Remaining Useful Life Prediction for Complex Systems With Multiple Indicators Based on Particle Filter and Parameter Correlation

SHAOWEI CHEN [1], (Member, IEEE), **MEINAN WANG**[1], **DENGSHAN HUANG**[1],
PENGFEI WEN[1], (Graduate Student Member, IEEE), **SHENGYUE WANG**[1],
AND SHUAI ZHAO [2], (Member, IEEE)

[1]School of Electronics and Information, Northwestern Polytechnical University, Xi'an 710072, China
[2]Department of Energy Technology, Aalborg University, 9220 Aalborg, Denmark

Corresponding author: Shuai Zhao (szh@et.aau.dk)

ABSTRACT In practical applications, the failure of large-scale complex equipment is often caused by the simultaneous degradation of multiple components. It is necessary to predict the remaining useful life (RUL) of the equipment with multiple degradation indicators. This article proposes a new joint-RUL-prediction method in the presence of multiple degradation indicators based on parameter correlation. The stochastic process model is established for each degradation indicator, and the model parameters are estimated by kernel smoothing particle filter (KS-PF) and maximum likelihood estimation (MLE). Meanwhile, to facilitate the dependencies between multiple degradation indicators, correlations of the degradation model parameter between multiple degradation indicators are established in KS-PF. In addition, optimal tuning (OT) is introduced to choose the best kernel parameter. A case study on the Commercial Modular Aero-Propulsion System Simulation (C-MAPSS) dataset is applied to verify the proposed method, the experiment shows that the proposed joint-RUL-prediction method based on parameter correlation possesses a superior prediction performance compared with that by using a single degradation indicator.

INDEX TERMS RUL prediction, particle filter, multiple degradation indicators, parameter correlation, kernel smoothing.

I. INTRODUCTION

As an essential part of prognostics and health management (PHM), remaining useful life (RUL) prediction can effectively help to improve the reliability of equipment, avoid the huge loss caused by equipment failure, and reduce the cost of equipment maintenance. The failure of large-scale complex equipment is often caused by the simultaneous degradation of its multiple components [1], [2]. The multiple degradations of different components are identified as multiple degradation indicators, and the RUL prediction based on multiple degradation indicators is expected to capture the health status of equipment effectively.

The RUL prediction method can be classified into physics model-based methods, empirical model-based methods, artificial intelligent (AI) methods, and hybrid methods [3]. The empirical model-based methods are the most popular one

The associate editor coordinating the review of this manuscript and approving it for publication was Yu Liu [id].

among the four categories and have been widely used in recent years. When the empirical model-based methods are used for RUL prediction, an empirical model describing the degradation process of equipment is established first, and then the model parameters and degradation states are updated using the degradation measurements, consequently deriving the RUL. Stochastic process-based models such as Wiener process, Gamma process, and Markov chain-based method are commonly used [4], [5] as the degradation models and they are more interpretable compared with other models [6]–[9]. The Bayesian filter methods such as Kalman filter (KF) and particle filter (PF) are widely used for model updating in recent years [10]–[17]. However, most of the existing empirical model-based methods only involve a single degradation indicator, which is insufficient for characterizing the complex degradation process of equipment in most cases. Consequently, the empirical model-based RUL prediction involving multiple degradation indicators should be developed necessarily for practical applications.

For equipment with multiple performance characteristics, or the systems which have multiple components, the multiple performance characteristics or components usually degrade simultaneously when the equipment or system degenerate, and the multiple indicators can be identified to indicate the degradation degree of different performance characteristics or components. And there may be dependencies between these indicators since they contain the information of the same degenerate equipment or system. Here the failure definition of the series systems is considered in this article, i.e., for equipment with multiple degradation indicators, the first hitting time when any degradation indicator exceeds the corresponding failure threshold is regarded as the End-of-Life (EoL) [1], [18], [19]. Based on this idea, this article intends to use the stochastic process model and PF to realize the RUL prediction of equipment with multiple degradation indicators. A multivariate degradation model needs to be established first, and the challenging part is how to characterize the dependencies between the multiple degradation indicators reasonably.

The dependencies are established on the stochastic term of the degradation process in some literature [1], [4], [18]–[20]. Peng *et al.* [1] established multivariate degradation models with copula functions and multiple stochastic processes, where the copula functions were used to establish the dependencies on the increments of degradation states. Rodriguez-Picon *et al.* [20] established different multivariate degradation models considering copula functions and different combinations of stochastic processes as marginal distributions, the dependencies were established on the degradation states, and then the product reliability under heterogeneous models was studied on the crack propagation data. Fang *et al.* [18] proposed a bivariate stochastic process model, with the dependencies establishing on the degradation states, and the Bayesian method was used for parameter estimation. Xi *et al.* [19] characterize the dependencies among different degradation indicators using a diffusion coefficient matrix, and a multivariate degradation model for multiple indicators was established. Then the Monte Carlo simulation method was used to predict the RUL. In addition to the above methods in which the dependencies are established on the multivariate degradation models for multiple indicators, Sun *et al.* [21] directly established the dependencies on multiple RUL probability distributions estimated in terms of different degradation indicators.

Since the model parameters can affect the long-term degradation trend of indicators, we consider establishing the dependencies on the model parameters in this article, and the long-term dependencies among different indicators are expected to be characterized accordingly. With the establishment of correlations of model parameters, the degradation states of indicators which are estimated using model parameters can capture the long term dependencies among different indicators effectively. Therefore, a joint-RUL-prediction method based on the stochastic process model and PF is proposed in this article, characterizing the dependencies between

multiple degradation indicators as the correlations between parameters of the models built for the corresponding indicators. The long-term dependencies among different indicators are taken into consideration using the correlation between parameters, and the multivariate degradation modeling in the proposed method can consequently provide accurate information of dependencies for RUL prediction.

In the proposed method, parameters in the degradation model can be divided into 1) the parameters of degradation trajectory and 2) the fluctuation parameters, according to their physical implications. Maximum likelihood estimation (MLE) is used on the training units to estimate the fluctuation parameters offline. Then kernel smoothing particle filter (KS-PF) [22], [23] is used to estimate the degradation trajectory parameters as well as the degradation states simultaneously online. In addition, optimal tuning (OT) [24] is adopted in KS-PF, i.e., OTKS-PF to obtain the optimal kernel parameters at each monitoring time when the degradation state and the parameters are estimated. To characterize the dependencies between multiple degradation indicators, the correlations between the model parameters are introduced to generate the degradation trajectory parameters in OTKS-PF. Then the joint probability distribution of the multidimensional degradation state of indicators is estimated, and the RUL can be consequently predicted with the pre-specified failure thresholds. The respective RUL predicted based on a single degradation indicator and multiple degradation indicators are comprehensively compared by using the Commercial Modular Aero-Propulsion System Simulation (C-MAPSS) dataset for the illustration of the effectiveness and superiority of the proposed method.

The remainder of this article is organized as follows. Section II describes the establishment of the marginal degradation models for each degradation indicator and the estimation method for fluctuation parameters. OTKS-PF is introduced in section III, to simultaneously estimate the degradation states and model parameters on a single degradation indicator. The joint-RUL-prediction method based on parameter correlation with multiple degradation indicators is proposed in section IV. Section V performs a case study on the C-MAPSS dataset, and the RUL predicted based on a single degradation indicator and multiple degradation indicators respectively are compared. Section VI gives a conclusion.

II. DEGRADATION MODELING AND FLUCTUATION PARAMETERS ESTIMATION

A. MARGINAL DEGRADATION MODEL

The uncertainty of the degradation processes is mainly caused by four sources of variability, i.e., the temporal variability, the unit-to-unit variability, the nonlinear variability, and the measurement variability [8]. The temporal variability denotes the inherent uncertainty of the degradation process. The unit-to-unit variability represents the heterogeneity of degradation processes of different units. The nonlinear variability means that the degradation rate can change over time.

The measurement variability represents the uncertainty caused by measurement error. Therefore, the following stochastic process models are established for each marginal degradation process:

$$X_k(t) = \mu_k(\boldsymbol{\theta}_k, t) + \sigma_k B(t), \tag{1}$$

$$Y_k(t) = X_k(t) + \varepsilon_k, \varepsilon_k \sim \mathcal{N}\left(0, v_k^2\right), \tag{2}$$

where $k = 1, 2, \ldots, K$ represents the kth degradation indicator. $X_k(t)$ is the degradation state of the kth degradation indicator at time t. $\mu_k(\boldsymbol{\theta}_k, t)$ is the mean function of the stochastic process $X_k(t)$, and $\sigma_k B(t)$ is the stochastic term of $X_k(t)$. $\mu_k(\boldsymbol{\theta}_k, t)$ is a differentiable and monotonical function which is defined as the degradation trajectory, and it may be nonlinear. $\boldsymbol{\theta}_k = [a_k, b_k, \cdots]^T$ is a vector of the degradation trajectory parameters, which determines the degradation rate of different units. Therefore, $\mu_k(\boldsymbol{\theta}_k, t)$ describes the nonlinear variability of the degradation process, and $\boldsymbol{\theta}_k$ characterizes the unit-to-unit variability by specifying the degradation trajectories of different equipment. $B(t)$ is a standard Brownian motion process and σ_k is the volatility parameter. $\sigma_k B(t)$ has normal increments, i.e., $\sigma_k(B(t_1) - B(t_2)) \sim \mathcal{N}\left(0, \sigma_k^2(t_1 - t_2)\right)$ (for $0 \le t_2 \le t_1$), and it describes the temporal variability by representing the inherent uncertainty of degradation process (1). $Y_k(t)$ is the measurement of the kth degradation indicator. ε_k is a zero-mean Gaussian random variable with the variance v_k^2, which can be seen as the measurement noise representing the measurement variability. Then the marginal degradation models are established for each degradation indicator and the models can characterize all the categories of uncertainty that may exist in equipment degradation processes.

