain Structure and Function 226(4), 1067–1098.

Hutchison, R. M., T. Womelsdorf, E. A. Allen, P. A. Bandettini, V. D. Calhoun, M. Corbetta, S. D. Penna, J. Duyn, G. Glover, J. Gonzalez-Castillo, et al. (2013). Dynamic functional connectivity: promises, issues, and interpretations. NeuroImage 80, 360–378.

Inclan, C. and G. C. Tiao (1994). Use of cumulative sums of squares for retrospective detection of changes of variance. Journal of the American Statistical Association 89(427), 913–923.

James, N. A. and D. S. Matteson (2015). ecp: An r package for nonparametric multiple change point analysis of multivariate data. Journal of Statistical Software, Articles 62(7), 1–25.

Joe, H. (1996). Families of m-variate distributions with given margins and m(m-1)/2 bivariate dependence parameters. Lecture Notes-Monograph Series 28, 120–141.

Kao, C., L. Trapani, G. Urga, et al. (2018). Testing for instability in covariance structures. Bernoulli 24(1), 740–771.

Kirch, C., B. Muhsal, and H. Ombao (2015). Detection of changes in multivariate time series with application to EEG data. Journal of the American Statistical Association 110(511), 1197–1216.

Kundu, S., J. Ming, J. Pierce, J. McDowell, and Y. Guo (2018). Estimating dynamic brain functional networks using multi-subject fmri data. NeuroImage 183, 635–649.

Li, Y.-N., D. Li, and P. Fryzlewicz (2019). Detection of multiple structural breaks in large covariance matrices. Preprint on webpage at `stats.lse.ac.uk/fryzlewicz/wbscov/wbscov.pdf`.

Matteson, D. S. and N. A. James (2014). A nonparametric approach for multiple change point analysis of multivariate data. Journal of the American Statistical Association 109(505), 334–345.

Menon, V. (2011). Large-scale brain networks and psychopathology: a unifying triple network model. Trends in cognitive sciences 15(10), 483–506.

Morales Napoles, O. (2016, 03). About the number of vines and regular vines on n nodes. Technical report, Institutional Repository.

Ofori-Boateng, D., Y. R. Gel, and I. Cribben (2021). Nonparametric anomaly detection on time series of graphs. Journal of Computational and Graphical Statistics, 1–12.

Ondrus, M., E. Olds, and I. Cribben (2021). Factorized binary search: change point detection in the network structure of multivariate high-dimensional time series. arXiv preprint arXiv:2103.06347.

Pace-Schott, E. F. and D. Picchioni (2017). Chapter 51 - neurobiology of dreaming. In M. Kryger, T. Roth, and W. C. Dement (Eds.), Principles and Practice of Sleep Medicine (Sixth Edition) (Sixth Edition ed.)., pp. 529–538.e6. Elsevier.

Politis, D. and J. Romano (1994, 12). The stationary bootstrap. Journal of The American Statistical Association - J AMER STATIST ASSN 89, 1303–1313.

Schröder, A. L. and H. Ombao (2019). Fresped: Frequency-specific change-point detection in epileptic seizure multi-channel eeg data. Journal of the American Statistical Association 114(525), 115–128.

Schwarz, G. et al. (1978). Estimating the dimension of a model. The annals of statistics 6(2), 461–464.

Sklar, M. (1959). Fonctions de repartition an dimensions et leurs marges. Publ. inst. statist. univ. Paris 8, 229–231.

Sundararajan, R. R. and M. Pourahmadi (2018). Nonparametric change point detection in multivariate piecewise stationary time series. Journal of Nonparametric Statistics 30(4), 926–956.

Tracy, J. I. and S. B. Boswell (2008). Mesial temporal lobe epilepsy: a model for understanding the relationship between language and memory. In Handbook of the neuroscience of language, pp. 319–328. Elsevier.

Tzourio-Mazoyer, N., B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, and M. Joliot (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.

Vuong, Q. (1989, 02). Likelihood ratio tests for model selection and non-nested hypotheses. Econometrica 57, 307–33.

Wehbe, L., B. Murphy, P. Talukdar, A. Fyshe, A. Ramdas, and T. Mitchell (2014). Simultaneously uncovering the patterns of brain regions involved in different story reading subprocesses. PloS one 9(11), e112575.

Xiong, X. and I. Cribben (2021). vccp: Vine Copula Change Point Detection in Multivariate Time Series. R package version 0.1.0.

Zhang, W., N. A. James, and D. S. Matteson (2017). Pruning and nonparametric multiple change point detection. In 2017 IEEE International Conference on Data Mining Workshops (ICDMW), pp. 288–295.

Supplementary Materials: Beyond linear dynamic functional connectivity: a vine copula change point model

Xin Xiong
Department of Biostatistics
Harvard T. H. Chan School of Public Health
and
Ivor Cribben *
Department of Accounting and Business Analytics
Alberta School of Business

August 7, 2021

1 Kendall's τ and copula formulae

Definition 1 (**Kendall's τ**) *Let (X_1, Y_1) and (X_2, Y_2) be two independent pairs of random variables with joint cdf F and marginal distributions F_X and F_Y. Kendall's τ is given by*

$$\tau(X, Y) = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0]. \tag{1}$$

The estimate of Kendall's τ is given by :

$$\hat{\tau} = \frac{N_{conc} - N_{disc}}{\sqrt{N_{conc} + N_{extra_y}} * \sqrt{N_{conc} + N_{extra_x}}} \tag{2}$$

where N_{conc} is the number of concordant pairs (the orderings of the two Xs is the same as the ordering of the two Ys), N_{disc} is the number of discordant pairs (the ordering of the

*To whom correspondence should be addressed.

two Xs is opposite from the ordering of the two Ys) and N_{extra_x}, N_{extra_y} are the numbers of extra-X pairs (i, j) such that $X_i = X_j, i \neq j$ and extra-Y pairs (i, j) such that $Y_i = Y_j, i \neq j$, respectively.

The formulas for the copulae we consider in this work are given by:

1. **Gauss Copula**

$$C(u_1, u_2) = \Phi_2(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \rho)$$

$$c(u_1, u_2) = \frac{1}{\sqrt{1-\rho^2}} \exp(-\frac{\rho^2(x_1^2 + x_2^2) - 2\rho x_1 x_2}{2(1-\rho^2)})$$

$$\tau = f(\rho) = \frac{2}{\pi} \arcsin(\rho)$$

where $x_1 = \Phi^{-1}(u_1)$, $x_2 = \Phi^{-1}(u_2)$, and $\Phi_2(., ., \rho)$ is the bivariate normal cdf with two standard univariate marginal cdfs and a Pearson correlation coefficient, ρ.

2. ***t* Copula**

$$c(u_1, u_2) = \frac{2}{2\pi dt_\nu(x_2)\sqrt{1-\rho^2}} \left(1 + \frac{x_1^2 + x_2^2 - 2\rho x_1 x_2}{\nu(1-\rho^2)}\right)^{-\frac{\nu+2}{2}}$$

$$\tau = f(\rho) = \frac{2}{\pi} \arcsin(\rho)$$

where $dt_\nu(x_2) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sqrt{\pi\nu}}(1 + \frac{x^2}{\nu})^{-\frac{\nu+2}{2}}$, $x_1 = t_\nu^{-1}(u_1)$, $x_2 = t_\nu^{-1}(u_2)$, and $t_\nu^{-1}(.)$ is the inverse cdf of univariate t distribution with ν degrees of freedom.

