



**HAL**  
open science

# Reliability assessment of complex mechatronic systems using a modified nonparametric belief propagation algorithm

Xiaopin Zhong, Mohamed Ichchou, Alexandre Saidi

► **To cite this version:**

Xiaopin Zhong, Mohamed Ichchou, Alexandre Saidi. Reliability assessment of complex mechatronic systems using a modified nonparametric belief propagation algorithm. Reliability Engineering and System Safety, 2010, 95 (11), pp.1174-1185. 10.1016/j.ress.2010.05.004 . hal-02075618

**HAL Id: hal-02075618**

**<https://hal.science/hal-02075618>**

Submitted on 23 Oct 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Reliability assessment of complex mechatronic systems using a modified nonparametric belief propagation algorithm

X. Zhong<sup>a</sup>, M. Ichchou<sup>a,\*</sup>, A. Saidi<sup>b</sup>

<sup>a</sup> Laboratoire de Tribologie et Dynamique des Systèmes, École Centrale de Lyon-36, Avenue Guy de Collongues, 69130 Ecully, France

<sup>b</sup> Laboratoire d'Informatique en Image et Systèmes d'Information, École Centrale de Lyon-36, Avenue Guy de Collongues, 69130 Ecully, France

Various parametric skewed distributions are widely used to model the time-to-failure (TTF) in the reliability analysis of mechatronic systems, where many items are unobservable due to the high cost of testing. Estimating the parameters of those distributions becomes a challenge. Previous research has failed to consider this problem due to the difficulty of dependency modeling. Recently the methodology of Bayesian networks (BNs) has greatly contributed to the reliability analysis of complex systems. In this paper, the problem of system reliability assessment (SRA) is formulated as a BN considering the parameter uncertainty. As the quantitative specification of BN, a normal distribution representing the stochastic nature of TTF distribution is learned to capture the interactions between the basic items and their output items. The approximation inference of our continuous BN model is performed by a modified version of nonparametric belief propagation (NBP) which can avoid using a junction tree that is inefficient for the mechatronic case because of the large treewidth. After reasoning, we obtain the marginal posterior density of each TTF model parameter. Other information from diverse sources and expert priors can be easily incorporated in this SRA model to achieve more accurate results. Simulation in simple and complex cases of mechatronic systems demonstrates that the posterior of the parameter network fits the data well and the uncertainty passes effectively through our BN based SRA model by using the modified NBP.

## 1. Introduction

Reliability analysis is an essential element during the development of mechatronic systems. The attempt to improve system reliability makes the system reliability assessment (SRA) an ongoing research topic. A number of methods for SRA have been developed, most of which estimate the system reliability using only the data of components since it is much easier and less expensive to test the components/subsystems than the entire system. In general, the authors believe that SRA remains a difficult problem for the following reasons.

- *Complexity*: Products are built increasingly complex due to added functionalities. It becomes more difficult to propagate component information to the system level in those systems with high complexity.
- *Expensiveness*: of testing. Testing a whole complex system tends to be impossible. Reliability assessment at the system

level requires a good tool to efficiently combine the information collected from various levels and diverse sources.

- *Functional and temporal dependency*: A failure might be caused by more than one mutually-dependent event, such as shared causes, exclusive events and standby redundancies. The exploitation of explicit dependence is itself a challenge.
- *Uncertainty* due to variability of model, also known as epistemic uncertainty. This arises from a lack of knowledge about the probability distribution of time-to-failure data.

SRA is clearly a problem of system decision with inherent ambiguities. The statistical inference provides a promising methodology for such decision problems by finding the “best guess” for the interesting items given the existing observation. As one of the most popular approaches over the last decade, Bayesian networks (BNs) have many advantages over the classical reliability formalisms when applied to SRA [1,2]. BNs can deal with the problem of dependency under not only temporal but also functional relations. The BN formalism can accept and integrate the diverse data from all kind of sources, such as the prior specifications given by manufacturers, the historical TTF data and the current testing data. In a BN, we can easily propagate the uncertainties, which would be difficult and even impossible with conventional logic models, e.g. fault trees and

\* Corresponding author. Tel.: +33 04 72 18 62 30; fax: +33 04 72 18 91 44.

E-mail addresses: xiaopin.zhong@ec-lyon.fr (X. Zhong), mohamed.ichchou@ec-lyon.fr (M. Ichchou), alexandre.saidi@ec-lyon.fr (A. Saidi).

event trees. Moreover, the methodology of BNs has been so well-studied that we have many options for modeling and reasoning in various situations.

However, the previous research on BNs has not paid enough attention to the parameter uncertainty of failure models that are usually parametric skewed distributions. There are obviously some unobservable items (e.g. no test data) in a complex mechatronic system, thus it is a challenge to estimate the model parameters of those items. The main contribution of this research is that we propose a continuous BN model to estimate the optimal configuration of time-to-failure (TTF) model parameters for a complex mechatronic system using data from diverse sources, and we develop a modified nonparametric belief propagation algorithm to obtain the marginal density in non-singly-connected BNs without the use of a junction tree.

The remainder of this paper is organized as follows. After reviewing the related works in Section 2, we formulate in Section 3 our basic model of system reliability assessment. In Section 4, the nonparametric belief propagation is modified and applied to the inference of the basic model. The simulation results and discussion shown in Section 5 demonstrate that our model is effective and efficient. Finally, some conclusions and perspectives are offered in Section 6.

## 2. Related works

In this section, the related methods of system reliability assessment (SRA) are reviewed and classified, especially those using the BN based approaches. Since there exist various methodologies for modeling and analyzing system reliability, the authors also refer the readers to a more detailed classification by Mihalache [3] and an alternative review by Zio [4].

In the past decades, many methods have been applied to SRA with the help of specific mathematical tools, such as fault trees (FTs) [5], reliability block diagrams [6], petri nets [7], method of moments [8] and so on. Among these approaches, the FT formalism is undoubtedly the most widely used one due to its efficiency and the capability of system abstraction. However, it cannot cope with the temporal or functional dependency among components. To deal with the temporal interrelation, people proposed Markov chain (MC) method [9,10] and dynamic fault trees (DFTs) [11,12]. The problem of combinatorial explosion is nevertheless intractable when applying those two well-known methods to highly-complex mechatronic systems. Moreover, they work only under the assumption that the TTF of components obey the exponential distribution law. In recent years, Bayesian reliability analysis has proved to be a powerful tool providing important methodological advantages over the conventional techniques [13]. The BN methodology is strongly recommended [1] as an effective and versatile SRA tool that is capable of modeling both temporal and functional dependency. Another important advantage over traditional approaches is the ability to combine the information from diverse sources [14]. BNs become attractive due to the ease of building a BN from FT [15] and directly from existing data [16].

In the authors' view, the published approaches based on BNs can be classified into three categories according to the variable type in the network. (1) *Working state based BNs*, also known as *event based BNs*. The working state usually indicates *fail/work* which is inherited from binary decision diagrams (BDD) and FTs. The node variables are thus defined in discrete space. Most published research work [17,18,2] falls into this category. For more flexibility in modeling the working condition, Weber [19] introduced the multi-state variable to describe the degradative states of items in dynamic Bayesian networks (DBNs). These

methods calculate the distribution of system state deterministically by using conditional probability tables (CPTs) established by experts in advance. However, the temporal dependency might be time-varying due to the existence of component deterioration and environment change. An unchangeable CPT may mismatch the time-varying temporal dependency. Furthermore, the working state based BNs are not flexible enough to analyze the risk in continuous time space because the estimation must be re-executed at each timestamp concerned. (2) *TTF model based BNs*. These BNs assign a probabilistic description to the TTF data of an item, see the simple example of binary distribution discussed by Langseth [1]. Boudali [20] presented a non-parametric discrete-time TTF model, and Boudali [21] succeeded in modeling a continuous-time TTF in close-form without considering the model uncertainty. However it is still non-trivial to directly model the TTF because of the complexity of modeling an arbitrary probability density in continuous space. Johnson [22] and Anderson-Cook [23] modeled the distribution parameters of TTF as a continuous unknown variable, such as the scale and the shape of a 2D Weibull density. This facilitates passing information through the network and the reliability analysis at system level based on the characteristics of the skewed distributions, e.g. Weibull analysis. Therefore, the uncertainty of the model parameters must unavoidably be considered. The problem of SRA is hence regarded as an inference of the continuous model parameters in a BN. We notice three advantages compared with the first category: any parametric distribution can be used to model the TTF model parameters uniformly or differently; no CPT needs to be completed, the dependency is specified according to the problems or learned from examples; after statistical inference, the posterior finds the "best guess" solution and fully describes the uncertainty relative to the TTF model parameters. (3) *Hybrid state based BNs*. Neil [24] and Marquez [25] have successfully modeled TTF distribution by continuous random variables as well as by discrete random variables (e.g. working states of items). In fact, the discrete variables can be directly appended into the BNs of the second category when the TTF models of items are ready. In all the research works above, little study was made of parameter uncertainty in the TTF model for system reliability, though it has long been investigated in relation to component reliability [26,27].

On the other hand, a number of statistical inference algorithms have been proposed to perform reasoning in a graph. Exact inference in discrete state BNs is possible with some exact marginalization algorithms for trees, such as variable elimination, belief propagation [28] and junction trees [29]. However, the integral in continuous state space still raises some difficulties. Until recently a number of discretization methods have been suggested to perform the inference in the continuous graph [30], such as mixture of truncated exponentials [31] and dynamic discretization [24,32]. Most of these methods rely on the junction tree algorithm in the case of non-singly-connected BN. However, performing an inference in the junction tree with treewidth  $k_{tr}$  will have at least one computation which requires time exponential in  $k_{tr}$  due to the marginalization process over supernodes. Therefore the junction tree based methods become less attractive for mechatronic systems because the subsystems in a mechatronic system, especially the electronic devices, commonly have a large number of components (input items), which leads to a large treewidth in the generated junction tree. We found that the sampling based nonparametric belief propagation (NBP) [33] is efficient and applicable to our BN model because of the flexible uncertainty representation by means of kernel density estimation (KDE, an open-source toolbox of KDE has been provided by Ihler [34]). Unlike the other inference methods for loopy graphs, the NBP does not propagate information in a

junction tree, but instead it iterates the message updates in a loopy graph until convergence.