It is considered that σ_k^2 and v_k^2 are only related to the operating conditions and measuring conditions of equipment. In other words, σ_k^2 and v_k^2 are the same and $\boldsymbol{\theta}_k$ is different for different equipment which runs under a single operating condition. Therefore, σ_k^2 and v_k^2 are defined as fluctuation parameters which are different from $\boldsymbol{\theta}_k$, and they are estimated using different methods in this article.

Let $x_{k,j} = X_k(t_j)$, $y_{k,j} = Y_k(t_j)$ and $\Delta t = t_j - t_{j-1}$, where $0 \le j \le M$, and M is the number of measurements. Considering the kth degradation indicator, the marginal models (1) and (2) can be transformed into the state transition equation combined with the measurement equation for PF:

State transition equation $f(x_{k,j} \mid x_{k,j-1})$:

$$\begin{cases} x_{k,j} = x_{k,j-1} + \Delta\mu_k(t_j) + \omega_{k,j}, \\ \Delta\mu_k(t_j) = \mu_k(\boldsymbol{\theta}_k, t_j) - \mu_k(\boldsymbol{\theta}_k, t_{j-1}), \\ \omega_{k,j} = \sigma_k(B(t_j) - B(t_{j-1})) \sim \mathcal{N}\left(0, \sigma_k^2\Delta t\right). \end{cases} \tag{3}$$

Measurement equation $h(y_{k,j} \mid x_{k,j})$:

$$\begin{cases} y_{k,j} = x_{k,j} + \varepsilon_k, \\ \varepsilon_k \sim \mathcal{N}\left(0, v_k^2\right). \end{cases} \tag{4}$$

When (3) and (4) are used on multiple degradation indicators, let $\mathbf{x}_j = [x_{1,j}, \cdots, x_{K,j}]^T$ denote the states of all

degradation indicators at time t_j, and $\mathbf{y}_j = [y_{1,j}, \cdots, y_{K,j}]^T$ denote the measurements. Then the state transition equation and the measurement equation are converted to $f(\mathbf{x}_j \mid \mathbf{x}_{j-1})$ and $h(\mathbf{y}_j \mid \mathbf{x}_j)$ respectively. For equipment with K degradation indicators, if the kth degradation indicator $X_k(t)$ exceeds its failure threshold H_k first, and that time is regarded as the failure time. The estimated RUL at time t_j is defined as

$$\begin{aligned} \text{RUL} = \inf\{\tau : X_1(t_j + \tau) \ge H_1 \text{ or } \cdots \\ \text{or } X_K(t_j + \tau) \ge H_K \mid x_{1,1:j}, \cdots, x_{K,1:j}\}. \end{aligned} \tag{5}$$

B. ESTIMATION OF FLUCTUATION PARAMETERS

The parameters $\boldsymbol{\theta}_k$, σ_k^2, and v_k^2 in (3) and (4) for each degradation indicator need to be estimated. For parameter estimation, the fluctuation parameters σ_k^2 and v_k^2 are estimated offline using MLE based on the training units. Then PF is used to estimate $\boldsymbol{\theta}_k$ and the degradation state $X_k(t)$ online simultaneously for each testing unit. The RUL prediction result based on the kth degradation indicator can be obtained after the two-step parameter estimation. The estimation of the fluctuation parameters is described in detail as follows.

Here the superscript '$*$' is used to represent the variables for the training units. Let $\mathbf{X}_k^* = [x_{k,1}^*, \cdots, x_{k,M}^*]^T$, $\mathbf{Y}_k^* = [y_{k,1}^*, \cdots, y_{k,M}^*]^T$, and $\mathbf{u}_k^* = [\mu_k(\boldsymbol{\theta}_k^*, t_1), \cdots, \mu_k(\boldsymbol{\theta}_k^*, t_M)]^T$ represent the states, measurements, and degradation trajectories of the kth degradation indicator respectively for training units. Then the distributions of $\mathbf{X}_k^*$ and $\mathbf{Y}_k^*$ can be derived by using the properties of standard Brownian motion:

$$\mathbf{X}_k^* \sim \mathcal{N}\left(\mathbf{u}_k^*, \sigma_k^2\mathbf{Q}\right), \tag{6}$$

$$\mathbf{Y}_k^* \sim \mathcal{N}\left(\mathbf{X}_k^*, v_k^2\boldsymbol{I}_M\right), \tag{7}$$

where $\mathbf{Q} = [\min\{t_i, t_j\}]_{1 \le i,j \le M}$ and $\boldsymbol{I}_M$ is an identity matrix of order M.

The log-likelihood functions of the fluctuation parameters are expressed as follows:

$$\begin{aligned} L\left(\sigma_k^2 \mid \mathbf{X}_k^*, \boldsymbol{\theta}_k^*\right) \\ = -\frac{M}{2}\ln(2\pi) - \frac{1}{2}\ln\left|\sigma_k^2\mathbf{Q}\right| \\ - \frac{1}{2}\left(\mathbf{X}_k^* - \mathbf{u}_k^*\right)^T\left(\sigma_k^2\mathbf{Q}\right)^{-1}\left(\mathbf{X}_k^* - \mathbf{u}_k^*\right), \end{aligned} \tag{8}$$

and

$$\begin{aligned} L\left(v_k^2 \mid \mathbf{Y}_k^*, \mathbf{X}_k^*\right) \\ = -\frac{M}{2}\ln(2\pi) - \frac{1}{2}\ln\left|v_k^2\boldsymbol{I}_M\right| \\ - \frac{1}{2}\left(\mathbf{Y}_k^* - \mathbf{X}_k^*\right)^T\left(v_k^2\boldsymbol{I}_M\right)^{-1}\left(\mathbf{Y}_k^* - \mathbf{X}_k^*\right). \end{aligned} \tag{9}$$

The measurements in $\mathbf{Y}_k^*$ of each training unit are collected, while the degradation states in $\mathbf{X}_k^*$ are usually implicit. In order to estimate the fluctuation parameters, the degradation state $\mathbf{X}_k^*$ can be distinguished through some appropriate methods first, such as the simple moving average (SMA), the

weighted moving average (WMA), exponential moving average (EMA), etc., and then v_k^2 can be estimated by maximizing the log-likelihood function (9). After a proper functional expression $\mu_k(\cdot)$ is used to characterize the degradation trajectory, the least-squares method is used to fit the degradation state $\mathbf{X}_k^*$ of training units to obtain the estimated $\boldsymbol{\theta}_k^*$ as well as the degradation trajectory. σ_k^2 can be finally estimated by maximizing (8).

III. SIMULTANEOUS ESTIMATION OF STATE AND PARAMETER WITH OTKS

To present the proposed method clearly, the RUL prediction with a single degradation indicator is discussed in this section first, and OTKS-PF is introduced. When the kth degradation indicator is used for RUL prediction, after the fluctuation parameters σ_k^2 and v_k^2 are estimated, $\boldsymbol{\theta}_k$ and $X_k(t)$ are jointly updated according to the new measurements using KS-PF. Moreover, the kernel parameter that is used in KS-PF is optimized online by introducing OT. Then the RUL can be predicted based on the estimated parameters and states.

A. PARTICLE FILTER

PF is a Bayesian method for state estimation. Let x_j be the state of equipment at time t_j that needs to be estimated in PF, and y_j be the measurement at time t_j. After establishing the state transition equation and the measurement equation for equipment, the posterior probability $p(x_j \mid y_{1:j})$ can be obtained through two steps: prediction and update. For prediction,

$$p(x_j \mid y_{1:j-1}) = \int p(x_j \mid x_{j-1}) p(x_{j-1} \mid y_{1:j-1}) \, dx_{j-1}. \tag{10}$$

and for update,

$$p(x_j \mid y_{1:j}) = \frac{p(y_j \mid x_j) p(x_j \mid y_{1:j-1})}{p(y_j \mid y_{1:j-1})}, \tag{11}$$

where the normalizing constant is

$$p(y_j \mid y_{1:j-1}) = \int p(y_j \mid x_j) p(x_j \mid y_{1:j-1}) \, dx_j. \tag{12}$$

Since the integral needs to be calculated in (10) and (12), it is difficult to implement the above optimal Bayesian filtering algorithms for nonlinear and non-gaussian systems. To solve the problem, the Monte Carlo sampling is introduced into PF to represent the probability distribution by particles that sampled from it, so the posterior probability of x_j can be estimated as a set of weighted particles (samples) $\{x_j^i, w_j^i\}_{i=1}^N$, where $i = 1, \cdots N$ and N is the number of particles. x_j^i is the ith particle of x_j and w_j^i is the weight associated with it. The posterior probability distribution of x_j is estimated as $p(x_j \mid y_{1:j}) \approx \sum_{i=1}^N w_j^i \delta(x_j - x_j^i)$, where $\delta(\cdot)$ is the Dirac delta function.

B. KERNEL SMOOTHING

The degradation state $x_{k,j}$ and the degradation trajectory parameter $\boldsymbol{\theta}_{k,j}$ can be jointly updated according to the measurement $y_{k,j}$, where $\boldsymbol{\theta}_{k,j} = [a_{k,j}, b_{k,j}, \cdots]^T$ are the estimated parameters of $\boldsymbol{\theta}_k$ at time t_j. Here the most commonly used approach is the augment PF, in which the parameter $\boldsymbol{\theta}_{k,j}$ and the system state $x_{k,j}$ are all considered as the hidden states which are estimated in PF, and the artificial evolution is introduced to realize the evolution of parameters in the particles. However, this approach will increase the variance of particles, and the convergence of PF will be delayed [22].

