A Multivariate Student-t Process Model for Dependent Tail-weighted Degradation Data

Ancha Xu[a,b], Guanqi Fang[a], Liangliang Zhuang[a,*], Cheng Gu[a]

[a] *School of Statistics and Mathematics, Zhejiang Gongshang University, Zhejiang 310018, China*
[b] *Collaborative Innovation Center of Statistical Data Engineering, Technology & Application Zhejiang Gongshang University, Zhejiang 310018, China*

Abstract

Traditionally, Gaussian assumption, implied by the Wiener process, is widely admitted for modeling degradation processes. However, when degradation data exhibit heavy tails, this assumption is not suitable. To overcome this limitation, this article proposes a novel class of tail-weighted multivariate degradation model, which is built upon Student-t process. The model is able to account for both between-unit variability and process dependency, while allows adjusting the tail heaviness through tuning the parameter of the degree of freedom. For reliability assessment, we derive the system reliability function and present an efficient Monte Carlo method for its evaluation. Further, we introduce an expectation-maximization algorithm for parameter estimation and design a bootstrap method for interval estimation. Comprehensive simulation studies are conducted to validate the effectiveness of the inference method. Finally, the proposed methodology is applied to analyze two real-world degradation datasets.

Keywords: Reliability; Bootstrap; EM algorithm; heavy tail; multivariate degradation.

*Corresponding author

Email address: 22010040022@pop.zjgsu.edu.cn (Liangliang Zhuang)

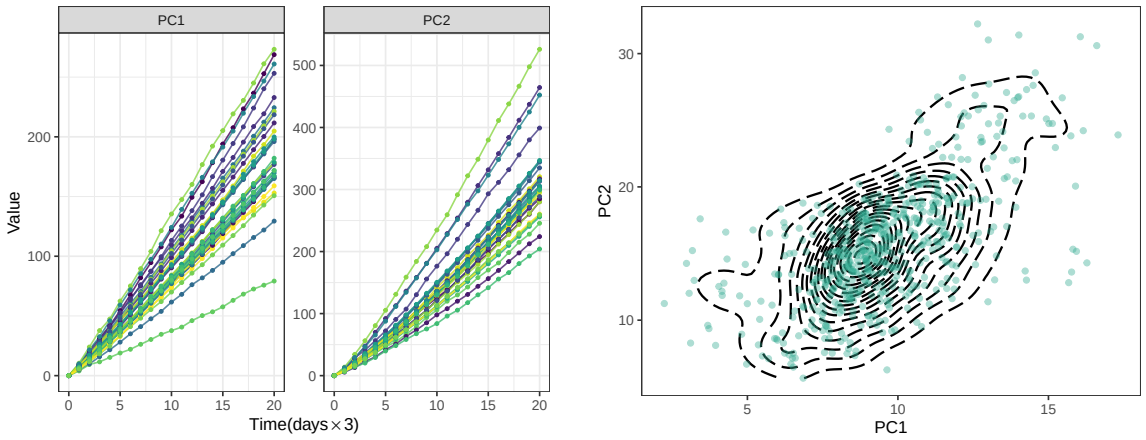
1. Introduction

1.1. Background

In engineering practice, degradation tests are commonly utilized to promptly identify the real-time quality status, enabling reliability assessments for long-lifespan products (Zhao et al., 2021; Chen et al., 2019). Examples of typical application include laser devices (Meeker et al., 2022), batteries (Zhao et al., 2023; Peng et al., 2019), and fatigue cracks (Fang et al., 2022). To accurately assess the overall quality of these products, the examination of multiple performance characteristics (PCs) is often entailed. And these PCs tend to exhibit strong correlations due to the stochastic nature of degradation processes. Furthermore, a product may encounter rare events during its operation. For example, electronic devices can be vulnerable to interference from nearby equipment or other sources of electromagnetic emissions. Abrupt voltage pulses, such as power fluctuations or lightning strikes, can have a significant impact on the operation of electronic devices (Peng and Cheng, 2020). When such rare events occur, from a statistical point of view, the PCs will exhibit heavy-tail characteristics. In response to these concerns, our objective is to develop a tail-weighted multivariate degradation model. The subsequent example delves into these points in greater detail, providing further incentive for our investigation.

1.2. Motivation

Permanent magnet brakes (PMBs), a type of braking system, use permanent magnets to generate the braking force. It achieves braking by clamping a magnetic yoke through actuating the rotor (Kou et al., 2021). To investigate the product's reliability, 30 specimens of PMBs as braking servo motors in injection molding machines were subjected to a degradation test in a typical high-temperature environment. In this study, the malfunction of PMBs, characterized by the loss of magnetization (demagnetization), is represented by two PCs: braking torque (PC1) and response time (PC2). The degradation of these two PCs was measured at three-day intervals throughout the two-month experimental period. Figure 1(a) demonstrates the degradation paths of the PCs, while Figure 1(b) presents a scatter plot of degradation increments and fitted empirical contour lines. Clearly, the scatter plot indicates a positive correlation between the two PCs and the non-elliptical trend in the tail region of the contour plane implies non-Gaussian features with tail dependency. Furthermore, we fit the data at two arbitrary time points ($t = 9$ and 18) with both normal and Student's t distributions, and Figure 2 displays quantile-quantile (Q-Q) plots for each PC. It



(a) Degradation paths (b) Contour plots

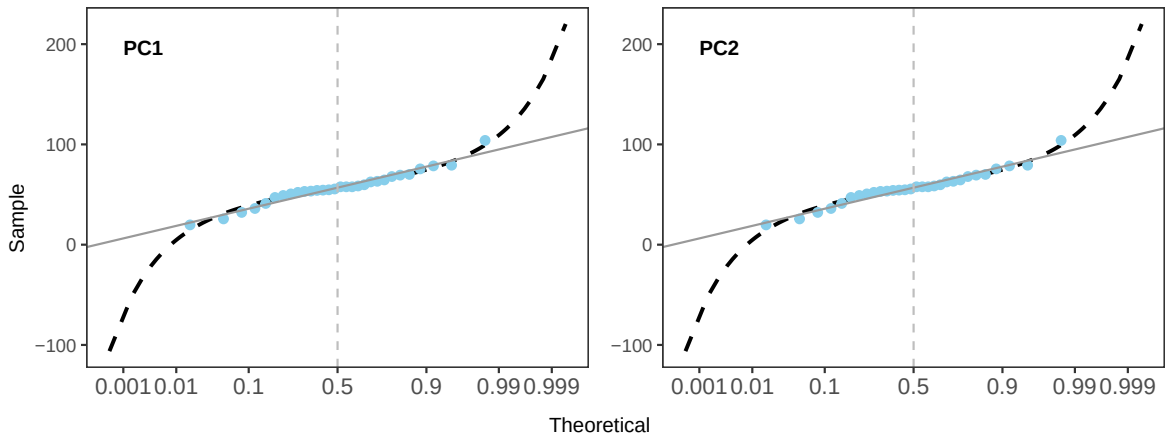
Figure 1: PMB dagradation data.

is obvious that the sample data points exhibit a curved (or S-shaped) pattern, suggesting the skewness and the departure from normality. And their proximity to the theoretical lines of the Student's t distribution (i.e., dashed lines) confirms this trend. Thus, all the aforementioned phenomena motivate us to establish a model for modeling tail-weighted multivariate degradation processes.