From the review above, we see that the previous research has not paid enough attention to the stochastic nature of parametric TTF models in system reliability. In this paper, we formulate a new BN model to cope with the parameter uncertainty of the SRA problem. Our proposed BN model falling into the second category is an extension of the formalism of Bayesian component reliability [13], see Section 3 for the basic model. In contrast with conventional methods, we do not predefine the quantitative part of BNs, but learn the dependency, which is derived from probabilistic noisy gates, from some sample testing data. This model is efficiently solved by a modified version of the NBP algorithm.

### 3. A new model for SRA

Unlike the continuous time-to-failure (TTF) models in [21,35], our method for SRA uses the general BN framework to model the failure parameters of items in the complex system and their dependency. This model allows us to estimate the parametric TTF distributions of all items given the observed failure data as evidence.

As the qualitative part of a BN model, the directed acyclic graph (DAG) can be constructed from the fault tree model by Bobbio's method [15]. However in our model the network variables denote neither the working state of items nor the items' TTF, but the items' failure parameters which are function outputs of the noisy gates rather than the deterministic gates. By "noisy" we do not mean the uncertainty of the output event (success/failure) as do the methods in other published work [15,35], rather we mean the uncertainty of the model parameters of the output failure distribution. For example, a simple OR-gate system C is shown in Fig. 1a with two basic independent components A and B. A, B and C have exponential failure distributions with the failure rates  $\lambda_a$ ,  $\lambda_b$  and  $\lambda_c$  respectively. Their failure times are denoted by  $t_a$ ,  $t_b$  and  $t_c$  respectively. If the OR-gate is deterministic, C fails when either A or B fails. Then C fails in the time interval  $(0, t]$  with the probability  $Pr(t_c \leq t) = Pr(t_a \leq t \cup t_b \leq t) = Pr(t_a \leq t) + Pr(t_b \leq t) - Pr(t_a \leq t)Pr(t_b \leq t)$ . Notice that for an exponential failure distribution  $Pr(t) = 1 - e^{-\lambda t}$ . Therefore the failure rate of C can be obtained from the equation above, i.e.  $\lambda_c^d = \lambda_a + \lambda_b$ , where the superscript *d* represents the deterministic case. Analogously in the case of the deterministic AND-gate,  $Pr(t_c \leq t) = Pr(t_a \leq t \cap t_b \leq t)$ , then  $\lambda_c^d = -\frac{1}{t} \ln[e^{-\lambda_a t} + e^{-\lambda_b t} - e^{-(\lambda_a + \lambda_b)t}]$ . We say that the gate is noisy or uncertain, then the failure rate is contaminated by a noise, such as  $\lambda_c = \lambda_c^d + w_c$  where  $w_c$  is an additive noise. This parameter uncertainty can be perceived by the stochastic nature of failure density in reliability analysis, see Fig. 2. Indeed, this uncertainty is common in a mechatronic system for at least the following reasons.

First, not all failure modes can be found because of the large number of components in the electronic and mechanical subsystems. The modes not found can have more or less influence on the gates. Second, the multi-parameter skewed distributions are widely used for a mechatronic system due to its various failure modes. For some multi-parameter distributions, the deterministic output parameters may not have any solution because we have only one equation (the output failure distribution). Third, the effect of environment is usually unmeasurable, especially for the electronic and mechanical subsystems.

None of the previous research has addressed this noise which is difficult to measure. Fortunately, this uncertainty can generally be described by a statistical model on the test data acquired from the basic components as well as from the system. The SRA problem can then be formulated as a probabilistic reasoning process for the failure parameter of each item given the test data.

As analyzed above, it is difficult to model the parameter uncertainty directly. For this reason we turn to the TTF density that is a function of the parameters. We denote the density by  $f_i(X_i, t)$ ,  $i \in \mathcal{I}n$ , where  $\mathcal{I}n$  is the label set of the nodes in the BN and  $X_i = x_i$  is a unified notation for all types of multidimensional parameters. For brevity the time is hereafter omitted in density expressions. For instance, the exponential TTF density of  $X_c$  in Fig. 1c is  $f_c(X_c = \lambda_c) = f_c(\lambda_c) = \lambda_c e^{-\lambda_c t}$ ,  $t \geq 0$ .

Fig. 2 shows the stochastic nature of failure density in reliability analysis. Thus a probability distribution is expected to represent the stochastic nature of  $f_i(X_i)$ ,  $i \in \mathcal{I}n$ . Notice that a higher probability should be assigned to an uncertain density that is closer to the deterministic density. Then some symmetrical distributions, such as the normal and the triangular distributions, can be selected when no other prior information is known [26]. In this paper, the normal distribution is used because of its flexibility. Further the "closer" is measured based on a distance

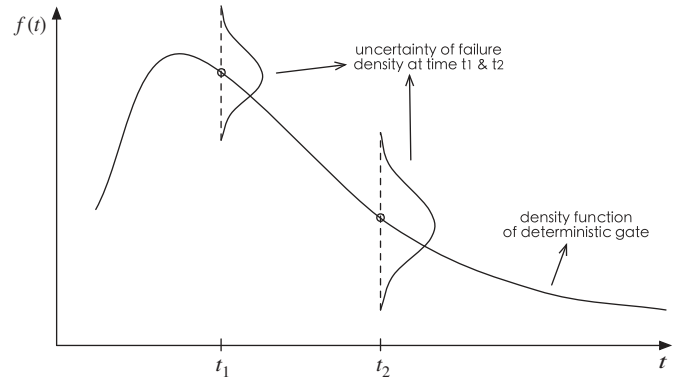


Fig. 2. The stochastic nature of failure density in reliability analysis.

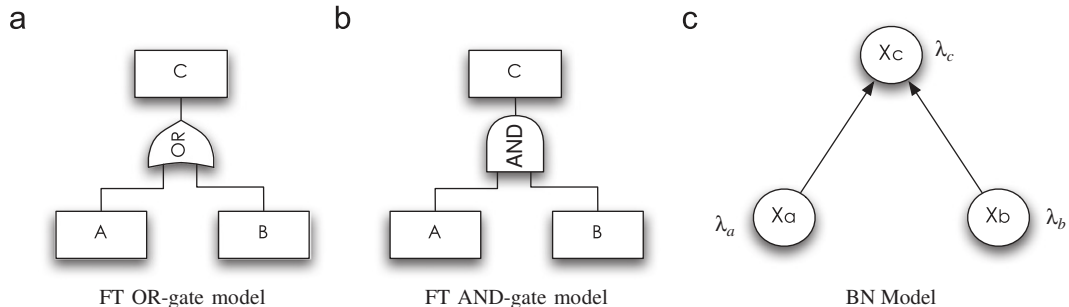


Fig. 1. The simplest OR-gate and AND-gate models described by Fault tree (FT) and Bayesian network (BN). The BN model is converted from the FTs by the method in [15]. (a) FT OR-gate model, (b) FT AND-gate model and (c) BN model.

function  $d(\cdot, \cdot)$  between probability densities. Thus the parameter uncertainty can be reduced to the distance uncertainty,

$$d(f_i(X_i), f_i^d) \sim \mathcal{N}(0, \sigma_i), \quad i \in \mathcal{I}n, \quad (1)$$

where  $f_i^d$  denotes the deterministic output density function of the gate going to node  $i$ .  $\mathcal{N}(0, \sigma_i)$  is a normal distribution with zero mean and standard deviation  $\sigma_i$  which will be discussed at the end of this section. Note that a distance cannot be negative. A one-sided normal distribution is then considered. Then given  $\sigma_i$ , we have

$$\Pr[d(f_i(X_i), f_i^d) | \sigma_i] = \frac{2}{\sqrt{2\pi}\sigma_i} \exp\left\{-\frac{d^2(f_i(X_i), f_i^d)}{2\sigma_i^2}\right\}, \quad i \in \mathcal{I}n. \quad (2)$$

In fact this ‘‘one-sided’’ transformation does not affect the simulation results of those parameters because of the normalization stage in a probabilistic inference algorithm. The formulation above plays the role of output parameter prediction based on the parameters of the basic items.

For the gates with more than two basic items, the deterministic output density functions  $f_i^d$  can still be derived without difficulty [13]. For a series structure, the system functions only if all components function, i.e. for  $i \in \mathcal{I}n$ ,

$$f_i^d = \sum_{j \in \mathcal{B}(i)} f_j(x_j) \cdot \prod_{k \in \mathcal{B}(i) \setminus j} (1 - F_k(x_k)), \quad (3)$$

where  $\mathcal{B}(i)$  denotes the label set of node  $i$ 's basic component nodes and  $F_k(x_k)$  is the cumulative probability function of  $f_k(x_k)$ .  $\mathcal{B}(i) \setminus j$  means the set  $\mathcal{B}(i)$  except  $j$ . For a parallel structure, the system functions if any component functions, i.e. for  $i \in \mathcal{I}n$ ,

$$f_i^d = \sum_{j \in \mathcal{B}(i)} f_j(x_j) \cdot \prod_{k \in \mathcal{B}(i) \setminus j} F_k(x_k). \quad (4)$$

Series and parallel structures are special cases of  $k$ -of- $n$  systems, which can be re-structured as  $C_n^k$  parallel subsystems of  $k$  series components. Hence any  $k$ -of- $n$  system can also be analyzed by a repeated calculation of Eqs. (3) and (4).

