Journal of Manufacturing Systems 48 (2018) 157–165

Contents lists available at ScienceDirect

Journal of Manufacturing Systems

journal homepage: www.elsevier.com/locate/jmansys

A novel critical point detection method for mechanical deformation in tightening processes

Juan Du[a], Xi Zhang[a,*], Xinyi Xu[b], Jianjun Shi[c]

[a] Department of Industrial Engineering and Management, Peking University, Beijing 100871, China
[b] Department of statistics, The Ohio State University, Columbus, OH 43210-1247, USA
[c] H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA

ARTICLE INFO

Keywords:
Critical point detection
Tightening processes
Mechanical deformation
State space model
Particle filter
Information

ABSTRACT

The mechanical deformation of workpieces due to tightening is a common phenomenon in most assembly processes. Such deformation is typically characterized on the basis of a few critical points from sensing signals during process monitoring. Our previous study focus on improving critical point detection accuracy by establishing a state space model and a two-stage particle filter algorithm. The state variables are estimated in the first stage and the critical point is estimated in the second stage. These two stages are recursively estimated until the estimation of critical point converges. However, such method usually requires a large amount of computational efforts which may not be affordable in practice. To effectively identify critical points as well as meet the timeliness of detection, we improve the estimation algorithm by leveraging the quantification of state changes and estimating the critical point in the first stage. In this way, the critical point can be identified within one stage, thereby significantly reducing computation costs. The results from a real case study indicate that our proposed method delivers efficient critical point detection performance for process monitoring.

1. Introduction

In many tightening processes, process failure resulting from excessive mechanical deformation is one of the abnormal process conditions. This type of deformation usually occurs when a local loading force exceeds the specified limits, and the contact surface of workpieces is unable to achieve the desired engineering precision. Examples can be found in many assembly processes, such as oil pipe connection and engine piston ring fitting [1], which require a high precision in assembly to guarantee the product quality. In practice, such product quality is typically examined by testing major functions offline. For example, the leak-proofness of connected oil pipes is determined by performing a hydraulic test after the pipe connection, and unqualified products are then reworked or discarded depending on leakage severity. This offline testing procedure is usually conducted in the final stage in workshop. Consequently, the nonconforming pipes occurred at certain intermediate stages need to follow the operational path and pass through many other stages before the final testing, resulting in many nonvalue-added activities involved in manufacturing systems.

The advancement of sensor technologies and computational capabilities provides an unprecedented opportunity to acquire valuable information for product quality assurance in a variety of manufacturing processes. Automatic process monitoring and fault detection by using sensing signals has significantly improved production efficiency and reduced the nonconforming rate of products, such as in forging, stamping, spot welding, and engine assembly process [2]. In these sensing signals, the profiles caused by local mechanical deformation can be regarded as a typical class of nonlinear profiles. Specifically, the occurrence of deformation can be identified by detecting associated critical points within these profiles. These critical points are generally interpretable from an engineering perspective, and in this paper they are referred in particular to the points indicating the occurrence of mechanical deformation. For example, in the oil pipe connection process shown in Fig. 1, an observable abrupt shift in the signal is mainly caused by a local shoulder deformation in the sealing procedure that generates a large elastic stress. This critical point named “shoulder point” represents the occurrence of this deformation in the torque signal. In practice, the shoulder point is one of the major quality characteristics of connected pipes. A pipe connection with a torque value within the specified threshold at the shoulder point, is treated as an eligible connection. Incorrect detection of shoulder points may lead to a failure in identifying unqualified connections. Consequently, nonconforming pipes will be released to the oil field, and the risk of leakage during oil well drilling, pumping, and transferring tends to increase. It

* Corresponding author.
E-mail address: xi.zhang@pku.edu.cn (X. Zhang).

https://doi.org/10.1016/j.jmsy.2018.07.007
Received 25 April 2018; Received in revised form 18 July 2018; Accepted 24 July 2018

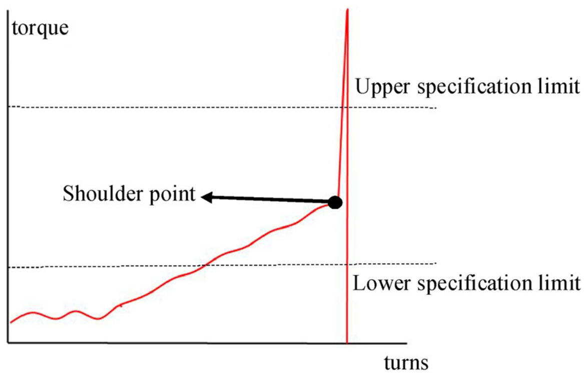

Fig. 1. The nominal torque curve over screwing-on turns.

is reported that the annual cost from the failure of oil pipes reaches half a billion dollars, and two-thirds of these failures are caused by pipe connections [3]. Thus, accurately detecting critical points from sensing signals to monitor the assembly process is of importance in product quality assurance.

There are three main challenges for automatic detection of critical points in tightening processes: (1) multiple types of nonlinear profiles simultaneously appear in a given signal, which significantly increase the number of potential critical points and mask the true critical point. Fig. 2 provides four examples of torque signals in pipe connection processes. Various nonlinear profiles that are generated from the connection process exist in the torque signals in the same connection machine, and the points along the curves show the potential change points in the torque signals from the engineering practice; (2) lateral oscillations caused by mechanical return differences frequently appear in original signals (Fig. 2(d)), and such signals could not be directly regarded as functional data. This characteristic may make some methods, such as conventional time series analysis, fail to detect critical points; (3) original signals usually exhibit different lengths as a result of inconsistent upstream process settings, leading to the mutable locations of critical points. Fig. 2 shows four example signals with different lengths. This variation poses another challenge in critical point detection because the locations of critical points may vary significantly.

Extensive research has been conducted on critical point detection methods. These methods are generally categorized into two classes: engineering-based methods and data-driven methods.

1.1. Engineering-based methods

In the use of engineering-based methods, the basic physical knowledge of a given process is usually studied through engineering and experimental analysis before model development. For instance, the tonnage signals of the stamping process can be divided into multiple segments, which can be physically interpreted by studying the stamping mechanism; the process faults from different operations can then be directly identified by analyzing the profiles or critical points within the