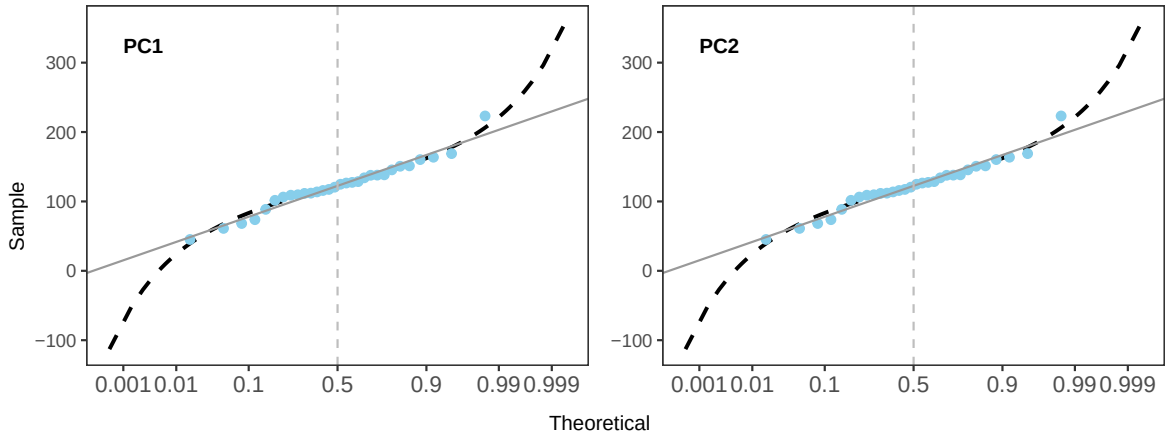
1.3. Related literature

In the existing literature, comprehensive studies have been conducted on the degradation analysis of single PC products, employing frameworks such as general path models and stochastic process models. The introduction of the general path model to degradation analysis was pioneered by Lu and Meeker (1993), and subsequent research can be found in (Bae and Kvam, 2004; Fan et al., 2012). Unlike the general path model, stochastic process models efficiently handle the introduced randomness from inherent variability and environmental factors in the degradation process. Three primary classes of stochastic degradation processes are widely discussed in the literature—the Wiener process (Zhai and Ye, 2018; Zhang et al., 2023; Zhai et al., 2024), the Gamma process (Yao et al., 2024; Chen and Ye, 2018), and the inverse Gaussian (IG) process (Peng et al., 2014; Zhuang et al., 2024). A comprehensive overview of degradation models can be found in (Ye and Xie, 2014; Si et al., 2011).

As products become more complex, scholars are increasingly delving into the degradation analysis of sophisticated systems with two or more PCs. For example, Wang et al.



(a) $t = 9$



(b) $t = 18$

Figure 2: Normal Q-Q plot for $t = 9$ and 18 of PMB data, where the black dotted line is the Student's t distribution, and the grey solid line is the normal distribution.

(2015) constructed a multi-dimensional Wiener degradation model to make a reliability analysis for fatigue crack-size (FCS) data. Lu et al. (2020) proposed a multivariate general path model to analyze degradation data involving multiple PCs. The model's applications are extended to both coating degradation data and Device-B data. Fang et al. (2022) presented a multivariate IG process with correlated random effects, and presented theoretical properties for the analysis of system failure time. A bivariate degradation model was introduced by Xu et al. (2018), where Wiener processes were employed to model the marginal degradation processes. Additionally, a shared random effect was incorporated to capture the dependence between the two degradation processes. A common stochastic time scale was introduced by

Zhai and Ye (2023) to characterize dependence stemming from the dynamic operating environment. Zheng et al. (2023) explored reliability analysis using a bivariate degradation model that accounts for random initial states and their correlation with degradation rates. Jiang et al. (2023) utilized copula functions to capture dependencies among PCs and conducted reliability analysis on a mechanical system. They also developed a predictive model for remaining useful life. In addition to the aforementioned papers, other works related to multivariate degradation processes include Fang and Pan (2023); Sun et al. (2021); Liu et al. (2021). For a comprehensive overview, please refer to Hong et al. (2018); Kang et al. (2020).

Previous research has made noteworthy strides in multivariate degradation modeling, primarily emphasizing the Wiener process owing to its favorable physical interpretability and mathematical properties. Nevertheless, dynamic fluctuations in environmental conditions can induce anomalous degradation increments in PCs, deviating from the assumption of normality, as illustrated in the motivating example in Section 1.2. Persisting with modeling the Wiener process for such degradation data may result in suboptimal fitting, given its potential challenges in accommodating the non-normal characteristics present in the degradation data. To the best of our knowledge, there is currently a scarcity of scholarly investigations that account for the heavy-tailed characteristics observed during the multiple-dependent degradation processes. To bridge this research gap, our goal is to propose a robust statistical model to characterize the degradation processes of multiple dependent PCs. Compared to the Wiener process, the tail-weighted process exhibits greater flexibility and holds significant advantages in handling and adapting to outliers, particularly in the fields of finance and economics. However, its application within the reliability domain is notably scarce, especially in characterizing the degradation paths of complex systems. Peng and Cheng (2020) first constructed a univariate Student-t model to analyze the degradation data of a single PC. Our research extends this model to a multivariate setting, enabling the simultaneous modeling of the interdependent degradation processes of multiple PCs. The main contributions and innovations of this article are as follows:

- Built upon the Student-t process, a multivariate degradation model that can account for both between-unit variability and process dependency has been proposed. The model is robust enough to accommodate tail-weighted data by allowing for the adjustment of tail heaviness through tuning the parameter of the degrees of freedom.

- The model possesses some nice statistical properties so that a tractable expectation

maximization (EM) algorithm triggered by a nonlinear least squares (NLS) step has been developed for parameter estimation. Additionally, we have designed a bootstrap approach for interval estimation.

- We have derived the system reliability function and propose a Monte Carlo (MC) method to address the computational challenges associated with its evaluation. Thorough investigations on two case studies are conducted to demonstrate the applicability of the proposed approach.

1.4. Overview

This article is structured as follows: Section 2 introduces the tail-weighted multivariate degradation model and utilizes the MC method for reliability estimation. In Section 3, we propose a two-stage parameter estimation method, and the performance of the estimation method is demonstrated through numerical simulation in Section 4, and two examples illustrating the application of the method are provided in the subsequent Section 5. Finally, our findings are summarized in Section 6.

2. Tail-weighted multivariate degradation model

2.1. A new multivariate degradation model

Consider a system that has p PCs. Let $\boldsymbol{Y}(t) = (Y_1(t), Y_2(t), \ldots, Y_p(t))'$ represent the degradation values of the p PCs of a system at time t ($t \geq 0$), where $Y_j(t)$ denotes the degradation value of the j-th PC at time t. The model considered in this paper is formulated as follows:

$$\begin{cases} Y_j(t) = \theta_j \Lambda_j(t) + \dfrac{\delta_j}{\sqrt{\tau}} W_j(\Lambda_j(t)), j = 1, \ldots, p, \\ \boldsymbol{\Theta} = (\theta_1, \theta_2, \ldots, \theta_p)' \sim \mathcal{N}_p(\boldsymbol{\eta}, \boldsymbol{\Sigma}_0/\tau), \\ \tau \sim \mathcal{G}(\nu/2, 2/\nu), \end{cases} \tag{1}$$

where θ_js are the drift parameters, δ_js denote the diffusion parameters, and $\Lambda_j(t)$s signify the time scale transformation functions, which are non-negative, monotonically increasing functions of time t. $W_i(\cdot)$s are the standard Brownian motions and are independent of each other. The vector $\boldsymbol{\Theta}$ captures the random effects assigned to the drift parameters, addressing the interdependence among the degradation processes of the p PCs. We assume $\boldsymbol{\Theta}$ follows a p-dimensional normal distribution with mean $\boldsymbol{\eta} = (\eta_1, \eta_2, \ldots, \eta_p)'$ and covariance matrix

$\boldsymbol{\Sigma}_0/\tau$. Here, $\eta_j > 0$, and $\boldsymbol{\Sigma}_0 = (\sigma_{ij})_{p\times p}$ is a positive definite matrix. We denote $\sigma_{jj} = \sigma_j^2$, is the j-th element on the diagonal of the matrix $\boldsymbol{\Sigma}_0$. The probability of θ_j being negative can be negligibly small when $\eta_j/\sigma_j \gg 0$, which is often a reasonable assumption in practice (Lu and Meeker, 1993; Peng and Cheng, 2020). The latent variable τ acts as a scale-weighted parameter, adjusting the tail flatness of the joint distribution $\boldsymbol{Y}(t)$ at time t. We assume that τ follows a gamma distribution with shape parameter $\nu/2$ and scale parameter $2/\nu$, and the probability density function (pdf) of τ is given by:

$$f_{\mathcal{G}}(\tau) = \frac{\tau^{\nu/2-1}}{\Gamma(\nu/2)(2/\nu)^{\nu/2}} \exp(-\nu\tau/2),\ \tau, \nu > 0,$$

where $\Gamma(\cdot)$ is the gamma function.