So far we can obtain the dependency between the output parameter and the parameters of basic items in Eqs. (2)–(4). Finally the quantitative part of our BN model, i.e. the conditional probability distribution (CPD), is defined by

$$\Pr(X_i = x_i | X_j = x_j)_{j \in \mathcal{B}(i)} = \Pr[d(f_i(X_i), f_i^d) | \sigma_i], \quad i \in \mathcal{I}n. \quad (5)$$

There are many choices for the distance function between probability densities [36], such as Kullback–Leibler (KL) divergence and Euclidean distance. The following example demonstrates the effect of the dependency described above.

**Example.** A simple exponential-distributed series system consists of two components, whose FT model has the same structure as that of Fig. 1a. The TTF of node A is modeled as exponential density with failure rate  $\lambda_a = 0.68$  and that of node B is modeled as Weibull distribution with shape  $k_b = 1.5$  and scale  $\lambda_b = 1$ . We discretize the densities (time interval =  $10^{-3}$ s) and take the KL divergence as distance metric. The distribution of the system parameter ( $X_c$ ) given the parameters of the basic items, i.e.  $\Pr(X_c = \lambda_c | X_a = \{\lambda_a\}, X_b = \{k_b, \lambda_b\})$ , is then calculated from Eq. (5) and shown in Fig. 3 for different deviation value  $\sigma_c$ . From Fig. 3 we observe that the smaller  $\sigma_c$  is, the sharper the distribution is. In other words, the gate tends to be deterministic with the decrease of  $\sigma_c$ . While  $\sigma_c$  increases, the information from the basic items to the system tends to be cut off by the gate because  $\Pr(X_c | X_a = \{\lambda_a\}, X_b = \{k_b, \lambda_b\})$  tends to be uniform and no uncertainty can pass through a uniform CPD.

The remainder of this dependency model is the standard deviation  $\sigma_i$  in Eq. (2) that encodes the dependency between the parameters of the basic items and those of the higher level

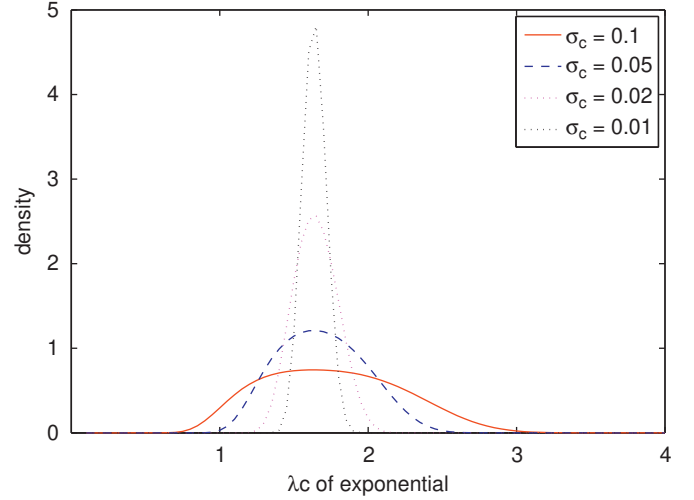


Fig. 3. The distribution of output parameter given the parameters of the basic items.

item. In order to achieve the best distinguishing capacity of Eq. (2), this standard deviation can be estimated by a parameter estimation technique. The maximum likelihood estimation (MLE) approach is selected in our implementation because of its ease of use. If there exists adequate sample data, MLE yields

$$\sigma_i^* = \operatorname{argmax}_{\sigma} \prod_j \Pr(y_j | \sigma_i), \quad i \in \mathcal{I}n, \quad (6)$$

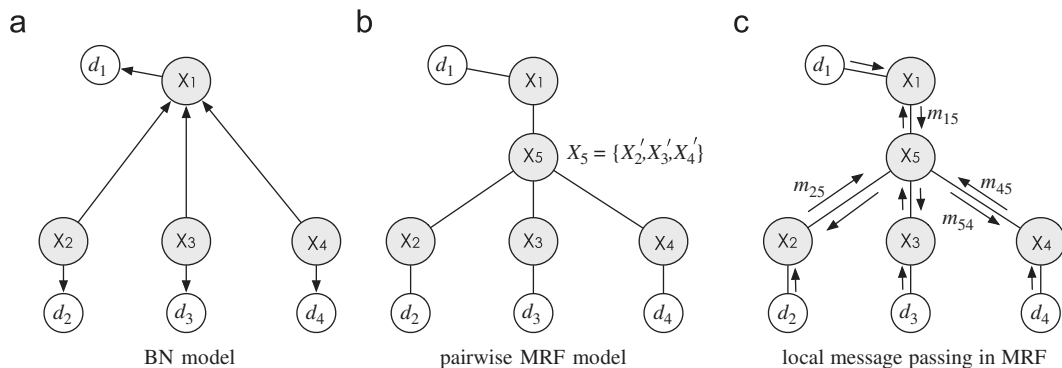
where  $y_j = d(f_i^{(j)}, f_i^d)$  is called pseudo-observation that represents the distance between a sample candidate density  $f_i^{(j)}$  and the deterministic density  $f_i^d$ . The above formulation can be converted to logarithmic interpretation to avoid the trouble of exponential overflow. In practice, we reuse this parameter  $\sigma$  for the same type of components or subsystems. The readers are referred to Section 5 for a simulation example of this standard deviation estimation.

#### 4. Approximate inference

After modeling, the Bayesian network needs a process of probabilistic reasoning, i.e. an approximate inference in the continuous networks. The inference task of a BN is to calculate the posterior marginal probability of each hidden variable given the observed variables. We have various methods to perform the approximate inference in a continuous network [30]. As described in Section 2 the conventional inference methods for non-singly-connected continuous graphs depends on a junction tree, such as dynamic discretization in [35,32]. However, the computation of exponential in treewidth makes the junction tree based methods less attractive for mechatronic applications. Recently a nonparametric belief propagation (NBP) in Markov random fields (MRFs) has attracted our attention because of its reasoning capacity and efficiency in the general loopy graphs. In particular, it is a linear time algorithm proportional to the number of hidden variables. To facilitate the inference process of NBP, we need to convert the BN to a corresponding MRF. Then following a brief introduction of NBP, a modification on the NBP is made to adapt the NBP to the newly-generated MRF.

##### 4.1. NBP algorithm in pairwise MRF

According to Weiss' method [37], arbitrary BN can be converted to an undirected graphical model, i.e. pairwise MRF. An



**Fig. 4.** An example of the conversion from (a) a BN to (b) the corresponding pairwise MRF and (c) the local message passing in the MRF. The shaded circles represent the hidden nodes and the white circles denote the observed ones. (a) BN model, (b) pairwise MRF model, and (c) local message passing in MRF.

example of this conversion is demonstrated in Fig. 4. The method is to add a cluster node for the parent nodes that share a common child in the BN, add the connection between the cluster node and the parents/child, then replace the directed arcs with undirected ones. In the example, the variable for the newly-created node is denoted by  $X_5$  which is a compound variable of the parents, i.e.  $X_5 = \{X_{5,2}, X_{5,3}, X_{5,4}\}$ . We say  $X_5$  and  $X_4$  are *consistent* with each other if  $x_{5,4} = x_4$  and *inconsistent* if  $x_{5,4} \neq x_4$ . For convenience in the pairwise MRF, we still call node 1 the child of node 5 and node 2 (or 3, 4) a parent of node 5. The child set and the parent set of a cluster node, say  $i$ , are denoted by  $Ch(i)$  and  $Pa(i)$  respectively. Furthermore  $\mathcal{I}c$  refers to the label set of the cluster nodes and  $d_i$  denotes the corresponding test data (e.g. TTF) of node  $i$ ,  $i \in \mathcal{I}n$ .

Besides the qualitative specification, the corresponding MRF and the original BN should have the same joint probability distribution, which is a quantitative part of the graphical model. Suppose that all probabilities are positive, the Hammersley-Clifford theorem [28] guarantees that the joint probability distribution can be factorized into a product of the compatibility functions of the cliques over the graph. Thus the joint density of a pairwise MRF is factorized as follows

$$Pr(X, D) \propto \prod_{i \in \mathcal{I}} \phi_i(X_i) \prod_{i \in \mathcal{I}c} \prod_{j \in \Gamma(i)} \psi_{ji}(X_j, X_i), \quad (7)$$

where  $\mathcal{I} = \{\mathcal{I}n, \mathcal{I}c\}$ .  $X$  and  $D$  denote the joint variable and joint observation respectively.  $\Gamma(i)$  denotes the label set of neighbors of node  $i$ . The quantitative specification of a pairwise MRF is then composed of  $\phi_i(\cdot)$  and  $\psi_{ji}(\cdot, \cdot)$ , named as first-order potential and second-order clique potential respectively. They encode the dependency between nodes and are set in the following way so that the joint probability distribution of the MRF is identical to that of the BN:  $\phi_i(X_i) = Pr(d_i|X_i)$  represents the likelihood of observed data  $d_i$  with respect to variable  $X_i$ , implying the local evidence;  $\psi_{ji}(X_j, X_i)$  is the child's conditional probability given the corresponding cluster variable if it is a potential between a cluster node and the corresponding child (e.g. node 5 and node 1 in Fig. 4b); otherwise,  $\psi_{ji}(X_j, X_i)$  is set to one if the compound node is consistent with the parent node and zero otherwise. For the example in Fig. 4b,  $\psi_{51}(X_5, X_1) = Pr(X_1|X_5)$ ,  $\psi_{25}(X_2, X_5) = 1$  if  $X_{5,2} = X_2$  ( $X_2$  and  $X_4$  are consistent with each other) and 0 when  $X_{5,2} \neq X_2$  (inconsistent). Finally  $Pr(X, D) = \prod_i Pr(d_i|X_i) \cdot Pr(X_1|X_2, X_3, X_4)$ , which is identical to that of the BN model.

