A latent process approach to change-point detection of mixed-type observations

Shuyu Chu, Xueying Liu, Achla Marathe & Xinwei Deng

Published online: 28 Jun 2023.

Submit your article to this journal ⃗

View related articles ⃗

View Crossmark data ⃗

Taylor & Francis
Taylor & Francis Group

Check for updates

A latent process approach to change-point detection of mixed-type observations

Shuyu Chu[a], Xueying Liu[a], Achla Marathe[b], and Xinwei Deng[a]

[a]Department of Statistics, Virginia Tech, Blacksburg, Virginia; [b]Biocomplexity Institute, University of Virginia, Charlottesville, Virginia

ABSTRACT

Mixed-type observations, such as continuous measurements, discrete counts, and binary outcomes, are commonly present in many applications. The change-point detection with mixed-type observations is challenging since it is difficult to quantify the hidden association among mixed-type observations. In this work, we propose a latent process method to model the mixed observations in a joint manner, and effectively detect the changes. Bayesian parameter estimation and inference are developed for the proposed method by combining the discrete particle filter (DPF) and sequential Monte Carlo (SMC) algorithms. Such an algorithm can efficiently update the high dimensional proposal distribution and can exploit the discrete and continuous natures of the latent processes simultaneously. The performance of the proposed method is illustrated by several numerical examples and a case study of civil-unrest data.

1. Introduction

In various applications, mixed-type observations are widely present, where continuous and discrete observations are often collected together to evaluate the system performance. For example, in a wafer lapping manufacturing process (Deng and Jin 2015; Kang et al. 2018), the total thickness variation is a continuous quality response to characterize the range of the wafer thickness. The conformity of site total indicator readings (STIR) is a binary response measurement to indicate whether the STIR is larger than the tolerance or not. In the service industry, the average time of job completion and the daily base job frequency are continuous and discrete measurements to monitor its service process in terms of freight amount (Ning and Tsung 2012). In an example involving civil unrest (O'brien 2010; Ramakrishnan et al. 2014), social events are recorded with several measurements, including the event frequency during a certain time interval and a continuous quantity that records the average tone of the events. The multiple measurements from a system are often highly correlated and thus should be jointly considered to effectively evaluate the performance of the system.

The change-point problems commonly occur in many processes. Detection of change-points is an important task for monitoring the performance of a system. We consider change-points to be those time points which divide the whole data set into distinct homogeneous segments. Generally, change-point detection techniques can be categorized by offline or online methods, and parametric or non-parametric methods. This work restricts its attention to methods which are offline and parametric. The change-point detection has been developed for over sixty years, with early work including Page (1954), Shiryaev (1963), and Hinkley and Hinkley (1970). Numerous applications are shown on a wide range of disciplines, such as manufacturing processes (Ge and Smyth 2000), bioinformatics (Lio and Vannucci 2000; Erdman and Emerson 2008), healthcare (Gao et al. 2019), cybersecurity (Young and Kuo 2001), and finance (Spokoiny 2009). For a more general overview of change-point detection methods, please refer to Müller and Siegmund (1994), Chen and Gupta (2011), Truong, Oudre, and Vayatis (2020), and Xie et al. (2021).

For the continuous observations, a variety of frequentist and Bayesian approaches are developed to detect change-points in the literature, including Hinkley and Hinkley (1970), Chen and Gupta (1997), Gupta and Chen (1996), Andrieu, Doucet, and Holenstein (2010), Cappé, Moulines, and Rydén

(2009), Fearnhead (1998), Frühwirth-Schnatter (2006), and Whiteley, Andrieu, and Doucet (2010). The frequentist approaches mainly focus on the likelihood-ratio based or penalized likelihood methods (Li, Tsung, and Zou 2013; Zhang et al. 2010), while the Bayesian approaches rely on the specification of a prior for the number and position of change-points (Eckley, Fearnhead, and Killick 2011). Among various methodology, the Bayesian methods using switching state-space model (SSSM) show promising performance. The Markov chain Monte Carlo (MCMC) algorithm is often used in such methods for efficient sampling of the posterior distribution for inference (Cappé, Moulines, and Rydén 2009; Frühwirth-Schnatter 2006; Eckley, Fearnhead, and Killick 2011). The SSSM assumes that the observations are generated from a latent Markov process, which can be continuous or discrete. If the latent process is continuous, one can use the sequential Monte Carlo (SMC) algorithm to design efficient high-dimensional proposals within the MCMC scheme (Andrieu, Doucet, and Holenstein 2010). If the latent process is discrete, the discrete particle filter (DPF) is an efficient algorithm for the estimation and inference of the latent parameters, see Fearnhead (1998), Whiteley, Andrieu, and Doucet (2010), and Cappé, Moulines, and Rydén (2009).

However, the change-point problem has not received much attention for cases when there are both discrete or mixed observations. Few works have focused on the change-point detection problem for mixed-type observations. In terms of modeling of mixed-type observations, it is mentioned in Chen and Brown (2013) that the linear Gaussian state-space model can be extended to accommodate the discrete or mixed observation. It is also known that the copula model could provide a unified framework to model statistical dependencies among continuous, discrete, or mixed-type random variables (Chen 2013). For example, de Leon and Wu (2011) developed a copula-based regression model for binary and continuous observations, where a latent variable formulation is adopted for the binary observations. However, the model in de Leon and Wu (2011) is specially for binary and continuous observations, and it is not straightforward to be extended to other types of discrete observations. For the aforementioned modeling methods, it is also not clear how to combine these modeling methods with the change-point detection.

There are several frequentist approaches proposed to address the change-point problem for mixed-type data. Ning and Tsung (2012) develop a density-based statistical process control scheme to detect process changes in mixed-type observations. Their key idea is to transform the multi-dimensional observations into a one-dimensional measurement using a local outlier factor (LOF). However, such a method requires several predetermined parameters such as control limits, which could highly depend on the quantity and quality of the data. In Qiu (2008), the mixed-type data are all converted into binary or categorical variables and their distributions are estimated using the log-linear modeling approach. Thus the change-point is detected based on changes in the estimated distributions. But such an approach requires a reasonable amount of high-quality in-control data to give an accurate estimation of the in-control distribution, which might be difficult to obtain in practice.

In this work, we propose a Bayesian approach for the change-point detection with mixed-type observations. Specifically, a so-called mixed switching state-space model (*mixed SSSM*) is proposed for the mixed-type observation with both continuous and discrete observations. The proposed model contains two latent Markov processes, the continuous and discrete latent processes, such that they can jointly describe the hidden dynamics in the continuous and discrete observations. In particular, a sequence of indicator variables is introduced to indicate the occurrence of significant changes in the data sequence. To enable efficient estimation of the posterior distributions of the latent processes, we develop an effective estimation algorithm by combining sequential Monte Carlo and discrete particle filter algorithms iteratively. The contributions of this work are as follows. First, we propose a *mixed SSSM* model for quantifying the relationship between continuous and discrete variables under various settings of underlying distributions. Second, a combined particle filter algorithm is developed for detecting change-point(s) effectively and estimating parameters efficiently. Third, the proposed method can conduct efficient Bayesian inference through the MCMC technique.

The rest of this work is organized as follows. In Section 2, we detail the proposed *mixed SSSM* model for mixed-type data. Section 3 elaborates the efficient sampling algorithm, as well as the pseudo-code of the particle marginal Metropolis-Hastings sampler. Section 4 reports several numerical examples to demonstrate the performance of the proposed schemes. A real case study of civil protest is used to illustrate the implementation of the proposed approach in Section 5. We conclude this work in Section 6 with some discussion.

2. Switching state-space models for mixed-type data

2.1. Notation

For notation convenience, the capital letters are used for random variables and lowercase letters denote their values. Suppose there are two sequences of observations $\{Y_n; n = 1, \cdots, T\} \subset \mathcal{Y}^{\mathbb{N}}$ and $\{Z_n; n = 1, \cdots, T\} \subset \mathcal{Z}^{\mathbb{N}}$, where $T \geqslant 1$ is the length of the sequence. Here the Y_n is a continuous variable and Z_n is a non-negative discrete variable. We also assume that the pair Y_n and Z_n are dependent with each other in some unobserved manner. Denote $0 < \tau_1 < \tau_2 < \cdots < \tau_k < T$ to be k change-point locations. It implies that the observations $\{Y_n, Z_n\}$ are homogeneous within each segment $[\tau_j, \tau_{j+1}]$ and heterogeneous across segments, where $j = 0, 1, \cdots, k$, and denote $\tau_0 = 0, \tau_{k+1} = T$.

To quantify the relationship between Y_n and Z_n, two latent processes are considered, a continuous latent process $\{X_n\}_{n \geqslant 1}$ and a discrete latent process $\{I_n\}_{n \geqslant 1}$. The values of X_n lie in a real-valued set $\mathcal{X}$ and the values of I_n belong to a finite set $\mathcal{I}$. Denote $|\mathcal{I}|$ as the cardinality of $\mathcal{I}$, then $\mathcal{I} = \{1, \cdots, |\mathcal{I}|\}$. The process $\{X_n\}_{n \geqslant 1}$ is to explain the dependence between Y_n and Z_n, while the process $\{I_n\}_{n \geqslant 1}$ indicates the states of the observations. That is, change-point occurs when two consecutive sets of observations are in different states. Under the context of change-point detection, we assume that I_n takes at least two different possible values. The value of I_n maintains the same within each segment and varies across different segments. Note that if I_n only takes one value, then there will be no change point.

Hereafter, let us denote $\boldsymbol{y}_{1:T}$ as an observed vector $\boldsymbol{y}_{1:T} = (y_1, \cdots, y_T)'$ of the random vector $\boldsymbol{Y}_{1:T} = (Y_1, \cdots, Y_T)'$. Similarly, one can define $\boldsymbol{z}_{1:T}, \boldsymbol{x}_{1:T}, \boldsymbol{i}_{1:T}, \boldsymbol{Z}_{1:T}, \boldsymbol{X}_{1:T}$, and $\boldsymbol{I}_{1:T}$. Denote the support of $\boldsymbol{I}_{1:n}$ as $\mathcal{I}^n$. As n increases, the possible paths of $\boldsymbol{I}_{1:n}$, which is $|\mathcal{I}|^n$, grows exponentially. Thus, we define a parameter N_1 as the maximum number of support points of $\boldsymbol{I}_{1:n}$ at time step n. Similarly, N_2 is defined as the number of sampling points of $\boldsymbol{X}_{1:n}$ at time n. Let us also denote the indicator function as $\delta_A(x)$, which takes value of 1 if $x \in A$, and 0 otherwise.

2.2. The proposed model

Recall that the observed sequences are $\{Y_n\}$ and $\{Z_n\}$, and the latent sequences are $\{I_n\}$ and $\{X_n\}$, where $n \geqslant 1$. The main objective of this work is to find the possible change-point location(s), $0 < \tau_1 < \tau_2 < \cdots < \tau_k < T, 1 \leqslant k \leqslant T$, where the latent discrete process I_n changes from one state to another.

We assume that both latent processes are Markov processes with initial values as $I_1 \sim \nu_{\boldsymbol{\theta}}(\cdot)$ and $X_1 \sim \mu_{\boldsymbol{\theta}}(\cdot)$, respectively. Their transition probability densities are denoted as $f_{\boldsymbol{\theta}}^I$ and $f_{\boldsymbol{\theta}}^X$,

$$I_{n+1}|(I_n = i) \sim f_{\boldsymbol{\theta}}^I(\cdot|i), \tag{1}$$

$$X_{n+1}|(X_n = x) \sim f_{\boldsymbol{\theta}}^X(\cdot|x), \tag{2}$$

where $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ is a vector of static parameters in the model. Here $f_{\boldsymbol{\theta}}^I$ is in fact a stochastic transition matrix. Given $\{I_n\}$ and $\{X_n\}$, we assume that $\{Y_n\}$ and $\{Z_n\}$ are conditionally independent, where the distributions of Y_n and Z_n only depend on the current latent observations of I_n and X_n, for all $n \geqslant 1$, similar to the convention in Whiteley, Andrieu, and Doucet (2010) and Andrieu, Doucet, and Holenstein (2010). This is because of the Markov independence property that knowing the state at any time makes the past, present and future observations statistically independent. By denoting $g_{\boldsymbol{\theta}}(y|x, i)$ and $h_{\boldsymbol{\theta}}(z|x, i)$ as the common marginal probability densities of Y_n and Z_n given the latent processes, we can have

$$Y_n|(X_1 = x_1, \cdots, X_n = x_n, I_1 = i_1, \cdots, I_n = i_n)$$
$$\sim g_{\boldsymbol{\theta}, i_n}(\cdot|x_n), \tag{3}$$

and,

$$Z_n|(X_1 = x_1, \cdots, X_n = x_n, I_1 = i_1, \cdots, I_n = i_n)$$
$$\sim h_{\boldsymbol{\theta}, i_n}(\cdot|x_n). \tag{4}$$

It is seen that the relationship between Y_n and Z_n is well defined through (3) and (4). Changes in the two latent processes I_n and X_n have a direct impact on Y_n and Z_n simultaneously. Thus the mixed-type observations are closely connected through the latent processes. Correspondingly, the complete formulation of the proposed switching state-space model for mixed-type observations is described through (1) to (4). We call the proposed model as the *mixed SSSM*.

The *mixed SSSM* generalizes the original SSSM to observations containing both continuous and discrete variables through two types of latent processes. In the proposed model, the continuous latent process not only controls the latent dynamics but also is developed as a bridge to connect mixed-type observations. Moreover, the proposed model retains the Markov properties and gives more flexibility on data structures. However, due to the mixture property, the *mixed SSSM* is clearly non-linear and non-Gaussian, which makes the traditional particle filter algorithm

no longer appropriate for the model estimation and inference.

