12-14-2017

Bayesian Methods in Brain Connectivity Change Point Detection with EEG Data and Genetic Algorithm

Bing Liu

BAYESIAN METHODS IN BRAIN CONNECTIVITY CHANGE POINT DETECTION WITH EEG DATA AND GENETIC ALGORITHM

by

BING LIU

Under the Direction of Jing Zhang, PhD

ABSTRACT

Human brain is processing a great amount of information everyday, and our brain regions are organized optimally for this information processing. There have been increasing number of studies focusing on functional or effective connectivity in human brain regions in the last decade. In this dissertation, Bayesian methods in Brain connectivity change point detection are discussed. First, a review of state-of-the-art Bayesian-inference-based methods applied to functional magnetic resonance imaging (fMRI) data is carried out, three methods are reviewed and compared. Second, the Bayesian connectivity change point model is extended to change point analysis in electroencephalogram (EEG) data, and the ability of EEG measures of frontal and temporo-parietal activity during mindfulness therapy to track response to dysfunctional anxiety patients' treatment is tested successfully. Then an

optimized method for Bayesian connectivity change point model with genetic algorithm (GA) is proposed and proved to be more efficient in change point detection. And due to the good parallel performance of GA, the change point detection method can be parallelized in GPU or multi-processor computers as a future work. Furthermore, a more advanced Bayesian bi-cluster connectivity change point model is developed to simultaneously detect change point of each subject within a group, and cluster subjects into different groups according to their change point distribution and connectivity dynamics. The method is also validated on experimental datasets. After discussing brain change point detection, a review of Bayesian analysis of complex mutations in HBV HCV and HIV studies is also included as part of my Ph.D. work. Finally, conclusions are drawn and future work is discussed.

INDEX WORDS: Bayesian inference, Brain network, Change point, Connectivity dynamics, Markov chain Monte Carlo, Electroencephalography, Genetic algorithm

BAYESIAN METHODS IN BRAIN CONNECTIVITY CHANGE POINT DETECTION WITH EEG DATA AND GENETIC ALGORITHM

by

BING LIU

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

in the College of Arts and Sciences

Georgia State University

2017

BAYESIAN METHODS IN BRAIN CONNECTIVITY CHANGE POINT DETECTION WITH EEG DATA AND GENETIC ALGORITHM

by

BING LIU

Committee Chair: Jing Zhang

Committee: Yi Pan

Xin Qi

Ruiyan Luo

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

December 2017

DEDICATION

This dissertation is dedicated to my parents Jingdong Liu and Yufang Liu.

ACKNOWLEDGEMENTS

This dissertation work would not have been possible without the support of many people. First, I want to express my gratitude to my advisor Dr. Jing Zhang, who gave me lots of suggestions and guidance while I was doing this dissertation research. Her insights and knowledge always inspired me and I truly learned a lot from her. I'd like to give my special thanks to Dr. Yi Pan for sharing his valuable knowledge and all the help in our collaboration. My special thanks to Dr. Xin Qi and Dr. Ruiyan Luo, for reviewing my dissertation and providing helpful advices.

Special thanks to Dr. Jessica Turner and Dr. Matthew Turner for their generous help in our EEG study and providing the EEG data for my dissertation research.

Special thanks to Dr. Xuan Guo, Dr. Xiuchun Xiao, Xueli Xiao, and Le Chen for their valuable contributions in our collaboration.

My gratitude also goes to our Chair Dr. Guantao Chen, who offered me advice and help. To Dr. Gengsheng Qin, who is always patient and gives me courage. To Dr. Yichuan Zhao, who organized wonderful statistics colloquium which provide us the opportunities to connect academia and industry. And to my undergraduate advisor, Dr. Danhui Yi, who led me into the field of statistics and gave me opportunities to practice while learning.

Special thanks to all the faculty, staff and students in the Department of Mathematics and Statistics at Georgia State University.

Special thanks to Brain and Behavior Seed Grant from Georgia State University.

Furthermore, I would like to thank my all friends who helped me when I faced difficulties.

Last but not least, I want to thank my dear parents Jingdong Liu and Yufang Liu, my girlfriend Yuyin Shi and my relatives who always support me once I make my decision to do something. I cannot achieve anything without your unconditional support. I love you all!

Thank you all who helped me and let's make our outlook brighter!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

- BBCCPM - Bayesian Bi-cluster Connectivity Change Point Model
- BCCPM - Bayesian Connectivity Change Point detection Model
- BMCPM - Bayesian Magnitude Change Point Model
- BVP - Bayesian Variable Partition
- CI - Confidence Interval
- DBVPM - Dynamic Bayesian Variable Partition Model
- EEG - Electroencephalography
- fMRI - Functional magnetic resonance imaging
- GA - Genetic algorithm
- GSU - Georgia State University
- HBV - Hepatitis B virus
- HCV - Hepatitis C virus
- HIV - human immunodeficiency virus
- i.i.d. - independent identically distributed
- MCMC - Markov chain Monte Carlo
- M-H - Metropolis-Hastings
- PTSD - Post-traumatic Stress Disorder
- RMS - Recursive Model Selection
- ROI - Region of interest

CHAPTER 1

INTRODUCTION

1.1 Brain Connectivity

Our brain processes a large number of information everyday, and most of the information is processed by different but related brain regions. The paradigm in neuroscience is that the anatomical and functional connections between our brain regions are organized optimally for information processing [39]. Brain connectivity often refers to three patterns of links, dependencies or interactions: anatomical connectivity, functional connectivity, effective connectivity [79].

- Anatomical connectivity, also notes as anatomical links or structure connectivity, which is in a form of a network of structural connections linking neurons or elements of neuronal systems. At shorter time scales (seconds to minutes), the anatomical connections are relatively stable, while at longer time scales (hours to days), it is likely to observe some plasticity.

- Functional connectivity, often called statistical dependencies, on the other hand, is based on statistical concepts. It captures the dependency between distributed and remote neuronal units. Theses statistical dependencies may be inferred by estimating the correlation or covariance, spectral coherence or phase-locking, and they are highly time dependent.

- Effective connectivity, also called causal interactions, describes the effect from one neural on another directionally. It can be considered as a combination of structural and functional connectivities. Oftentimes, these casual effects can be analyzed by time series analysis.

There have been increasing number of studies focusing on functional or effective connectivity in the last decade, mostly from functional neuroimaging techniques like functional magnetic resonance imaging (fMRI), electroencephalogram (EEG), magnetoencephalogram (MEG), positron emission tomography (PET), and so on. The predominant technique, however, studies functional magnetic resonance imaging data and the blood-oxygen-level-dependent (BOLD) signal [39].

1.2 Brain Connectivity Change Points Detection and Dynamics Exploration

Recently in the research area of functional magnetic resonance imaging (fMRI), an increasing interest comes in the dynamic connectivity of brain regions which communicate one from another. Lindquist *et al.* [48] proposed a statistical method, named as Hierarchical Exponentially Weighted Moving Average (HEWMA), on fMRI data to detect the state change of BOLD signals in response to stimulus. Cribben *et al.* [12] developed a Dynamic Connectivity Regression method, which is a data-driven method for detecting functional connectivity change points. Independent component analysis (ICA) method, including dynamic spatial ICA, was developed by Sakoğlu *et al.* [75] to investigate the connectivity dynamics. Also, sliding-time-window-based approaches were designed to capture the dynamics of brain functional interactions across different time windows [41][103][67][53][102][1]. For example, Allen *et al.* [1] described an approach to assess whole-brain functional connectivity (FC) dynamics based on spatial ICA, sliding-time-window correlation analysis, and k-means clustering of windowed correlation matrices, and it revealed unanticipated FC states which were strongly different from stationary connectivity patterns. Li *et al.* [41] derived functional connectomes (FCs) to characterize brain conditions from resting-state fMRI data and then FCs were divided into quasi-stable segments temporally via a sliding-time-window approach. In [103], a novel framework was designed by the combination of sliding-window approach and multiview spectral clustering to extract temporally dynamic functional connectome patterns for resting-state networks, and the four detected clusters are believed to play critical roles in functional brain dynamics during resting states. In [67], a novel algorithmic framework based

on hidden Markov models was presented to cluster and label the brains functional states, represented by a large-scale functional connectivity matrix and derived via an overlapping sliding-time-window approach. The framework achieved decent classification performance on the data including 25 ADHD (attention-deficit/hyperactivity disorder) patients and 49 normal controls. Lv *et al.* [53] also adopted the sliding-window-based method and employed a dynamic programming strategy to infer functional information transition routines on structural networks and identified the hub routers that participate in these routines most frequently.

Inspired by the multivariate graphical models based on Bayesian networks, which has been shown to be robust and reliable in estimating functional interactions and less sensitive to noise in the fMRI signals [73][81][77], recently, several Bayesian-inference-based methods were proposed to infer global functional interactions within brain networks and their temporal transition boundaries [46][43][101][64][65][45]. By evaluating and estimating both simulated and real data, Bayesian-inference-based methods are proved to be more powerful approaches for analysis fMRI data comparing to the methods mentioned above.

1.3 Main Contributions

In this dissertation, the main contributions are summarized as follows:

- A review of state-of-the-art Bayesian-inference-based methods applied to functional magnetic resonance imaging (fMRI) data:
  - Bayesian Magnitude Change Point Model;
  - Bayesian Functional Connectivity Change Point Model;
  - Dynamic Bayesian Variable Partition Model;
- An extension of Bayesian connectivity change point model to EEG data change point analysis:
  - detect brain connectivity change points from EEG measurement signals;

  - design multiple experiments to validate the model on EEG data;
  - apply the model to EEG data from dysfunctional anxiety patients;
  - test the ability of EEG measures of frontal and temporo-parietal activity during mindfulness therapy to track response to treatment.

- An optimized method for Bayesian connectivity change point model with genetic algorithm:
  - combine Bayesian connectivity change point model with a modified genetic algorithm;
  - optimize the evolutionary procedure to improve the detection accuracy;
  - decrease the time consumption/computational cost;
  - validate the optimized method on simulated and real data.

- A Bayesian bi-cluster connectivity change point model:
  - detect change point of each subject within a group, and simultaneously
  - cluster subjects into different groups according to their change point distribution with bayesian framework;
  - develop a two level MCMC scheme to sample from the posterior distribution;
  - validate the proposed method on experimental datasets;.

- Other topics: a review of Bayesian analysis of complex mutations in HBV HCV and HIV studies:
  - review the Bayesian-inference-based methods applied to Hepatitis B virus (HBV), Hepatitis C virus (HCV) and the human immunodeficiency virus (HIV) studies with a focus on the detection of the viral mutations and various problems which are correlated to these mutations;

- summarize the Bayesian Variable Partition (BVP) model, and the Recursive Model Selection (RMS) procedure, which are designed to detect the mutations and to further infer the detailed dependence structure among the interactions;
- summarize of the Bayesian methods applications toward these viruses studies, where several important and useful results have been discovered;

1.4 Structure of the Rest of the Dissertation

The remainder of the dissertation focuses on the brain connectivity change point detection methods on both fMRI and EEG data. Chapter 2 reviews the current novel Bayesian statistical inference methods for change point detection in fMRI data. Chapter 3 extends one of the Bayesian change point detection methods from fMRI data to EEG data and focuses on a preliminary research in testing the ability of EEG measures of frontal and temporo-parietal activity during mindfulness therapy to track response to treatment, as an evaluation of EEG as a physiological aid in therapy. Chapter 4 introduces an optimized method for Bayesian Connectivity Change Point Model with genetic algorithm, which improves the estimate efficiency in detecting brain connectivity change points. Chapter 5 presents a more advanced bi-cluster Bayesian inference model to simultaneously detect change points and cluster subjects/patients according to their connectivity change point patterns, which potentially can be used to classify healthy subjects and diseased patients. After discussing the change point analysis, some other Bayesian methods applied in the fields of HBV, HCV and HIV researches are also discussed and reviewed in Chapter 6 as part of this dissertation. Finally, Chapter 7 summarizes the conclusions and discusses future work.

CHAPTER 2

A REVIEW OF BAYESIAN INFERENCE FOR FUNCTIONAL DYNAMICS EXPLORING IN FMRI DATA

The material in Sections 2.1 through 2.7 of this chapter draws heavily on the author's previously published review article [31].

2.1 Bayesian Inference

Bayesian inference is a statistical inference method in which Bayes' Rule is used to update the estimated probability for a hypothesis as evidence is acquired, and it is a novel method for combining prior knowledge with observed information to answer the questions that researchers are usually interested in, like "what is the probability of getting lung cancer for certain patients who does not smoke?". It is a natural way to combine multiple experiments information and fit realistic but complicated models. On the other hand, Bayesian inference often has more computational cost and requires at least one of elicitation of real subjective probability distributions of prior beliefs. The good thing is that sensitivity analysis shows that the choice of prior does not strongly affect the inference.

Bayes' Rule [90]

If θ is a discrete random variable with a probability mass function (p.m.f.), then

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{\sum_i p(y|\theta_i)p(\theta_i))} \tag{2.1}$$

When θ is continuous, Bayes' Rule has the following form

$$p(\theta|y) = \frac{p(\theta, y)}{p(y)} = \frac{p(y|\theta)p(\theta)}{p(y)} = \frac{p(y|\theta)p(\theta)}{\int_i p(y|\theta)p(\theta))d\theta} \tag{2.2}$$

Bayes's Rule is often written as $p(\theta|y) \propto p(\theta)p(y|\theta)$, when treated as a function of θ

for a fixed y, and $p(y|\theta)$ is the likelihood. So Bayes' Rule is considered as the posterior is proportional to the product of prior and likelihood,

$$\begin{aligned} Posterior &\propto Prior \times Likelihood \\ p(\theta|y) &\propto p(\theta) \times p(y|\theta) \end{aligned} \tag{2.3}$$

Basics of the Bayesian Inference [90]

When we use Bayesian statistics to make inferences, consider

1. Setting up a probability model;

2. Applying the probability theory and the Bayes' Rule.

For example, let $(X_1, ..., X_n)$ be an independent and identically distributed sample from Binomial distribution $Bin(n, \pi)$, where n is the sample size and π is the probability of success. We have $x|\pi \sim Bin(n, \pi)$. And the likelihood can be written as

$$p(x|\pi) = \binom{n}{x} \pi^x (1-\pi)^{n-x}; \pi \in [0, 1] \tag{2.4}$$

If we want to make inference on π given x and n, a prior distribution $p(\pi)$ for π is needed, for example, we can choose a uniform distribution such that, $\pi \sim U(0, 1)$:

$$p(\pi) = \begin{cases} 1, & \text{if } 0 \leq \pi \leq 1 \\ 0, & \text{otherwise} \end{cases} \tag{2.5}$$

Then by applying the Bayes' Rule, we get

$$\begin{aligned} p(x, \pi) &= \binom{n}{x} \pi^x (1-\pi)^{n-x} \\ p(x) &= \int_0^1 \binom{n}{x} \pi^x (1-\pi)^{n-x} d\pi = \frac{1}{n+1} \\ p(\pi|x) &= \frac{p(x, \pi)}{x} = (n+1) \binom{n}{x} \pi^x (1-\pi)^{n-x} \end{aligned} \tag{2.6}$$

Conjugate Prior [90]

A prior probability distribution is said to be conjugate to the sampling density if the resulting posterior distribution is a member of the same parametric family as the prior. For example, binomial likelihood × beta prior = beta posterior.

Bayesian Analysis on fMRI Data

An fMRI data series consists of values at specific voxel of the image at some time point t. These series are collected as a T-dimensional vector $\vec{y} = (y_1, y_2, ..., y_T)$. Suppose we record the data based on m regions of interest (ROIs), then the dataset is an $m \times T$ matrix Y. An example can be found in Figure 2.1 [31].

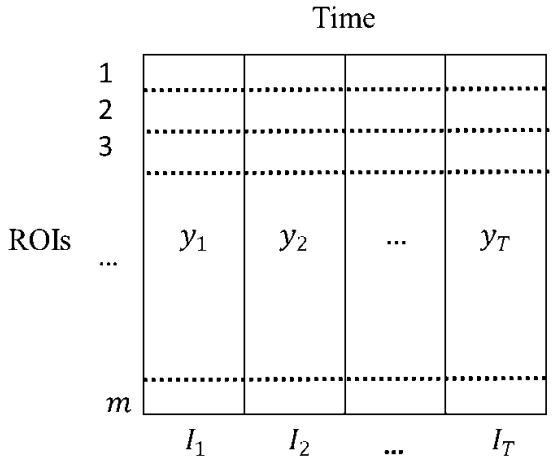

Figure 2.1. Illustration of data matrix Y with m ROIs and T time points, and a block indicator vector $\vec{I}$, where y_i is the values of all m ROIs at the time point i.

Our purpose is to use Bayesian analysis to make inference on the dynamic functional connectivity change points. Define a block indicator $\vec{I} = (I_1, I_2, ..., I_T)$ to indicate the locations of change points in the $m \times T$ dataset. And this indicator becomes our parameter of interest. Follow the Bayesian framework, we assume the prior of $\vec{I}$ is Bernoulli(0.5) so that $P(\vec{I}) = \prod_{t=1}^{T} p(I_i)$. If $\vec{I}$ is given, we can calculate the likelihood of the data as $p(Y|\vec{I})$. Thus,

the posterior distribution of $p(\vec{I}|Y)$ can be obtained by

$$p(\vec{I}|Y) \sim p(\vec{I})p(Y|\vec{I}) \tag{2.7}$$

Then a Markov chain Monte Carlo (MCMC) scheme can be used to sample the posterior with a random initial block indicator.

Following these procedures, several models have been established and applied in change point detection research. In the following sections in this chapter, three Bayesian models are reviewed: Bayesian magnitude change point model, Bayesian connectivity change point model, and Bayesian variable partition model for detecting functional interaction and transition patterns.

2.2 Bayesian Magnitude Change Point Model

Lian et al. proposed a Bayesian Magnitude Change Point Model (BMCPM) to detect group-wise consistent magnitude change points [46]. A key feature of BMCPM is the capability to consider the group-wise fMRI signals of corresponding cortical landmarks across a population of subjects and optimally determines the change boundaries. Magnitude change points are defined as the temporal points dividing ROI data matrix into blocks which exhibit substantial differences in brain states from each other. Figure 2.2 [31] demonstrates the basic idea in BMCPM. One temporal change point located at time point T_{100} partitions the ROI data matrix into two time blocks with different distributions.

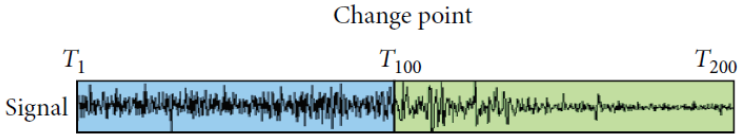

Figure 2.2. An ROI signal with one magnitude change point at time point T_{100}.

Given a vector $\vec{a} = (a_1, a_2, ...a_t)$ i.i.d. from a normal distribution with mean μ and variance σ^2, $a \sim N(\mu, \sigma^2)$, t is the dimension of vector $\vec{a}$. We use a conjugate prior Normal-

Inverse-Chi-Square (N-Inv-χ^2) [28] for unknown mean and variance, then the posterior distribution is in the same family as the prior probability distribution. Thus, by assuming a conjugate prior N-Inv-$\chi^2(\mu_0, \sigma_0^2/\kappa_0, \nu_0, \sigma_0^2)$ for (μ_0, σ_0^2), the posterior distribution of (μ, σ^2) will be N-Inv-$\chi^2(\mu_t, \sigma_t^2/\kappa_t, \nu_t, \sigma_t^2)$. Then the probability of $a_1, a_2, ...a_t$ can be calculated as

$$p(a_1, a_2, ...a_t) = \left(\frac{1}{2\pi}\right)^{t/2} \sqrt{\frac{\kappa_0}{\kappa_t}} \frac{\Gamma(\nu_t/2)}{\Gamma(\nu_0/2)} \frac{(\nu_0\sigma_0^2/2)^{\nu_0/2}}{(\nu_t\sigma_t^2/2)^{\nu_t/2}} \tag{2.8}$$

Based on (2.8), given a data matrix $A = (\vec{a}_1, \vec{a}_2, ..., \vec{a}_m)$, where $\vec{a}_i$ is a vector with the data i.i.d. from the normal distribution as described before, and $\vec{a}_i$'s are independent. Then, the probability of A can be calculated as

$$p(A) = \prod_{i=1}^{m} p(\vec{a}_i) \tag{2.9}$$

where $p(\vec{a}_i)$ is computed as in (2.8)

To make inferences on the magnitude temporal change points, we define a block indicator vector $\vec{I} = (I_1, I_2, ..., I_T)$, $I_i = 1$ if the i^{th} observation is the beginning point of a block, and $I_i = 0$ otherwise. Note that $I_1 = 1$ as the first location is always being considered as a change point. Let Y_i denote the data of m ROIs at the i^{th} location. Then for data matrix Y defined as in previous section, the likelihood given a block indicator vector can be calculated as

$$p(Y|\vec{I}) = \prod_{i=1}^{\sum I_j} p(Y_i) \tag{2.10}$$

where $P(Y_i)$ can be calculated using (2.9). Then the posterior of $\vec{I}$ can be obtained by

$$p(\vec{I}|Y) \propto p(\vec{I})p(Y|\vec{I}) \tag{2.11}$$

A Bernoulli distribution with parameter θ which could be adjusted to reflect our knowledge about the prior distribution of I is used, i.e., $p(\vec{I}) \propto \theta^{\sum I_j}(1-\theta)^{(T-\sum I_j)}$

2.3 Bayesian Functional Connectivity Change Point Model

In order to analyze the joint probabilities among the nodes of brain networks between different time periods, a Bayesian Connectivity Change Point Model (BCCPM) [43] was proposed to determine the temporal boundary where there is an abrupt change of multivariate functional interactions in the brain networks. Different from BMCPM, which considers ROIs independent of each other, BCCPM infers the boundaries of temporal blocks via a unified Bayesian framework by analyzing the dynamics of multivariate functional interactions.

Suppose a vector $(b_1, b_2, ..., b_t)$ i.i.d. from an m-dimensional multivariate normal distribution s.t. $b_i \sim N(\vec{\mu}, \Sigma)$, $i = 1, 2, ..., t$, where t is the number of vectors and m is the dimension of vector b_i, $\vec{\mu}$ denotes the m-dimensional mean vector and Σ denotes the $m \times m$ covariance matrix. For multivariate normal distribution with unknown mean and unknown covariance, we can use a conjugate prior Normal-Inverse-Wishart (N-Inv-Wishart) [28]. Therefore, assuming a conjugate prior N-Inv-Wishart$(\mu_0, \Lambda_0/\kappa_0, \nu_0, \Lambda_0)$ for $(\vec{\mu}, \Sigma)$, the posterior will be N-Inv-Wishart$(\mu_t, \Lambda_t/\kappa_t, \nu_t, \Lambda_t)$.

As we are interested in the posterior distribution of the configuration, the joint probability of $b_1, b_2, ..., b_t$ can be calculated as

$$p(b_1, b_2, ..., b_t) = \left(\frac{1}{2\pi}\right)^{mt/2} \left(\frac{\kappa_0}{\kappa_T}\right)^{m/2} \frac{\Gamma_m(\nu_t/2)}{\Gamma_m(\nu_0/2)} \frac{(det(\Lambda_0))^{\nu_0/2}}{(det(\Lambda_t))^{\nu_t/2}} 2^{mt/2} \tag{2.12}$$

where Γ_m is the multivariate Gamma distribution.

The model aims to detect the connectivity change points which separate the temporal segments where the joint probabilities have underlying differences among the m ROIs between different time periods. Figure 2.3 [31] shows the idea of BCCPM.

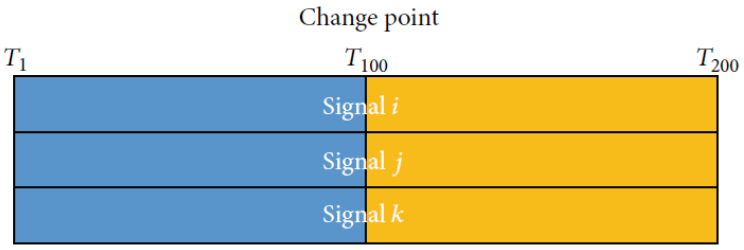

Figure 2.3. Three ROI signals with one connectivity change point at time point T_{100}. The multivariate normal distribution inside the block T_1-T_{100} of color blue is different from the distribution of signal in the rest block of color orange.

A block indicator $\vec{I} = (I_1, I_2, ..., I_T)$ is defined similarly as in BMCPM. The marginal likelihood of the data matrix $Y = (y_1, y_2, ..., y_T)$ can be calculated as

$$p(Y|\vec{I}) = \prod_{i=1}^{\sum I_j} p(Y_i) \tag{2.13}$$

where $p(Y_i)$ can be calculated by (2.12). Similar to BMCPM, the posterior distribution of the configuration $p(\vec{I}|Y)$ can be calculated by (2.11). Note that one important assumption here is that the temporal segments (blocks) indicated by $\vec{I}$ are mutually independent to each other across T.

2.4 Bayesian Change Point Model Using One-Level MCMC Scheme

In BMCPM and BCCPM, Metropolis-Hastings (MH) scheme is used for calculating the Bayesian inference. A one-level MH (MCMC) scheme is provided as follows with a randomly initialized block indicator $\vec{I}^0$ and a user-defined iteration number N:

1. Generate a new block indicator $\vec{I}^*$ by randomly switching the value of one element from 0 to 1 or from 1 to 0 in $\vec{I}^{n-1}$. Then calculate $p(\vec{I}^*|Y)$ by (2.11);

2. Generate a random number u from Uniform(0,1) and update $\vec{I}^n$ by

$$\vec{I}^n = \begin{cases} \vec{I}^*, & \text{if } u \leq min\left[1, \frac{p(\vec{I}^*|Y)}{p(\vec{I}^{n-1}|Y)}\right] \\ \vec{I}^{n-1}, & \text{otherwise} \end{cases} \tag{2.14}$$

3. Iterate step 1 and 2 until n reaches N;

4. Finally, the posterior probabilities for each time point being a change point can be calculated from the MCMC samples excluding the burn-in samples.

By default, the initial parameters $\mu_0, \kappa_0, \nu_0, \sigma_0^2, \Lambda_0$ are fixed constants. The number of iterations N can be determined by the trace plot of posterior probability or the Gelman and Rubin scale reduction factor [29].

2.5 Dynamic Bayesian Variable Partition Model

Zhang et al. [101] proposed a Dynamic Bayesian Variable Partition Model (DBVPM) to simultaneously infer global functional interactions within brain networks and their temporal transition boundaries. Two dependence structures, chain- and V-dependence structures, are designed to capture all the conditional independence global structure. An example can be found in Figure 2.4 [31], the DBVPM aims to simultaneously infer the temporal change points and the dependence structure inside the temporal block.

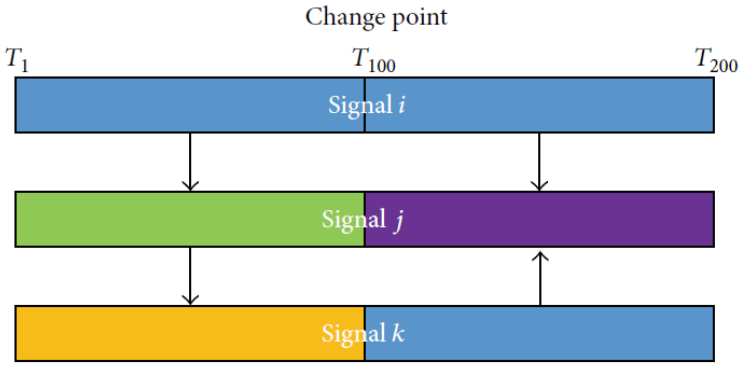

Figure 2.4. Three ROIs signals with one temporal change point at time point T_{100}. Left block shows a chain dependence structure with signals $i \to j \to k$ and the right block shows a V-dependence structure with signals $i \to j \leftarrow k$.

2.5.1 Chain-Dependence Model

A set of variables X_G has a chain-dependence model if the index set G can be grouped into three subgroups U, V, and W such that X_U and X_W are independent given X_V, i.e. $X_U \to X_V \to X_W$. The joint distribution of the chain-dependence model is given as

$$p(X_G) = p(X_U)p(X_V|X_U)p(X_W|X_V) = \frac{F(X_V, X_U)F(X_W, X_V)}{F(X_V)} \tag{2.15}$$

where $F(X_V, X_U, ...)$ is the joint probability function of $(X_V, X_U, ...)$

2.5.2 V-Dependence Model

A set of variables X_G has a V-dependence model if the index set G can be grouped into three subgroups U, V, and W such that X_U and X_W are mutually independent, i.e. $X_U \to X_V \leftarrow X_W$. The joint distribution of the V-dependence model is given as

$$p(X_G) = p(X_U)p(X_W)p(X_V|X_U, X_W) = F(X_U)F(X_W)\frac{F(X_V, X_U, X_W)}{F(X_U, X_W)} \tag{2.16}$$

where $F(X_V, X_U, ...)$ is the joint probability function of $(X_V, X_U, ...)$

2.5.3 Dynamic Bayesian Variable Partition Model

DBVPM uses the similar Bayesian inference as in BCCPM but the number of dimension in the multivariate normal distribution is not fixed among the temporal order. Given $c_1, c_2, ..., c_t$ i.i.d. from r-dimensional multivariate normal distribution $N(\vec{\mu}, \Sigma)$, $i = 1, 2, ..., t$, we only need to change the dimension from m to r from (2.12), and r is determined by the joint variables.

To infer the dependence structure (chain or V structures) in DBVPM, two more indicator vectors are introduced. An indicator vector $\vec{I}^G = (I_1^G, I_2^G, ..., I_m^G)$ is used to denote the grouping of the index G of ROIs to subgroups U, V, and W, where $I_i^G = j$ means i^{th} ROI is in subgroup j ($j = 0$ means U, $j = 1$ means V, and $j = 2$ means W). Another binary indicator I^{CV} is used to denote the dependence structures (chain or V structures). Now the posterior distribution for observations in i^{th} temporal block can be calculated as

$$p(\vec{I}^G, I^{CV}|Y_i) \propto p(Y_i|\vec{I}^G, I^{CV})p(\vec{I}^G)p(I^{CV}) \tag{2.17}$$

where Y_i is the under the same definition as before, $p(Y_i|\vec{I}^G, I^{CV})$ is calculated by (2.15) when $I^{CV} = 0$ and by (2.16) when $I^{CV} = 1$.

Combine with the previous temporal block indicator $\vec{I} = (I_1, I_2, ..., I_T)$, let $\vec{I}^{CV} = (I_1^{CV}, I_2^{CV}, ..., I_{\sum I_i}^{CV})$ be the structure indicator vector and $\vec{I}^G = (\vec{I}_1^G, \vec{I}_2^G, ..., \vec{I}_{\sum I_i}^G)$, where $\vec{I}^G = (I_1^G, I_2^G, ..., I_m^G)$ with $I_i^G = 0, 1, 2$, be the partition indicator vector in the i^{th} block. Then the posterior distribution of data matrix $Y = (y_1, y_2, ..., y_T)$ can be calculated by

$$p(\vec{I}, \vec{I}^G, \vec{I}^{CV}|Y) \propto p(Y|\vec{I}, \vec{I}^G, \vec{I}^{CV})p(\vec{I}, \vec{I}^G, \vec{I}^{CV}) \tag{2.18}$$

where $p(Y|\vec{I}, \vec{I}^G, \vec{I}^{CV}) = \prod p(Y_i|\vec{I}_i^G, I_i^{CV})$ and $p(\vec{I}, \vec{I}^G, \vec{I}^{CV}) = p(\vec{I}) \prod p(\vec{I}_i^G, I_i^{CV}|I_i)$. A uniform prior can be used on $p(\vec{I})$ and $p(\vec{I}_i^G, I_i^{CV}|I_i)$.

2.6 DBVPM Using Two-Level MCMC Scheme

In DBVPM, a two-level MH (MCMC) scheme is applied to sample from the posterior distribution of the temporal blocks and the dependence structure within each block. The lower level MCMC samples from the posterior distribution of dependency structures within a block given the temporal block boundaries; the higher level MCMC samples from the posterior distribution of block boundaries. Specifically, the lower level MCMC involves changing between the chain structure and V structure, and updating the group lables of each variable of ROI. The likelihood functions can be evaluated by using (2.15), (2.16), and (2.18). The higher level MCMC involves dividing one block into two, merging two neighboring blocks into one, and shifting the value in the block indicator. In each iteration in higher level MCMC, each block runs though a lower level MCMC. A dependency structure is sampled for each block in the higher level proposal and the log likelihood of the proposal can be calculated by summing up all the log likelihood probabilities of each block. Details of this two level MCMC and calculation can be found in [101].

2.7 Summary

To summarize the three models we reviewed, the major differences lie in the assumptions of the relationships among ROIs. For BMCPM, no explicit connection is assumed among the ROIs, so it used one-dimensional normal distribution model. For BCCPM, the ROIs are linked together one way or another so multivariate normal distribution is employed. For DBVPM, the chain- and V- dependence structures are used to capture more complex connections among ROIs. These Bayesian inference methods have been applied to solve three problems in the anaylsis of fMRI data: 1) detecting magnitude change points; 2) detecting functional connectivity change points; and 3) detecting change points and identifying functional interaction patterns. The DBVPM can identify not only all possible change points but also the functional interaction patterns, but the convergence speed is much slower comparing with BMCPM and BCCPM as it nests two levels of MCMC scheme. Detailed computational

cost comparison can be found in [31].

CHAPTER 3

BRAIN CONNECTIVITY CHANGE POINT DETECTION ON EEG DATA WITH BCCPM

This research is collaboration work with Dr. Jessica Turner and Dr. Matthew Turner in the Department of Psychology at Georgia State University. The author truly thanks Dr. Jessica Turner for her support with the EEG data provided for this dissertation research (See Appendix B for her Letter of Support). Special thanks to both Dr. Jessica Turner and Dr. Matthew Turner for their generous help.

3.1 Introduction and Contributions

In the previous chapter, we reviewed the state-of-art Bayesian inference methods applied to fMRI data in exploring the brain dynamics. In this chapter, we extend the application of one of the three models, the Bayesian connectivity change point detection model (BCCPM), onto the change point analysis of Electroencephalography (EEG) data to determine network dynamics over time. In particular, the concept of Bayesian inference by using BCCPM is applied to find the change points on EEG data in order to test the ability of EEG measures of frontal and temporo-parietal activity during mindfulness therapy to track response to treatment, as preliminary evaluation for EEG as a physiological aid in therapy [86].

In the EEG literature of the studying anxiety and related disorders, research focuses on the measures of quantitative EEG, which include spectral band amplitude, individual peak frequency, and individualized bandwidths [6]. A more complex approach studies the "microstates", which are very short periods of stability in the EEG signal across different scalp locations [89][36]. These microstates fluctuate in spatial arrangement and duration, and comparing with healthy population, they have been shown to vary in subjects with mood and panic disorders [37]. The dynamic functional interaction patterns also differ in subjects with

psychiatric disorders. In [64], BCCPM was successfully applied in differentiating attention-deficit/hyperactivity disorder (ADHD) children from normal control group on fMRI datasets.

In this research, we are able to apply the network analysis of BCCPM to the EEG signals. BCCPM can identify these brain networks of signals which are "coherently interacting over time" and detect the change points [86]. Although the BCCPM was originally "developed and refined on fMRI data" [86], we successfully developed and validated this model for discovering the brain functional interaction patterns through EEG data. Also, we are able to use the results in EEG data to support that the subjects become less anxious as the therapy sessions continue and that the dynamics change is with response to the treatment.

3.2 Introduction to EEG Data

Electroencephalography (EEG) is an electrophysiological monitoring method to record electrical activity of the brain; it is typically noninvasive, with the electrodes placed along the scalp to record brain activities [91]. EEG measures voltage fluctuations resulting from ionic current within the neurons of the brain [62]. In clinical contexts, EEG refers to recording of the brain's spontaneous electrical activity over a period of time [62], from multiple electrodes placed on scalp.

EEG is "used extensively in neuroscience, cognitive science and cognitive psychology, neurolinguistics, and psychophysiology research" [91]. Research on mental health and mental disabilities like attention deficit hyperactivity disorder (ADHD) is also becoming more and more widely known using EEG [91].

There are some advantages of EEG over the other methods such as fMRI, magnetoencephalography (MEG), positron emission tomography (PET), Single-photon emission computed tomography (SPECT), magnetic resonance spectroscopy (MRS), Near-infrared spectroscopy (NIRS) and Event-related optical signal (EROS), etc. Despite the relatively poor spatial sensitivity of EEG, it has several advantages over the others [91]:

- EEG data recording's hardware costs are significantly lower than most of other techniques [87].

- EEG recording devices can be used in more places than fMRI, MEG, PET, SPECT or MRS. Those techniques need sizable and immobile equipments. For instance, fMRI requires the use of a 1-ton magnet in a magnetically shielded room; MEG requires equipment consisting of liquid helium-cooled detectors that can be used only in shielded rooms, altogether costing upwards of several million dollars [32].

- EEG is relatively tolerant of subject movement, unlike most other neuroimaging techniques. There even exist methods for minimizing, and even eliminating movement artifacts in EEG data [63].

- EEG does not have any sound, which allows for better study of the responses to auditory stimuli [91].