In the belief propagation (BP) algorithm, we use a so-called "message" from a hidden node (say  $j$ ) to another one (say  $i$ ) to describe how likely node  $i$  will be in one state given the knowledge of node  $j$ , i.e.  $m_{ji}(X_i)$ , see Fig. 4c for message passing paths. The belief at node  $i$  denotes how likely node  $i$  should be in one state given the information from neighbors and from its local evidence. It is then proportional to the product of local evidence

and all the messages coming to node  $i$ , i.e.

$$b_i(X_i) \propto \phi_i(X_i) \cdot \prod_{j \in \Gamma(i)} m_{ji}(X_i). \quad (8)$$

Accordingly, the estimates of minimum mean square error (MMSE) and maximum a posterior (MAP) are obtained based on this belief. Furthermore, the messages are updated iteratively by the rule

$$m_{ji}^t(X_i) = \alpha \sum_{X_j} \psi_{ji}(X_j, X_i) \cdot \phi_j(X_j) \cdot \prod_{k \in \Gamma(j) \setminus i} m_{kj}^{t-1}(X_j), \quad (9)$$

where  $t$  indicates the iteration index and  $\alpha$  is a normalization constant. The summation is replaced by an integral with respect to  $X_j$  for a continuous network. The BP is also known as a sum-product algorithm. It has been proven [38] that in the case of a singly-connected graph, the BP in fact converges and gives the exact marginal probabilities for all the hidden nodes. Although the BP cannot guarantee the convergence in the loopy case, the BP has been applied with experimental success [39].

Until recently, the techniques of Monte Carlo sampling and kernel density estimation has provided an efficient way to perform BP inference in loopy continuous graphs, named nonparametric belief propagation (NBP, [33]). It is called nonparametric because the message is represented by a mixture of Gaussian which is a sample-based or kernel density estimate

$$m_{ji}(X_i) = \sum_{n=1}^N w_{ji}^{(n)} \cdot \mathcal{N}(X_i; \mu_{ji}^{(n)}, A_{ji}), \quad (10)$$

where  $w_{ji}^{(n)}$  is a weight associated with  $n$ th kernel mean  $\mu_{ji}^{(n)}$ , and  $\sum_n w_{ji}^{(n)} = 1$ .  $N$  is the number of kernel means used for density estimation.  $A_{ji}$  is the common bandwidth. The belief (Eq. (8)) and message updating (Eq. (9)) are implemented by using sampling-based approximations in two stages: message product and message propagation.

In the first stage, the product of  $m$   $N$ -component messages yields a new mixture of Gaussian with  $N^m$  components. Each component and its weight can be obtained by the following equations

$$\mathcal{N}(\bar{X}; \bar{\mu}, \bar{A}) \propto \prod_{i=1}^m \mathcal{N}(\bar{X}; \mu_i, A_i), \quad (11a)$$

$$\bar{A}^{-1} = \sum_{i=1}^m A_i^{-1}, \quad \bar{A}^{-1} \bar{\mu} = \sum_{i=1}^m A_i^{-1} \mu_i, \quad (11b)$$

$$\bar{w} = \prod_{i=1}^m w_i \mathcal{N}(\bar{X}; \mu_i, A_i) / \mathcal{N}(\bar{X}; \bar{\mu}, \bar{A}). \quad (11c)$$

To avoid the exponentially-large number of mixture components, a Gibbs sampler [40] is employed to draw  $N$  independent samples as



kernel means of the message product. In our experience, about one hundred kernels make it possible to achieve satisfying approximations of the 2D and 3D densities even for a rare event. It is not surprising that this approximation is much faster than the other stochastic simulation methods, such as Markov chain Monte Carlo (MCMC) which requires many more samples to stochastically update the density representation.

In the second stage, the NBP propagates each component of a message by approximating the integral in Eq. (9). Before the propagation, the product of  $\int \psi_{ji}(X_j, X_i) dX_i$  and  $\phi_j(X_j)$  are regarded as a marginal influence and calculated as  $\xi(X_j)$ ,

$$\xi(X_j) = \phi_j(X_j) \int_{X_i} \psi_{ji}(X_j, X_i) dX_i. \quad (12)$$

For message  $m_{ji}(X^i)$ , each sample  $x_j^{(n)}$  drawn from the product  $\xi(X_j)\phi_j(X_j)\prod_{k \in \Gamma(j), i} m_{kj}(X_j)$  in the first stage is propagated to node  $i$  by sampling  $X_i$  from  $\psi_{ji}(x_j^{(n)}, X_i)$  using importance sampling or MCMC techniques. After obtaining those output samples as kernel means of the updated message, we can choose a kernel bandwidth  $A_i$  by leave-one-out likelihood cross-validation [41]. We refer the readers to [33] for the detailed generic NBP algorithm.

#### 4.2. Modified NBP

We notice that the generic NBP may encounter two problems in the newly-generated MRF when propagating the message between a cluster node and its parents, because of the specification of second-order clique potentials generated by the conversion from BN to MRF. Without loss of generality, we suppose node  $s \in \mathcal{I}c$  is a cluster node and node  $p \in \mathcal{P}a(s)$  is one of its parents. Like the example in Fig. 4b  $X_5 = \{X_{5,2}, X_{5,3}, X_{5,4}\}$ , we denote by  $X_{s,p}$  the corresponding component of node  $p$  in compound variable  $X_s$ . Then we discuss the message propagation stage for the constructions of  $m_{sp}(X_p)$  and  $m_{ps}(X_s)$  as follows respectively.

First, since the second-order clique potential between  $X_s$  and  $X_p$  is a delta function, any common sampling technique for  $X_s \sim \psi_{ps}(X_p, X_s)$  or  $X_p \sim \psi_{sp}(X_s, X_p)$  is unfeasible. We suggest making them consistent with each other instead of sampling. For example, given a message product sample  $x_s^{(n)}$  of  $X_s$  we set  $x_p^{(n)} = x_{s,p}^{(n)}$  as a new sample of  $X_p$ . Vice versa, when given a sample  $x_p^{(n)}$  of  $X_p$  we set  $x_{s,p}^{(n)} = x_p^{(n)}$ . For the other components in  $X_s$ , sampling techniques can still be used. However,  $\psi_{ps}(X_p, X_s)$  does not provide any information about the other components except that of node  $p$ . Any sampling method will sample from a uniform distribution for the other components.

The second concern is that sampling the other components from uniform distributions will degenerate the message product  $\prod_{p \in \mathcal{P}a(s)} m_{ps}(X_s)$ . We use a two-component example for further explication. When a system, say  $X_3$ , consists of two basic components, say  $X_1$  and  $X_2$ , we can refer  $X_4 = \{X_{4,1}, X_{4,2}\}$  to the compound node. Then the kernel means of message  $m_{14}(X_4)$  are sampled as described in the first issue, i.e.  $x_{4,1}^{(n)} = x_1^{(n)}$  and  $X_{4,2}$  sampled from uniform distribution. The kernels of message  $m_{24}(X_4)$  are obtained in the same manner. Before propagating those messages to node 3, a message product is required. Fig. 5 shows a possible kernel of  $m_{14}(X_4)$  and two possible kernels of  $m_{24}(X_4)$ . According to the rule of message product in Eqs. (11a)–(11c), their product is also a normal function whose mean is a linear interpolation of two kernels. Obviously, the important region for the following message propagation is a small area where the samples get high probability in both components, such as the ellipse region in Fig. 5. This is because concentrating the samples in this area can reduce the estimation variance and failing to do so leads to a biased estimation of the density [42]. However in this two-component example, most of these product

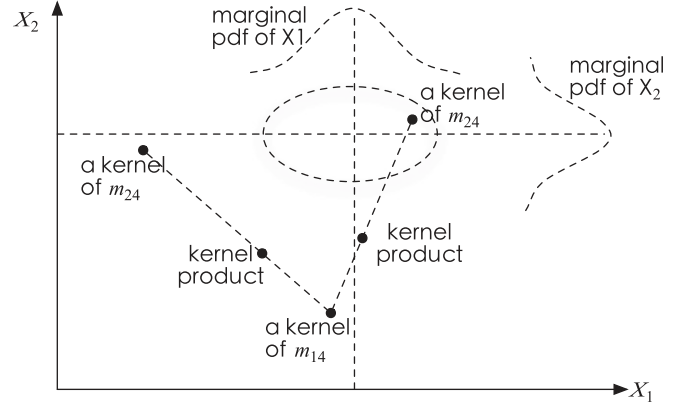


Fig. 5. The degeneration problem of message product in a two-component case.

Table 1

Algorithm of modified NBP for pairwise MRFs converted from BN by Weiss' method [37].

1. Initialization :
  - (1) sample randomly  $N$  kernel means for all messages,
  - (2) initialize the messages  $m_{ij}^0, i \in \mathcal{I}, j \in \Gamma(i)$  as uniform distributions, i.e.  $w_{ji}^{(n)} = 1/N$ ,
  - (3) initialize the iteration step  $t=0$  and the maximum iteration number  $T$ .
2. Updating the messages and the beliefs for  $i \in \mathcal{I}, j \in \Gamma(i)$  : given input messages  $m_{kj}^t(X_j) = \{x_{kj}^{(n)}, A_{kj}^{(n)}, w_{kj}^{(n)}\}_{n=1}^N$ , where  $k \in \Gamma(j), i$ , construct  $m_{ji}^{t+1}(X_i)$  as follows
  - (1) determine the marginal influence of Eq. (12) to get  $\xi(X_j)$ ,
  - (2) draw  $N$  samples  $\{x_j^{(n)}\}_{n=1}^N$  from the product  $\xi(X_j)\prod_k m_{kj}^t(X_j)$  by using Eqs. (11a)–(11c) and the Gibbs sampler depicted in [33],
  - (3) for  $n=1, \dots, N$ , update kernel means :
    - if  $j \in \mathcal{I}c$  (or  $i \in \mathcal{I}c$ ) and  $i \in \text{Ch}(j)$  (or  $j \in \text{Ch}(i)$ ), draw a sample from  $x_j^{(n)} \sim \psi_{ji}(X_j = x_j^{(n)}, X_i)$  by importance sampling or MCMC technique,
    - if  $j \in \mathcal{I}c$  and  $i \in \mathcal{P}a(j)$ ,  $x_{ji}^{(n)} = x_{ji}^{(n)}$ ,
    - if  $i \in \mathcal{I}c$  and  $j \in \mathcal{P}a(i)$ ,  $x_{ji}^{(n)} = x_j^{(n)}$ , the remaining components are sampled from previous beliefs by Gibbs sampler, i.e.  $x_{ji,r}^{(n)} \sim \phi_r(X_r)\prod_{u \in \Gamma(r)} m_{ur}^t(X_r), r \in \mathcal{P}a(i) \setminus j$ ,
  - (4) Construct  $m_{ji}^{t+1}(X_i) = \{x_{ji}^{(n)}, A_{ji}^{(n)}, w_{ji}^{(n)}\}_{n=1}^N$  :
    - set  $w_{ji}^{(n)}$  as the importance weights (if they exist) generated by sampling techniques,
    - select a bandwidth  $A_{ji}$  by using the method in [41].
3. If  $t+1 < T$ , set  $t=t+1$  and go to step 2.
4. Compute the MMSE and MAP estimations, for  $i \in \mathcal{I}n$ :
  - (1) compute the beliefs and normalize them,  $b_i(X_i) \propto \phi_i(X_i)\prod_{k \in \Gamma(i)} m_{ki}^T(X_i)$ ,
  - (2) find the optimum,  $\hat{x}_i^{\text{MAP}} = \text{argmax}_{X_i} b_i(X_i)$ ,  $\hat{x}_i^{\text{MMSE}} = \sum_n b_i(X_i) x_i^{(n)}$ .