In this article, kernel smoothing (KS) is introduced into PF to solve the problem. The evolution of parameters in the particles $\boldsymbol{\theta}_{k,j}^i \mid \boldsymbol{\theta}_{k,j-1}^i$ is achieved by two steps: shrinkage and perturbation. For the shrinkage,

$$\tilde{\boldsymbol{\theta}}_{k,j}^i = \sqrt{1 - s^2} \boldsymbol{\theta}_{k,j-1}^i + \left(1 - \sqrt{1 - s^2}\right) \bar{\boldsymbol{\theta}}_{k,j-1}, \tag{13}$$

and for the perturbation,

$$\boldsymbol{\theta}_{k,j}^i = \tilde{\boldsymbol{\theta}}_{k,j}^i + \boldsymbol{\xi}, \boldsymbol{\xi} \sim \mathcal{N}\left(0, s^2 \boldsymbol{V}_{j-1}^k\right), \tag{14}$$

where $\boldsymbol{\theta}_{k,j}^i = [a_{k,j}^i, b_{k,j}^i, \cdots]^T$ is the ith particle of $\boldsymbol{\theta}_k$ at time t_j, and $\tilde{\boldsymbol{\theta}}_{k,j}^i$ is the shrinkage of $\boldsymbol{\theta}_{k,j}^i$. $\{\boldsymbol{\theta}_{k,j-1}^i, w_{k,j-1}^i\}_{i=1}^N$ is the particle set of $\boldsymbol{\theta}_k$ with the particle weights at time t_{j-1}. $\bar{\boldsymbol{\theta}}_{k,j-1}$ and $\boldsymbol{V}_{j-1}^k$ are the mean and variance of $\boldsymbol{\theta}_{k,j-1}$ based on $\{\boldsymbol{\theta}_{k,j-1}^i, w_{k,j-1}^i\}_{i=1}^N$, where $\bar{\boldsymbol{\theta}}_{k,j-1} = [\bar{a}_{k,j-1}, \bar{b}_{k,j-1}, \cdots]^T$, and $\boldsymbol{V}_{j-1}^k = \text{diag}(V_{a,j-1}^k, V_{b,j-1}^k, \cdots)$ is a diagonal matrix in which the entries outside the main diagonal are all zero. $s \in [0, 1]$ is the kernel parameter.

After the shrinkage, each particle of parameters shrinks towards the mean of the parameters' particles at the previous moment. The larger the kernel parameter s is, the greater the volume of the shrinkage is. When $s = 1$, each particle of parameters will shrink to the mean and all the particles will be the same, so the variance of the particles will decrease due to the shrinkage. Then the artificial random walk is added to ensure the diversity of particles in the perturbation step. Finally, the variance of particles keeps unchanged and the diversity of particles is also kept after the shrinkage and perturbation. Let $q = \sqrt{1 - s^2}$, and the state transition equation (3) then can be deduced as

$$\begin{cases} x_{k,j} = x_{k,j-1} + \Delta\mu_k(t_j) + \omega_{k,j}, \\ \boldsymbol{\theta}_{k,j}^i \mid \boldsymbol{\theta}_{k,j-1}^i \\ \sim \mathcal{N}\left(q\boldsymbol{\theta}_{k,j-1}^i + (1-q)\bar{\boldsymbol{\theta}}_{k,j-1}, (1-q^2)\boldsymbol{V}_{j-1}^k\right), \\ \Delta\mu_k(t_j) = \mu_k(\boldsymbol{\theta}_{k,j}, t_j) - \mu_k(\boldsymbol{\theta}_{k,j-1}, t_{j-1}), \\ \omega_{k,j} \sim \mathcal{N}\left(0, \sigma_k^2 \Delta t\right). \end{cases} \tag{15}$$

Using the sequential importance resampling (SIR) framework with the state transition equation (15) and the measurement equation (4), $X_k(t)$ and $\boldsymbol{\theta}_k$ are jointly estimated online by the following steps:

1. Initialization:

Let $j = 0$, and the initial distributions are set as $p(x_{k,0})$ and $p(\boldsymbol{\theta}_{k,0})$.

2. State estimation:

Let $j = j + 1$. If the measurement $y_{k,j}$ exists, perform the following steps, otherwise terminate this procedure.

1) Prediction: the particles $\{x_{k,j|j-1}^i, \boldsymbol{\theta}_{k,j|j-1}^i, w_{k,j|j-1}^i\}_{i=1}^N$ are sampled from the proposal distributions $p(x_{k,j}|x_{k,j-1})$ and $p(\boldsymbol{\theta}_{k,j}|\boldsymbol{\theta}_{k,j-1})$, and the prior probability of $x_{k,j}$ and $\boldsymbol{\theta}_{k,j}$ can be represented by the weighted particles.

2) Update: the weights of the particles are updated according to the measurement $y_{k,j}$ at time t_j:

$$\widetilde{w}_{k,j}^i = w_{k,j|j-1}^i \cdot p\left(y_{k,j} \mid x_{k,j|j-1}^i\right), \tag{16}$$

and the weights are normalized as

$$\hat{w}_{k,j}^i = \frac{\widetilde{w}_{k,j}^i}{\sum_{i=1}^N \widetilde{w}_{k,j}^i}. \tag{17}$$

3) Resampling: the number of effective particles is approximated as $N_{\text{neff}} \approx 1/\sum_{i=1}^N \left(\hat{w}_{k,j}^i\right)^2$. If N_{neff} is less than a given threshold, the resample method is implemented, otherwise resampling is not performed. After the particle resampling, the updated particles are obtained as $\{x_{k,j}^i, \boldsymbol{\theta}_{k,j}^i, w_{k,j}^i\}_{i=1}^N$, and the posterior probability distribution of $x_{k,j}$ and $\boldsymbol{\theta}_{k,j}$ can be consequently represented by the weighted particles.

4) Iteratively perform Steps 2-3.

C. OPTIMAL TUNING

The value of the kernel parameter s affects the scale of shrinkage and perturbation of the particles, so it is important to select an appropriate value for s. s is usually set as a constant in the present literature [1], [22]. Instead, we adopt OT here, to select the optimal kernel parameter s_j at each monitoring time t_j.

The estimation accuracy of PF is related to the proposal distribution. The optimal proposal distribution should minimize the variance of the updated weights of particles, and it can also be interpreted that a good proposal distribution should make the prior distribution $p\left(x_j|y_{1:j-1}\right)$ as similar as possible to the posterior distribution $p\left(x_j|y_{1:j}\right)$ [24], [25]. KL divergence is an index to measure the matching degree of two probability distributions, so the optimal s_j is chosen in OT to make the KL divergence between the prior distribution and the posterior distribution minimal [24].

According to PF, the prior distribution of the state of the kth degradation indicator at time t_j is estimated as

$$\begin{aligned} Q_{k,j} &= p\left(x_{k,j} \mid y_{k,1:j-1}\right) \\ &\approx \sum_{i=1}^N w_{k,j|j-1}^i \delta\left(x_{k,j} - x_{k,j|j-1}^i\right), \end{aligned} \tag{18}$$

and the posterior distribution is

$$P_{k,j} = p\left(x_{k,j} \mid y_{k,1:j}\right) \approx \sum_{i=1}^N w_{k,j}^i \delta\left(x_{k,j} - x_{k,j}^i\right). \tag{19}$$

The KL divergence between the prior distribution and the posterior distribution is

$$\begin{aligned} KL\left(Q_{k,j} \parallel P_{k,j}\right) &= \int p\left(x_{k,j} \mid y_{k,1:j-1}\right) \\ &\times \log \frac{p\left(x_{k,j} \mid y_{k,1:j-1}\right)}{p\left(x_{k,j} \mid y_{k,1:j}\right)} dx_{k,j}. \end{aligned} \tag{20}$$

We can directly derive to (21) by substituting (11), (12), and (18) into (20) using the sifting property of the Dirac delta function.

$$\begin{aligned} KL\left(Q_{k,j} \parallel P_{k,j}\right) &= \sum_{i=1}^N w_{k,j|j-1}^i \\ &\times \log \frac{\sum_{i=1}^N w_{k,j|j-1}^i p\left(y_{k,j}|x_{k,j|j-1}^i\right)}{p\left(y_{k,j}|x_{k,j|j-1}^i\right)}. \end{aligned} \tag{21}$$

Substitute (16) into (17), and the following equation can be derived:

$$\frac{\sum_{i=1}^N w_{k,j|j-1}^i p\left(y_{k,j}|x_{k,j|j-1}^i\right)}{p\left(y_{k,j}|x_{k,j|j-1}^i\right)} = \frac{w_{k,j|j-1}^i}{\hat{w}_{k,j}^i}, \tag{22}$$

therefore, substitute (22) into as (21) as

$$KL\left(Q_{k,j} \parallel P_{k,j}\right) = -\sum_{i=1}^N w_{k,j|j-1}^i \log\left(\frac{\hat{w}_{k,j}^i}{w_{k,j|j-1}^i}\right). \tag{23}$$

Finally, the optimal kernel parameter for the kth degradation indicator at monitoring time t_j is derived as

$$s_{k,j} = \arg\min_{s_{k,j} \in [0,1]} \left[-\sum_{i=1}^N w_{k,j|j-1}^i \log\left(\frac{\hat{w}_{k,j}^i}{w_{k,j|j-1}^i}\right)\right]. \tag{24}$$

IV. RUL PREDICTION WITH MULTIPLE DEGRADATION INDICATORS BASED ON PARAMETER CORRELATION

To characterize the dependencies between multiple degradation indicators, we establish the correlations between parameters of the state transition equation, and the dependencies are also taken into consideration in OTKS-PF for state estimation of multiple indicators.

The commonly used RUL prediction methods with multiple degradation indicators establish the dependencies on the stochastic term of the degradation process, i.e., it is considered that there are correlations between $\sigma_1 B_1(t), \cdots, \sigma_K B_K(t)$ in (1) which cause the dependencies between $X_1(t), \cdots, X_K(t)$. Here we consider that the hidden dependencies between different degradation indicators are caused by the correlations between $\boldsymbol{\theta}_1, \cdots, \boldsymbol{\theta}_K$, and the RUL prediction method with multiple degradation indicators based on parameter correlation is proposed.