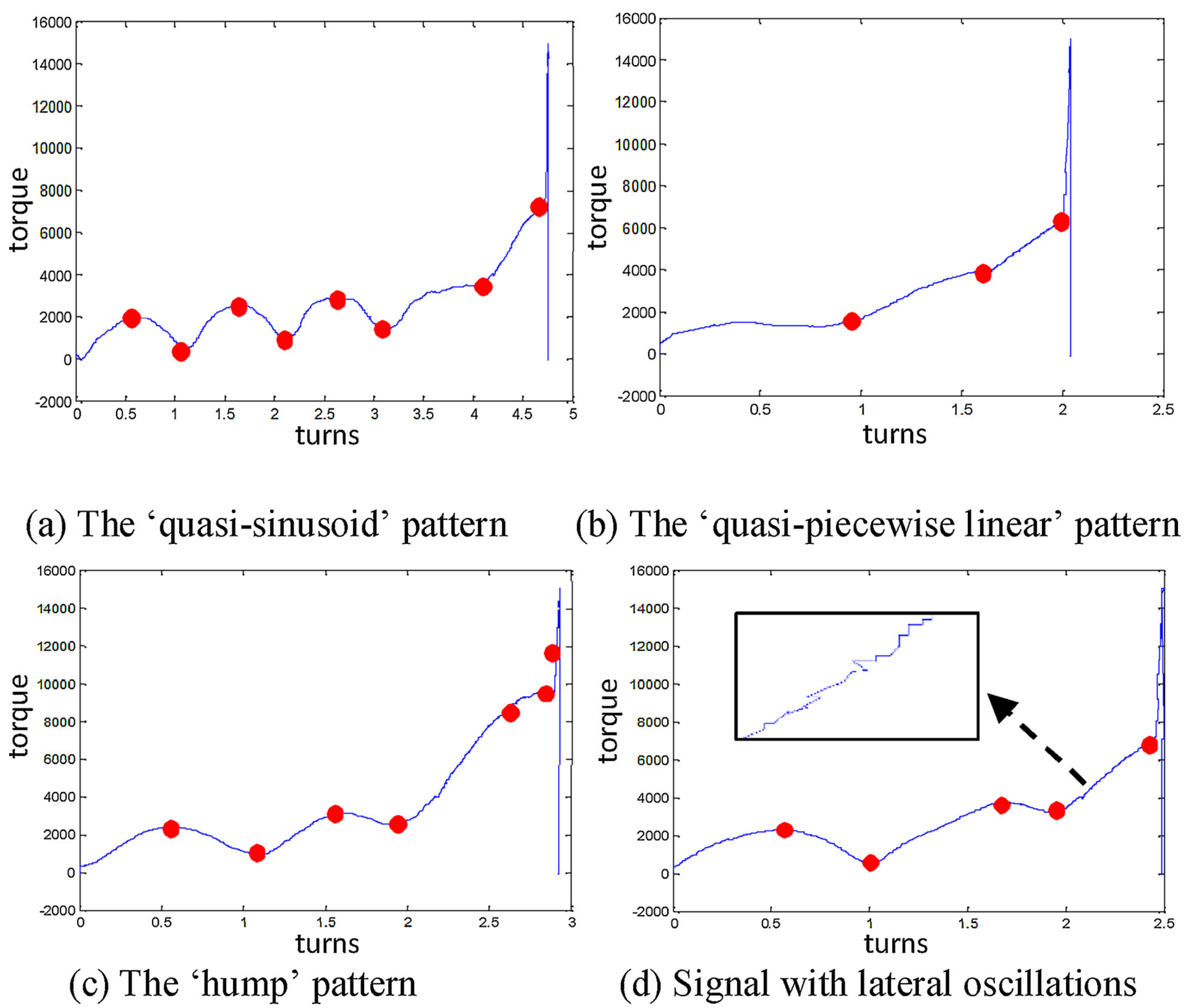

Fig. 2. Examples of typical profiles in torque signal, and the critical points easily submerge under different nonlinear patterns of torque signals.

associated segments [4]. Another example is that the critical point of oil pipe connection in torque signals could be identified by using empirical methods of engineering practice [34,35]. Besides this, the end point of wafer polishing in optical signals could be identified by matching a given signal to a reference image obtained from physical experiments [5]. A few other works that utilized similar methods could be found in the areas of hot forming, chemical vapor decomposition, and coating and granulation [6–8]. These methods generally provide a basic overview and explanations on changes in fault and process conditions under certain situations from a physical perspective. However, one common limitation is that these methods may fail to achieve repeatable successes under situations that could not be explained directly with associated engineering experience.

1.2. Data-driven methods

In data-driven methods, statistical models for critical point detection are proposed by incorporating sensing signals as the functional process variables. These methods consist of two streams: parametric methods and nonparametric/semiparametric methods.

Parametric methods are mainly used in studies on the distribution of signals and associated parameters to characterize the condition of the manufacturing process. Canonical critical point detection has been carried out using parametric methods [9–11], in which the hypothesis test of time series data is initially performed for change point detection. For example, a sequential change point detection method for the normal distribution uses the generalized likelihood ratio statistics [12]. In addition, a two-phase regression model is adopted in undocumented change point detection with the use of F_{max} test [13]. These methods are acceptable when the signals completely comply with the model's constraints; however the model assumption may not be fully satisfied in practice. Another track of change point detection method involves the use of statistical control charts. For example, exponential weighted moving average charts [14] and Gaussian process model-based T^2 control charts were used to detect the occurrence of a shift in time series data [15]. Moreover, analysis based on segmented signals such as piecewise linear approach [13] and piecewise constant approach [16] were also developed for change point detection. These methods are mainly employed to detect changes in signal distribution from collected data. However, the signals in complex manufacturing processes usually contain various nonlinear profiles caused by different types of mechanical deformation. Hence, directly using specific distributions is unsuitable to characterize these scenarios.

To overcome the limitations of parametric methods, nonparametric and semi-parametric models for critical point detection have been proposed. Mann-Whitney test [17], relative density ratio test [18], graph-based method [19], and kernel change detection method [20] were proposed to detect the abrupt distribution change in a data sequence. These methods are model-free and powerful but may fail to detect critical points of the deformation signals with existence of mixed types' nonlinear profiles. Other alternative nonparametric methods for critical point detection include spectral analysis [21–24]; nevertheless, spectral analysis partially loses its interpretation if the original sequence data is not in the time domain. A general critical point detection method, such as the Kullback–Leibler information discrimination, was proposed under the local stationarity assumption [25]. This method occasionally fails because the local stationarity of signals could not be always satisfied in many industrial practices. To deal with this problem, a state space model has been introduced with appropriate state estimation algorithms to capture the critical information from the signals. Relevant works include the end point detection of chemical mechanical polishing processes [26] and ultrasonic-cavitation based nanoparticle dispersion processes [27]. These data-driven methods provide a feasible way to search over the critical points while the signal contains a few of nonlinear patterns and traditional types of noises. However, in most of mechanical assembling processes, such as tightening processes, featured points due to various types of mechanical deformation may exist widely within the signal and vary a lot, and these methods may fail to identify the critical ones among these points. Meanwhile, critical point detection methods without consideration of the physical mechanism usually lead to incorrect decisions when signals contain condition-irrelevant nonlinear patterns caused by some latent process factors. Therefore, it is desirable to develop a generic critical-point detection method appropriately incorporating mechanical engineering knowledge.

In our previous study [36], we proposed a critical point detection method for the premium pipe connection by considering the physical interpretations of the connection mechanism and developed a two-stage recursive particle filter algorithm for critical point identification. The first stage of this particle filter algorithm estimates the states and the inference on the location of critical point is made in the second stage. To assure the accuracy of critical point detection, these two stages are recursively conducted until this algorithm converges under the designed criteria. This method requires a large amount of computations and may not meet the requirements of production cycle time in tightening processes. To address this issue, we propose a new particle filter algorithm by quantifying state changes, which is directly used for inference on the location of critical points in the first stage. In this way, the second stage and iterations of previous two-stage particle filter algorithm are non-necessary, which significantly reduces the computational time to meet online monitoring requirement. The framework of the proposed methodology is shown in Fig. 3.