kernel means ( $N^m$  in total) are spread out instead of being concentrated into the important region. Therefore there will be a degeneration of the sample set for the message product. It occurs in the multi-component case in the same way.

Our strategy to overcome this degeneration is to obtain the samples of the other components from their corresponding beliefs instead of from uniform distributions. For the above two-component example, given all the messages of previous time step i.e.  $m_{ji}^{t-1}(X_i), i \in \mathcal{I}, j \in \Gamma(i)$ , and the kernel mean set of current message product for node 1 i.e.  $\{x_j^{(n)}\}_{n=1}^N$ , the kernel means of current message  $m_{14}(X_4)$ , denoted by  $x_4^{(n)} = \{x_{4,1}^{(n)}, x_{4,2}^{(n)}\}$  where  $n=1, \dots, N$ , are sampled as follows:  $x_{4,1}^{(n)} = x_1^{(n)}$ ; draw a sample  $x_{4,2}^{(n)}$  from  $\phi_2(X_2)\prod_{k \in \Gamma(2), 1} m_{k2}^{t-1}(X_2)$ . In fact  $\Gamma(2) \setminus 1 = \emptyset$  for this two-component example. The multi-component case can be deduced by analogy. Now the samples of message product are certainly concentrated into the important region by this means.

In summary, a modified version of NBP is then proposed for our pairwise MRF and Table 1 shows the steps of the algorithm. In

practice, our simulation demonstrates that one hundred kernels is enough to achieve good convergence results.

### 5. Simulation analysis: application to mechatronic systems

In this section, we demonstrate simulations on a hypothetical mechatronic system: an active vehicle suspension (AVS). We start by giving a brief description of the complex system. Our model is then applied to its passive subsystem with/without observation and the entire system itself. These simulations are all implemented based on the Kernel Density Estimation (KDE) Matlab toolbox [34].

The AVS system is intended to support the vehicle body and reduce body vibration from the road surface. Fig. 6 shows the fault tree of some important components/subsystems. Apart from the linkage which is considered to be un failing, the parallel system is composed of two subsystems: a passive subsystem and an actuator subsystem. When the two subsystems both function, the suspension works in semi-active mode. It fails if none of the subsystem functions. But then it works in passive mode if only the

actuator subsystem fails and in active mode if only the actuator subsystem functions. Moreover, the passive subsystem works in a series structure with the spring and damper components. So does the actuator subsystem whose failure can be provoked by a series mechanical part or by a series electronic part.

This FT description (Fig. 6) is then converted to BN representation (Fig. 7a), by the transition algorithm of Bobbio [15]. For simplicity, only constant failure rates of the root components are taken from the reliability data base [43] and given in Table 2, where the unit of failure rate has been converted to failure-per-kilometer by assuming a constant velocity. Moreover, their observations are sampled from the corresponding exponential distribution of km-to-failure (KTF).

#### 5.1. Passive subsystem

##### 5.1.1. Learning the dependency

The passive subsystem is a simple series case with only three nodes. The failure rate (FR) of the subsystem is thus  $0.785 \times 10^{-7}$  F/km when decided deterministically. Given its true FR we can obtain the optimal deviation parameter  $\sigma_2^*$  of node 2 by

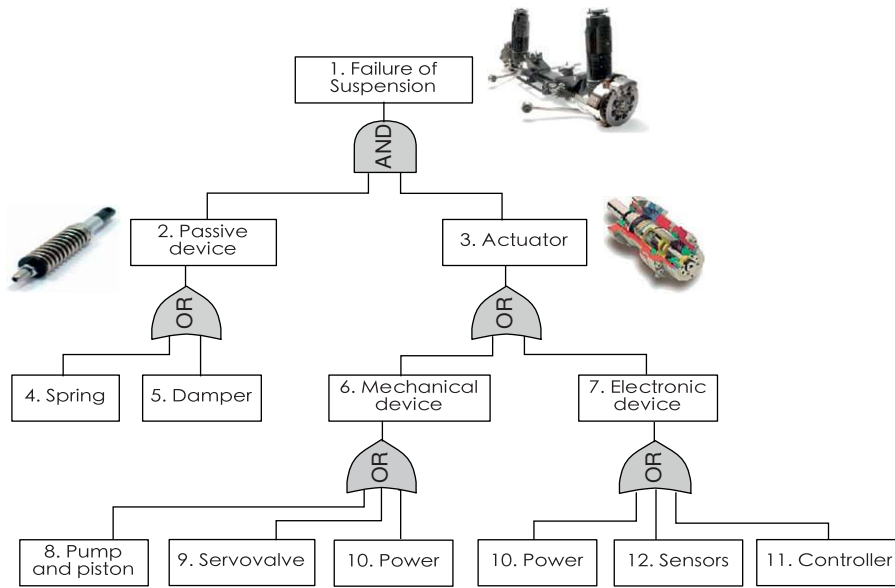


Fig. 6. A fault tree example of an active vehicle suspension.

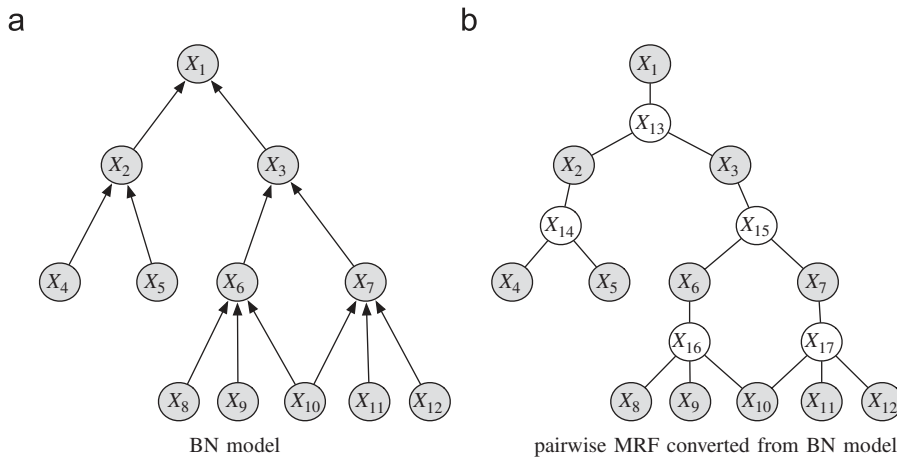


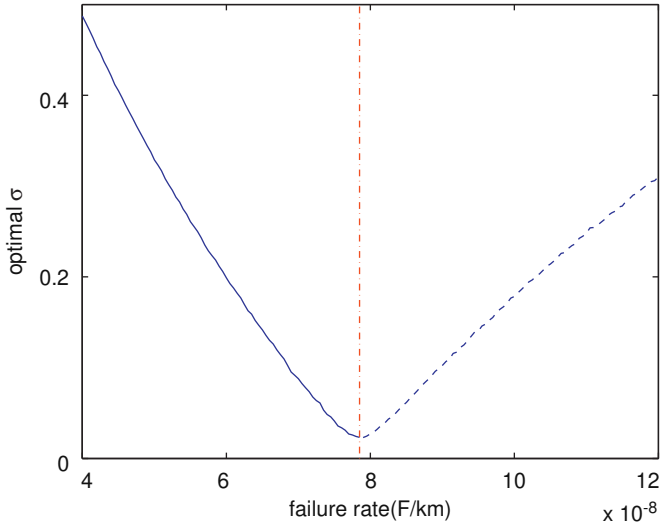
Fig. 7. The corresponding BN model and pairwise MRF model of an active vehicle suspension. The gray circles represent hidden nodes and the white ones denote cluster nodes. (a) BN model and (b) pairwise MRF converted from BN model.



**Table 2**

Failure rates and observed km-to-failure (KTF) of components.