Let $\boldsymbol{\Sigma}(t) = \text{diag}\{\Lambda_1(t), \ldots, \Lambda_p(t)\}$ and $\boldsymbol{\Omega}_\delta = \text{diag}\{\delta_1^2, \ldots, \delta_p^2\}$. Based on the model (1), by integrating out $\boldsymbol{\Theta}$ and τ, the joint distribution of $\boldsymbol{Y}(t)$ at time t is

$$\boldsymbol{Y}(t) \sim \mathcal{T}_p\left(\boldsymbol{\Lambda}_\eta(t), \boldsymbol{U}(t), \nu\right), \tag{2}$$

where $\mathcal{T}_p(\cdot, \cdot, \nu)$ denotes the p-dimensional Student's t distribution with ν degrees of freedom, $\boldsymbol{\Lambda}_\eta(t) = (\eta_1\Lambda_1(t), \ldots, \eta_p\Lambda_p(t))'$ and $\boldsymbol{U}(t) = \boldsymbol{\Sigma}(t)\boldsymbol{\Sigma}_0\boldsymbol{\Sigma}(t) + \boldsymbol{\Sigma}(t)\boldsymbol{\Omega}_\delta$, It is worth noting that when $\nu = 1$, the Student's t distribution becomes a multivariate Cauchy distribution, and when $\nu \to \infty$, the Student's t distribution reduces to a multivariate normal distribution.

Assume that we take m measurements for the j-th PC, and the measurement time $\boldsymbol{t}_{(m)} = (t_1, \ldots, t_m)'$. The corresponding degradation values of the j-th PC are denoted as $Y_j(\boldsymbol{t}_{(m)}) = (Y_j(t_1), \ldots, Y_j(t_m))'$. By integrating out of θ_j and τ, the joint distribution of $Y_j(\boldsymbol{t}_{(m)})$ is

$$Y_j(\boldsymbol{t}_{(m)}) \sim \mathcal{T}_m\left(\eta_j\Lambda_j(\boldsymbol{t}_{(m)}), V_j(\boldsymbol{t}_{(m)}), \nu\right), \tag{3}$$

where $\Lambda_j(\boldsymbol{t}_{(m)}) = (\Lambda_j(t_1), \ldots, \Lambda_j(t_m))'$, $V_j(\boldsymbol{t}_{(m)}) = \sigma_j^2\Lambda_j(\boldsymbol{t}_{(m)})\Lambda_j(\boldsymbol{t}_{(m)})' + \delta_j^2 Q(\boldsymbol{t}_{(m)})$ and $Q(\boldsymbol{t}_{(m)}) = [\min\{\Lambda_j(t_{s_1}), \Lambda_j(t_{s_2})\}]_{1\le s_1, s_2 \le m}$. The mean and variance of the degradation process $Y_j(t)$ are $\eta_j\Lambda_j(t)$, and $\nu/(\nu-2)V_j(t)$, respectively. Thus, at any given time t, the Pearson correlation coefficient between $Y_{j_1}(t)$ and $Y_{j_2}(t)$ can be expressed as

$$\rho\left[Y_{j_1}(t), Y_{j_2}(t)\right] = \rho_{j_1j_2}(t) = \frac{\sigma_{j_1j_2}}{\sqrt{\sigma_{j_1}^2 + \delta_{j_1}^2/\Lambda_{j_1}(t)}\sqrt{\sigma_{j_2}^2 + \delta_{j_2}^2/\Lambda_{j_2}(t)}}, \tag{4}$$

where $\sigma_{j_1j_2} = \rho_{j_1j_2}\sigma_{j_1}\sigma_{j_2}$. This equation reveals the dynamic correlation between any two PCs. Initially, there is no correlation between the components (i.e., $\rho_{j_1j_2}(0) = 0$), but as

time progresses and degradation advances, the correlation gradually strengthens, eventually approaching a long-term stable value $\rho_{j_1 j_2}(\infty) = \rho_{j_1 j_2}$.

Remark 1: Based on equations (2) and (3), it can be observed that model (1) possesses the capability to detect outliers, both within a single PC and across multiple PCs. The parameter ν serves as a control parameter that determines the probability of outlier presence when monitoring the degradation of these PCs. Varying values of ν allow for adjustment in the sensitivity of the model towards outliers. A smaller value of ν enhances robustness against outliers due to heavier tails in Student's t distribution. Consequently, even if there are outliers present in the data, reasonable estimates for model parameters can still be obtained. Conversely, a larger value of ν reduces sensitivity towards outliers and focuses more on modeling normal behavior within the data. This may result in ineffective capture and modeling of outliers, leading to inaccurate or misleading outcomes. Therefore, careful selection of ν enables a balance between robustness against outliers while maintaining an appropriate level of sensitivity towards normal behavior. This facilitates effective monitoring and analysis of degradation patterns across multiple PCs by considering both normal and outlier behaviors. This paper, based on data-driven methods, uses a two-stage statistical inference approach to estimate model parameters and the degrees of freedom ν, as detailed in Section 3.

2.2. Reliability analysis

Let ω_j denote the failure threshold level for the j-th PC. Consequently, the lifetime of the j-th PC is defined as the time at which its degradation first surpasses the failure threshold (Di Nardo et al., 2001). That is,

$$T_j = \inf\{t : Y_j(t) \geq \omega_j\}.$$

Given θ_j and τ, $Y_j(t)$ follows a Wiener process. As a result, the lifetime T_j conforms to a transformed inverse Gaussian distribution, specifically, $\Lambda_j(T_j) \sim \mathcal{IG}(\omega_j/\theta_j, \omega_j^2\sqrt{\tau}/\delta_j)$. The conditional pdf of T_j is expressed as

$$f_j(t|\theta_j, \tau) = \frac{\omega_j}{\sqrt{2\pi\delta_j^2\Lambda_j^3(t)/\tau}} \exp\left\{-\frac{(\omega_j - \theta_j\Lambda_j(t))^2\tau}{2\delta_j^2\Lambda_j(t)}\right\} \frac{\mathrm{d}\Lambda_j(t)}{\mathrm{d}t}. \tag{5}$$

The joint pdf for $T_1, T_2, \ldots, T_p$ can then be calculated as:

$$f(t_1, t_2, ..., t_p) = \int\int \prod_{j=1}^{p} f_j(t_j|\theta_j, \tau) f(\boldsymbol{\Theta}|\tau) f(\tau) \mathrm{d}\boldsymbol{\Theta} \mathrm{d}\tau. \tag{6}$$

We assume a competitive relationship among the p PCs, considering the system to have failed when any PC reaches the failure threshold level. Thus, the lifetime of the system is defined as:

$$T_\omega = \inf\{t : Y_1(t) \geq \omega_1 \text{ or } \cdots \text{ or } Y_p(t) \geq \omega_p\} = \min\{T_1, \ldots, T_p\}.$$

Then the reliability of the system at time t is

$$\begin{aligned} R_{T_\omega}(t) &= P\{T_\omega > t\} = P\{T_1 > t, \ldots, T_p > t\} \\ &= \int_t^{+\infty} \cdots \int_t^{+\infty} f(t_1, t_2, ..., t_p) \mathrm{d}t_1 \cdots \mathrm{d}t_p. \end{aligned} \tag{7}$$

Considering the computational challenges associated with evaluating the integral in (7), we employ an MC method to estimate $R_{T_\omega}(t)$. In essence, the MC approach involves generating random samples from the distribution of interest, in this case, the joint distribution of $T_1, T_2, \ldots, T_p$ as defined in (6). Then we can find the lifetime of the system, and further estimate the reliability function $R_{T_\omega}(t)$. The procedure of the MC method can be implemented through Algorithm 1.