The rest of this paper is organized as follows. Section 2 introduces the model formulation for critical point detection via the two-phase state space model. Section 3 presents the proposed new particle filter algorithm for critical point detection. In Section 4, a case study on an oil pipe connection process is provided to validate our proposed approach. The discussions are provided in Section 5. Section 6 concludes the paper.

2. A two-phase state space model for critical point detection

In this section, the critical point detection problem is formulated as a two-phase state space model with a critical point parameter. To clarify the problem, we use an oil pipe connection process as a demonstration. As explained in Fig. 1, the torque curve in the connection process is regarded as two linear parts separated by a critical point according to the relationship between the stress and Young's modulus in material mechanics [28]. Thus, in light of physical interpretation of

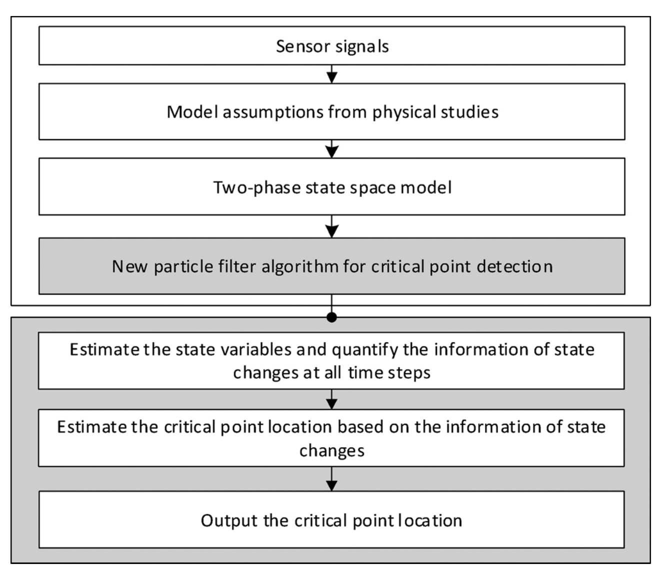

Fig. 3. The framework of the proposed methodology.

mechanical deformation, we set the following assumptions for our proposed approach:

(1) A single critical point existing among multiple potential points can capture occurrence of deformation in tightening processes.
(2) The nominal curve behind the original signal in each segment is linearly presented in terms of the principles in material mechanics. Specifically, in oil pipe connection, the first segment is referred to as thread engagement procedure and the second segment is referred to as shoulder contact procedure.

Notably, the nominal curve in assumption (2) denotes the curve theoretically derived from the physical analysis of mechanical deformation [28], which is piecewise linear. Although the obtained torque signals in Fig. 2 present various nonlinear patterns, the basic physical mechanism beneath each signal is identical to the nominal curve. Although other nonlinear models, such as polynomial models, can be used to formulate this problem instead of a two-phase piecewise linear model, these models cannot satisfactorily address the critical point detection problem with mechanical deformation profiles. The reasons are explained as follows: (1) The physics behind mechanical deformation profiles can be theoretically represented by a piecewise linear model with a parameter regarding the critical point location. Thus, nonlinear models without physical interpretation, such as polynomial models, could not successfully address the critical point due to mechanical deformation during the tightening process. (2) Nonlinear models, such as quadratic or basis-function models, might detect a series of points which are mixed by some local pseudo change points caused by system noises or numerous latent process factors, leading to confusion in explanation of these change points. Therefore, a two-phase piecewise linear state space model is developed to concisely address the critical point detection problem in this study.

Specifically, the two-phase state space model is proposed as follows:

$$y_k = \begin{cases} a_k + b_k t_k + \varepsilon_k & t_k < c \\ a_0 + b_0 c + b_k (t_k - c) + \varepsilon_k & t_k \geq c \end{cases} \tag{1}$$

$$(a_k,\ b_k) = \begin{cases} (a_{k-1},\ b_{k-1})\ \textit{with probability}\ 1-p \\ (a_0 + b_0 c,\ b_{k-1} + \delta)\ \textit{with probability}\ p \end{cases} \tag{2}$$

where y_k, t_k and ε_k represent the torque observation, screwing turn observation, and Gaussian noise at time k, respectively. For this tightening process, the variance of the noise appears stable according to the continuous historical data observations. Thus, in this research, we assumed the variance of the noise term as a constant. Parameter c denotes the location of the critical point in the x -axis, which represents the screwing turns while process condition changes. a_k and b_k are the intercept and slope of the current linear segment at time k. Specifically, a_0 and b_0 are the intercept and slope before the critical point. p is the transition probability between current state and new state. δ is the increment of the slope b_k at the critical point. For torque signals in oil pipe connection processes, observation y_k dramatically increases in the second segment once the pipe edge touches the shoulder of the casing. Accordingly, δ should be a large positive number. In general, (1) can be regarded as the measurement model, and (2) can be treated as the prediction model which indicates the transitions between state variables of previous and current time. Since the measurement model has a piecewise linear structure, some traditional tools, such as Kalman filter, are not effective to be used in its parameter estimation. A particle filter, as an alternative method, is suitable to deal with this problem [29]. However, as there is an unknown parameter c of the critical point that does exist in the measurement model, the direct estimation of the distributions of variables using a traditional particle filter algorithm is not applicable. In the next section, a new particle filter algorithm is proposed to tackle this problem.

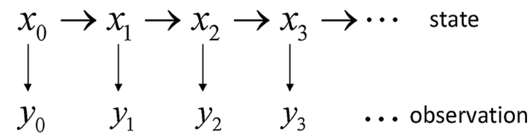

Fig. 4. The general view of a state space model.

3. Particle filter algorithm for critical point detection

The proposed particle filter is built upon the basic ideas of a particle filter. Therefore, we will briefly review the basic particle filter algorithm in Subsection 3.1, and then introduce the novel proposed particle filter in Subsection 3.2.

3.1. The basic concepts of the particle filter algorithm

In a general state space model, the state x_i transits over time, and the observation y under the state x_i can be obtained as shown in Fig. 4. Generally, a two-step procedure is established for the state estimation in a state space model: update procedure and prediction procedure.