Item name	FR ( $F/10^7$ km)	ObservedKTF ( $\times 10^7$ km)
$X_4$ . Spring	0.105	12.29, 3.47, 5.75, 7.80, 18.72, 1.47, 8.95, 42.04, 12.61, 2.60
$X_5$ . Damper	0.68	1.76, 0.25, 2.62, 1.21, 1.40, 0.19, 0.84, 1.25, 2.73, 2.97
$X_8$ . Pmp & pst.	11.23	0.04, 0.09, 0.06, 0.08, 0.23, 0.13, 0.19, 0.15, 0.13, 0.08
$X_9$ . Servovalve	1.51	0.62, 0.94, 0.60, 1.55, 1.34, 0.04, 0.03, 0.37, 1.87, 0.96
$X_{10}$ . Power	0.34	8.83, 0.30, 0.17, 2.09, 2.10, 3.19, 0.31, 2.93, 6.46, 0.73,
$X_{11}$ . Controller	0.091	11.44, 2.17, 45.86, 34.57, 19.53, 4.75, 3.43, 4.77, 8.75, 6.63
$X_{12}$ . Sensors	2.6	0.47, 0.11, 0.64, 0.14, 0.65, 0.38, 0.18, 0.10, 0.97, 0.03,
$X_2$ . Passive dev.	0.785	0.43, 5.21, 0.28, 2.48, 0.93, 1.73, 1.26, 0.52, 2.26, 1.62,
		2.06, 2.08, 1.42, 0.16, 0.95, 1.15, 2.18, 0.04, 1.14, 0.21,
		0.618, 1.24, 0.16, 0.30, 0.97, 0.26, 0.13, 1.07, 1.39, 0.65

**Fig. 8.** Monte Carlo simulation of learning  $\sigma$  in Eq. (6), where  $X_4=0.105 \times 10^{-7}$  F/km,  $X_5=0.68 \times 10^{-7}$  F/km.

Monte Carlo simulation of Eq. (6). Thanks to the easy manipulation of discretized probability density function (PDF), the density distance measure in Eq. (2) is defined as a city block difference between two densities,

$$d(r, q) = s_t \cdot \sum_i |r_i - q_i|, \quad (13)$$

where  $r_i$  and  $q_i$  are discretized values of the two probability densities  $r$  and  $q$ .  $s_t$  is the constant interval of the discretization. A curve of  $\sigma_2^*$  with respect to actual FR  $\lambda_2^{true}$  is illustrated in Fig. 8 after maximum likelihood estimation (MLE). For a series structure, the FRs greater than the deterministic one (sum of the components' FR) are not feasible, thus plotted by a dashed line. It is also found that the optimal  $\sigma$  is monotonically decreasing with respect to the distance from observed FR to the deterministic one that is indicated in the figure by dash-dot straight line. This is sound because the increasing FR implies decreasing dependency (see in Eq. (2)) when it is less than the deterministic one.

### 5.1.2. SRA of unobserved subsystem

Without loss of generality, the KTF data is modeled by a one-parameter exponential distribution. Using the i.i.d. KTF data observed in Table 2, the likelihoods of the observation with respect to FR are calculated as follows:

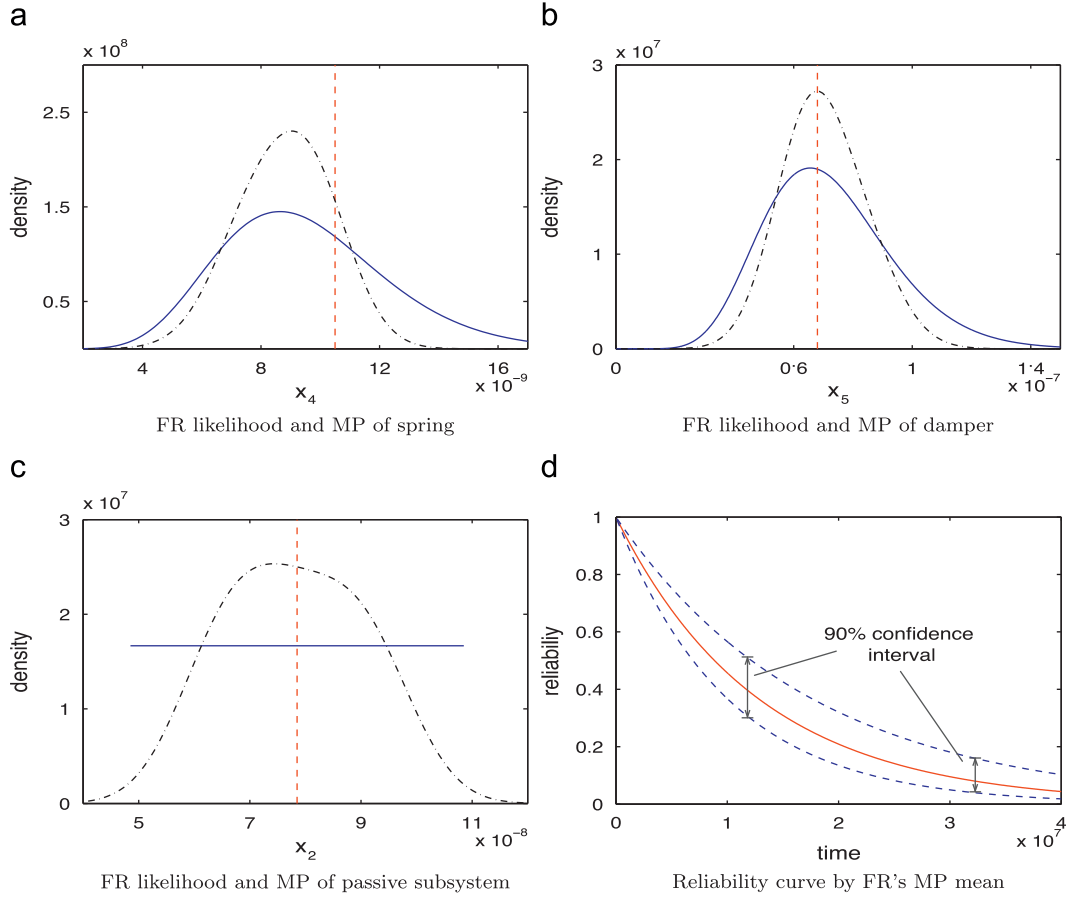
$$L_i(d_i|x_i) \propto \prod_{n=1}^{10} \Pr(d_i^{(n)}|x_i), \quad i = 4, 5, \quad (14)$$

where  $d_i^{(n)}$  denotes the  $n$ th sample KTF in the data set of  $d_i$ . We show the normalized likelihoods with solid line in Fig. 9a and b respectively. When the passive subsystem is unobserved and has no priors, the likelihood curve with respect to FR is assumed to be uniform in the range of  $[0.485 \times 10^{-7}$  F/Km,  $1.085 \times 10^{-7}$  F/Km], shown as a solid straight line in Fig. 9c. After 10 iterations of inference by our modified NBP, the maximum marginal criteria yields the marginal posterior (MP) distributions. They are plotted in Fig. 9a, b and c with dash-dot lines for the spring (node 4), the damper (node 5) and the passive subsystem (node 2) respectively. Notice that the posteriors are all sharper than the normalized likelihoods even when the priors do not exist. This implies that the network has been correctly driven by the dependency defined in Eq. (2). To compare more easily, the actual FRs of the components and the deterministic FR of the subsystem are all indicated in those figures by dashed straight lines.

The comparison of those estimates is also reported in Table 3. The results are identical to the densities in Fig. 9. In this table MLE comes from normalized likelihood density, while the minimum mean square estimation (MMSE), maximum a posterior (MAP) and the quantiles are evaluated by using the posterior distribution. Notice that the MMSE results are much closer to the actual values than the MLE results. We also obtain the MMSE and the MAP of subsystem, which is not feasible for MLE. These confirm the effectiveness of our model. Furthermore, Fig. 9d presents the reliability of the subsystem given by  $R_2(t) = \exp[-x_2 t]$  with the mean of the marginal posterior (solid line) and 90% confidence intervals (dashed lines) respectively.

### 5.1.3. SRA of observed subsystem

In the following example, we kept the configuration of Section 5.1.2 unchanged except that the subsystem acquired 30 KTF data given as *Passive device* in Table 2. We use a Weibull distribution with two parameters for the subsystem node. The likelihood of the observed data with respect to FR can be estimated in the same way as done in Section 5.1.2. It is plotted in Fig. 10c by a surface in the 2D Weibull parameter space: *shape* and  $1/\text{scale}$  (for comparing the FR because  $\text{scale} = \text{FR}^{-1}$  when the *shape* is set at one). After 10 iterations of our modified NBP, we likewise obtain the FR's marginal posteriors (MP) of the components and the Weibull's marginal posterior surface of the subsystem in Fig. 10. Notice that the marginal posteriors in Fig. 10a and b are sharper than those of Section 5.1.2 because we have received more evidence for the network. In the same way, the posterior in Fig. 10d has less covariance than the normalized likelihood does in Fig. 10c. Moreover, MAP of the posterior is much closer to the actual value than the MLE, see Fig. 10d and c where the straight solid line indicates the actual FR ( $0.785 \times 10^{-7}$  F/Km). These results demonstrate that our model is validated by the inference.



**Fig. 9.** Normalized likelihood and marginal posterior (MP) of FR in the passive subsystem. In (a), (b) and (c) solid lines stand for the likelihoods, dash-dot lines for the posteriors and dashed lines for the actual values. In (d) the solid line represents the reliability of the subsystem by FR's posterior mean and the dashed lines for those of 5% and 95% quantiles of the posterior, indicating the 90% confidence intervals.

**Table 3**  
Comparison of FR estimations in passive unobserved subsystem ( $F/10^7$  km).

Item name	Actual	MLE	MMSE	MAP	Quantiles	
					5%	95%
$X_4$ , Spring	0.105	0.086	0.088	0.093	0.064	0.111
$X_5$ , Damper	0.68	0.704	0.703	0.679	0.519	0.920
$X_2$ , Passive dev.	0.785	—	0.780	0.741	0.561	0.999

The goodness-of-fit is discussed based on the statistic of Bayesian  $\chi^2$  presented in [13] where  $K = 4 \approx 30^{0.4}$  equal probability bins are suggested for our case. We draw repeatedly from the marginal posterior of Weibull parameter space, say  $(a, b)$ , and compute the corresponding cumulative probability of observations' likelihood,  $F_{Lik}(d_i|(a, b))$ . We decide which bin the draw should be assigned to, and calculate the Bayesian  $\chi^2$  statistic, then compare it against the 0.95 quantile of  $\chi^2_3$  ( $\chi^2_{3,0.95} \approx 7.82$ ). We observed that 13% of these values exceed the quantile, which indicates that the marginal posterior fits the subsystem's data observed reasonably well.