3. **Frank Copula**

$$C(u_1, u_2) = -\frac{1}{\theta} \log(1 + \frac{(\exp(-\theta u_1) - 1)(\exp(-\theta u_2) - 1)}{\exp(-\theta) - 1})$$

$$\tau = f(\theta) = 1 - \frac{4}{\theta} + 4 \cdot \frac{D_1(\theta)}{\theta}$$

where $D_1(\theta) = \int_0^\theta \frac{x/\theta}{\exp(x)-1}$ and $\theta \in (-\infty, +\infty)$. As $\theta \to -\infty$, the two variables linked by the Frank copula become more negatively correlated. As $\theta \to +\infty$, the two variables become more positively correlated.

4. **Clayton Copula**

$$C(u_1, u_2) = (u_1^{-\theta} + u_2^{-\theta} - 1)^{-\frac{1}{\theta}}$$

$$\tau = f(\theta) = \frac{\theta}{\theta + 1}$$

where $\theta \in (-1, \infty) \backslash 0$. Clayton copulas are only used to describe positive dependence. As $\theta \to 0$, the two variables become more independent, and as $\theta \to \infty$, the two variables become more positively correlated.

5. **Gumbel Copula**

$$C(u_1, u_2) = \exp(-((-\log u_1)^{\theta} + (-\log u_1)^{\theta})^{\frac{1}{\theta}}$$

$$\tau = f(\theta) = 1 - \frac{1}{\theta}$$

where $\theta \in (1, \infty)$. Gumbel copulas are only used to describe positive dependence. When $\theta = 1$, the two variables are independent, and as $\theta \to \infty$, the two variables become more positively correlated.

2 Anxiety fMRI data set

This fMRI data set (SET data) was taken from an anxiety-inducing experiment. Before the start of the experiment, participants were told that they would be given 2 minutes in the scanner to present a 7-minute speech, though there was a small chance that they might not be randomly selected to give the speech. Once the fMRI acquisition began, subjects rested for 2 min and then an instruction slide showing the topic - 'Why you are a good friend' could be seen for 15 s. When the slide disappeared, subjects went through a 2min silent brainstorming which was interrupted by another instruction slide that informed participants they would not have to give the speech. Then they rested for 2min which completed the functional run. Data was acquired and preprocessed the same way as in the previous work (Wager et al., 2009; Cribben et al., 2012, 2013). The experimental data set consists of 5 ROIs: (1) visual cortex, (2) left superior temporal sulci, (3) ventral striatum, (4) right

superior sulci and (5) ventromedial PFC that were created by averaging the voxel time series across the entire region. $N = 26$ subjects were involved in the experiment, which lasted for $T = 215$, with TR = 2 s. Our VCCP methodology weas applied to individual subjects to learn if the stressor onset triggered a change point in the FC network between the 5 ROIs.

3 Results

3.1 Anxiety fMRI data

In Figure 1, we display the results from applying all segmentation methods of VCCP (with the Vuong test) to the SET fMRI data set. The solid vertical lines indicate the times of the showing and of the removal of the visual cues. We compare the change points detected across the segmentation methods using both the Gauss and the 5-copula family. Similar to the simulation results, the best method segmentation method is NBS, finding the most change points that are close to the times of the showing and of the removal of the visual cues (change points are more tightly clustered around the four vertical lines). OBS and MOSUM have a similar performance to NBS, however MOSUM detects a smaller number of change points across all subjects. WBS has the worst performance. The 5-copula family has a marginal superior performance over the Gauss copula across all segmentation methods of VCCP, but many of the change points align, indicating that the change points are due to changes in the Gauss copula.

In Figure 2, we display the results from applying all segmentation methods of VCCP (with the SB) to the SET fMRI data set. The results are very similar to applying the Vuong test (Figure 1). The main difference is that the SB is more computationally intense due to the resampling procedure.

As discussed in Section 2.3 (main file), the parameter δ, represents the the minimal distance between two candidate change points. It needs to be sufficiently large enough to estimate a stable D-vine structure, but also small enough in order to not miss candidate change points. In Figure 3, we explore the change points detected from applying the

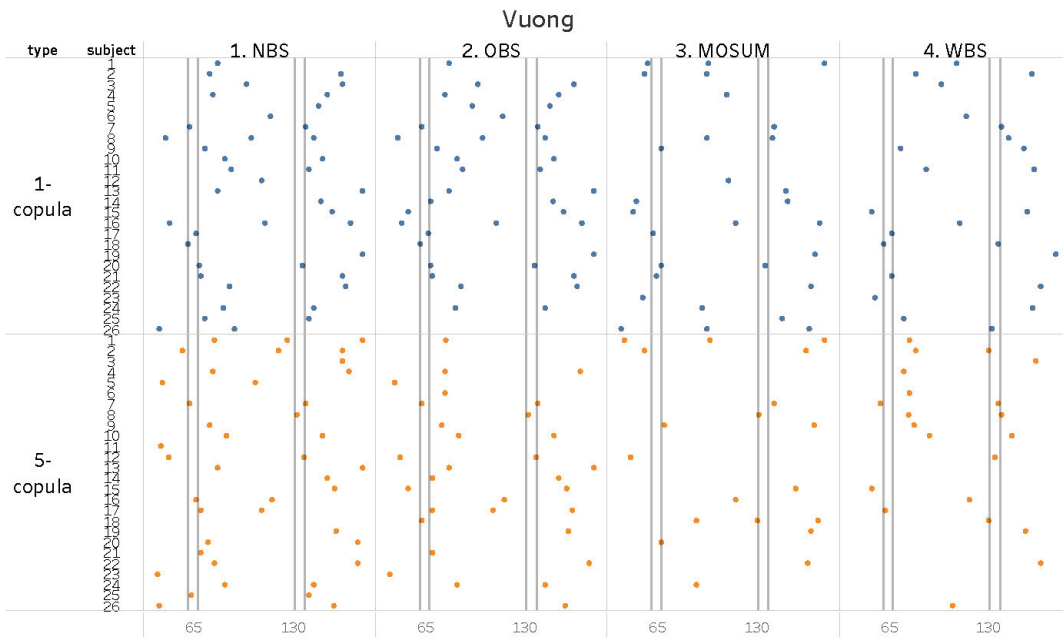

Figure 1: The results from applying all segmentation methods of VCCP (with the Vuong test) to the SET fMRI data set. NBS, OBS, MOSUM, and WBS denote the adapted binary segmentation, old binary segmentation, moving sum, and the wild binary segmentation methods, respectively. We set the minimum distance between change points to be $\delta = 40$ and used the Gauss and the 5-copula family. The solid vertical lines indicate the times of the showing and of the removal of the visual cues.

combination NBS.D.V (NBS segmentation, D-vine, and the Vuong test) from VCCP to the SET fMRI data set using values of $\delta = 20, 30, 40$. As expected, the number of change points decreases as δ increases. This inverse relationship is reasonable as once a candidate point t_0 is detected, we exclude other possible change points in the range $[t_0 - \delta, t_0 + \delta]$ to ensure sufficient sample size during the vine copula construction. Remarkably, even when we explore a limited range ($\delta = 40$), NBS can identify change points close to the times of the showing and of the removal of the visual cues. The $\delta = 40$ appears to provide and the best and most stable results. The 5-copula family has a superior performance over the Gauss copula across all δ values, but many of the change points align, indicating that the change points are due to changes in the Gauss copula. The results from using the SB instead of the Vuong test are very similar.