Example: A linear mixed Gaussian-Poisson SSSM.

Suppose the discrete latent process $\{I_n\}$, where $I_n \in \{0, 1\}$, is a Markov chain with transition matrix P_I as

$$P_I = \begin{bmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{bmatrix}. \tag{5}$$

For the continuous latent process $\{X_n\}$, consider the following transition relationship,

$$x_{n+1} = \phi x_n + \sigma I_n V_n, \tag{6}$$

where $\{V_n\}$ are independent and identically distributed (i.i.d.) with standard normal distribution $\mathcal{N}(0, 1)$. For the initial distributions, set $I_1 \sim Bern(0.5)$ and $X_1 \sim \mathcal{N}(0, 1)$.

Given the observed continuous process $\{X_n\}$ and the observed discrete process $\{Z_n\}, n = 1, \cdots, T$, consider the mixed Gaussian-Poisson SSSM as follows,

$$y_n = x_n + \gamma V_n, \tag{7}$$

$$z_n \sim Poisson\left(\alpha \sqrt{|x_n|}\right). \tag{8}$$

It means that, given I_n and X_n, the continuous variable Y_n follows a Gaussian distribution, $\mathcal{N}(x_n, \gamma^2)$ and the discrete variable Z_n follows a Poisson distribution with parameter $\alpha \sqrt{|x_n|}$. Clearly, the static parameter vector θ contains $\phi, \sigma, \gamma,$ and α in this example.

2.3. Model inference

Based on (3) and (4), we can obtain,

$$g_\theta(\mathbf{y}_{1:T} | \mathbf{x}_{1:T}, \mathbf{i}_{1:T}) = \prod_{n=1}^T g_{\theta, i_n}(y_n | x_n), \tag{9}$$

$$h_\theta(\mathbf{z}_{1:T} | \mathbf{x}_{1:T}, \mathbf{i}_{1:T}) = \prod_{n=1}^T h_{\theta, i_n}(z_n | x_n). \tag{10}$$

Consequently, we get the joint distribution of $\mathbf{y}_{1:T}$ and $\mathbf{z}_{1:T}$ given $\mathbf{x}_{1:T}, \mathbf{i}_{1:T}$,

$$p_\theta(\mathbf{y}_{1:T}, \mathbf{z}_{1:T} | \mathbf{x}_{1:T}, \mathbf{i}_{1:T}) = g_{\theta, i_{1:T}}(\mathbf{y}_{1:T} | \mathbf{x}_{1:T}) \times h_{\theta, i_{1:T}}(\mathbf{z}_{1:T} | \mathbf{x}_{1:T})$$

$$= \prod_{n=1}^T g_{\theta, i_n}(y_n | x_n) h_{\theta, i_n}(z_n | x_n). \tag{11}$$

Conditional on the observed data $\mathbf{y}_{1:T}, \mathbf{z}_{1:T}$ for $T \geqslant 1$, the goal is to conduct the Bayesian inference of all the unknown parameters $(\theta, \tau_{1:k})$, especially the change-point locations $\tau_{1:k}$. If $\theta \in \Theta$ is known, it is easy to see that the Bayesian inference mainly relies on the posterior density,

$$p_\theta(\mathbf{x}_{1:T}, \mathbf{i}_{1:T} | \mathbf{y}_{1:T}, \mathbf{z}_{1:T}) \propto p_\theta(\mathbf{x}_{1:T}, \mathbf{i}_{1:T}, \mathbf{y}_{1:T}, \mathbf{z}_{1:T})$$

$$= f_\theta^I(\mathbf{i}_{1:T}) f_\theta^X(\mathbf{x}_{1:T} | \mathbf{i}_{1:T}) g_\theta(\mathbf{y}_{1:T} | \mathbf{x}_{1:T}, \mathbf{i}_{1:T}) h_\theta(\mathbf{z}_{1:T} | \mathbf{x}_{1:T}, \mathbf{i}_{1:T})$$

$$= \nu_\theta(i_1) \mu_\theta(x_1) \prod_{n=2}^T f_\theta^I(i_n | i_{n-1}) f_{\theta, i_{n-1}}^X(x_n | x_{n-1})$$

$$\prod_{n=1}^T g_{\theta, i_n}(y_n | x_n) h_{\theta, i_n}(z_n | x_n). \tag{12}$$

If θ is unknown, we can assign a suitable prior density $p(\theta)$ and then the Bayesian inference can be conducted based on the joint density

$$p(\theta, \mathbf{x}_{1:T}, \mathbf{i}_{1:T} | \mathbf{y}_{1:T}, \mathbf{z}_{1:T}) \propto p_\theta(\mathbf{x}_{1:T}, \mathbf{i}_{1:T}, \mathbf{y}_{1:T}, \mathbf{z}_{1:T}) p(\theta). \tag{13}$$

To calculate the posterior distributions of latent processes $\{I_n\}$ and $\{X_n\}$, we propose to alternately combine the DPF and SMC algorithms, which will be detailed in the next section. As the model defined through (1) to (4) can be non-linear or non-Gaussian, there are often times no explicit expressions for those posterior densities, $p_\theta(\mathbf{x}_{1:T}, \mathbf{i}_{1:T}, \mathbf{y}_{1:T}, \mathbf{z}_{1:T})$ and $p(\theta, \mathbf{x}_{1:T}, \mathbf{i}_{1:T} | \mathbf{y}_{1:T}, \mathbf{z}_{1:T})$, making exact inference difficult in practice. Thus it is natural to resort to approximations, where the particle Markov chain Monte Carlo (PMCMC) method provides a flexible Bayesian framework to address such difficulties.

3. Particle MCMC algorithm for mixture SSSM

It is known that particle filter algorithms are commonly used to conduct efficient Bayesian inferences under the latent processes context. In the literature, the sequential Monte Carlo (SMC) algorithm was proposed for the continuous latent process (Andrieu, Doucet, and Holenstein 2010), while the discrete particle filter (DPF) algorithm was specially developed for the discrete latent process (Whiteley, Andrieu, and Doucet 2010; Fearnhead 1998; Fearnhead and Clifford 2003). However, neither of these two algorithms can work well individually for the *mixed SSSM* containing both continuous and discrete observations. To the best of our knowledge, there is little work focusing on problems with both continuous and discrete observations. To address this challenge, we propose a new algorithm, so-called "combined DPF & SMC" algorithm, by taking advantages of both SMC and DPF algorithms such that we can jointly estimate the unknown parameter θ and change-point(s) $\tau_1, \cdots, \tau_k$. We will detail the combined DPF & SMC algorithm in Section 3.1. Using the posterior densities estimated from Section 3.1, Section 3.2

will elaborate how to update unknown parameters and latent processes using the particle marginal Metropolis-Hastings (PMMH) sampler.

3.1. Combined DPF & SMC algorithm

A combined DPF & SMC algorithm is developed to make Bayesian inference for the *mixed SSSM*, conditional upon the mixed observations $y_{1:T}, z_{1:T}$ and treating the static parameter θ and both the latent processes $X_{1:T}, I_{1:T}$ unknown. The joint density of these unknowns is shown in (13). It is easy to see that this posterior can be factorized as follows,

$$p(\theta, x_{1:T}, i_{1:T}|y_{1:T}, z_{1:T}) = p(\theta, i_{1:T}|y_{1:T}, z_{1:T})p_\theta(x_{1:T}|y_{1:T}, z_{1:T}, i_{1:T}) \quad (14)$$

where

$$p(\theta, i_{1:T}|y_{1:T}, z_{1:T}) = \frac{p_\theta(y_{1:T}, z_{1:T}|i_{1:T})p(i_{1:T}|\theta)p(\theta)}{\int_\Theta \sum_{i'_{1:T} \in \mathcal{I}^T} p_\theta(y_{1:T}, z_{1:T}|i'_{1:T})p(i'_{1:T}|\theta)p(\theta) \, d\theta} \quad (15)$$

However, the difficulty in computing $p(\theta, i_{1:T}|y_{1:T}, z_{1:T})$ comes from the exponentially increasing realizations of $i_{1:T}$, which is $|\mathcal{I}|^T$. Even if θ is fixed, it would be very expensive to consider all the possible realizations of $i_{1:T}$ for a modest value of T (Whiteley, Andrieu, and Doucet 2010). Thus consider an approximated computation of $p_\theta(i_{1:T}|y_{1:T}, z_{1:T})$ sequentially *via* the recursive relationship,

takes a value s_n which is a subset of $\mathcal{I}^n$. As n increases, the possible paths of $I_{1:n}$ grows exponentially. Then N_1 acts as a pruning parameter to prevent the support from growing too big, through resampling techniques when $|S_n| > N_1$. On the other hand, the value of N_1 controls the precision of the DPF algorithm. A larger value of N_1 will lead to more accurate (on average) approximation for the target distribution $p_\theta(i_{1:T}|y_{1:T}, z_{1:T})$, and it has been shown that the DPF algorithm works efficiently even with a moderate number of particles (Whiteley, Andrieu, and Doucet 2010; Chen and Liu 2000; Doucet, Godsill, and Andrieu 2000, Doucet, Gordon, and Kroshnamurthy 2001).

Since the objective is to find the change-point(s) locations, we need the approximation of the posterior distribution of $I_{1:n}$, which is approximated by a set of N_1 weighted particles as follows,

$$\hat{p}_\theta(i_{1:n}|y_{1:n}, z_{1:n}) := \sum_{k=1}^{|S_n|} W_n^k \delta_{I_{1:n}^k}(i_{1:n}), \quad (17)$$

where W_n^k is a so-called normalized importance weight associated with the kth particle $I_{1:n}^k$ at time n such that $\sum_{k=1}^{|S_n|} W_n^k = 1$. The delta indicator function $\delta_{I_{1:n}^k}(\cdot)$ takes value of 1 if $i_{1:n} \in I_{1:n}^k$, and 0 otherwise. Note that the above approximation requires the estimation of the conditional marginal likelihood $p_{\theta, i_{1:n}}(y_{1:n}, z_{1:n}|y_{1:n-1}, z_{1:n-1})$. However, due to its non-linear or/and non-Gaussian characteristics, the com-

$$p_\theta(i_{1:n}|y_{1:n}, z_{1:n}) = \frac{p_\theta(y_n, z_n|y_{1:n-1}, z_{1:n-1}, i_{1:n})f_\theta^I(i_n|i_{1:n-1})p_\theta(i_{1:n-1}|y_{1:n-1}, z_{1:n-1})}{\sum_{i'_{1:n} \in \mathcal{I}^n} p_\theta(y_n, z_n|y_{1:n-1}, z_{1:n-1}, i'_{1:n})f_\theta^I(i'_n|i'_{1:n-1})p_\theta(i_{1:n-1}|y_{1:n-1}, z_{1:n-1})}. \quad (16)$$

However, the computation involved in (16) still increases exponentially in n. Thus the DPF approximates the posterior distributions $\{p(\theta, i_{1:n}|y_{1:n}, z_{1:n}); 1 \leqslant n \leqslant T\}$ sequentially using a collection of $N_1|\mathcal{I}|$ weighted trajectories (so-called "particles"),

$$\left\{ I_{1:n}^{(k)}; k = 1, \cdots, N_1|\mathcal{I}| \right\},$$

where $|\mathcal{I}|$ is the cardinality of $\mathcal{I}$ and N_1 is the maximum number of discrete particles sampled at each time n.

Note that for a given n and $\theta \in \Theta$, the support of $p_\theta(i_{1:n}|y_{1:n}, z_{1:n})$ is $\mathcal{I}^n$. Specifically, at each time step n, we consider to resample N_1 of the $N_1|\mathcal{I}|$ trajectories and then adjust their weights accordingly. Let us denote by $S_1, S_2, \cdots, S_T$ the random support sets generated from the resampling step. In particular, each S_n

monly used Kalman filter techniques (Kalman 1960) cannot work here. Therefore, given θ and $I_{1:n} = i_{1:n}$, the SMC technique is adopted to deal with the continuous latent process $\{X_n\}_{n \geqslant 1}$ to approximate the density $p_{\theta, i_{1:n}}(y_{1:n}, z_{1:T}|y_{1:n-1}, z_{1:n-1})$.

The SMC algorithm aims to approximate the joint posterior density of the continuous latent process $X_{1:n}$, which is $p_{\theta, i_{1:n}}(x_{1:n}|y_{1:n}, z_{1:n})$ by a set of N_2 weighted random samples or particles by a discrete density probability distribution,

$$\hat{p}_{\theta, i_{1:n}}(x_{1:n}|y_{1:n}, z_{1:n}) := \sum_{k=1}^{N_2} \tilde{W}_n^k \delta_{X_{1:n}^k}(x_{1:n}), \quad (18)$$

where $\tilde{W}_n^k$ is the normalized weight for the kth particle at time n such that $\sum_{k=1}^{N_2} \tilde{W}_n^k = 1$, and $\tilde{W}_n^k = \frac{\tilde{w}_n(X_{1:n}^k)}{\sum_{m=1}^{N_2} \tilde{w}_n(X_{1:n}^m)}$

with $\tilde{w}_n(X_{1:n}^k)$ being the corresponding unnormalized weights. The $\delta_{X_{1:n}^k}(\cdot)$ takes value of 1 if $x_{1:n} \in X_{1:n}^k$, and 0 otherwise. Here N_2 is the number of particles $X_{1:n}^k$ sampled at each time n. Similar to N_1, the choice of N_2 controls the precision of the SMC algorithm. Thus, N_1 and N_2 together controls the precision of the combined DPF & SMC algorithm. Then, at time T, we can obtain the estimation of $p_{\theta, i_{1:T}}(y_{1:T}, z_{1:T})$ by SMC as,

$$\hat{p}_{\theta, i_{1:T}}(y_{1:T}, z_{1:T}) = \hat{p}_{\theta, i_1}(y_1, z_1) \prod_{n=2}^{T} \hat{p}_{\theta, i_{1:n}}(y_n, z_n | y_{1:n-1}, z_{1:n-1}),$$

$$\hat{p}_{\theta, i_{1:n}}(y_n, z_n | y_{1:n-1}, z_{1:n-1}) = \frac{1}{N_2} \sum_{k=1}^{N_2} \tilde{w}_n(x_{1:n}^k).$$

(19)

For the detailed procedure of the DPF or SMC algorithm, please refer to Andrieu, Doucet, and Holenstein (2010), Fearnhead (1998), and Whiteley, Andrieu, and Doucet (2010).