Notice that in this example the newly-observed data is successfully combined into the network and this can yield an up-to-date posterior estimation. In the same way, other types of new data, such as expert prior and historical data, can also be integrated to the network without difficulty. More information is fused, more-accurate estimation can be achieved. This can further

help to identify the system deterioration by sequential updating after replacing the past observation with the new one.

## 5.2. The entire mechatronic system

In this example, the new SRA model is tested on the entire system by the NBP and the modified NBP respectively. In this system only the root nodes are observed and given in Table 2 (all data except the passive subsystem). As a deterministic function, a series structure of exponential family still has an output of exponential distribution, but a parallel structure does not. Therefore we use a Weibull distribution for the system node and the exponential family for the subsystems. The actual FR of the nodes are given in Table 4, where the optimal deviation is leaned by Eq. (6) with some simulated observation by the density distance measure in Eq. (13). Note that a loop (formed by node 3, 6, 7, and 10) exists in the actuator subsystem. Hence the inference algorithm needs more iterations to converge. After 60 iterations in the modified NBP, we obtain the MMSE and MAP results of the system as well as those of the subsystems, shown in Table 4. Notice that the errors of MMSE results are all less than 10% even without any observation of those nodes. The further reliability assessment can be accomplished by using the corresponding TTF distribution model, as done in Fig. 9d.

We also show every-five-iterations the convergence process of the system's Weibull parameters in Fig. 11, where the points indicate the MMSE values by NBP and modified NBP respectively.

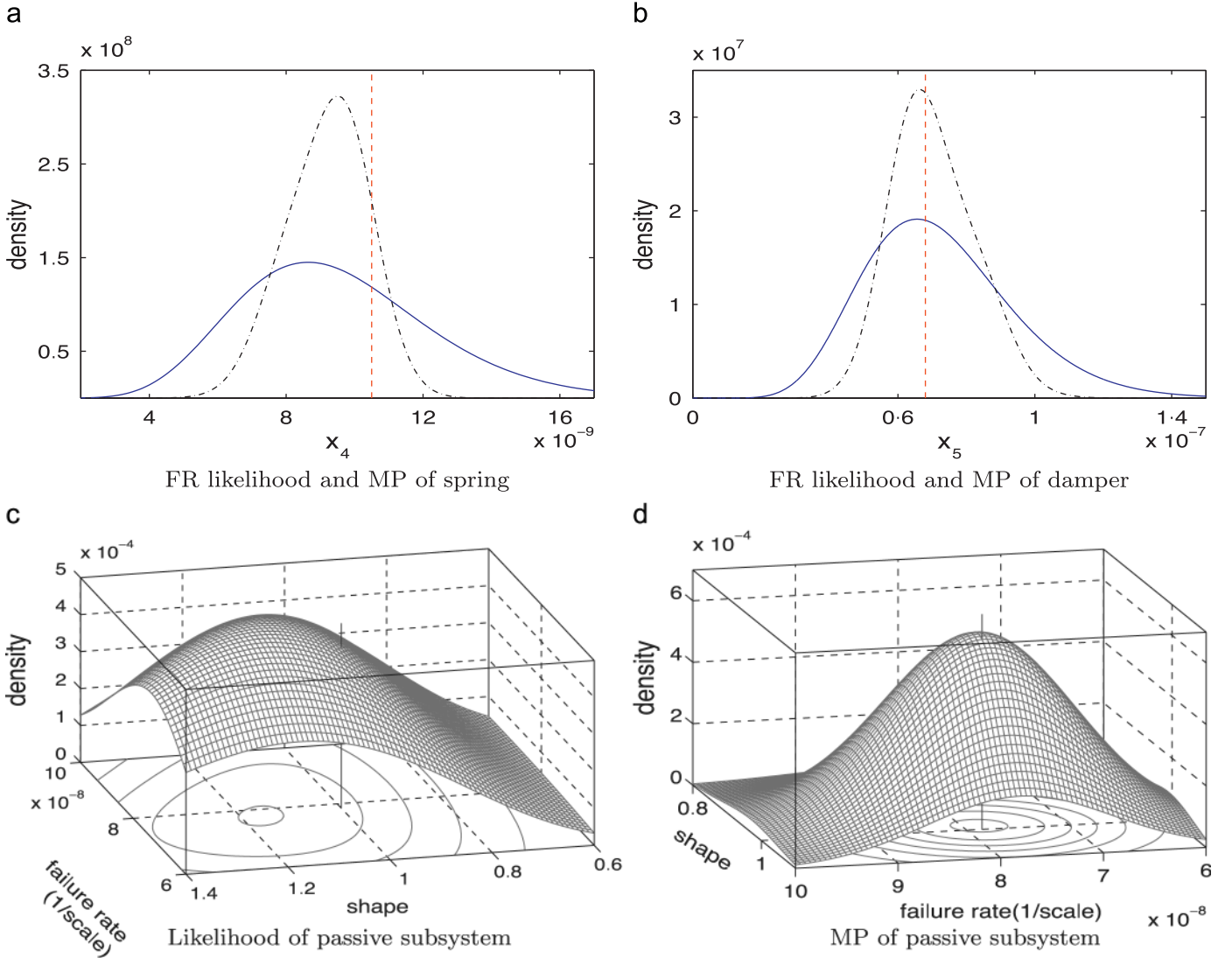


Fig. 10. Likelihood and marginal posterior (MP) of passive subsystem ( $X_2$ ) with respect to Weibull parameters (*shape* and *1/scale*) after 10 iterations of inference.

**Table 4**  
Comparison of FR estimation in the suspension system.

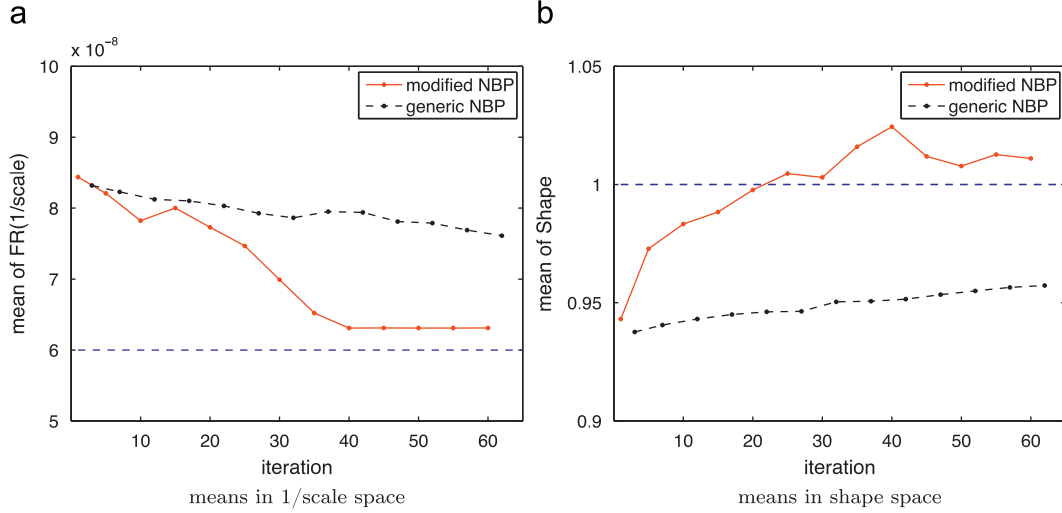
Item name	Actual FR	$\sigma^*$	MMSE	MAP
$X_1$ . Syst. (Weib.)	$0.6 \times 10^{-7}$	0.0890	(1.0110, $0.62 \times 10^{-7}$ )	(1.0201, $0.62 \times 10^{-7}$ )
$X_2$ . Passive	$0.65 \times 10^{-7}$	0.1407	$0.70 \times 10^{-7}$	$0.72 \times 10^{-7}$
$X_3$ . Actuator	$14.0 \times 10^{-7}$	0.0705	$13.5 \times 10^{-7}$	$12.9 \times 10^{-7}$
$X_6$ . Mech. dev.	$12.5 \times 10^{-7}$	0.0411	$13.3 \times 10^{-7}$	$14.7 \times 10^{-7}$
$X_7$ . Elec. dev.	$2.8 \times 10^{-7}$	0.0628	$3.0 \times 10^{-7}$	$3.1 \times 10^{-7}$

It is clear that our modification makes NBP converge much faster. All these results demonstrate again that our dependency model is effective by using the normal distribution in Eq. (2).

## 6. Conclusion

In this research, the system reliability assessment (SRA) has been formulated as a statistical inference problem of Bayesian networks (BNs). In contrast to conventional methods, we model the time-to-failure of the system/components by the parametric distributions whose parameters are considered as random

variables in the BN, such as failure rate of exponential family, shape and scale of 2D Weibull family and so on. Further, the parameter uncertainty of output function is modeled by a normal distribution whose standard deviation can be learned from samples. All of these uncertainties constitute the quantitative specification of the BN. To avoid the use of junction trees which are inefficient for the large-treewidth mechatronic systems, a modified nonparametric belief propagation (NBP) is developed to perform the inference in the BNs. For reasoning in a continuous BN, this provides an alternative solution to the other methods, such as mixture of truncated exponentials (MTE), dynamic discretization (DD) and Markov chain Monte Carlo (MCMC).



**Fig. 11.** Scale and shape means of system Weibull model in 60 iterations by the NBP and the modified NBP algorithm. The dashed lines denote the actual values.

Experimental results demonstrate the efficiency of our NBP modification and the effectiveness of our SRA method: the posterior fits the observed data well, and the uncertainty is effectively propagated through our BN based SRA model. Based on the parametric time-to-failure (TTF) representation, our BN model can perform further analysis, such as sensitivity analysis, diagnosis analysis. One future work is to consider model selection for unobservable items. It is another future work to apply our SRA method to the reliable active control issue in complex structural problems [44–46].