Using the results from Figure 3 ($\delta = 40$), we plot the networks for subjects 2, 7 and

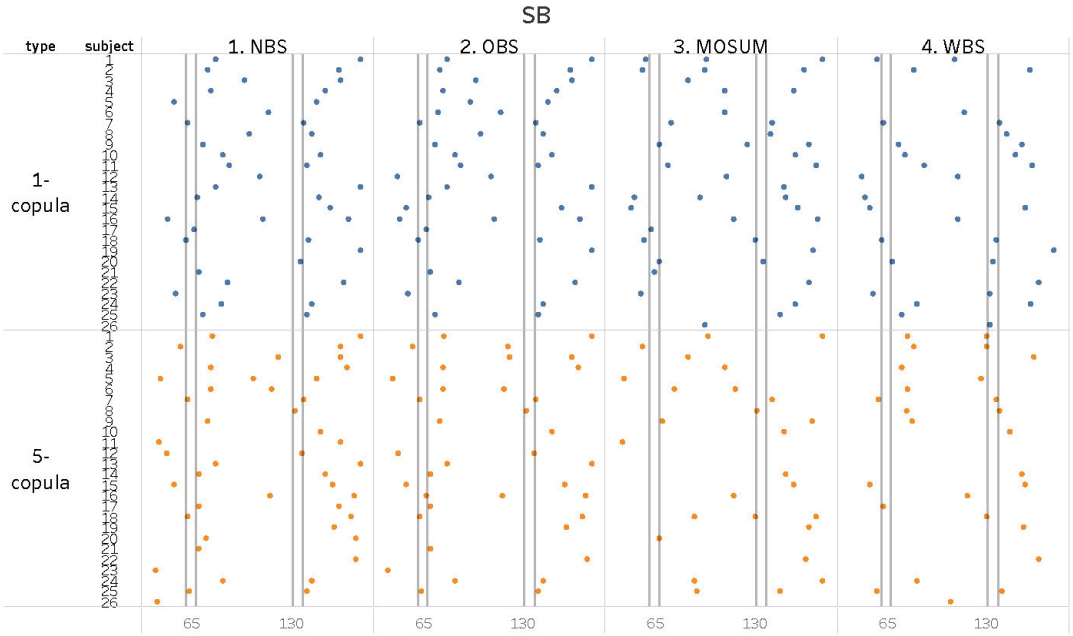

Figure 2: The results from applying all segmentation methods of VCCP (with the SB) to the SET fMRI data set. NBS, OBS, MOSUM, and WBS denote the adapted binary segmentation, old binary segmentation, moving sum, and the wild binary segmentation methods, respectively. We used the Gauss copula and the 5-copula family. The solid vertical lines indicate the times of the showing and of the removal of the visual cues.

25 in the SET data, with change points detected by NBS.D.V and the 5-copula family in Figure 4. The main graph of each partition corresponds to the Kendall's τ network, while the two sub-graphs on the top corners summarize the tail dependence. All partition specific networks (apart from the last partition for subject 7) contain at least an edge with heavy tail dependence, some of them are only displayed in dashed lines in the main graph due to an insignificant Kendall's τ. Since the tail dependence cannot be described by Gaussian models, changes in these non-Kendall-but-heavy-tail correlations may be neglected if researchers build their change point detection model on the Gaussianity assumption. This may be the reason for a sparser change point detection using the Gauss copula than the results from the 5-copula family. For example, in Figure 3 the Gauss copula is only able to detect the two change points for subject 2, while the 5-copula family can detect three change points. For subject 7, the change points using the Gauss copula and the 5-copula family are identical indicating that the change points are due to the Gauss copula with little tail dependence.

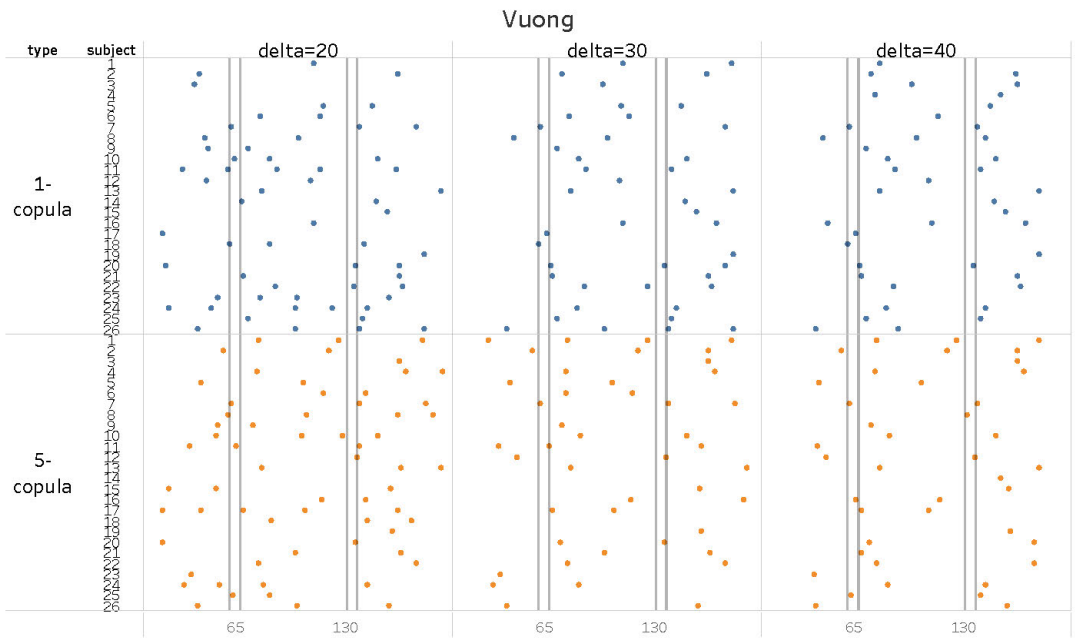

Figure 3: The results from applying the combination NBS.D.V (with Gauss and 5-copula family) from VCCP to the SET fMRI data set. The minimum distance between change points is set to $\delta = 20, 30, 40$. The solid vertical lines indicate the times of the showing and of the removal of the visual cues.

The middle panel of Figure 4 verifies this.

In Figure 4, we also label the best copula type for each edge. Only a small portion of them are denoted as Gauss copulas (N), while most of the edges are best fitted with Clayton (C) or Gumbel (G) copulas, consistent with their heavy tail dependence. During different periods, some best-fitting copula types remain the same, such as the F copula linking ventromedial PFC (vPFC) and right superior temporal sulci (R.STS), and Gumbel copula linking the right superior temporal sulci (R.STS) and the left superior temporal sulci (L.STS) in subject 25. However, some edges experience changes not only in connecting strength but also in copula types. The diversity and the variability of copula types in the networks once again prove the superiority of a vine copula model with multiple copula candidates.