- EEG is not exposed to high-intensity (> 1 tesla) magnetic fields, as in some of the other techniques, especially MRI and MRS. These can cause a variety of undesirable issues with the recorded data, and also prohibit use of these techniques with participants who have metal implants in their body, such as metal-containing pacemakers [76].

- EEG has a "better understanding of what signal is measured as compared to other research techniques, i.e. the BOLD response in MRI" [91].

EEG can detect changes over milliseconds, which is excellent considering an action potential take approximately 0.5-130 milliseconds to propagate across a single neuron, depending on the type of neuron [2]. Other methods like fMRI and PET have time resolution between seconds and minutes [91]. EEG measures the brain activity through electrical activity directly while other methods record changes in blood flow (like fMRI) or metabolic activity (like PET), which are indirect markers for brain activity [91].

3.3 Data Acquisition and Processing

3.3.1 Data Acquisition Device

In this research, the EEG data are recorded using the EPOC/EPOC+ system by Dr. Jessica Turner and Dr. Matthew Turner. The EMOTIV EPOC+ 14 Channel Mobile EEG system is a low-cost research grade EEG system designed for practical contextualized research and advanced brain control interface (BCI) applications [86] [16]. It has 14 channel electrodes and a fixed selection of electrodes placements on the scalp [86]. It provides us the access to the raw EEG signals [16]. "The EPOC/EPOC+ records at 2048 Hz internally, resolution is 14 or 16 bits per channel and the frequency response is 0.16 43 Hz" [86]. The system has an "open-source interface to the raw real-time data signals" [86]. Figure 3.1 [15] shows an example of the EPOC/EPOC+ headset device kit with 14 electrodes that can be placed on scalp for signal recording.

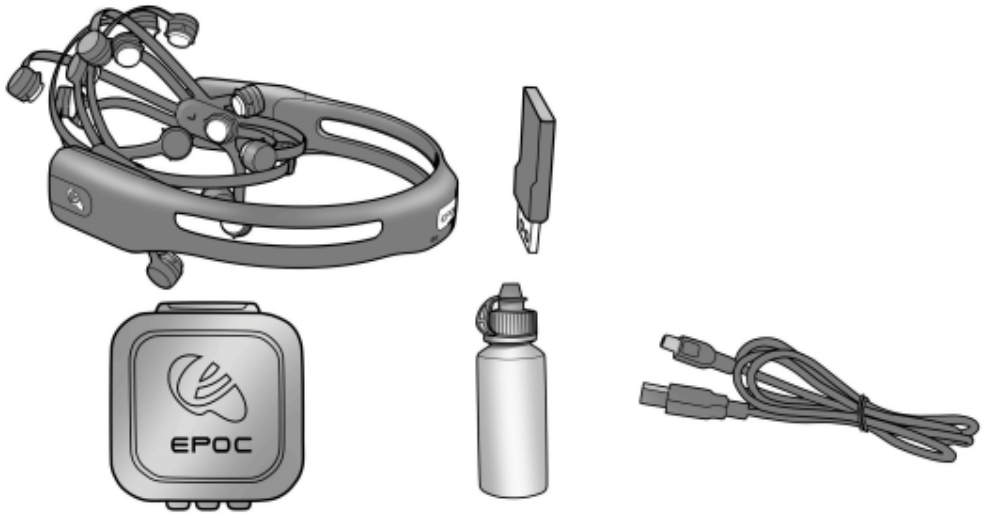

Figure 3.1. EPOC/EPOC+ headset device with 14 electrodes. Image from EMOTIV EPOC User Manual.

The recording system can be used in different environments so the data can be collected in the clinic or a convenient but suitable room [86]. It requires the use of saline instead of gel

on the electrodes, so the starting preparation almost does not cost any time and the recording can start off when it's placed on the scalp (demonstration can be found in Figure 3.2 [15]). Besides the recording headset, it also comes with a USB proprietary wireless receiver that is compatible with major platforms including Windows, OS X, Linux, Android, and iOS [86] [16]. Thanks to the wireless design, the recording subject can be seated anywhere within the range [86]. The company designs the product for use in brain training and video games and researchers have used the device in neurofeedback and raw EEG signal recording outside of standard laboratories [4] [14] [85].

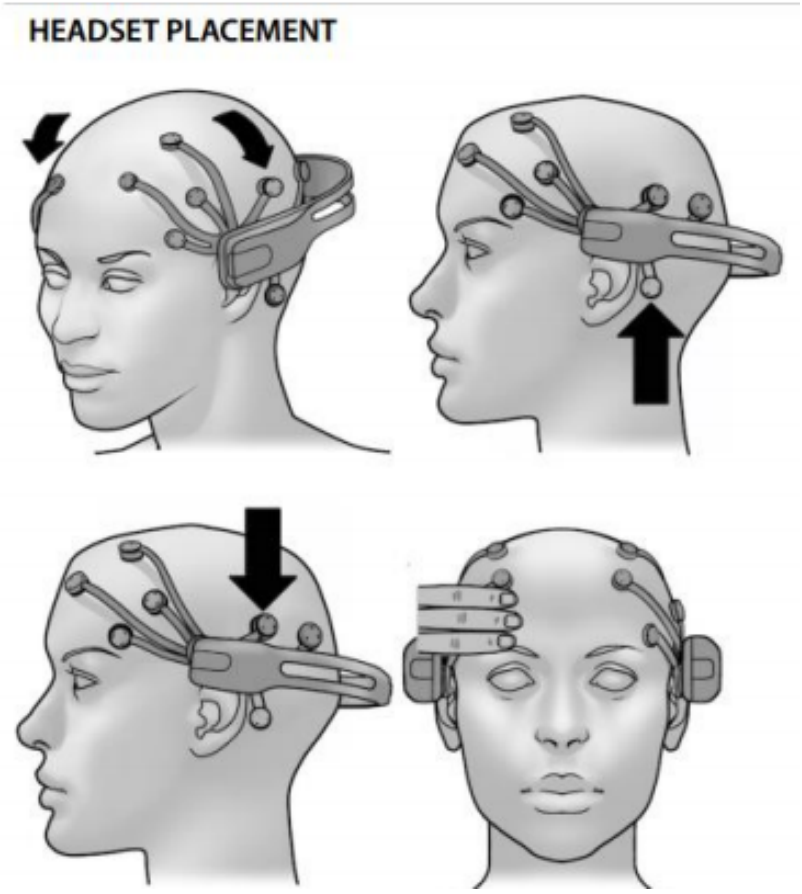

Figure 3.2. EPOC/EPOC+ headset placement demonstration. Image from EMOTIV EPOC User Manual.

3.3.2 Data Preprocessing

The collected raw EEG data is pre-processed by using EEGLAB, an interactive Matlab toolbox for processing continuous and event-related EEG data [84]. After the pre-processing, the EEG dataset will become a matrix with size $n \times T$, where n is the number of electrodes (in our EEG recording, $n = 14$), and T is the total number of time points in the recording. The dataset is saved in text file.

3.4 Experimental Results

In this section, four settings of experiments are designed and EEG data are collected in different sessions separately. The first design has 3 minutes' EEG and 6 blocks of activities, each of them lasts for 30 seconds; the second design has 4 minutes' recording and 4 blocks of activities, each of them lasts for 1 minute; The third design has 2 minutes and 30 seconds' EEG and 4 blocks of activities (30 seconds, 30 seconds, 1 minute and 30 seconds each); and the last design also has 4 minutes' recording and 4 blocks of activities, each of them lasts for 1 minute. All experiment designs show good results after we apply the BCCPM model, which allows us to use BCCPM in detecting change points in EEG data in additional to fMRI data on which the BCCPM was original designed and validated.

Note that in BCCPM, there is one tuning parameter p in the program allowing us to try different values to detect change points with different sensitivity. By default, the p is set to be 0, if we increase p to positive values, there will be more change points detected from the EEG data; if we decrease p to negative values, say -1000, -2000, etc., there will be fewer change points detected. In the experimental data, choosing the optimal p is also of interest, so multiple values has been applied: 0, -1000, -2000, -4000, -6000. We start from 0 because there are already more than wanted change points detected at this level. Also note that we repeat the MCMC five times for each of the recorded data and choose the best result.

3.4.1 Experimental Design

The **first** experiment setting has 3 minutes recording with the following activities (a brief summary can be found in Table 3.1):

- The first 30 seconds: listen to music
- The second 30 seconds: solve basic mathematical operations problems in mind
- The third 30 seconds: read paragraphs from a novel
- Repeat the first 30 seconds: listen to music
- Repeat the second 30 seconds: solve basic mathematical operations problems in mind
- Repeat the third 30 seconds: read paragraphs from a novel

Table 3.1. Experiment 1 (in seconds)

0 - 30	30 - 60	60 - 90	90 - 120	120 - 150	150 - 180
music	math	novel	music	math	novel

The **second** setting has 4 minutes recording with the following activities (Table 3.2):

- The first minute: listen to music
- The second minute: read paragraphs from a novel
- Repeat the first minute: listen to music
- Repeat the second minute: read paragraphs from a novel

Table 3.2. Experiment 2 (in minutes)

0 - 1	1 - 2	2 - 3	3 - 4
music	novel	music	novel

The **third** setting has 2 minutes and 30 seconds recording with the following activities (summary shown in Table 3.3):

- The first 30 seconds: listen to live news on the radio
- The second 30 seconds: solve basic mathematical operations problems in mind
- Another one minute: listen to live news on the radio (minor change rather than repeating for the same time period)
- Repeat the second 30 seconds: solve basic mathematical operations problems in mind

Table 3.3. Experiment 3 (in seconds)

0 - 30	30 - 60	60 - 120	120 - 150
news	math	news	math

The **fourth** setting has 4 minutes recording with the following activities (Table 3.4):

- The first minute: listen to live news on the radio
- The second minute: solve basic mathematical operations problems in mind
- Repeat the first minute: listen to live news on the radio
- Repeat the second minute: solve basic mathematical operations problems in mind

Table 3.4. Experiment 4 (in minutes)

0 - 1	1 - 2	2 - 3	3 - 4
news	math	news	math

3.4.2 Results

All the four experimental results show good detection of the designed change points. As mentioned at the beginning of this subsection, different values of tuning parameter p are applied and the results are compared, we can observe that as p decreases from 0 to -1000 to -2000, the number of change points decreases accordingly.

Results of the first experiment are shown in Figures 3.3 to 3.13. When $p = 0$, the convergence of the markov chains are shown in Figure 3.3 and the change points detected by BCCPM are shown in Figure 3.4. The top left plot in Figure 3.4 has the best result according to the convergence trace plot (highest curve) and it's enlarged as in Figure 3.5. Now let's take Figure 3.5 as an example, right before or after each of the red dotted lines (locations of the designed change points), there is always a spike, which indicates a change point detected by BCCPM at that location. At the same time, we may also observe there are many other spikes, which are change points detected within each block at $p = 0$. Those could be simply the change of brain activity related to specific task within each designed activity block, and the number will decrease as p goes to -1000, -2000,..., -6000. When $p = -1000$, the results are shown in Figures 3.6 to 3.8. Now we can observe that the number of change points decreases. When $p = -2000$, the results are shown in Figures 3.9 to 3.11, the number of change points keeps decreasing to an optimal stage where we can observe the change point pattern much more clearly. In this study, we keep decreasing the p to -4000 and -6000, to show that the model can detect the actual designed change points without other possible change points not directly related to the designed blocks; for simplicity, only the best results are shown in Figures 3.12 to 3.13. Other figures are listed in Appendix A from A.1 to A.4.

Results of the second experiment are shown in Figure 3.14 to 3.24, A.5 to A.8; results of the third experiment are shown in Figure 3.25 to 3.35, A.9 to A.12; and results of the fourth experiment are shown in Figure 3.36 to 3.46, A.13 to A.16, respectively.

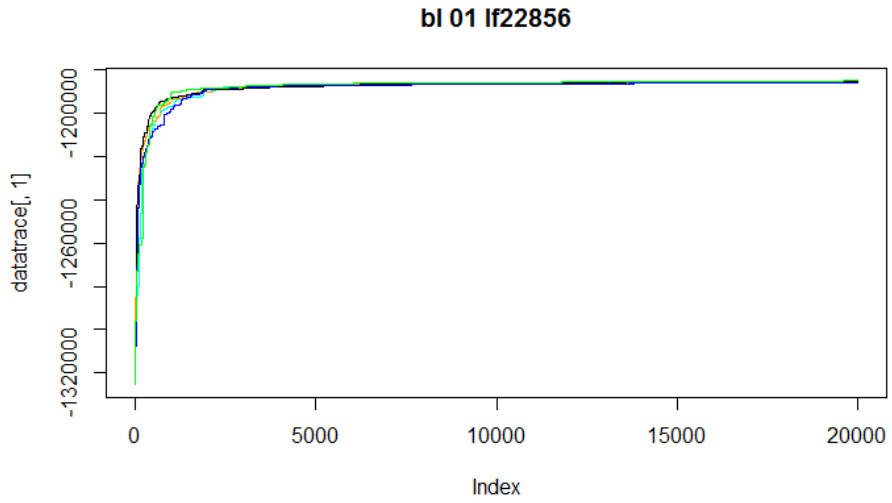

Figure 3.3. Experiment 1: Traceplot shows the convergence of MCMC chains (p=0).

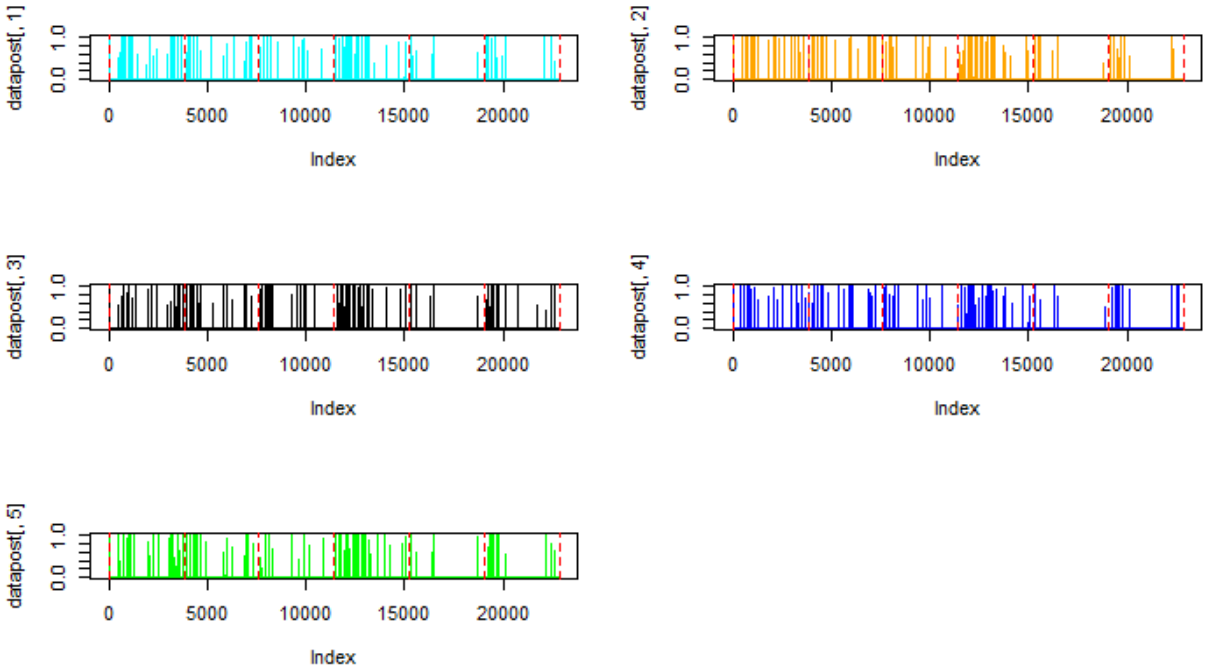

Figure 3.4. Experiment 1: Change points detected by BCCPM for five repeated MCMC chains (p=0). Red dotted lines are the locations of designed change points.

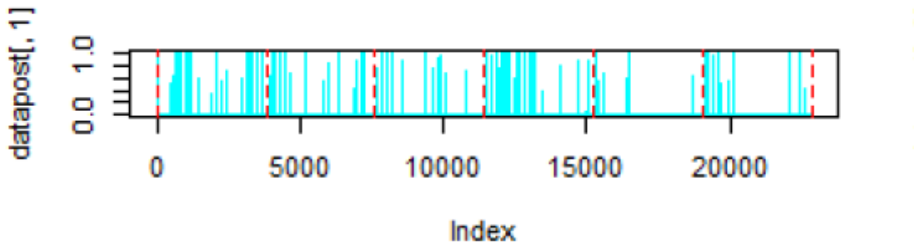

Figure 3.5. Experiment 1: Change points detected by BCCPM for the best detection result (p=0). Red dotted lines are the locations of designed change points.

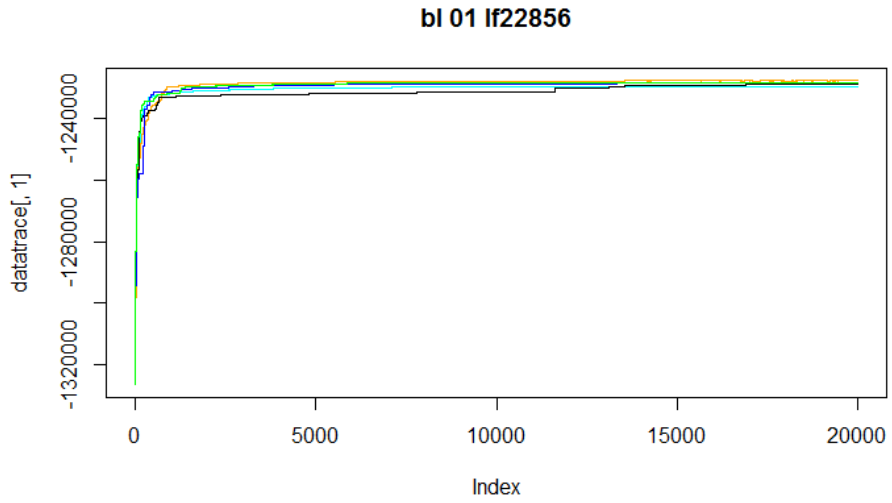

Figure 3.6. Experiment 1: Traceplot shows the convergence of MCMC chains (p=-1000).

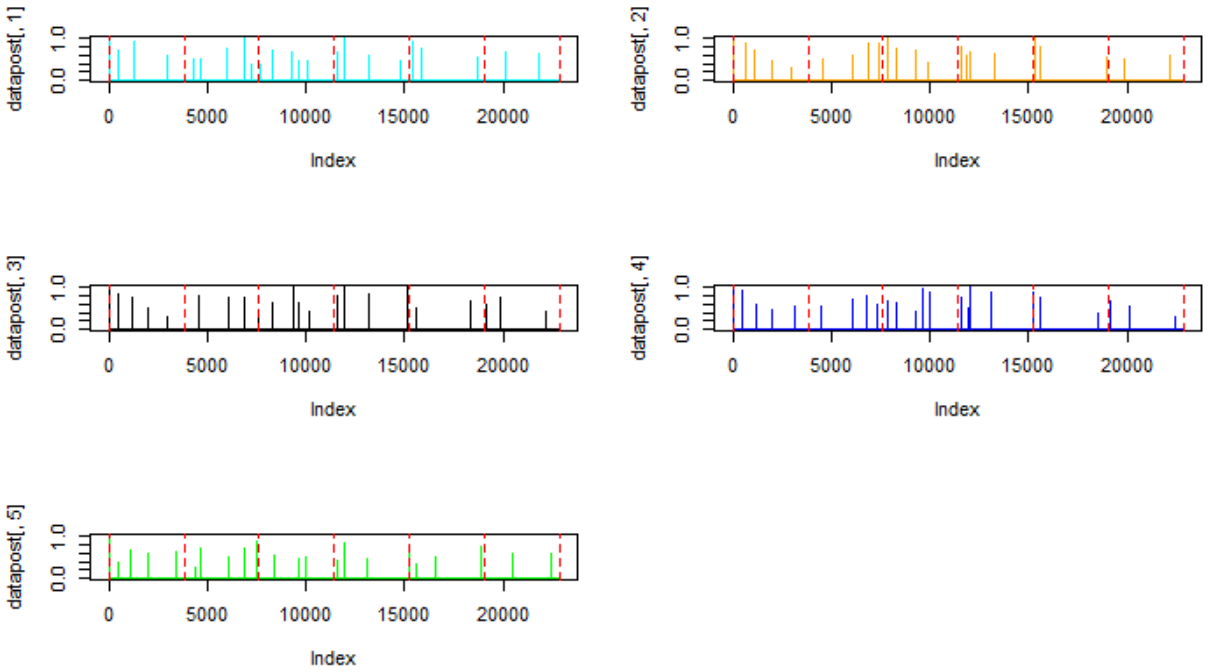

Figure 3.7. Experiment 1: Change points detected by BCCPM for five repeated MCMC chains (p=-1000). Red dotted lines are the locations of designed change points.

Figure 3.8. Experiment 1: Change points detected by BCCPM for the best detection result (p=-1000). Red dotted lines are the locations of designed change points.

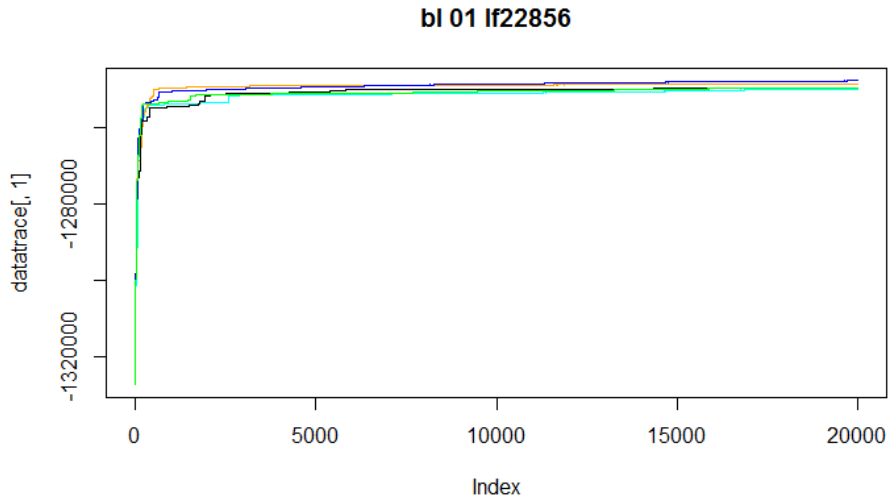

Figure 3.9. Experiment 1: Traceplot shows the convergence of MCMC chains (p=-2000).

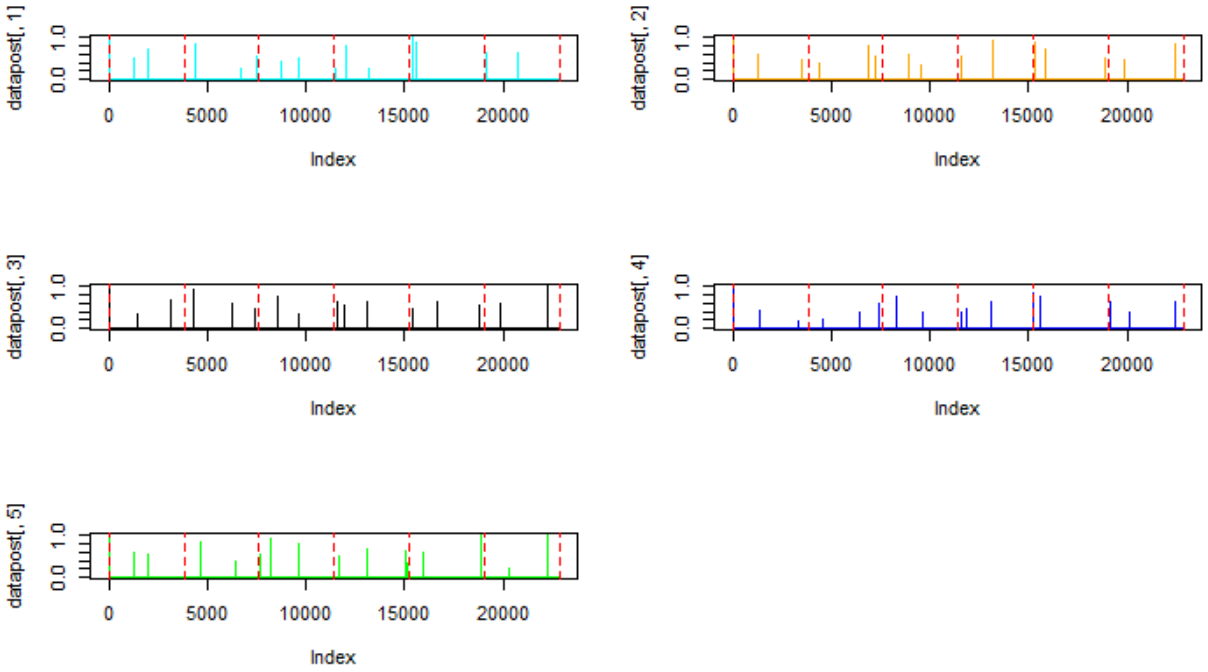

Figure 3.10. Experiment 1: Change points detected by BCCPM for five repeated MCMC chains (p=-2000). Red dotted lines are the locations of designed change points.

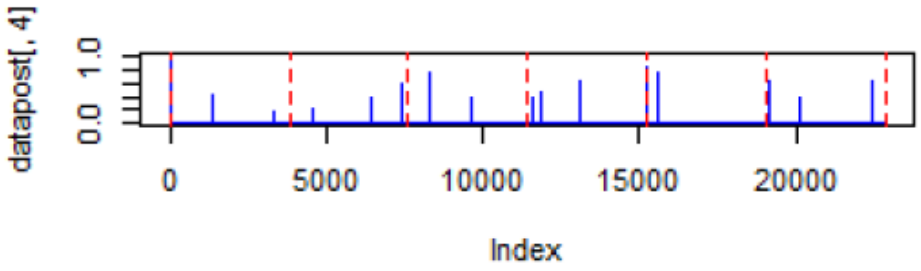

Figure 3.11. Experiment 1: Change points detected by BCCPM for the best detection result (p=-2000). Red dotted lines are the locations of designed change points.

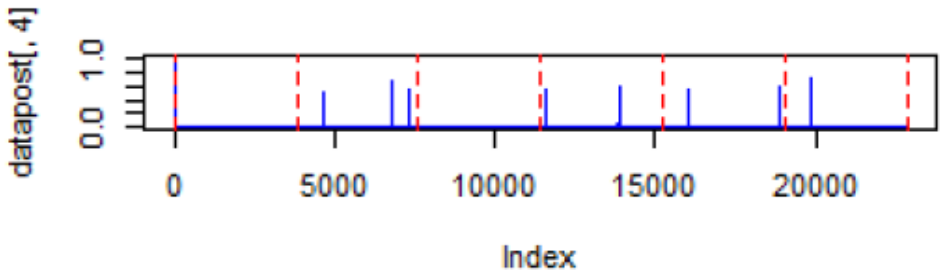

Figure 3.12. Experiment 1: Change points detected by BCCPM for the best detection result (p=-4000). Red dotted lines are the locations of designed change points.

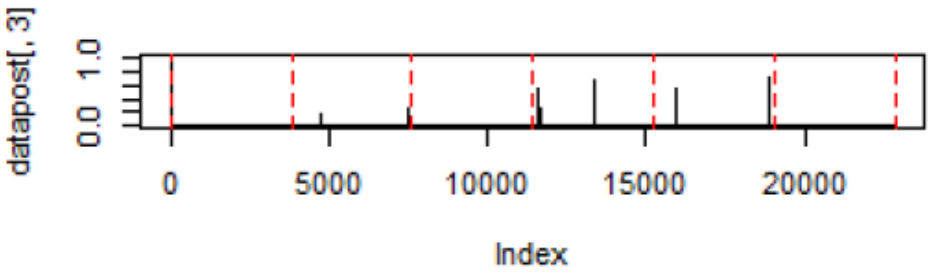

Figure 3.13. Experiment 1: Change points detected by BCCPM for the best detection result (p=-6000). Red dotted lines are the locations of designed change points.

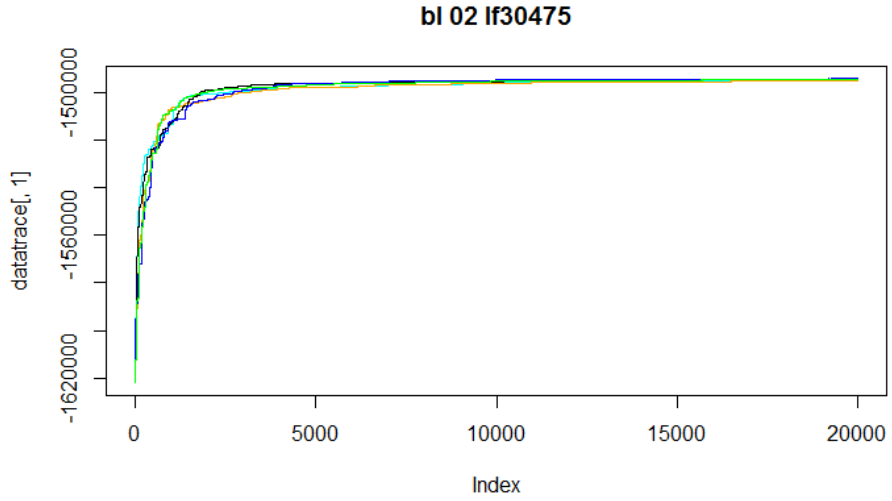

Figure 3.14. Experiment 2: Traceplot shows the convergence of MCMC chains (p=0).

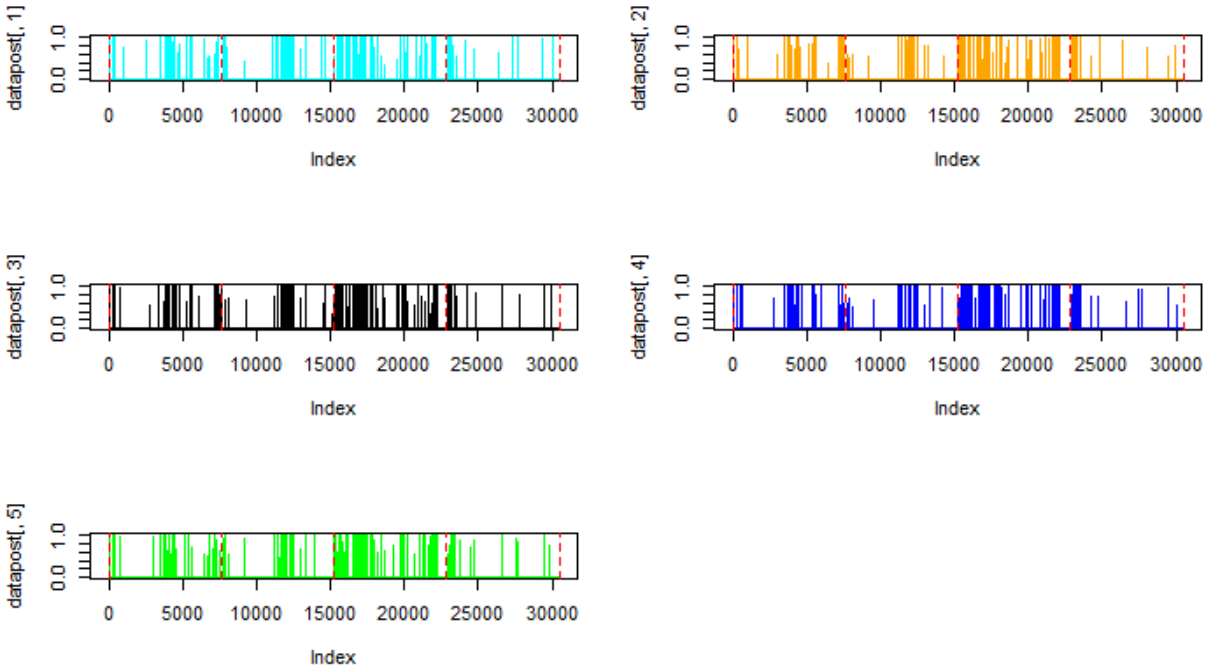

Figure 3.15. Experiment 2: Change points detected by BCCPM for five repeated MCMC chains (p=0). Red dotted lines are the locations of designed change points.

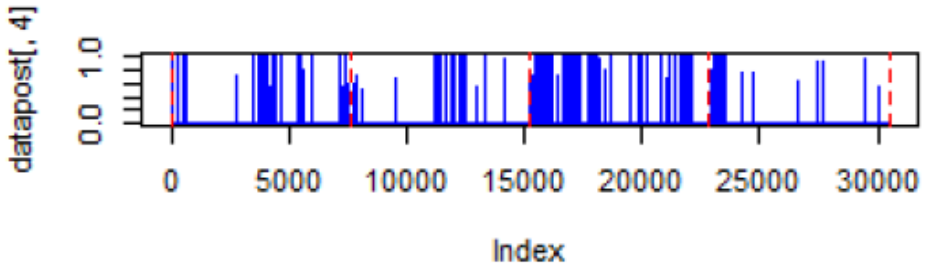

Figure 3.16. Experiment 2: Change points detected by BCCPM for the best detection result (p=0). Red dotted lines are the locations of designed change points.

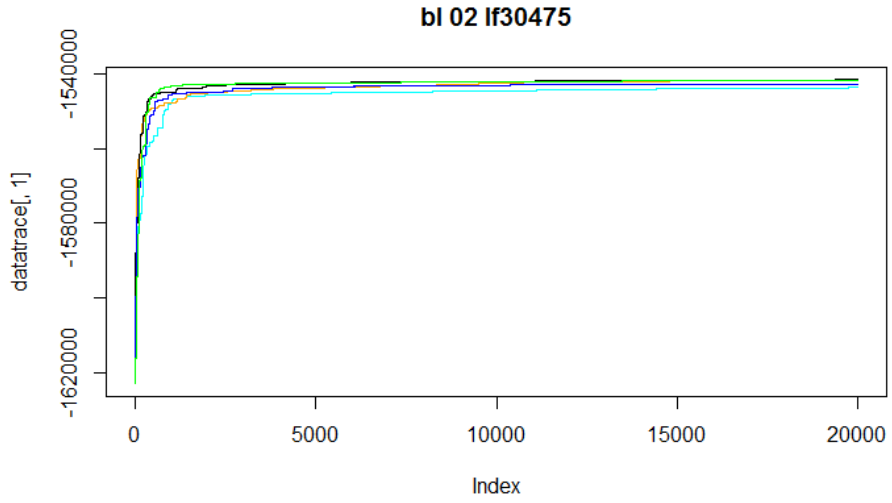

Figure 3.17. Experiment 2: Traceplot shows the convergence of MCMC chains (p=-1000).

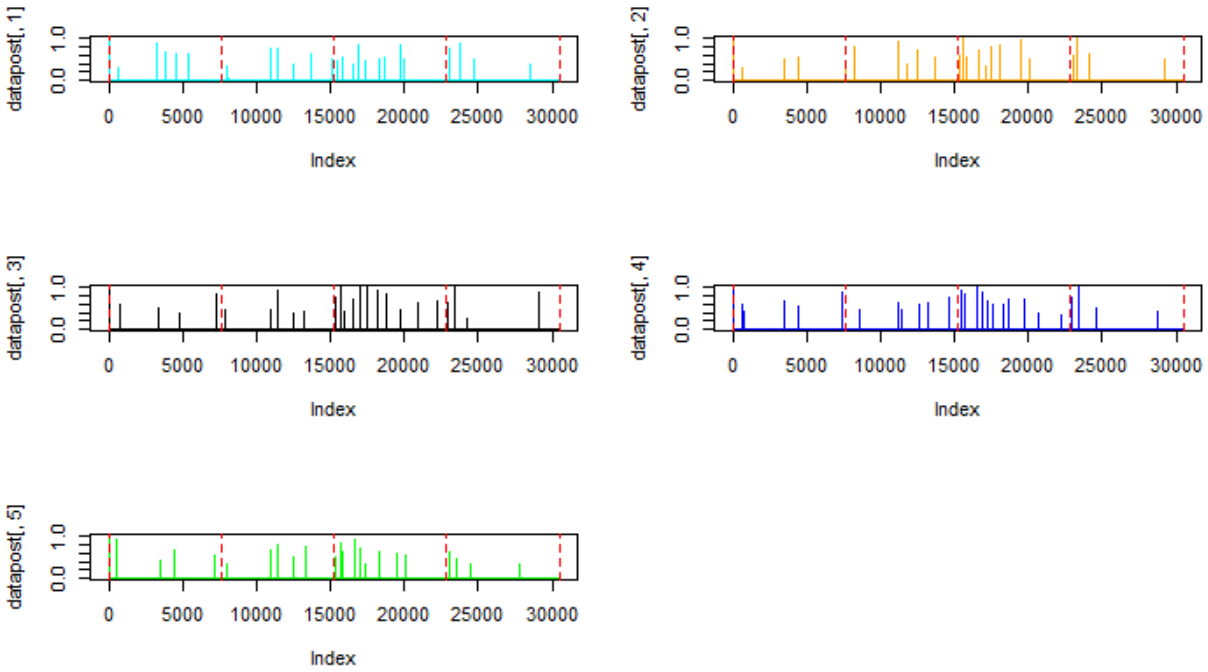

Figure 3.18. Experiment 2: Change points detected by BCCPM for five repeated MCMC chains (p=-1000). Red dotted lines are the locations of designed change points.