Algorithm 1: Reliability function estimation by MC approach.

Input: $t, \nu, \boldsymbol{\Sigma}_0, \eta_j, \omega_j, \Lambda_j(t)$, and $\delta_j, j = 1, \ldots, p$.

Output: $R_{T_\omega}(t)$.

for $q = 1$ **to** Q **do**

- Generate τ^* from $\mathcal{G}(\nu/2, 2/\nu)$;
- Generate $\boldsymbol{\Theta}^* = \left(\theta_1^*, \ldots, \theta_p^*\right)'$ following $\mathcal{N}_p(\boldsymbol{\eta}, \boldsymbol{\Sigma}_0/\tau^*)$;
- Given the generated θ_j^* and τ^*, generate T_j from (5), denoted as T_j^*;
- Obtain $T_{\omega q}^* = \min\left\{T_1^*, \ldots, T_p^*\right\}$.

end

Estimate $R_{T_\omega}(t)$ by $\sum_{q=1}^{Q} I_{\{T_{\omega q}^* \geq t\}}/Q$, where $I_{\{\cdot\}}$ denotes the indicator function.

This MC methodology provides a computationally feasible solution to the analytical integration complexity encountered in this context. By circumventing the need for explicit analytical solutions, we leverage statistical sampling to gain insights into the system's reliability, making the computational process more manageable and adaptable to complex models.

3. Statistical inference

Assuming that there are n systems involved in an experiment, the degradation of each system's PCs is measured at different time points: $t_{i,1}, \ldots, t_{i,m_i}$, and the degradation values of the p PCs at time $t_{i,k}$ are denoted as $\boldsymbol{Y}_{i,k} = (Y_{i,1,k}, \ldots, Y_{i,p,k})'$, where $i = 1, \ldots, n$ and $k = 1, \ldots, m_i$. Let $\Delta\boldsymbol{Y}_{i,k} = \boldsymbol{Y}_{i,k} - \boldsymbol{Y}_{i,k-1}$, where $t_{i,0} = 0$ and $\boldsymbol{Y}_{i,0} = \boldsymbol{0}$. Based on model (1), we know that

$$\begin{cases} \Delta\boldsymbol{Y}_{i,k}|\boldsymbol{\Theta}_i, \tau_i \sim \mathcal{N}_p\left(\Delta\boldsymbol{\Sigma}(t_{i,k})\boldsymbol{\Theta}_i, \dfrac{\boldsymbol{\Omega}_\delta}{\tau_i}\Delta\boldsymbol{\Sigma}(t_{i,k})\right), \\ \boldsymbol{\Theta}_i \sim \mathcal{N}_p(\boldsymbol{\eta}, \boldsymbol{\Sigma}_0/\tau_i), \\ \tau_i \sim \mathcal{G}(\nu/2, 2/\nu), \end{cases} \tag{8}$$

where $\Delta\boldsymbol{\Sigma}(t_{i,k}) = \boldsymbol{\Sigma}(t_{i,k}) - \boldsymbol{\Sigma}(t_{i,k-1})$. For the j-th time scale transformation function $\Lambda_j(t)$, we assume a parametric form with an unknown parameter γ_j, represented as $\Lambda_j(t;\gamma_j)$. The choice of the specific form for $\Lambda_j(t;\gamma_j)$ can be determined based on engineering experience or empirical investigation. Commonly used forms include the power law function $\Lambda_j(t;\gamma_j) = t^{\gamma_j}$ and the exponential function $\Lambda_j(t;\gamma_j) = \exp(\gamma_j t) - 1$. Let $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, ..., \gamma_p)'$. Then the model parameters are $\boldsymbol{\Phi} = (\boldsymbol{\eta}, \boldsymbol{\Omega}_\delta, \boldsymbol{\Sigma}_0, \boldsymbol{\gamma}, \nu)$. Let $\Delta\boldsymbol{Y}_i = \{\Delta\boldsymbol{Y}_{i,k}, k = 1, \ldots, m_i\}$ and $\mathbb{Y} = \{\Delta\boldsymbol{Y}_i, i = 1, \ldots, n\}$. Then the likelihood function of $\boldsymbol{\Phi}$ is expressed as:

$$\begin{aligned} \ell(\mathbb{Y}|\boldsymbol{\Phi}) = \prod_{i=1}^{n} \int\int \Big[\prod_{k=1}^{m_i} & \frac{\tau_i^{p/2}}{(2\pi)^{p/2}\left|\boldsymbol{\Omega}_\delta\Delta\boldsymbol{\Sigma}(t_{i,k})\right|^{1/2}} \\ & \times \exp\left\{-\frac{\tau_i}{2}(\Delta\boldsymbol{Y}_{i,k} - \Delta\boldsymbol{\Sigma}(t_{i,k})\boldsymbol{\Theta}_i)'\left(\boldsymbol{\Omega}_\delta\Delta\boldsymbol{\Sigma}(t_{i,k})\right)^{-1}(\Delta\boldsymbol{Y}_{i,k} - \Delta\boldsymbol{\Sigma}(t_{i,k}\boldsymbol{\Theta}_i)\right\}\Big] \\ & \times \frac{\tau_i^{p/2}}{(2\pi)^{p/2}|\boldsymbol{\Sigma}_0|^{1/2}} \exp\left\{-\frac{\tau_i}{2}(\boldsymbol{\Theta}_i - \boldsymbol{\eta})'\boldsymbol{\Sigma}_0^{-1}(\boldsymbol{\Theta}_i - \boldsymbol{\eta})\right\} \\ & \times \frac{\tau_i^{\nu/2-1}}{\Gamma(\nu/2)(2/\nu)^{\nu/2}} \exp\left\{-\frac{\nu}{2}\tau_i\right\} \mathrm{d}\boldsymbol{\Theta}_i \mathrm{d}\tau_i. \end{aligned} \tag{9}$$

Given the complexity of (9) and its involvement in intractable integration, the direct maximization of $\log \ell(\mathbb{Y}|\boldsymbol{\Phi})$ to obtain the estimate of $\boldsymbol{\Phi}$ proves to be a challenging task. In light of this complexity, the EM algorithm stands out as a powerful tool for handling high-dimensional parameter estimation problems. By alternating between the E-step and M-step, one can manage latent variables and optimize complex likelihood functions (Dempster et al., 1977; Wu, 1983). It is worth noting that the parameters in this model typically manifest

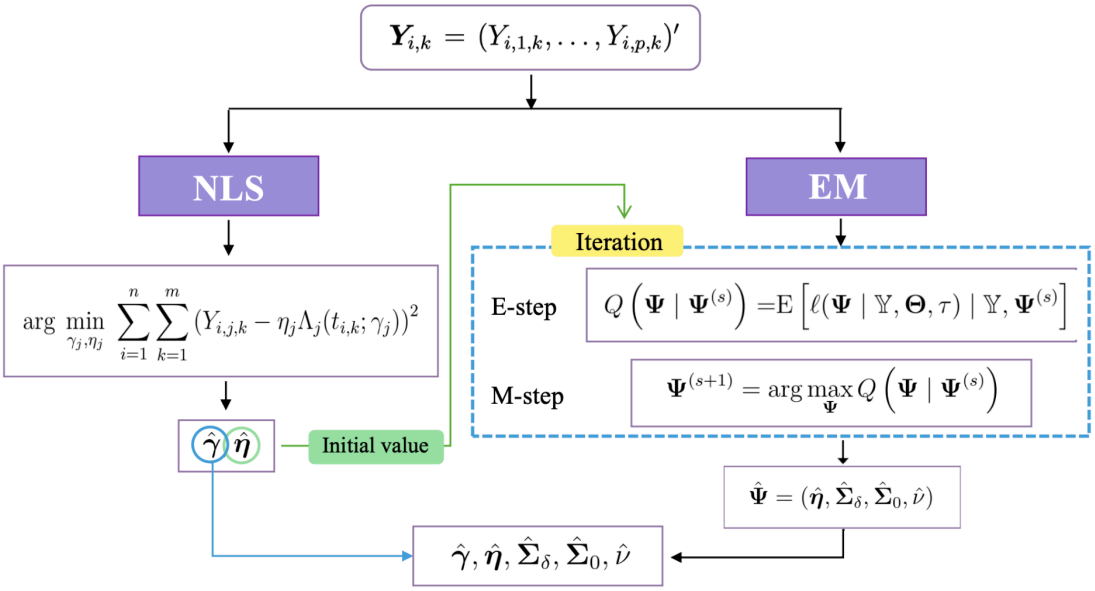


Figure 3: Proposed two-stage algorithm for model parameter estimation.