Overall, for the anxiety-inducing experiment, the NBS segmentation method performs best with consistent change point detection close to time points $(61-67.5)$ and $(131-137.5)$. The performance of the SB and the Vuong test are similar. While the 5-copula family

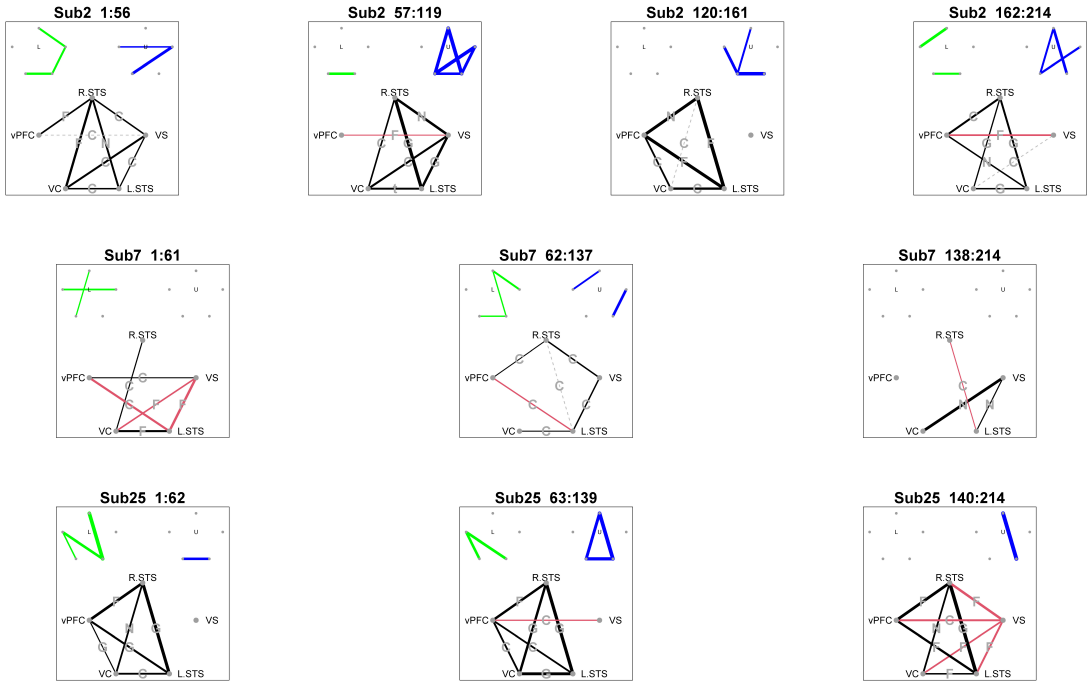

Figure 4: The networks for subjects 2, 7, and 25 in the SET data, with change points detected by NBS.D.V and the 5-copula family. Black and red lines in the main networks correspond to positive and negative Kendall's τ, with a grey label marking the best-fitting copula type for each edge. Only edges that are found to be significant in Kendall's correlation test or with heavy heavy tail dependence are shown. The green sub-network on the top left corner represents lower tail dependence, and the blue sub-network on the top right corner represents upper tail dependence. Dashed lines indicate ROIs with heavy tail dependence but insignificant Kendall's τ.

allowed for more flexibility in the model and change points in the network, it also increases the computational time.

3.2 Simulations

Table 1 shows the results from the baseline case, where the data is multivariate normal (MVN) with no change point. The divisive methods in the **ecp** R package perform best, achieving the nominal Type I error. However, NBS and OBS perform similarly, they do not find many FP change points, but when candidate change points are at the edges of the time course their performance deteriorates. MOSUM and WBS, on the contrary, seem to be overly sensitive and hence have a higher Type I error rate. The worst combination is

MOSUM.D.V, which detects 39 FP change points.

Table 1: The results from applying all variations of VCCP and the two best performing non-parametric methods in the **ecp** R package to the MVN data with no change points (Simulation 1) over 100 simulated data sequences. NBS, OBS, MOSUM, and WBS denote the adapted binary segmentation, old binary segmentation, moving sum, and the wild binary segmentation methods, respectively. The D denotes the D-vine. FP and Num.det denote the false positive rate and number of change points detected, respectively.

	Stationary Bootstrap				Vuong Test				divisive	
	NBS.D	OBS.D	MOSUM.D	WBS.D	NBS.D	OBS.D	MOSUM.D	WBS.D	alpha=1	alpha=2
FP.31-60	0.01	0.01	0.05	0.03	0	0	0.05	0.05	0.02	0.01
FP.61-90	0.01	0.01	0.02	0	0	0	0.06	0.03	0	0.01
FP.91-120	0	0	0.03	0.03	0	0	0.07	0.05	0	0.01
FP.121-150	0	0	0.05	0.02	0	0	0.07	0.05	0.02	0.02
FP.151-180	0.01	0.01	0.03	0.01	0.01	0.01	0.02	0.03	0.01	0.01
FP.181-210	0	0	0.02	0.03	0	0	0.04	0.04	0	0
FP.211-240	0.01	0.01	0.03	0.04	0.01	0.01	0.05	0.03	0	0
FP.241-270	0.08	0.09	0.02	0	0.07	0.06	0.03	0.04	0	0
Num.det	0.12	0.13	0.25	0.16	0.09	0.08	0.39	0.32	0.05	0.06

Table 2 shows the results from the MVN data with 4 change points (Simulation 2). Again, the NBS and OBS segmentation methods shared the best performance with each perfectly detecting the four true change points within the bias range for each iteration, especially in combination with SB. NBS.D.SB and OBS.D.SB have the smallest scaled Hausdorff distance, indicating the detected change points are located closest to the true change points. Other combinations such as MOSUM.D present good results, but with a higher FP rate and larger scaled Hausdorff distance. WBS appears to be too conservative with inferior accuracy with both the SB and the Vuong test. Overall, e.cp3o and ks.cp3o_delta have inferior performance to all variations of VCCP with lower TP, higher TN and higher FP. Both e.cp3o and ks.cp3o detected more than 4 change points on average.

In Simulation 3, the nonzero edges are fixed over the entire time course with the edge weights varying between change points. This is a very difficult simulation but NBS in combination with both the SB and the Vuong test perform well, identifying almost all 4 change points across the 100 iterations (Table 3). Overall, NBS.D.V has the best performance, having the smallest scaled Hausdorff distance and the highest TP rate. However, NBS.D.SB has almost the same performance but with a smaller number of FPs. MOSUM has the second best performance across the segmentation methods, identifying over 3 change points

Table 2: The results from applying all variations of VCCP and the two best performing non-parametric methods in the **ecp** R package to the MVN data with 4 change points (Simulation 2) over 100 simulated data sequences. NBS, OBS, MOSUM, and WBS denote the adapted binary segmentation, old binary segmentation, moving sum, and the wild binary segmentation methods, respectively. The D denotes the D-vine. TP, FN, FP, Cand.num, Num.det, and d_H denote the true positive rate, false negative rate, false positive rate, number of candidate change points, number of change points detected, scaled Hausdorff distance, respectively. ks.cp3o_d is short for ks.cp3o_delta.

	Stationary Bootstrap				Vuong test				e.cp3o	ks.cp3o_d
	NBS	OBS	MOSUM	WBS	NBS	OBS	MOSUM	WBS	alpha=1 Kmax=8	Kmax=8
TP.61	1	1	1	0.84	1	0.97	1	0.8	0.54	0.47
TP.121	1	1	1	1	1	0.96	1	0.98	0.52	0.42
TP.181	1	1	1	0.98	1	0.96	1	1	0.52	0.47
TP.241	1	1	1	0.73	0.99	0.96	1	0.82	0.37	0.4
FN.61	0	0	0	0	0	0.01	0	0	0.44	0.46
FN.121	0	0	0	0	0	0.01	0	0	0.29	0.34
FN.181	0	0	0	0	0	0	0	0	0.26	0.42
FN.241	0	0	0	0.03	0.01	0.02	0	0.02	0.37	0.49
FP.0-60	0	0	0.01	0.02	0	0	0.01	0.03	0.4	0.48
FP.61-120	0	0	0.02	0.01	0.02	0.02	0.02	0.03	0.83	0.74
FP.121-180	0	0	0	0	0.09	0.07	0	0	0.8	0.79
FP.181-240	0	0	0.03	0.01	0.02	0.02	0.04	0.02	0.57	0.73
FP.241-300	0	0	0.01	0.01	0.01	0.01	0.01	0.05	0.3	0.58
Cand.num	4.67	4.57	4.12	3.73	4.67	4.47	4.12	3.77	8.84	9.94
Num.det	4	4	4.07	3.6	4.13	3.97	4.08	3.7	4.97	5.05
d_H	0.03	0.03	0.08	0.39	0.1	0.23	0.09	0.34	0.77	0.54

on average, while WBS outperformed OBS. The OBS.D.V combination has the worst performance, with no TP rate exceeding 65% and with a large scaled Hausdorff distance. It also has a high FN rate at the last change point ($t = 241$), which may be due to the Vuong test's conservative behavior. In terms of the non-parametric methods, e.cp3o and ks.cp3o_delta perform best. However, they are all inferior to all variations of VCCP, with the FP rates being particularly poor.