A pseudo code of the combined DPF & SMC is provided below. To alleviate the notational burden, we adopt the similar convention used in Andrieu, Doucet, and Holenstein (2010). Whenever the index k is used, it means "for all $k \in \{1, \cdots, N_2\}$" for the continuous particle $\{X_n\}_{n \geq 1}$. The dependence of weights on θ is also omitted for convenience.

Algorithm 1 (Combined DPF & SMC Algorithm).

Step 1: at time $n = 1$,

a. Set $\mathbf{S}_1 = \mathcal{I}$ and for each $i_1 \in \mathcal{I}$, obtain $\hat{p}_{\theta, i_1}(y_1, z_1)$ using the SMC algorithm as follows,

$$\tilde{w}_1(x_1^k) = g_{\theta, i_1}(y_1 | x_1^k) h_{\theta, i_1}(z_1 | x_1^k),$$
$$\hat{p}_{\theta, i_1}(y_1, z_1) = \frac{1}{N_1} \sum_{k=1}^{N_2} \tilde{w}_1(x_1^k).$$

(20)

b. Compute and normalize the weights for discrete particles. For each $i_1 \in \mathcal{I}$,

$$w_1(i_1) = v_\theta(i_1) \hat{p}_{\theta, i_1}(y_1, z_1),$$
$$W_1(i_1) = \frac{w_1(i_1)}{\sum_{i_1' \in \mathcal{I}} w_1(i_1')}.$$

(21)

Step 2: at times $n = 2, \cdots, T$,

a. If $|\mathbf{S}_{n-1}| \leqslant N_1$ set $C_{n-1} = \infty$ otherwise set C_{n-1} to the unique solution of

$$\sum_{i_{1:n-1} \in \mathbf{S}_{n-1}} 1 \wedge C_{n-1} W_{n-1}(i_{1:n-1}) = N_1$$

(22)

b. Maintain the L_{n-1} trajectories in $\mathbf{S}_{n-1}$ which have weights strictly superior to $1/C_{n-1}$, then apply the stratified resampling mechanism to the other trajectories to yield $N_1 - L_{n-1}$ survivors. Set $\mathbf{S}_{n-1}'$ to the set of surviving and maintained trajectories.

c. Set $\mathbf{S}_n = \mathbf{S}_{n-1}' \times \mathcal{I}$.

d. For each $i_{1:n} \in \mathbf{S}_n$, obtain $\hat{p}_{\theta, i_{1:n}}(y_n, z_n | y_{1:n-1}, z_{1:n-1})$ using the SMC algorithm by,

$$\tilde{w}_n(x_{1:n}^k) = g_{\theta, i_{1:n}}(y_n | x_n^k) h_{\theta, i_{1:n}^k}(z_n | x_n^k),$$
$$\hat{p}_{\theta, i_{1:n}}(y_n, z_n | y_{1:n-1}, z_{1:n-1}) = \frac{1}{N_1} \sum_{k=1}^{N_1} \tilde{w}_n(x_{1:n}^k).$$

(23)

e. Compute and normalize the weights. For each $i_{1:n} \in \mathbf{S}_n$,

$$w_n(i_{1:n}) = f_\theta^I(i_n | i_{1:n-1}) \hat{p}_{\theta, i_{1:n}}(y_n, z_n | y_{1:n-1}, z_{1:n-1})$$
$$\frac{W_{n-1}(i_{1:n-1})}{1 \wedge C_{n-1} W_{n-1}(i_{1:n-1})},$$
$$W_n(i_{1:n}) = \frac{w_n(i_{1:n})}{\sum_{i_{1:n}' \in S_n} w_n(i_{1:n}')}.$$

(24)

In addition, the proposed combined DPF & SMC algorithm also provides an estimate of the marginal likelihood $p_\theta(y_{1:T}, z_{1:T})$ given by

$$\hat{p}_\theta(y_{1:T}, z_{1:T}) = \hat{p}_\theta(y_1, z_1) \prod_{n=2}^{T} \hat{p}_\theta(y_n, z_n | y_{1:n-1}, z_{1:n-1}),$$

(25)

where

$$\hat{p}_\theta(y_1, z_1) = \sum_{i_1 \in \mathcal{I}} w_1(i_1),$$
$$\hat{p}_\theta(y_n, z_n | y_{1:n-1}, z_{1:n-1}) = \sum_{i_{1:n} \in \mathbf{S}_n} w_n(i_{1:n}), n > 1.$$

(26)

Note that $p_\theta(i_{1:n} | y_{1:n}, z_{1:n})$ can be computed exactly when n is small. However, when n is large enough with $|\mathbf{S}_{n-1}| > N_1$, the stratified resampling mechanism is employed to prune the set of trajectories. One can refer to Whiteley, Andrieu, and Doucet (2010) for more details about the resampling techniques.

3.2. Particle marginal metropolis–hastings sampler

A popular choice for sampling from $p(\theta, X_{1:T}, I_{1:T} | y_{1:T}, z_{1:T})$ is the particle marginal Metropolis-Hastings (PMMH) widely recommended in literatures for efficient parameter estimation and

inference (Andrieu, Doucet, and Holenstein 2010; Whiteley, Andrieu, and Doucet 2010). The PMMH sampler can jointly update the unknown static parameters $\boldsymbol{\theta}$, and two latent processes $\boldsymbol{X}_{1:T}$ and $\boldsymbol{I}_{1:T}$ by using the proposal density below,

$$
\begin{aligned}
q\{(\boldsymbol{\theta}^*, \boldsymbol{x}_{1:T}^*, \boldsymbol{i}_{1:T}^*)|(\boldsymbol{\theta}, \boldsymbol{x}_{1:T}, \boldsymbol{i}_{1:T})\} \\
= q(\boldsymbol{\theta}^*|\boldsymbol{\theta})p_{\boldsymbol{\theta}^*}(\boldsymbol{x}_{1:T}^*, \boldsymbol{i}_{1:T}^*|\boldsymbol{y}_{1:T}, \boldsymbol{z}_{1:T}).
\end{aligned}
$$

In this scenario, the proposed $\boldsymbol{x}_{1:T}^*$ and $\boldsymbol{i}_{1:T}^*$ are perfectly "adapted" to the proposed $\boldsymbol{\theta}^*$, and the only degree of freedom of the algorithm is $q(\boldsymbol{\theta}^*|\boldsymbol{\theta})$. The resulting MH acceptance ratio is given by

$$
\begin{aligned}
&\frac{p(\boldsymbol{\theta}^*, \boldsymbol{x}_{1:T}^*, \boldsymbol{i}_{1:T}^*|\boldsymbol{y}_{1:T}, \boldsymbol{z}_{1:T})}{p(\boldsymbol{\theta}, \boldsymbol{x}_{1:T}, \boldsymbol{i}_{1:T}|\boldsymbol{y}_{1:T}, \boldsymbol{z}_{1:T})} \frac{q\{(\boldsymbol{\theta}, \boldsymbol{x}_{1:T}, \boldsymbol{i}_{1:T})|(\boldsymbol{\theta}^*, \boldsymbol{x}_{1:T}^*, \boldsymbol{i}_{1:T}^*)\}}{q\{(\boldsymbol{\theta}^*, \boldsymbol{x}_{1:T}^*, \boldsymbol{i}_{1:T}^*)|(\boldsymbol{\theta}, \boldsymbol{x}_{1:T}, \boldsymbol{i}_{1:T})\}} \\
&= \frac{p_{\boldsymbol{\theta}^*}^*(\boldsymbol{y}_{1:T}, \boldsymbol{z}_{1:T})p(\boldsymbol{\theta}^*)}{p_{\boldsymbol{\theta}}(\boldsymbol{y}_{1:T}, \boldsymbol{z}_{1:T})p(\boldsymbol{\theta})} \frac{q(\boldsymbol{\theta}|\boldsymbol{\theta}^*)}{q(\boldsymbol{\theta}^*|\boldsymbol{\theta})}.
\end{aligned}
\tag{27}
$$

The ratio in (27) suggests that the PMMH algorithm effectively targets the marginal density $p(\boldsymbol{\theta}|\boldsymbol{y}_{1:T}, \boldsymbol{z}_{1:T}) \propto p_{\boldsymbol{\theta}}(\boldsymbol{y}_{1:T}, \boldsymbol{z}_{1:T})p(\boldsymbol{\theta})$, justifying the terminology of marginal Metropolis-Hastings. Moreover, it bypasses the difficulty of sampling from $p(\boldsymbol{\theta}, \boldsymbol{x}_{1:T}, \boldsymbol{i}_{1:T}|\boldsymbol{y}_{1:T}, \boldsymbol{z}_{1:T})$ by sampling from $p(\boldsymbol{\theta}|\boldsymbol{y}_{1:T}, \boldsymbol{z}_{1:T})$, which is typically defined on a much smaller space and can be approximated using Algorithm 1. The proposed PMMH sampler is summarized as follows.

Algorithm 2 (PMMH Sampler).

Step 1: *initialization, $j = 0$,*

a. *set $\boldsymbol{\theta}(0)$ arbitrarily and*
b. *run **Algorithm 1** targeting both $p_{\boldsymbol{\theta}(0)}(\boldsymbol{y}_{1:T}, \boldsymbol{z}_{1:T})$ and $p_{\boldsymbol{\theta}(0)}(\boldsymbol{i}_{1:T}|\boldsymbol{y}_{1:T}, \boldsymbol{z}_{1:T})$, sample $\boldsymbol{i}_{1:T}(0) \sim \hat{p}_{\boldsymbol{\theta}(0)}(\cdot|\boldsymbol{y}_{1:T}, \boldsymbol{z}_{1:T})$ and let $\hat{p}_{\boldsymbol{\theta}(0)}(\boldsymbol{y}_{1:T}, \boldsymbol{z}_{1:T})$ denote the marginal likelihood estimate.*

Step 2: *for iteration $j \geqslant 1$,*

a. *sample $\boldsymbol{\theta}^* \sim q\{\cdot|\boldsymbol{\theta}(j-1)\}$,*
b. *run **Algorithm 1** targeting both $p_{\boldsymbol{\theta}^*}(\boldsymbol{y}_{1:T}, \boldsymbol{z}_{1:T})$ and $p_{\boldsymbol{\theta}^*}(\boldsymbol{i}_{1:T}|\boldsymbol{y}_{1:T}, \boldsymbol{z}_{1:T})$, sampling $\boldsymbol{i}_{1:T}^* \sim \hat{p}_{\boldsymbol{\theta}^*}(\cdot|\boldsymbol{y}_{1:T}, \boldsymbol{z}_{1:T})$, and let $\hat{p}_{\boldsymbol{\theta}^*}(\boldsymbol{y}_{1:T}, \boldsymbol{z}_{1:T})$ denote the marginal likelihood estimate, and*
c. *with probability*

$$
1 \wedge \frac{\hat{p}_{\boldsymbol{\theta}^*}(\boldsymbol{y}_{1:T}, \boldsymbol{z}_{1:T})p(\boldsymbol{\theta}^*)}{\hat{p}_{\boldsymbol{\theta}(j-1)}(\boldsymbol{y}_{1:T}, \boldsymbol{z}_{1:T})p\{\boldsymbol{\theta}(j-1)\}} \frac{q\{\boldsymbol{\theta}(i-1)|\boldsymbol{\theta}^*\}}{q\{\boldsymbol{\theta}^*|\boldsymbol{\theta}(j-1)\}}
\tag{28}
$$

set $\boldsymbol{\theta}(j) = \boldsymbol{\theta}^*, \boldsymbol{i}_{1:T}(j) = \boldsymbol{i}_{1:T}^*$ and $\hat{p}_{\boldsymbol{\theta}(j)}(\boldsymbol{y}_{1:T}, \boldsymbol{z}_{1:T}) = \hat{p}_{\boldsymbol{\theta}^*}(\boldsymbol{y}_{1:T}, \boldsymbol{z}_{1:T})$; otherwise set $\boldsymbol{\theta}(j) = \boldsymbol{\theta}(j-1), \boldsymbol{i}_{1:T}(j) = \boldsymbol{i}_{1:T}(j-1)$ and $\hat{p}_{\boldsymbol{\theta}(j)}(\boldsymbol{y}_{1:T}, \boldsymbol{z}_{1:T}) = \hat{p}_{\boldsymbol{\theta}(j-1)}(\boldsymbol{y}_{1:T}, \boldsymbol{z}_{1:T})$.

---

4. Simulation study

In this section, simulation studies are conducted to evaluate the performance of the proposed method for change-point detection. Four scenarios of mixed-type data based on the *mixed SSSM* are considered: (S1) mixed Gaussian-Bernoulli, (S2) mixed Gaussian-Poisson, (S3) mixed Gaussian-Gaussian, and (S4) mixed Gaussian-Noncentral *t*. For each scenario, we consider one change-point as well as multiple change-points. Moreover, we also consider three different locations where the change occurs: at the beginning of the time period, in the middle of the time period, and at the end of the time period. Each simulation setting is repeated for 50 iterations.