as vectors and matrices, introducing additional intricacies into the estimation process. Furthermore, managing parameters associated with the time-scale transformation function $\Lambda_j(t)$ during iterations may present specific challenges.

Consequently, we propose a novel two-stage parameter estimation method to solve this problem, as illustrated in Figure 3. Based on the data $\mathbb{Y}$, we initially employ the NLS method to estimate the parameters $\boldsymbol{\gamma}$ and $\boldsymbol{\eta}$ (see Section 3.1 for details). Here, $\boldsymbol{\gamma}$ serves as the final estimate, while $\boldsymbol{\eta}$ can be utilized as the initial value for the subsequent steps. For the estimation of other parameters $\boldsymbol{\Psi} = (\boldsymbol{\eta}, \boldsymbol{\Sigma}_\delta, \boldsymbol{\Sigma}_0, \nu)$, we utilize the EM algorithm, considering $\boldsymbol{\Theta}_i$ and τ_i as missing data, achieved through iterative execution of the E-step and M-step until convergence is reached (Wan and Bai, 2024), as detailed in Section 3.2. Ultimately, following the two-stage algorithm, we obtain point estimates for all parameters in the proposed model.

3.1. Nonlinear least squares estimation

NLS is a statistical method used for estimating the parameters of nonlinear regression models. The estimation is obtained by minimizing the sum of squared residuals (SSR), where residuals are the differences between the observed values and the values predicted by the model. In our study, $E[Y_j(t)] = \eta_j\Lambda_j(t;\gamma_j)$, which indicates that the relationship between the degradation values of the j-th PC and time t. Given the observed data $\{\boldsymbol{Y}_{i,k}, i =$

$1, \ldots, n, k = 1, \ldots, m_i\}$, we define the SSR for the j-th PC as follows:

$$SSR_j = \sum_{i=1}^{n} \sum_{k=1}^{m} \left(Y_{i,j,k} - \eta_j \Lambda_j(t_{i,k}; \gamma_j)\right)^2, j = 1, \ldots, p. \tag{10}$$

The estimate $(\hat{\gamma}_j, \hat{\eta}_j)$ can be obtained by minimizing SSR_j, that is,

$$(\hat{\gamma}_j, \hat{\eta}_j) = \arg \min_{\gamma_j, \eta_j} SSR_j. \tag{11}$$

Estimates $(\hat{\gamma}_j, \hat{\eta}_j)$ can be derived using optimization techniques like Gauss-Newton, Levenberg-Marquardt, or gradient-based methods, which are favored for their rapid convergence, and suitability for complex, high-dimensional nonlinear optimization problems. Once these estimates $\{(\hat{\gamma}_j, \hat{\eta}_j), j = 1, \ldots, p\}$ are obtained, $\Lambda_j(t; \hat{\gamma}_j)$ can be treated as a known function, and $\hat{\eta}_j$ can serve as an initial value in the EM algorithm. To streamline the discussion, we will use the notation $\Lambda_j(t)$ to represent $\Lambda_j(t, \hat{\gamma}_j)$ in the subsequent subsections.

3.2. EM algorithm

To alleviate the complexity of integration in (9), we consider $(\boldsymbol{\Theta}, \boldsymbol{\tau}) = \{\boldsymbol{\Theta}_i, \tau_i, i = 1, \ldots, n\}$ as the missing data. Let $\Delta\Lambda_j(t_{i,k}) = \Lambda_j(t_{i,k}) - \Lambda_j(t_{i,k-1})$. Based on the complete data $\{\mathbb{Y}, \boldsymbol{\Theta}, \boldsymbol{\tau}\}$, the log-likelihood function of $\boldsymbol{\Psi}$ is

$$\begin{aligned} \ell(\mathbb{Y}, \boldsymbol{\Theta}, \boldsymbol{\tau} | \boldsymbol{\Psi}) = \sum_{i=1}^{n} \Bigg\{ \ell_c + \Big(\frac{(m_i + 1)p + \nu}{2} - 1 \Big) \ln \tau_i - m_i \sum_{j=1}^{p} \ln \delta_j \\ - \frac{1}{2} \sum_{j=1}^{p} \sum_{k=1}^{m_i} \ln \Delta\Lambda_j(t_{i,k}) \Bigg\} - \frac{1}{2} \sum_{i=1}^{n} \tau_i \left(\sum_{k=1}^{m_i} \ell_{i,k} + \ell_{i,0} + \nu \right), \end{aligned} \tag{12}$$

where

$$\begin{aligned} \ell_c &= -\frac{(m_i + 1)p}{2} \ln 2\pi - \frac{1}{2} \ln |\boldsymbol{\Sigma}_0| - \ln \Gamma(\frac{\nu}{2}) + \frac{\nu}{2} \ln(\frac{\nu}{2}), \\ \ell_{i,0} &= (\boldsymbol{\Theta}_i - \boldsymbol{\eta})' \boldsymbol{\Sigma}_0^{-1} (\boldsymbol{\Theta}_i - \boldsymbol{\eta}), \\ \ell_{i,k} &= (\Delta \boldsymbol{Y}_{i,k} - \Delta\boldsymbol{\Sigma}(t_{i,k}) \boldsymbol{\Theta}_i)' \big(\boldsymbol{\Omega}_\delta \Delta\boldsymbol{\Sigma}(t_{i,k}) \big)^{-1} (\Delta \boldsymbol{Y}_{i,k} - \Delta\boldsymbol{\Sigma}(t_{i,k}) \boldsymbol{\Theta}_i). \end{aligned}$$

The EM algorithm acquires the estimate for $\boldsymbol{\Psi}$ through iterative execution of the E-step and M-step until convergence is achieved. Assuming that the optimal solution in the M-step at the s-th iteration is denoted as $\boldsymbol{\Psi}^{(s)}$, in the subsequent $(s + 1)$-th iteration, the E-step

involves computing the following Q-function:

$$
\begin{aligned}
Q\left(\boldsymbol{\Psi} \mid \boldsymbol{\Psi}^{(s)}\right) =& \mathrm{E}\left[\ell(\boldsymbol{\Psi} \mid \mathbb{Y}, \boldsymbol{\Theta}, \tau) \mid \mathbb{Y}, \boldsymbol{\Psi}^{(s)}\right] \\
=& \sum_{i=1}^{n}\left\{\ell_c + \left(\frac{(m_i+1)p+v}{2} - 1\right)\mathrm{E}\left[\ln \tau_i | \Delta \boldsymbol{Y}_i, \boldsymbol{\Psi}^{(s)}\right] - m_i \sum_{j=1}^{p} \ln \delta_j \right. \\
& \left. - \frac{1}{2}\sum_{j=1}^{p}\sum_{k=1}^{m_i} \ln \Delta\Lambda_j(t_{i,k})\right\} - \frac{1}{2}\sum_{i=1}^{n}\left\{\sum_{k=1}^{m_i}\mathrm{E}\left[\tau_i \ell_{i,k} | \Delta \boldsymbol{Y}_i, \boldsymbol{\Psi}^{(s)}\right]\right. \\
& \left. + \mathrm{E}\left[\tau_i \ell_{i,0} | \Delta \boldsymbol{Y}_i, \boldsymbol{\Psi}^{(s)}\right] + \nu \mathrm{E}[\tau_i | \Delta \boldsymbol{Y}_i, \boldsymbol{\Psi}^{(s)}]\right\},
\end{aligned} \tag{13}
$$