We also compare the results in Table 2 (MVN data with 4 change points) with the results in Table 3 (MVN data with 4 weak change points), and find that the results for the latter are inferior due to the less distinctive changing patterns. The number of detected change points declines from 3.27 to 2.58 on average across all variations of VCCP. For the latter data set, the change point located at time point $t = 61$ is the most difficult to detect, where the highest TP rate is 61% from NBS.D.SB. While the FN rate remains low in the former data, it increases for the latter data, especially using MOSUM, where the FN rate for the first partition exceeds 0.3 when combined with the SB. Since MOSUM is based on segmenting the reduced BIC series, the soaring FN rate may be due to the selection of the fixed bandwidth G. Due to its conservative behavior, WBS does not perform well in either data set. However, it did have low FN and FP rates and high TP rates at the change points at $t = 121$ and $t = 241$ in the latter data set.

In Simulation 4, the MVN data is composed of networks that contain 7 small change points (changes in edge structure) and not weak change points (changes in edge strength). NBS and OBS with the SB have the best performance, with OBS marginally outperforming NBS. MOSUM and WBS have a similar performance in terms of scaled Hausdorff distance with MOSUM detecting more change points on average. Overall the SB marginally outperforms the Vuong test. However, for MOSUM, the Vuong test outperforms the SB. The worst TP and FN rates are for the last change point at time point $t = 526$ for the Vuong test, while the change point at $t = 301$ is difficult to detect using the SB.

With respect to the non-parametric methods, all variations of VCCP outperform them. In particular, for this data set, best performing methods, e.cp3o_delta E.cp3o, have low TP, high FN and very high FP rates.

Table 3: The results from applying all variations of VCCP and the two best performing non-parametric methods in the **ecp** R package to the MVN data with 4 weak change points (Simulation 3) over 100 simulated data sequences. NBS, OBS, MOSUM, and WBS denote the adapted binary segmentation, old binary segmentation, moving sum, and the wild binary segmentation methods, respectively. The D denotes the D-vine. TP, FN, FP, Cand.num, Num.det, and d_H denote the true positive rate, false negative rate, false positive rate, number of candidate change points, number of change points detected, scaled Hausdorff distance, respectively. ks.cp3o_d is short for ks.cp3o_delta.

	Stationary Bootstrap				Vuong test				e.cp3o	ks.cp3o_d
	NBS	OBS	MOSUM	WBS	NBS	OBS	MOSUM	WBS	alpha=1 Kmax=8	Kmax=8
TP.61	0.61	0.51	0.43	0.39	0.6	0.42	0.45	0.29	0.48	0.41
TP.121	0.87	0.87	0.89	0.93	0.89	0.63	0.85	0.91	0.54	0.49
TP.181	0.61	0.72	0.67	0.45	0.67	0.47	0.66	0.38	0.44	0.49
TP.241	0.96	0.95	0.96	0.93	0.96	0.65	0.91	0.84	0.29	0.53
FN.61	0.16	0.12	0.31	0.14	0.17	0.09	0.29	0.09	0.29	0.5
FN.121	0.04	0.01	0	0.01	0.02	0.07	0.04	0.01	0.28	0.39
FN.181	0.25	0.12	0.12	0.09	0.19	0.08	0.13	0.17	0.32	0.27
FN.241	0.01	0.01	0	0.04	0.01	0.12	0.05	0.09	0.35	0.38
FP.0-60	0.13	0.14	0.06	0.03	0.14	0.05	0.05	0.01	0.5	0.57
FP.61-120	0.1	0.17	0.09	0.03	0.15	0.09	0.11	0.03	0.78	0.76
FP.121-180	0.09	0.09	0.15	0.12	0.12	0.07	0.16	0.13	0.67	0.75
FP.181-240	0.04	0.1	0.06	0.04	0.13	0.1	0.1	0.05	0.55	0.74
FP.241-300	0.03	0.04	0.06	0.02	0.06	0.02	0.05	0.01	0.42	0.54
Cand.num	5.38	3.55	4.48	2.9	5.38	3.44	4.48	3.25	8.46	9.74
Num.det	3.38	3.55	3.31	2.9	3.67	2.47	3.29	2.58	4.76	5.3
Hausdorff	0.64	0.52	0.66	0.79	0.52	1.57	0.81	1.00	1.00	0.54

Table 4: The results from applying all variations of VCCP and the two best performing non-parametric methods in the **ecp** R package to the MVN data with 7 small change points (Simulation 4) over 100 simulated data sequences. NBS, OBS, MOSUM, and WBS denote the adapted binary segmentation, old binary segmentation, moving sum, and the wild binary segmentation methods, respectively. The D denotes the D-vine. TP, FN, FP, Cand.num, Num.det, and d_H denote the true positive rate, false negative rate, false positive rate, number of candidate change points, number of change points detected, scaled Hausdorff distance, respectively. e.cp3o_d is short for e.cp3o_delta.

	Stationary Bootstrap				Vuong test				e.cp3o_d	ks.cp3o
	NBS	OBS	MOSUM	WBS	NBS	OBS	MOSUM	WBS	alpha=1 Kmax=14	Kmax=14
TP.76	1	1	0.67	0.57	1	0.96	1	0.56	0.4	0.53
TP.151	1	1	0.42	0.97	1	0.96	0.96	0.96	0.36	0.46
TP.226	0.96	0.97	0.72	0.92	0.95	0.96	0.87	0.86	0.39	0.39
TP.301	0.73	0.68	0.44	0.44	0.71	0.44	0.49	0.3	0.47	0.43
TP.376	0.96	0.94	0.74	0.85	0.93	0.66	0.78	0.73	0.49	0.32
TP.451	1	1	0.99	1	1	1	1	1	0.44	0.17
TP.526	0.78	0.84	0.57	0.23	0.65	0.54	0.32	0.21	0.39	0.18
FN.76	0	0	0.02	0.02	0	0.02	0	0	0.01	0.23
FN.151	0	0	0	0	0	0	0.01	0	0.62	0.29
FN.226	0.02	0	0	0	0.03	0	0.03	0.05	0.55	0.34
FN.301	0.07	0	0.12	0.02	0.09	0.01	0.24	0.02	0.48	0.34
FN.376	0	0	0.01	0.01	0.04	0.26	0.11	0.05	0.48	0.49
FN.451	0	0	0	0	0	0	0	0	0.56	0.48
FN.526	0.12	0.02	0.09	0.25	0.19	0.09	0.27	0.21	0.21	0.34
FP.0-75	0.03	0.02	0.28	0.02	0.11	0.03	0.04	0.01	0.6	0.62
FP.76-150	0	0.02	0	0.02	0.13	0.08	0.1	0.02	0.81	0.75
FP.151-225	0.04	0.02	0.07	0.01	0.16	0.07	0.05	0.04	0.9	0.82
FP.226-300	0.18	0.16	0.13	0.03	0.19	0.11	0.12	0.03	0.89	0.73
FP.301-375	0.13	0.08	0.13	0.06	0.18	0.05	0.15	0.05	0.87	0.56
FP.376-450	0.03	0.04	0.08	0.02	0.07	0.02	0.1	0	0.85	0.39
FP.451-525	0.05	0.07	0.02	0.01	0.24	0.13	0.05	0.05	0.83	0.31
FP.526-600	0.13	0.08	0.13	0.06	0.18	0.05	0.15	0.05	0.87	0.56
Cand.num	10.57	7.92	6.42	5.76	10.44	7.04	8.58	5.47	20.14	16.75
Num.det	6.91	6.92	5.4	5.21	7.46	6.05	6.14	4.84	9.34	8.38
d_H	0.38	0.33	0.95	0.96	0.52	0.75	0.84	1.03	0.44	2.30

We also put the performance of all the non-parametric methods from Simulation 1 to Simulation 7 in table 5, 6, 7, 8, 9, 10, and 11, respectively.