In this subsection, we assume that the distribution of c is known as *a prior*, and the state $\boldsymbol{x}_k$ only contains two parameters a_k and b_k, i.e. $\boldsymbol{x}_k = (a_k,\ b_k)$. Hence, state estimation can be conducted by:

$$p(\boldsymbol{x}_k | y_{1:k}) = \frac{p(y_k | \boldsymbol{x}_k)}{p(y_k | y_{1:k-1})} p(\boldsymbol{x}_k | y_{1:k-1}) \tag{3}$$

where $p(\boldsymbol{x}_k|y_{1:k})$ is the state posterior probability given a sequence of observations y_1 to y_k, and $p(y_k|\boldsymbol{x}_k)$ is the probability of the current observation y_k given the state $\boldsymbol{x}_k$, which can be estimated from (1). $p(y_k|y_{1:k-1})$ is the marginal distribution, which can be obtained by calculating $\sum_{\boldsymbol{x}_k} p(y_k|\boldsymbol{x}_k)p(\boldsymbol{x}_k|y_{1:k-1})$. In the implementation, calculating $p(y_k|y_{1:k-1})$ can be avoided by converting $p(y_k|\boldsymbol{x}_k)p(\boldsymbol{x}_k|y_{1:k-1})$ to a density function. In the state prediction procedure, the general form of the state prediction is:

$$p(\boldsymbol{x}_{k+1}|y_{1:k}) = \int p(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k)p(\boldsymbol{x}_k|y_{1:k})d\boldsymbol{x}_k \tag{4}$$

where $p(\boldsymbol{x}_{k+1}|y_{1:k})$ is the prior distribution of state variable at time $k+1$ given the observations y_1 to y_k, and $p(\boldsymbol{x}_{k+1}|\boldsymbol{x}_k)$ is the state transition probability, which can be obtained from the prediction model (see (2)). Consequently, (3) provides the state posterior distribution for (4), and (4) provides the state prior distribution at time $k+1$ for (3). It is quite challenging to obtain the analytical solution of the current state from (3) and (4) if the measurement model is piecewise linear. In such a case, the particle filter, which generates the particles from the prior distribution and estimates the posterior distribution via sampling techniques, is a powerful tool to use. In particular, particle filters employ a large number of samples as particles to approximate the probability distribution function $f(\boldsymbol{x})$ of the state posterior probability $p(\boldsymbol{x}_{0:k}|y_{1:k})$. Denote χ_k as $\chi_k := \boldsymbol{x}_k^{[1]}, \boldsymbol{x}_k^{[2]}, \ldots, \boldsymbol{x}_k^{[m]}, \ldots, \boldsymbol{x}_k^{[M]}$, i.e., χ_k is a set of particles of state $\boldsymbol{x}_k$. Each particle $\boldsymbol{x}_k^{[m]}$ is a realization of the state estimation at time k, and M is the total number of particles in the particle set χ_k. Ideally, the probability of a state $\boldsymbol{x}_k$ shown in the particle set χ_k is proportional to $p(\boldsymbol{x}_k|y_{1:k})$ when M becomes infinite. In practice, a large value of M is often used to guarantee an accurate approximation. It should be noticed that sampling particles from $f(\boldsymbol{x})$ directly may not be applicable because $f(\boldsymbol{x})$ can be from any arbitrary complex distribution, such as non-Gaussian, or multimodal distributions. Thus, an importance sampling technique, which samples particles from a known probability distribution function $g(\boldsymbol{x})$ to approximate the target probability distribution function $(\boldsymbol{x})$, is used to tackle this issue. Here we denote $I(x_k^{[m]} \in A)$ as the indicator function, and then the integral of $g(\boldsymbol{x})$ under a Borel set A can be approximated by

$$\frac{1}{M}\sum_{m=1}^{M} I(\boldsymbol{x}_k^{[m]} \in A) \xrightarrow{M\to+\infty} \int_A g(\boldsymbol{x})dx \tag{5}$$

where $\boldsymbol{x}_k^{[m]}$ is the m^{th} sample from $g(\boldsymbol{x})$. Define the importance sampling weight as:

$$w_k^{[m]} = \frac{f(\boldsymbol{x}_k^{[m]})}{g(\boldsymbol{x}_k^{[m]})}. \tag{6}$$

As a result, the integral of f under A can be approximated by

$$\left[\sum_{m=1}^{M} w_k^{[m]}\right]^{-1} \sum_{m=1}^{M} I(\boldsymbol{x}_k^{[m]} \in A) w_k^{[m]} \xrightarrow{M\to+\infty} \int_A f(\boldsymbol{x}) d\boldsymbol{x} \tag{7}$$

A notable property of importance sampling is that it allows the probability distribution function g to converge to f as M becomes infinity. In particle filter, the particles in χ_{k-1} are sampled from $p(\boldsymbol{x}_{k-1}|y_{1:k-1})$, and we select distribution function g as

$$p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1})\, p(\boldsymbol{x}_{0:k-1}y_{1:k-1}). \tag{8}$$

Hence, the state posterior distribution can be obtained as

$$p(\boldsymbol{x}_{0:k}|y_{1:k}) \propto p(y_k|\boldsymbol{x}_k) p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1})\, p(\boldsymbol{x}_{0:k-1}y_{1:k-1}) \tag{9}$$

where the derivation procedure of (9) is given in Appendix A. Notably, at time $k = 1$, the state prior distribution for $\boldsymbol{x}_k$ is empirical distribution from engineering knowledge, while at time $k > 1$, the state prior distribution for $\boldsymbol{x}_k$ can be formulated as:

$$p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1})\, p(\boldsymbol{x}_{0:k-1}|y_{1:k-1}) \tag{10}$$

with the importance sampling weight

$$w_k = \frac{f(\boldsymbol{x}_k)}{g(\boldsymbol{x}_k)} \propto \frac{p(y_k|\boldsymbol{x}_k) p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}) p(\boldsymbol{x}_{0:k-1}|y_{1:k-1})}{p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}) p(\boldsymbol{x}_{0:k-1}|y_{1:k-1})} \propto p(y_k|\boldsymbol{x}_k) \tag{11}$$

The basic particle filter algorithm is established based on the above formulae. However, particle filter has a common degeneracy problem [31], which appears that only a few particles have dominant weights while most particles have small weights after a few iterations. Resampling is a popular approach to tackle the degeneracy problem by eliminating the particles with small weights and focusing on the particles with significant weights [32]. Therefore, the resampling step is essential in particle filter. On the basis of the above discussions, the basic particle filter algorithm is given in Appendix B.

3.2. The proposed particle filter algorithm

Although the resampling method can deal with the particle degeneracy problem, randomness is introduced to such procedure due to the particle impoverishment and further increases the estimation variance of state parameters [30]. To deal with the particle resampling issue in the proposed two-phase state space model, we introduce an integrated sampling strategy as below. Specifically, given that transition probability p is included in our two-phase state-space model, stratified sampling is necessary to sample the state prior distribution $p(\boldsymbol{x}_k|y_{1:k-1})$ at time k from the state posterior distribution $p(\boldsymbol{x}_{k-1}|y_{1:k-1})$ with probability $1-p$ at previous time $k-1$ and the new state distribution $h(\boldsymbol{x})$ with probability p (see Eq. (2)). Afterward, a low-variance sampling technique is used to replace the traditional resampling technique in the resampling step to obtain particles with large weights. Compared with independent sampling, low-variance sampling reduces the variance and computation load in the resampling step of the particle filter. Low-variance sampling has a complexity of $O(M)$, whereas regular independent sampling has a complexity of $O(MlogM)$ [30]. In the following discussions, we will briefly introduce these two sampling techniques first, and then present the details of the proposed particle filter algorithm.