The state estimation based on any single degradation indicator using OTKS-PF has been introduced in section III, and then the RUL prediction can be conducted based on the estimated states and parameters. Consider the RUL prediction method with multiple degradation indicators, the prediction

framework is still the same as that with a single indicator. After the establishment of marginal degradation models and the estimation of fluctuation parameters, the proper multivariate state transition equation and the measurement equation can be established according to the marginal degradation models. Then OTKS-PF can be expanded considering the correlations between $\boldsymbol{\theta}_1, \cdots, \boldsymbol{\theta}_K$ for the case of multiple indicators, specifically, the correlations are introduced in the generating process of parameters' particles in the kernel smoothing step. The parameters $\boldsymbol{\theta}_1, \cdots, \boldsymbol{\theta}_K$ and the system states $X_1(t), \cdots, X_K(t)$ are jointly estimated online using OTKS-PF, and the joint probability distribution $p\left(x_{1j}, \cdots, x_{Kj}\right)$ of the degradation state $\mathbf{x}_j$ is obtained. At the prediction stage, the last obtained particles (the joint probability distribution) are propagated using the state transition equation $f\left(\mathbf{x}_j \mid \mathbf{x}_{j-1}\right)$ until equipment fails, i.e., any propagated indicator exceeds its corresponding failure threshold. Finally, the probability distribution of RUL can be obtained and the joint RUL prediction with multiple degradation indicators based on OTKS-PF is consequently derived.

A. JOINT STATE ESTIMATION OF MULTIPLE DEGRADATION INDICATORS BASED ON PARAMETER CORRELATION

It is considered that the $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \cdots, \boldsymbol{\theta}_K$ in (1) are linearly correlated with each other, and these correlations cause the dependencies between different degradation indicators when equipment degrades. And it should be noted that the linear or nonlinear correlations between multiple degradation indicators $X_1(t), X_2(t), \cdots, X_K(t)$ can be characterized using a combination of the linearly correlated parameters and the mapping relationship between $\boldsymbol{\theta}_k$ and $X_k(t)$ defined in (1). Therefore, compared with the RUL prediction method with a single indicator, the proposed joint estimation method needs to estimate ont only the parameters $\boldsymbol{\theta}_k$, σ_k^2, and v_k^2, but also needs to estimate the Pearson correlation coefficients between $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \cdots, \boldsymbol{\theta}_K$. Assume that the models for all degradation indicators are set as $\mu_k\left(\boldsymbol{\theta}_k, t\right) = a_k \exp\left(b_k \cdot t\right)$, where $\boldsymbol{\theta}_k = [a_k, b_k]^T$. The Pearson correlation coefficients between a_{k_1} and a_{k_2}, i.e., ρ_{k_1,k_2}^a and that between b_{k_1} and b_{k_2}, i.e., ρ_{k_1,k_2}^b can be estimated on the training units respectively, where $1 \leq k_1, k_2 \leq K$

To characterize the correlations between parameters, this approach generates the parameters' particles of different indicators from a multivariate Gaussian distribution directly to realize the evolution of parameters. This multivariate Gaussian distribution is the proposal distribution of the parameters with the correlations between each component.

Let $\Delta \mathbf{u}_j = \left[\Delta \mu_1(t_j), \cdots, \Delta \mu_K(t_j)\right]^T$ and $\boldsymbol{\omega}_j = \left[\omega_{1,j}, \cdots, \omega_{K,j}\right]^T$. And let $\boldsymbol{a}_j^i = [a_{1,j}^i, \cdots, a_{K,j}^i]^T$, $\boldsymbol{b}_j^i = [b_{1,j}^i, \cdots, b_{K,j}^i]^T$, and $\mathbf{x}_j^i = [x_{1,j}^i, \cdots, x_{K,j}^i]^T$ be the ith particle of parameter $\boldsymbol{a}$, parameter $\boldsymbol{b}$ and the state of multiple degradation indicators $\mathbf{x}$ at time t_j, respectively, where $\boldsymbol{a} = [a_1, \cdots, a_K]^T$, $\boldsymbol{b} = [b_1, \cdots, b_K]^T$, and $\mathbf{x} = [X_1(t), \cdots, X_K(t)]^T$. The state transition equation $f\left(\mathbf{x}_j \mid \mathbf{x}_{j-1}\right)$ of the proposed joint estimation method with

multiple indicators is

$$
\begin{cases}
\mathbf{x}_j = \mathbf{x}_{j-1} + \Delta \mathbf{u}_j + \boldsymbol{\omega}_j, \\
\boldsymbol{a}_j^i \mid \boldsymbol{a}_{j-1}^i \sim \mathcal{N}\left(q\boldsymbol{a}_{j-1}^i + (1-q)\bar{\boldsymbol{a}}_{j-1}, \left(1-q^2\right)\boldsymbol{V}_{a,j-1}\right), \\
\boldsymbol{b}_j^i \mid \boldsymbol{b}_{j-1}^i \sim \mathcal{N}\left(q\boldsymbol{b}_{j-1}^i + (1-q)\bar{\boldsymbol{b}}_{j-1}, \left(1-q^2\right)\boldsymbol{V}_{b,j-1}\right), \\
\Delta \mu_k(t_j) = \mu_k\left(\boldsymbol{\theta}_{k,j}, t_j\right) - \mu_k\left(\boldsymbol{\theta}_{k,j-1}, t_{j-1}\right), \\
\omega_{k,j} \sim \mathcal{N}\left(0, \sigma_k^2 \Delta t\right),
\end{cases}
\tag{25}
$$

where

$$
\boldsymbol{V}_{a,j-1}
= \begin{bmatrix}
V_{a,j-1}^1 & \cdots & \rho_{1,K}^a \sqrt{V_{a,j-1}^1 V_{a,j-1}^K} \\
\vdots & \ddots & \vdots \\
\rho_{K,1}^a \sqrt{V_{a,j-1}^K V_{a,j-1}^1} & \cdots & V_{a,j-1}^K
\end{bmatrix},
$$

$$
\boldsymbol{V}_{b,j-1}
= \begin{bmatrix}
V_{b,j-1}^1 & \cdots & \rho_{1,K}^b \sqrt{V_{b,j-1}^1 V_{b,j-1}^K} \\
\vdots & \ddots & \vdots \\
\rho_{K,1}^b \sqrt{V_{b,j-1}^K V_{b,j-1}^1} & \cdots & V_{b,j-1}^K
\end{bmatrix},
$$

$\bar{\boldsymbol{a}}_{j-1} = [\bar{a}_{1,j-1}, \cdots, \bar{a}_{K,j-1}]^T$, $\bar{\boldsymbol{b}}_{j-1} = [\bar{b}_{1,j-1}, \cdots, \bar{b}_{K,j-1}]^T$. $\{\mathbf{x}_{j-1}^i, \boldsymbol{a}_{j-1}^i, \boldsymbol{b}_{j-1}^i, w_{j-1}^i\}_{i=1}^N$ are the particles of $\mathbf{x}$, $\boldsymbol{a}$ and $\boldsymbol{b}$ with the corresponding weights estimated at time t_{j-1}. $\bar{a}_{k,j-1}$ and $V_{a,j-1}^k$ are the mean and variance of $a_{k,j-1}$ based on $\{a_{k,j-1}^i, w_{j-1}^i\}_{i=1}^N$. $\bar{b}_{k,j-1}$ and $V_{b,j-1}^k$ are the mean and variance of $b_{k,j-1}$ based on $\{b_{k,j-1}^i, w_{j-1}^i\}_{i=1}^N$. It should be noted that there is no correlation between $\boldsymbol{a}_j^i$ and $\boldsymbol{b}_j^i$ in (25).

It can be seen that when particles $\boldsymbol{a}_j^i$ and $\boldsymbol{b}_j^i$ are generated from the proposal distributions in the KS step, the proposal distributions are two multivariate Gaussian distributions with non-zero covariance, i.e., ρ_{k_1,k_2}^a and ρ_{k_1,k_2}^b which represent the correlations between indicators are contained in the covariance matrices of the multivariate Gaussian distributions. In this way, there are correlations of the parameters' particles between different indicators that are sampled from the proposal distribution, and then the weights of particles can be updated based on the newly collected measurements.

When the state transition equation (25) and the measurement equation (4) are used to estimate the degradation states and the parameters simultaneously with OTKS-PF, the steps are still the same as that in the section III-B, while the difference is that the update equation (16) is derived as

$$
\widetilde{w}_j^i = w_{j|j-1}^i \cdot \prod_{k=1}^K p\left(y_{k,j} \mid x_{k,j|j-1}^i\right),
\tag{26}
$$

and (24) is derived as

$$
s_j = \arg \min_{s_j \in [0,1]} \left[-\sum_{i=1}^N w_{j|j-1}^i \log\left(\frac{\hat{w}_j^i}{w_{j|j-1}^i}\right) \right],
\tag{27}
$$

so as to determine the optimal kernel parameter s_j in the OT step.

The weighted particles $\{\mathbf{x}_j^i, \mathbf{a}_j^i, \mathbf{b}_j^i, w_j^i\}_{i=1}^N$ can be obtained at time t_j using this joint estimation method with multiple degradation indicators based on parameter correlation, and the obtained particle $\mathbf{x}_j^i = [x_{1,j}^i, \cdots, x_{K,j}^i]^T$ representing the estimated states of multiple indicators is multidimensional with each dimension represents the state of the corresponding degradation indicator. There are correlations between the obtained $x_{k_1,j}^i$ and $x_{k_2,j}^i$, $a_{k_1,j}^i$ and $a_{k_2,j}^i$, $b_{k_1,j}^i$ and $b_{k_2,j}^i$ for any $k_1, k_2 \in [1, K]$. Therefore, the dependencies between different degradation indicators are established on the corresponding model parameters using the proposed method, and the joint probability distribution of degradation indicators $p(x_{1j}, \cdots, x_{Kj})$ is obtained.