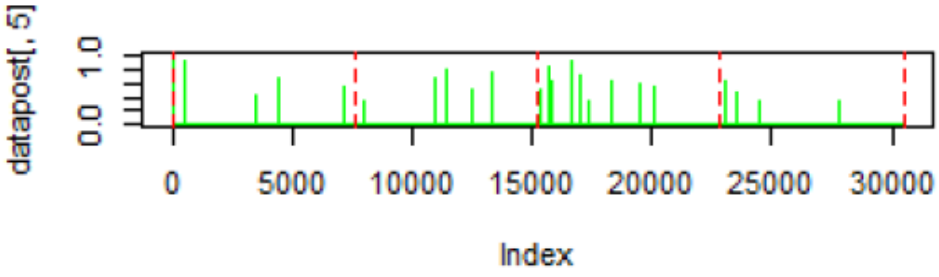

Figure 3.19. Experiment 2: Change points detected by BCCPM for the best detection result (p=-1000). Red dotted lines are the locations of designed change points.

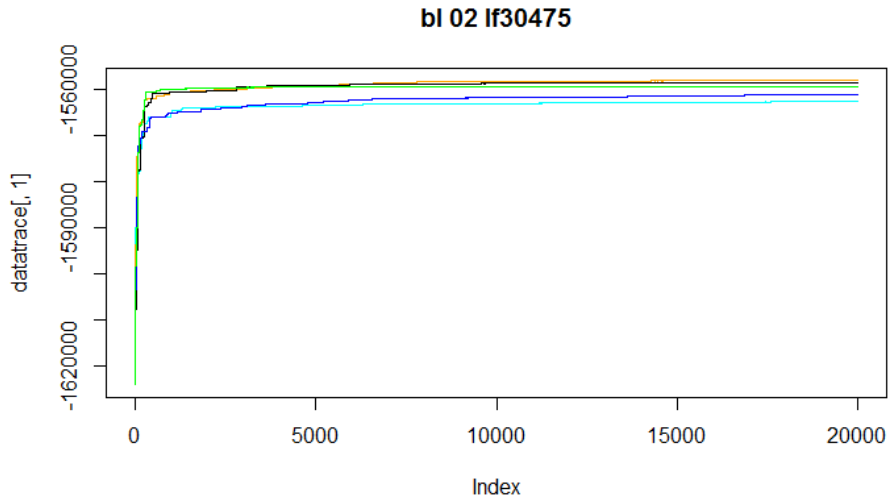

Figure 3.20. Experiment 2: Traceplot shows the convergence of MCMC chains (p=-2000).

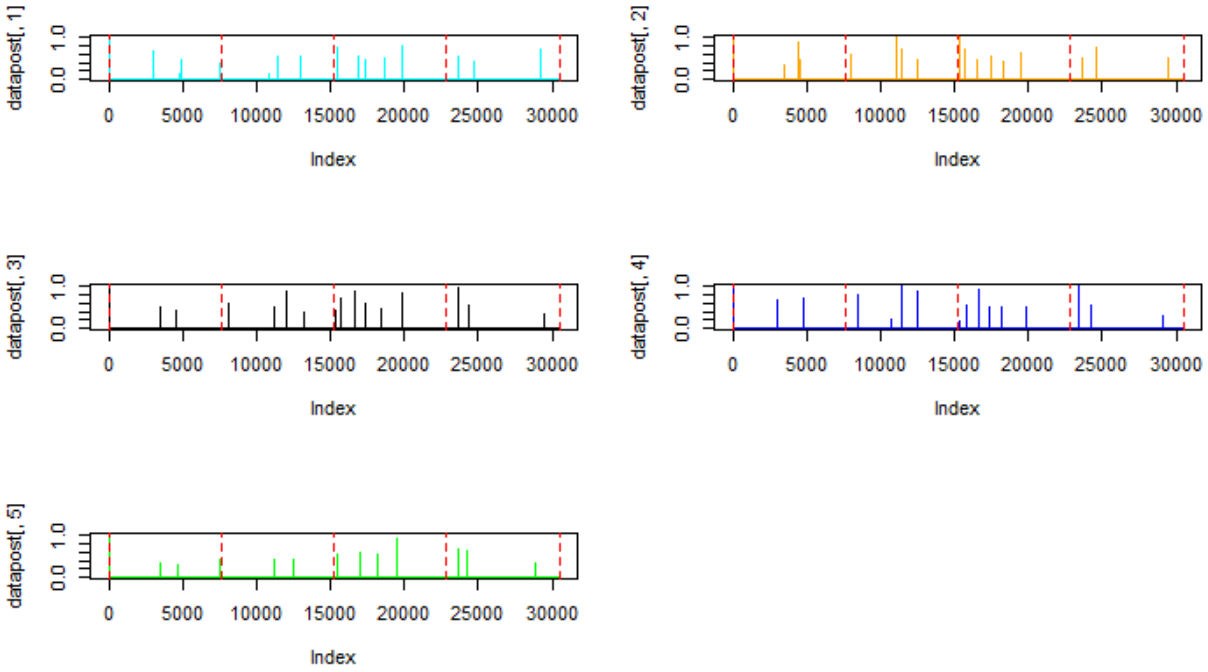

Figure 3.21. Experiment 2: Change points detected by BCCPM for five repeated MCMC chains (p=-2000). Red dotted lines are the locations of designed change points.

Figure 3.22. Experiment 2: Change points detected by BCCPM for the best detection result (p=-2000). Red dotted lines are the locations of designed change points.

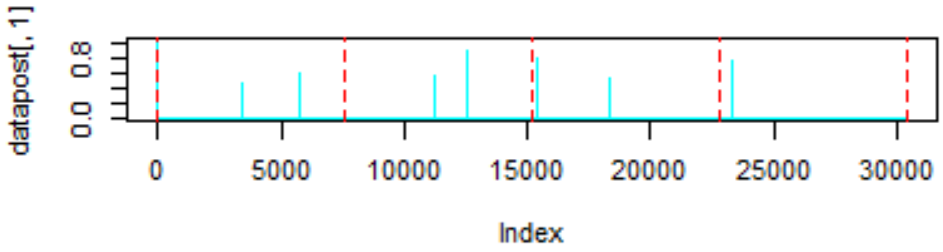

Figure 3.23. Experiment 2: Change points detected by BCCPM for the best detection result (p=-4000). Red dotted lines are the locations of designed change points.

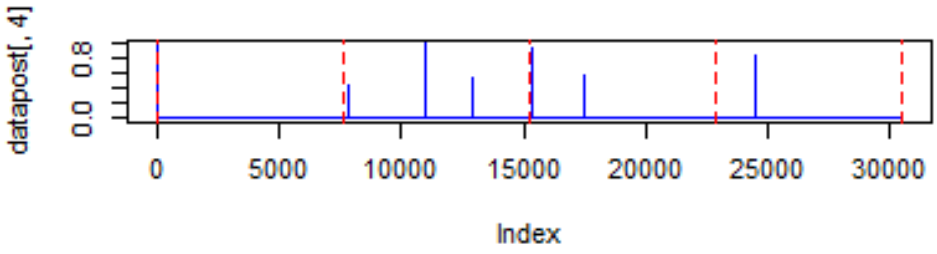

Figure 3.24. Experiment 2: Change points detected by BCCPM for the best detection result (p=-6000). Red dotted lines are the locations of designed change points.

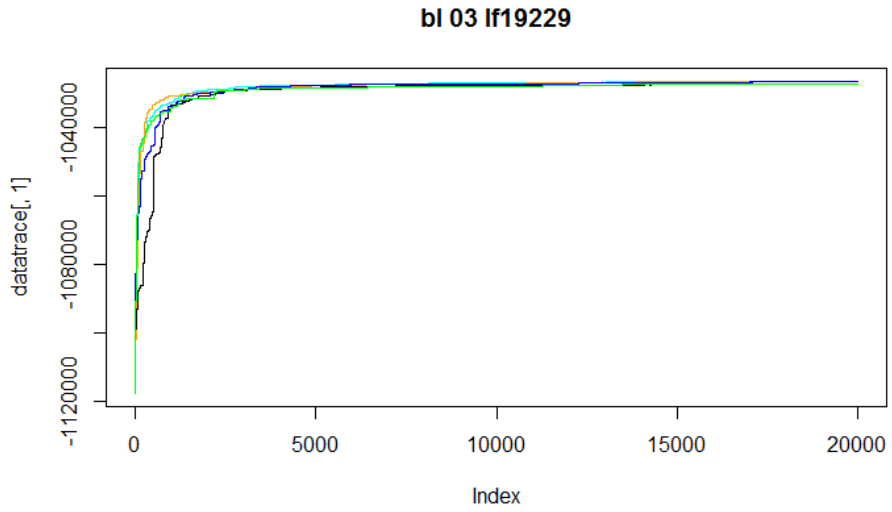

Figure 3.25. Experiment 3: Traceplot shows the convergence of MCMC chains (p=0).

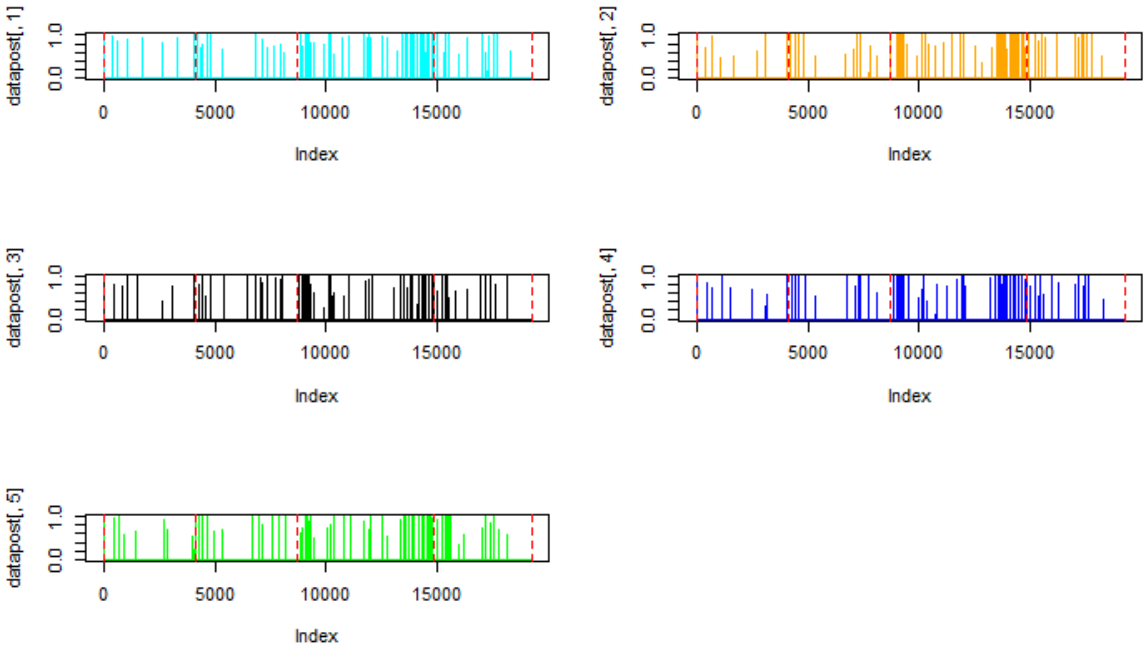

Figure 3.26. Experiment 3: Change points detected by BCCPM for five repeated MCMC chains (p=0). Red dotted lines are the locations of designed change points.

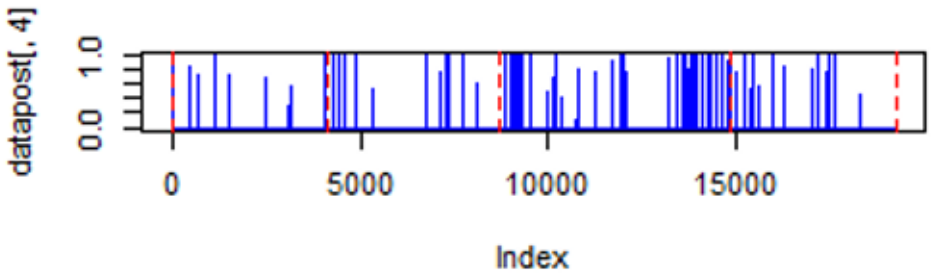

Figure 3.27. Experiment 3: Change points detected by BCCPM for the best detection result (p=0). Red dotted lines are the locations of designed change points.

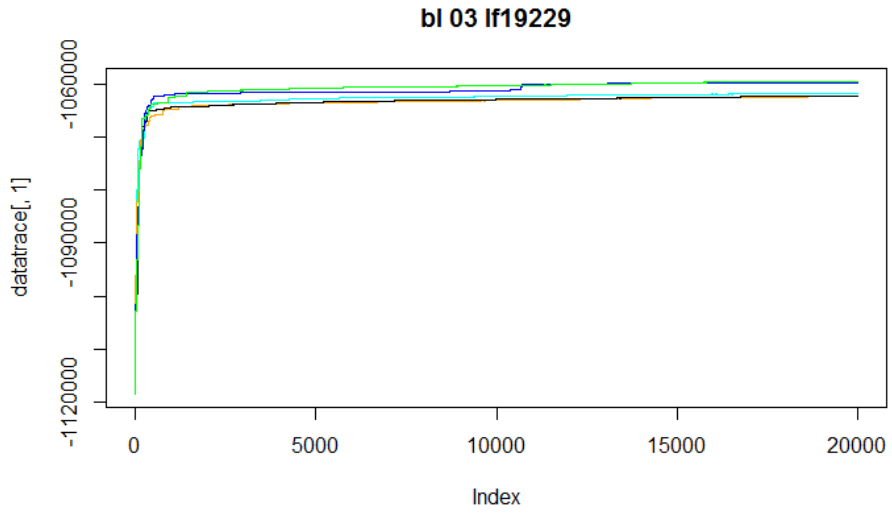

Figure 3.28. Experiment 3: Traceplot shows the convergence of MCMC chains (p=-1000).

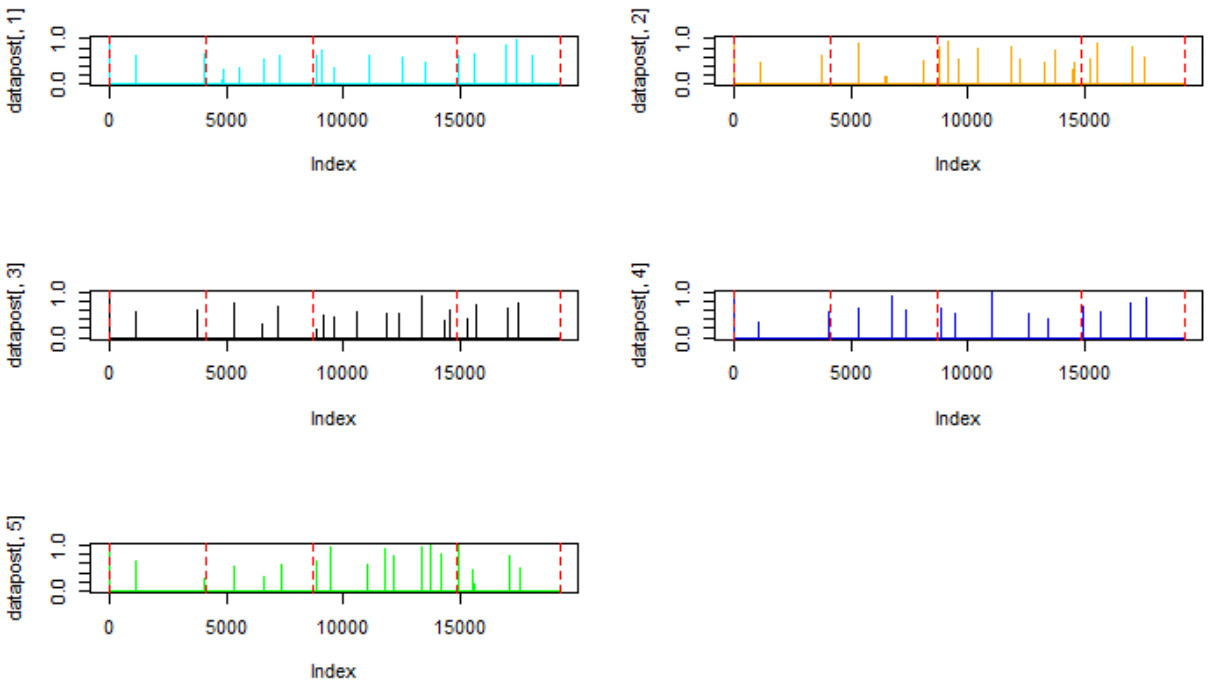

Figure 3.29. Experiment 3: Change points detected by BCCPM for five repeated MCMC chains (p=-1000). Red dotted lines are the locations of designed change points.

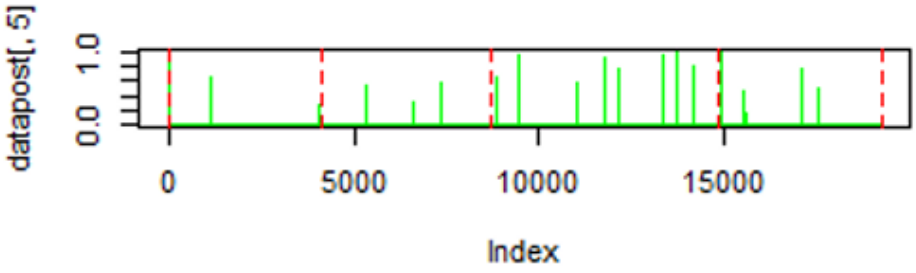

Figure 3.30. Experiment 3: Change points detected by BCCPM for the best detection result (p=-1000). Red dotted lines are the locations of designed change points.

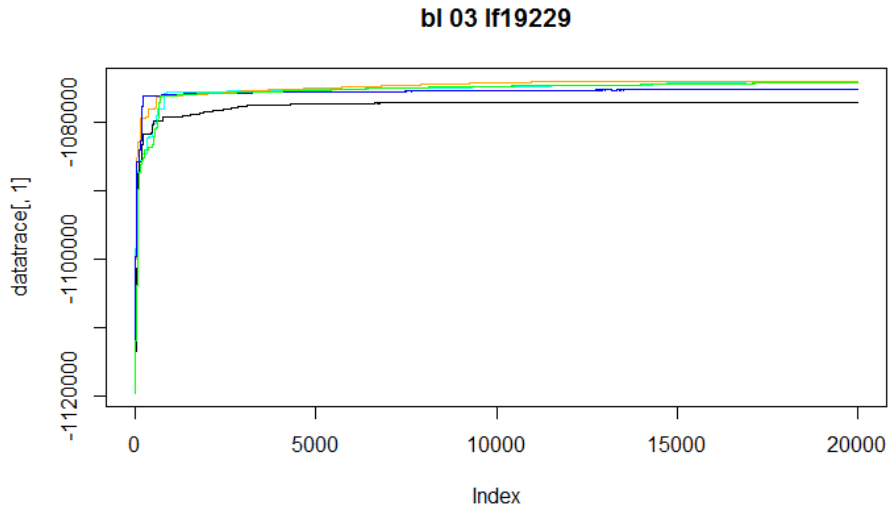

Figure 3.31. Experiment 3: Traceplot shows the convergence of MCMC chains (p=-2000).

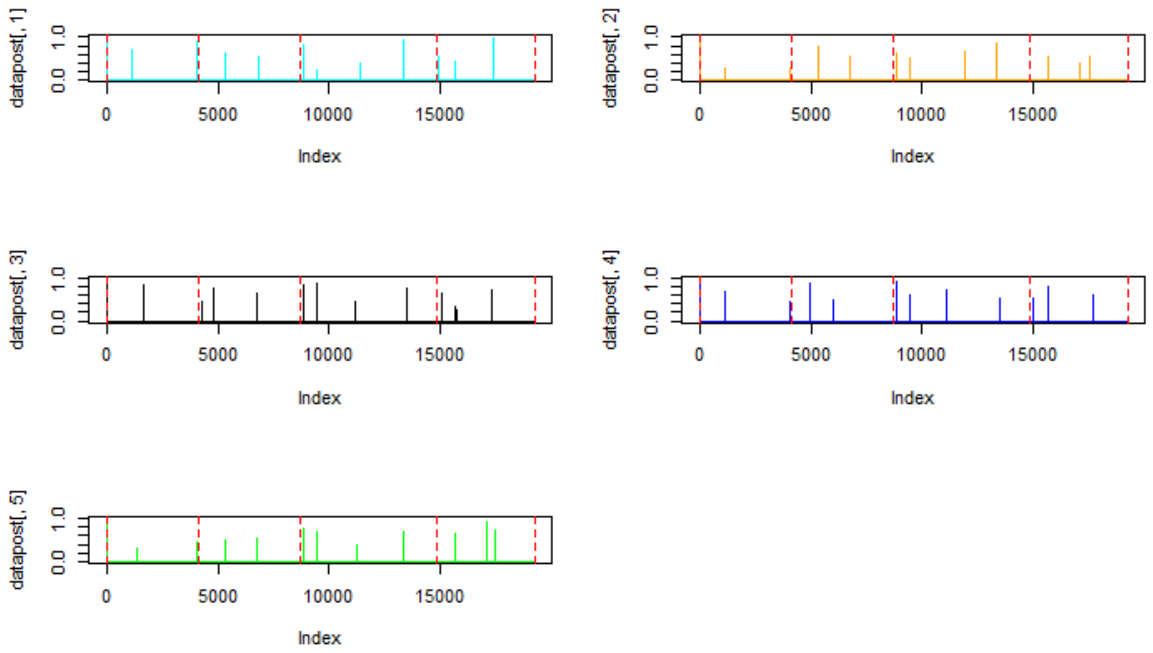

Figure 3.32. Experiment 3: Change points detected by BCCPM for five repeated MCMC chains (p=-2000). Red dotted lines are the locations of designed change points.

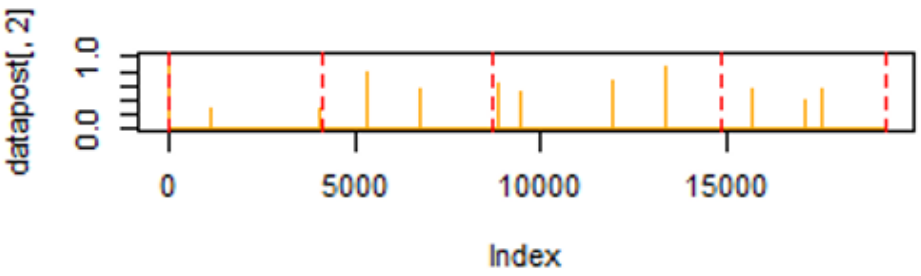

Figure 3.33. Experiment 3: Change points detected by BCCPM for the best detection result (p=-2000). Red dotted lines are the locations of designed change points.

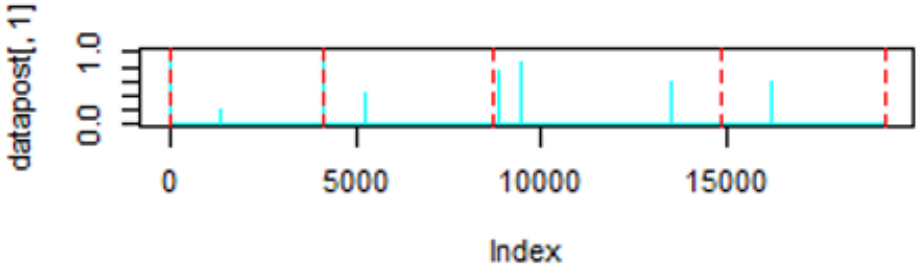

Figure 3.34. Experiment 3: Change points detected by BCCPM for the best detection result (p=-4000). Red dotted lines are the locations of designed change points.

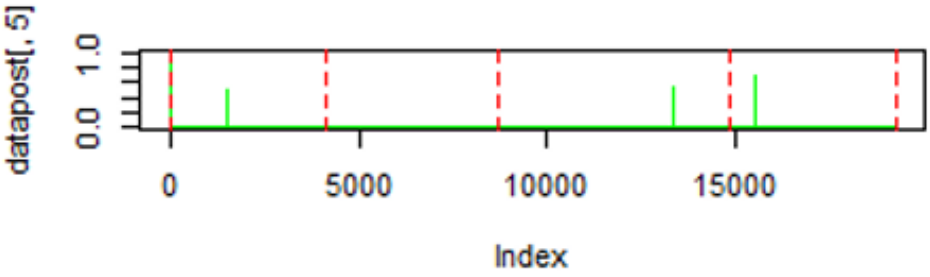

Figure 3.35. Experiment 3: Change points detected by BCCPM for the best detection result (p=-6000). Red dotted lines are the locations of designed change points.

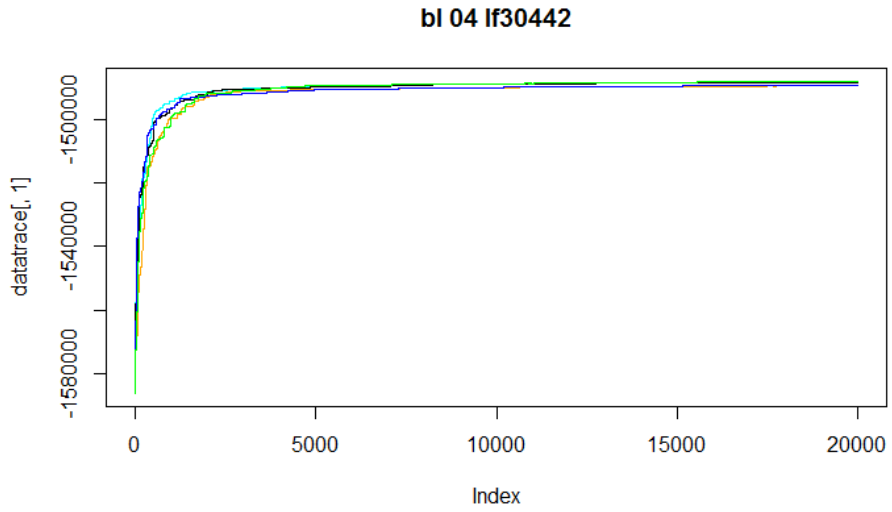

Figure 3.36. Experiment 4: Traceplot shows the convergence of MCMC chains (p=0).

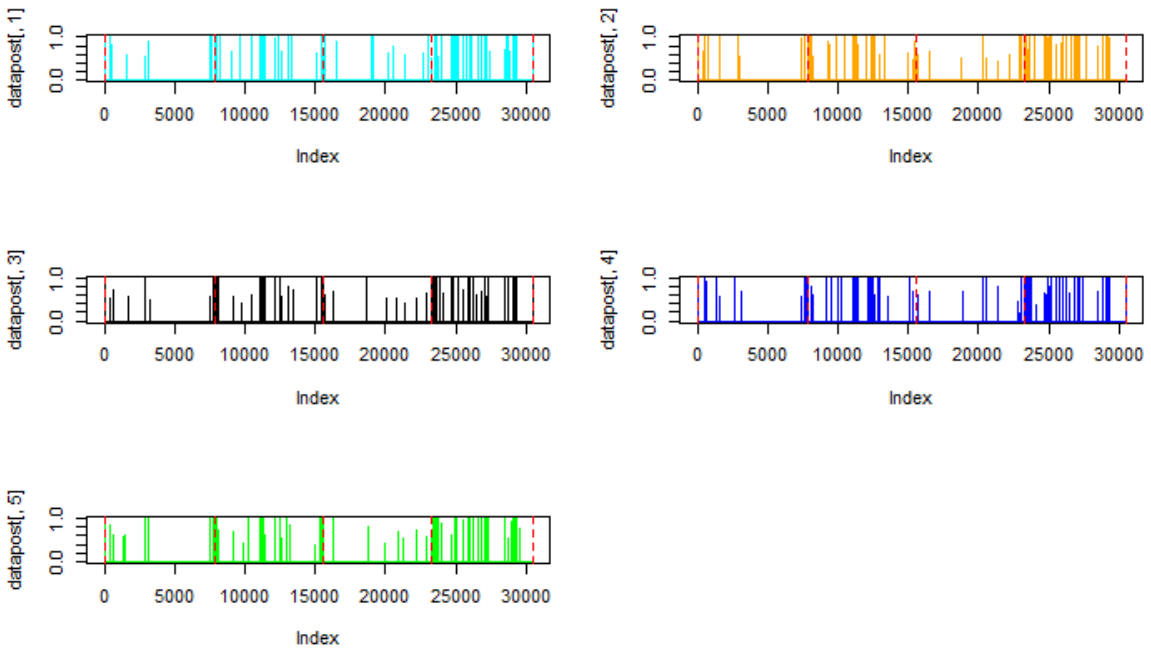

Figure 3.37. Experiment 4: Change points detected by BCCPM for five repeated MCMC chains (p=0). Red dotted lines are the locations of designed change points.

Figure 3.38. Experiment 4: Change points detected by BCCPM for the best detection result (p=0). Red dotted lines are the locations of designed change points.

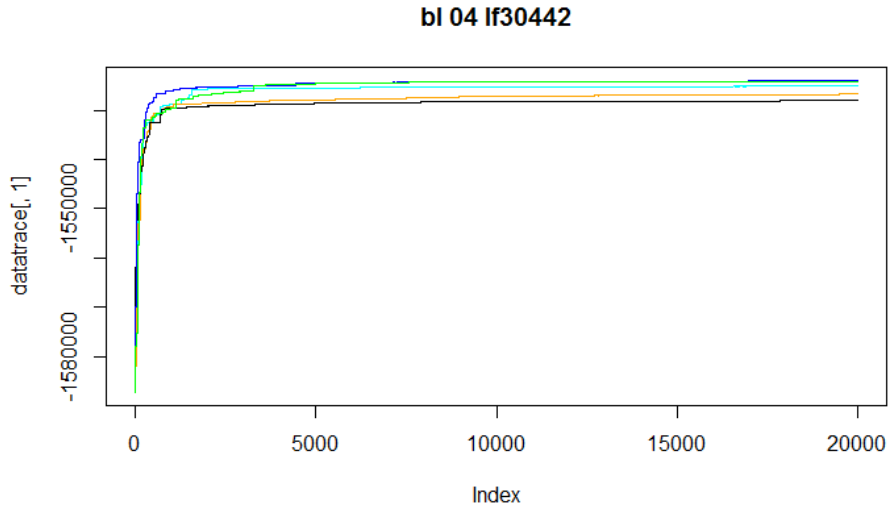

Figure 3.39. Experiment 4: Traceplot shows the convergence of MCMC chains (p=-1000).

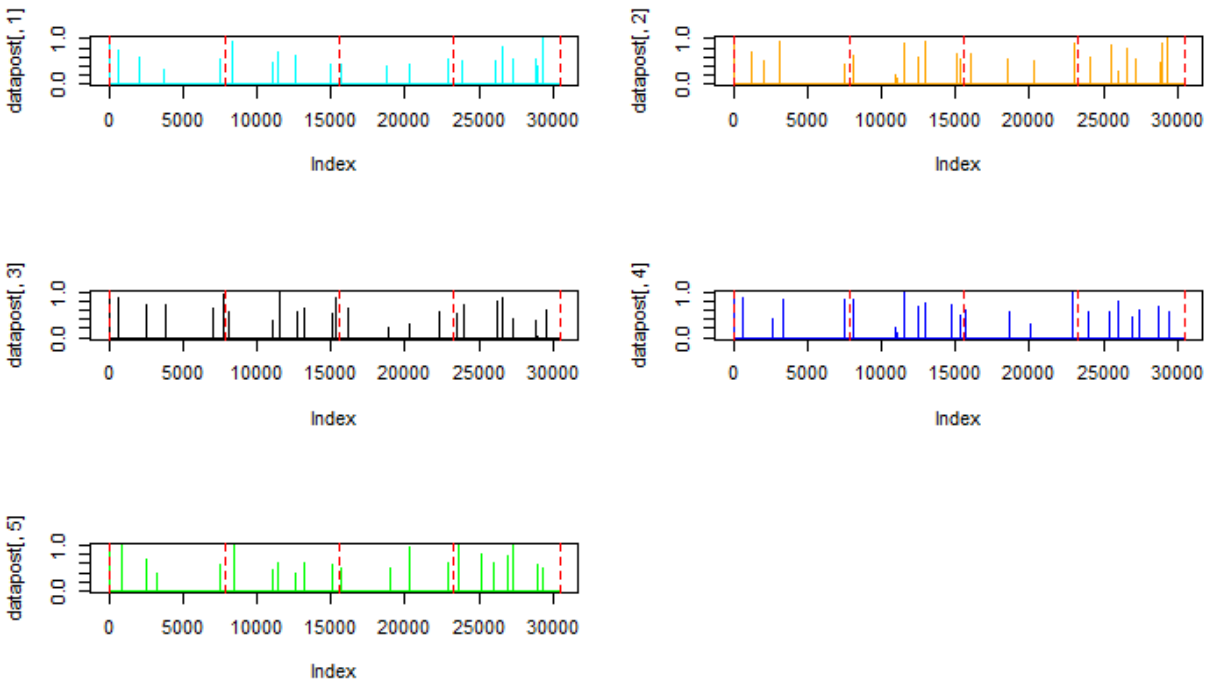

Figure 3.40. Experiment 4: Change points detected by BCCPM for five repeated MCMC chains (p=-1000). Red dotted lines are the locations of designed change points.

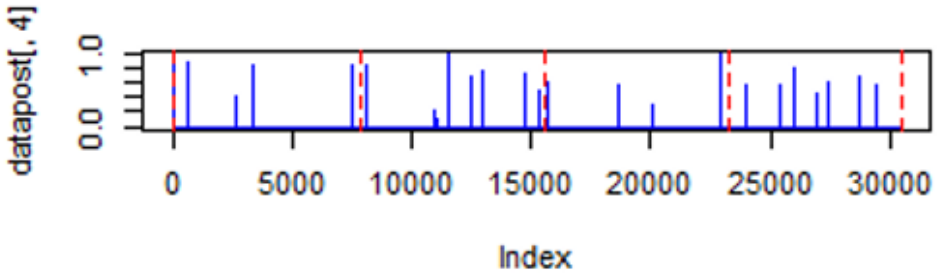

Figure 3.41. Experiment 4: Change points detected by BCCPM for the best detection result (p=-1000). Red dotted lines are the locations of designed change points.

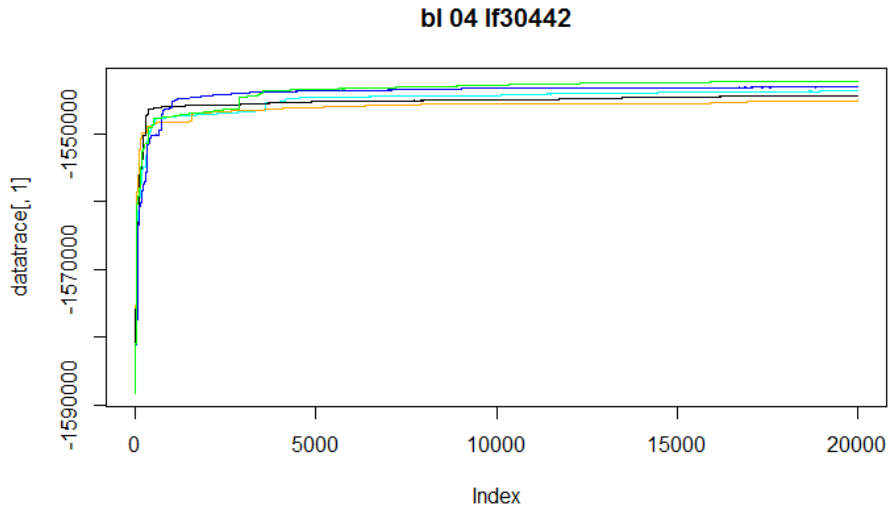

Figure 3.42. Experiment 4: Traceplot shows the convergence of MCMC chains (p=-2000).