where the expectations are calculated concerning the conditional distribution $f(\boldsymbol{\Theta}_i, \tau_i \mid \Delta \boldsymbol{Y}_i, \boldsymbol{\Psi}^{(s)})$. (13) involves four conditional expectations: $\mathrm{E}[\ln \tau_i | \Delta \boldsymbol{Y}_i, \boldsymbol{\Psi}^{(s)}]$, $\mathrm{E}[\tau_i \ell_{i,k} | \Delta \boldsymbol{Y}_i, \boldsymbol{\Psi}^{(s)}]$, $\mathrm{E}[\tau_i \ell_{i,0} | \Delta \boldsymbol{Y}_i, \boldsymbol{\Psi}^{(s)}]$, and $\mathrm{E}[\tau_i | \Delta \boldsymbol{Y}_i, \boldsymbol{\Psi}^{(s)}]$. We need the following results to derive the conditional expectations.

Theorem 1. *Given $\Delta \boldsymbol{Y}_i$, the conditional distribution of $\boldsymbol{\Theta}_i$ and τ_i can be decomposed into the following two distributions.*

(a) $\boldsymbol{\Theta}_i | \Delta \boldsymbol{Y}_i, \tau_i \sim \mathcal{N}_p\left(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_{\boldsymbol{\Theta}_i}/\tau_i\right)$, *where* $\boldsymbol{\Sigma}_{\boldsymbol{\Theta}_i} = \left[\boldsymbol{\Sigma}_0^{-1} + \boldsymbol{\Omega}_\delta^{-1}\boldsymbol{\Sigma}(t_{im_i})\right]^{-1}$ *and* $\boldsymbol{\mu}_i = \boldsymbol{\Sigma}_{\boldsymbol{\Theta}_i}(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\eta} + \boldsymbol{\Omega}_\delta^{-1}\boldsymbol{Y}_{i,m_i})$.

(b) $\tau_i | \Delta \boldsymbol{Y}_i \sim \mathcal{G}\left(\frac{m_i p+\nu}{2}, \frac{2}{K_{i,1}-K_{i,2}+\nu}\right)$, *where* $K_{i,1} = \sum_{k=1}^{m_i} \Delta \boldsymbol{Y}_{i,k}' \boldsymbol{\Omega}_\delta^{-1} \Delta \boldsymbol{\Sigma}^{-1}(t_{i,k}) \Delta \boldsymbol{Y}_{i,k} + \boldsymbol{\eta}^\top \boldsymbol{\Sigma}_0^{-1} \boldsymbol{\eta}$ *and* $K_{i,2} = \boldsymbol{\mu}_i' \boldsymbol{\Sigma}_{\boldsymbol{\Theta}_i}^{-1} \boldsymbol{\mu}_i$.

The proof of Theorem 1 is provided in the Appendix A. From Theorem 1 (b), $\mathrm{E}(\tau_i | \Delta \boldsymbol{Y}_i, \boldsymbol{\Psi}^{(s)})$ and $\mathrm{E}(\ln \tau_i | \Delta \boldsymbol{Y}_i, \boldsymbol{\Psi}^{(s)})$ used in the E-step can be easily computed, which are

$$
\begin{aligned}
\mathrm{E}(\tau_i | \Delta \boldsymbol{Y}_i, \boldsymbol{\Psi}^{(s)}) &= \frac{m_i p + \nu^{(s)}}{K_{i,1}^{(s)} - K_{i,2}^{(s)} + \nu^{(s)}}, \\
\mathrm{E}(\ln \tau_i | \Delta \boldsymbol{Y}_i, \boldsymbol{\Psi}^{(s)}) &= \psi\left(\frac{m_i p + \nu^{(s)}}{2}\right) - \ln\left(\frac{K_{i,1}^{(s)} - K_{i,2}^{(s)} + \nu^{(s)}}{2}\right),
\end{aligned} \tag{14}
$$

where $\psi(x) = \mathrm{d} \ln \Gamma(x)/\mathrm{d}x$ represents the digamma function, $K_{i,1}^{(s)}$ and $K_{i,2}^{(s)}$ are $K_{i,1}$ and $K_{i,2}$ with the parameters $\boldsymbol{\Psi}$ substituted by $\boldsymbol{\Psi}^{(s)}$.

Theorem 2. *Given the joint distributions of $\boldsymbol{\Theta}_i$ and τ_i in Theorem 1, if the solution in the M-step at the s-th iteration is $\boldsymbol{\Psi}^{(s)}$, then*

$$
\mathrm{E}\left(\tau_i \ell_{i,0} | \Delta \boldsymbol{Y}_i, \boldsymbol{\Psi}^{(s)}\right) = \mathrm{tr}\left(\boldsymbol{\Sigma}_0^{-1} \boldsymbol{\Sigma}_{\boldsymbol{\Theta}_i}^{(s)}\right) + \mathrm{E}\left(\tau_i | \Delta \boldsymbol{Y}_i, \boldsymbol{\Psi}^{(s)}\right)\left(\boldsymbol{\mu}_i^{(s)} - \boldsymbol{\eta}\right)' \boldsymbol{\Sigma}_0^{-1}\left(\boldsymbol{\mu}_i^{(s)} - \boldsymbol{\eta}\right),
$$

$$\mathrm{E}\left(\tau_i \ell_{i,k} | \Delta \boldsymbol{Y}_i, \boldsymbol{\Psi}^{(s)}\right) = \mathrm{tr}\left(\Delta\boldsymbol{\Sigma}(t_{i,k})\boldsymbol{\Omega}_\delta^{-1}\boldsymbol{\Sigma}_{\boldsymbol{\Theta}_i}^{(s)}\right) + \mathrm{E}\left(\tau_i | \Delta \boldsymbol{Y}_i, \boldsymbol{\Psi}^{(s)}\right)$$
$$\times \left(\boldsymbol{\mu}_i^{(s)} - \Delta\boldsymbol{\Sigma}^{-1}(t_{i,k})\Delta\boldsymbol{Y}_{i,k}\right)' (\boldsymbol{\Omega}_\delta \Delta\boldsymbol{\Sigma}^{-1}(t_{i,k}))^{-1} \left(\boldsymbol{\mu}_i^{(s)} - \Delta\boldsymbol{\Sigma}^{-1}(t_{i,k})\Delta\boldsymbol{Y}_{i,k}\right).$$

The proof of Theorem 2 is given in the Appendix B. Given the results in theorems 1 and 2, the Q-function can be completely determined. Then we update the optimal solution in the M-step at the $(s+1)$-th iteration as

$$\boldsymbol{\Psi}^{(s+1)} = \arg\max_{\boldsymbol{\Psi}} Q\left(\boldsymbol{\Psi} \mid \boldsymbol{\Psi}^{(s)}\right). \tag{15}$$

This can be accomplished by taking the partial derivatives of $Q\left(\boldsymbol{\Psi} \mid \boldsymbol{\Psi}^{(s)}\right)$ with respect to the parameters, and solving these equations.