B. RUL PREDICTION WITH MULTIPLE DEGRADATION INDICATORS

According to the joint estimation method with multiple indicators using OTKS-PF in Section IV-A, the particle set $\{\mathbf{x}_j^i, \boldsymbol{\theta}_j^i, w_j^i\}_{i=1}^N$ can be obtained at time t_j. Then the joint probability distribution of the state of multiple degradation indicators is

$$p(x_{1j}, \cdots, x_{Kj} \mid \mathbf{y}_{1:j}) \approx \sum_{i=1}^N w_j^i \delta\left(\mathbf{x}_j - \mathbf{x}_j^i\right). \quad (28)$$

The RUL corresponding to the ith particle which is predicted at monitoring time t_j is

$$l_j^i = \inf\{\tau : x_{1j}^i + \sum_{n=j+1}^{j+\tau} \left(\Delta\mu_1(t_n) + \omega_{1,n}\right) \geq H_1 \text{ or}$$

$$\cdots \text{ or } x_{Kj}^i + \sum_{n=j+1}^{j+\tau} \left(\Delta\mu_K(t_n) + \omega_{K,n}\right) \geq H_K\}, \quad (29)$$

and the probability density function (PDF) of RUL that is predicted at time t_j is

$$p(l_j \mid \mathbf{y}_{1:j}) \approx \sum_{i=1}^N w_j^i \delta\left(l_j - l_j^i\right). \quad (30)$$

V. EXPERIMENTAL VERIFICATION

The turbine engine data (C-MAPSS) that is published by NASA is applied to demonstrate the effectiveness of the proposed method for series systems, which contains four subsets (named *FD001 ∼ FD004*). Details of these subsets can be accessed in [26], and the subset involving a single failure mode and a single operating condition (*FD001*) is adopted here. The proposed joint-RUL-prediction method with multiple degradation indicators (OTKS-PF-Joint) is used to predict the RUL of the engines, and the prediction result is compared with the RUL result that obtained with a single degradation indicator (OTKS-PF-Single).

A. DATASET DESCRIPTION AND PREPROCESSING

The dataset contains degradation data from 100 simulated engines, each of them runs from a specific state to system
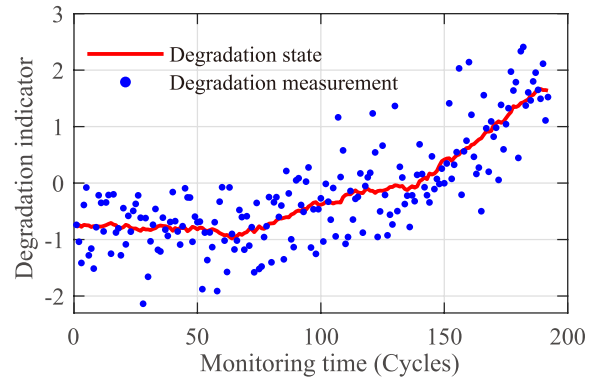
(a) Degradation indicator #1.
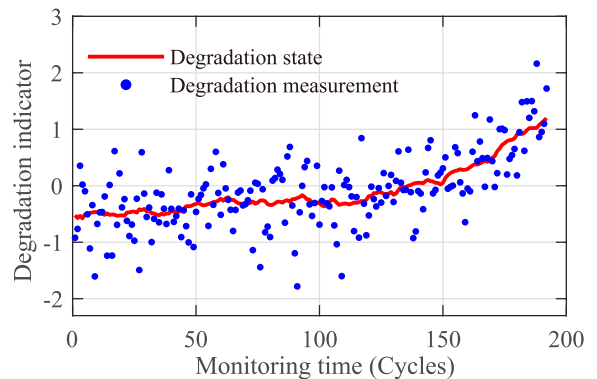
(b) Degradation indicator #2.

FIGURE 1. Degradation data of engine #1.

failure under a single operating condition and single fault mode. The degradation status of each engine is measured by 21 sensors.

The degradation data are preprocessed first, which includes standardization and dimension reduction. Z-score standardization is carried out on the original data for all the 21 sensors, and PCA is used for dimension reduction to remove features that provide little degradation information. The obtained two features using PCA, i.e., the first two principal components, retain 87.6% of the variance of the original data, so the first two principal components are chosen as the two degradation indicators which characterize the degradation of engines. Then the proposed joint-RUL-prediction method with two degradation indicators is conducted to predict the RUL of engines, and the RUL prediction result based on a single indicator is compared with that based on two indicators.

The stochastic process models are established as (1) and (2) for each degradation indicator, where $k = 1, 2$ represents the kth degradation indicator. Then SMA is used to obtain the smoother curves regarded as the states of the degradation indicators $X_k^*(t)$, $k = 1, 2$. Figure. 1 shows the measurements and the degradation states of the engine #1.

To demonstrate the effectiveness of the proposed method for series systems, the failure thresholds of indicator #1 and indicator #2 are set as $H_1 = 1.22$ and $H_2 = 0.85$, respectively. And the failure time of the engine is defined as the first hitting time that any indicator exceeds the corresponding

failure threshold, which is consistent with the failure definition of the series systems. For all the 100 engines, 5-fold cross-validation is adopted to divide the testing units and the training units and five experiments are conducted using different training units and testing units. The run-to-failure data of training units are already collected, while partial data of each testing unit are removed to simulate the practical cases, where there are only partial data for the practicer to derive RUL. In each experiment, the RUL of the testing units is predicted at different monitoring time. The following analysis takes the first experiment as an example due to the limit of the length of this article, and the results of all the five experiments are also illustrated.

B. DEGRADATION MODELING AND PARAMETER ESTIMATION

The marginal degradation models are established according to (1) and (2), and the degradation trajectory $\mu_k(\cdot)$ should be determined first. Liu *et al.* [27] predicted the RUL on this dataset with the exponential model, here we also establish the exponential model for degradation indicators as

$$\mu_k(\boldsymbol{\theta}_k, t) = a_k + b_k \cdot \exp(c_k \cdot t). \tag{31}$$

Two strategies are considered when the model (31) is used for testing units, and the commonly used strategy (strategy I) is that a_k, b_k and c_k are all considered as unknown parameters for testing units which should be estimated by PF. However, the estimation error of parameters will be large because of the interaction between b_k and c_k in PF [22], and considering the consistency of different units given the same operating condition and failure mode [3], the degradation processes of different units are expected to be similar. Therefore, strategy II can be chosen by defining b_k or c_k as constants which are calculated using training units for the testing units.

Degradation data of training units are used to verify strategy I first. Consistent with section V-A, the measurements $Y_k^*(t)$, $k = 1, 2$ of the training units are processed by SMA, and the degradation states are obtained as $X_k^*(t)$, $k = 1, 2$. Matlab curve fitting toolbox is used to fit the degradation states of the two indicators with the model (31) to get the fitting parameters a_k^*, b_k^*, and c_k^* using the least-squares method, and Fig. 2 is the boxplot that illustrates the distribution of the fitting parameters.

It can be seen from Fig. 2 that, different from the parameters a_k^* and c_k^*, there are a number of outliers in the distributions of b_1^* and b_2^*, which indicates b_k^* may not reflect the consistency of different units under the same failure mode and operating condition accurately. Moreover, according to [28], model parameters can be categorized as fixed-effect and random effect parameters, and the former describes the homogeneity among the units. Therefore, strategy II is considered by defining the parameters b_1 and b_2 as two mixed-effect parameters which are estimated as the median of the fitting parameters: $b_1 = 0.085$, $b_2 = 0.056$, and a_k^* and c_k^* are random for testing units to describe the heterogeneity among the units. For this experiment (the first experiment),
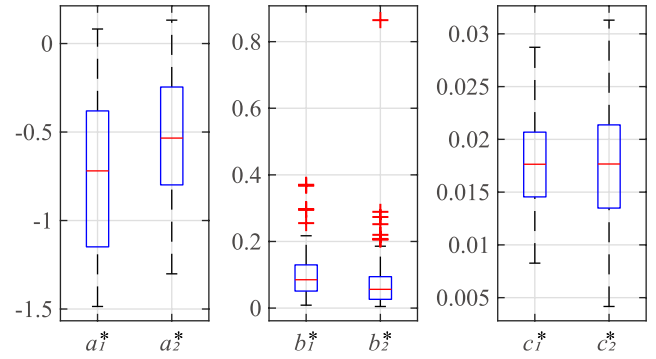

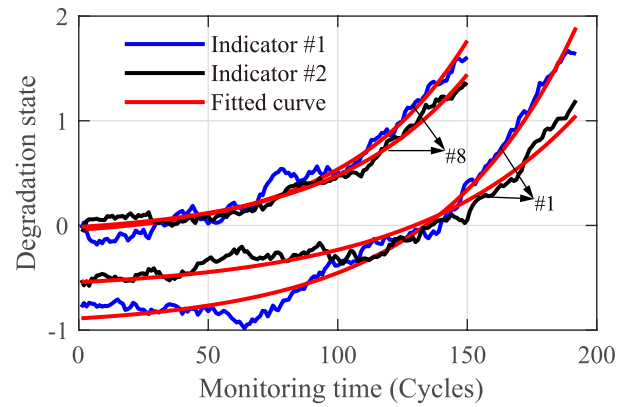**FIGURE 2.** Boxplot of fitting parameters using model (32) with strategy I.

FIGURE 3. Degradation data and fitted curves of engine #1 and engine #8.

TABLE 1. Estimation results of fluctuation parameters.