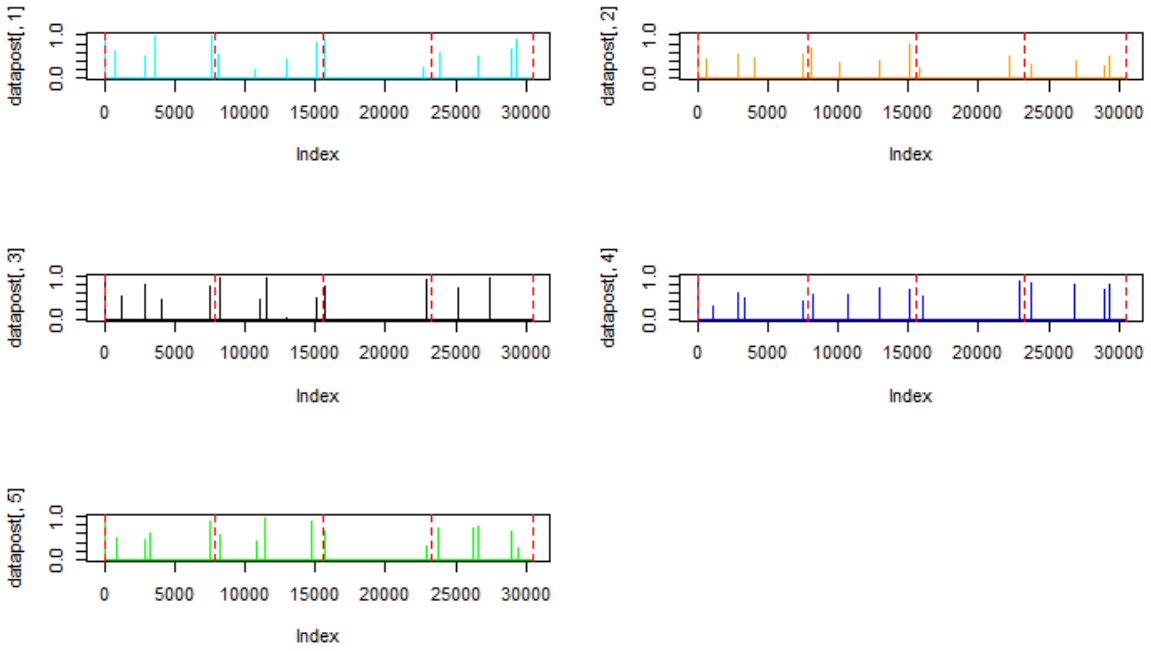

Figure 3.43. Experiment 4: Change points detected by BCCPM for five repeated MCMC chains (p=-2000). Red dotted lines are the locations of designed change points.

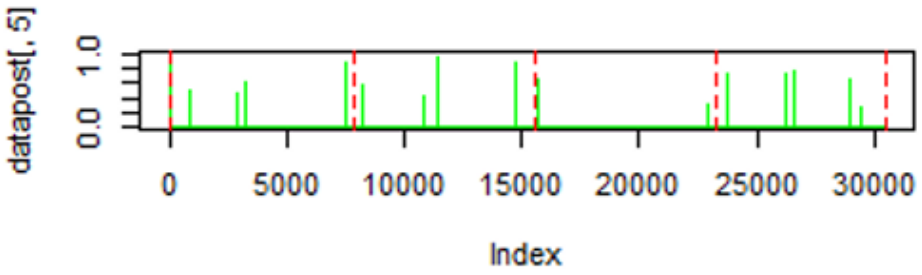

Figure 3.44. Experiment 4: Change points detected by BCCPM for the best detection result (p=-2000). Red dotted lines are the locations of designed change points.

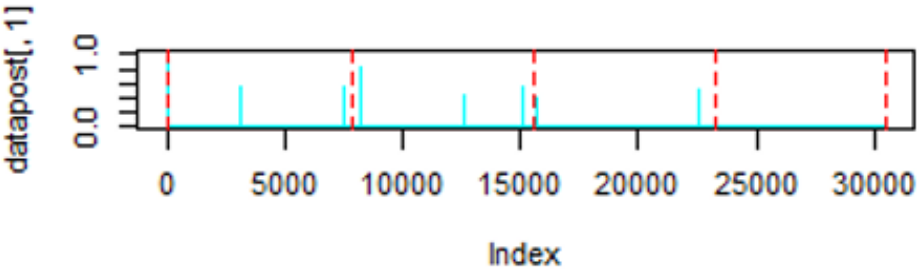

Figure 3.45. Experiment 4: Change points detected by BCCPM for the best detection result (p=-4000). Red dotted lines are the locations of designed change points.

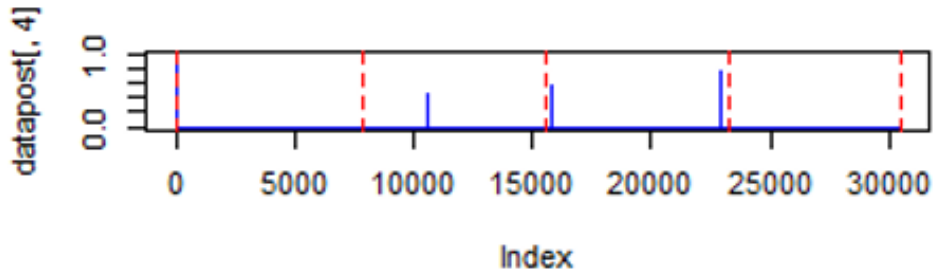

Figure 3.46. Experiment 4: Change points detected by BCCPM for the best detection result (p=-6000). Red dotted lines are the locations of designed change points.

3.5 Results on EEG Data from Mindfulness Therapy for Dysfunctional Anxiety Patients

After we validated the application of BCCPM towards EEG data through experimental designs, we finally applied the model to our real data analysis. Results indicate that: 1) as the mindfulness therapy goes from session 1 to later sessions, the change points in EEG data from the subjects are decreasing, this pattern overtime indicates the subjects become less anxious as the therapy sessions goes on; 2) At the same time, with paired data analysis, we observe that the change-point patterns are different for different subjects in the same session. This implies that the dynamics change is with response to treatment.

3.5.1 Recorded Data Description

In the study, there are 8 pairs of subjects in multiple mindfulness therapy sessions. A list of paired subjects is shown in Table 3.5 with recording date, subject number and session number.

Table 3.5. List of paired subjects

Recording Date	***Subject***	***Session***
10/26/2015	1001	1
	1003	1
11/02/2015	1002	2
	1006	2
11/30/2015	1002	5
	1005	5
12/07/2015	1002	6
	1006	6
12/14/2015	1001	7
	1005	7
02/01/2016	1012	1
	1013	1
02/22/2016	1009	4
	1012	4
03/21/2016	1012	8
	1013	8

3.5.2 Results

As briefly discussed at the beginning of this subsection, there are two main results we can find form change point analysis of this EEG data. First, as the mindfulness therapy goes from session 1 to later sessions, the number of change points in EEG data is decreasing, this pattern overtime indicates the subjects become less anxious as the therapy sessions goes on. This can be shown in Figure 3.47, as the therapy session goes on from 1 to 4 to 7 for subject 1, the number of change points decreases, which indicates that as the therapy continues, the subject become less anxious in terms of brain activities changes. The same trend is discovered in all the subjects. Note that in this section, p is set at -2000 for best results.

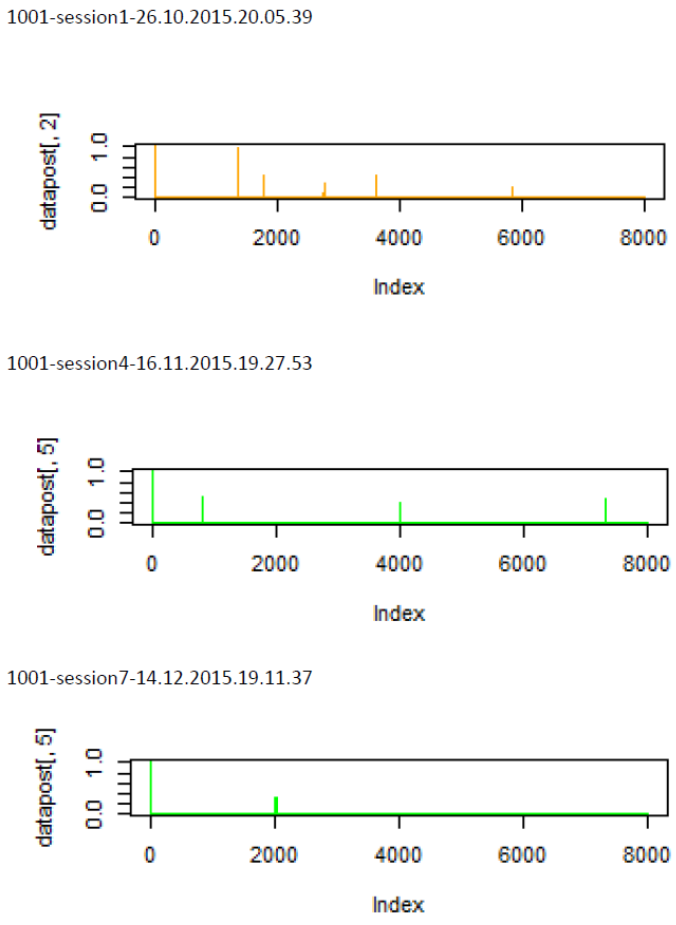

Figure 3.47. Comparison of number of change points of Subject 1 from therapy session 1 to 4 to 7. The number of change points decreases as sessions go on (p=-2000).

In order to measure this trend quantitatively, we used Poisson process and estimated the parameter rate which can be interpreted as the average number of points per some unit of extents such as length, area, volumn, or time (which is our case). The Poisson rate describes the average number of change points per time in our study and the results are summarized in Table 3.6. We can see that the rate decreases for each subject as session goes on for all the subjects (except for subject 1006). The standard errors are shown in round brackets ().

Table 3.6. Poisson rates of test subjects in different sessions (p=-2000).

Subject	***Session***	***Rate (Std. Error)***
1001	1	0.00075 *(0.00031)*
1001	4	0.00050 *(0.00025)*
1001	7	0.00025 *(0.00018)*
1002	2	0.000625 *(0.00028)*
1002	5	0.000375 *(0.00022)*
1005	5	0.00125 *(0.0004)*
1005	7	0.000875 *(0.00033)*
1006	2	0.0005 *(0.00025)*
1006	6	0.0005 *(0.00025)*

Second, the change-point patterns are different for two subjects in the same session. This implies that the dynamics change is with response to treatment. An example is shown in Figure 3.48. For Subject 1 and Subject 5 in session 7, the change point patterns are quite different, which implies the change points are not associated with the therapy instructions given by the therapist, instead it shows the true brain activities changes of the individual subjects.

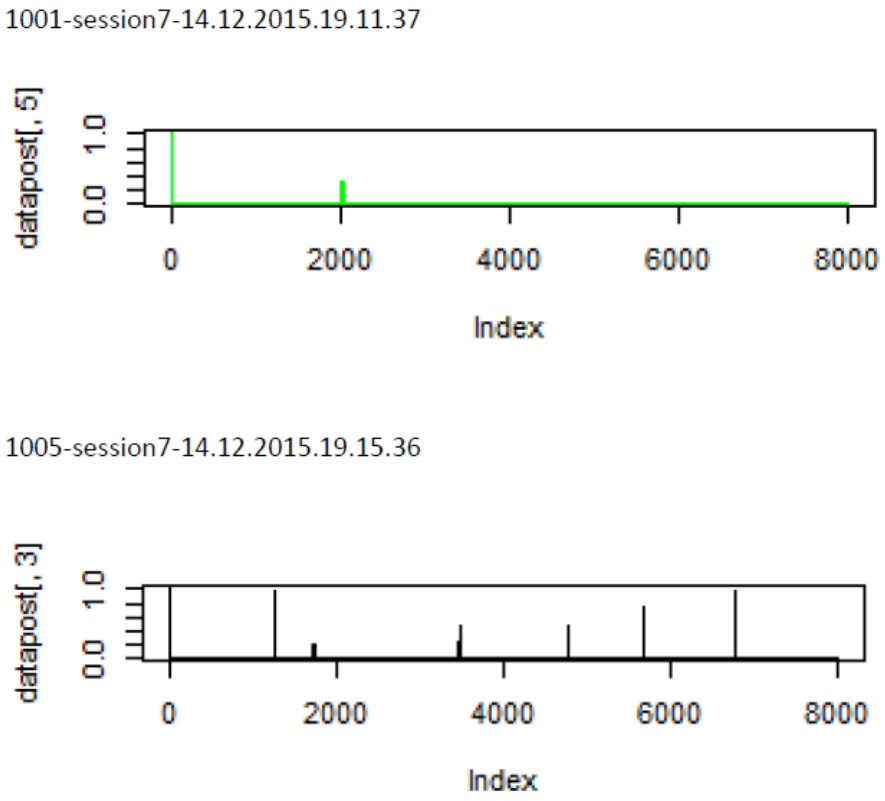

Figure 3.48. In session 7, Subject 1 and Subject 5 have different change point patterns (p=-2000).

3.6 Summary

In this chapter, we successfully extend the application of the Bayesian connectivity change point detection model (BCCPM) onto the change point analysis of Electroencephalography (EEG) data to determine network dynamics over time. The ability of EEG measures of frontal and temporo-parietal activity during mindfulness therapy to track response to treatment as preliminary evaluation for EEG as a physiological aid in therapy is successfully tested by employing the concept of Bayesian inference using BCCPM.

CHAPTER 4

AN OPTIMIZED BAYESIAN FUNCTIONAL CONNECTIVITY CHANGE POINT MODEL WITH GENETIC ALGORITHM

The material in this chapter is from the author's research articles [98][97].

This research was supported by the Brains-Behavior Seed grant and Molecular Basis of Disease(MBD) from Georgia State University. It's a collaborative work with Dr. Xiuchun Xiao, Dr. Jing Zhang, Dr. Yi Pan and Xueli Xiao.

The paper [98] is to appear in the *Journal of Computational Biology*; Xiuchun Xiao and Bing Liu are joint first authors, and Jing Zhang and Yi Pan are joint corresponding authors of this paper. The paper [97] has been published and can be found at `https://doi.org/10.1007/978-3-319-59575-7_28` .

4.1 Introduction

Understanding functional localizations and exploring functional interactions within the brain is an ongoing challenge in the area of neuroscience [68][52]. Among several methods, neuroimaging is an efficient way to achieve this task. Functional magnetic resonance imaging(fMRI) is a functional neuroimaging method [92]. By quantifying blood flow using MRI technology, fMRI data can be used to measure human brain activities [92][22][51].

In the recent years, multiple neuroscience researches on neuronal network-level activities using fMRI dataset have invoked increasing number of attentions [101][66]. Modeling functional connectivity and abrupt boundaries among regions of interest(ROIs) in fMRI data has been generally considered as a powerful way to investigate brain functional interactions [67][101][64][57].

Inspired by the successes of signal analysis technologies, sliding time window, short-time Fourier transform(STFT), and wavelet transform(WT) based schemes are proposed

to detect brain functional dynamics in recent literatures [1][101][5][23]. Unfortunately, these frameworks are limited due to some difficulties such as how to decide the length of the sliding window. Sparse representation is widely used in applications such as image processing, audio processing, and document analysis [68][96][56][71][104][54]. Recently, there are several researches that utilize sparse representation for fMRI signal analysis, and some excellent results are achieved. However, sparse representation is based on the typical assumptions that the neural integration of those components is linear and the components of each voxels fMRI signal are sparse [68]. Generally, these basic assumptions might not be easy to reach for fMRI data analysis. From a technical perspective, several other very popular tools for signal analysis can also be developed for exploring brain functional interactions, such as Markov random field (MRF) models, principal component analysis (PCA), independent component analysis (ICA), autoregressive spatial models [67][68][54][65].

In some recent studies, several Bayesian inference based methods have been proposed for exploring brain functional dynamics. These methods are less sensitive to fMRI noises and more reliable in estimating functional interactions [80]. Lian, et.al proposed a Bayesian Connectivity Change Point Model(BCCPM) to detect change points(defined as abrupt boundaries of functional interactions in brain networks) in fMRI [43]. It uses Bayesian inference method to calculate the probability of multivariate time series while some certain time points are assumed as change points among fMRI data. It obtains significant performance. However, it is based on Markov Chain and Monte Carlo (MCMC) strategy and not easy to be speeded up with multi-processor computers or GPU.

In this chapter, in order to develop a higher accuracy algorithm that can be parallelized in multi-processor computers or GPU in the future realization, we investigate most Bayesian inference based methods for exploring brain functional interactions [46][43][101][45][31][80][44][97]. By combining Bayesian connectivity change point model, we propose a modified genetic algorithm to optimize the evolutionary procedure to improve the detection accuracy and decrease the time consumption. We test GA-Based-BCCPM on several synthesized datasets, and the experimental results verify that our new model

produces results with higher accuracy in less time. Also, we apply the new model to real block-designed task-based fMRI dataset, and excellent results are obtained.

4.2 Modified Genetic Algorithm and Bayesian Inference Theory

In this section, we mainly describe the modified genetic algorithm and Bayesian connectivity change point model(BCCPM). The modified genetic algorithm will be utilized as an optimization strategy and Bayesian inference theory will be used to calculate fitness of each individual in genetic algorithm.

4.2.1 Modified Genetic Algorithm

As aforementioned, we need an optimization method to achieve better performance such as higher accuracy and lower time consumption. Moreover, the potential parallelized realization is an important factor we may concern.

Genetic algorithm (GA) is a kind of evolutionary method inspired by the process of natural selection [50][69]. Commonly, we represent the possible solutions of a problem with indicator vectors called individuals. Then, we evolve the initial individuals to the next generation using selection, crossover, and mutation operators.

Actually, genetic algorithm may utilize very different selection, crossover, and mutation strategies for different problems. To effectively explore brain functional interactions, we design a series of special selection, crossover, and mutation operators as follows,

Selection and crossover:

Step 1 Randomly produce an integer $n \in 1 \sim n_0$.

Step 2 Randomly produce two different integers a_1 and $a_2 \in 1 \sim n$, and select the a_1-th and a_2-th individuals in the sorted i-th generation.

Step 3 Randomly produce two different integers b_1 and $b_2 \in 1 \sim N$ as the selected positions.

Step 4 Crossover the selected individuals at the selected positions.

Mutation:

Step 5 Randomly produce a floating point number $u \in [0, 1]$.

Step 6 if $u > u_0$, go to Step 8.

Step 7 Randomly produce integers $c_1, c_2, , c_s \in [0, N]$.

Step 8 Change the c_1-th, c_2-th, ... , and c_s-th position from 1 to 0 or from 0 to 1.

Step 9 If all the individuals have been generated, stop; otherwise, go to Step 1.

Figure 4.1 [98] illustrates the structure and flowchart of the modified genetic algorithm. The main procedures of the algorithm can be summarized as follows:

Firstly, we randomly initialize some binary indicator vectors to represent different distributions of change points. The dimensions of these indicator vectors are the same as the number of time points in the fMRI data. From the view of genetic algorithm, these indicator vectors can be regarded as individuals of the initial population.

We then calculate Bayesian posterior probability and use it as the fitness of everyone.

Finally, we evolve individuals of the current generation toward the next higher fitness generation by a series of modified genetic operators, i.e., special selection, special crossover, and special mutation strategies.

After several evolutionary procedures, individuals in the final generation may have outstanding fitness and the highest fitness individual can represent the most likely change point distribution in the corresponding fMRI data. Thus, the most probable change points distribution could be resolved.

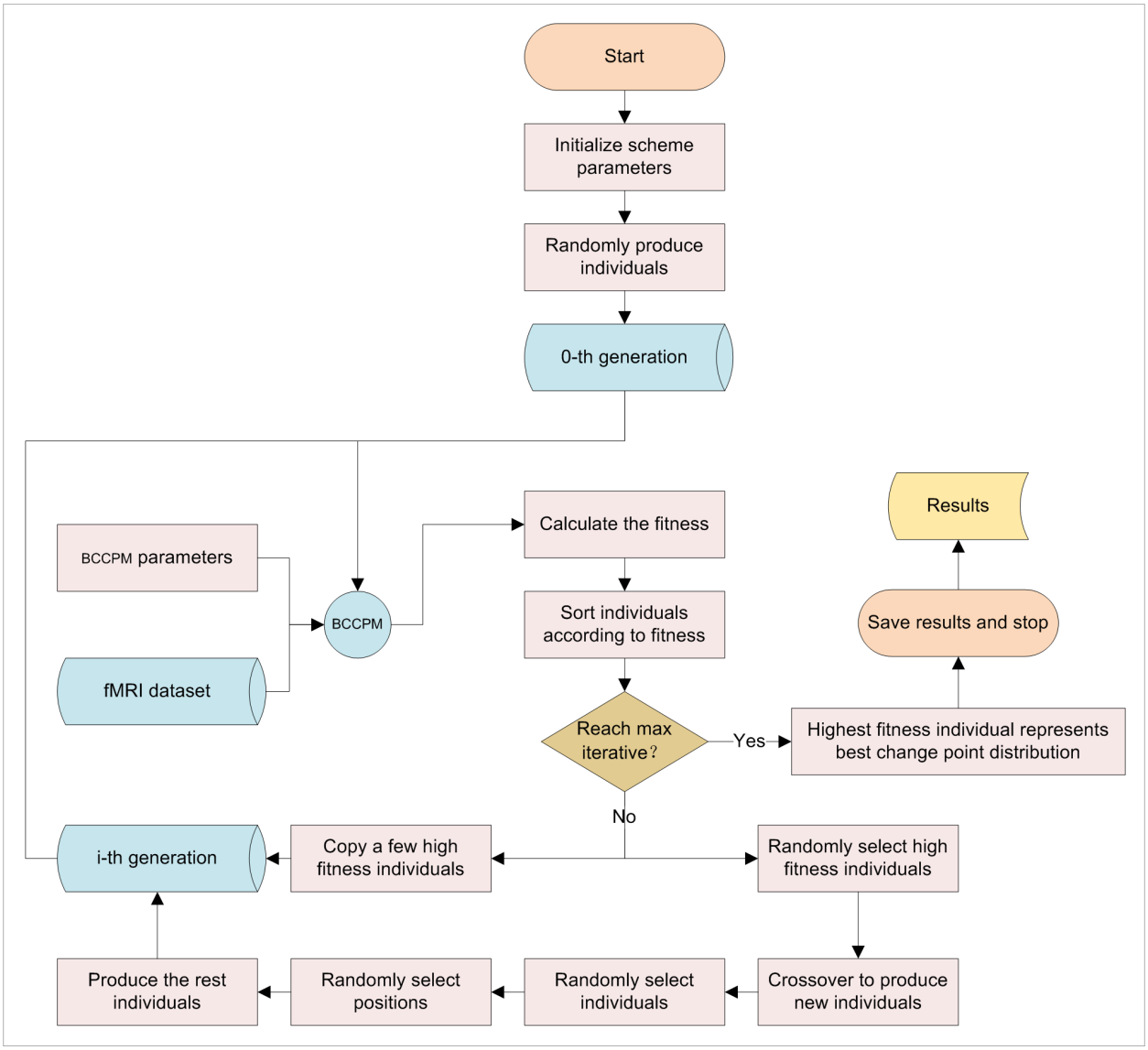

Figure 4.1. Flowchart of the modified genetic algorithm.

Moreover, if we carefully inspect every step illustrated in Figure 4.1, we can see that the modified genetic algorithm is different from a traditional one. Besides typical selection, crossover, and mutation operators, it has a sorting operation and a duplication operation. These two operations exert very important influence on all other operators, and may directly improve the performance. Commonly, the general genetic algorithm may select some excellent individuals to produce the next generation by using evolutionary operators such

as crossover and mutation. However, in our method, we copy a few best individuals directly to the next generation. This strategy could make sure the optimal solution will not be changed by the coming evolutional operators and also could make them exert more influence on generating the new individuals. In fact, we think this strategy is fair because the excellent individuals may have long life or survive long in the natural world. Figure 4.2 [97] illustrated the influence more intuitively. Carefully inspect Figure 4.2 (b) to (c), we can see that all the evolutional operators used to produce individuals from the i-th generation to (i+1)-th generation should use the sorted fitness results.

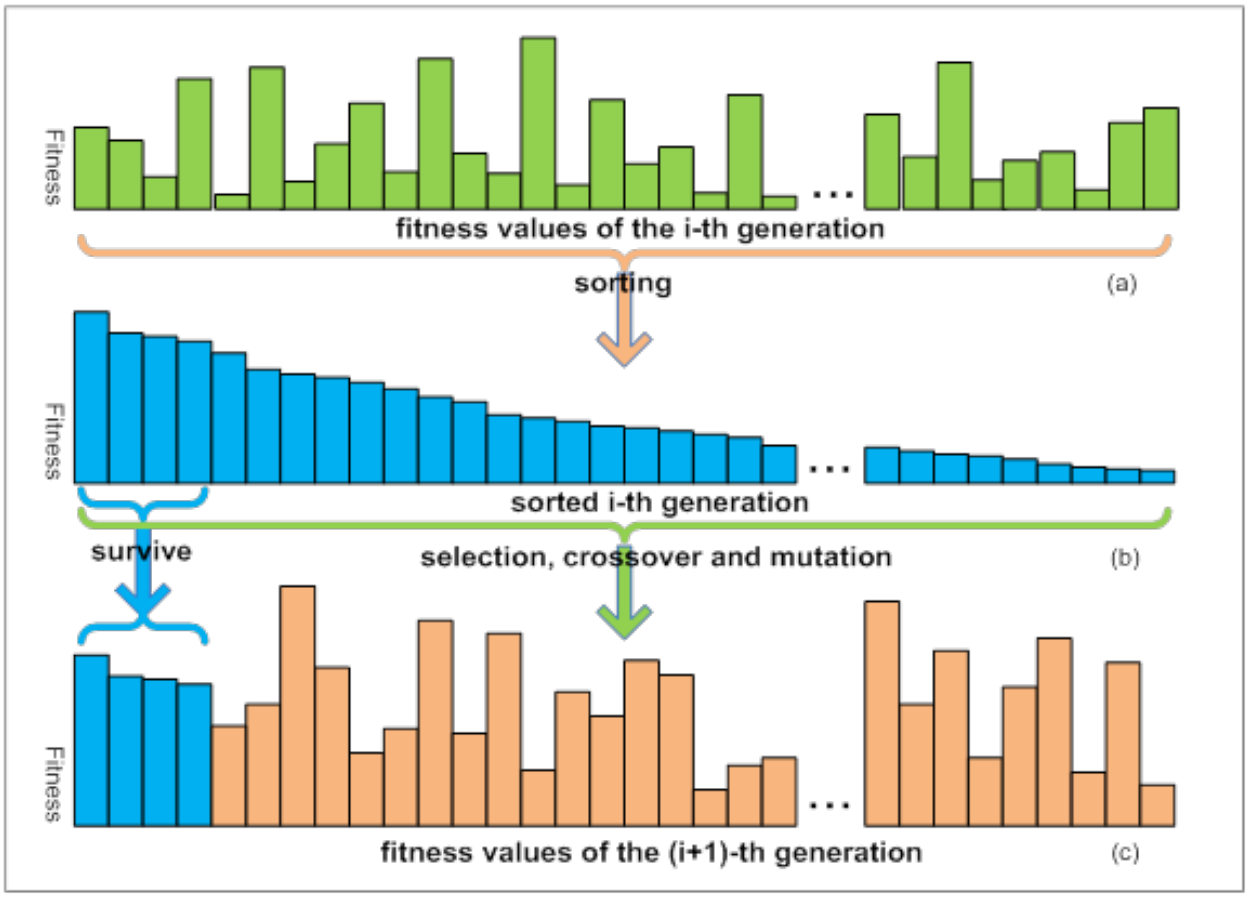

Figure 4.2. The fitness values influence: sorted fitness values exert more influence on evolutional operators in each iteration.

4.2.2 Bayesian Connectivity Change Poing Model

In this subsection, we will review the Bayesian connectivity change point model (BC-CPM) again, which was firstly proposed by Lian et al in literature [43]. We will use Bayesian inference theory to calculate fitness of each individual in genetic algorithm.

Given an $R \times T$ dataset $X = (x_1, x_2, ..., x_T)$, in which T is the number of observations and R is the number of ROIs, we are interested in if there are some differences in the joint probabilities within these ROIs between different time periods.

We define a block indicator vector as,

$$\vec{I} = (I_1, I_2, ..., I_T) \tag{4.1}$$

where $I_k = 1$ if the k-th observation x_k is a change point at the beginning of a temporal block, $I_k = 0$ otherwise.

Now, suppose a set of vectors $x_1, x_2, ..., x_T$ are i.i.d. (independent and identically distributed) from R-dimensional multivariate normal distribution, i.e. $x \sim N(\mu, \Sigma)$, $t = 1, 2, ..., T$, where T denotes the number of vectors, R denotes the dimension of vectors, μ denotes the R-dimensional mean vector, and Σ denotes the $R \times R$ covariance matrix. The conjugate prior distribution of (μ, Σ) is N-Inv-Wishart [28], and the posterior distribution of (μ, Σ) based on the data $X = (x_1, x_2, ..., x_T)$ is also N-Inv-Wishart [28]. Therefore, we can calculate the probability of as follows,

$$P(x_1, x_2, ..., x_T) = \frac{P(x_1, x_2, ..., x_T; \mu, \Sigma)}{P(\mu, \Sigma | x_1, x_2, ..., x_T)} = \left(\frac{1}{2\pi}\right)^{RT/2} \left(\frac{\kappa_0}{\kappa_T}\right)^{R/2} \frac{\Gamma_R(\frac{\nu_T}{2})}{\Gamma_R(\frac{\nu_0}{2})} \frac{(det(\Lambda_0))^{\nu_0/2}}{(det(\Lambda_T))^{\nu_T/2}} 2^{RT/2} \tag{4.2}$$

where $\Gamma_R(z)$ is the multivariate gamma function:

$$\Gamma_R(z) = \pi^{R(R-1)/4} \prod_{j=1}^{R} \Gamma(z + (1-j)/2) \tag{4.3}$$

Consider the block indicator vector in Equation (4.1), the likelihood of the data matrix $X = (x_1, x_2, ..., x_T)$ is:

$$p(X|\vec{I}) = \prod_{b=1}^{\sum I_k} p(X_b) \tag{4.4}$$

where X_b is the temporal observations that belong to b-th block and $p(X_b)$ can be calculated according to Equation (4.2). The temporal blocks are independent from each other; therefore, the posterior distribution of the configuration is:

$$p(\vec{I}|X) \propto p(\vec{I})p(X|\vec{I}) \tag{4.5}$$

where $p(\vec{I}) = \prod_{t=1}^{T} p(I_t)$ and $p(I_t) \sim \text{Bern}(0.5)$.

It is worth noting that Equation (4.4) will be regarded as the fitness function of the proposed genetic algorithm to calculate the fitness of every new individual generated by the evolutionary operators.

4.3 Simulation Study and Real Data Application

In this section, intensive simulation study is carried out to evaluate and validate MCMC-Based-BCCPM[22] and GA-Based-BCCPM on several simulation datasets. Furthermore, the application results on a real emotion processing task-based fMRI dataset are compared and summarized.

4.3.1 Simulation Study I

In the first set of simulation study, we perform several experiments between MCMC-Based-BCCPM and GA-Based-BCCPM for network (a)-(f) illustrated in Figure 4.3 [97]. We repeat every simulation experiments for 5 times and save all results to calculate their average performance. For the purpose of fairness, all parameters of BCCPM are set as same values. The iterative number of GA-Based-BCCPM is set as 100 while GA-Based-BCCPM is set as 20000 to synchronously achieve good convergence and detection results for both the two

methods.

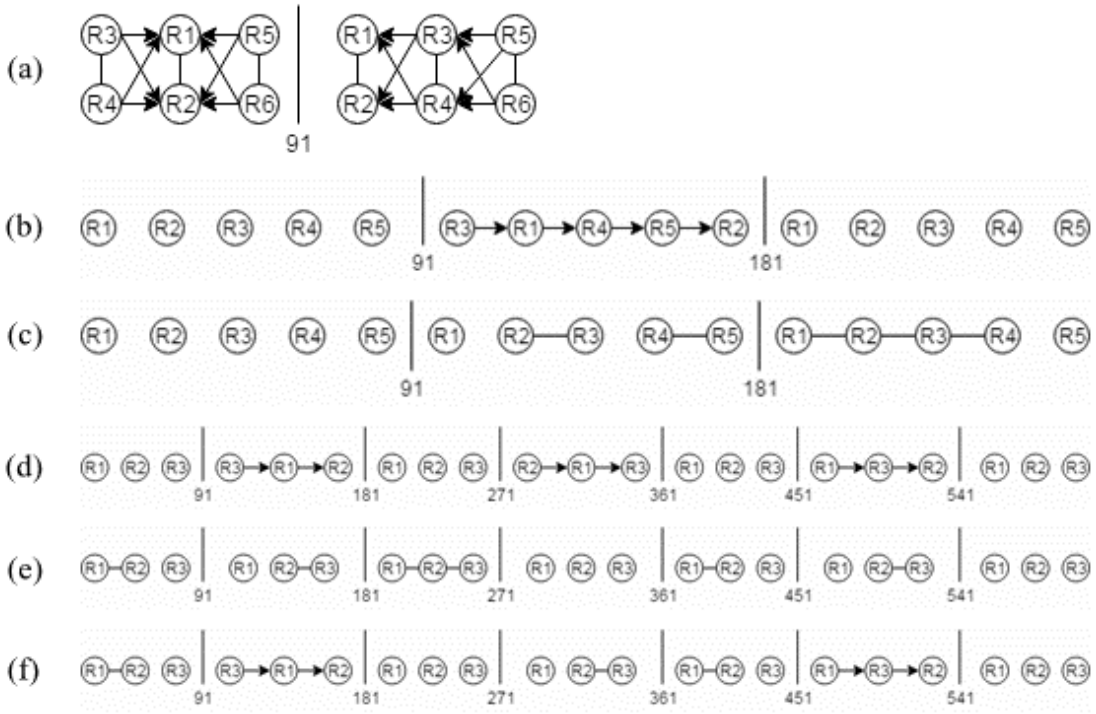

Figure 4.3. Six different structures of dynamic networks and their change-point-distributions.

Figure 4.4 [97] illustrates the convergence curve of the MCMC-Based and GA-Based-BCCPM. Figure 4.4 [97](a)-(f) denote the convergence curves of network (a)-(f), respectively; and the results of MCMC-Based and GA-Based-BCCPM are listed on the left and right column, respectively. Carefully observe the results of left and right columns, we can see the convergence curve of MCMC-Based-BCCPM vibrates even when it has reached its highest peak, while GA-Based-BCCPM only climbs for the highest peak. Obviously, the good results of proposed method may benefit from the modified GA, since it will always copy individuals with highest fitness values directly into the next generation.

Figure 4.5 [97] illustrates Change point detection results for network (a)-(f) of the MCMC-Based and GA-Based-BCCPM. We can observe that GA-Based-BCCPM outperforms MCMC-Based-BCCPM for all six networks in detection precision. Even more, GA-

Based-BCCPM does not miss or mistake change points in any network.

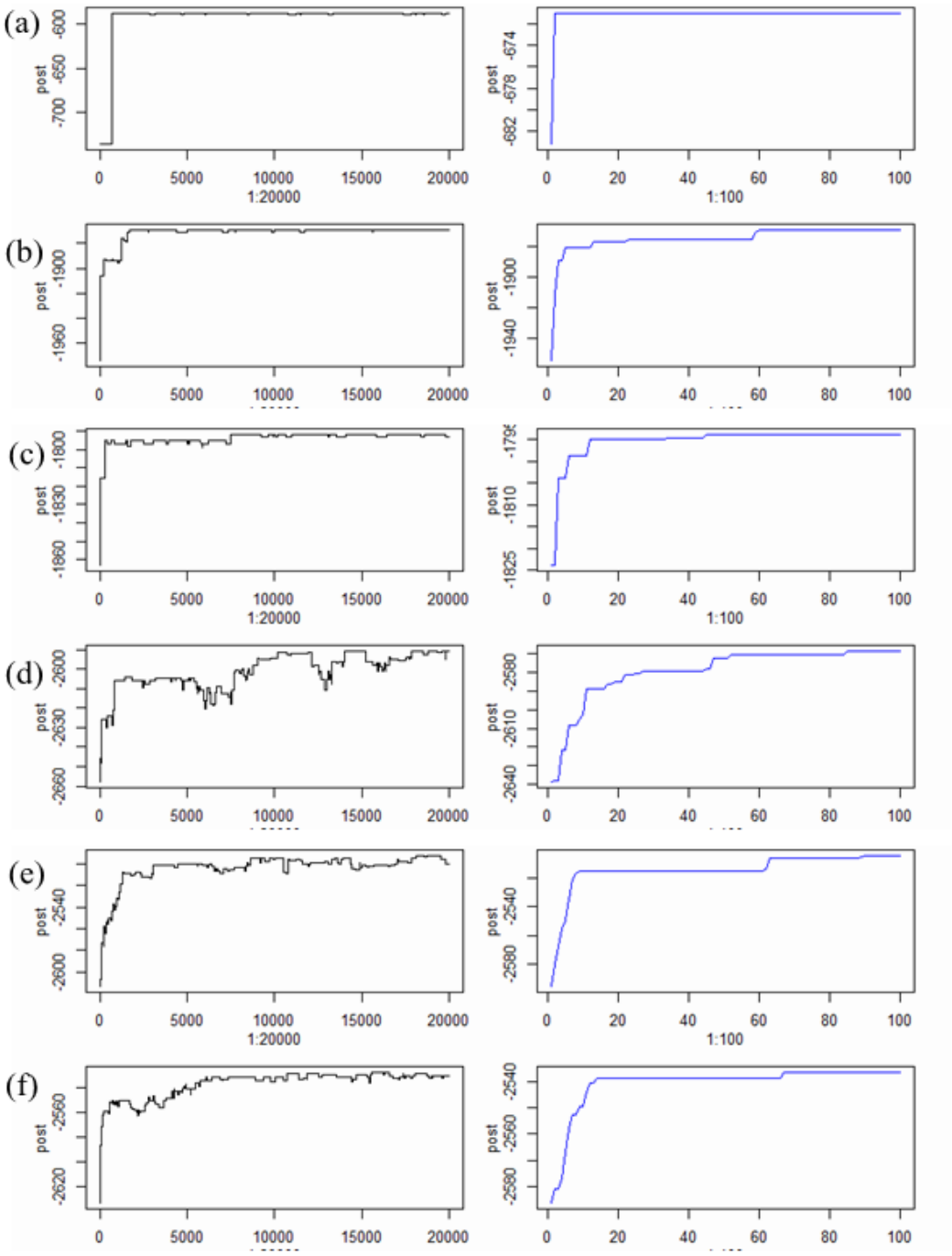

Figure 4.4. Convergence curve for network (a)-(f) (Left column is results of MCMC-Based-BCCPM. Right column is results of GA-Based-BCCPM).