Theorem 3. *Given the solution in the M-step at the s-th iteration is $\boldsymbol{\Psi}^{(s)}$, the solution of (15) can be updated as follows:*

$$\boldsymbol{\eta}^{(s+1)} = \frac{\sum_{i=1}^n \boldsymbol{\mu}_i^{(s)} \mathrm{E}\left(\tau_i | \boldsymbol{\Delta Y}_i, \boldsymbol{\Psi}^{(s)}\right)}{\sum_{i=1}^n \mathrm{E}\left(\tau_i | \boldsymbol{\Delta Y}_i, \boldsymbol{\Psi}^{(s)}\right)},$$
$$\boldsymbol{\Sigma}_0^{(s+1)} = \frac{\sum_{i=1}^n \left[\boldsymbol{\Sigma}_{\boldsymbol{\Theta}_i}^{(s)} + \mathrm{E}(\tau_i | \Delta\boldsymbol{Y}_i, \boldsymbol{\Psi}^{(s)})(\boldsymbol{\mu}_i^{(s)} - \boldsymbol{\eta}^{(s+1)})(\boldsymbol{\mu}_i^{(s)} - \boldsymbol{\eta}^{(s+1)})'\right]}{n},$$
$$\boldsymbol{\Omega}_\delta^{(s+1)} = \frac{1}{\sum_{i=1}^n m_i} \sum_{i=1}^n \sum_{k=1}^{m_i} \Big[\Delta\boldsymbol{\Sigma}(t_{i,k})\boldsymbol{\Sigma}_{\boldsymbol{\Theta}_i}^{(s)} + \mathrm{E}(\tau_i | \Delta\boldsymbol{Y}_i, \boldsymbol{\Psi}^{(s)})$$
$$\times (\boldsymbol{\mu}_i^{(s)} - \Delta\boldsymbol{\Sigma}^{-1}(t_{i,k})\Delta\boldsymbol{Y}_{i,k})(\boldsymbol{\mu}_i^{(s)} - \Delta\boldsymbol{\Sigma}^{-1}(t_{i,k})\Delta\boldsymbol{Y}_{i,k})'\Big].$$

The proof of Theorem 3 is given in the Appendix C. Besides, the update of ν can be implemented by maximizing the following function:

$$-2\ln\Gamma(\nu/2) + \nu\ln(\nu/2) + \frac{\nu}{n}\sum_{i=1}^n \left[\mathrm{E}\left(\ln\tau_i | \boldsymbol{\Delta Y}_i, \boldsymbol{\Psi}^{(s)}\right) - \mathrm{E}\left(\tau_i | \boldsymbol{\Delta Y}_i, \boldsymbol{\Psi}^{(s)}\right)\right]. \tag{16}$$

Given the initial values $\boldsymbol{\Psi}^{(0)}$, the EM algorithm can then be executed until convergence according to a specified criterion. Consequently, the ML estimate for $\boldsymbol{\Psi}$ can be obtained. Formally, the EM algorithm in our study can be implemented through algorithm 2.

3.3. Interval estimation

In addition to point estimation, there exists a significant interest in establishing interval estimates for the model parameters $\boldsymbol{\Psi}$. These interval estimates take into account

Algorithm 2: Implementation of the proposed EM algorithm.

Input: $\mathbb{Y}$,$\boldsymbol{\Psi}^{(0)}$, ϵ.

Output: $\hat{\boldsymbol{\Psi}} = \{\hat{\boldsymbol{\eta}}, \hat{\boldsymbol{\Sigma}}_\delta, \hat{\boldsymbol{\Sigma}}_0, \hat{\nu})$.

while $| \, | \, \boldsymbol{\Psi}^{(s+1)} - \boldsymbol{\Psi}^{(s)} \, | \, | \geq \epsilon$ **do**

E-step:

Compute $\mathrm{E}[\tau_i|\Delta \boldsymbol{Y}_i, \boldsymbol{\Psi}^{(s)}]$ and $\mathrm{E}\left[\ln \tau_i|\Delta \boldsymbol{Y}_i, \boldsymbol{\Psi}^{(s)}\right]$ by (14);

Compute $\mathrm{E}\left[\tau_i \ell_{i,0}|\Delta \boldsymbol{Y}_i, \boldsymbol{\Psi}^{(s)}\right]$ and $\mathrm{E}\left[\tau_i \ell_{i,k}|\Delta \boldsymbol{Y}_i, \boldsymbol{\Psi}^{(s)}\right]$ by Theorem 2;

M-step:

Update $\boldsymbol{\Psi}^{(s+1)}$ by Theorem 3 and (16).

end

uncertainties and variability in the estimation process, providing a plausible range for the parameters (Luo et al., 2020). This is particularly important as it allows the practitioners to assess the precision and reliability of their parameter estimates. To construct these interval estimations, it is common practice to utilize asymptotic theories. However, due to the complexity involved in evaluating the Fisher information matrix for the proposed model, we instead adopt the bootstrap method to obtain reliable interval estimates. The outline of the bootstrap procedure is provided in Algorithm 3. After obtaining the B bootstrap estimates $\left\{\hat{\boldsymbol{\Psi}}^*_1, \ldots, \hat{\boldsymbol{\Psi}}^*_B\right\}$, we can proceed to construct an approximate $100(1-\alpha)\%$ bootstrap confidence interval for a function of the parameters $h(\boldsymbol{\Psi})$. The interval estimation is constructed as follows:

$$\left[h\left(\hat{\boldsymbol{\Psi}}^*\right)_{(\alpha B/2)}, h\left(\hat{\boldsymbol{\Psi}}^*\right)_{((1-\alpha/2)B)}\right],$$

where $h\left(\hat{\boldsymbol{\Psi}}^*\right)_{(b)}$ denotes the b-th order statistic among $\left\{h\left(\hat{\boldsymbol{\Psi}}^*\right)_1, \ldots, h\left(\hat{\boldsymbol{\Psi}}^*\right)_B\right\}$.

4. Simulation studies

This section evaluates the estimation performance of model parameters, correlation coefficients, and reliability through numerical simulations. We consider degradation with two or three PCs (i.e., $p = 2$ or 3). Both linear degradation paths (i.e., $\Lambda(t) = t$) and nonlinear degradation paths (i.e., $\Lambda(t) = t^\gamma$) are examined. Four combinations of p and $\Lambda(t)$ as outlined in Table 1, which also includes the corresponding parameter values. To characterize the heavy-tail features of the degradation values, we set the degrees of freedom

Algorithm 3: Bootstrap algorithm procedure.

Input: Point estimate $\hat{\boldsymbol{\Psi}}$ and $\hat{\boldsymbol{\gamma}}$.

Output: B resamples of the estimate $\left\{\hat{\boldsymbol{\Psi}}_1^*, \ldots, \hat{\boldsymbol{\Psi}}_B^*\right\}$.

for $b = 1$ **to** B **do**
  for $i = 1$ **to** n **do**
    Generate $\tilde{\tau}_i$ from $\mathcal{G}\left(\hat{v}/2, 2/\hat{v}\right)$;
    Generate $\tilde{\boldsymbol{\Theta}}$ from $\mathcal{N}_p(\hat{\boldsymbol{\eta}}, \hat{\boldsymbol{\Sigma}}_0/\tilde{\boldsymbol{\tau}})$;
    for $k = 1$ **to** m_i **do**
      Given $\tilde{\boldsymbol{\Theta}}_i$ and $\tilde{\tau}_i$, generate $\Delta\tilde{\boldsymbol{Y}}_{i,k}$ from $\mathcal{N}_p\left(\Delta\boldsymbol{\Sigma}(t_{i,k})\tilde{\boldsymbol{\Theta}}_i, \frac{\boldsymbol{\Omega}_\delta}{\tilde{\tau}_i}\Delta\boldsymbol{\Sigma}(t_{i,k})\right)$;
    end
  end
  Obtain the bootstrapped degradation data $\widetilde{\mathbb{Y}}$;
  Obtain $\hat{\boldsymbol{\Psi}}_b^*$ based on $\tilde{\mathbb{Y}}$ using the proposed EM algorithm.
end

to $\nu = 5$. Assuming periodic measurements of degradation for n units at $t = 5, 10, ..., 5m_i$, where all m_i are uniform ($m_1 = \cdots = m_n = m$), we assess the influence of sample size on inference with variations in $n = 10, 20, 30$ and the number of observations $m = 10, 20, 30$. For each setting, we conduct 1000 replications of data generated from the simulated model and fit them using the EM algorithm. The stopping criterion for iterations is set with an error tolerance of $\epsilon = 10^{-5}$. The simulations run on a notebook computer with an Intel Core i7 processor operating at 2.3 GHz and 16GB of RAM under the Windows 11 operating system. The computation time for parameter estimation in all sample size combinations is within one minute, demonstrating satisfactory performance.