Parameter	σ_1^2	σ_2^2	ν_1^2	ν_2^2
Estimated value	0.0011	9.17e-04	0.2649	0.2432

model (31) can be redefined by using strategy II:

$$\mu_1(\boldsymbol{\theta}_1, t) = a_1 + 0.085 * \exp(c_1 \cdot t),$$
$$\mu_2(\boldsymbol{\theta}_2, t) = a_2 + 0.056 * \exp(c_2 \cdot t). \tag{32}$$

Figure. 3 shows the fitting result of two engines in the training set using strategy II, and it can be seen that the model fits the degradation data well using strategy II.

As a result, strategy II is used for the testing units in this experiment. Then (8) and (9) are used to estimate the fluctuation parameters σ_k^2 and ν_k^2 on the training units according to the method in section II-B. Table 1 shows the parameters estimated based on the training units.

C. JOINT RUL PREDICTION BASED ON PARAMETER CORRELATION WITH TWO INDICATORS

After the establishment of the degradation models and the estimation of the fluctuation parameters, OTKS-PF is used for joint state estimation of the two degradation indicators involving parameter correlation. According to (25), the following state transition equation $f(\mathbf{x}_j, \boldsymbol{c}_j \mid \mathbf{x}_{j-1}, \boldsymbol{c}_{j-1})$ is established

for the two-dimensional degradation indicators of the engine:

$$\begin{cases} \mathbf{x}_j = \mathbf{x}_{j-1} + \Delta \mathbf{u}_j + \boldsymbol{\omega}_j, \\ \mathbf{c}_j^i \mid \mathbf{c}_{j-1}^i \sim \mathcal{N}\left(q\mathbf{c}_{j-1}^i + (1-q)\,\bar{\mathbf{c}}_{j-1}, \left(1-q^2\right)\mathbf{V}_{c,j-1}\right), \\ \Delta \mu_k(t_j) = b_k \cdot \left(\exp\left(c_{k,j} \cdot t_j\right) - \exp\left(c_{k,j-1} \cdot t_{j-1}\right)\right), \\ \omega_{k,j} \sim \mathcal{N}\left(0, \sigma_k^2 \Delta t\right), \end{cases}$$

$$(33)$$

where $k = 1, 2$ represents the kth degradation indicator.

$$\mathbf{V}_{c,j-1} = \begin{bmatrix} V_{c,j-1}^1 & \rho_{1,2}^c \sqrt{V_{c,j-1}^1 V_{c,j-1}^2} \\ \rho_{2,1}^c \sqrt{V_{c,j-1}^2 V_{c,j-1}^1} & V_{c,j-1}^2 \end{bmatrix},$$

$\mathbf{c}_j^i = \left[c_{1,j}^i, c_{2,j}^i\right]^T$, $b_1 = 0.085$, $b_2 = 0.056$ (for the first experiment). $\bar{\mathbf{c}}_{j-1} = \left[\bar{c}_{1,j-1}, \bar{c}_{2,j-1}\right]^T$, and $\{\mathbf{c}_{j-1}^i, w_{j-1}^i\}_{i=1}^N$ are the particles of $\mathbf{c} = [c_1, c_2]^T$ with their corresponding weights estimated at time t_{j-1}. $\bar{c}_{k,j-1}$ and $V_{c,j-1}^k$ are the mean and variance of $c_{k,j-1}$ based on $\{c_{k,j-1}^i, w_{j-1}^i\}_{i=1}^N$.

It can be seen from (33) that only the increment of the state is involved in the state transition process, so a_1 and a_2 do not need to be estimated in (33). For each training unit, the parameters a_k^* and c_k^* are obtained by fitting the degradation states $X_k^*(t)$ with model (32), and $\mathbf{c}_u^* = \left[c_{1,u}^*, c_{2,u}^*\right]^T$ is obtained for all taining units, where $u = 1, 2, \cdots, 80$ representing the uth unit. The Pearson correlation coefficient between c_1 and c_2 is estimated as

$$\rho_{1,2}^c = \frac{\sum_{u=1}^{80}\left(c_{1,u}^* - \bar{c}_1^*\right)\left(c_{2,u}^* - \bar{c}_2^*\right)}{\sqrt{\left[\sum_{u=1}^{80}\left(c_{1,u}^* - \bar{c}_1^*\right)^2\right]\left[\sum_{n=1}^{80}\left(c_{2,u}^* - \bar{c}_2^*\right)^2\right]}},$$

$$(34)$$

where $\bar{c}_k^* = \sum_{u=1}^{80}\left(c_{k,u}^*\right)/80$, $k = 1, 2$, and it is calculated as $\rho_{1,2}^c = 0.9873$ according to (34). Figure. 4 shows the scatter plot of the parameters c_1^* and c_2^* of all the training units. It can be seen that there is a significant correlation between the parameters c_1 and c_2 from Fig. 4 and the calculated value of $\rho_{1,2}^c$, which indicates that the dependence between degradation indicators is evident and can be characterized in the correlation of parameters.

The OTKS-PF method in Section IV is used to conduct the joint estimation of $X_1(t)$, $X_2(t)$, c_1, and c_2 online using the state transition equation (33) and the measurement equation (4). The initial distribution of states and parameters $p(\mathbf{x}_0)$ and $p(\mathbf{c}_0)$ are set as Gaussian distributions, i.e., $\mathbf{x}_0 \sim \mathcal{N}(\boldsymbol{\mu}_x, \boldsymbol{\Sigma}_x)$ and $\mathbf{c}_0 \sim \mathcal{N}(\boldsymbol{\mu}_c, \boldsymbol{\Sigma}_c)$, where $\boldsymbol{\mu}_x = \left[\bar{x}_{1,1}, \bar{x}_{2,1}\right]^T$, $\boldsymbol{\mu}_c = [\bar{c}_1, \bar{c}_2]^T$, $\boldsymbol{\Sigma}_x = \text{diag}(\sigma_{x,1}, \sigma_{x,2})$, and $\boldsymbol{\Sigma}_c = \text{diag}(\sigma_{c,1}, \sigma_{c,2})$. $\bar{x}_{k,1}$ and $\sigma_{x,k}$ are the mean and variance of $x_{k,1}^*$ of all training units, respectively, where $x_{k,1}^*$ is the degradation state of the kth indicator at time t_1. $\bar{c}_k$ and $\sigma_{c,k}$ are the mean and variance of c_k^* of all training units, respectively.

Figure. 5 shows the parameter estimation results of engine #5 using the proposed OTKS-PF-Joint and the OTKS-PF-Single respectively. The black lines represent the posterior
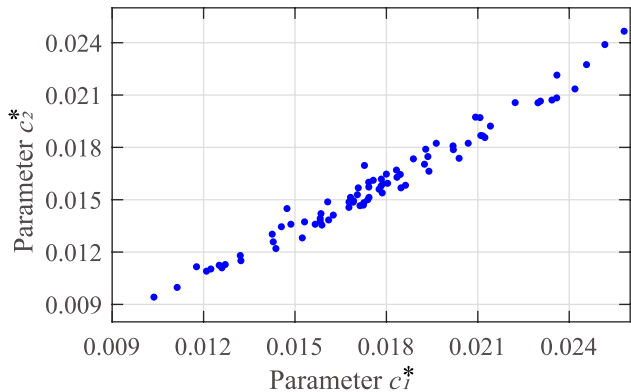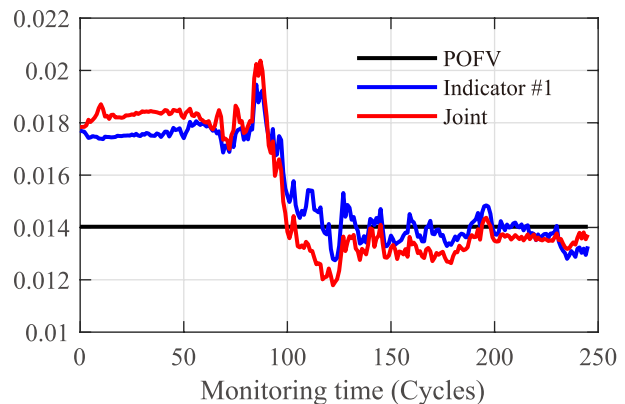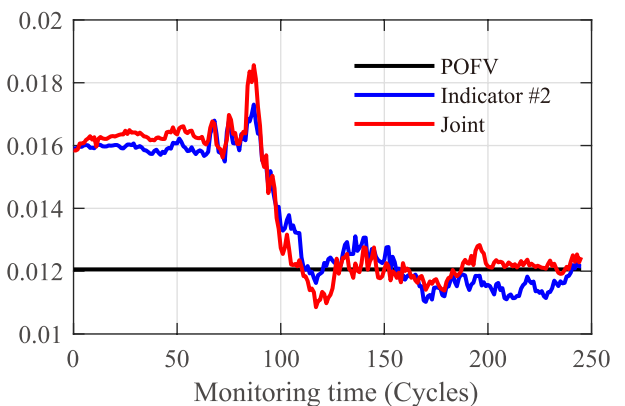

FIGURE 4. Scatter plot of the fitting parameters of training units.

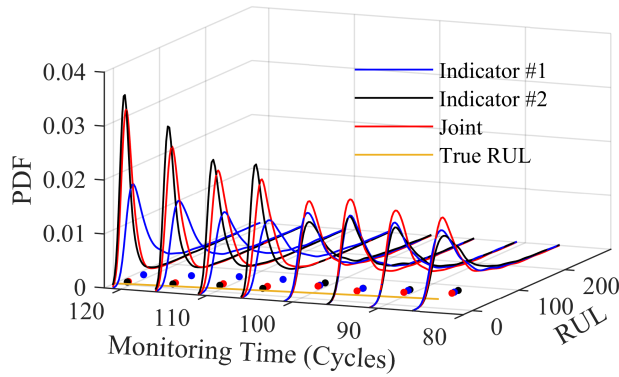(a) Estimation of parameter c_1 using indicator #1 and OTKS-PF-Joint.

(b) Estimation of parameter c_2 using indicator #2 and OTKS-PF-Joint.