(1) Stratified sampling. Stratified sampling is an efficient sampling method to partition the population into non-overlapping groups when the groups within the overall population follow different distributions [33]. Here, by using the stratified sampling, state prior distribution $p(\boldsymbol{x}_k|y_{1:k-1})$ is sampled with a deterministic sample size m_0 from $h(\boldsymbol{x})$ at time k, and $M-m_0$ sample size from state distribution $p(\boldsymbol{x}_{k-1} = \boldsymbol{x}|y_{1:k-1})$ at time $k-1$, respectively. The state prior distribution of stratified sampling is

$$p(\boldsymbol{x}_k = \boldsymbol{x}|y_{1:k-1}) = ph(\boldsymbol{x}) + (1-p)p(\boldsymbol{x}_{k-1} = \boldsymbol{x}|y_{1:k-1}) \tag{12}$$

Denote $\boldsymbol{x}_{k,0}^{[1]}, \boldsymbol{x}_{k,0}^{[2]},\dots, \boldsymbol{x}_{k,0}^{[m_0]}$ to be randomly sampled from $h(\boldsymbol{x})$ and $\boldsymbol{x}_{k,1}^{[1]}, \boldsymbol{x}_{k,1}^{[2]},\dots, \boldsymbol{x}_{k,1}^{[M-m_0]}$ to be the random samples from $p(\boldsymbol{x}_{k-1}|y_{1:k-1})$. In practice, the number of the allocated particles is usually in proportion to the transition probability to reduce sampling variance, that is, $m_0 = pM$ [27].

(2) Low variance sampling. The key idea of low variance sampling is that the selection of samples involves a sequential stochastic process and is thus not performed independently in the resampling process. In the original particle filter and the stratified sampling-based particle filter, M random numbers are used in particle selection. On the contrary, in low variance sampling, only one random number r is uniformly selected in the interval $[0, 1/M]$ and is then added by $1/M$ repeatedly once a particle is selected. Any real number $u = r + (m-1)\frac{1}{M}$ in the interval [0, 1] exactly corresponds to one particle selected. The corresponding relationship between the sampling particle i and u is shown in the following equation:

$$i = \underset{j}{\operatorname{argmin}} \sum_{m=1}^{j} w_k^{[m]} \geq u \tag{13}$$

Two advantages emerge in the implementation of the low variance sampling strategy: (1) the resampled particles preserve higher diversity than that without low variance sampling, and (2) the complexity of the low variance sampling strategy is lower than the independent sampler during the entire resampling procedure [30]. Given that computation load is a critical challenge in particle filter, an efficient resampling strategy will significantly improve the time efficiency in real applications. The limitation of low variance sampling technique also exists when the state variables are unchanged, i.e., $\boldsymbol{x}_k = \boldsymbol{x}_{k-1}$. In such situation, one should never resample because the resampling will lead to M identical particles ($\boldsymbol{x}_k^{[m]} \sim p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}^{[m]})$) and lose the particle diversity, thereby leading to inaccurate estimations of state variables. Noteworthy, the combined stratified sampling has an advantage to bring at least m_0 new particles into the state estimation, which will increase the particle diversity and thus reduce the particle impoverishment introduced by the resampling. In this way, the limitation of the low variance sampling can be avoided.

Fig. 5 shows our integrated sampling strategy for the proposed particle filter algorithm. At time when $k > 1$, stratified sampling is used to draw particles from priors and state posterior distributions at the previous time step, and low variance sampling is performed in the

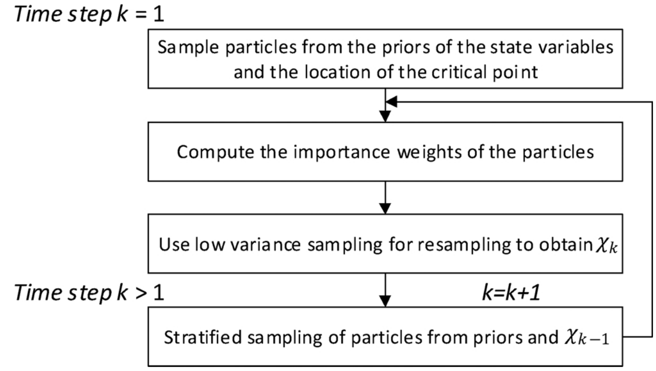

Fig. 5. The flowchart of the integrated sampling technique.

resampling step to resample the particles of large importance weights.

Through above sampling strategy, the state variables can be directly estimated. Since the critical point is located at the places where a significant change of state variables occurs, we need to know the exact change quantity from the state variables to identify the critical point. Here, we propose to introduce the information concept from Shannon [37] to quantify the change information of the state variables and incorporate such information into our proposed algorithm for critical point detection. In the following paragraphs, we will first introduce the information concept and then illustrate how to use such information for critical point identification.

Let E_k denote the event "the change of states occurs at time step k", and the information quantity of event E_k is defined as $I_n(E_k) = \log\left(\frac{1}{P_{E_k}}\right)$. Notably, such definition shows that the information quantity of a deterministic event and an impossible event is zero and infinity, respectively. Hence, in our problem context, $I_n(E_k) = \infty$ indicates the critical point does not occur at time step k while $I_n(E_k) = 0$ indicates the critical point has occurred no later than time step k. Thus, a possible way to identify the critical point is to exploit the information quantity of state variable changes. The particle distribution in the state space exhibits the change information, which is the key for change detection. We propose to group the particles according to their state values for quantification of $I_n(E_k)$. From the engineering perspective of the tightening process, the critical point occurs once the shoulder contact occurs between the pipe and the casing. Hence, the state variable b_k experiences a dramatic increase after the critical point occurs. So we expect to group the state variables based on their state values and obtain the proportion of particles which achieves dramatic increase in state values.

Let s denote the threshold to partition the particles into two groups. The first group contains the particles with b_k values smaller than s, and the other particles are separated into the second group. s can be given by the engineering specifications based on the types of connections between the pipe and the casing. After grouping, P_{E_k} and $I_n(E_k)$ can be calculated. $I_n(E_k)$ is almost infinity before the critical point and will be decreasing after the critical point occurs. When $I_n(E_k)$ becomes zero, the critical point definitely occurs before time step k. P_{E_k} can be estimated as

$$P_{E_k} = Pr(b_k \geq s \,|\, y_{1:k}) \approx \sum_{m=1}^{M} w_k^{[m]} I_{\{b_k \geq s\}}(\boldsymbol{x}_k^{[m]}) \tag{14}$$

Based on the above discussions, the specific particle filter algorithm is given as follows:

A New Particle Filter Algorithm for Critical Point Detection

At time step $k = 1$,
(1) Sample particles from priors
 For $m = 1$ to M
 Sample $\boldsymbol{x}_1^{[m]} \sim g_1(\boldsymbol{x})$, $m = 1, 2, \ldots, M$
 Sample $c^{[m]} \sim g_2(c)$, $m = 1, 2, \ldots, M$
 Compute $w_1^{[m]} = p(y_1 | \boldsymbol{x}_1^{[m]}, c^{[m]})$
 $I_n(E_k) = \infty$
(2) Low variance sampling for resampling step
 Let r be a random number that is uniformly selected in the interval $[0, 1/M]$
 $d = w_k^{[1]}$
 $i = 1$
 For $m = 1$ to M
  $u = r + (m-1)\dfrac{1}{M}$
  If $u > d$
   $i = i + 1$
   $d = d + w_k^{[i]}$

Here, $g_1(\boldsymbol{x})$ and $g_2(c)$ denote the prior distributions of state $\boldsymbol{x}_k = (a_k, b_k)$ and $c^{[m]}$ and $h(x)$ can be derived from Eq. (2). In the particle filter algorithm, the state variables are estimated in each time step. The state change information quantity $I_n(E_k)$ at any time can be obtained. Hence, the location of critical point can be obtained according to the point where continuous decrease of $I_n(E_k)$ occurs.