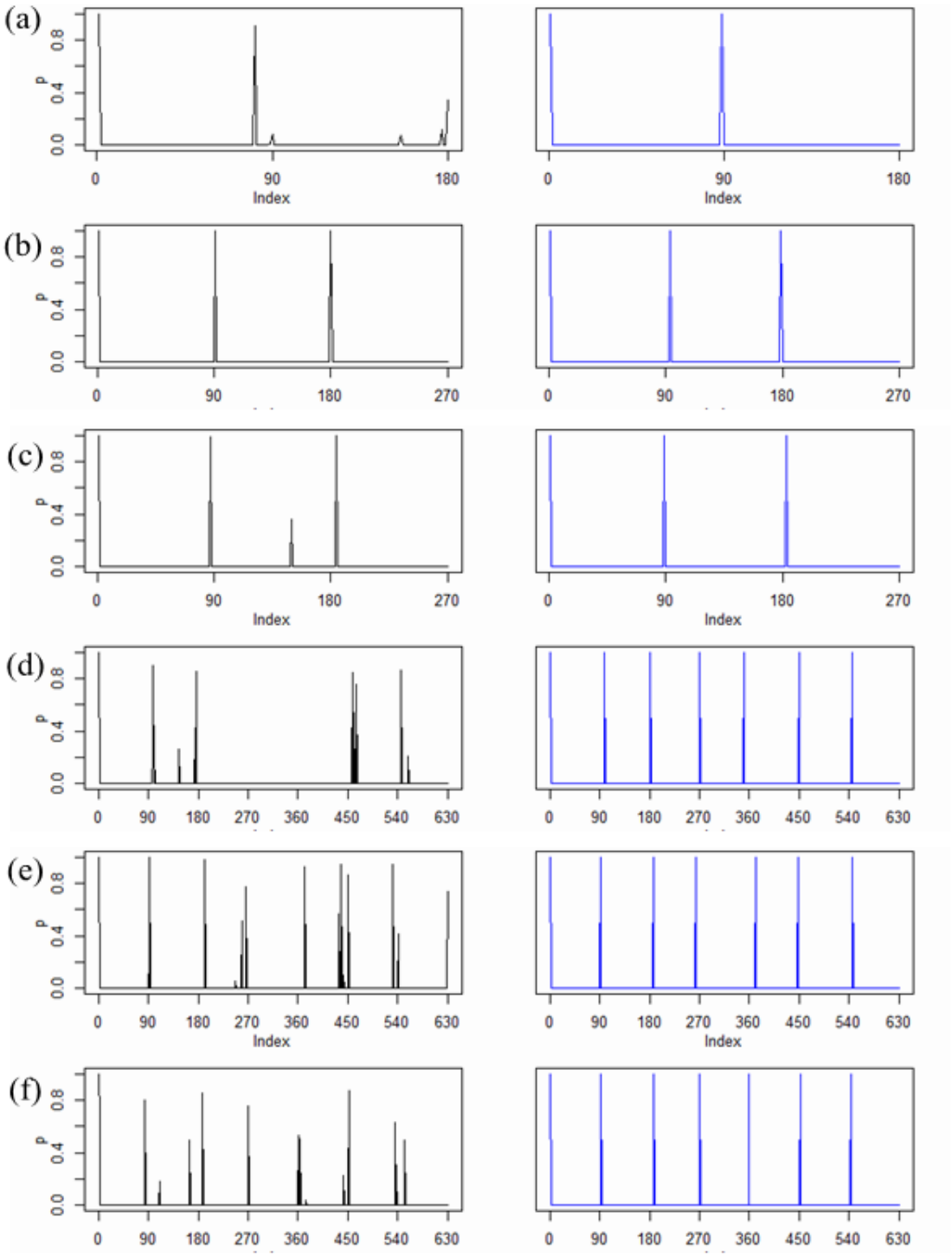

Figure 4.5. Change point detection results for network (a)-(f) (Left column is results of MCMC-Based-BCCPM. Right column is results of GA-Based-BCCPM).

Finally, we take the time consumption into our consideration, the running time of MCMC-Based and GA-Based-BCCPM are listed in Table 4.1 [97]. The environment of our experiment is as follows: operating system: Windows 10 Pro; system type: 64-bit operating system x64-based processor; CPU: Intel(R) Core(TM) i7-6600U CPU@2.6GHz 2.81Hz; memory: 12 GB. The last row of Table 4.1 [97] is the average running time and the other rows are networks (a)-(f). It is very easy to see that GA-Based-BCCPM is faster than MCMC-Based-BCCPM in all networks, and of course the average time consumption.

Table 4.1. Comparison of computational cost (time in ms)

	MCMC-Based-BCCPM	GA-Based-BCCPM
Network (a)	5796.8	3631.2
Network (b)	6206.0	3568.6
Network (c)	7825.2	4391.0
Network (d)	6268.8	4231.4
Network (e)	6218.2	4231.2
Network (f)	6472.0	4256.4
Average	6464.5	4051.6

4.3.2 Simulation Study II

In the second set of simulation study, eight datasets are generated based on eight different structures of dynamic networks (Figure 4.6 (a - h)) to testify that GA-Based-BCCPM can effectively detect the change points [98]. There are various types of interaction patterns in the eight different structures of dynamic networks to make sure the applicability of the model. Multiple experiments are completed to validate the proposed model. The designed change points of eight different dynamic networks are shown in Figure 4.6 by using a vertical solid line with a position number below it.

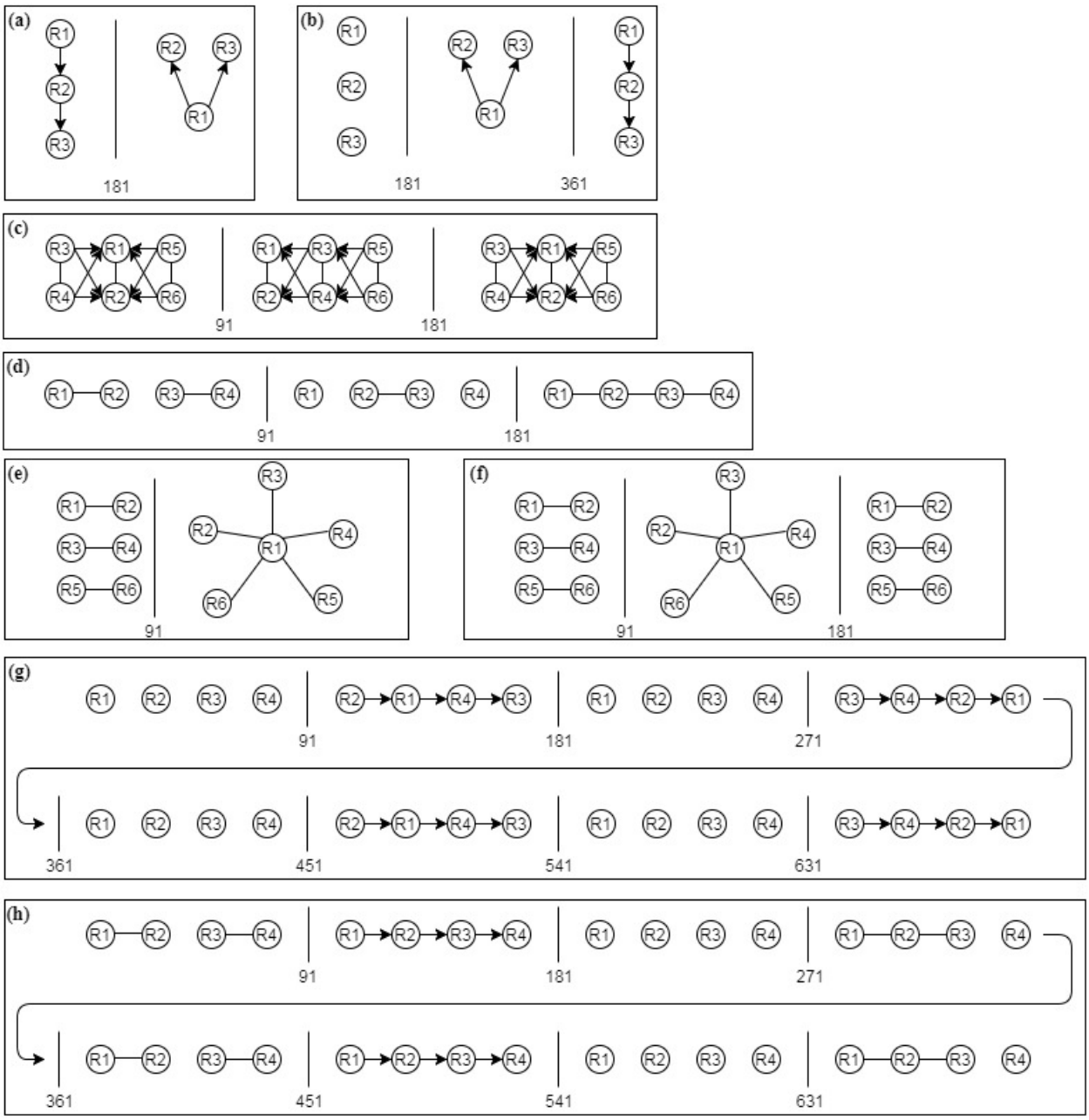

Figure 4.6. Eight different structures of dynamic networks and their change-point distributions.

For each of the eight structures of dynamic networks, we repeat the experiments for 50 times and all the results are saved to calculate their average performance. The number of repetitions of GA-Based-BCCPM is set as 100 and the number of repetitions of MCMC-

Based-BCCPM is set as 20000 to achieve good convergence and change points detection results for both methods. Note that all parameters of BCCPM are set at the same values for the sake of fairness.

The simulation results are summarized in Figure 4.7 [98] and Figure 4.8 [98] to compare the convergence curves and change point detection results. In Figure 4.7 [98], the convergence curves of the two methods are displayed on the left and right columns respectively from network (a) to network (h). We can observe that there are some vibrations within the highest peaks of the convergence curves in MCMC-Based-BCCPM. However, thanks to the advantage of the proposed modified genetic algorithm, which copies the subjects with highest fitness directly to the next generation, GA-Based-BCCPM performs better as it always climbs to its peak monotonically. In Figure 4.8 [98], the change point detection results of the two methods are shown on the left and right columns respectively from networks (a) to (h). It is obvious that GA-Based-BCCPM has better performance than MCMC-Based-BCCPM, as we can observe the proposed GA-Based-BCCPM catches all the designed change points precisely, while MCMC-Based-BCCPM mistakes some positions in networks (b), (c), (d), (f), and (h).

On top of that, the computational cost of each method is again taken into our consideration. The running time (in milliseconds (ms)) of MCMC-Based-BCCPM and the proposed GA-Based-BCCPM are recorded and summarized in Table 4.2 [98]. We can easily see that the proposed GA-Based-BCCPM has lower computational cost in all eight networks. Note that our computing environment is: Windows 10 Pro; Processor: Intel Core i7-4770K CPU @ 3.50GHz; Installed RAM: 16.0 GB; System type: 64-bit operating system, x64-based processor.

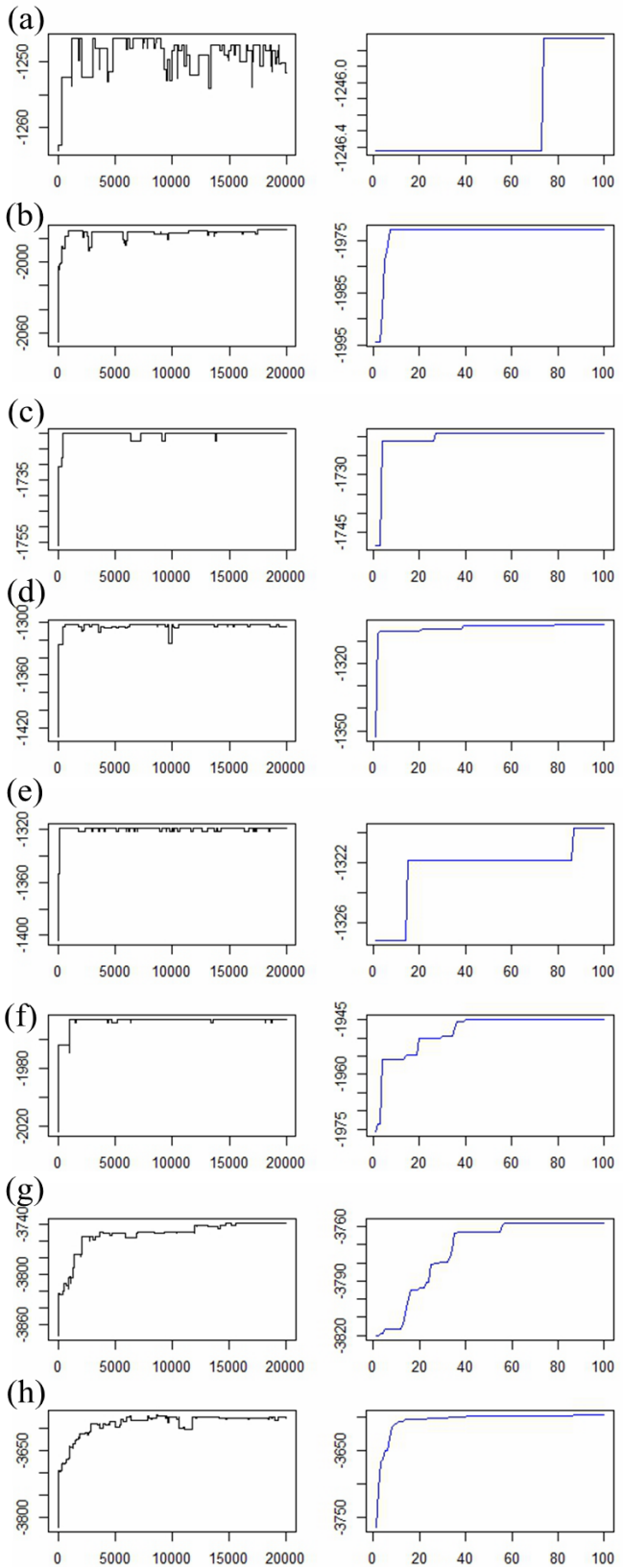

Figure 4.7. Convergence curves for networks (a)-(h) (Left column shows results from MCMC-Based-BCCPM. Right column shows results from GA-Based-BCCPM).

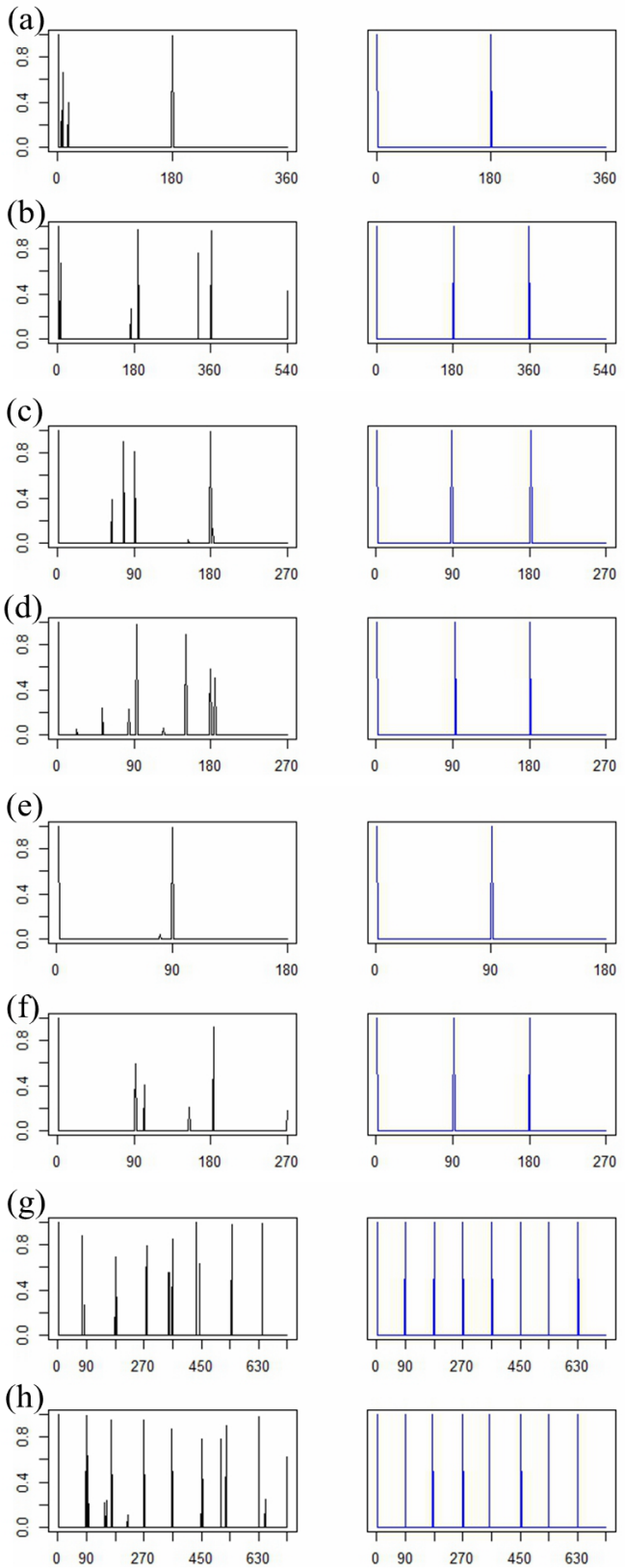

Figure 4.8. Change point detection results for network (a)-(h) (Left column shows results from MCMC-Based-BCCPM. Right column shows results from GA-Based-BCCPM).

Table 4.2. Comparison of computational cost (time in ms)

	MCMC-Based-BCCPM	GA-Based-BCCPM
Network (a)	3572.2	3168.6
Network (b)	5284.6	3618.8
Network (c)	8687.4	4468.6
Network (d)	5000.2	2790.4
Network (e)	6009.4	3206.4
Network (f)	8784.6	4525.0
Network (g)	7890.2	5706.0
Network (h)	8218.8	5728.2

4.3.3 Real Data Application

In this study, we applied the GA-Based-BCCPM on a real emotion processing task-based fMRI data and compared the results with MCMC-Based-BCCPM. These 9 fMRI datasets are acquired from an Emotion Processing task on a Siemens Skyra 3T scanner [35]. The total scan length was 176 frames (i.e. time points). The datasets were preprocessed by authors in [43] using the open source DICCCOL tools in [105]. From each of the scanned subjects brain, fMRI time series are extracted from 358 DICCCOL Region of Interest (ROIs) . Details of the data processing can be found in [43][35]. The resulting size of each dataset is 358*176, where 358 is the total number of ROIs and 176 is the total number of frames (i.e. time points). The application results are summarized in Table 4.3 [98]. Similar to [43], we define change points that are detected within 5 time points as aligned with boundaries, and change points that are detected within 10 time points as partially aligned. From the results, we can see that our proposed GA-Based-BCCPM detects change points more precisely than MCMC-Based-BCCPM among most of the 9 subjects. On average, out of the 7 expected change points, MCMC-Based-BCCPM detects 75% (aligned and partially aligned), and GA-Based-BCCPM detects 81%(aligned and partially aligned).

Additionally, change points that are not aligned with any boundaries are also shown in Table 4.3 [98]. The average number of not aligned change points detected by MCMC-Based-BCCPM is 2.44, while those detected by GA-Based-BCCPM is 1.67. It also indicates that the GA-Based-BCCPM has better performance compared with MCMC-Based-BCCPM. Furthermore, these not aligned change points could be used to infer the dynamic changes happened within the designed blocks.

Table 4.3. Results on real data application

Subject	Model	# of change points detected	# of change points aligned with boundaries	# of change points partially aligned with boundaries	# of change points not aligned with any boundaries
1	MCMC-Based	7	1	2	4
	GA-Based	7	5	1	1
2	MCMC-Based	7	3	2	2
	GA-Based	8	5	2	1
3	MCMC-Based	8	5	1	2
	GA-Based	7	2	2	3
4	MCMC-Based	7	4	1	2
	GA-Based	7	2	3	2
6	MCMC-Based	8	4	2	2
	GA-Based	7	6	1	0
7	MCMC-Based	8	2	3	3
	GA-Based	8	4	3	1
8	MCMC-Based	8	4	2	2
	GA-Based	7	2	2	3
9	MCMC-Based	8	4	2	2
	GA-Based	8	2	4	2
10	MCMC-Based	8	3	2	3
	GA-Based	7	2	3	2

To better understand the functional interactions between the ROIs within the temporal blocks found by our method, the PC algorithm [78] is applied to search the patterns which can represent the underlying causal structure from the data. Since subject 6 has the most aligned change points detected by our method, we will use subject 6 as an example to show

the functional interaction patterns searched by the PC algorithm. In Figure 4.9 [98], the top graph shows that the change points detected by GA-Based-BCCPM align with designed task blocks and only the last one is not aligned per our definition. From these temporal blocks, three of them are labeled as I, II and III, which represent a face block, a shape block and a fixation block in the experiment respectively. The bottom row in Figure 4.9 shows the functional interaction patterns searched by the PC algorithm for I. face block, II. shape block, and III. fixation block. Its obvious that the interaction patterns are different from each other in the two task blocks and the fixation block.

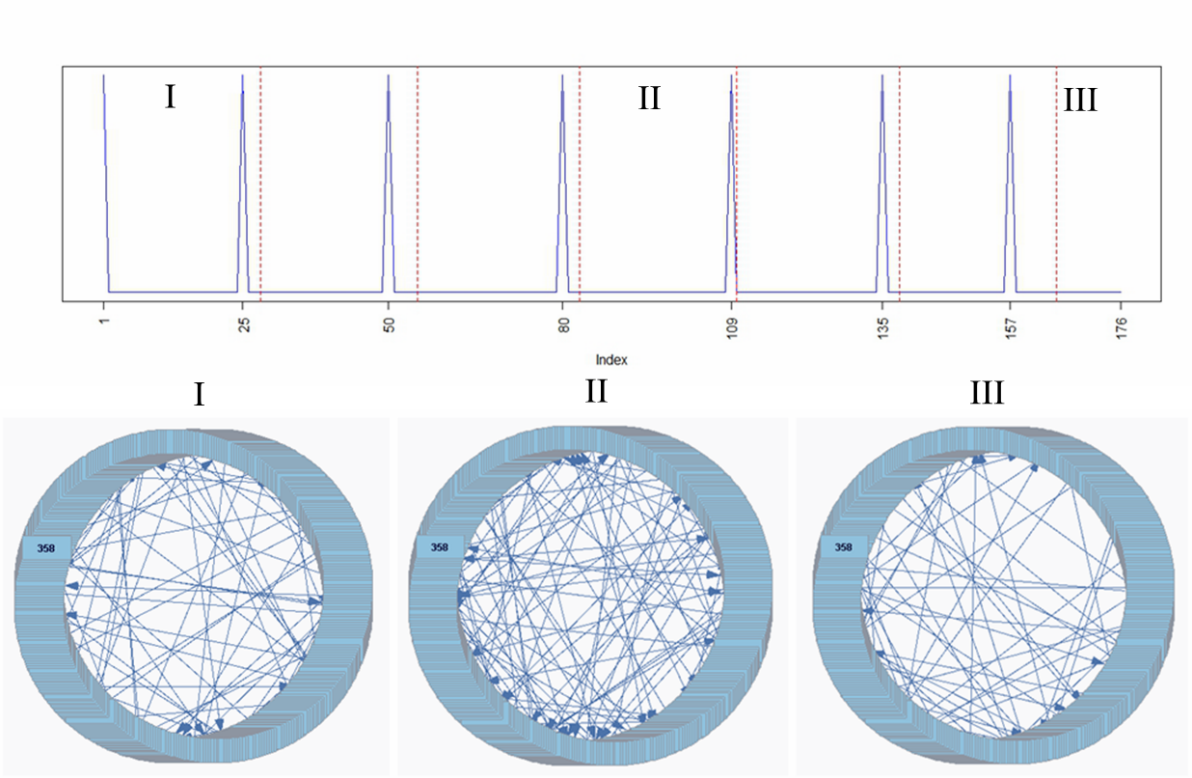

Figure 4.9. Top: Detected change point shown in blue ridges with designed blocks split by red dotted lines of subject 6. Bottom: Functional interaction patterns searched by PC algorithm from three temporal blocks: I. face, II. shape, III. fixation.

4.4 Conclusion and Future Work

Finally, we conclude our method and discuss further work in Section 4. In this chapter, an optimized method for Bayesian connectivity change point model using genetic algorithm is presented. We apply it to detect change point in synthesized and real fMRI data, and excellent results are obtained.

In the future, we will combine our method with other Bayesian inference model, such as BMCPM and DBVPM. Furthermore, for the potential parallel realization of the genetic algorithm, GA-Based-BCCPM could be easily implemented in a parallel mode and run efficiently within GPU or multi-processor computers, thus, we will also investigate its GPU or multi-processor version.

CHAPTER 5

BAYESIAN BI-CLUSTER CONNECTIVITY CHANGE POINT MODEL

Human brain's functional dynamics have been demonstrated with recent studies and detecting the functional connectivity change points of single subject is also carried out by different researches. However, the clustering of multiple subjects with finding change points at the same time is still very challenging. To contribute in this area, this work presents a novel Bayesian bi-cluster change-point model (BBCCPM). This model simultaneously infers the dynamics of functional brain interactions as well as the cluster of different subjects based on the boundaries of temporally quasi-stable blocks. The proposed model analyzes the joint probabilities among multiple subjects with whole brain ROIs between different time-periods and applies the MCMC scheme to sample the posterior probability distribution of each time point as being a change point as well as different subject cluster scheme. Finding the change points can help investigate temporal functional brain dynamics, and grouping them can help us distinguish the differences of the brain dynamics among multiple subjects and lead to further research on finding the reasons behind it. The BBCCPM has been evaluated and validated by experimental datasets and good results are achieved.

This is collaboration work with Dr. Xuan Guo, Dr. Xiuchun Xiao, Dr. Yi Pan, and Dr. Jing Zhang.

5.1 Introduction

In recent year, there are lots of neuroimaging researches on functional Magnetic Resonance Imaging (fMRI) data. In particularly, various Bayesian-inference-based methods have been designed to detect magnitude or functional connectivity change points and furthermore led to infer the interaction patterns based on the temporal blocks [31]. As discussed in Lindquist M.A., et al. (2007)[48], the detection of functional brain state-related change

points without a known timing information has been an important consideration, as psychological processes could not be specified in advance; at the same time, the onset time and the duration of the response may vary considerably across subjects. Besides functional connectivity, the study on functional network connectivity, which focuses on the interactions in network level, estimates the clusters of brain regions having similar functionalities and has been applied to many diseases to examine brain network differences between healthy and diseased brains [9][3][106].

However, we are also very interested in clustering of multiple subjects with finding change points at the same time, and this is still very challenging. In previous studies, multivariate graphical causal models based on Bayesian networks are more robust and reliable in estimating functional interactions and less sensitive to noise in the fMRI signals [73], and Lian, et al. has developed a novel Bayesian Bayesian Connectivity Change Point Model (BCCPM) [43] to detect the change points by finding the boundaries of temporal blocks via a unified Bayesian framework via the analysis of the dynamics of multivariate functional interactions. In this work, we present a novel Bayesian bi-cluster change-point model (B-BCCPM). This model simultaneously infers the dynamics of functional brain interactions as well as the cluster of different subjects based on the boundaries of temporally quasi-stable blocks. The proposed model analyzes the joint probabilities among multiple subjects with whole brain ROIs between different time-periods and applies the MCMC scheme to sample the posterior probability distribution of each time-point as being a change point as well as different subjects' clustering scheme.

5.2 Methodology

5.2.1 Bayesian Bi-cluster Connectivity Change Point Model

Fundamentals of Bayesian inference

For T vectors $y_1, y_2, ..., y_T$ i.i.d (independent and identically distributed) from r-

dimensional multivariate normal distribution

$$y_t \sim N(\mu, \Sigma), t = 1, 2, ..., T$$

where T is the number of vectors, r is the dimension of vector y_t, μ is the r-dimensional mean vector, and Σ is the $r \times r$ covariance matrix. When μ and Σ are both unknown, the prior distribution of (μ, Σ), which is conjugate, is the N-Inv-Wishart$(\mu_0, \Lambda_0/\kappa_0, \nu_0, \Lambda_0)$ [28]:

$$\mu|\Sigma \sim N(\mu_0, \Sigma/\kappa_0)$$
$$\Sigma \sim Inv - Wishart(\nu_0, \Lambda_0)$$

The posterior distribution of (μ, Σ) given the data $y_1, y_2, ..., y_T$ is the same type of N-Inv-Wishart$(\mu_T, \Lambda_T/\kappa_T, \nu_T, \Lambda_T)$, with:

$$\mu_T = \frac{\kappa_0}{\kappa_0 + T}\mu_0 + \frac{T}{\kappa_0 + T}\bar{y}$$
$$\kappa_T = \kappa_0 + T$$
$$\nu_T = \nu_0 + T$$
$$\Lambda_T = \Lambda_0 + S + \frac{\kappa_0 T}{\kappa_0 + T}(\hat{y} - \mu_0)(\hat{y} - \mu_0)^T$$
$$S = \sum_{i=1}^{T}(y_i - \hat{y})(y_i - \hat{y})^T$$

Here S is na $r \times r$ matrix. Then we can calculate the probability of $y_1, y_2, ..., y_T$ as follows:

$$P(y_1, y_2, ..., y_T) = \frac{P(y_1, y_2, ..., y_T; \mu, \Sigma)}{P(\mu, \Sigma|y_1, y_2, ..., y_T)} = \left(\frac{1}{2\pi}\right)^{rT/2}\left(\frac{\kappa_0}{\kappa_T}\right)^{r/2}\frac{\Gamma_r(\nu_T/2)}{\Gamma_r(\nu_0/2)}\frac{(det(\Lambda_0))^{\nu_0/2}}{(det(\Lambda_T))^{\nu_T/2}}2^{rT/2} \tag{5.1}$$

where Γ_r is the multivariate gamma function:

$$\Gamma_r(z) = \pi^{r(r-1)/4}\prod_{j=1}^{r}\Gamma(z + (1-j)/2)$$

Baysian bi-cluster connectivity change point model

Given an $r \times T$ data matrix $Y = (y_1, y_2, ..., y_T)$, in which T is the number of observations and r is the number of ROIs as shown in the sample matrix (Figure 5.1), we are interested in if there are some differences in the joint probabilities within the r ROIs between different time periods and the locations of change points from each other.

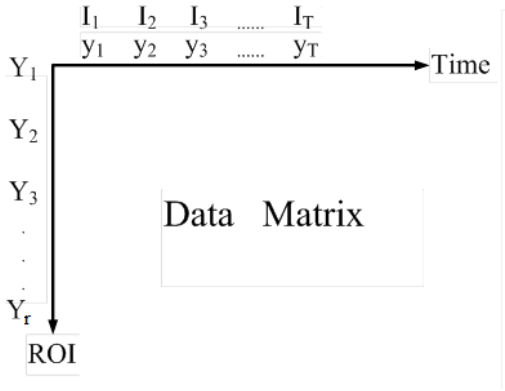

Figure 5.1. Data matrix of Y and block indicator vector I, where y_t are the values of all ROIs at time t (the t-th column in the matrix), Y_j are the values of the j-th ROI at all times (the j-th row in the matrix) and I_t is a block indicator (identifying the change points) at time t

Now we define a block indicator vector:

$$\vec{I} = (I_1, I_2, ..., I_T)$$

in which $I_t = 1$ if the t-th observation y_t is a change point, $I_t = 0$ otherwise. Then the T observations would be divided into $\sum_{t=1}^{T} I_t$ blocks, in which the starting time point I_1 is always considered as a change point. The marginal likelihood of the data matrix $Y = (y_1, y_2, , y_T)$ can be represented as follows:

$$p(Y|\vec{I}) = \prod_{b=1}^{\sum I_t} p(Y_b) \tag{5.2}$$

where Y_b are the observations belonging to b-th block and $p(Y_b)$ can be calculated according to Equation (5.1). We assume statistical independence among the blocks. Therefore, the posterior distribution of $p(\vec{I}|Y)$ can be easily obtained by:

$$p(\vec{I}|Y) \propto p(\vec{I})p(Y|\vec{I})$$

where $p(\vec{I}) = \prod_{t=1}^{T} p(I_t)$ and $p(I_t) \sim \text{Bern}(0.5)$.

Likelihood of multiple subjects.

Now say for N subjects data matrix $\vec{Y} = (Y^{(1)}, Y^{(2)}, ..., Y^{(N)})$, $Y^{(n)}$ is $r \times T$ ROI data matrix for the n-th subject. Based on Equation (5.2), assuming all the subjects follow the same block partition, the marginal likelihood of $\vec{Y}$ is,

$$p(\vec{Y}|\vec{I}) = \prod_{b=1}^{\sum I_t} \prod_{n=1}^{N} p(Y_b^{(n)}) \tag{5.3}$$

where $Y_b^{(n)}$'s are the observations belonging to b-th block of the n-th subject and $p(Y_b^{(n)})$ can be calculated according to Equation (5.1).

Clustering of subjects based on block partition.

The main purpose of the proposed method is to cluster multiple subjects into different groups, and subjects within each group follow the same dynamics of functional brain interaction (i.e. the same block partition). Please note that the total number of groups of subjects S is unknown but between 1 and N. Figure 5.2 illustrate an example of the proposed idea, where the first two subjects belong to one group and the third sub-ject belongs to a second group.

We define a specific representation of clustering group structures: Each subject receives a unique label: 1,2,...,N. Each clustering group of subjects also obtains a unique label which is the lowest label among all the subjects it contains. For example, given the clustering group structure illustrated in Figure 5.2, the Group 1 (G_1) containing subjects 1 and 2 is

labeled 1 and the Group 2 (G_2) consisting of subject 3 is labeled 3. There is no group label 2 for such grouping structure. This representation of group labels makes sure each distinct group structure will receive a unique labeling. Therefore the same clustering group structure cannot be labeled in two different ways.

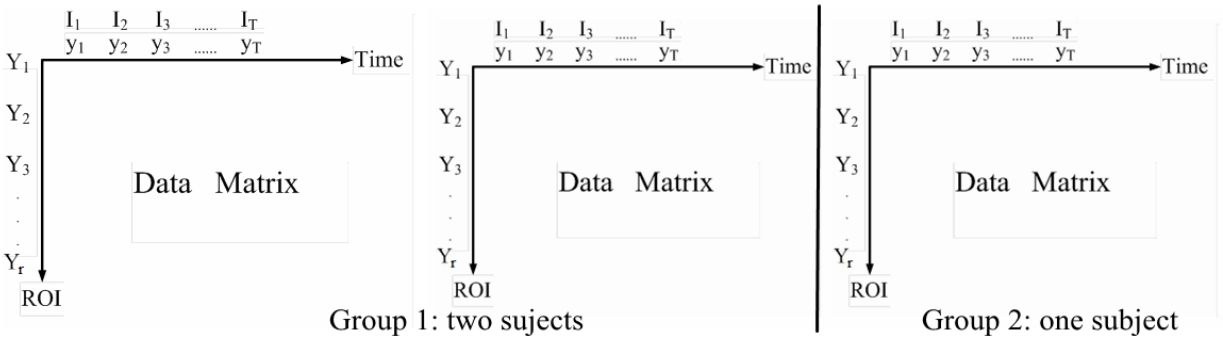

Figure 5.2. Three subjects in two clustering groups.

Based on the above representation of group labels, we define a subject indicator vector $\vec{J} = (J_1, J_2, ..., J_N)$ where J_n denotes the label of grouping which the n-th subject belongs to. Each subject only belongs to one group, meaning there is no overlapping between different clustering groups. Still taking the above clustering group structure of 3 subjects in Figure 5.2 as an example, its subject indicator vector is $\vec{J} = (1, 1, 3)$ according to the above clustering. The 3 subjects are clustered into two groups.

Recall that we totally have S (unknown) groups, thus we define a block indicator matrix $\mathbf{I} = (\vec{I}_{G_1}, \vec{I}_{G_2}, ..., \vec{I}_{G_S})$ where $\vec{I}_{G_s}$ is a block indicator vector for s-th clustering groups G_s as defined previously. Therefore, in each clustering group G_s, all the subjects are segmented into multiple blocks by the block indicator vector $\vec{I}_{G_s}$.