FIGURE 5. Online estimation of parameter c of engine #5 using OTKS-PF-Single and OTKS-PF-Joint.

optimal fitted values (POFV). After all the measurements are collected, the POFVs are obtained by fitting the degradation states estimated by using SMA. It can be seen that the estimation result of OTKS-PF-Joint is similar to that of OTKS-PF-Single. The estimated parameters can gradually converge to the POFV with the increase of the number of collected measurements.
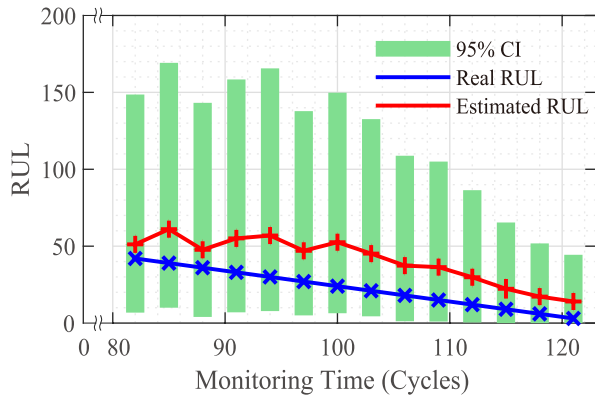
Figure. 6 and Fig. 7 respectively show the RUL prediction results of engine #19 and engine #68 in the testing set. Figure. 6-a and Fig. 7-a show the predicted PDFs and RULs
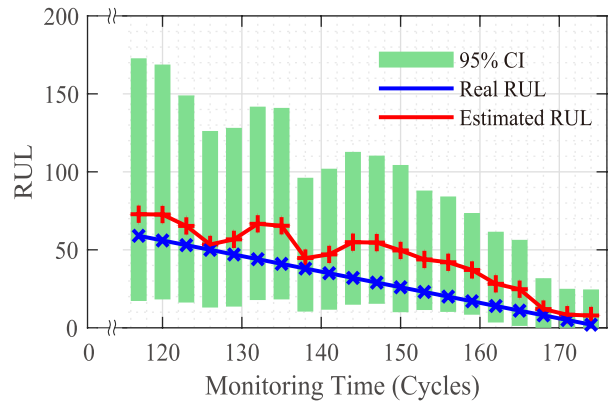
(a) RUL prediction result using OTKS-PF-Single and OTKS-PF-Joint.
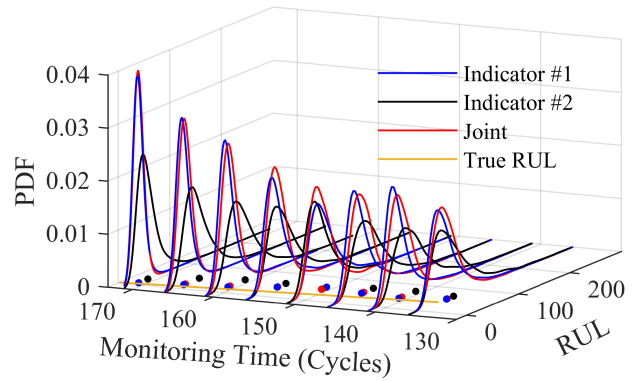
(b) RUL prediction result using OTKS-PF-Joint.

FIGURE 6. RUL prediction of engine #19 using OTKS-PF-Single and OTKS-PF-Joint.

(a) RUL prediction result using OTKS-PF-Single and OTKS-PF-Joint.

(b) RUL prediction result using OTKS-PF-Joint.

FIGURE 7. RUL prediction of engine #68 using OTKS-PF-Single and OTKS-PF-Joint.

estimated using indicator #1, indicator #2, and OTKS-PF-Joint respectively. It can be seen from Fig. 6-a that the predicted PDF of RUL with OTKS-PF-Joint keeps closer tracking to the predicted PDF based on indicator #2 than that based on indicator #1. Such a situation changes and the predicted PDF of RUL with OTKS-PF-Joint keeps closer to the indicator #1 in Fig. 7-a. These are because the degradation indicator #2 of engine #19 exceeds its corresponding failure threshold first, while the degradation indicator #1 of engine #68 exceeds the failure threshold first. The prediction results in Fig. 6-a and Fig. 7-a demonstrate that no matter which indicator exceeds its corresponding failure threshold first, the PDF of RUL predicted by the proposed OTKS-PF-Joint can track the PDF that predicted based on that indicator in an adaptive way. As a result, the proposed method will provide accurate RUL prediction for equipment according to the definition of equipment failure, i.e., the first hitting time that any indicator exceeds the corresponding failure threshold is regarded as the EoL.

Subsequently, the above prediction results of the proposed method which are shown in Fig. 6-a and Fig. 7-a are analyzed in detail. At the prediction stage, each particle is propagated using the state transition equation, and the predicted failure time of the equipment based on each particle is defined as the first hitting time that any dimension of the particle reaches

the corresponding failure threshold. Therefore, the predicted PDF of RUL with multiple degradation indicators is obtained and the predicted RUL value will be accurate.

Figure. 6-b and Fig. 7-b respectively show the RUL prediction result of engine #19 and engine #68 using OTKS-PF-Joint. The widths of the 95% confidence intervals become narrower as the monitored equipment running, which means that the uncertainty of prediction decrease gradually with the number of collected measurements increasing. It also can be seen that the predicted RUL is getting closer to the real one as the monitored equipment running, which indicates that the accuracy of RUL prediction increases when the monitored equipment runs closer to its EoL.

To compare the RUL prediction results of the proposed OTKS-PF-Joint and the OTKS-PF-Single intuitively, the RUL of testing units is predicted at each cycle after the engine runs beyond two-thirds of its whole life so that several illustrative prediction results can be obtained for each testing unit. Then the mean of the several prediction results is calculated. Figure. 8 shows the mean prediction error and the mean width of the 95% confidence interval for the testing units.

It can be seen that the prediction errors of the proposed method are generally lower than that of the RUL prediction via OTKS-PF-Single. The 95% confidence interval widths
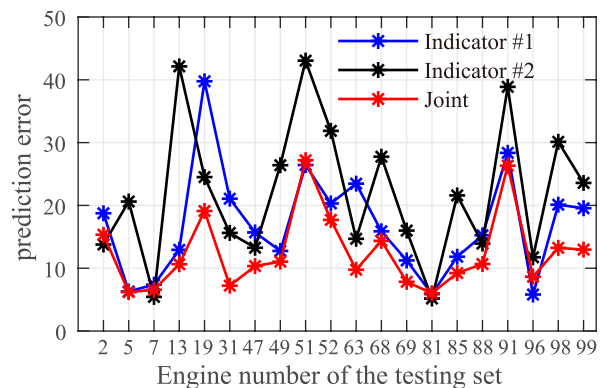
(a) Mean prediction errors.
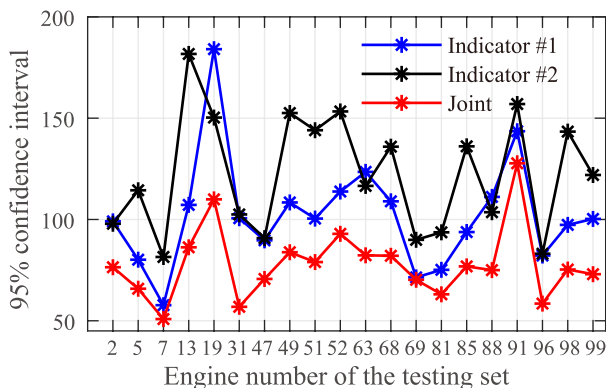
(b) Mean widths of 95% confidence intervals.

FIGURE 8. The mean prediction errors and mean widths of 95% confidence intervals of testing units.

TABLE 2. Prediction result of OTKS-PF-Joint and OTKS-PF-Single.

Experiment	1	2	3	4	5
Method	Mean prediction error				
Indicator #1	16.93	24.31	21.63	19.64	14.71
Indicator #2	22.00	25.29	21.66	22.45	15.18
Joint	12.51	17.26	16.56	15.27	9.76
Method	Mean width of 95% confidence interval				
Indicator #1	102.42	124.76	117.67	110.60	102.99
Indicator #2	122.50	148.18	134.66	141.50	128.01
Joint	77.82	92.27	89.58	83.21	74.97

obtained by the proposed method are also narrower than that with a single indicator. This result confirms the validity of the proposed joint-RUL-prediction method with multiple indicators based on parameter correlation. This result confirms the validity of the proposed joint-RUL-prediction method with multiple indicators based on parameter correlation.

For each experiment, the mean prediction errors and the mean widths of 95% confidence interval for all testing units are averaged, and Table 2 gives the average prediction result using the proposed OTKS-PF-Joint and the OTKS-PF-Single respectively for all 5 experiments of the 5-fold cross-validation. The results demonstrate that the proposed method is superior to OTKS-PF-Single in the accuracy and the uncertainty of the prediction.

VI. CONCLUSION AND DISCUSSION

In this article, a joint-RUL-prediction method with multiple degradation indicators based on parameter correlation is proposed by combining the stochastic process model and PF. The dependencies between different degradation indicators are characterized as the correlations between their parameters, and the covariance matrices of model parameters between different degradation indicators are introduced into KS-PF to quantify such a correlation. Besides, the OT approach is used to choose the optimal kernel parameter in KS. Then the joint probability distribution of multidimensional degradation state can be estimated via OTKS-PF considering parameter correlation, and the RUL of equipment is consequently predicted with multiple degradation indicators. A case study regarding engine degradation is performed, and the result shows that the proposed method is superior to the prediction method with a single indicator in prediction accuracy and prediction uncertainty.

The proposed method is designed for the product with multiple performance characteristics, or the system with multiple components, in which the dependencies between multiple degradation indicators can be characterized as the correlations of model parameters. In addition to the series systems, the proposed method can also be applied to other systems with differnt failure structures, such as the parallel systems or the combination systems, since the degradation states of multiple degradation indicators can be estimated simultaneously by the proposed method.