Table 5: The results from applying 5 non-parametric methods in the **ecp** R package to the MVN data with no change points (Simulation 1) over 100 simulated data sequences. TP, FN, FP, Cand.num, Num.det, and d_H denote the true positive rate, false negative rate, false positive rate, number of candidate change points, number of change points detected, scaled Hausdorff distance, respectively.

	divisive		e.cp3o				e.cp3o_delta				ks.cp3o		ks.cp3o_delta	
alpha	1	2	1	1	2	2	1	1	2	2				
Kmax			2	4	2	4	2	4	2	4	2	4	2	4
FP.31-60	0.02	0.01	0.17	0.28	0.17	0.26	0	0	0	0	0.08	0.45	0.21	0.43
FP.61-90	0	0.01	0.09	0.29	0.07	0.3	0	0.19	0	0.16	0.1	0.51	0.17	0.34
FP.91-120	0	0.01	0.06	0.56	0.06	0.57	0	0.81	0.05	0.82	0.15	0.47	0.1	0.28
FP.121-150	0.02	0.02	0	0.42	0.02	0.39	0.52	0.06	0.44	0.08	0.17	0.37	0.14	0.34
FP.151-180	0.01	0.01	0.18	0.45	0.16	0.49	0.48	0.16	0.51	0.23	0.25	0.28	0.16	0.36
FP.181-210	0	0	0.14	0.38	0.19	0.37	0	0.77	0	0.64	0.14	0.15	0.09	0.25
FP.211-240	0	0	0.16	0.15	0.14	0.14	0	0.23	0	0.3	0.08	0.18	0.06	0.26
FP.241-270	0	0	0.2	0.14	0.19	0.14	0	0	0	0.05	0.03	0.1	0.07	0.19
Num.det	0.05	0.06	1	2.67	1	2.66	1	2.22	1	2.28	1	2.51	1	2.45

Table 6: The results from applying 5 non-parametric methods in the **ecp** R package to the MVN data with 4 change points (Simulation 2) over 100 simulated data sequences. TP, FN, FP, Cand.num, Num.det, and d_H denote the true positive rate, false negative rate, false positive rate, number of candidate change points, number of change points detected, scaled Hausdorff distance, respectively.

	divisive		e.cp3o				e.cp3o_delta				ks.cp3o		ks.cp3o_delta	
alpha	1	2	1	1	2	2	1	1	2	2				
Kmax			4	8	4	8	4	8	4	8	4	8	4	8
TP.61	0.08	0.03	0.38	0.54	0.33	0.46	0.16	1	0.15	1	0.34	0.61	0.23	0.47
TP.121	0.05	0.02	0.32	0.52	0.28	0.47	0.13	1	0.15	0.97	0.36	0.54	0.22	0.42
TP.181	0.06	0.04	0.25	0.52	0.27	0.44	0.11	1	0.2	0.98	0.17	0.5	0.25	0.47
TP.241	0.01	0.01	0.11	0.37	0.15	0.4	0.11	1	0.19	1	0.09	0.35	0.16	0.4
FN.61	0	0	0.21	0.44	0.16	0.36	0.98	0.48	0.85	0.58	0.25	0.47	0.21	0.46
FN.121	0.01	0	0.26	0.29	0.27	0.27	0.92	0.49	0.7	0.59	0.3	0.4	0.1	0.34
FN.181	0	0	0.33	0.26	0.23	0.26	0.88	0.49	0.63	0.59	0.36	0.43	0.11	0.42
FN.241	0	0	0.12	0.37	0.14	0.32	0.93	0.49	0.78	0.61	0.11	0.53	0.13	0.49
FP.0-60	0.03	0.04	0.19	0.4	0.22	0.42	0	0	0	0	0.26	0.52	0.21	0.48
FP.61-120	0.1	0.05	0.53	0.83	0.54	0.84	0.74	1	0.7	1	0.57	0.83	0.43	0.74
FP.121-180	0.1	0.09	0.41	0.8	0.44	0.81	0.36	1	0.37	1	0.38	0.72	0.36	0.79
FP.181-240	0.03	0.01	0.26	0.57	0.25	0.65	0.85	1	0.69	1	0.24	0.67	0.4	0.73
FP.241-300	0.02	0.01	0.08	0.3	0.11	0.32	0	0	0	0	0.15	0.47	0.26	0.58
Cand.num	1.53	1.31	4.86	8.84	5	8.66	6.21	14.99	5.92	14.73	5.01	10.2	4.07	9.94
Num.det	0.49	0.31	2.55	4.97	2.61	4.91	2.37	7	2.37	6.97	2.49	5.25	2.42	5.05
d_H	3.46	3.68	1.82	0.77	1.81	0.79	0.70	0.48	0.75	0.47	1.73	0.81	1.40	0.54
Time	0.76	0.59	372.95	458.21	266.53	368.92	0.06	0.07	0.05	0.06	0.23	0.26	0.01	0.02

Table 7: smallThe results from applying 5 non-parametric methods in the **ecp** R package to the MVN data with 4 weak change points (Simulation 3) over 100 simulated data sequences. TP, FN, FP, Cand.num, Num.det, and d_H denote the true positive rate, false negative rate, false positive rate, number of candidate change points, number of change points detected, scaled Hausdorff distance, respectively.