Given a subject indicator vector $\vec{J}$ and a block indicator matrix $\mathbf{I}$, the marginal likelihood of the data $\vec{Y} = (Y^{(1)}, Y^{(2)}, ..., Y^{(N)})$ is:

$$p(\vec{Y}|\vec{J}, \mathbf{I}) = \prod_{s=1}^{S} p(\vec{Y}_{G_s}|\vec{J}, \vec{I}_{G_s}) \tag{5.4}$$

where $\vec{Y}_{G_s}$ is the data of all the subjects belongs to s-th clustering group G_s and the likelihood $p(\vec{Y}_{G_s}|\vec{J}, \vec{I}_{G_s})$ can be calculated by Equation (5.3) given the block indicator vector $\vec{I}_{G_s}$. Different clustering groups are independent, so the posterior distribution is

$$p(\vec{J}, \mathbf{I}|\vec{Y}) \propto p(\vec{J}, \mathbf{I})p(\vec{Y}|\vec{J}, \mathbf{I}) \tag{5.5}$$

where $p(\vec{J}, \mathbf{I}) = p(\vec{J})p(\mathbf{I}) = p(\vec{J}) \prod_{s=1}^{S} p(\vec{I}_{G_s})$, and we use independent uniform priors for $p(\vec{J})$ and $p(\vec{I}_{G_s})$. By substituting Equation (5.4) into Equation (5.5), we have

$$p(\vec{J}, \mathbf{I}|\vec{Y}) \propto p(\vec{J}, \mathbf{I}) \prod_{s=1}^{S} p(\vec{Y}_{G_s}|\vec{J}, \vec{I}_{G_s}) \tag{5.6}$$

5.2.2 Two-level MCMC Scheme

In this section, a two-level MCMC scheme [42] is proposed to sample from the posterior distribution of the cluster grouping structures and block boundaries within each clustering group, as illustrated in Figure 5.3. The lower level MCMC samples from the posterior distribution of the block boundaries within each clustering group, and the higher level MCMC samples from the posterior distribution of the clustering group structures. The posterior distribution of the configuration can be evaluated using the formula (Equation (5.3) and (5.6)) described previously.

In the higher level MCMC, three proposals for updating subject indicator vector $\vec{J}$ are: 1) randomly selecting one clustering group and dividing it into two smaller clustering groups; 2) randomly selecting two clustering groups and merging them together; 3) randomly selecting two subjects and switching their clustering group memberships.

In the lower level MCMC, within each clustering group, there are also three proposals for updating block indicator vector $\vec{I}_{G_s}$: 1) randomly selecting one block and dividing it into two smaller blocks; 2) selecting two consecutive blocks and merging them together; 3) randomly selecting one block and shift its boundary to the left or the right.

In the higher level MCMC, after proposing a new subject indicator vector, we need to

re-label the subject indicator to satisfy our specific representation of clustering groups. For example, given $\vec{J} = (1, 1, 3)$, we randomly select two subjects 1 (the first) and 3 (the third), and switch their clustering group memberships and have $\vec{J} = (3, 1, 1)$. After re-labeling, a new group indicator vector $\vec{J}^* = (1, 2, 2)$ is generated.

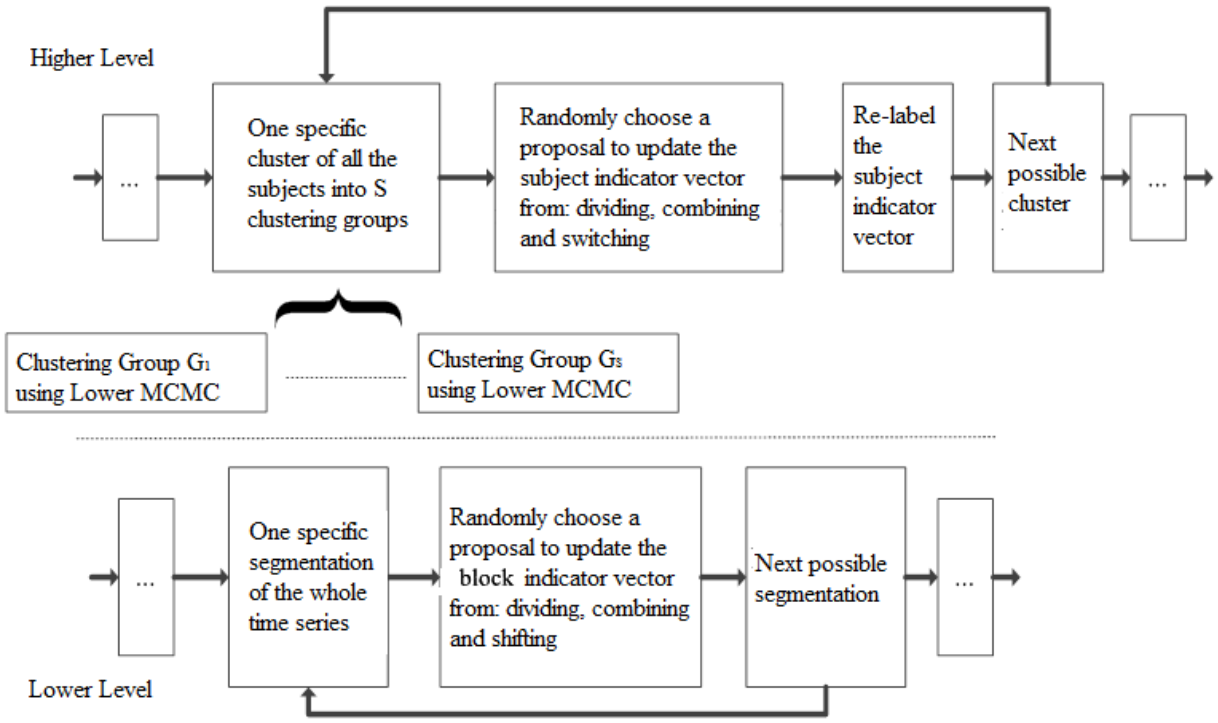

Figure 5.3. Work flow of the two-level MCMC scheme.

5.3 Experimental Results

In this section, we are going to validate our proposed Bayesian bi-cluster change point model on experimental datasets.

The experimental datasets includes five subjects with two clustering groups: Subject 1 and 2 in one group, and Subjects 3, 4, and 5 in another group. Figure 5.4 shows the change point distributions and clustering structure. The first group has 3 temporal blocks with change points at 101 and 201, and the second group has 2 temporal blocks with change

point at 151; both groups have 3 ROI's and 300 time points.

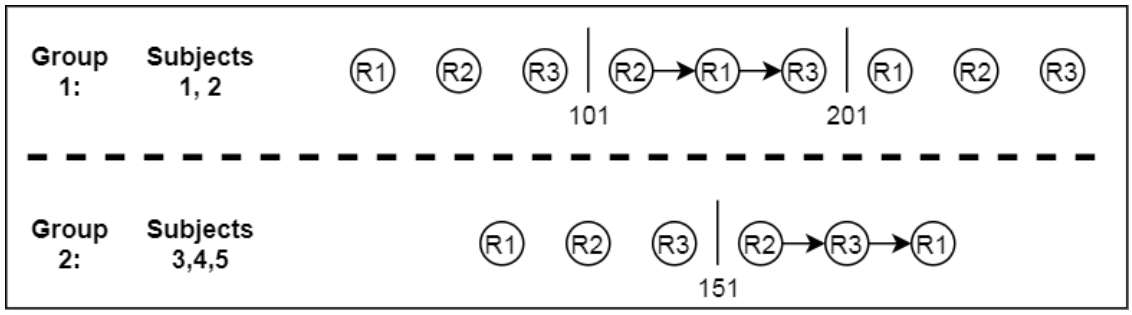

Figure 5.4. Experimental design: Five subjects with two group clusterings and their change point distributions.

The repetition of the lower MCMC is set at 500 and the repetition of upper MCMC is set at 200 (larger numbers for these two repetitions are expected for more complicated change point distributions and group clusters). The results are good as the proposed method detects the change points correctly, and at the same time, clusters the five subjects into two groups. Figure 5.5 shows the overall convergence trace plot of the 2-level MCMC. Figures 5.6 - 5.9 shows the convergence trace plots of the lower MCMC in the last repetition of the upper MCMC, and the corresponding change point detection results in group 1 and 2.

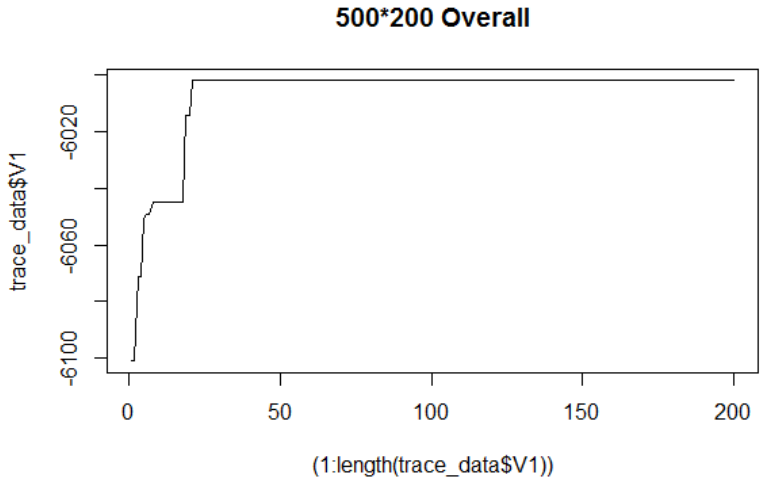

Figure 5.5. Overall convergence curve.

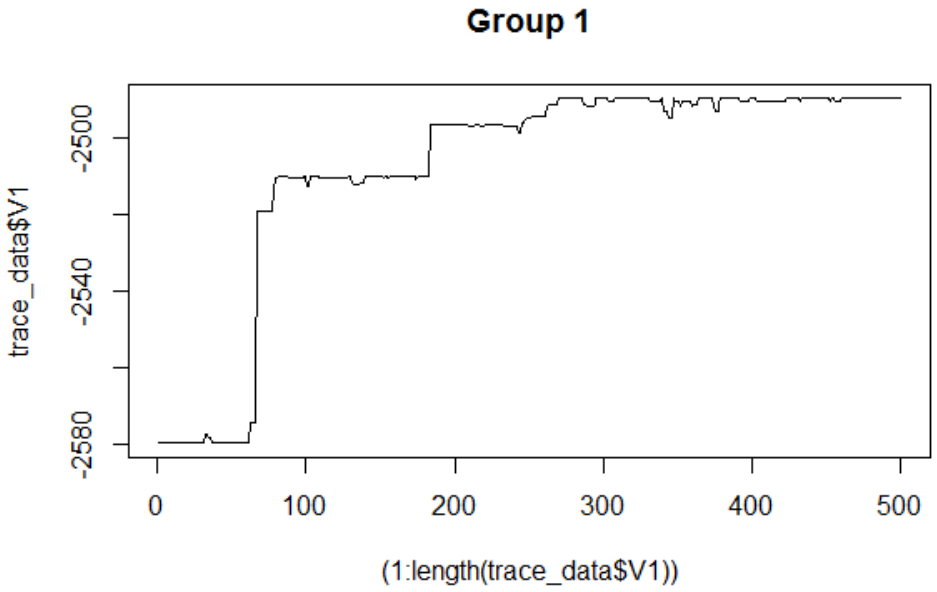

Figure 5.6. Convergence curve for lower level MCMC in Group 1 (Subjects 1,2).

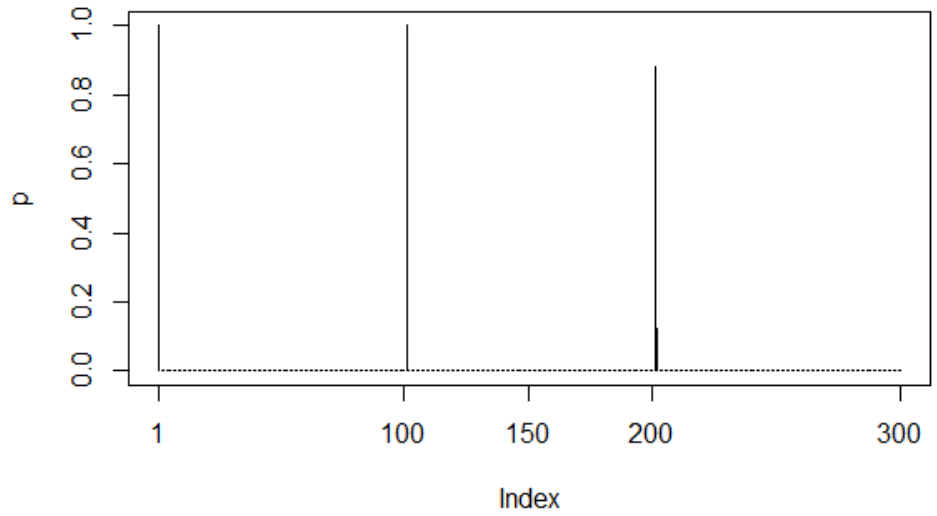

Figure 5.7. Change point detection result for Group 1 (Subjects 1,2).

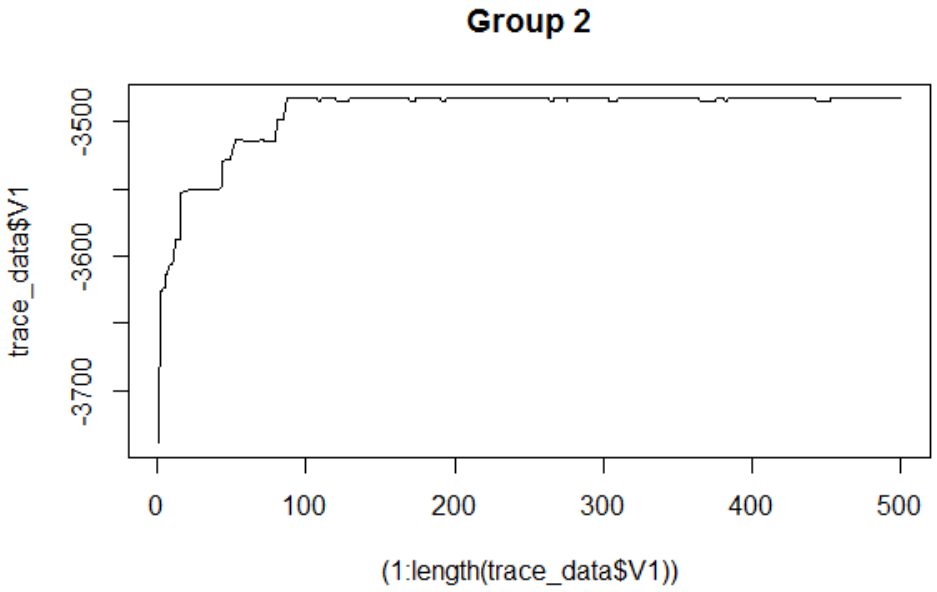

Figure 5.8. Convergence curve for lower level MCMC in Group 2 (Subjects 3,4,5).

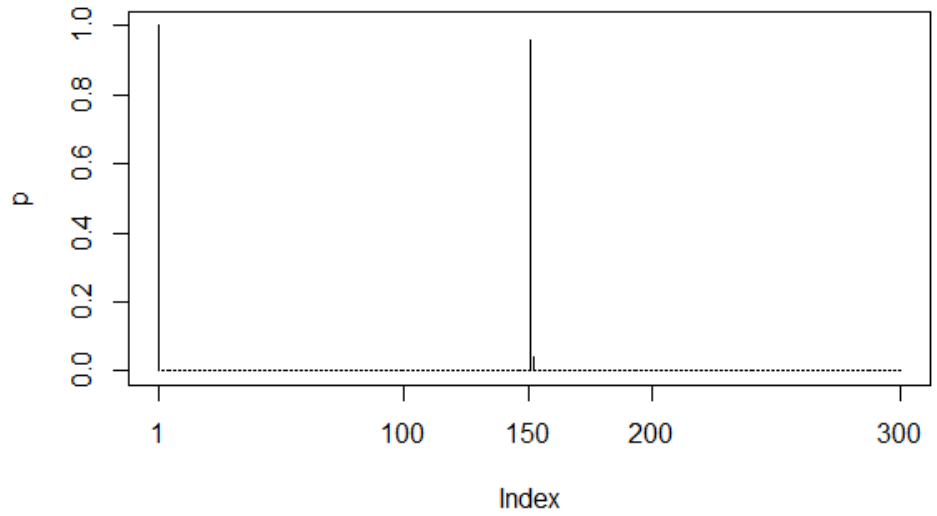

Figure 5.9. Change point detection result for Group 2 (Subjects 3,4,5).

The result of the subject indicator vector is $\vec{J} = (1, 1, 3, 3, 3)$, which correctly clusters the five subjects into two groups, the first two in group 1 and the later three subjects in group 2. The probability of getting this subject indicator is calculated as 100% without burn-in period.

5.4 Conclusion and Future Work

In this chapter, we present a novel Bayesian bi-cluster connectivity change-point model (BBCCPM), which can simultaneously infer the dynamics of functional brain interactions as well as the cluster of different subjects based on the boundaries of temporally quasi-stable blocks. The method has been evaluated on a set of simulated datasets and good results are obtained. In the future, the method may be applied to study the brain dynamics in different group of people and has the potential to classify healthy and diseased patients according to their brain dynamics. As this method utilizes a two-level MCMC scheme, it takes a lot of more time than one-level MCMC. So the modified genetic algorithm may also be employed to achieve good results with less computational cost.

CHAPTER 6

OTHER TOPICS: BAYESIAN ANALYSIS OF COMPLEX MUTATIONS IN HBV HCV AND HIV STUDIES

The material in this chapter is from the author's review article [49]. It's a collaborate work with Shishi Feng, Dr. Yi Pan and Dr. Jing Zhang. The article has been submitted and under review.

In this article, we aim to provide a thorough review of the Bayesian-inference-based methods applied to Hepatitis B virus (HBV), Hepatitis C virus (HCV) and the human immunodeficiency virus (HIV) studies with a focus on the detection of the viral mutations and various problems which are correlated to these mutations. Although these interacting mutation patterns are extremely difficult to efficiently uncover and interpret, the use of Bayesian statistical modeling provides an unprecedented opportunity to solve these problems. Here we summarize a novel statistical approach, the Bayesian Variable Partition (BVP) model, and the Recursive Model Selection (RMS) procedure, which are designed to detect the mutations and to further infer the detailed dependence structure among the interactions. The BVP and RMS in which Markov Chain Monte Carlo (MCMC) methods are used have been widely applied in a number of HBV, HCV and HIV studies in the recent years. We will also provide a summary of the Bayesian methods applications toward these viruses studies, where several important and useful results have been discovered. We envisage the applications to other infectious diseases and cancer cells of more modified Bayesian methods will be following with important medical results before long.

6.1 Introduction

Per historical data, there are up to 30 million people across the world who are infected with Hepatitis B virus (HBV) and up to 600 thousand die every year [33][95]. Among

infected adults, less than 5% of otherwise healthy persons who are infected as adults will develop chronic infection, and $20\% - 30\%$ of adults who are chronically infected will develop cirrhosis and/or liver cancer; and rate is higher in younger populations: $80\% - 90\%$ of infants infected during the first year of life develop chronic infections, and $30\% - 50\%$ of children infected before the age of 6 years develop chronic infections [95]. There are about 1/3 of chronic infected subjects will have irreversible liver damage and it lead to cirrhosis and hepatocellular carcinoma; and the other 2/3 infected subjects will retain the virus in their body and become highly infectious though asymptomatic [61]. In total, up to 25% of subjects with chronic infected HBV die from the complications due to the disease [61].

HBV is a member of the Hepadnaviridae family, and it comprises an icosahedral protein capsid surrounding the viral DNA, with a lipoprotein viral envelope [8][94][59][88]. The virus DNA is organized in 4 open reading frames (ORF): S, which stands for surface and it encodes HBsAg; C, which stands for core and it encodes HBcAg and HBeAg; P, which stands for polymerase and it encodes DNA polymerase; and X, which encodes an X protein, whose precise function is currently unclear [94][88]. Two highly immunogenic proteins consist the nucleocapsid, they are HBcAG and HBeAg, and a less immunogenic surface antigen HBsAg is in the viral envelop [94][59][88]. In patients with chronic hepatitis B, serum HBV-DNA reflects the disease progression as well as the transition across the different stages of the disease [60]. Identifying HBsAg mutations correlated with different levels of serum HBV-DNA in HBV chronically infected patients naive to anti-HBV drugs is one of the interests of HBV studies. In the meantime, Occult HBV infection (OBI) is a threat for the safety of blood-supply, and has been associated with the onset of HBV-related hepatocellular carcinoma and lymphomagenesis. The genetic markers in HBsAg (particularly in D-genotype, the most common in Europe) significantly associated with OBI in vivo are missing, so the correlation between HBaAg-mutations and OBI and its impact on HBsAg detection is also of importance [83]. The above problems can be solved by using Bayesian framework.

Hepatitis C virus (HCV) is a single-strand RNA virus and has been classified into

at least six genotypes with several subtypes in each. The responses patterns of different genotypes to interferon-based therapy are diverse with them spreading in different regions [72]. In previous clinical experience, IFN and ribavirin combined therapy has a significantly higher rate of sustained response in chronic HCV patients compared with interferon-based therapy which has only less than 20% sustained response [13][58].

Some variations in the HCV sequences have the ability of interfering the effective functioning of IFN-based therapies. Among all these variations, the ones in the NS5A region [55][19] are the main subject in our review. NS5A is a nonstructural protein that can lead to IFN therapy resistance by impacting the function of an important mediator of IFN response called dsRNA dependent protein kinase (PKR) [27][26]. NS5A region has 1344 base pairs linking to 448 amino acid and constitutes several regions: the membrane attachment region (aa 1 - 236), the carboxyl region (aa 237 - 448), and the regions within the carboxyl end which included PKRbd (aa 237 - 302), variable region 4 (V4; aa 310-330), variable region 4 (V3; aa 381-409), the region between V3 and V4 (aa 331-380), and the downstream region of V3 (aa 410-448) [25].

In general, mutations in NS5A region have been proposed to be related to therapy resistance by Enomoto and Sato [20] and other researchers [18][30]. However, the relation between mutations in NS5A region and IFN resistance remains ambiguous because of contradictory results obtained in studies concerning PKR binding domain in NS5A [11]. Thus, a better understanding of the role of NS5A region plays in antiviral resistance to IFN therapy will contribute greatly to the development of treatment strategies against HCV.

The human immunodeficiency virus (HIV) is an enveloped virus with a single-stranded RNA genome and is the cause of the Acquired Immunodeficiency Syndrome (AIDS) which killed more than 20 million people since 1980s [70][40][93]. The replication cycle of HIV-1 virus consists of 13 important steps, beginning with the attachment step and ending with the protease-mediated mutation process. The attachment step marks the entry of virus into host cell by the fusion of membranes of the cell and virus [17][21]. A trimer of gp120 and

gp41 heterodimers forms the only protein envelope on the viral surface. The HIV-1s delivery of genome in to the host cell is an extremely intricate process in which a collaborative interaction of the envelope glycoprotein gp120 with the CD4 receptor and with chemokine receptors is required. The chemokine receptors mainly refer to CC chemokine receptor type 5 (CCR5) and C-X-C chemokine receptor type 4 (CXCR4) [7].

These receptors can be used to classify HIV-1 virus since the ability of virus to use the CCR5 and CXCR4 co-receptor differs from each other. It has been proposed by previous studies that R5-reopic viruses which can only use the CCR5 co-receptor, are the predominant in majority of newly HIV-1 infected patients and are generally responsible for the initial infection. Meanwhile, CXCR4 co-receptor usage is observed more often in advanced stages of disease [7][74]. And among the domains of HIV-1 gp120, the V3 loop is the primary determinant for HIV-1 co-receptor usage [34]. Thus, in order to provide more valuable information for the development of anti-HIV-1 drugs targeting on inhibiting the entry of CCR5-tropic HIV-1 strains into host cell, we keep our focus on defining the V3 genetic determinants and the structural features underlying the ability of HIV-1 to use the CCR5 and CXCR4 co-receptors. Moreover, understanding the detailed interaction mutation patterns related to drug-resistance in V3, is also of great importance to develop effective treatment against HIV [38].

Zhang et al. [100] proposed an innovative method for investigating mutation interactions of HIV after certain drug treatment. This method has been used in detecting genome-wide associations on HBV and HCV as well. In this article, we will provide a thorough review of the Bayesian Variable Partition (BVP) model, the Recursive Model Selection (RMS), and their real applications in HBV, HCV and HIV studies.

6.2 Bayesian Methods in HBV, HCV and HIV Studies

In this section, we will first summarize and generalize the Bayesian statistical models applied to HBV, HCV and HIV studies in terms of finding the virus sequence mutations and the difference in two (or three) different groups of patients. Then a summary of important

and interesting results found by applying these methods will be carried out for HBV, HCV and HIV studies. An introduction to basics of Bayesian inference can be found in Section 2.1 as reference.

6.2.1 Bayesian Variable Partition Model

Zhang et al. (2010) [100] first developed the Bayesian Variable Partition (BVP) model to detect and understand combinatorial mutation patterns responsible for HIV drug resistance. Up to now, this method has been successfully applied in various virus studies.

Generally, suppose we have two data sets in the form of matrices, say $A = [A_1, ..., A_m]$ (of dimension $n_A \times m$) and $B = [B_1, ..., B_m]$ (of dimension $n_B \times m$), respectively (each row is a sequence, each column is a position of amino acid sequence). The number of sequences in two groups are denoted using n_A and n_B, and m denotes the number of positions. On top of that, we establish the following four assumptions for the distribution of the positions from the two groups [24]:

- H1: The identity of the **independent** positions, where group A and group B data share the **same** probability distribution.
- H2: The identity of the **independent** positions, where group A and group B data have **different** probability distributions.
- H3: The identity of the **dependent** positions, where group A and group B data share the **same** probability distribution.
- H4: The identity of the **dependent** positions, where group A and group B data have **different** probability distributions.

From these hypotheses, we are interested in positions from H2 and H4 particularly. Therefore, we will start with the positions from H2. Given that the position i is from H2, and we assume there are c_i possible values (amino acids) at position i, and for every sequence in group A, we have p_1 for the first value, p_2 for the second, ..., $p_(c_i)$ for the last value, and

$\sum_{j=1}^{c_i} p_j = 1$. Then we can calculate the likelihood for data set A at position i as,

$$P(A_i|p_1, p_2, ..., p_{c_i}, H2) = \prod_{j=1}^{c_i} p_j^{n_j} \tag{6.1}$$

where n_j denotes the number of sequence with the j-th value in A_i. At the same time, we have $p_j^{'}$ for the j-th value in group B, and $\sum_{j=1}^{c_i} p_j^{'} = 1$. So the likelihood for group B at position i is,

$$P(B_i|p_1^{'}, p_2^{'}, ..., p_{c_i}^{'}, H2) = \prod_{j=1}^{c_i} (p_j^{'})^{n_j^{'}} \tag{6.2}$$

where $n_j^{'}$ is the number of sequence with the j-th value in B_i.

Under the assumption of H2, $p_j \neq p_j^{'}$, since we dont know the true values of p_j or $p_j^{'}$, we assume they are random and a Dirichlet prior is applied on them.

$$p \sim Dirichlet(\alpha_1, \alpha_2, ..., \alpha_{c_i}) : P(p_1, p_2, ..., p_{c_i}|H2, \alpha_1, \alpha_2, ..., \alpha_{c_i}) = \frac{1}{B(\alpha)} \prod_{j=1}^{c_i} p_j^{\alpha_j - 1} \tag{6.3}$$

where $B(\alpha) = \frac{\prod_{j=1}^{c_j} \Gamma(\alpha_j)}{\Gamma(\sum_{j=1}^{c_j} \alpha_j)}$, $\alpha = (\alpha_1, \alpha_2, ..., \alpha_{c_i})$ and $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$;

$$p^{'} \sim Dirichlet(\alpha_1^{'}, \alpha_2^{'}, ..., \alpha_{c_i}^{'}) : P(p_1^{'}, p_2^{'}, ..., p_{c_i}^{'}|H2, \alpha_1^{'}, \alpha_2^{'}, ..., \alpha_{c_i}^{'}) = \frac{1}{B(\alpha^{'})} \prod_{j=1}^{c_i} (p_j^{'})^{\alpha_j^{'} - 1} \tag{6.4}$$

where $B(\alpha^{'}) = \frac{\prod_{j=1}^{c_j} \Gamma(\alpha_j^{'})}{\Gamma(\sum_{j=1}^{c_j} \alpha_j^{'})}$, $\alpha^{'} = (\alpha_1^{'}, \alpha_2^{'}, ..., \alpha_{c_i}^{'})$ and $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$.
Then we have

$$P(A_i, p_1, p_2, ..., p_{c_i}|H2) = \prod_{j=1}^{c_i} p_j^{n_j} \times Dirichlet(\alpha_1, \alpha_2, ..., \alpha_{c_i}) = \frac{1}{B(\alpha)} \prod_{j=1}^{c_i} p_j^{n_j + \alpha_j - 1} \tag{6.5}$$

$$P(B_i, p_1^{'}, p_2^{'}, ..., p_{c_i}^{'}|H2) = \prod_{j=1}^{c_i} (p_j^{'})^{n_j^{'}} \times Dirichlet(\alpha_1^{'}, \alpha_2^{'}, ..., \alpha_{c_i}^{'}) = \frac{1}{B(\alpha^{'})} \prod_{j=1}^{c_i} (p_j^{'})^{n_j^{'} + \alpha_j^{'} - 1} \tag{6.6}$$

By integrating out p and p respectively, we get

$$P(A_i|H2) = \int_p P(A_i, p_1, p_2, ..., p_{c_i}|H2)dp = \prod_{j=1}^{c_j} \frac{\Gamma(n_j + \alpha_j)}{\Gamma(\alpha_j)} \frac{\Gamma(\sum_{j=1}^{c_j} \alpha_j)}{\Gamma(\sum_{j=1}^{c_j}(n_j + \alpha_j))} \tag{6.7}$$

$$P(B_i|H2) = \int_p P(B_i, p_1^{'}, p_2^{'}, ..., p_{c_i}^{'}|H2)dp^{'} = \prod_{j=1}^{c_j} \frac{\Gamma(n_j^{'} + \alpha_j^{'})}{\Gamma(\alpha_j^{'})} \frac{\Gamma(\sum_{j=1}^{c_j} \alpha_j^{'})}{\Gamma(\sum_{j=1}^{c_j}(n_j^{'} + \alpha_j^{'}))} \tag{6.8}$$

And then

$$P(A_i, B_i|H2) = P(A_i|H2)P(B_i|H2) \tag{6.9}$$

When under H1, we have $p_j = p_j^{'}$, so we can obtain

$$\begin{aligned} P(A_i, B_i|H1) &= \int_p P(A_i, B_i, p_1, p_2, ..., p_{c_i}|H1)dp = \int_p \frac{1}{B(\alpha)} \prod_{j=1}^{c_i} p_j^{n_j + n_j^{'} + \alpha_j - 1} \\ &= \prod_{j=1}^{c_j} \frac{\Gamma(n_j + n_j^{'} + \alpha_j)}{\Gamma(\alpha_j)} \frac{\Gamma(\sum_{j=1}^{c_j} \alpha_j)}{\Gamma(\sum_{j=1}^{c_j}(n_j + n_j^{'} + \alpha_j))} \end{aligned} \tag{6.10}$$

For hypothesis H4, let's assume there are c possible number of combinations of the dependent positions. Likewise, suppose for every sequence in group A, we have p_1 for the first combination, p_2 for the second combination, ..., p_c for the last combination, and $\sum_{j=1}^{c} p_j = 1$; for every sequence in group B, we have $p_1^{'}$ for the first combination, $p_2^{'}$ for the second combination, ..., $p_c^{'}$ for the last combination, and $\sum_{j=1}^{c} p_j^{'} = 1$. Then, we have:

$$P(\text{dependent positions in A}|H4) = \prod_{j=1}^{c} \frac{\Gamma(n_j + \alpha_j)}{\Gamma(\alpha_j)} \frac{\Gamma(\sum_{j=1}^{c} \alpha_j)}{\Gamma(\sum_{j=1}^{c}(n_j + \alpha_j))} \tag{6.11}$$

$$P(\text{dependent positions in B}|H4) = \prod_{j=1}^{c} \frac{\Gamma(n_j^{'} + \alpha_j^{'})}{\Gamma(\alpha_j^{'})} \frac{\Gamma(\sum_{j=1}^{c} \alpha_j^{'})}{\Gamma(\sum_{j=1}^{c}(n_j^{'} + \alpha_j^{'}))} \tag{6.12}$$

where n_j and $n_j^{'}$ are the numbers of the j-th combination in A and B separately, and then

$$P(\text{dependent positions in A, B}|H4) = P(\text{dependent positions in A}) \times P(\text{dependent positions in B}) \tag{6.13}$$

Now under H3, we have $p_j = p_j^{'}$, so similarly we have:

$$P(\text{dependent positions in A, B}|H3) = \prod_{j=1}^{c} \frac{\Gamma(n_j + n_j^{'} + \alpha_j)}{\Gamma(\alpha_j)} \frac{\Gamma(\sum_{j=1}^{c} \alpha_j)}{\Gamma(\sum_{j=1}^{c}(n_j + n_j^{'} + \alpha_j))} \quad (6.14)$$

We define an indicator vector $I = [I_1, I_2, , I_m]$ to indicate the hypothesis group of m different positions belong to, where $I_i = 1$ means position i is from H1, $I_i = 2$ means position i is from H2, $I_i = 3$ means position i is from H3, and at last $I_i = 4$ means that the position i is from H4.

Then, as we are interested in the inference of I, so we want to find the posterior distribution of I, given the data sets A and B, i.e. $P(I|A, B)$. Applying the Bayes theorem, we obtain:

$$P(I|A, B) = \frac{P(I)P(A, B|I)}{\sum_{\text{all possible I}} P(I)P(A, B|I)} \quad (6.15)$$

Therefore,

$$P(I|A, B) \sim P(I)P(A, B|I) \quad (6.16)$$

Based on H1, H2, H3, and H4, we have

$$P(A, B|I) = \prod_{I_i=1,2} P(A_i, B_i|I_i) \times P(\text{dependent positions from H3}) \times P(\text{dependent positions from H4}) \quad (6.17)$$

In practice, we also need to assume the prior for I. For example, we may assume most positions should be in H1 and H3, then we set $P(I_i = 2) = P(I_i = 4) = 0.01$, and $P(I) = \prod_{i=1}^{m} P(I_i)$.

6.2.2 Bayesian Partition on Dual Usage of Co-receptor Model

To detect and understand genetic and structural features in HIV-1 B subtype V3 underlying HIV-1 co-receptor usage, Chen et al [10] developed a Bayesian Partition on Dual Usage of Co-receptor Model (BPDUCM) to define V3 genetic determinants either independently or interactively associated with the usage CCR5 co-receptor only, CXCR4 co-receptor only

or dual of CCR5/CXCR4 co-receptor.

This method is based on three datasets - CCR5 only, CXCR4 only and dual usage. Suppose there are N_t sequences from CXCR4-using viruses, N_u from CCR5-using viruses and N_w from dual-using. Each sequence is of p-residues long. Let $X = (X_1, X_2, , X_p)$ be the observation of sequences. X_j is a column vector that contains $N = N_t + N_u + N_w$ observations at the j-th position. Set dataset indicator $Y = (Y_1, Y_2, , Y_N)$ represent the status of co-receptor usage of each sequence: $Y_i = 0$ if ith sequence is from CCR5, $Y_i = 1$ if CXCR-4, and $Y_i = 2$ if dual-using. The goal is to describe the complicated relationship between the sequence observations (X) and the dataset indicator (Y). Basically, we partitioned the p positions into K groups according to their relationship to Y. Each of the K groups represent one relationship between X and Y. Denote with $I = (I_1, I_2, , I_p)$ as the group indicator, $I_j = k, (j = 1, .., p$ and $k = 1, , K)$ means j-th position is partitioned into the k-th group. Given Y, we want to infer I when X is observed and when we have p and K as fixed. The likelihood is $P(X|I, Y)$, and the posterior probability is $P(I|X, Y)$, we have

$$P(I|X, Y) \propto P(I|Y)P(X|I, Y) \tag{6.18}$$

assuming I is independent from Y, $P(I|Y) = P(I)$.

6.2.3 Metropolis-Hastings Algorithm

Then the Markov chain Monte Carlo (MCMC) is used to sample from the posterior probability $P(I|A, B)$ (or $P(I|X, Y)$) via the Metropolis-Hastings (M-H) algorithm to infer which variables are associated with the treatment status, group indicators, etc. The procedure of M-H algorithm is as follows [24]:

1. Initialization: randomly assign a starting value $I^{(t)}$ to I, here t=0;

2. Proposal: propose a new I as follows: randomly choose one $I_i^{(t)}$ and change it to other values with equal probabilities, set new I as y;

3. Evaluation: evaluate the posterior. Since the proposal is symmetric, the acceptance probability is $\alpha(I^{(t)}, y) = min\{1, P(I = y|A, B)/P(I = I^{(t)}|A, B)\}$;

4. Update: generate u from Uniform(0,1) and set

$$I^{t+1} = \begin{cases} y, & \text{if } u \leq \alpha(I^t, y) \\ I^{(t)}, & \text{otherwise} \end{cases} \tag{6.19}$$

5. If $t \geq N$, stop; otherwise set t=t+1 and go to step (2) and repeat this procedure. (N is the total number of iterations.)

6.2.4 Recursive Model Selection

In the studies, the next step is to apply the Recursive Model Selection (RMS) [38] procedure to infer the detailed dependence structure among the interacting positions generated by the Bayesian variable partition model. The strategy is to apply a model selection of two cruder models recursively until the data does not support more detailed models. One of the two is the chain-dependence model and the other is the V-dependence model.