REFERENCES

[1] W. Peng, Z.-S. Ye, and N. Chen, "Joint online RUL prediction for multivariate deteriorating systems," *IEEE Trans. Ind. Informat.*, vol. 15, no. 5, pp. 2870–2878, May 2019.

[2] S. Zhao, S. Chen, F. Yang, E. Ugur, B. Akin, and H. Wang, "A composite failure precursor for condition monitoring and remaining useful life prediction of discrete power devices," *IEEE Trans. Ind. Informat.*, vol. 17, no. 1, pp. 688–698, Jan. 2021.

[3] Y. Lei, N. Li, L. Guo, N. Li, T. Yan, and J. Lin, "Machinery health prognostics: A systematic review from data acquisition to RUL prediction," *Mech. Syst. Signal Process.*, vol. 104, pp. 799–834, May 2018.

[4] X. Jia, L. Xing, and G. Li, "Copula-based reliability and safety analysis of safety-critical systems with dependent failures," *Qual. Rel. Eng. Int.*, vol. 34, no. 5, pp. 928–938, Jul. 2018.

[5] S. Zhao, V. Makis, S. Chen, and Y. Li, "Evaluation of reliability function and mean residual life for degrading systems subject to condition monitoring and random failure," *IEEE Trans. Rel.*, vol. 67, no. 1, pp. 13–25, Mar. 2018.

[6] Q. Zhai and Z.-S. Ye, "RUL prediction of deteriorating products using an adaptive Wiener process model," *IEEE Trans. Ind. Informat.*, vol. 13, no. 6, pp. 2911–2921, Dec. 2017.

[7] N. Li, Y. Lei, L. Guo, T. Yan, and J. Lin, "Remaining useful life prediction based on a general expression of stochastic process models," *IEEE Trans. Ind. Electron.*, vol. 64, no. 7, pp. 5709–5718, Jul. 2017.

[8] Y. Lei, N. Li, and J. Lin, "A new method based on stochastic process models for machine remaining useful life prediction," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 12, pp. 2671–2684, Dec. 2016.

[9] X. Wang, C. Hu, X. Si, Z. Pang, and Z. Ren, "An adaptive remaining useful life estimation approach for newly developed system based on nonlinear degradation model," *IEEE Access*, vol. 7, pp. 82162–82173, 2019.

[10] C. Sbarufatti, M. Corbetta, M. Giglio, and F. Cadini, "Adaptive prognosis of lithium-ion batteries based on the combination of particle filters and radial basis function neural networks," *J. Power Sources*, vol. 344, pp. 128–140, Mar. 2017.

[11] Y. Wu, W. Li, Y. Wang, and K. Zhang, "Remaining useful life prediction of lithium-ion batteries using neural network and bat-based particle filter," *IEEE Access*, vol. 7, pp. 54843–54854, 2019.

[12] L. Zhang, Z. Mu, and C. Sun, "Remaining useful life prediction for lithium-ion batteries based on exponential model and particle filter," *IEEE Access*, vol. 6, pp. 17729–17740, 2018.

[13] X. Li, L. Peng, L. Gao, D. Bi, X. Xie, and Y. Xie, "A robust hybrid filtering method for accurate battery remaining useful life prediction," *IEEE Access*, vol. 7, pp. 57843–57856, 2019.

[14] M. Jouin, R. Gouriveau, D. Hissel, M.-C. Péra, and N. Zerhouni, "Particle filter-based prognostics: Review, discussion and perspectives," *Mech. Syst. Signal Process.*, vols. 72–73, pp. 2–31, May 2016.

[15] M. Bressel, M. Hilairet, D. Hissel, and B. O. Bouamama, "Extended Kalman filter for prognostic of proton exchange membrane fuel cell," *Appl. Energy*, vol. 164, pp. 220–227, Feb. 2016.

[16] J. Liu, E. Zio, and Y. Hu, "Particle filtering for prognostics of a newly designed product with a new parameters initialization strategy based on reliability test data," *IEEE Access*, vol. 6, pp. 62564–62573, 2018.

[17] Q. Qin, S. Zhao, S. Chen, D. Huang, and J. Liang, "Adaptive and robust prediction for the remaining useful life of electrolytic capacitors," *Microelectron. Rel.*, vol. 87, pp. 64–74, Aug. 2018.

[18] G. Fang, R. Pan, and Y. Hong, "Copula-based reliability analysis of degrading systems with dependent failures," *Rel. Eng. Syst. Saf.*, vol. 193, Jan. 2020, Art. no. 106618.

[19] X. Xi, M. Chen, and D. Zhou, "Remaining useful life prediction for multi-component systems with hidden dependencies," *Sci. China Inf. Sci.*, vol. 62, no. 2, Feb. 2019, Art. no. 022202.

[20] L. A. Rodríguez-Picón, V. H. Flores-Ochoa, L. C. Méndez-González, and M. A. Rodríguez-Medina, "Bivariate degradation modelling with marginal heterogeneous stochastic processes," *J. Stat. Comput. Simul.*, vol. 87, no. 11, pp. 2207–2226, Jul. 2017.

[21] F. Sun, N. Wang, X. Li, and W. Zhang, "Remaining useful life prediction for a machine with multiple dependent features based on Bayesian dynamic linear model and copulas," *IEEE Access*, vol. 5, pp. 16277–16287, 2017.

[22] Y. Hu, P. Baraldi, F. Di Maio, and E. Zio, "A particle filtering and kernel smoothing-based approach for new design component prognostics," *Rel. Eng. Syst. Saf.*, vol. 134, pp. 19–31, Feb. 2015.

[23] Y. Hu, P. Baraldi, F. Di Maio, and E. Zio, "Online performance assessment Method_newline for a model-based prognostic approach," *IEEE Trans. Rel.*, vol. 65, no. 2, pp. 718–735, Jun. 2016.

[24] A. Tulsyan, B. Huang, R. B. Gopaluni, and J. F. Forbes, "On simultaneous on-line state and parameter estimation in non-linear state-space models," *J. Process Control*, vol. 23, no. 4, pp. 516–526, Apr. 2013.

[25] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 174–188, Feb. 2002.

[26] W. Yu, I. Y. Kim, and C. Mechefske, "Remaining useful life estimation using a bidirectional recurrent neural network based autoencoder scheme," *Mech. Syst. Signal Process.*, vol. 129, pp. 764–780, Aug. 2019.

[27] K. Liu, A. Chehade, and C. Song, "Optimize the signal quality of the composite health index via data fusion for degradation modeling and prognostic analysis," *IEEE Trans. Autom. Sci. Eng.*, vol. 14, no. 3, pp. 1504–1514, Jul. 2017.

[28] C. J. Lu and W. O. Meeker, "Using degradation measures to estimate a Time-to-Failure distribution," *Technometrics*, vol. 35, no. 2, pp. 161–174, May 1993.

SHAOWEI CHEN (Member, IEEE) is currently an Associate Professor with the School of Electronics and Information, Northwestern Polytechnical University, Xi'an, China, where he is also the Dean of the Department of Telecommunication Engineering and the Director of the Signal Processing and Control Technology Laboratory. He is also the Principal Investigator of several projects supported by the Aeronautical Science Foundation of China, Beijing, China. His research interests include fault diagnosis, sensors, condition monitoring, and prognosis of electronic systems. He is a Senior Member of the Chinese Institute of Electronics. He is selected to receive several Provincial and Ministerial Science and Technology Awards.

MEINAN WANG received the B.S. degree in telecommunication engineering from Northwestern Polytechnical University, Xi'an, China, in 2017, where she is currently pursuing the M.S. degree in telecommunication and information system with the School of Electronics and Information. Her research interests include condition monitoring and residual life prediction of complex systems.

DENGSHAN HUANG received the B.S. degree in radar engineering and the M.S. degree in signals, circuits and systems from Xidian University, Xi'an, China, in 1982 and 1985, respectively.

From 1994 to 1995, he was a Visiting Scholar with The University of Tokyo, Tokyo, Japan. He is currently a Professor with the School of Electronics and Information, Northwestern Polytechnical University, Xi'an. His research interests include system informatics, signal processing, and condition monitoring of complex systems.

PENGFEI WEN (Graduate Student Member, IEEE) received the B.S. degree in telecommunication engineering from the School of Electronics and Information, Northwestern Polytechnical University, Xi'an, China, in 2017, where he is currently pursuing the Ph.D. degree in information and telecommunication engineering. His research interests include statistical modeling, data and information fusion for data analytics, and prognostics.

SHENGYUE WANG received the B.S. degree in electronic science and technology from Northeastern University, Shenyang, China, in 2019. She is currently pursuing the M.S. degree in telecommunication and information system with the School of Electronics and Information, Northwestern Polytechnical University, Xi'an, China. Her research interests include the reliability of power electronic converter systems and the machine learning.

SHUAI ZHAO (Member, IEEE) received the B.S. degree in telecommunication engineering, the M.S. degree in telecommunication and information system, and the Ph.D. degree in information and telecommunication engineering from Northwestern Polytechnical University, Xi'an, China, in 2011, 2014, and 2018, respectively.

From September 2014 to September 2016, he was a Visiting Ph.D. Student with the Department of Mechanical and Industrial Engineering, University of Toronto, Toronto, ON, Canada, with the scholarship from the China Scholarship Council (CSC). In August 2018, he was a Visiting Scholar with the Power Electronics and Drives Laboratory, Department of Electrical and Computer Science, The University of Texas at Dallas, Richardson, TX, USA. He is currently a Postdoctoral Researcher with the Center of Reliable Power Electronics (CORPE), Department of Energy Technology, Aalborg University, Denmark. His research interests include condition monitoring, data analytics, residual life prediction, and health assessment of power electronic systems.

• • •