4. Case study

In this section, a real case study from the pipe connection process are conducted to validate the effectiveness of our proposed methods. In our implementation of the proposed method, the priors of the parameters were first specified. We assume that the state vector $\boldsymbol{x}$ follows a normal distribution $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$; parameter $\boldsymbol{\mu}$ can be estimated from the coefficients of the piecewise linear regression in the original signals. Meanwhile, we can reasonably assume that the covariance matrix is with the structure $\boldsymbol{\Sigma} = diag(\sigma_0^2, \sigma_1^2)$. We set the prior distribution of c follows a beta distribution $beta(\alpha, \beta)$. In practice, the prior distribution of c does not greatly impact the detection results. For transition probability p, the same rule in Ref. [27] is applied, and we set the transition probability $p = 0.2$ for fair comparison. The number of particles $M = 500$ and the threshold $s = 20000$ from engineering specifications are set for real case study.

Our proposed approach is validated by using the torque signals collected in a real pipe connection process. A total of 180 samples of torque and screwing turn signals from four continuous batches of connected pipes were collected. All the collected signals have different lengths, and various nonlinear profiles exist with lateral oscillations. The critical point locations were annotated by experienced technicians. A few typical examples are selected from our 180 samples to demonstrate the performance of our method (Fig. 6). The critical points that characterize the connection conditions are accurately identified from the different nonlinear profiles in the torque signals.

Given that the signal sampling rate is 20 Hz, we defined a range of

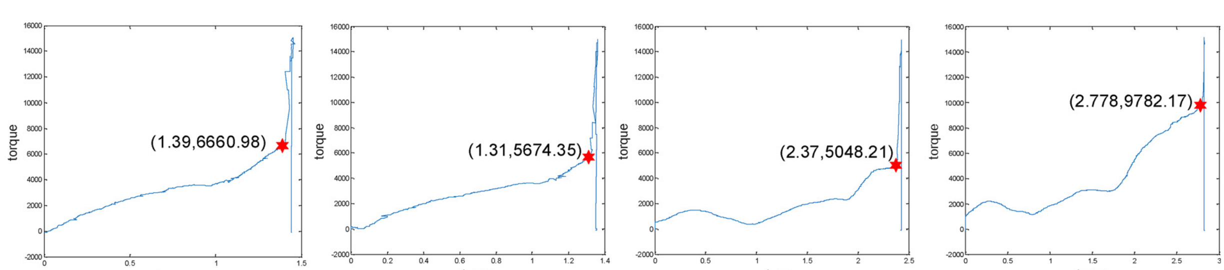

Fig. 6. Examples of critical point detection in torque signals with nonlinear profiles.

Table 1
The SDRs of our method and other methods in tightening processes.

Batch Number	Bach size	Our method	Method in [36]	Wu's Method	Cubic spline	piece linear Regression	VAM's Method	SDM	VRT
1	50	0.96	0.96	0.38	0.64	0.92	0.22	0.56	0.14
2	50	0.96	0.90	0.20	0.54	0.88	0.62	0.40	0.26
3	50	0.96	0.96	0.30	0.54	0.84	0.42	0.34	0.36
4	30	1.00	0.93	0.23	0.37	1.00	1.00	0.50	0.63
Mean	–	0.97	0.94	0.28	0.54	0.90	0.52	0.44	0.32

Table 2
The RMSEs of our method and other methods in tightening processes.

Batch Number	Bath size	Our Method	Method in [36]	Wu's Method	Cubic spline	piece linear Regression	VAM's Method	SDM	VRT
1	50	0.017	0.012	0.048	0.278	0.182	0.050	0.027	1.411
2	50	0.016	0.019	0.064	0.602	0.287	0.037	0.039	1.323
3	50	0.028	0.029	0.055	0.518	0.324	0.052	0.038	1.457
4	30	0.013	0.018	0.046	0.050	0.011	0.011	0.032	0.559
Mean	–	0.019	0.020	0.054	0.397	0.222	0.041	0.034	1.257

Table 3
Computation Time(s) of each batch of our method and other methods in tightening processes.

Batch Number	Bath size	Our Method	Method in [36]	Wu's Method	Cubic spline	piece linear Regression	VAM's Method	SDM	VRT
1	50	43.685	166.125	413.025	9.6396	19.8925	–	1.3859	1.2007
2	50	48.630	165.860	436.900	8.5529	19.6415	–	1.2209	1.1512
3	50	41.610	149.845	390.865	8.7191	18.4493	–	1.1243	1.0716
4	30	20.856	77.526	187.230	4.4507	8.9525	–	0.4443	0.5746

screwing turns $[-0.04, 0.03]$ as a criterion for successful detection; that is, the difference between detected and real turns should be within the specified range in the successful detection of the critical point. The successful detection rate (SDR) of our proposed method in this case study is provided in Table 1. To measure the deviation between the true screwing turn value and the detected screwing turn value, we define root mean square error (RMSE) of screwing turns, and the performance is shown in Table 2. We use MATLAB 2017b running on a CORE i5 Intel processor to compute the proposed method. The computation time is listed in Table 3. As shown in these tables, our proposed method achieves a satisfactory performance on both detection accuracy and detection time.

We first compared our proposed method with our previous method [36] by using the same dataset. As shown in Tables 1 and 2, the accuracy rate and the RMSE of our method is 3.2% higher and 5% smaller than the method in [36] respectively. Moreover, the proposed method is 72.3% faster than the method in [36], which meets the requirement of the online monitoring for the fast pipeline production. Hence, our new proposed methodology not only improve the computation speed but also keep the high successful detection rate.

Our method was also compared with Wu's method [27] by using the same dataset. Wu's method uses a particle filter combined with stratified sampling to infer the distribution of the latest change point. The detection power of Wu's method is much lower than that of our proposed method in terms of the detection rate, and the RMSE, but much higher in the computation time. Wu's method is more sensitive to the multiple nonlinear patterns in the signal and detects a series of false critical points, which makes the detection ineffective in our problem context. We randomly pick one sample to show the detection results of Wu's method and our method in Fig. 7, which clearly show that Wu's method fails to detect the critical point in this context.