Chain-Dependence Model. A set of variables X_G has a chain-dependence model if the index set G can be grouped into three subgroups U, V, and W such that X_U and X_W are independent given X_V, i.e. $X_U \to X_V \to X_W$. The joint distribution of the chain-dependence model is given as

$$p(X_G) = p(X_U)p(X_V|X_U)p(X_W|X_V) = \frac{F(X_V, X_U)F(X_W, X_V)}{F(X_V)} \tag{6.20}$$

where $F(X_V, X_U, ...)$ is the joint probability function of $(X_V, X_U, ...)$. Illustration of chain-dependence model is shown in Figure 6.1 (reproduced according to [38]).

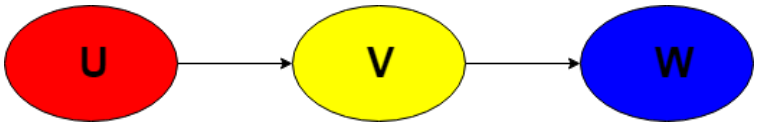

Figure 6.1. The chain-dependence model structure. Colors indicate different sets of variables.

V-Dependence Model. A set of variables X_G has a V-dependence model if the index set G can be grouped into three subgroups U, V, and W such that X_U and X_W are mutually independent, i.e. $X_U \rightarrow X_V \leftarrow X_W$. The joint distribution of the V-dependence model is given as

$$p(X_G) = p(X_U)p(X_W)p(X_V|X_U, X_W) = F(X_U)F(X_W)\frac{F(X_V, X_U, X_W)}{F(X_U, X_W)} \tag{6.21}$$

where $F(X_V, X_U, ...)$ is the joint probability function of $(X_V, X_U, ...)$. Illustration of V-dependence model is shown in Figure 6.2 (reproduced according to [38]).

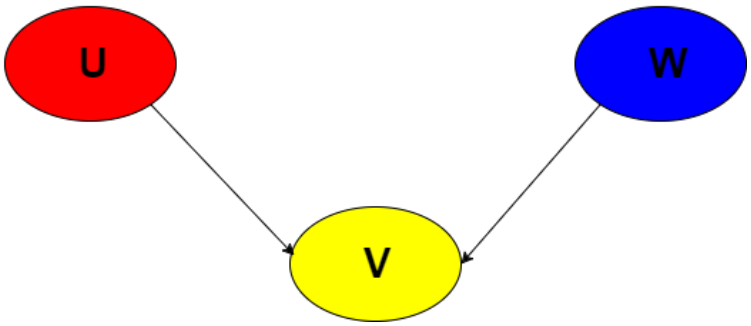

Figure 6.2. The V-dependence model structure. Colors indicate different sets of variables. Notice that variables in U are marginally independent of the variables in W.

Note that in these two models, only set W is allowed to be empty, in which case these models become the saturated model.

We can use a model indicator $I^{CV} = (I_1^{CV}, I_2^{CV}, , I_L^{CV})$ to imply the membership of the L positions with $I_j^{CV} = 0$ representing the chain-dependence model and $I_j^{CV} = 1$ indicating the

V-dependence model. If we use S to denote the set partition, then the posterior distribution of I^{CV} and S is

$$P(S, I^{CV}|\text{data}) \propto P(\text{data}|S, I^{CV})P(S)P(I^{CV}) \tag{6.22}$$

One can set equal priors for I^{CV} and S. Then we can use the MCMC algorithm again to sample from the posterior and find the optimal model type and variable selection. The procedure is applied recursively until only single-variable nodes are available.

Then we can apply BVP and RMS sequentially to the data of the different groups to make inferences on the mutations.

6.3 Applications of Bayesian Methodology to HBV, HCV and HIV studies

6.3.1 Applications in HBV studies

The Bayesian methods described has been applied to multiple HBV related studies including detecting correlation between specific mutations in the C-terminus domain of HBV surface antigen and low level of serum HBV-DNA in patients with chronic HBV infection, HBV amino acid sequence mutations in occult infections, the correlation between HBsAg markers and occult HBV infection and detection. A summary of the results from these studies can be found in Table 6.1. Note that one of the advantages is that the Bayesian-based method showed the ability of analyzing high-order combinations of positions [47].

6.3.2 Applications in HCV studies

By applying Bayesian Variable Partition (BVP) model and Recursive Model Selection (RMS) method to multiple controlled datasets, some interesting findings were discovered to help understanding the HCV drug response and resistance related mutations.

Lets concentrate on NS5A region particularly for HCV genotype 1a in [25]. In NS5A region there are 1344 base pairs, linking to 448 amino acids. The Bayesian methods were applied to the pretreatment sequences of response (47 sequences) and non-response (29 sequences) samples. The result gives us a reliable idea of the mutation mechanism of positions

49, 349, and 199, 209, 242, 398 which have the highest frequencies. Detailed results can be found in Table 6.2.

Table 6.1. A summary of results from applications of Bayesian methods to HBV studies

HBV genotype	**Mutations discovered by Bayesian methods**	**In correlation with**	**Comments**
D (and/or A)	M197T,-S204N-Y206C/H-F220L	serum HBV-DNA<2000IU/ml	These mutations were localized in the HBsAg C-terminus, known to be
D (and/or A)	Y206C/H and/or F220L	lower median (IQR) HBsAg-levels and lower median (IQR) transaminases	involved in virion and/or HBsAg secretion
C (HBV and OBI)	RT mutation V173L	drug resistance in patients receiving antiviral treatments, such as adefovir and lamivudine; HBV vaccine escape	Details results can be found in [82]
C (HBV and OBI)	H126Q, H126Q+138R	OBI samples	
D	20 HBsAg-mutations	occult HBV D-genotype infection in vivo	Details results can be found in [83]

Table 6.2. Single positions result summary

Position	**Result**	**Comments**
49, 349	statistically different in response and non-response patients and are independent of other positions.	Position 49 is in membrane attachment region; Position 349 is in the region between V3 and V4;
199, 209, 242, 398	dependent and demonstrate significant difference in response and non-response patients	Positions 199 and 209 are in membrane attachment region; Position 242 is in ISDR region; Position 398 is in V3 region. These positions may have some biological influence on drug resistance to IFN and ribavirin [25].

While analyzing single positions as above is helpful, a lot of positions are not mutating independently. Figure 6.3 [25] shows the interacting positions detected by BVP in response samples and Figure 6.4 [25] shows the interacting positions detected by BVP in non-response samples. And some significant discoveries can be found in Table 6.3.

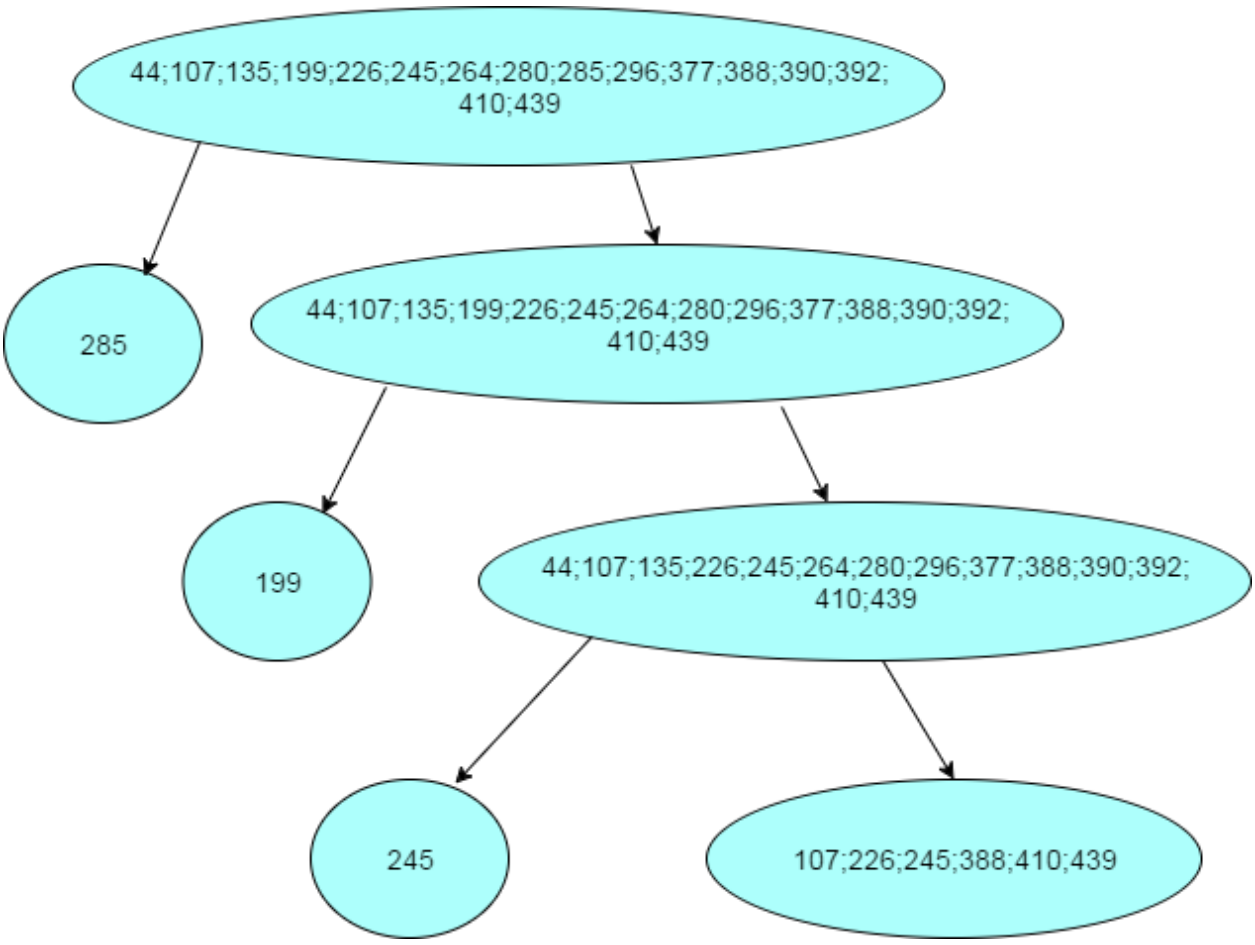

Figure 6.3. Flowchart of detected mutation positions and position combinations in the pretreatment sequence of patients who respond to the treatment.

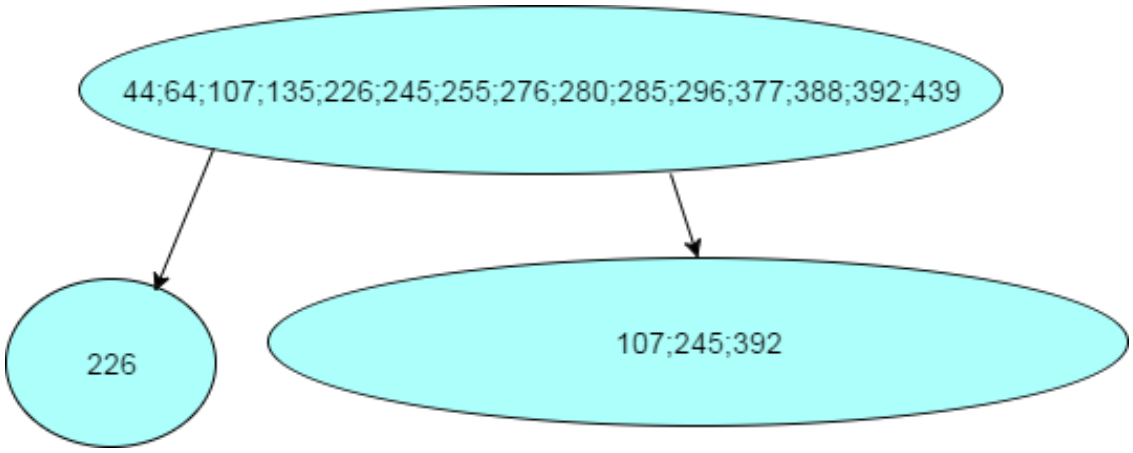

Figure 6.4. Flowchart of detected mutation positions and position combinations in the pretreatment sequence of patients who don't respond to the treatment.

Table 6.3. Dependence structure inferred by RMS in detail

Position	**amino acid(s)**	**Result**	**Comments**
285	E	frequency is 13.8% in non-response samples and 8.5% in the response samples	
199	L	frequency decreases from 100% to 87.2%, from non-response samples to response samples	
226	M	frequency decreases from 20.75 to 14.9%, from non-response samples to response samples	
107, 226, 288, 410, 439	EMIAE	does not exist in response samples	indicates that those positions combined may be a distinguishing factor for response and non-response patients
107, 226, 288, 410, 439	KEIAG, TMVAG, TLIAE	only exist in non-response samples	

6.3.3 Applications in HIV studies

The Bayesian methods summarized in previous chapter has also been successfully applied to multiple HIV drugs in both single-drug treatments and multiple-drug treatments. A compact summary of the results of such Bayesian analysis was carried out in Table 6.4 [100][99]. We can observe that several statistically significant interaction patterns among resistance causing mutations have been discovered using the Bayesian methods. It is of importance that the molecular basis of multiple interacting mutations found by RMS was analyzed with MD simulations and free energy calculations [99]. Therefore, this is another example of the statistical study where biological processes underling drug resistance can be extracted from the discovered independence groups.

Table 6.4. A summary of results from applications in HIV drug resistance studies

Drugs	**Antiretroviral effect**	**Mutation interactions discovered by Bayesian methods**	**Comments**
Indinavir (IDV)	Protease inhibitor	{24,47{32{46⊥54|82}}} {10,71}{73,90}	Interesting group {46,54,82}[1]
Nevirapine	Non-nucleoside RT inhibitor	{106}{188}{103?181}{190}	Weak interactions
Zidovudine	Nucleoside analog RT inhibitor	{41,210,215}{67,219}{70}	Further biochemical investigations needed[2]
IDV, NFV	Protease inhibitors	{24,54,82}{30,88}{73,90}	6 positions disappeared[3]
IDV, SQV	Protease inhibitors	{61,71}{46,54,82}{73,90}	Other details ambiguous
IDV, NFV, SQV	Protease inhibitors	{30,88}{73,90}{24,46,54,82}	Ambiguous structure in 3rd group

Epistatic mutations discovered with BVP approach are partitioned using RMS algorithm. Independence groups are enclosed in brackets. "?" indicates inconclusive result.

[1] Sequential mutation acquisition in this group leads to conditional independence. The results were confirmed by the MD simulations

[2] It is not possible to study the structural basis of mutations using MD simulations for Zidovudine

[3] When compared to single-drug treatment profiles

6.4 Summary and Discussion

In this review article, we presented and summarized three important applications of the Bayesian inference paradigm. First of all, in HBV studies, the evidence has been provided that there exists some specific HBsAg-mutations which correlate with its replicative potential, particularly, the state of low level serum HBV-DNA and HBsAg [60][83][47][82]. At second, several independent HCV-drug-resistance-related mutations and interacting mutation patterns have been detected [25][24]. Moreover, a detailed understanding of complex interacting mutation patterns and new genetic determinants underlying co-receptor usage in HIV-1 have been revealed [100][10][99].

The Bayesian statistical analysis of viral genetic characteristics described above is an advanced and innovative method that can connect statistical modeling with molecular dynamic simulations [24], thus to detect interacting mutations. However, certain significant issues should be addressed in more detailed and be paid more attention to, such as the emergence of bias caused by multiple subpopulations in the data and the decreased sensitivity of the BVP algorithm caused by the transmitted resistance occurrence [38]. Moreover, many factors that may affect the results of the above studies about the three viruses have been ignored since the proposed statistical method is only designed as a baseline analysis. For instance, further studies might be needed to strengthen the correlation between HBsAg mutations and low serum HBV-DNA due to the overlapping of HBsAg and RT genes [60].

Despite all other ignored possibilities, the Bayesian method proposed here has given us several valuable information that will contribute to not only further studies in related areas but also the development of antiviral treatment.

CHAPTER 7

CONCLUSIONS AND FUTURE WORK

In this dissertation, Bayesian methods in brain connectivity change point detection and network dynamics exploration are discussed. These methods include existing methods like BMCPM [46], BCCPM [43], and DBVPM [101], which are state-of-art Bayesian-inference-based methods applied to fMRI data; as well as newly proposed optimized BCCPM with genetic algorithm and bi-cluster connectivity change point model. These successful Bayesian methods can provide good application in brain connectivity exploration and potential usage in studying brain-related diseases.

With the good results achieved from these methods, especially the excellent results obtained from BCCPM in ADHD studies [64], it has been further extended and successfully applied to change point detections in EEG data. This extension of BCCPM helps to test the ability of EEG measurements of frontal and temporo-parietal activity during mindfulness therapy to track response to treatment.

The modified method for BCCPM with genetic algorithm aims to improve the computation efficiency in the MCMC scheme that was used in the original BCCPM with similar or better accuracy. The modified genetic algorithm can optimize the evolutionary process to improve the detection accuracy and at the same time decrease the time consumption in computation. The modified method has been proved to have the claimed advantages through validation results from both simulated experimental designs and real datasets.

In order to cluster/distinguish subjects into different groups (e.g., healthy group and diseased group), the BBCCPM is proposed to simultaneously detect change points in each of the subjects within a group, and cluster subjects into groups according to their change point distributions. Currently, we applied a 2-level MCMC scheme to sample from the posterior distribution to make inference on both the change point detection and clustering. This

method is validated on experimental datasets and achieve good results in both inferences.

To summarize these existing and proposed methods, we present a table with the models' abilities of detecting change points and/or clustering subjects in Table 7.1. In application, if one has simple fMRI data and tries to infer magnitude change points, then BMCPM [46] should be used as it's simple and takes less time to converge than BCCPM [43] and DBVPM [101]; if one wants to detect the functional connectivity change points, we now have two choices: BCCPM and GA-BCCPM, the latter will have better performance when the dimension of the data is larger; if one wants to infer the functional interaction patterns, then DBVPM [101] can be applied; if one wants to cluster subjects according to change point distributions into different groups such as diseased/non-diseased, BBCCPM is a good choice now.

Table 7.1. Summary of existing and proposed models

	BMCPM[46]	BCCPM[43]	DBVPM[101]	GA-BCCPM	BBCCPM
Updating scheme	One-level MCMC	One-level MCMC	Two-level MCMC	Genetic Algorithm	Two-level MCMC
Ability to infer magnitude change points	Yes	Yes	Yes	Yes	Yes
Ability to infer functional connectivity change points	No	Yes	Yes	Yes	Yes
Ability to infer functional ineraction patterns	No	No	Yes	No	No
Ability to cluster subjects according to change point distributions	No	No	No	No	Yes
Computational cost	low	medium	high	low	high

Besides the brain connectivity change point analysis, Bayesian inferences are widely applied in different areas, and one of them is the analysis of complex mutations in viruses such as HBV, HCV, and HIV. We reviewed the Bayesian methods applied to these studies with

a focus on detection the viral mutations and various problems related to these mutations. In particular, BVP and RMS are discussed and results from applications of these two methods on HBV, HCV and HIV are summarized. We believe that the Bayesian methods will contribute to not only further studies in related areas but also the development of antiviral treatment.

In the future, we will continue on several approaches: 1. Redesign and modify the evolutionary process in other change point detection models such as BMCPM, DBVPM and BBCCPM to improve the computational efficiency and detection accuracy. For example, in DBVPM and BBCCPM, 2-level MCMC are used to sample from the posterior distribution of parameters of interest. This makes the estimation process very time-consuming, if we can utilize the good features of genetic algorithm, these models' efficiency will be improved. 2. Further extend these Bayesian methods in EEG data analysis. EEG data has its own advantage that the recording can be done using simple devices like the EPOC/EPOC+ system. With easy accessible data, we can study more interesting brain-related disorders or diseases through the analysis of EEG data. 3. As machine learning and artificial intelligence are very hot topics with potential applications in many different fields, we are also exploring machine learning techniques like auto-encoding and deep learning, and combine them with probability models to have better prediction and clustering performance. In the big picture, we hope the Bayesian inference methods will be applied to more data analysis areas and make contributions to improving human health.

REFERENCES

[1] E. A. Allen, E. Damaraju, S. M. Plis, et al. Tracking whole-brain connectivity dynamics in the resting state. *Cerebral Cortex*, 24(3):663–676, 2014.

[2] J. Anderson. *Cognitive Psychology and Its Implications (Hardcover) (6th ed.)*. New York, NY: Worth Publishers, 2004.

[3] M. Assaf, K. Jagannathan, V. Calhoun, et al. Temporal sequence of hemispheric network activation during semantic processing: a functional network connectivity analysis. *Brain and Cognition*, 70(2):238–246, 2009.

[4] N. A. Badcock, P. Mousikou, Y. Mahajan, et al. Validation of the emotiv epoc((r)) eeg gaming system for measuring research quality auditory erps. *PeerJ*, 1, e38, 2013.

[5] D. S. Bassett, A. Meyer-Lindenberg, S. Achard, et al. Adaptive reconfiguration of fractal small-world human brain functional networks. *Proceedings of the National Academy of Sciences*, 103(51):19518–19523, 2006.

[6] O. M. Bazanova and D. Vernon. Interpreting eeg alpha activity. *Neuroscience & Biobehavioral Reviews*, 44:94–110, 2014.

[7] E. A. Berger, R. W. Doms, E. M. Fenyo, et al. A new classification for hiv-1. *Nature*, 391:240, 1998.

[8] C. R. Bourne, S. P. Katen, M. R. Fulz, et al. A mutant hepatitis b virus core protein mimics inhibitors of icosahedral capsid self-assembly. *Biochemsitry*, 48:1736–1742, 2009.

[9] V. D. Calhoun, K. A. Kiehl, and G. D. Pearlson. Modulation of temporally coherent brain networks estimated using ica at rest and during cognitive tasks. *Human Brain Mapping*, 29(7):828–838, 2008.

[10] M. Chen, V. Svicher, and A. Artese. Detecting and understanding genetic and structural features in hiv-1 b subtype v3 underlying hiv-1 co-receptor usage. *Bioinformatics*, 29(4):451–460, 2013.

[11] M. J. Clemens and A. Elia. The double-stranded rnadependent protein kinase pkr: Structure and function. *Journal of Interferon & Cytokine Research*, 17(9):503–524, 1997.

[12] I. Cribben, R. Haraldsdottir, Atlas L. Y., et al. Dynamic connectivity regression: determining state-related changes in brain connectivity. *Neuroimage*, 61(4):907–20, 2012.

[13] G. L. Davis, R. Esteban-Mur, V. Rustgi, et al. Interferon alfa-2b alone or in combination with ribavirin for the treatment of relapse of chronic hepatitis c. international hepatitis interventional therapy group. *The New England Journal of Medicine*, 339(21):1493–1499, 1998.

[14] M. Duvinage, T. Castermans, M. Peteau, et al. Performance of the emotiv epoc headset for p300-based applications. *Biomedical engineering online*, 12, 56, 2013.

[15] EMOTIV. Emotiv epoc user manual, 2014.

[16] EMOTIV. Emotiv epoc+ - 14 channel wireless eeg headset, 2016. [Online at `https://www.emotiv.com/epoc/`; accessed 26-June-2017].

[17] A. Engelman and P. Cherepanov. The structural biology of hiv-1: mechanistic and therapeutic insights. *Nature Reviews Microbiology*, 10:279–290, 2012.

[18] N. Enomoto, I. Sakuma, Y. Asahina, et al. Comparison of full-length sequences of interferon sensitive and resistant hepatitis c virus 1b. sensitivity to interferon is conferred by amino acid substitutions in the ns5a region. *Journal of Clinical Investigation*, 96(1):224–230, 1995.

[19] N. Enomoto, I. Sakuma, Y. Asahina, et al. Mutations in the nonstructural protein 5a gene and response to interferon in patients with chronic hepatitis c virus 1b infection. *The New England Journal of Medicine*, 334(2):77–81, 1996.

[20] N. Enomoto and C. Sato. Clinical relevance of hepatitis c virus quasispecies. *Journal of Viral Hepatitis*, 2(6):267–272, 1995.

[21] C. Flexner. Hiv drug development: the next 25 years. *Nature Reviews Drug Discovery*, 6:959–966, 2007.

[22] K. J. Friston, P. Fletcher, O. Josephs, et al. Event-related fmri: characterizing differential responses. *Neuroimage*, 7(1):30–40, 1998.

[23] K. J. Friston, P. Jezzard, and R. Turner. Analysis of functional mri timeseries. *Human Brain Mapping*, 1(2):153–171, 1994.

[24] Y. Fu, G. Chen, L. Fu, and J. Zhang. Investigating genotype 1a hcv drug resistance in ns5a region via bayesian inference. *Tsinghua Science and Technology*, 20(5):484–490, 2015.

[25] Y. Fu, G. Chen, X. Guo, et al. Analyzing the effects of pretreatment diversity on hcv drug treatment responsiveness using bayesian partition methods. *Journal of Bioinformatics and Proteomics Review*, 1(1):1–6, 2015.

[26] M. J. Gale, C. M. Blakely, B. Kwieciszewski, et al. Control of pkr protein kinase by hepatitis c virus nonstructural 5a protein: Molecular mechanisms of kinase regulation. *Molecular and Cellular Biology*, 18(9):5208–5218, 1998.

[27] M. J. Gale, M. J. Korth, N. M. Tang, et al. Evidence that hepatitis c virus resistance to interferon is mediated through repression of the pkr protein kinase by the nonstructural 5a protein. *Virology*, 230(2):217–227, 1997.

[28] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin. *Bayesian Data Analysis, 2nd Edition*. Chapman & Hall, London, UK, 2014.

[29] A. Gelman and D.B. Rubin. Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4):457–472, 1992.

[30] M. Gerotto, F. Dal Pero, D. G. Sullivan, et al. Evidence for sequence selection within the non-structural 5a gene of hepatitis c virus type 1b during unsuccessful treatment with interferon-alpha. *Journal of Viral Hepatitis*, 6(5):367–372, 1999.

[31] X. Guo, B. Liu, L. Chen, et al. Bayesian inference for functional dynamics exploring in fmri data. *Computational and Mathematical Methods in Medicine*, 2016(3279050):9, 2016.

[32] M. Hämäläinen, R. Hari, R. J. Ilmoniemi, et al. Magnetoencephalography-theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of Modern Physics*, 65(2):413, 1993.

[33] Hepatitis_B_Foundation. Hepatitis b foundation statistics, 2009. [Online at `http://www.hepb.org/statistics.htm`; accessed 15-Mar-2017].

[34] T. L. Hoffman and R. W. Doms. Hiv-1 envelope determinants for cell tropism and chemokine receptor use. *Molecular Membrane Biology*, 16:57–65, 1999.

[35] Human-Connectome-Project. Hcp protocols, 2013. [Online at `http://protocols.humanconnectome.org/HCP/3T/task-fMRI-protocol-details.html`; accessed 26-July-2017].

[36] A. Khanna, A. Pascual-Leone, C. M. Michel, et al. Microstates in resting-state eeg: current status and future directions. *Neuroscience & Biobehavioral Reviews*, 49:105–113, 2015.

[37] M. Kikuchi et al. Eeg microstate analysis in drug-naive patients with panic disorder. *PloS one*, 6, e22912, 2011.

[38] I Kozyryev and J. Zhang. *Chapter 22 Bayesian Analysis of Complex Interacting Mutations in HIV Drug Resistance and Cross-Resistance*. Shanghai Jiao Tong University Press, Shanghai and and Springer Science+Business Media Dordrecht, 2015.

[39] E. W. Lang, A. M. Tom, I. R. Keck, et al. Brain connectivity analysis: A short survey. *Computational Intelligence and Neuroscience*, 2012, 2012.

[40] T. Lengauer and T. Sing. Bioinformatics-assisted anti-hiv therapy. *Nature Reviews Microbiology*, 4:790–797, 2006.

[41] X. Li, D. Zhu, X. Jiang, et al. Dynamic functional connectomics signatures for characterization and differentiation of ptsd patients. *Human Brain Mapping*, 35(4):1761–1778, 2014.

[42] Z. Lian, X. Li, Y. Pan, et al. Dynamic bayesian brain network partition and connectivity change point detection. *Computational Advances in Bio and Medical Sciences (ICCABS), 2015 IEEE 5th International Conference on*, 2012.

[43] Z. Lian, X Li, J Xing, et al. Exploring functional brain dynamics via a bayesian connectivity change point model. *Proceedings of the IEEE 11th International Symposium on Biomedical Imaging (ISBI 14)*, pages 600–603, 2014.

[44] Z. Lian, X. Li, T. Young, et al. Dynamic network partition via bayesian connectivity bipartition change point model. *Proceedings of the IEEE 11th International Symposium on Biomedical Imaging (ISBI 14)*, pages 545–548, 2014.

[45] Z. Lian, X. Li, H. Zhang, et al. Detecting cell assembly interaction patterns via bayesian based change-point detection and graph inference model. *Proceedings of the IEEE 11th International Symposium on Biomedical Imaging (ISBI 14)*, pages 17–20, 2014.

[46] Z. Lian, J. Lv, J. Xing, et al. Generalized fmri activation detection via bayesian magnitude change point model. *Proceedings of the IEEE 11th International Symposium on Biomedical Imaging (ISBI 14)*, pages 21–24, 2014.

[47] Z. Lian, Q. N. Tian, and Y. Liu. Detecting hepatitis b viral amino acid sequence mutations in occult hepatitis b infections via bayesian partition model. *Journal of Proteomics and Bioinformatics*, S6:005, 2013.

[48] M. A. Lindquist, C. Waugh, and T. D. Wager. Modeling state-related fmri activity using change-point theory. *Neuroimage*, 35(3):1125–1141, 2007.

[49] B. Liu, S. Feng, Y. Pan, and J. Zhang. Bayesian analysis of complex mutations in hbv hcv and hiv studies. *Computational and Mathematical Methods in Medicine*, 2017. submitted and under review.

[50] Z. Liu, Y. Kong, and B Su. An improved genetic algorithm based on the shortest path problem. *2016 IEEE International Conference on Information and Automation (ICIA)*, pages 328–332, 2016.

[51] N. K. Logothetis, J. Pauls, M. Augath, et al. Neurophysiological investigation of the basis of the fmri signal. *Nature*, 412(6843):150, 2001.

[52] J. Lv, X. Jiang, X. Li, et al. Assessing effects of prenatal alcohol exposure using group-wise sparse representation of fmri data. *Psychiatry Research: Neuroimaging*, 233(2):254–268, 2015.

[53] P. Lv, L. Guo, X. Hu, et al. Modeling dynamic functional information flows on large-scale brain networks. *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2013, Springer*, 2013.

[54] C. Lyu, X. Li, J. Lv, et al. Identifying group-wise consistent sub-networks via spatial sparse representation of natural stimulus fmri data. *Biomedical Imaging (ISBI), 2016 IEEE 13th International Symposium*, pages 62–65, 2016.

[55] A. Macdonald and M. Harris. Hepatitis c virus ns5a: Tales of a promiscuous protein. *Journal of General Virology*, 85:2485–2502, 2004.

[56] J. Mairal, M. Elad, and G. Sapiro. Sparse representation for color image restoration. *IEEE Transactions on Image Processing*, 17(1):53–69, 2008.

[57] D. Malonek and A Grinvald. Interactions between electrical activity and cortical microcirculation revealed by imaging spectroscopy: implications for functional brain mapping. *Science*, 272(5261):551, 1996.

[58] J. G. McHutchison, S. C. Gordon, E. R. Schiff, et al. Interferon alfa-2b alone or in combination with ribavirin as initial treatment for chronic hepatitis c. hepatitis interventional therapy group. *The New England Journal of Medicine*, 339(21):1485–1492, 1998.

[59] J. P. Miguet and D. Dhumeaux. *Progress in Hepatology 93*. France: Jogn Libbey Eurotext, 1993.

[60] C. Mirabelli, M. Surdo, F. Van Hemert, et al. Specific mutations in the c-terminus domain of hbv surface antigen significantly correlate with low level of serum hbv-dna in patients with chronic hbv infection. *Journal of Infection*, 70:288–298, 2015.

[61] M. Nettleman. Hepatitis b. [Online at `http://www.emedicinehealth.com/hepatitis_b/article_em.htm`; accessed 15-Mar-2017].

[62] E. Niedermeyer and F. L. da Silva. *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Lippincott Williams & Wilkins, 2004.

[63] S. O'Regan, S. Faul, and W. Marnane. Automatic detection of eeg artefacts arising from head movements. *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*, pages 6353–6, 2010.

[64] J. Ou, Z. Lian, L. Xie, et al. Atomic dynamic functional interaction patterns for characterization of adhd. *Human Brain Mapping*, 35(10):5262–5278, 2014.

[65] J. Ou, L. Xie, C. Jin, et al. Characterizing and differentiating brain state dynamics via hidden markov models. *Brain Topography*, 28(5):666–679, 2015.

[66] J. Ou, L. Xie, X. Li, et al. Atomic connectomics signatures for characterization and differentiation of mild cognitive impairment. *Brain Imaging and Behavior*, 9(4):663–677, 2015.

[67] J. Ou, L. Xie, P. Wang, et al. Modeling brain functional dynamics via hidden markov models. *Neural Engineering (NER), 2013 6th International IEEE/EMBS Conference*, pages 569–572, 2013.

[68] R. E. Passingham, K. E. Stephan, and R Kötter. The anatomical basis of functional localization in the cortex. nature reviews. *Neuroscience*, 3(8):606, 2002.

[69] S. Pei, J. Wang, W. Cui, et al. Codelet scheduling by genetic algorithm. *Trustcom/BigDataSE/I SPA, 2016 IEEE*, pages 1492–1499, 2016.

[70] M. Pirmohamed and D. J. Back. The pharmacogenomics of hiv therapy. *The Pharmacogenomics Journal*, 1:243–253, 2001.

[71] M. D. Plumbley, T. Blumensath, L. Daudet, et al. Sparse representations in audio and music: from coding to source separation. *Proceedings of the IEEE*, 98(6):995–1005, 2010.

[72] T. Poynard, V. Leroy, M. Cohard, et al. Meta-analysis of interferon randomized trials in the treatment of viral hepatitis c: Effects of dose and duration. *Hepatology*, 24(4):778–789, 1996.

[73] J. D. Ramsey, S. J. Hanson, and C. Glymour. Multi-subject search correctly identifies causal connections and most causal directions in the dcm models of the smith et al. simulation study. *NeuroImage*, 58(3):838–848, 2011.

[74] R. R. Regoes and S. Bonhoeffer. The hiv coreceptor switch: a population dynamical perspective. *Trends in Microbiology*, 13:269–277, 2005.

[75] U. Sakoğlu, G. D. Pearlson, K. A. Kiehl, et al. A method for evaluating dynamic functional network connectivity and task-modulation: application to schizophreni-

a. *Magnetic Resonance Materials in Physics, Biology and Medicine*, 23(5-6):351–366, 2010.

[76] J. F. Schenck. The role of magnetic susceptibility in magnetic resonance imaging: Mri magnetic compatibility of the first and second kinds. *Medical Physics*, 23(6):815–50, 1996.

[77] S. M. Smith, K. L. Miller, G. Salimi-Khorshidi, et al. Network modelling methods for fmri,. *NeuroImage*, 54(2):875–891, 2011.

[78] P. Spirtes and C. Glymour. An algorithm for fast recovery of sparse causal graphs. *Social Science Computer Review*, 9(1):62–72, 1991.

[79] Olaf Sporns. Brain connectivity. *Scholarpedia*, 2(10):4695, 2007.

[80] I. H. Stevenson, J. M. Rebesco, N. G. Hatsopoulos, et al. Bayesian inference of functional connectivity and network structure from spikes. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 17(3):203–213, 2009.

[81] J. Sun, X. Hu, X. Huang, et al. Inferring consistent functional interaction patterns from natural stimulus fmri data. *NeuroImage*, 61(4):987–999, 2012.

[82] V. Svicher, V. Cento, M. Bernassola, et al. Specific hbsag genetic determinants are associated with occult hbv infection in vivo and hbsag detection. *The 2nd International HIV & Hepatitis Virus Drug Resistance Workshop ANTIVIRAL THERAPY*, 16(4):A85, 2011.

[83] V. Svicher, V. Cento, M. Bernassola, et al. Novel hbsag markers tightly correlate with occult hbv infection and strongly affect hbsag detection. *Antiviral Research*, 93:288–298, 2012.

[84] Swartz_Center_for_Computational_Neuroscience. What is eeglab?, 2017. [Online at `https://sccn.ucsd.edu/eeglab/index.php`; accessed 17-September-2017].

[85] K. P. Thomas, A. P. Vinod, and G. Cuntai. Design of an online eeg based neurofeedback game for enhancing attention and memory. *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*, 2013.

[86] J. A. Turner. Brains & behavior program seed grant application, 2016. [B&B Program, Georgia State University].

[87] P. M. Vespa, V. Nenov, and M. R. Nuwer. Continuous eeg monitoring in the intensive care unit: Early findings and clinical efficacy. *Journal of Clinical Neurophysiology*, 16(1):1–13, 1999.

[88] A. F. Voevodin and P. A. Marx. *Sunuab virology*. USA: Wiley-Blackwell, 2009.

[89] J. Wackermann, D. Lehmann, C. M. Michel, et al. Adaptive segmentation of spontaneous eeg map series into spatially defined microstates. *International journal of psychophysiology : official journal of the International Organization of Psychophysiology*, 14:269–283, 1993.