## Acknowledgment

The authors would like to thank China Scholarship Council (CSC) 2007153069 for the financial support of this research and gratefully acknowledge the anonymous referees for all their insightful comments and very helpful suggestions.

## Appendix A. Notation

$\mathcal{I}n$	label subset of original BN nodes
$\mathcal{I}c$	label subset of cluster nodes in pairwise MRF
$\mathcal{I}$	label set of all nodes in pairwise MRF, and $\mathcal{I} = \{\mathcal{I}n, \mathcal{I}c\}$
$X$	joint hidden variable
$D$	joint observation
$N_n$	total number of hidden variables in pairwise MRF
$X_i$	variable of the node $i$
$x_i$	a value of variable $X_i$
$d_i$	observation of the node $i$
$X_{i,j}$	corresponding component of variable $X_j$ in compound variable $X_i$
$\phi_i(X_i)$	potential of unitary clique in a MRF
$\psi_{ij}(X_i, X_j)$	potential of pairwise clique in a MRF
$\mathcal{B}(i)$	label set of node $i$ 's basic components
$\mathcal{P}a(i)$	label set of parent nodes of cluster node $i$ in MRF
$\Gamma(i)$	label set of neighbor nodes of node $i$
$f_i(X_i)$	TTF density of node $i$
$f_i^d$	deterministic TTF density of node $i$
$F_i(\cdot)$	cumulative probability distribution of node $i$
$\sigma$	dependency parameter

$C_n^k$	combination: the number of $k$ -combinations from the set with $n$ elements
$m_{ij}(X_j)$	message from node $i$ to node $j$
$x_{ij}^{(n)}$	the $n$ th kernel mean in message $m_{ij}(X_j)$
$x_{ij,r}^{(n)}$	corresponding component of variable $X_r$ in the $n$ th kernel mean of message $m_{ij}(X_j)$
$w_{ij}^{(n)}$	weight associated with $n$ th kernel mean in message $m_{ij}(X_j)$
$A_{ij}$	common bandwidth of message $m_{ij}(X_j)$

## References

- [1] Langseth H, Portinale L. Bayesian networks in reliability. *Reliability Engineering and System Safety* 2007;92(1):92–108.
- [2] Wilson AG, Huzurbazar AV. Bayesian networks for multilevel system reliability. *Reliability Engineering and System Safety* 2007;92(10):1413–20.
- [3] Mihalache AG. Modelisation et évaluation de la fiabilité des système mechatronics: application sur système embarqué. PhD thesis, Université d'Angers, 2007.
- [4] Zio E. Reliability engineering: old problems and new challenges. *Reliability Engineering and System Safety* 2009;94(2):125–41.
- [5] Rauzy A. New algorithms for fault trees analysis. *Reliability Engineering & Systems Safety* 1993;40(3):203–11.
- [6] Bennetts RG. Analysis of reliability block diagrams by boolean techniques. *IEEE Transactions on Reliability* 1982;31(2):159–66.
- [7] Dutuit Y, Chatelet E, Signoret JP, Thomas P. Dependability modelling and evaluation by using stochastic Petri nets: application to two test cases. *Reliability Engineering and System Safety* 1997;55(2):117–24.
- [8] Zhao YG, Ang AHS. System reliability assessment by method of moments. *Journal of Structural Engineering* 2003;129(10):1341–9.
- [9] Whittaker JA, Thomason MG. A Markov chain model for statistical software testing. *IEEE Transactions on Software Engineering* 1994;20(10):812–24.
- [10] Koutras MV. On a Markov chain approach for the study of reliability structures. *Journal of Applied Probability* 1996;33(2):357–67.
- [11] Dugan JB, Bavuso SJ, Boyd MA. Dynamic fault-tree models for fault-tolerant computer systems. *IEEE Transactions on Reliability* 1992;41(3):363–77.
- [12] Čepin M, Mavko B. A dynamic fault tree. *Reliability Engineering and System Safety* 2002;75(1):83–91.
- [13] Hamada MS, Wilson AG, Reese CS, Harry FM. Bayesian reliability. Springer; 2008. Series in Statistics.
- [14] Wilson AG, McNamara LA, Wilson GD. Information integration for complex systems. *Reliability Engineering and System Safety* 2007;92(1):121–30.
- [15] Bobbio A, Portinale L, Minichino M, Ciancamerla E. Improving the analysis of dependable systems by mapping fault trees into Bayesian networks. *Reliability Engineering and System Safety* 2001;71(3):249–60.
- [16] Doguc O, Ramirez-Marquez JE. A generic method for estimating system reliability using Bayesian networks. *Reliability Engineering and System Safety* 2009;94(2):542–50.
- [17] Torres-Toledano JG, Sucar LE. Bayesian networks for reliability analysis of complex systems. *Lecture notes in artificial intelligence*, vol. 1484, 1998. p. 195–206.

- [18] Mahadevan S, Zhang R, Smith N. Bayesian networks for system reliability reassessment. *Structural Safety* 2001;23(3):231–51.
- [19] Weber P, Jouffe L. Complex system reliability modelling with dynamic object oriented Bayesian networks (DOOBN). *Reliability Engineering and System Safety* 2006;91(2):149–62.
- [20] Boudali H, Dugan JB. A discrete-time Bayesian network reliability modeling and analysis framework. *Reliability Engineering and System Safety* 2005; 87(3):337–49.
- [21] Boudali H, Dugan JB. A continuous-time Bayesian network reliability modeling, and analysis framework. *IEEE Transactions on Reliability* 2006; 55(1):86–97.
- [22] Johnson VE, Graves TL, Hamada M, Reese CS. A hierarchical model for estimating the reliability of complex systems. In: *Bayesian statistics 7: proceedings of the seventh Valencia international meeting, 2003*. p. 199–213.
- [23] Anderson-Cook CM, Graves T, Hamada M, Hengartner N, Johnson V, Reese CS, et al. Bayesian stockpile reliability methodology for complex systems with application to a missile stockpile. Technical report, Los Alamos National Laboratory LA-UR-05-1742, 2005.
- [24] Neil M, Taylor M, Marquez D, Fenton N, Hearty P. Modelling dependable systems using hybrid Bayesian networks. *Reliability Engineering and System Safety* 2008;93(7):933–9.
- [25] Marquez D, Neil M, Fenton N. A new Bayesian network approach to reliability modelling. In: *Proceedings of mathematical methods in reliability (MMR07)*, Glasgow, July 2007.
- [26] Zafropoulos EP, Dialynas EN. Reliability and cost optimization of electronic devices considering the component failure rate uncertainty. *Reliability Engineering and System Safety* 2004;84(3):271–84.
- [27] Cohen AC. Maximum likelihood estimation in the Weibull distribution based on complete and on censored samples. *Technometrics* 1965;7(4):579–88.
- [28] Pearl J, Shafer G. Probabilistic reasoning in intelligent systems: networks of plausible inference. CA: Morgan Kaufmann San Mateo; 1988.
- [29] Heckerman D. A tutorial on learning with Bayesian networks. Technical report, Microsoft Research, MSR-TR-95-06, March 1995.
- [30] Langseth H, Nielsen TD, Rumí R, Salmerón A. Inference in hybrid Bayesian networks. *Reliability Engineering and System Safety* 2009;94(10):1499–509.
- [31] Moral S, Rumpf R, Salmerón A. Mixtures of truncated exponentials in hybrid Bayesian networks. In: *Proceedings of symbolic and quantitative approaches to reasoning with uncertainty: 6th European conference, ECSQARU 2001*. Springer Verlag; 2001. p. 156–67.
- [32] Neil M, Taylor M, Marquez D. Inference in hybrid Bayesian networks using dynamic discretization. *Statistics and Computing* 2007;17(3):219–33.
- [33] Sudderth EB, Ihler AT, Freeman WT, Willsky AS. Nonparametric belief propagation. In: *IEEE conference on computer vision and pattern recognition*, vol. 1, 2003. p. 605–12.
- [34] <<http://www.ics.uci.edu/~ihler/code/>>.
- [35] Marquez D, Neil M, Fenton N. Improved reliability modeling using Bayesian networks and dynamic discretization. *Reliability Engineering and System Safety* 2010;95(4):412–25.
- [36] Cha SH. Comprehensive survey on distance/similarity measures between probability density functions. *International Journal of Mathematical Models and Methods in Applied Sciences* 2007;1(4):300–7.
- [37] Weiss Y, Freeman WT. On the optimality of solutions of the max-product belief-propagation algorithm in arbitrary graphs. *IEEE Transactions on Information Theory* 2001;47(2):736–44.
- [38] MacKay DJC. *Information theory, inference, and learning algorithms*. Cambridge University Press; 2003.
- [39] Murphy K, Weiss Y, Jordan MI. Loopy belief propagation for approximate inference: an empirical study. In: *Proceedings of uncertainty in AI, Citeseer*, 1999. p. 467–75.
- [40] Casella G, George EI. Explaining the Gibbs sampler. *American Statistician* 1992;46(3):167–74.
- [41] Silverman BW. *Density estimation for statistics and data analysis*. Chapman Hall/CRC; 1986.
- [42] Ripley BD. *Stochastic simulation*. New York: Wiley; 1987.
- [43] Lyonnet P. *Ingénierie de la fiabilité*, Lavoisier Paris, 2006.
- [44] Ichchou MN, Jemai B, Bellon L, Jezequel L. Active rubber mounts by means of piezoelectric actuators, experimental work. *Smart Materials and Structures* 2001;10(5):1095–9.
- [45] Jemai B, Ichchou MN, Jezequel L, Noe M. An assembled plate active control damping set-up: optimization and control. *Journal of Sound and Vibration* 1999;225(2):327–43.
- [46] Jemai B, Ichchou MN, Jezequel L. Double panel partition AVNC by means of optimized piezoceramic structural boundary control. *Journal of Vibration and Control* 2002;8(5):595–617.