We also compared our method with the cubic spline with one free interior knot [38] and the piecewise linear regression [13] with two pieces. For the cubic spline method, the location of the knot is regarded as the critical point. For the piecewise linear regression method, the connection point between two pieces is regarded as the critical point. The successful detection rate (SDR), the RMSE, and the computation time are listed in Tables 1–3, respectively. Although the cubic spline method and piecewise linear regression method are faster than our method, both SDR and RMSE indicate our method is superior in the detection power of the critical point.

We finally compared our method with the VAM method [34], which is the current practice in steel industry, the slope detection method (SDM) [39], and the variance ratio test (VRT) method [40]. In SDM, least squares are used to estimate the slope of a moving window. In VRT, we recognize the critical point whose location is with the largest variance ratio of two adjacent moving windows. Given that the resolution of the torque signal is 0.002, we set the window size to 10 to achieve an acceptable precision. The detection results are shown in Tables 1–3. The VAM detection results are provided by the steel plant since it is the current practice. This method can meet the online detection requirement reported from the steel company, although the exact computational time is not released by the plant. As shown in the tables, the mean SDR and the mean RMSE of VAM method are 46.39% lower and 115.79% higher than our method, respectively. The SDM and VRT methods are faster, however, the detection rates of both methods are much lower than our proposed method, and meanwhile their RMSEs are much larger. These results indicate that our proposed method outperforms these three methods.

5. Discussions

Motivated from the premium pipe connection problem, [36] proposed a pairwise critical point detection method over torque signals. Such method highlights the accuracy by designing a two-stage estimation algorithm since the change point detection power is the sole criterion in [36]. A large number of iterations between two stages are required to obtain the precise change location estimation. Hence, the method in [36] is hard to be fully implemented in production lines to meet the time requirements and leading to a lower production efficiency. Differently, our proposed method overcomes such flaw in [36], and leverages the Shannon information quantification of state changes

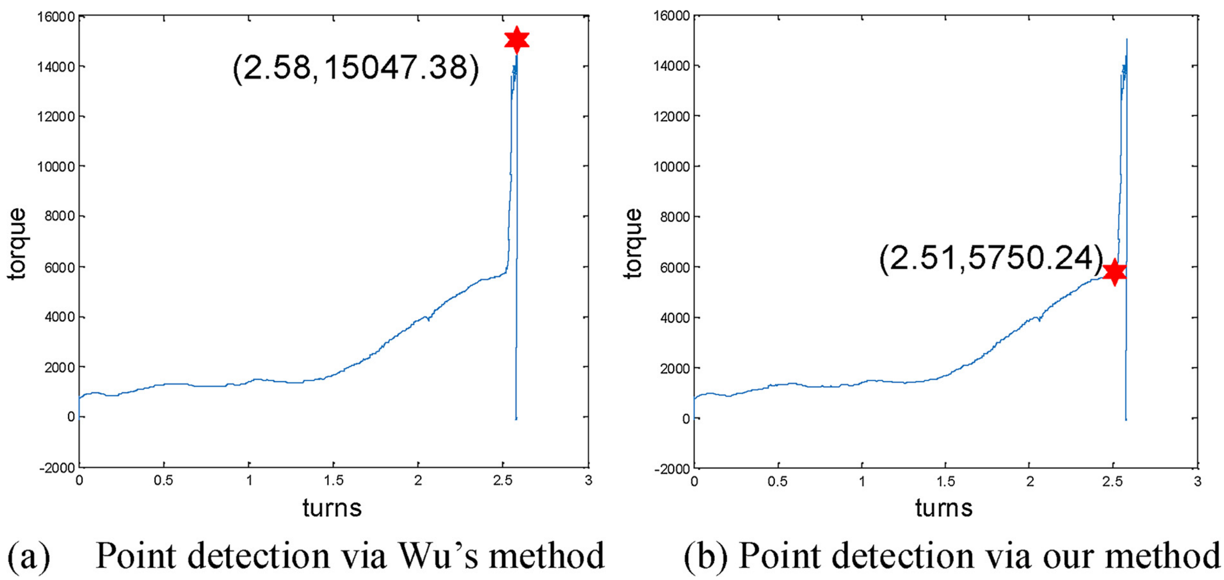

Fig. 7. An example of the comparison between Wu's method and our method.

and estimates the critical point through one stage, thereby significantly reducing computation costs. As shown in Table 3 on Page 9, our method achieves the comparable detection accuracy as [36] but is much faster than [36], which meets the online requirements of steel plants.

From the comparison studies in Section 4, our method outperforms other methods in terms of the detection accuracy and computation time, which can be used for online critical point detection in the tightening process. The key advantage of our method incorporates the engineering domain knowledge of the pipe tightening process into the statistical model and estimation algorithm. Specifically, the Shannon information quantifies the state change during the one-stage estimation algorithm. In this way, we can achieve a timely precise critical point detection. Due to the estimation of critical point depends on the information quantity of the state change, low signal-noise ratio may mask the state change, and further affect the effectiveness of our proposed method. However, for the torque signals in the tightening process, the signal-noise ratio is sufficiently high as shown in Fig. 2 and our method can be implemented effectively.

6. Conclusions

In this paper, a novel method is developed for critical point detection in tightening processes. In-situ sensing is used to measure the torque of a tightening process of pipes, which generates nonlinear profiles with sensing noises. Although the previous study can achieve good critical point detection performance, the algorithm is time consuming, thereby leading infeasibility for online monitoring of fast production lines in tightening processes. Development of an algorithm which achieves both high successful detection rate and low computation time is urgently needed but challenging. To address this problem, we propose a new particle filter algorithm by integrating sampling techniques and quantifying the information of state change for efficient critical point detection.

A real case study on a pipe tightening process was conducted to validate our method. The results show that the proposed method outperforms other existing methods in terms of SDR, RMSE and computation time. For our proposed method, the transition state in the state space model is pre-defined on the basis of engineering knowledge of mechanical deformation and the critical point is identified according to the information quantity of state changes. Thus, its computational time is much shorter than that of the particle filter-based models without incorporation of engineering knowledge. It is worthy to point out that the proposed methodology can be used for critical point detection in other manufacturing process by further investigations of the corresponding physical mechanism of mechanical deformation.

Acknowledgement

This work was partially supported by National Natural Science Foundation of China under Grant No. 71471005, 71690232, 61433001 and 71571003.