4.1. Data generation

For simplicity, we assume that Y_n is a continuous random variable following a Gaussian distribution, while Z_n is a discrete random variable following either Bernoulli in (S1) or Poisson distribution in (S2). Since the proposed method also can accommodate the situation of both Y_n and Z_n being continuous, we consider that Z_n follows Gaussian in (S3) or noncentral *t* distribution in (S4). Suppose $\mathcal{I} = \{0, 1\}$, which means that there are two states for the Markov chain, $\{I_n\}$. The transition matrix P_I is

$$
P_I = \begin{bmatrix} p_1 & 1-p_1 \\ 1-p_2 & p_2 \end{bmatrix},
\tag{29}
$$

where $P(I_n = 1|I_{n-1} = 0) = 1 - p_1, P(I_n = 0|I_{n-1} = 0) = p_1, P(I_n = 0|I_{n-1} = 1) = 1 - p_2, P(I_n = 1|I_{n-1} = 1) = p_2$, for $n > 1$. Note that p_1 and p_2 are also unknown parameters need to be estimated together with other unknown parameters. And I_n takes different values before and after each true change-point position τ_i, $i = 1, \cdots, k$. Four scenarios of data generation based on the *mixed SSSM* are listed as follows,

(S1) Mixed Gaussian-Bernoulli

$$
\begin{aligned}
x_{n+1} &= \phi x_n + \sigma I_n V_n, \\
y_n &= x_n + \gamma V_n, \\
z_n &\sim Bern(p(I_n)),
\end{aligned}
$$

where $p(I_n) \sim Unif[0.1 + 0.5I_n, 0.4 + 0.5I_n]$. Thus, if $I_n = 0$, the parameter p in the Bernoulli distribution

Table 1. Simulation results for one change-point scenario: the median of $\hat{\tau}$ over 50 iterations.

	S1			S2			S3			S4		
T	21	51	91	21	51	91	21	51	91	21	51	91
$\hat{\tau}$	22	53	93	23	53	93	23	53	93	23	53	93
MAD ($\hat{\tau}$)	1	1	1	0	0	0	1	0	0.5	1	0.5	1
Acceptance rate of MCMC	0.24	0.23	0.20	0.18	0.18	0.20	0.19	0.22	0.25	0.28	0.28	0.32

follows a uniform distribution with support $[0.1, 0.4]$. While, if $I_n = 1$, the support is $[0.6, 0.9]$.

(S2) **Mixed Gaussian-Poisson**

$$x_{n+1} = \phi x_n + \sigma I_n V_n,$$
$$y_n = x_n + \gamma V_n,$$
$$z_n \sim Poisson\left(\sqrt{|x_n|}\right).$$

(S3) **Mixed Gaussian-Gaussian**

$$x_{n+1} = \phi x_n + \sigma I_n V_n,$$
$$y_n = x_n + \gamma V_n,$$
$$z_n = x_n + V_n.$$

(S4) **Mixed Gaussian-Noncentral t**

$$x_{n+1} = \phi x_n + \sigma I_n V_n,$$
$$y_n = x_n + \gamma V_n,$$
$$z_n \sim T(df, x_n),$$

where $T(df, \delta)$ denotes the Noncentrol t distribution with degree of freedom df and noncentrality parameter δ. In the simulation example, $df = 4$ will be used.

In the above four scenarios, $\{V_n\}$ are i.i.d. standard normal random variables, and $\boldsymbol{\theta} = (\phi, \sigma, \gamma, p_1, p_2)$. For the initialization of the *mixed SSSM*, we set the initial distribution $X_1 \sim \mathcal{N}(0, 1)$. The initial distribution of I_n is Bernoulli with probability 0.01, i.e., $I_1 \sim Bern(0.01)$. Such a small value in the initial density is used to ensure that the initial state is always 0 ($I_1 = 0$) for easy comparison. Based on the true change-points $\tau_1, \cdots, \tau_k$, we first generate the state process $\boldsymbol{I}_{1:T}$, where T is set as $T = 100$. Then two sets of observations $(\boldsymbol{y}_{1:T}, \boldsymbol{z}_{1:T})$ are generated according to the corresponding *mixed SSSMs* in (S1)-(S4) with $\phi = 0.9, \sigma = 4, \gamma = 1$. Here we use a proper proposal for combined DPF & SMC sampling, i.e., $q_\theta(x_1) = \mu_\theta(x_1)$ and $q_\theta(x_n|y_n, z_n, x_{n-1}) = f_\theta^X(x_n|x_{n-1})$ for $n = 2, \cdots, T$. For the prior of $\boldsymbol{\theta}$, the following independent priors are used, $logit(\phi) \sim \mathcal{N}(\mu_\phi, \sigma_\phi^2)$, $\log(\sigma) \sim \mathcal{N}(\mu_\sigma, \sigma_\sigma^2)$, and $\gamma \sim \mathcal{N}(\mu_\gamma, \sigma_\gamma^2)$, where $logit(x) = \log\left(\frac{x}{1-x}\right)$. Such settings of priors are commonly used in the literature (Gelman et al. 1995; Kang et al. 2018, 2021; Chen et al. 2023). The values of hyper-parameters are set as $\mu_\phi = 1, \sigma_\phi^2 = 2, \mu_\sigma = 1.5, \sigma_\sigma^2 = 2,$ and $\mu_\gamma = 2, \sigma_\gamma^2 = 2$. For $0 < p_1 < 1$ and $0 < p_2 < 1$, the same Dirichlet

distribution $Dir([1\ 1])$ is used. Let $N_1 = 128$ and $N_2 = 200$ in Algorithm 1. The number of MCMC simulation is $N = 300$ with burn-in 200. A normal random-walk Metropolis-Hastings proposal is used in Algorithm 2 to update the parameters jointly, with the covariance of the proposal proportional to the identity matrix up to some small constant, such as 0.3.

4.2. Results for one change-point detection

For the case of only one change-point τ, the transition matrix needs certain restriction such that I_n can only change from 0 to 1 once, and not the other way around. This can be accomplished by setting $P(I_n = 0|I_{n-1} = 1) = 0$. Thus the transition matrix in (29) can be rewritten as

$$P_I = \begin{bmatrix} p & 1-p \\ 0 & 1 \end{bmatrix}. \tag{30}$$

Recall that the data contain $T = 100$ observations collected at locations $1, \cdots, T$, respectively. If the change-point location is τ, then the state becomes,

$$I_n = \begin{cases} 0 & \text{for} & n = 1, \cdots, \tau, \\ 1 & \text{for} & n = \tau + 1, \cdots, T. \end{cases} \tag{31}$$

For each scenario, we consider three different change-point locations $\tau = 21, 51, 91$, respectively. Here the unknown parameters $\boldsymbol{\theta} = (\phi, \sigma, \gamma)$ are set as $\phi = 0.9, \sigma = 4.0, \gamma = 1.0$. Table 1 reports the median of the estimated values of ϕ, σ, γ and change-point τ based on 50 iterations. The median absolute deviation (MAD) of the estimates τ is also reported based on 50 iterations. From the results in Table 1, it is seen that the estimated change-points are accurate with only one or two times delay for all scenarios. And the MADs are small ranging from zero to one time unit over all scenarios.

Moreover, the acceptance rates are reported in Table 1. The acceptance rates are well maintained around 20% to 32%, which is within a reasonable range for efficient MCMC updates. A high acceptance rate means that the Markov chain is moving slowly and not fully exploring the parameter space. On the other hand, a low acceptance rate indicates that the proposed samples are always rejected and the chain
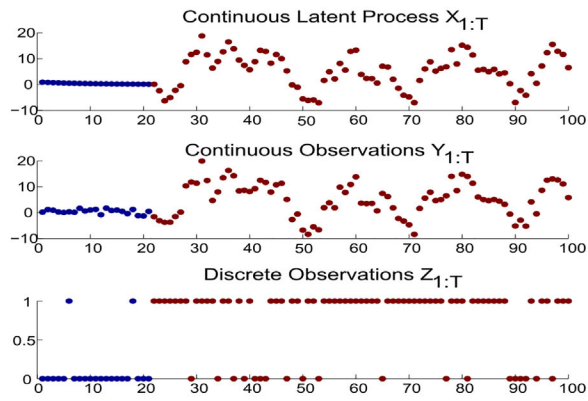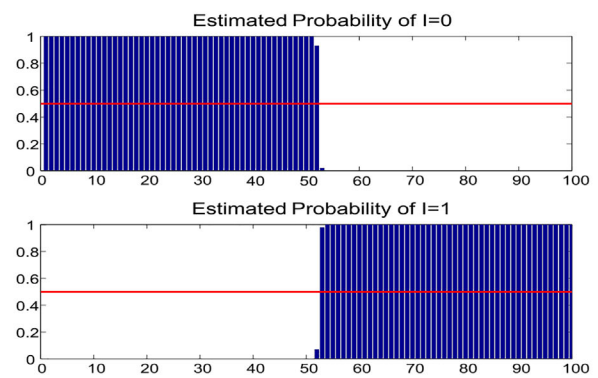
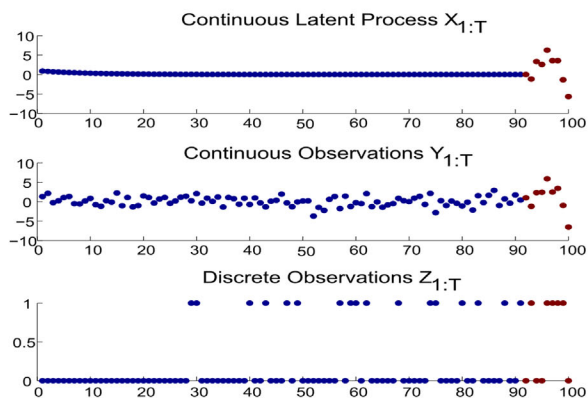(a) Simulation Data with $\tau = 21$
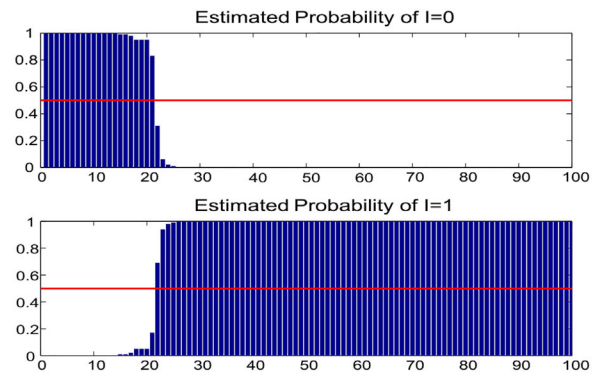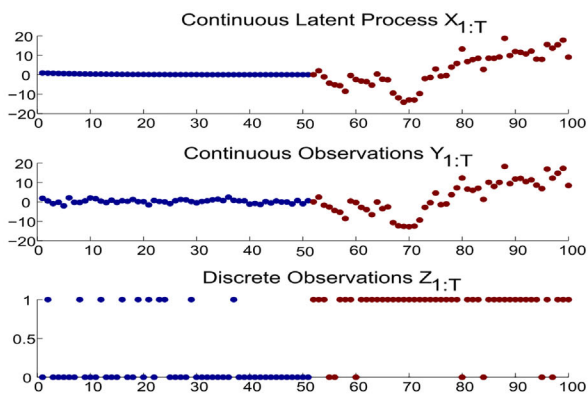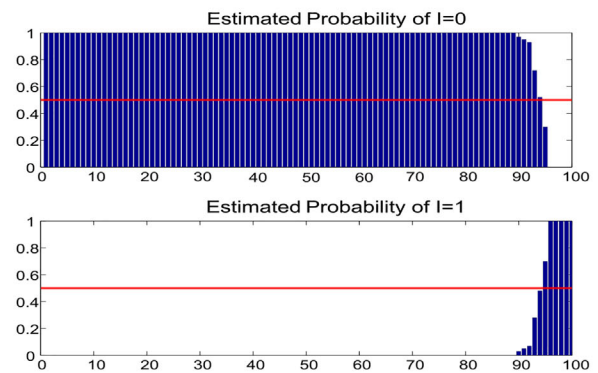
(b) Change point estimation

(c) Simulation Data with $\tau = 51$

(d) Change point estimation

(e) Simulation Data with $\tau = 91$

(f) Change point estimation

**Figure 1.** One change-point detection for Gaussian-Bernoulli (S1).

may fail to converge. An efficient Metropolis sampler should have an acceptance rate that is neither too high nor too low. Roberts, Gelman, and Gilks (1997) shows that for random-walk Metropolis algorithms,

the optimal acceptance probability for the Markov chain should be around 23% in high dimensions.

In addition, Figure 1 shows the simulated data and estimated probability of states in one simulation for

Table 2. Simulation results for one change-point scenario of different σ values in (S1): the median of $\hat{\tau}$ over 50 iterations.

	21			51			91		
τ									
σ	4	1	0.1	4	1	0.1	4	1	0.1
$\hat{\tau}$	22	21	21	53	52	51	93	93	92
MAD ($\hat{\tau}$)	1.5	3.0	5.2	1.5	3.0	3.0	1.5	2.2	4.4
Acceptance rate of MCMC	0.24	0.22	0.22	0.23	0.20	0.19	0.19	0.18	0.17

(a) $\tau = 21$　　　　(b) $\tau = 51$　　　　(c) $\tau = 91$
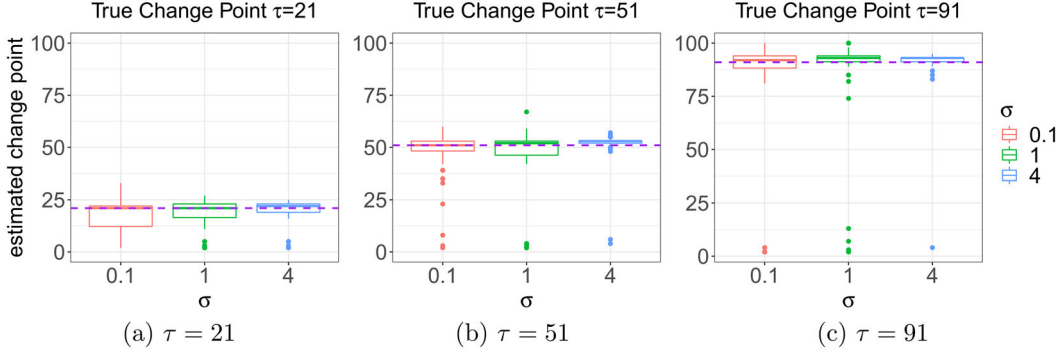
Figure 2. Boxplots of estimated change point in one change-point scenario for different σ values in (S1). Purple lines are the true change point $\tau = 21, 51, 91$.