	divisive		e.cp3o				e.cp3o_delta				ks.cp3o		ks.cp3o_delta	
alpha	1	2	1	1	2	2	1	1	2	2				
Kmax			4	8	4	8	4	8	4	8	4	8	4	8
TP.61	0.02	0.02	0.35	0.48	0.35	0.55	0.16	1.00	0.20	1.00	0.34	0.56	0.30	0.41
TP.121	0	0	0.33	0.54	0.3	0.5	0.09	1	0.12	1	0.28	0.49	0.26	0.49
TP.181	0.01	0	0.18	0.44	0.17	0.41	0.13	1	0.18	1	0.18	0.49	0.19	0.49
TP.241	0.01	0	0.13	0.29	0.14	0.27	0.15	1	0.11	1	0.13	0.38	0.26	0.53
FN.61	0	0	0.22	0.29	0.24	0.27	0.93	0.42	0.77	0.58	0.24	0.45	0.12	0.5
FN.121	0	0	0.3	0.28	0.28	0.22	0.95	0.42	0.82	0.59	0.27	0.47	0.14	0.39
FN.181	0	0	0.25	0.32	0.26	0.3	0.95	0.42	0.82	0.58	0.26	0.38	0.12	0.27
FN.241	0	0	0.16	0.35	0.17	0.35	0.95	0.42	0.9	0.58	0.17	0.54	0.12	0.38
FP.0-60	0.02	0.01	0.34	0.5	0.32	0.52	0	0	0	0	0.32	0.65	0.24	0.57
FP.61-120	0.01	0.01	0.55	0.78	0.53	0.79	0.79	1	0.75	1	0.53	0.77	0.37	0.76
FP.121-180	0	0	0.32	0.67	0.32	0.63	0.42	1	0.39	1	0.43	0.83	0.37	0.75
FP.181-240	0.01	0.02	0.22	0.55	0.25	0.62	0.78	1	0.74	1	0.24	0.65	0.31	0.74
FP.241-300	0	0	0.16	0.42	0.17	0.45	0	0	0	0	0.09	0.46	0.24	0.54
Cand.num	1.08	1.06	5.01	8.46	4.96	8.26	6.3	15	6.12	14.87	5.14	10.2	4.11	9.74
Num.det	0.08	0.06	2.63	4.76	2.55	4.8	2.44	7	2.41	7	2.5	5.35	2.53	5.3
d_H	3.91	3.94	1.97	1.00	2.03	1.03	0.69	0.47	0.73	0.48	1.76	0.70	1.35	0.54
Time	0.59	0.50	374.21	459.35	285.29	371.75	0.06	0.07	0.05	0.06	0.24	0.25	0.01	0.02

Table 8: smallThe results from applying 4 non-parametric methods in the **ecp** R package to the MVN data with 7 small change points (Simulation 4) over 100 simulated data sequences. TP, FN, FP, Cand.num, Num.det, and d_H denote the true positive rate, false negative rate, false positive rate, number of candidate change points, number of change points detected, scaled Hausdorff distance, respectively.

	divisive		e.cp3o_delta				ks.cp3o		ks.cp3o_delta	
alpha	1	2	1	1	2	2				
Kmax			7	14	7	14	7	14	7	14
TP.76	0.02	0.01	0.19	0.4	0.15	0.42	0.55	0.53	0.14	0.24
TP.151	0	0	0.16	0.36	0.13	0.21	0.43	0.46	0.2	0.32
TP.226	0	0	0.08	0.39	0.11	0.29	0.23	0.39	0.11	0.18
TP.301	0.01	0.01	0.14	0.47	0.17	0.36	0.09	0.43	0.14	0.36
TP.376	0.01	0.01	0.25	0.49	0.26	0.36	0.06	0.32	0.13	0.34
TP.451	0.01	0.01	0.14	0.44	0.17	0.28	0.04	0.17	0.19	0.25
TP.526	0	0	0.12	0.39	0.09	0.41	0.08	0.18	0.22	0.33
FN.76	0	0	0.56	0.01	0.48	0.05	0.25	0.23	0.13	0.4
FN.151	0	0	0.38	0.62	0.27	0.57	0.23	0.29	0.16	0.26
FN.226	0	0	0.36	0.55	0.28	0.62	0.39	0.34	0.09	0.43
FN.301	0	0	0.34	0.48	0.26	0.52	0.19	0.34	0.18	0.31
FN.376	0	0	0.32	0.48	0.18	0.52	0.13	0.49	0.15	0.3
FN.451	0	0	0.39	0.56	0.25	0.5	0.07	0.48	0.07	0.37
FN.526	0	0	0.63	0.21	0.54	0.21	0.04	0.34	0.11	0.41
FP.0-75	0.01	0.02	0	0.6	0	0.42	0.66	0.62	0.21	0.53
FP.76-150	0.01	0	0.38	0.81	0.44	0.83	0.75	0.75	0.39	0.64
FP.151-225	0.01	0.01	0.64	0.9	0.62	0.88	0.49	0.82	0.4	0.56
FP.226-300	0.01	0.01	0.42	0.89	0.43	0.84	0.18	0.73	0.32	0.63
FP.301-375	0.01	0.01	0.32	0.87	0.29	0.81	0.09	0.56	0.32	0.62
FP.376-450	0.01	0.01	0.57	0.85	0.5	0.86	0.1	0.39	0.36	0.64
FP.451-525	0	0	0.43	0.83	0.46	0.78	0.08	0.31	0.44	0.6
FP.526-600	0.01	0.01	0.32	0.87	0.29	0.81	0.09	0.56	0.32	0.62
Cand.num	1.11	1.1	9.75	20.14	9.32	19.58	8.93	16.75	7.18	14.76
Num.det	0.11	0.1	3.77	9.34	3.78	8.34	4.43	8.38	3.84	7.17
d_H	6.91	6.92	1.13	0.44	1.15	0.49	4.30	2.30	2.18	1.06
time	2.59	2.55	0.23	0.25	0.21	0.23	4.87	5.24	0.10	0.12

Table 9: The results from applying 5 non-parametric methods in the **ecp** R package to the VAR data with 4 change points (Simulation 5) over 100 simulated data sequences. TP, FN, FP, Cand.num, Num.det, and d_H denote the true positive rate, false negative rate, false positive rate, number of candidate change points, number of change points detected, scaled Hausdorff distance, respectively.

	divisive		e.cp3o				e.cp3o_delta				ks.cp3o		ks.cp3o_delta	
alpha	1	2	1	1	2	2	1	1	2	2				
Kmax			4	8	4	8	4	8	4	8	4	8	4	8
TP.61	0.12	0.09	0.31	0.41	0.3	0.39	0.18	1	0.14	1	0.37	0.57	0.19	0.39
TP.121	0.26	0.24	0.31	0.63	0.34	0.62	0.1	0.98	0.05	0.92	0.2	0.51	0.15	0.46
TP.181	0.28	0.26	0.37	0.62	0.38	0.6	0.12	0.99	0.1	0.87	0.11	0.5	0.22	0.5
TP.241	0.1	0.08	0.08	0.18	0.09	0.14	0.19	1	0.22	1	0.07	0.4	0.19	0.44
FN.61	0	0	0.18	0.42	0.21	0.45	0.81	0.61	0.58	0.62	0.25	0.5	0.14	0.53
FN.121	0.01	0	0.31	0.25	0.34	0.33	0.68	0.6	0.39	0.52	0.28	0.4	0.14	0.35
FN.181	0.02	0.01	0.13	0.3	0.16	0.37	0.68	0.62	0.44	0.57	0.23	0.43	0.16	0.29
FN.241	0.02	0	0.05	0.53	0.06	0.54	0.88	0.61	0.71	0.56	0.13	0.5	0.14	0.4
FP.0-60	0.47	0.45	0.39	0.55	0.35	0.53	0	0	0	0	0.45	0.7	0.44	0.74
FP.61-120	0.43	0.37	0.29	0.57	0.26	0.56	0.76	1	0.8	1	0.5	0.72	0.32	0.63
FP.121-180	0.69	0.62	0.47	0.82	0.5	0.8	0.61	1	0.68	1	0.3	0.73	0.24	0.71
FP.181-240	0.11	0.13	0.09	0.31	0.11	0.28	0.75	1	0.69	0.99	0.34	0.76	0.5	0.83
FP.241-300	0.25	0.2	0.12	0.24	0.14	0.3	0	0	0	0	0.16	0.47	0.21	0.52
Cand.num	3.77	3.55	4.72	8.9	4.73	8.99	6.31	14.8	5.99	14.09	5.09	10.64	4.07	9.77
Num.det	2.73	2.47	2.45	4.37	2.5	4.29	2.6	6.98	2.62	6.84	2.48	5.38	2.4	5.22
d_H	1.69	1.88	2.01	1.18	1.96	1.17	0.69	0.47	0.70	0.46	2.01	0.81	1.51	0.58
Time	1.65	1.37	429.68	468.26	317.88	375.62	0.07	0.06	0.05	0.05	0.25	0.25	0.01	0.02