[90] Wikipedia. Bayesian inference, 2017. [Online at `https://en.wikipedia.org/wiki/Bayesian_inference`; accessed 17-September-2017].

[91] Wikipedia. Electroencephalography, 2017. [Online at `https://en.wikipedia.org/wiki/Electroencephalography`; accessed 17-September-2017].

[92] Wikipedia. Functional magnetic resonance imaging, 2017. [Online at `https://en.wikipedia.org/wiki/Functional_magnetic_resonance_imaging`; accessed 1-June-2017].

[93] World_Health_Organization. [Online at `http://www.who.int/hiv/en/`].

[94] World_Health_Organization. Hepatitis b: The hepatitis b virus, 2002.

[95] World_Health_Organization. Hepatitis b: Fact sheet no. 204, 2016. [Online at `http://www.who.int/mediacentre/factsheets/fs204/en/`; accessed 15-Mar-2017].

[96] J. Wright, A. Y. Yang, A. Ganesh, et al. Robust face recognition via sparse representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(2):210–227, 2009.

[97] X. Xiao, B. Liu, J. Zhang, et al. Detecting change points in fmri data via bayesian inference and genetic algorithm model. *In: Cai Z., Daescu O., Li M. (eds) Bioinformatics Research and Applications. ISBRA 2017. Lecture Notes in Computer Science*, 2017. Springer, Cham.

[98] X. Xiao, B. Liu, J. Zhang, X. Xiao, and Y. Pan. An optimized method for bayesian connectivity change point model. *Journal of Computational Biology*, 2017, to appear.

[99] J. Zhang, T. Hou, Y. Liu, et al. Systematic investigation on interactions for hiv drug resistance and cross-resistance among protease inhibitors. *Journal of Proteome Science and Computational Biology*, 1(1):2, 2012.

[100] J. Zhang, T. Hou, W. Wang, and J. S. Liu. Detecting and understanding combinatorial mutation patterns responsible for hiv drug resistance. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, 107:1321, 2010.

[101] J. Zhang, X. Li, C. Li, et al. Inferring functional interaction and transition patterns via dynamic bayesian variable partition models. *Human Brain Mapping*, 35(7):3314–3331, 2014.

[102] X. Zhang, L. Guo, X. Li, et al. Characterization of task-free and task-performance brain states via functional connectome patterns. *Medical Image Analysis*, 17(8):1106–1122, 2013.

[103] X. Zhang, X. Li, C. Jin, et al. Identifying and characterizing resting state networks in temporally dynamic functional connectomes. *Brain Topography*, 27(6):747–765, 2014.

[104] M. Zhao, S. Li, and J. Kwok. Text detection in images using sparse representation with discriminative dictionaries. *Image and Vision Computing*, 28(12):1590–1599, 2010.

[105] D. Zhu, K. Li, L. Guo, et al. Dicccol: Dense individualized and common connectivity-based cortical landmarks. *Cerebral Cortex*, 23(4):786–800, 2013.

[106] D. Zhu, K. Li, P. D. Terry, et al. Connectome-scale assessments of structural and functional connectivity in mci. *Human Brain Mapping*, 35(7):2911–2923, 2014.

Appendix A

ADDITIONAL FIGURES FOR CHAPTER 3

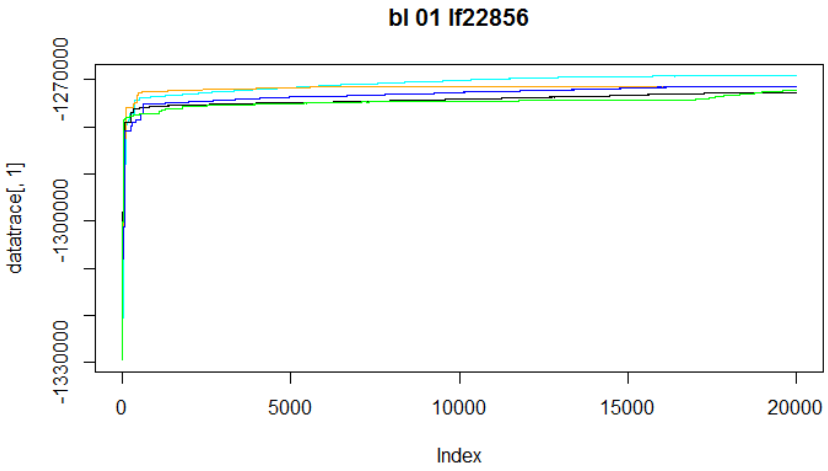

Figure A.1. Experiment 1: Traceplot shows the convergence of MCMC chains (p=-4000).

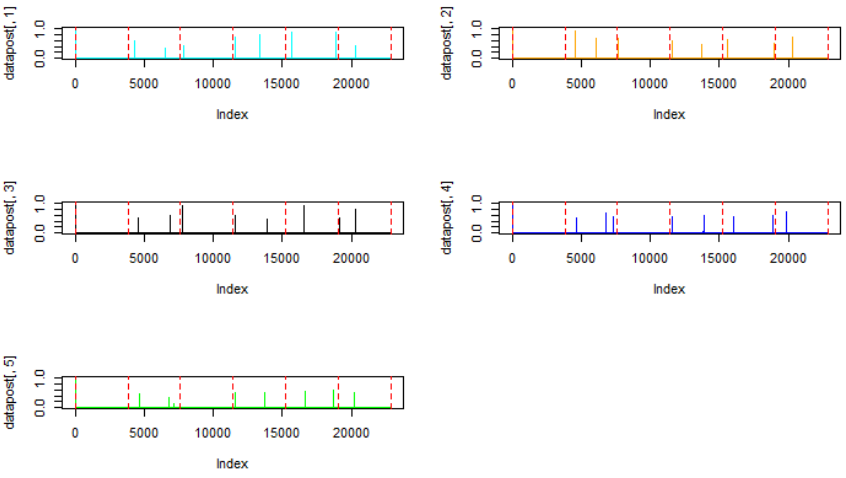

Figure A.2. Experiment 1: Change points detected by BCCPM for five repeated MCMC chains (p=-4000). Red dotted lines are the locations of designed change points.

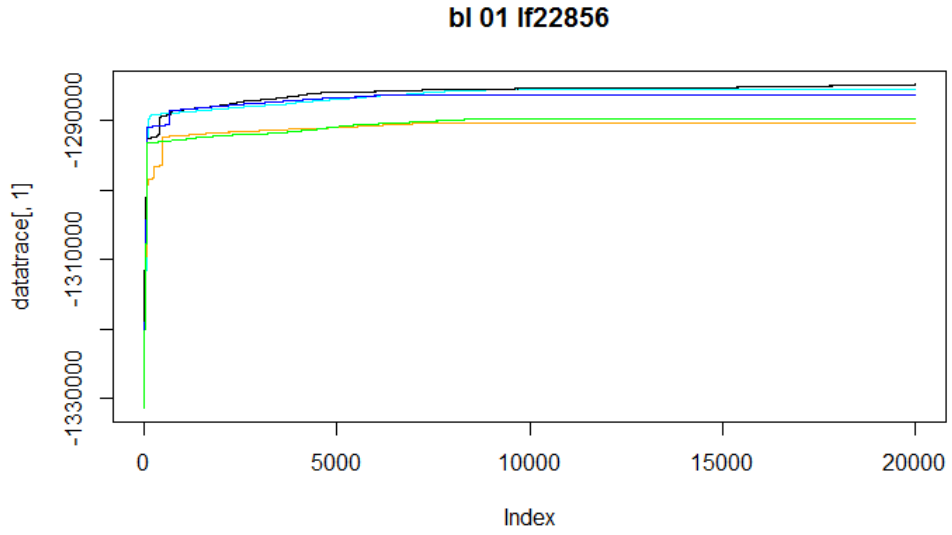

Figure A.3. Experiment 1: Traceplot shows the convergence of MCMC chains (p=-6000).

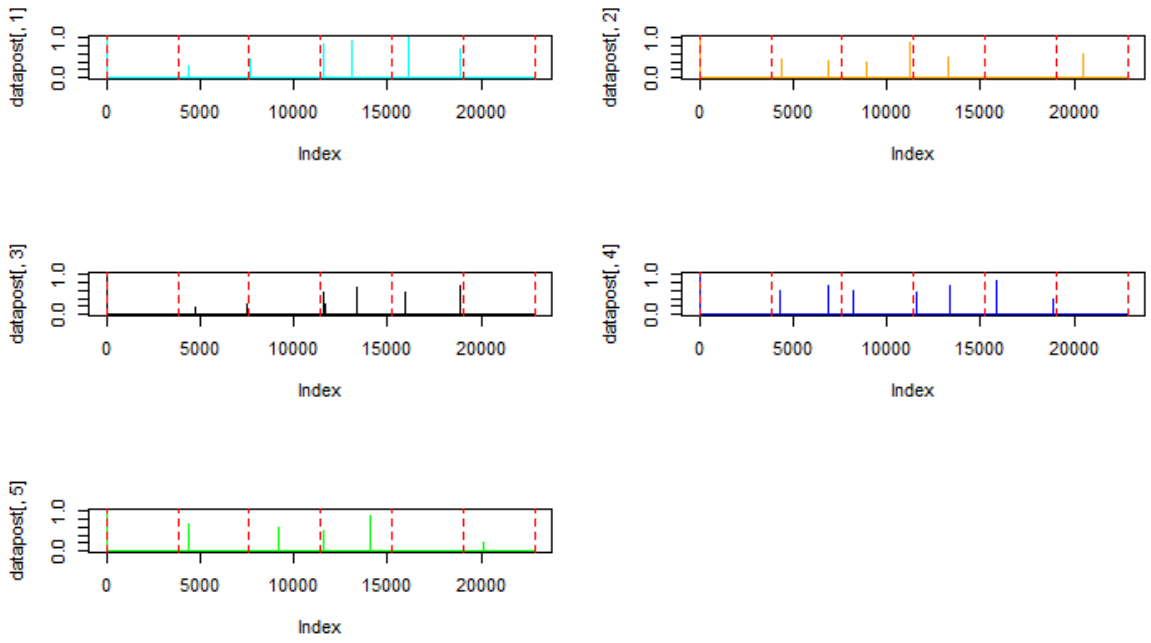

Figure A.4. Experiment 1: Change points detected by BCCPM for five repeated MCMC chains (p=-6000). Red dotted lines are the locations of designed change points.

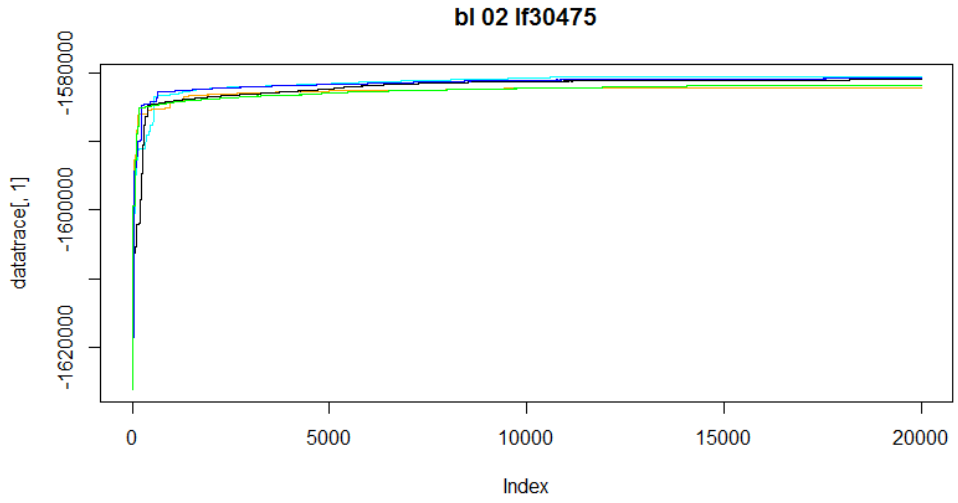

Figure A.5. Experiment 2: Traceplot shows the convergence of MCMC chains (p=-4000).

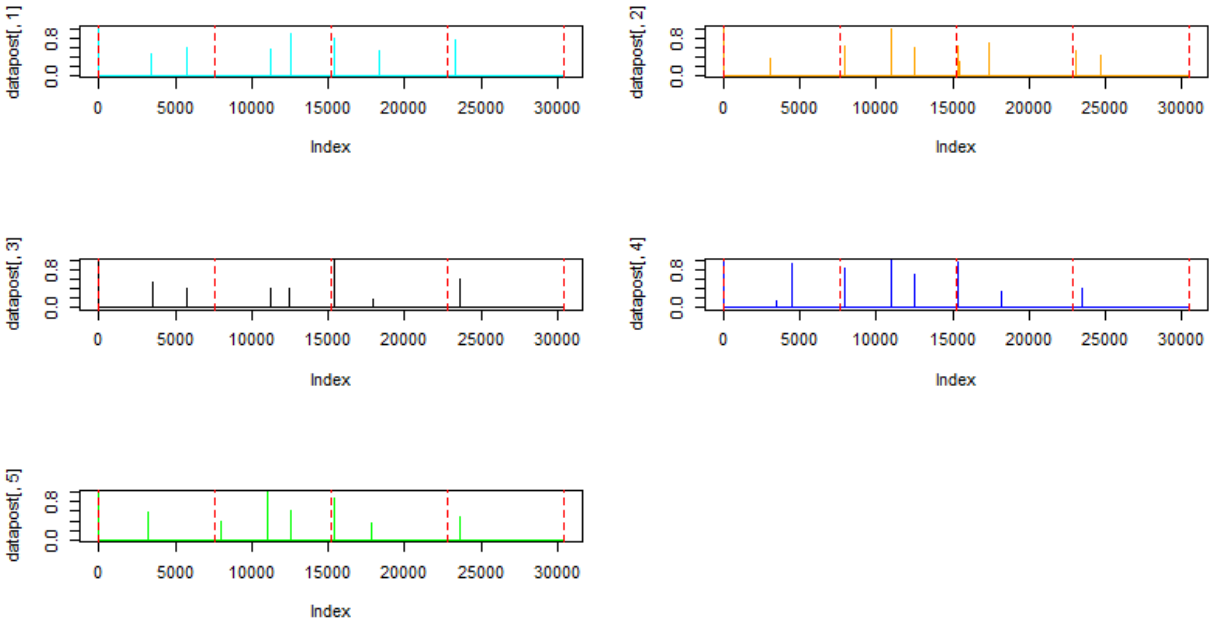

Figure A.6. Experiment 2: Change points detected by BCCPM for five repeated MCMC chains (p=-4000). Red dotted lines are the locations of designed change points.

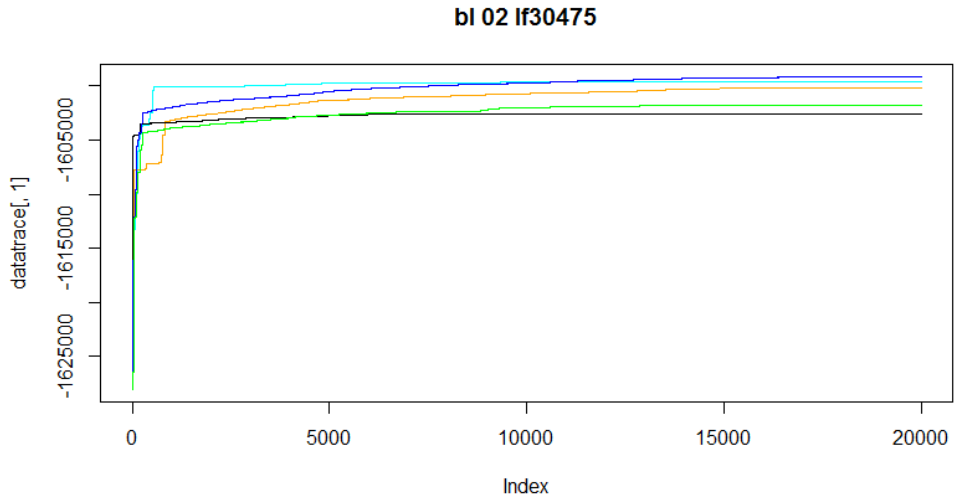

Figure A.7. Experiment 2: Traceplot shows the convergence of MCMC chains (p=-6000).

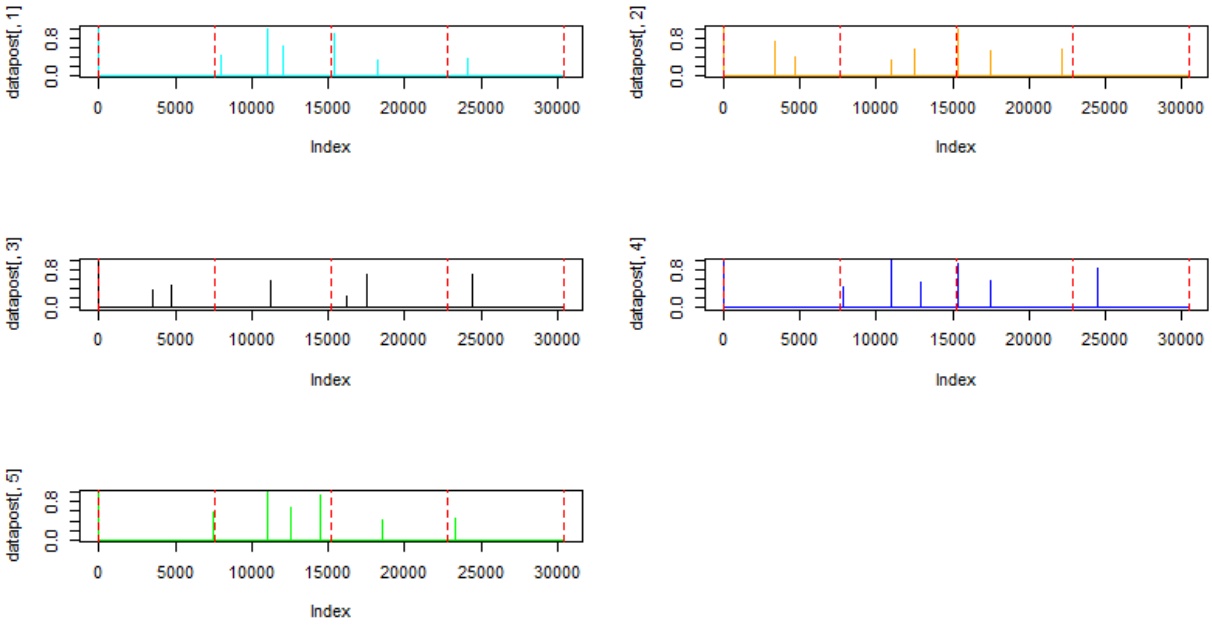

Figure A.8. Experiment 2: Change points detected by BCCPM for five repeated MCMC chains (p=-6000). Red dotted lines are the locations of designed change points.

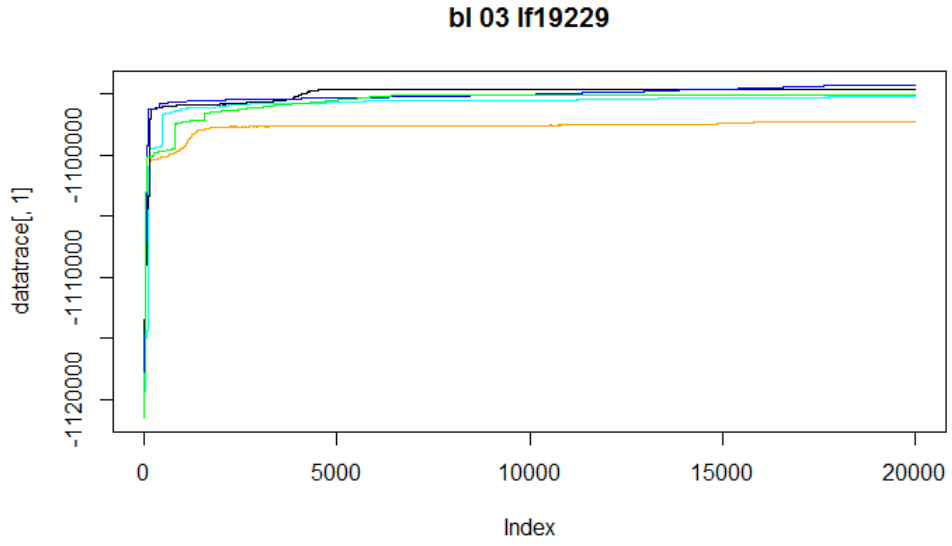

Figure A.9. Experiment 3: Traceplot shows the convergence of MCMC chains (p=-4000).

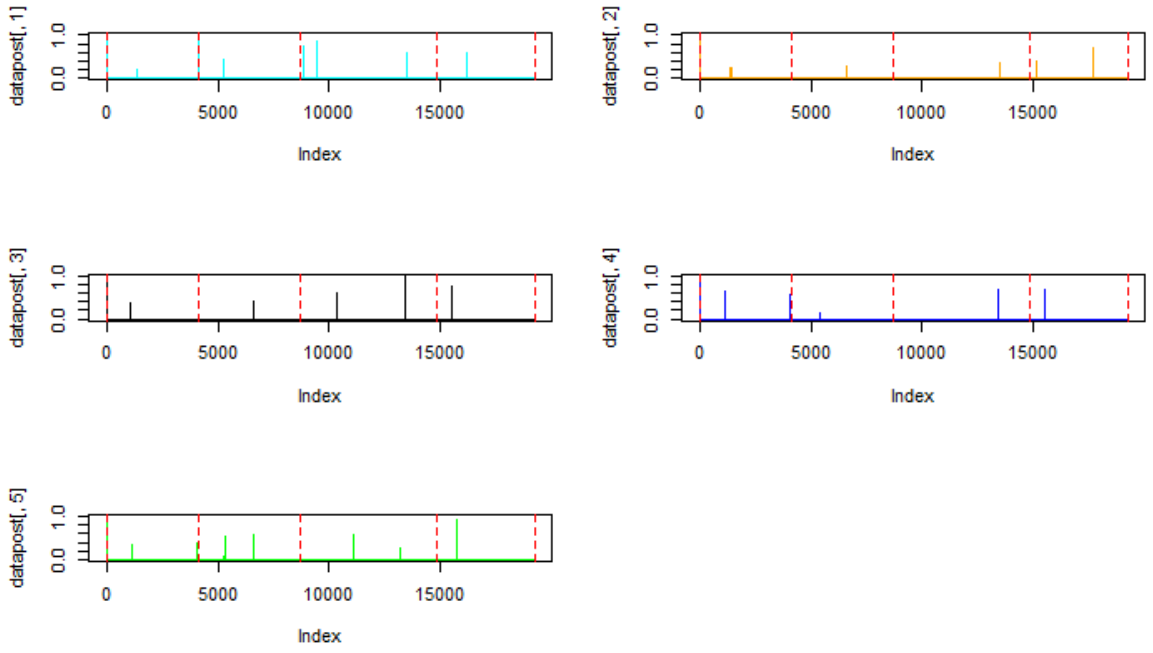

Figure A.10. Experiment 3: Change points detected by BCCPM for five repeated MCMC chains (p=-4000). Red dotted lines are the locations of designed change points.

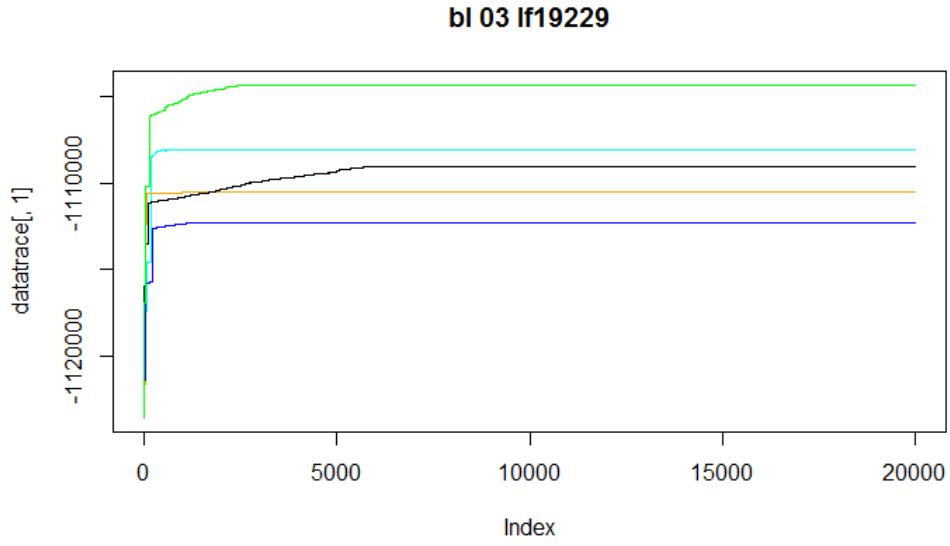

Figure A.11. Experiment 3: Traceplot shows the convergence of MCMC chains (p=-6000).

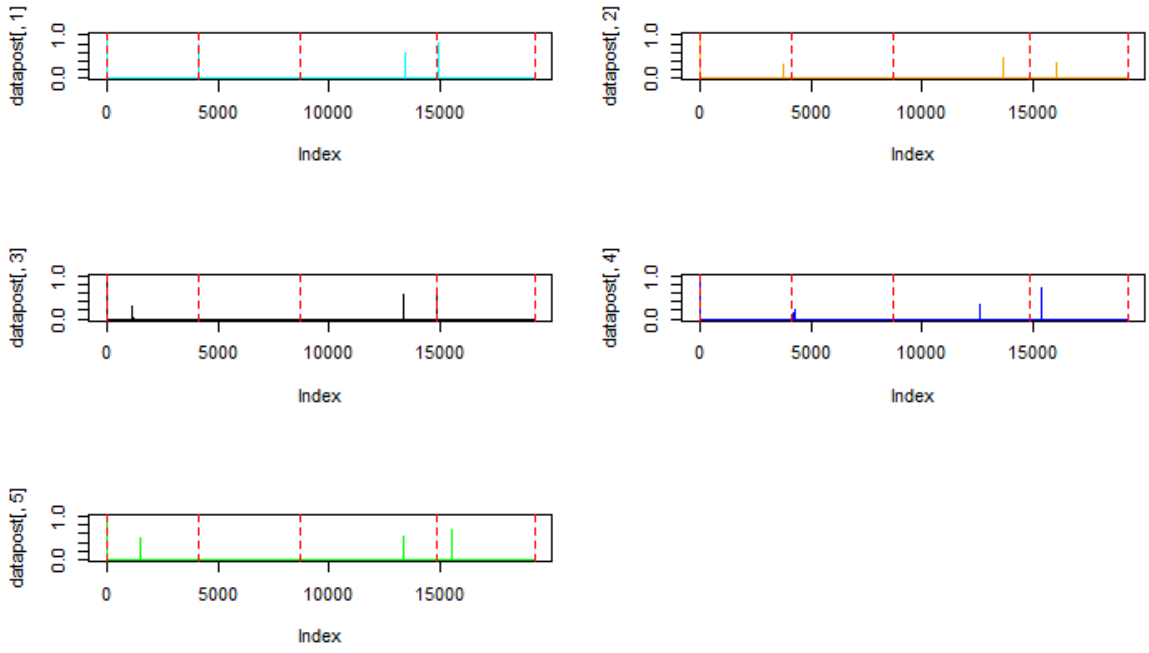

Figure A.12. Experiment 3: Change points detected by BCCPM for five repeated MCMC chains (p=-6000). Red dotted lines are the locations of designed change points.

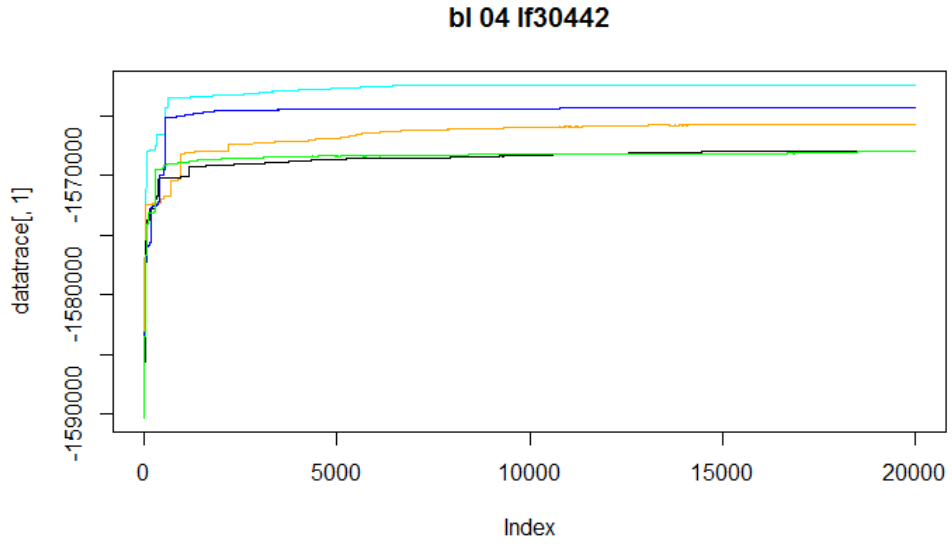

Figure A.13. Experiment 4: Traceplot shows the convergence of MCMC chains (p=-4000).

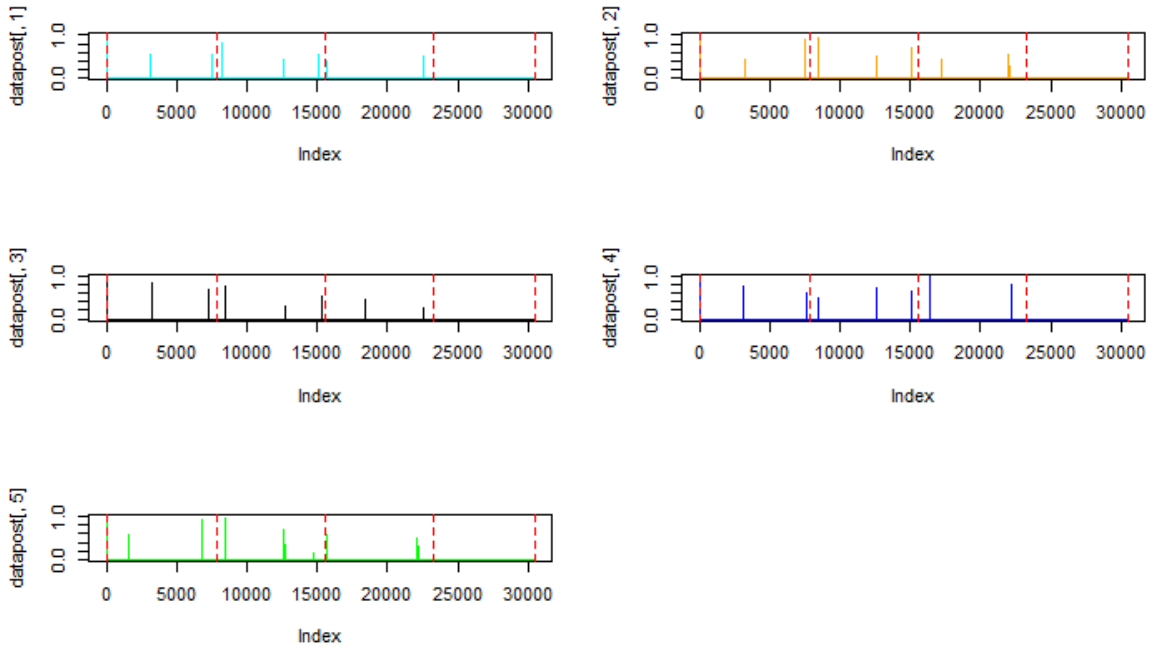

Figure A.14. Experiment 4: Change points detected by BCCPM for five repeated MCMC chains (p=-4000). Red dotted lines are the locations of designed change points.

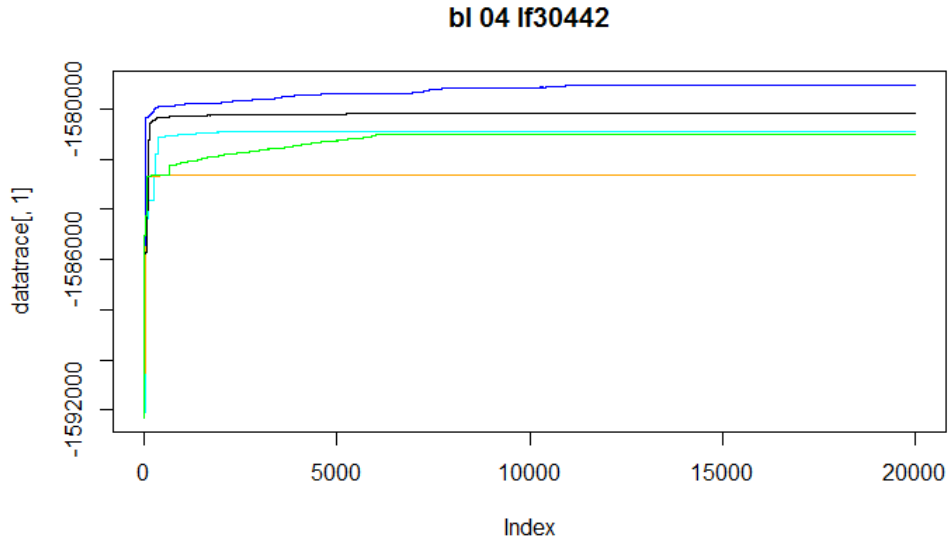

Figure A.15. Experiment 4: Traceplot shows the convergence of MCMC chains (p=-6000).

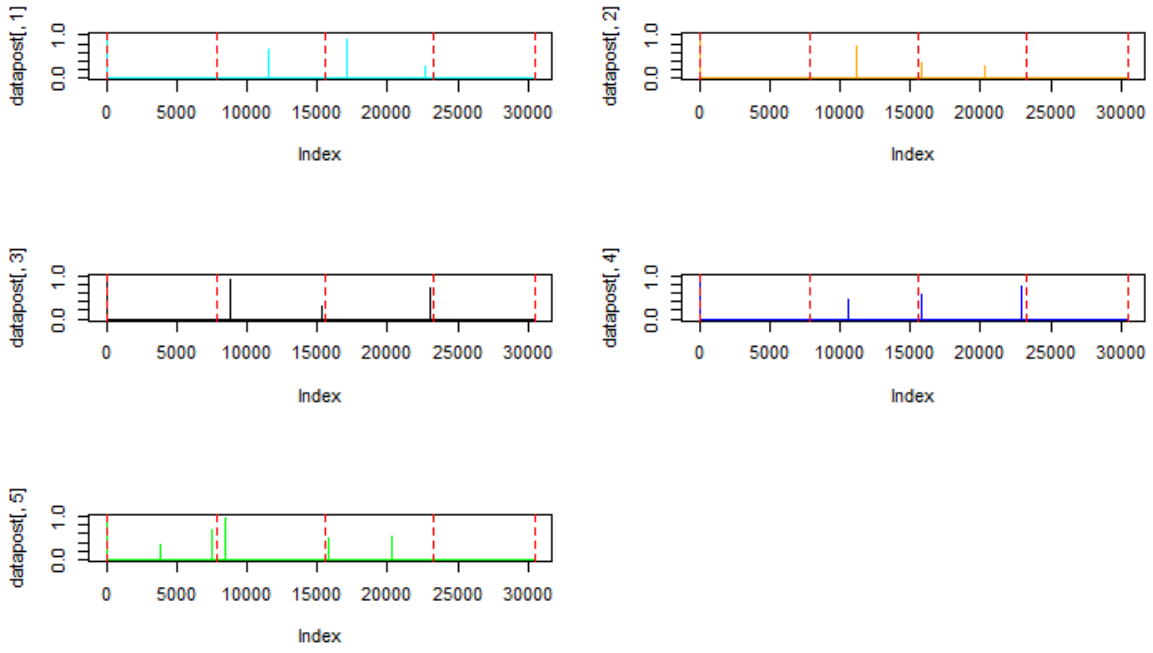

Figure A.16. Experiment 4: Change points detected by BCCPM for five repeated MCMC chains (p=-6000). Red dotted lines are the locations of designed change points.

Appendix B

LETTER OF SUPPORT FROM DR. JESSICA TURNER

DEPARTMENT OF PSYCHOLOGY
College of Arts and Sciences

P.O. Box 5010
Atlanta GA 30302-5010

Phone: 404/413-6200
Fax: 404/413-6207

Georgia State University

October 17, 2016

Dear Bing,

I am happy to write in support of your dissertation fellowship. You are proposing to use data collected as part of one of my studies; you can certainly use that data on the proviso that my co-PI Dr. Matthew Turner and I are included in manuscript preparation and authorship in any publications using these data, and given acknowledgement in other presentations as the source of the dataset. The other proviso is that you not share the data with others, as we are sharing it with you for your dissertation analyses only. We are happy to discuss other data sharing arrangements in the future.

The analysis you are proposing is a very good one to do on these data, and I wish you the best of success in it.

Best wishes,

Jessica Turner

Jessica A. Turner, Ph.D.
Associate Professor, Psychology and Neuroscience
Georgia State University
Atlanta, GA
(404) 413-6211
email: jturner63@gsu.edu

Associate Professor (Affiliate), Translational Neuroscience
Mind Research Network
Albuquerque, New Mexico
email: jturner@mrn.org

Figure B.1. Letter of Support from Dr. Jessica Turner