Table 3. Simulation results for two change-points scenario: the median of $\hat{\tau}$ over 50 iterations.

	S1			S2		
T	{21,51}	{51,91}	{21,91}	{21,51}	{51,91}	{21,91}
$\hat{\tau}$	{22, 52}	{52, 92}	{22, 92}	{22, 52}	{52, 92}	{22, 92}
MAD ($\hat{\tau}$)	{1, 1}	{1, 1}	{1, 1}	{0, 1}	{0, 1}	{1, 1}
Acceptance rate of MCMC	0.20	0.19	0.20	0.22	0.33	0.23
	S3			S4		
T	{21,51}	{51,91}	{21,91}	{21,51}	{51,91}	{21,91}
$\hat{\tau}$	{22, 52}	{52, 92}	{22, 92}	{22, 52}	{52, 92}	{22, 92}
MAD ($\hat{\tau}$)	{0, 0}	{0, 1}	{0, 1}	{0.5, 1}	{1, 2 }	{0, 1}
Acceptance rate of MCMC	0.26	0.28	0.28	0.31	0.35	0.34

(S1). Similar plots are reported in Figures B1–B3 for (S2)-(S4) in Appendix B. Taking Figure 1 as an example, the left panels of the figure include the continuous latent process $X_{1:T}$, the continuous observation $Y_{1:T}$ and the other observation sequence $Z_{1:T}$. For $1 \leqslant n \leqslant T$, the probabilities of $p(I_n = 1)$ and $p(I_n = 0) = 1 - p(I_n = 1)$ are calculated based on 100 estimates of the latent process $I_{1:T}$ from MCMC simulations, and their plots are shown in the right panels of the figure. These plots show that the process is in state 0 at the beginning with a high probability of $p(I_1 = 0)$, and the high probability value is maintained until a certain time point when it decreases drastically to almost 0. Thus we can identify the change point accurately.

Note that the values of N_1 and N_2 control the accuracy of the DPF and SMC algorithms, respectively. The larger their values are, the more accurate approximation of the target distribution will be. We have also conducted additional simulations to investigate the choice of N_1 and N_2 in Appendix A. Furthermore, we investigate the effect on the setting of σ for the performance of

the proposed methods under the mixed Gaussian-Bernoulli scenario. Specifically, for each value of τ, we examine three different values of $\sigma = 4, 1,$ and 0.1. Table 2 reports the median value for the estimated change-point τ and the median absolute deviation (MAD) of the estimates τ based on 50 iterations. The results indicate that the estimated change-points are generally accurate, with a delay of only one or two time steps. However, as σ decreases, the MADs increase, indicating that the variability of estimation increases as the magnitude of the change decreases. Such a pattern can also be observed by Figure 2, where the interquartile range for the case of $\sigma = 0.1$ is larger than other cases. Moreover, as σ decreases, the acceptance rates decrease. It indicates that the MCMC samples are more likely to be rejected, and the chain may fail to converge when the value of σ is small.

4.3. Results for multiple change-point detection

Under the multiple change-point scenario, the key difference from the one change-point detection is the

Table 4. Simulation results for three change-points scenario: the median of $\hat{\tau}$ over 50 iterations.

	S1			S2		
T	{21, 51, 91}	{17, 32, 77}	{21, 52, 72}	{21, 51, 91}	{14, 51, 90}	{37, 63, 88}
$\hat{\tau}$	{22, 52, 91}	{18, 33, 78}	{22, 53, 72}	{22, 52, 92}	{15, 52, 91}	{38, 64, 89}
MAD ($\hat{\tau}$)	{1, 1, 1}	{1, 1, 1}	{1, 1, 1}	{0, 1, 0}	{0, 1, 0.5}	{0, 1, 0}
Acc. rate	0.29	0.30	0.30	0.20	0.20	0.20
	S3			S4		
T	{21, 51, 91}	{12, 22, 75}	{42, 61, 94}	{21, 51, 91}	{7, 37, 64}	{41, 76, 85}
$\hat{\tau}$	{22, 52, 92}	{13, 23, 76}	{43, 62, 95}	{22, 52, 92}	{8, 38, 66}	{42, 77, 86}
MAD ($\hat{\tau}$)	{0, 1, 1.5}	{0, 0 0, 0}	{0, 0.5, 0}	{0.5, 1, 0}	{1, 2, 0.5}	{0.5, 2, 1}
Acc. rate	0.26	0.27	0.27	0.31	0.33	0.34

structure of the transition matrix P_I. Now the state parameter I_n can change from state 0 to 1 and then back to 0. For simplicity, let $p_1 = p_2$ in (29), and then the transition matrix can be specified as,

$$P_I = \begin{bmatrix} p & 1-p \\ 1-p & p \end{bmatrix}. \qquad (32)$$

The unknown parameters are still the same, $\theta = (\phi, \sigma, \gamma, p)$. We consider the cases of two change-points and three change-points, respectively. Three sets of true change-points (τ) are selected under each case. Tables 3 and 4 report the estimation results for two change-points and three change-points based on 50 iterations, respectively. In general, the estimated change-points are all close to the truth within one time unit delay. The largest value of MAD is about two time units. The estimations of static parameters are also very close to the true values. Figures 3 and 4 illustrate the simulated data and the change-point estimation for two and three change-points for one simulation under (S1), respectively.

5. Case study of civil unrest data

In this section, the proposed method is applied for change-point detection of the civil unrest data. Social events, such as strikes, protests and fights are happening every day across the globe. Millions of lives could be saved or protected if these kinds of social unrest events could be known in advance. Two datasets, ICEWS (Integrated Conflict Early Warning System) and GDELT (Global Database of Events, Language, and Tone) provide a global database of political and social events and have been used to develop predictive systems Cadena et al. (2015); Ramakrishnan et al. (2015, 2014).

GDELT is a global database of events which has been coded from vast quantities of publicly available text that is produced by the world's news media. ICEWS, on the other hand, is an early warning system designed to help US policy analysts understand, monitor and predict national and international crises to which the US might have to respond. These include

international and domestic crisis, ethnic and religious violence, as well as rebellion and insurgency. GDELT and ICEWS are based on similar, though not identical methods and sources. Both data sets use Conflict and Mediation Event Observations (CAMEO) coding for recording events. There are 20 event types in total, and it has an ordinal increase in cooperation as one goes from category 01 to 09, and an ordinal increase in conflict as one goes from 10 to 20. In this work, we will mainly focus on protests (the 14th event type) in three Latin America countries, Argentina, Brazil, and Venezuela, since these countries often encounter events related to civil unrest.

The mixed observation used for analysis is weekly binary observations indicating whether protests occurred during a week or not, and weekly continuous observations measuring the AverageTone of all protests in each week. The value of AverageTone is the average tone of all documents containing one or more mentions of an event. The value ranges from -10 and $+10$, with 0 indicating neutral. The events and tone are highly related because if a protest has an extremely negative average tone, it suggests a far more serious occurrence of civil unrest, which is likely to spread spatially and may even become violent.

The goal is to detect change times (in weeks) when more frequent or serious protests are going to happen. The time period of the data used is from January 2010 to April 2014. Within this time period, a sequence of protests that were clustered around a time period occurred in each of the three countries, as recorded in Wikipedia. These protests are: (1) the protests during September 2012, at Cacerolazo in Argentina, (2) protests in Brazil during April and July 2013, (3) protests in Venezuela started on February 12, 2014 and are ongoing. Details about these protests can be found in Wikipedia (2015), Wikipedia (2016a), and Wikipedia (2016b).
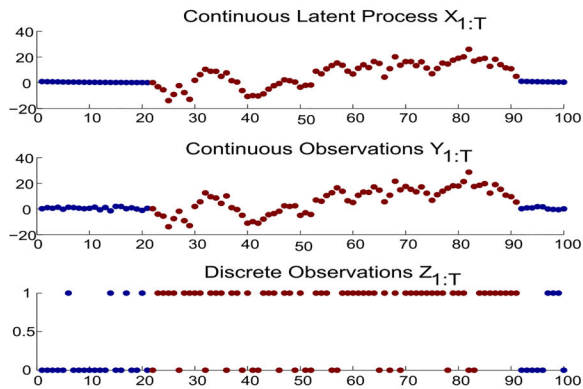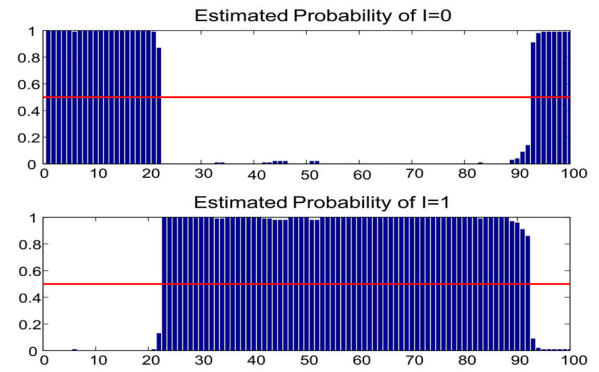
By applying the proposed method using the Gaussian-Bernoulli setting with one change-point defined in Section 4.1, the change-point in each
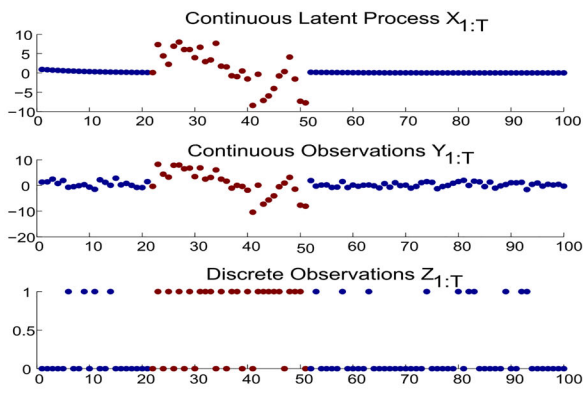
(a) Simulation Data with $\tau = 21, 51$
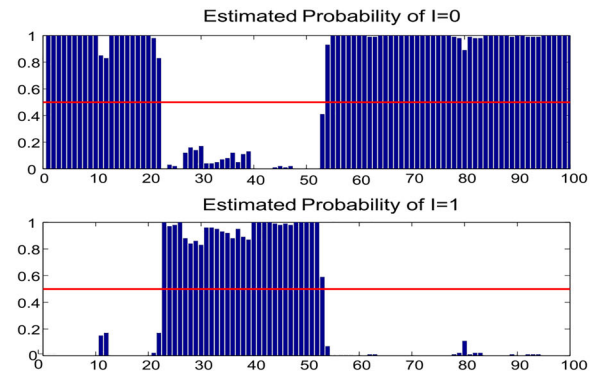
(b) Change point estimation
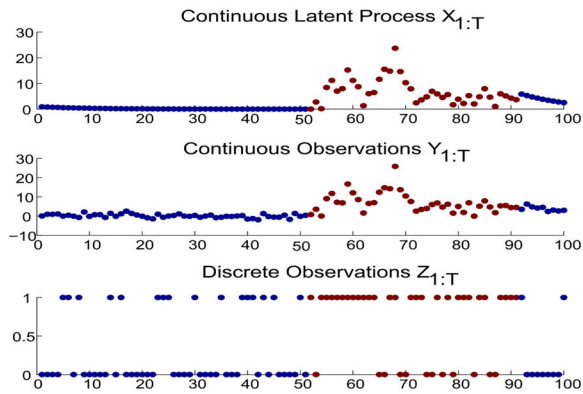
(c) Simulation Data with $\tau = 51, 91$
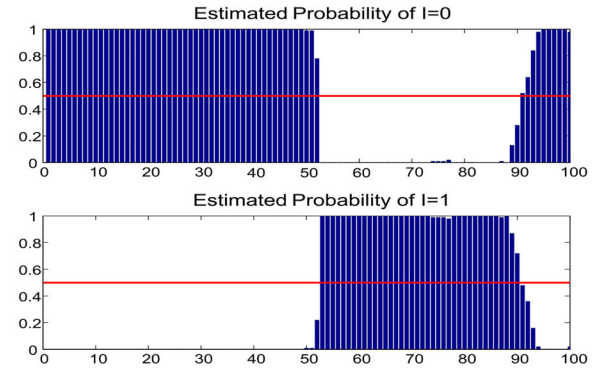
(d) Change point estimation