Table 10: The results from applying 5 non-parametric methods in the **ecp** R package to the VAR data with 4 weak change points (Simulation 6) over 100 simulated data sequences. TP, FN, FP, Cand.num, Num.det, and d_H denote the true positive rate, false negative rate, false positive rate, number of candidate change points, number of change points detected, scaled Hausdorff distance, respectively.

	divisive		e.cp3o				e.cp3o_delta				ks.cp3o		ks.cp3o_delta	
alpha	1	2	1	1	2	2	1	1	2	2				
Kmax			4	8	4	8	4	8	4	8	4	8	4	8
TP.61	0.54	0.50	0.57	0.70	0.56	0.71	0.24	1.00	0.33	1.00	0.54	0.70	0.64	0.82
TP.121	0.45	0.5	0.02	0.62	0.05	0.64	0.12	0.95	0.24	0.86	0.19	0.58	0.05	0.52
TP.181	0.11	0.11	0.01	0.08	0.01	0.1	0.16	0.92	0.2	0.75	0.09	0.21	0.07	0.15
TP.241	0.44	0.43	0.18	0.38	0.18	0.32	0.13	1	0.07	1	0.11	0.12	0.12	0.34
FN.61	0	0	0.33	0.18	0.35	0.17	0.72	0.66	0.6	0.64	0.37	0.28	0.2	0.07
FN.121	0.02	0.01	0.51	0.31	0.52	0.27	0.85	0.62	0.66	0.46	0.51	0.45	0.23	0.32
FN.181	0	0	0.03	0.44	0.04	0.36	0.81	0.64	0.72	0.53	0.09	0.51	0.04	0.52
FN.241	0	0.01	0.04	0.51	0.03	0.5	0.9	0.63	0.88	0.62	0.05	0.72	0.08	0.57
FP.0-60	0.67	0.69	0.67	0.88	0.67	0.85	0	0	0	0.02	0.62	0.82	0.65	0.82
FP.61-120	0.66	0.69	0.45	0.9	0.48	0.92	0.75	1	0.63	1	0.34	0.91	0.72	0.94
FP.121-180	0.4	0.42	0.04	0.41	0.02	0.44	0.41	1	0.4	0.99	0.09	0.61	0.1	0.43
FP.181-240	0.14	0.18	0.11	0.19	0.1	0.2	0.75	1	0.7	0.99	0.16	0.3	0.12	0.28
FP.241-300	0.48	0.47	0.34	0.55	0.35	0.55	0	0	0	0	0.29	0.43	0.23	0.66
Cand.num	4.91	5.02	4.89	8.69	4.88	8.5	6.11	14.34	5.79	12.86	5.12	9.53	4.14	8.6
Num.det	3.99	4.08	2.41	4.94	2.4	4.91	2.48	6.91	2.49	6.7	2.35	4.82	2.69	5.03
d_H	1.77	1.73	2.54	1.31	2.55	2.55	0.70	0.47	0.75	0.75	2.36	1.20	1.98	0.88
Time	1.93	1.95	476.90	522.94	385.06	431.19	0.07	0.07	0.05	0.06	0.24	0.27	0.02	0.02

Table 11: The results from applying 4 non-parametric methods in the **ecp** R package to the VAR data with 7 small change points (Simulation 7) over 100 simulated data sequences. TP, FN, FP, Cand.num, Num.det, and d_H denote the true positive rate, false negative rate, false positive rate, number of candidate change points, number of change points detected, scaled Hausdorff distance, respectively.

	divisive		e.cp3o_delta				ks.cp3o		ks.cp3o_delta	
alpha	1	2	1	1	2	2				
Kmax			7	14	7	14	7	14	7	14
TP.76	0.54	0.5	0.07	0.45	0.07	0.4	0.52	0.42	0.07	0.3
TP.151	0.13	0.12	0.12	0.22	0.08	0.17	0.18	0.27	0.04	0.1
TP.226	0.09	0.11	0.11	0.18	0.06	0.13	0.1	0.22	0.03	0.08
TP.301	0.54	0.56	0.08	0.26	0.04	0.32	0.06	0.23	0.04	0.26
TP.376	0.54	0.51	0.23	0.4	0.23	0.44	0.16	0.45	0.5	0.64
TP.451	0.46	0.36	0.11	0.28	0.05	0.29	0.09	0.41	0.04	0.28
TP.526	0.49	0.46	0.15	0.47	0.14	0.4	0.06	0.14	0.09	0.22
FN.76	0	0	0.54	0.1	0.4	0.17	0.06	0.34	0.24	0.38
FN.151	0	0	0.22	0.6	0.11	0.5	0.13	0.29	0.07	0.47
FN.226	0	0	0.12	0.55	0.08	0.52	0.09	0.34	0.04	0.41
FN.301	0	0	0.09	0.62	0.06	0.62	0.14	0.36	0.2	0.4
FN.376	0	0.01	0.13	0.49	0.12	0.33	0.33	0.35	0.12	0.1
FN.451	0	0	0.2	0.52	0.18	0.52	0.34	0.44	0.23	0.42
FN.526	0	0	0.36	0.19	0.34	0.26	0.17	0.33	0.11	0.48
FP.0-75	0.82	0.83	0.03	0.32	0.04	0.25	0.65	0.58	0.76	0.96
FP.76-150	0.22	0.22	0.3	0.69	0.29	0.65	0.38	0.55	0.05	0.14
FP.151-225	0.66	0.53	0.69	0.84	0.73	0.82	0.22	0.49	0.23	0.47
FP.226-300	0.48	0.46	0.26	0.84	0.28	0.81	0.12	0.31	0.05	0.16
FP.301-375	0.9	0.87	0.48	0.89	0.61	0.91	0.34	0.68	0.83	0.97
FP.376-450	0.89	0.89	0.51	0.82	0.56	0.87	0.31	0.67	0.84	0.95
FP.451-525	0.71	0.7	0.44	0.81	0.58	0.83	0.1	0.44	0.01	0.13
FP.526-600	0.9	0.87	0.48	0.89	0.61	0.91	0.34	0.68	0.83	0.97
Cand.num	10.64	10.14	8.94	19.04	8.75	18.44	8.48	16.4	7.07	14.69
Num.det	9.59	9.09	3.6	7.93	3.84	7.83	3.98	7.72	4.4	7.22
d_H	1.25	1.42	1.26	0.53	1.30	0.60	4.12	2.58	1.77	1.06
time	14.17	13.14	0.23	0.24	0.20	0.21	4.76	4.81	0.10	0.11

References

Cribben, I., R. Haraldsdottir, L. Atlas, T. Wager, and M. Lindquist (2012, 03). Dynamic connectivity regression: Determining state-related changes in brain connectivity. NeuroImage 61, 907–20.

Cribben, I., T. Wager, and M. Lindquist (2013). Detecting functional connectivity change points for single-subject fmri data. Frontiers in Computational Neuroscience 7, 143.

Wager, T., C. Waugh, M. Lindquist, D. Noll, B. Fredrickson, and S. Taylor (2009, 06). Brain mediators of cardiovascular responses to social threat: Part i: Reciprocal dorsal and ventral sub-regions of the medial prefrontal cortex and heart-rate reactivity. NeuroImage 47, 821–35.