Appendix A. Posterior distribution derivation for (9)

$$
\begin{aligned}
&p(\boldsymbol{x}_{0:k}|y_{1:k}) \propto p(y_k|\boldsymbol{x}_{0:k},\ y_{1:k-1})p(\boldsymbol{x}_{0:k}|y_{1:k-1})\\
&\propto p(y_k|\boldsymbol{x}_k)p(\boldsymbol{x}_{0:k}|y_{1:k-1})\\
&\propto p(y_k|\boldsymbol{x}_k)p(\boldsymbol{x}_k|\boldsymbol{x}_{0:k-1},\ y_{1:k-1})p(\boldsymbol{x}_{0:k-1}|y_{1:k-1})\\
&\propto p(y_k|\boldsymbol{x}_k)p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1})\,p(\boldsymbol{x}_{0:k-1}|y_{1:k-1})
\end{aligned}
$$

Appendix B. Basic particle filter algorithm based on importance sampling

At time $k = 1$,

(1) For $m = 1$ to M Sample $\boldsymbol{x}_1^{[m]} \sim g_1(\boldsymbol{x}),\ m = 1,\ 2,\ \ldots, M$ Compute $w_1^{[m]} = p(y_1|\boldsymbol{x}_1^{[m]})$ End

(2) For $m = 1$ to M Draw i with probability $\propto w_1^{[m]}$ Add $\boldsymbol{x}_1^{[i]}$ to χ_1 End At time $k \geq 2$,

(3) For $m = 1$ to M Sample $\boldsymbol{x}_k^{[m]} \sim p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}^{[m]}),\ m = 1,\ 2,\ ...,M$ Compute $w_k^{[m]} = p(y_k|\boldsymbol{x}_k^{[m]})$ End

(4) For $m = 1$ to M Draw i with probability $\propto w_k^{[i]}$ Add $\boldsymbol{x}_k^{[i]}$ to χ_k End

References

[1] Paynabar K, Jin J. Characterization of non-linear profiles variations using mixed-effect models and wavelets. IIE Trans 2011;43(4):275–90.
[2] Zhou S, Sun B, Shi J. An SPC monitoring system for cycle-based waveform signals using Haar transform. Autom Sci Eng IEEE Trans 2006;3(1):60–72.
[3] Yuan G, Yao Z, Wang Q, Tang Z. Numerical and experimental distribution of temperature and stress fields in API round threaded connection. Eng Fail Anal 2006;13(8):1275–84.
[4] Jin J, Shi J. Feature-preserving data compression of stamping tonnage information using wavelets. Technometrics 1999;41(4):327–39.
[5] Yi J, Xu C. Broadband optical end-point detection for linear chemical-mechanical planarization processes using an image matching technique. Mechatronics 2005;15(3):271–90.
[6] Chang Y, Meng Z, Liang L, Li X, Ma N, Hu P. Influence of hot press forming techniques on properties of vehicle high strength steels. J Iron Steel Res Int 2011;18(5):59–63.
[7] Dietz N, Straburg M, Woods V. Real-time optical monitoring of ammonia flow and decomposition kinetics under high-pressure chemical vapor deposition conditions. J Vac Sci Technol A 2005;23(4):1211–25.
[8] Silva C, Butzge J, Nitz M, Taranto O. Monitoring and control of coating and granulation processes in fluidized beds – a review. Adv Power Technol 2014;25(1):195–210.
[9] Page E. Continuous inspection schemes. Biometrika 1954;41:100–15.
[10] Hinkley DV. Inference about the change-point in a sequence of random variables. Biometrika 1970;1(1):1–17.
[11] Lee CB. Bayesian analysis of a change-point in exponential families with applications. Comput Stat Data Anal 1998;27(2):195–208.
[12] Siegmund D, Venkatraman ES. Using the generalized likelihood ratio statistic for sequential detection of a change-point. Ann Stat 1995;23:255–71.
[13] Lund R, Reeves J. Detection of undocumented changepoints: a revision of the two-phase regression model. J Clim 2002;15:2547–54.
[14] Lucas JM, Saccucci MS. Exponentially weighted moving average control schemes: properties and enhancements. Technometrics 1990;32(1):1–12.
[15] Huang X, Zhou Q, Zeng L, Li X. Monitoring spatial uniformity of particle distributions in manufacturing processes using the K function. Autom Sci Eng IEEE Trans 2017;14(2):1031–41.
[16] Harchaoui Z, Lévy-Leduc C. Multiple change-point estimation with a total variation penalty. J Am Stat Assoc 2012;105(492):1480–93.
[17] Pettitt AN. A non-parametric approach to the change-point problem. Appl Stat 1979;28(2):126–35.
[18] Liu S, Yamada M, Collier N, Sugiyama M. Change-point detection in time-series data by relative density-ratio estimation. Neural Netw 2013;43:72–83.
[19] Chen H, Zhang N. Graph-based change point detection. Ann Stat 2015;43(1):139–76.
[20] Desobry F, Davy M, Doncarli C. An online kernel change detection algorithm. Signal Process IEEE Trans 2005;53(8):2961–74.
[21] Ombao HC, Raz JA, Von Sachs R, Malow BA. Automatic statistical analysis of bivariate nonstationary time series. J Am Stat Assoc 2001;96(454):543–60.
[22] Choi H, Ombao H, Ray B. Sequential change-point detection methods for nonstationary time series. Technometrics 2012;50(1):40–52.
[23] Adak S. Time-dependent spectral analysis of nonstationary time series. J Am Stat Assoc 1998;93(444):1488–501.
[24] Olsen LR, Chaudhuri P, Godtliebsen F. Multiscale spectral analysis for detecting short and long range change points in time series. Comput Stat Data Anal 2008;52(7):3310–30.
[25] Last M, Shumway R. Detecting abrupt changes in a piecewise locally stationary time series. J Multivariate Anal 2008;99(2):191–214.
[26] Kong Z, Beyca O, Bukkapatnam ST, Komanduri R. Nonlinear sequential bayesian analysis-based decision making for end-point detection of chemical mechanical planarization (cmp) processes. Semicond Manuf IEEE Trans 2011;24(4):523–32.
[27] Wu J, Chen Y, Zhou S, Li X. Online steady-state detection for process control using multiple change-point models and particle filters. Autom Sci Eng IEEE Trans 2016;13(2):688–700.
[28] Mayne R, Margolis S. Introduction to engineering. New York: McGraw-Hill; 1982.
[29] Doucet A, Johansen AM. A tutorial on particle filtering and smoothing: fifteen years later. Handbook of Nonlinear Filtering 12. 2009. p. 656–704.
[30] Thrun S, Burgard W, Fox D. Probabilistic robotics. Cambridge, MA: MIT press; 2005.
[31] Kong A, Liu JS, Wong WH. Sequential imputations and Bayesian missing data problems. J Am Stat Assoc 1994;89(425):278–88.
[32] Smith AF, Gelfand AE. Bayesian statistics without tears: a sampling–resampling perspective. Amer Stat 1992;46(2):84–8.
[33] Cochran WG. Sampling techniques. New York: Wiley; 2007.
[34] VAM book. [Online]. Available: http://www.vamservices.com/Library/files/VAM%C2%AE%20Book.pdf. 2016 April.
[35] Ruehmann R, Ruark G. Shoulder yielding detection during pipe make up. Offshore Technology Conf. 2011. p. 1–11.
[36] Du J, Zhang X, Shi J. Pairwise critical point detection using torque signals in threaded pipe connection processes. J Manuf Sci Eng 2017;139(9):091002.
[37] Shannon CE. Communication theory of secrecy systems. Bell Labs Tech J 1949;28(4):656–715.
[38] Friedman J, Hastie T, Tibshirani R. The elements of statistical learning. New York: Springer series in statistics; 2001.
[39] Bethea RM, Rhinehart RR. Applied engineering statistics. Marcel Dekker; 1991.
[40] Crow EL, Davis FA, Maxfield MW. Statistics manual: with examples taken from ordnance development. Mineola: Dover Publications; 1960.