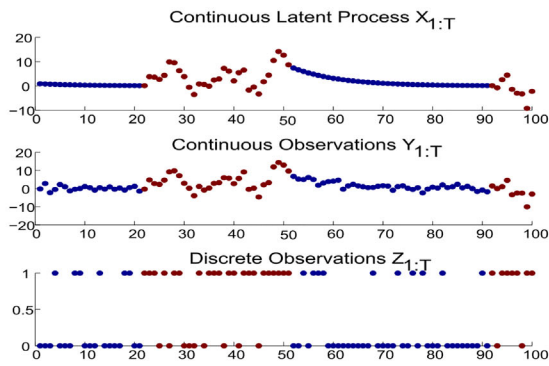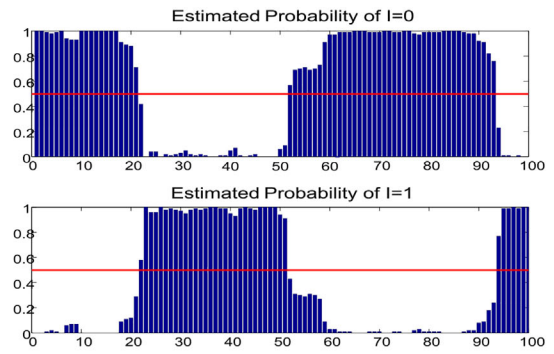(e) Simulation Data with $\tau = 21, 91$

(f) Change point estimation

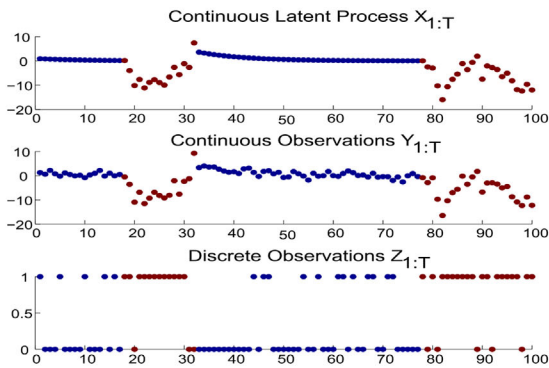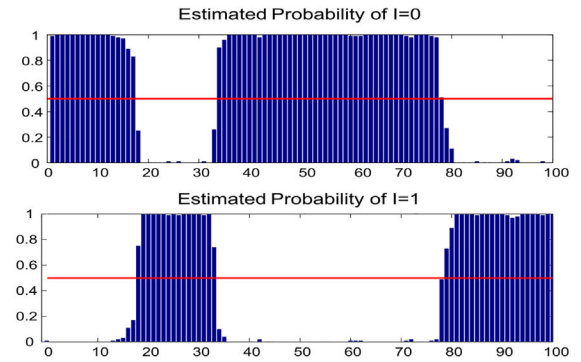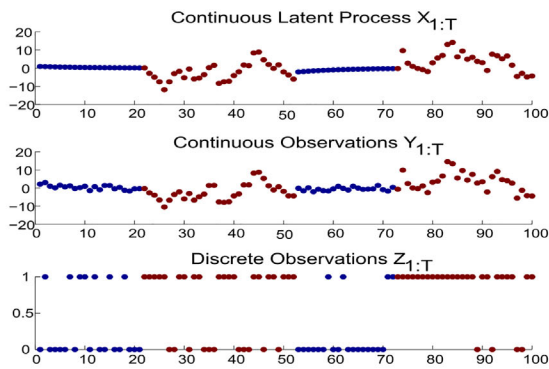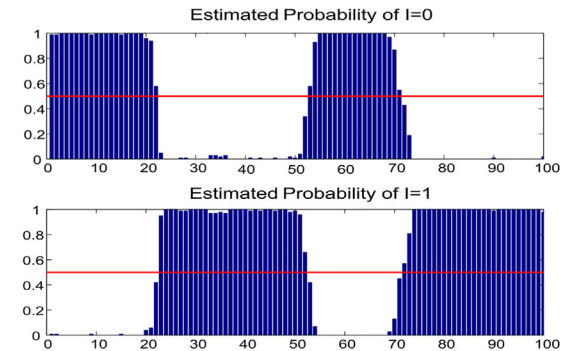**Figure 3.** Two change-points detection for Gaussian-Bernoulli (S1).

country is successfully detected. Figure 5 shows the mixed observations and the change-point detection plots for each country. The change-points indicate the weeks after which those three sequences of protests would happen. We summarize the analysis results as follows. (1) In Argentina, the estimated change-point occurred around week 09/09/2012, which is the same week as recorded in Wikipedia (2015). (2) In Brazil, the estimated week is 02/24/2013, and it is five weeks before the so-called Brazilian Spring movement (Wikipedia 2016a). (3) In Venezuela, the estimated week is around 01/26/2014, which is just one week before the February protests reported in Wikipedia (2016b).

(a) Simulation Data with $\tau = 21, 51, 91$

(b) Change-points estimation

(c) Simulation Data with $\tau = 17, 32, 77$

(d) Change-points estimation

(e) Simulation Data with $\tau = 21, 52, 72$

(f) Change-points estimation

Figure 4. Three change-points detection for Gaussian-Bernoulli (S1).

5.1. Comparison

Since the proposed method can make full use of both the discrete and continuous observations, it is expected to detect changes more efficiently than those only one type of observations. To show this advantage, we compare the proposed method with two conventional methods: (M1) SSSM only using continuous observations, and (M2) SSSM only using binary observations. The change-point detection results are reported in Table 5. From the results in the table, we can see that the M1 method, which only uses the continuous observations, cannot detect any changes at all for three countries. For the M2 method which only makes use of binary observations, changes can still be detected for Argentina and Venezuela, but not for Brazil. In contrast, the proposed *mixed SSSM* can always effectively detect changes for all three
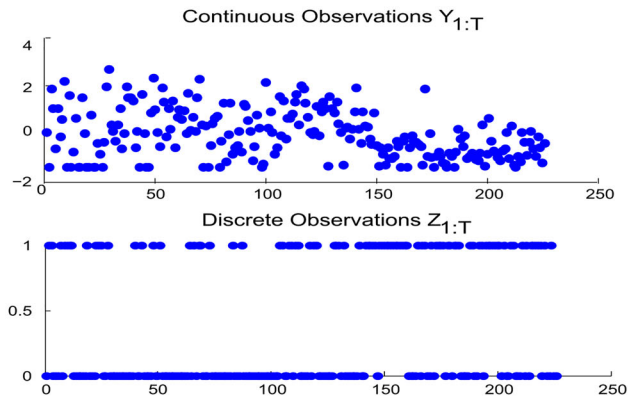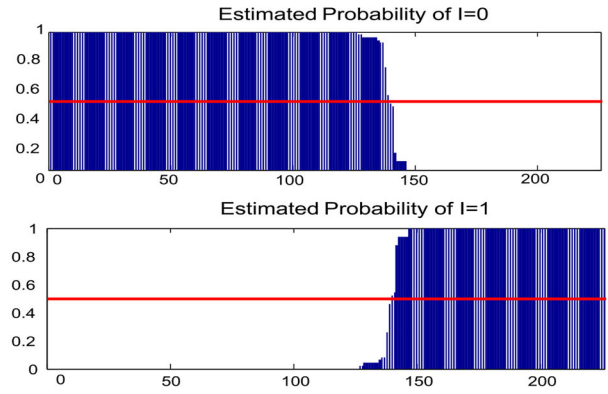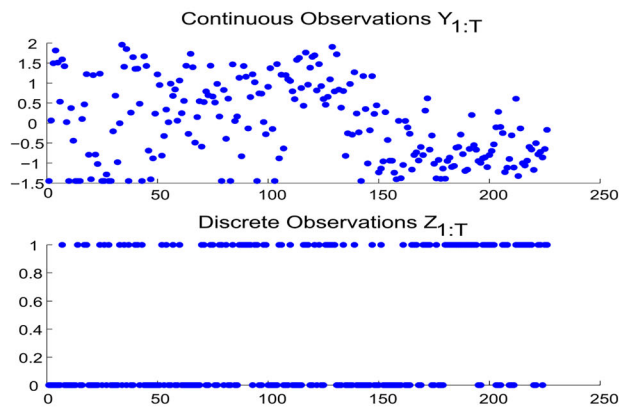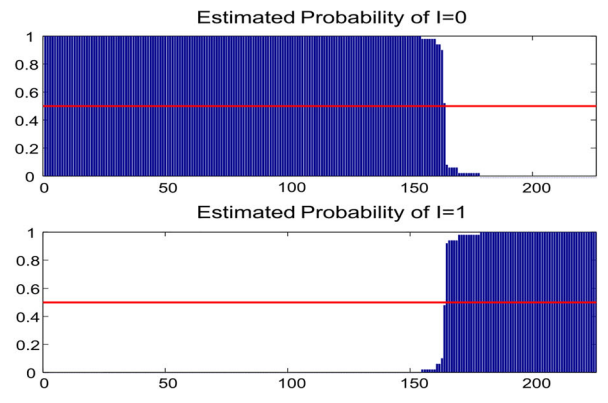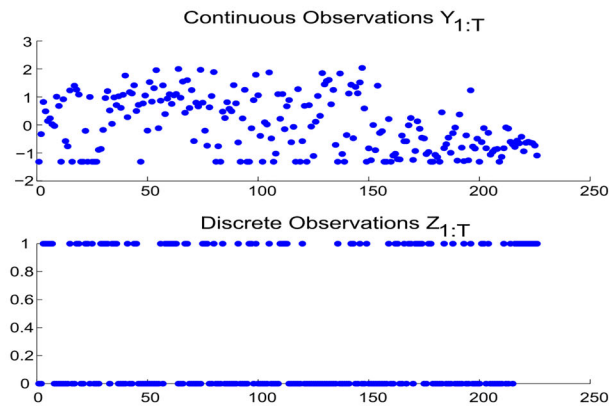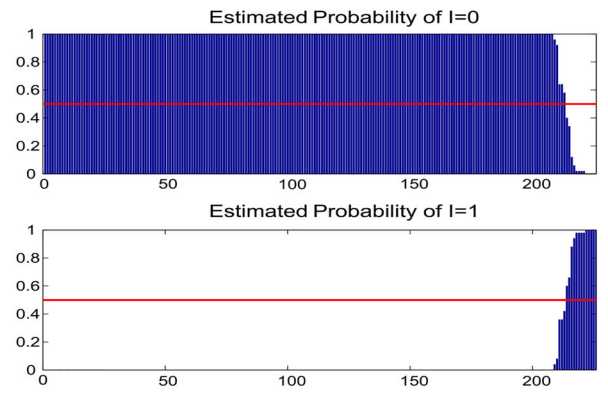
Figure 5. Protests detection for Argentina, Brazil, and Venezuela.

countries. One possible explanation is that the binary observations could contain sufficient information to detect the changes for Argentina and Venezuela. But for Brazil, neither the continuous nor the binary observations alone can provide enough knowledge about the big onset of protests. Thus, by combing these two types of information together, the proposed method can make a better detection of the change(s).

Table 5. Estimated change-points comparison for the proposed *mixed SSSM*, the M1 (SSSM only using continuous observations), and M2 (SSSM only using binary observations).

	Reported protest date	*Mixed SSSM*	M1	M2
Argentina	September 2012	week 09/09/2012	week 09/09/2012	None
Brazil	April 2013	week 02/24/2013	None	None
Venezuela	February 12, 2014	week 01/26/2014	week 01/26/2014	None

6. Discussion

The mixed-type data problem attracts broad interests because of its increasing popularity in modern society. However, dealing with mixed-type data is quite challenging due to the difficulty of quantifying the dependency between continuous and discrete variables in an appropriate manner. The proposed method models the mixed-type data by the latent processes in state-space models. Bayesian estimation and inference are conducted efficiently by developing a combined DPF & SMC algorithm.

The proposed method can accommodate various settings of mixed data such as mixed Gaussian-Bernoulli and mixed Gaussian-Poisson under different change-point scenarios. Both numerical examples and real case studies are analyzed to elaborate the merits of the proposed method in terms of estimation accuracy of parameter and change-points. Note that the current method has distribution assumptions for Gaussian and Bernoulli or Poisson. If the process is misspecified, the accuracy of the proposed method may be affected. It will be interesting to extend the proposed method to a robust change-point detection for the mixed-type observations. The proposed method can also be extended to the analysis of high-dimensional data. It may require multiple latent processes in the *mixed SSSM* to quantify the associations in the high-dimensional data.

Regarding the number of change-points, if it is known that there is only one change-point in the data, the transition matrix P_I is restricted in (30), such that only one estimated change-point is guaranteed. However, if there are more than one change-point, the construction in (32) may not give exactly the same number of change-points. In this situation, if the true number of change-points K is pre-specified, one can perform K-means clustering on the estimated change-points. However, in practice, the value of K is usually unknown. Knowledge from domain experts can provide a good guideline for the choice of K. Alternatively, one can consider other clustering methods such as hierarchical clustering to address the problem.

About the authors

Shuyu Chu is a Ph.D. Student in the Department of Statistics at Virginia Tech. Her research interests include change-point detection and modeling complex data.

Xueying Liu is a Ph.D. Student in the Department of Statistics at Virginia Tech. Her research interests focus on uncertainty quantification and big data analytics.

Achla Marathe is a professor at the Biocomplexity Institute and at the Department of Public Health Sciences, School of Medicine, at the University of Virginia. Her research interest include health economics, social epidemiology, data driven modeling of socially coupled systems, and energy markets.

Xinwei Deng is a Professor of Statistics and Data Science Faculty Fellow at Virginia Tech. His research interest focus on modeling and analysis of data with complex structures, machine learning, design of experiments, and uncertainty quantification.

Acknowledgments

Funding

ORCID

Xinwei Deng 🔟 http://orcid.org/0000-0002-1560-2405

References

Andrieu, C., A. Doucet, and R. Holenstein. 2010. Particle markov chain Monte Carlo methods. *Journal of the Royal Statistical Society Series B: Statistical Methodology* 72 (3): 269–342. doi:10.1111/j.1467-9868.2009.00736.x.

Cadena, J., G. Korkmaz, C. J. Kuhlman, A. Marathe, N. Ramakrishnan, and A. Vullikanti. 2015. Forecasting social unrest using activity cascades. *PloS One* 10 (6):e0128879. doi:10.1371/journal.pone.0128879.

Cappé, O., E. Moulines, and T. Rydén. 2009. Inference in hidden markov models. In Proceedings of EUSFLAT Conference, pp. 14–16.

Chen, Z. 2013. An overview of bayesian methods for neural spike train analysis. *Computational Intelligence and Neuroscience* 2013:1–17. doi:10.1155/2013/251905.

Chen, Z., and E. N. Brown. 2013. State space model. *Scholarpedia* 8 (3):30868. doi:10.4249/scholarpedia.30868.

Chen, J., and A. K. Gupta. 1997. Testing and locating variance changepoints with application to stock prices.

Journal of the American Statistical Association 92 (438): 739–47. doi:10.1080/01621459.1997.10474026.

Chen, J., and A. K. Gupta. 2011. *Parametric statistical change point analysis: With applications to genetics, medicine, and finance.* New York, NY: Springer.

Chen, X., X. Kang, R. Jin, and X. Deng. 2023. Bayesian Sparse regression for mixed multi-responses with application to runtime metrics prediction in fog manufacturing. *Technometrics* 65 (2):206–19. doi:10.1080/00401706.2022.2134928.

Chen, R., and J. S. Liu. 2000. Mixture Kalman filters. *Journal of the Royal Statistical Society Series B: Statistical Methodology* 62 (3):493–508. doi:10.1111/1467-9868.00246.

de Leon, A. R., and B. Wu. 2011. Copula-based regression models for a bivariate mixed discrete and continuous outcome. *Statistics in Medicine* 30 (2):175–85. doi:10.1002/sim.4087.

Deng, X., and R. Jin. 2015. QQ Models: Joint modeling for quantitative and qualitative quality responses in manufacturing systems. *Technometrics* 57 (3):320–31. doi:10.1080/00401706.2015.1029079.

Doucet, A., S. Godsill, and C. Andrieu. 2000. On sequential Monte Carlo sampling methods for Bayesian filtering. *Statistics and Computing* 10 (3):197–208. doi:10.1023/A:1008935410038.

Doucet, A., N. J. Gordon, and V. Kroshnamurthy. 2001. Particle filters for state estimation of jump Markov linear systems. *IEEE Transactions on Signal Processing* 49 (3):613–24. doi:10.1109/78.905890.

Eckley, I., P. Fearnhead, and R. Killick. 2011. Analysis of changepoint models. In *Bayesian Time Series Models*, ed. D. Barber, A. Cemgil, and S. Chiappa, 205–24. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511984679.011.

Erdman, C., and J. W. Emerson. 2008. A fast Bayesian change point analysis for the segmentation of microarray data. *Bioinformatics (Oxford, England)* 24 (19):2143–8. doi:10.1093/bioinformatics/btn404.

Fearnhead, P., and P. Clifford. 2003. On-line inference for hidden Markov models via particle filters. *Journal of the Royal Statistical Society Series B: Statistical Methodology* 65 (4):887–99. doi:10.1111/1467-9868.00421.

Fearnhead, P. 1998. Sequential Monte Carlo methods in filter theory. PhD thesis., University of Oxford.

Frühwirth-Schnatter, S. 2006. *Finite mixture and Markov switching models.* New York, NY: Springer.

Gao, Z., Z. Shang, P. Du, and J. L. Robertson. 2019. Variance change point detection under a smoothly-changing mean trend with application to liver procurement. *Journal of the American Statistical Association* 114 (526):773–81. doi:10.1080/01621459.2018.1442341.

Ge, X., and P. Smyth. 2000. Segmental semi-Markov models for change-point detection with applications to semiconductor manufacturing. Tech. rep., University of California at Irvine, March.

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin. 1995. *Bayesian data analysis.* Boca Raton, FL: Chapman and Hall/CRC.

Gupta, A., and J. Chen. 1996. Detecting changes of mean in multidimensional normal sequences with applications to literature and geology. *Computational Statistics* 11:211–21.

Hinkley, D. V., and E. A. Hinkley. 1970. Inference about the change-point in a sequence of binomial variables. *Biometrika* 57 (3):477–88. doi:10.1093/biomet/57.3.477.

Kalman, R. E. 1960. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering* 82 (1):35–45. doi:10.1115/1.3662552.

Kang, L., X. Kang, X. Deng, and R. Jin. 2018. A Bayesian hierarchical model for quantitative and qualitative responses. *Journal of Quality Technology* 50 (3):290–308. doi:10.1080/00224065.2018.1489042.

Kang, X., S. Ranganathan, L. Kang, J. Gohlke, and X. Deng. 2021. Bayesian auxiliary variable model for birth records data with qualitative and quantitative responses. *Journal of Statistical Computation and Simulation* 91 (16):3283–303. doi:10.1080/00949655.2021.1926459.

Li, J., F. Tsung, and C. Zou. 2013. Directional change-point detection for process control with multivariate categorical data. *Naval Research Logistics (NRL)* 60 (2):160–73. doi:10.1002/nav.21525.

Lio, P., and M. Vannucci. 2000. Wavelet change-point prediction of transmembrane proteins. *Bioinformatics (Oxford, England)* 16 (4):376–82. doi:10.1093/bioinformatics/16.4.376.

Müller, H.-G., and D. Siegmund. 1994. *Change-point problems.* Hayward, CA: Institute of Mathematical Statistics.

Ning, X., and F. Tsung. 2012. A density-based statistical process control scheme for high-dimensional and mixed-type observations. *IIE Transactions* 44 (4):301–11. doi:10.1080/0740817X.2011.587863.

O'brien, S. P. 2010. Crisis early warning and decision support: Contemporary approaches and thoughts on future research. *International Studies Review* 12 (1):87–104. doi:10.1111/j.1468-2486.2009.00914.x.

Page, E. 1954. Continuous inspection schemes. *Biometrika* 41 (1–2):100–15. doi:10.1093/biomet/41.1-2.100.

Qiu, P. 2008. Distribution-free multivariate process control based on log-linear modeling. *IIE Transactions* 40 (7):664–77. doi:10.1080/07408170701744843.

Ramakrishnan, N., P. Butler, S. Muthiah, N. Self, R. Khandpur, P. Saraf, W. Wang, J. Cadena, A. Vullikanti, and G. Korkmaz. 2014. 'Beating the news' with EMBERS: Forecasting civil unrest using open source indicators. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM*, 1799–808.

Ramakrishnan, N., C.-T. Lu, M. Marathe, A. Marathe, A. Vullikanti, S. Eubank, M. Roan, J. S. Brownstein, K. Summers, L. Getoor, et al. 2015. Model-based forecasting of significant societal events. *IEEE Intelligent Systems* 30 (5):86–90. doi:10.1109/MIS.2015.74.

Roberts, G. O., A. Gelman, and W. R. Gilks. 1997. Weak convergence and optimal scaling of random walk Metropolis algorithms. *The Annals of Applied Probability* 7 (1):110–20. doi:10.1214/aoap/1034625254.

Shiryaev, A. N. 1963. On optimum methods in quickest detection problems. *Theory of Probability & Its Applications* 8 (1):22–46. doi:10.1137/1108002.

Spokoiny, V. 2009. Multiscale local change point detection with applications to value-at-risk. *The Annals of Statistics* 37 (3):1405–36. doi:10.1214/08-AOS612.

Truong, C., L. Oudre, and N. Vayatis. 2020. Selective review of offline change point detection methods. *Signal Processing* 167:107299. doi:10.1016/j.sigpro.2019.107299.

Whiteley, N., C. Andrieu, and A. Doucet. 2010. Efficient Bayesian inference for switching state-space models using discrete particle Markov chain Monte Carlo methods. Tech. Rep. 10:04, Bristol Statistics Research Report.

Wikipedia. 2015. September 2012 cacerolazo in Argentina—Wikipedia, The Free Encyclopedia. Accessed December 21, 2015. https://en.wikipedia.org/w/index.php?title=September_2012_cacerolazo_in_Argentina&oldid=696215834, [Online].

Wikipedia. 2016a. 2013 protests in Brazil—Wikipedia, The Free Encyclopedia. Accessed October 16, 2016. https://en.wikipedia.org/w/index.php?title=2013_protests_in_Brazil&oldid=744677098, [Online].

Wikipedia. 2016b. 2014-16 Venezuelan protests—Wikipedia, The Free Encyclopedia. Accessed September 27, 2016. https://en.wikipedia.org/w/index.php?title=2014\%E2\%80\%9316_Venezuelan_protests&oldid=741406311. [Online].

Xie, L., S. Zou, Y. Xie, and V. V. Veeravalli. 2021. Sequential (quickest) change detection: Classical results and new directions. *IEEE Journal on Selected Areas in Information Theory* 2 (2):494–514. doi:10.1109/JSAIT.2021.3072962.

Young, Y. T., and L. Kuo. 2001. Bayesian binary segmentation procedure for a Poisson process with multiple changepoints. *Journal of Computational and Graphical Statistics* 10 (4):772–85. doi:10.1198/106186001317243449.

Zhang, N. R., D. O. Siegmund, H. Ji, and J. Z. Li. 2010. Detecting simultaneous changepoints in multiple sequences. *Biometrika* 97 (3):631–45. doi:10.1093/biomet/asq025.

Appendices A. Investigation on the choice of N_1 and N_2

It is known that the values of N_1 and N_2 control the precision of the DPF and SMC algorithms, respectively. The choice of N_1 and N_2 has been discussed in details for the DPF and SMC algorithms (Andrieu, Doucet, and Holenstein 2010; Whiteley, Andrieu, and Doucet 2010; Fearnhead 1998; Fearnhead and Clifford 2003; Chen and Liu 2000; Doucet, Godsill, and Andrieu 2000, Doucet, Gordon, and Kroshnamurthy 2001). Based on our empirical study, the proposed algorithm works efficiently with a moderate value of N_1 or N_2. The decision of using values ($N_1 = 128, N_2 = 200$) follows the suggestion from those papers.

We have conducted additional simulation for $N_1 = 512, N_2 = 512$ in mixed Gaussian-Bernoulli (S1) scenario. Table A1 and Figure A1 report the comparison results. The results suggest that increasing the values of N_1 and N_2 can lead to slightly more accurate estimations. Note that this improvement comes at the cost of increased computational time.

Table A1. Simulation results for one change-point scenario of different N_1 and N_2 values in (S1): the median of $\hat{\tau}$ over 50 iterations.

| τ	21		51		91	
$[N_1, N_2]$	[128, 200]	[512, 512]	[128, 200]	[512, 512]	[128, 200]	[512, 512]
$\hat{\tau}$	22	22	53	52	93	92
MAD ($\hat{\tau}$)	1.5	1.5	1.5	1.5	1.5	1
Time elapsed (s)	137.55	180.50	153.15	194.36	157.95	197.85
Acceptance rate of MCMC	0.24	0.25	0.23	0.22	0.20	0.19
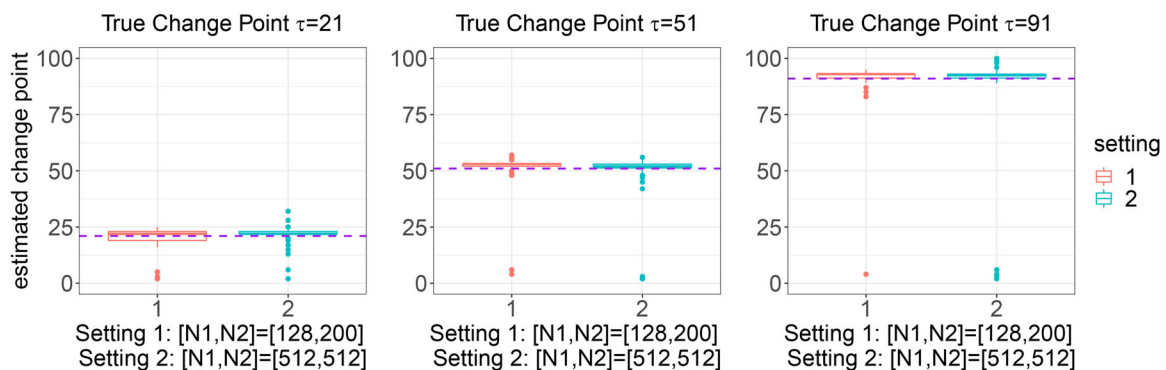
Figure A1. Boxplots of estimated change point in one change-point scenario for different N_1 and N_2 values in (S1). Purple lines are the true change point $\tau = 21, 51, 91$.
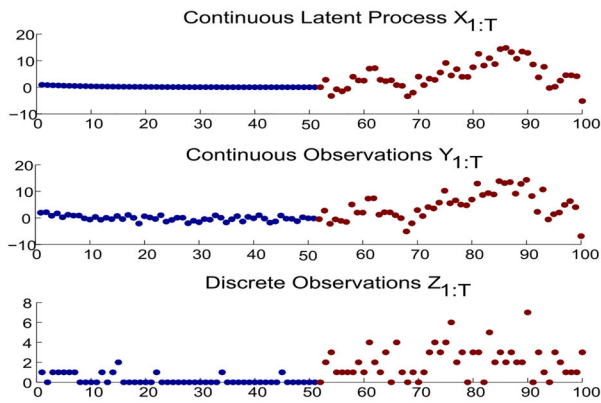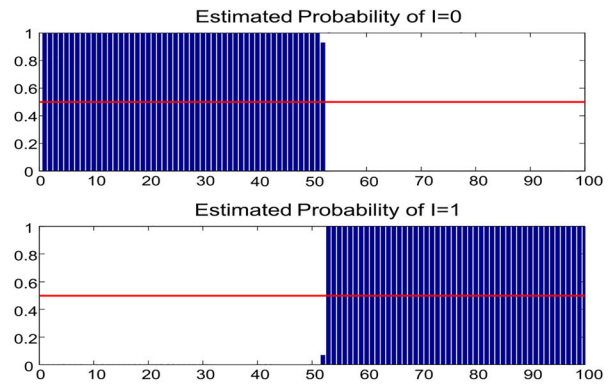
B. Additional simulation results
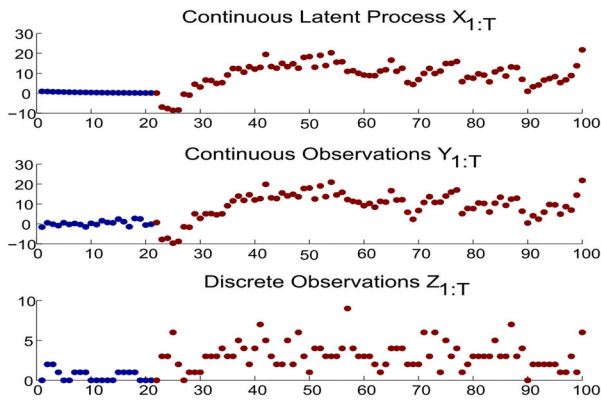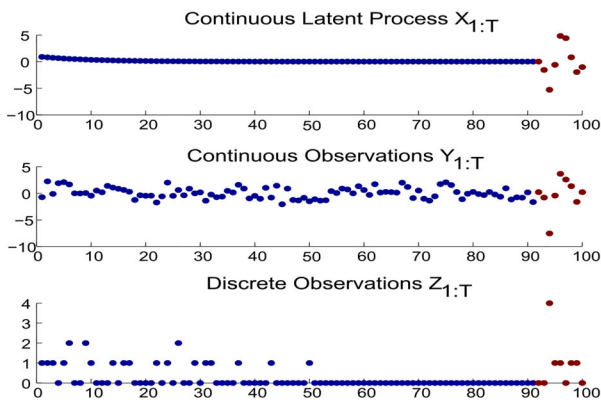

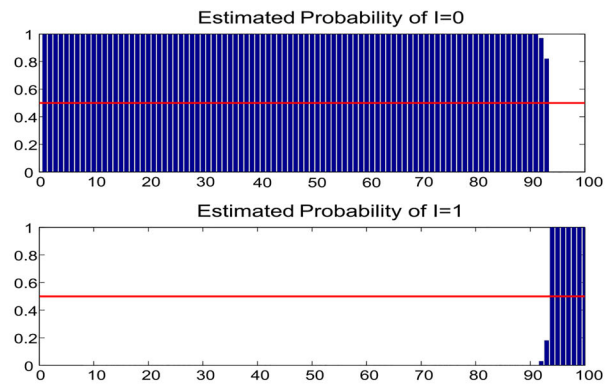(a) Simulation data with $\tau = 21$
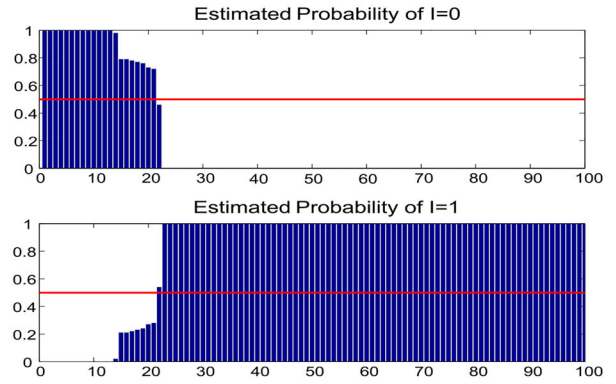
(b) Change-point estimation

(c) Simulation data with $\tau = 51$

(d) Change-point estimation

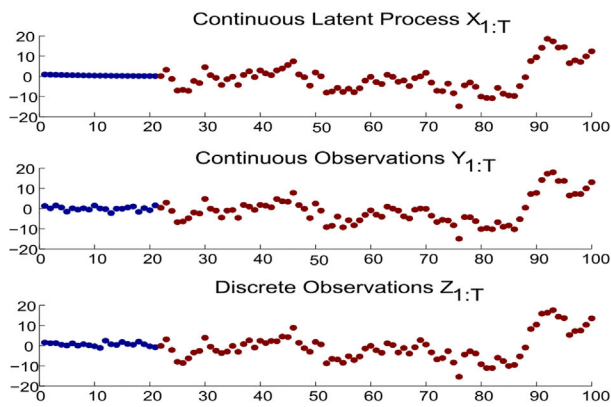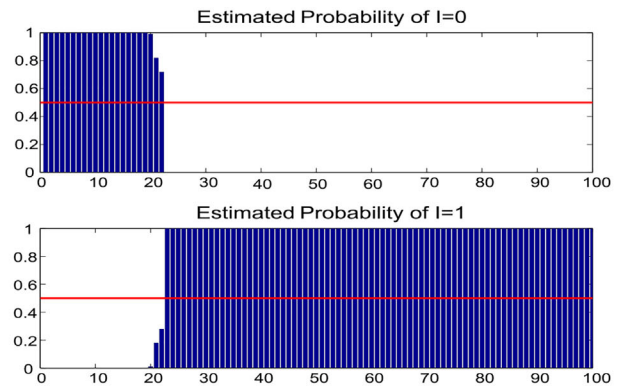(e) Simulation data with $\tau = 91$
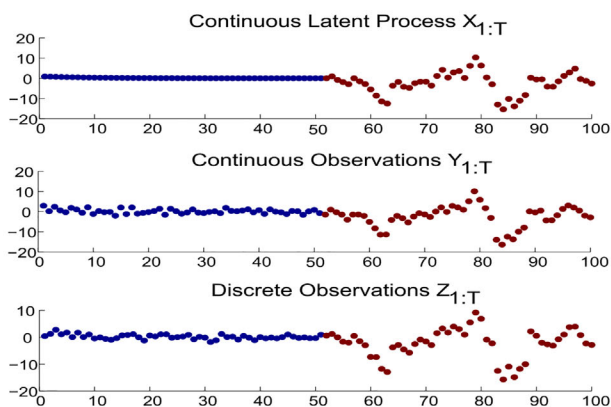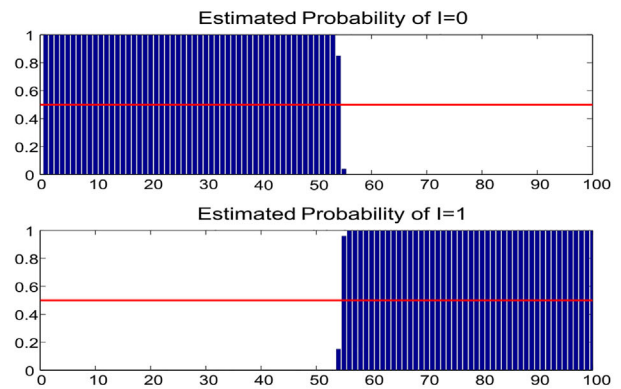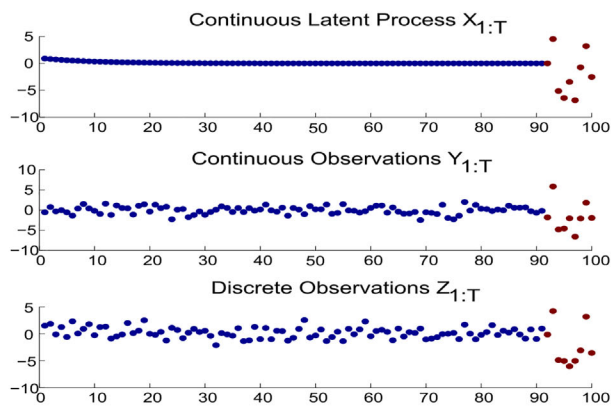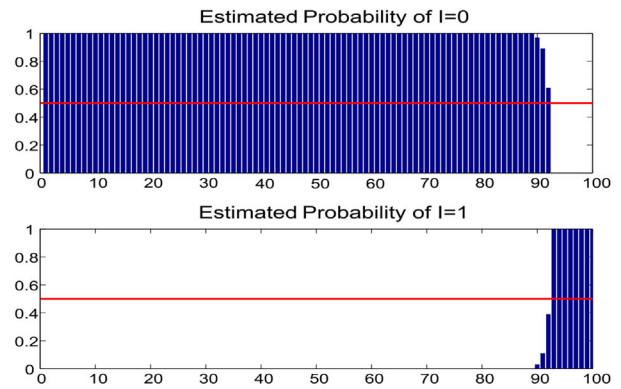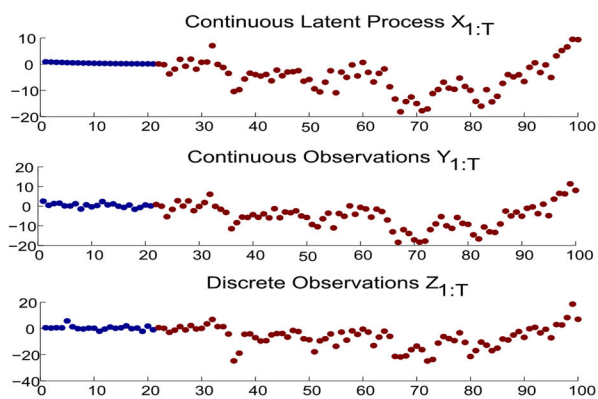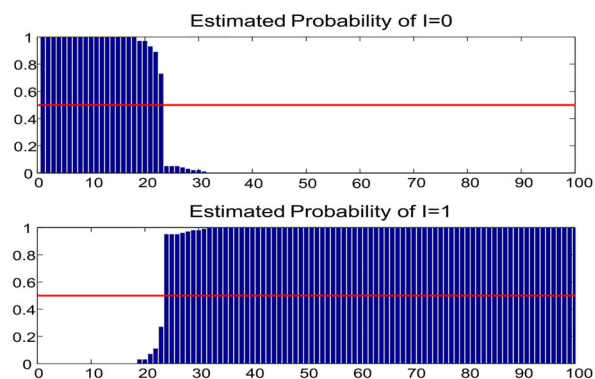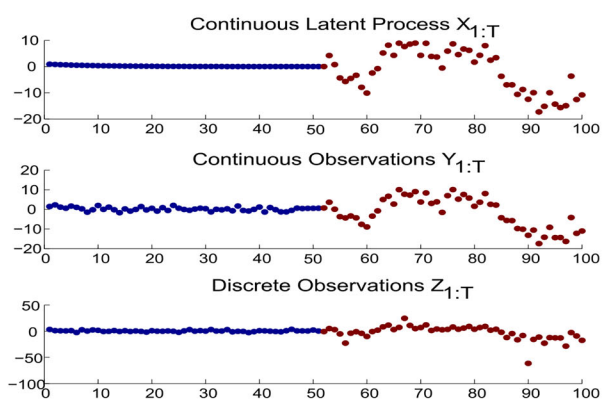
(f) Chang-point estimation

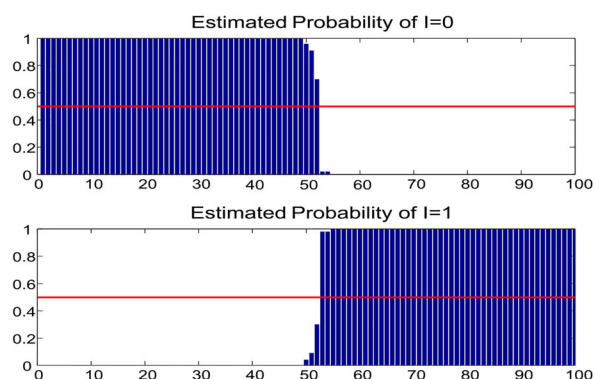**Figure B1.** One change-point detection for Gaussian-Poisson (S2).
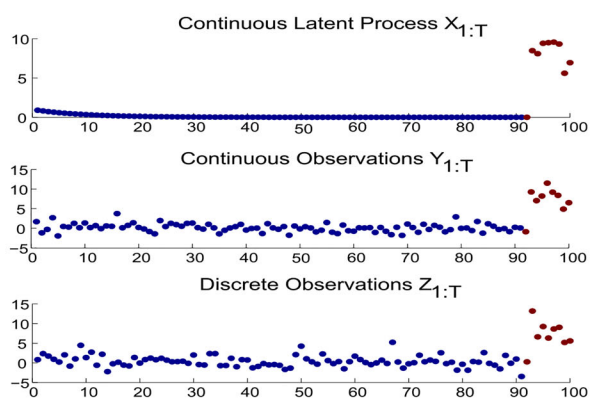
(a) Simulation data with $\tau = 21$ (b) Change-point estimation

(c) Simulation data with $\tau = 51$ (d) Change-point estimation

(e) Simulation data with $\tau = 91$ (f) Change-point estimation

Figure B2. One change-point detection for Gaussian-Gaussian (S3).

(a) Simulation data with $\tau = 21$
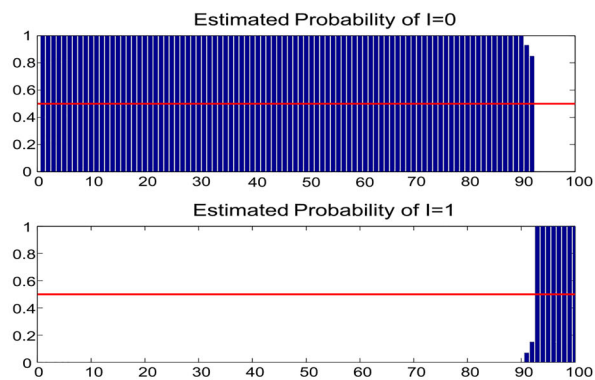
(b) Change point estimation

(c) Simulation data with $\tau = 51$

(d) Change point estimation

(e) Simulation data with $\tau = 91$

(f) Change point estimation

Figure B3. One change-point detection for Gaussian-Noncentral t (S4).