4.1. Performance evaluation of model parameters

To evaluate the overall quality of the EM estimation results, we calculate the root mean squared error (RMSE) for different combinations of (n, m). Due to space constraints and the consistency of conclusions, we provide scatter plots for scenarios I and III (with $p = 2$), illustrating results for various combinations of $\Lambda(t)$ and (n, m). Additional results for $p = 3$ are available in the Supplementary Section S1. These plots reveal minimal bias in all parameters. Moreover, as n increases with fixed m, RMSE decreases, and vice versa.

Table 1: Four combinations of p and $\Lambda(t)$, along with their corresponding parameter setting.

Scen.	$\Lambda(t)$	p	η_1	η_2	η_3	δ_1^2	δ_2^2	δ_3^2	σ_1^2	σ_2^2	σ_3^2	σ_{12}	σ_{13}	σ_{23}	γ_1	γ_2	γ_3
I	t	2	11	12	-	0.5	1.5	-	1	2	-	0.75	-	-	-	-	-
II		3	11	12	13	0.4	0.6	0.8	1	2.25	4	0.75	-1.0	1.2	-	-	-
III	t^γ	2	11	12	-	0.5	1.5	-	1	2	-	0.75	-	-	1.1	1.2	-
IV		3	11	12	13	0.4	0.6	0.8	1	2.25	4	0.75	-1.0	1.2	0.8	1	1.2

This trend suggests that the model exhibits satisfactory fitting performance, particularly for moderate values of n and m.

Validating the interval estimation performance of the bootstrap procedure as detailed in Section 4 with $B = 1000$, we implement it on every synthetic dataset generated in the preceding simulations. Figure 5 illustrates the coverage probability (CP) of the 95% confidence intervals (CIs) for the model parameters by the bootstrap in scenarios I and III, where the dashed line represents the 95% horizontal line. Notably, estimations for $\{\boldsymbol{\eta}, \boldsymbol{\delta}, \boldsymbol{\gamma}\}$ demonstrate excellent performance, consistently achieving CP values exceeding 90% across all combinations. While CP for $\boldsymbol{\sigma}$ may experience a slight reduction due to sampling randomness and the intricacies of matrix computations, the results remain within acceptable ranges. Notably, as both n and m increase, the CP for all parameters shows a continuous improvement.

4.2. Performance evaluation of correlation coefficients

Furthermore, we evaluate the performance of the proposed statistical inference methods in estimating correlation coefficients. We fix $m = 20$ to examine how the evaluation of correlation coefficient performance varies with different n, while other parameters remain consistent with Section 4.1. Under these settings, the average absolute correlation coefficient among PCs across the four scenarios is 0.482, characterizing the dependencies between PCs. Figure 6 displays the RMSE ($\times 10^{-2}$) of correlation estimates across different sample sizes and scenarios as time progresses. The RMSE of correlation estimates is relatively small and decreases with larger sample sizes, similar to the parameter estimation results. Additionally, as the correlation coefficients stabilize over time, the RMSE values also become stable.

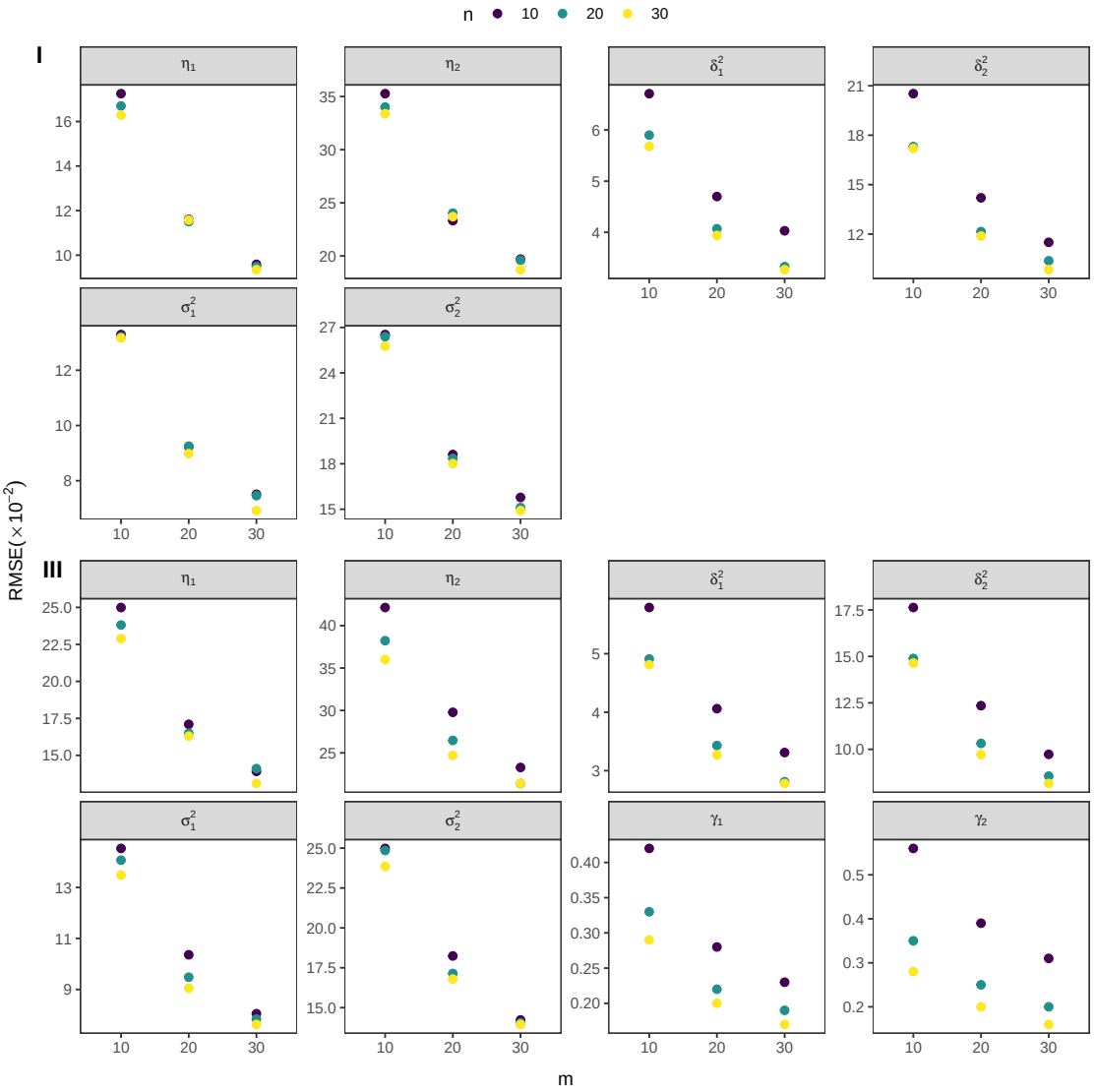


Figure 4: RMSE ($\times 10^{-2}$) for parameter estimators in scenarios I and III.

4.3. Effect of model misspecification

In this section, we perform a simulation study to evaluate the impact of model misspecification, specifically focusing on multivariate Wiener process models ($\nu \to \infty$) using EM algorithm for parameter estimation. With the simulated data in each replication, we perform model inferences and calculate the mean time to failure (MTTF) of the system (Luo et al., 2024),

$$\text{MTTF} = E(T) = \int_0^\infty R_{T_\omega}(t)\mathrm{d}t.$$

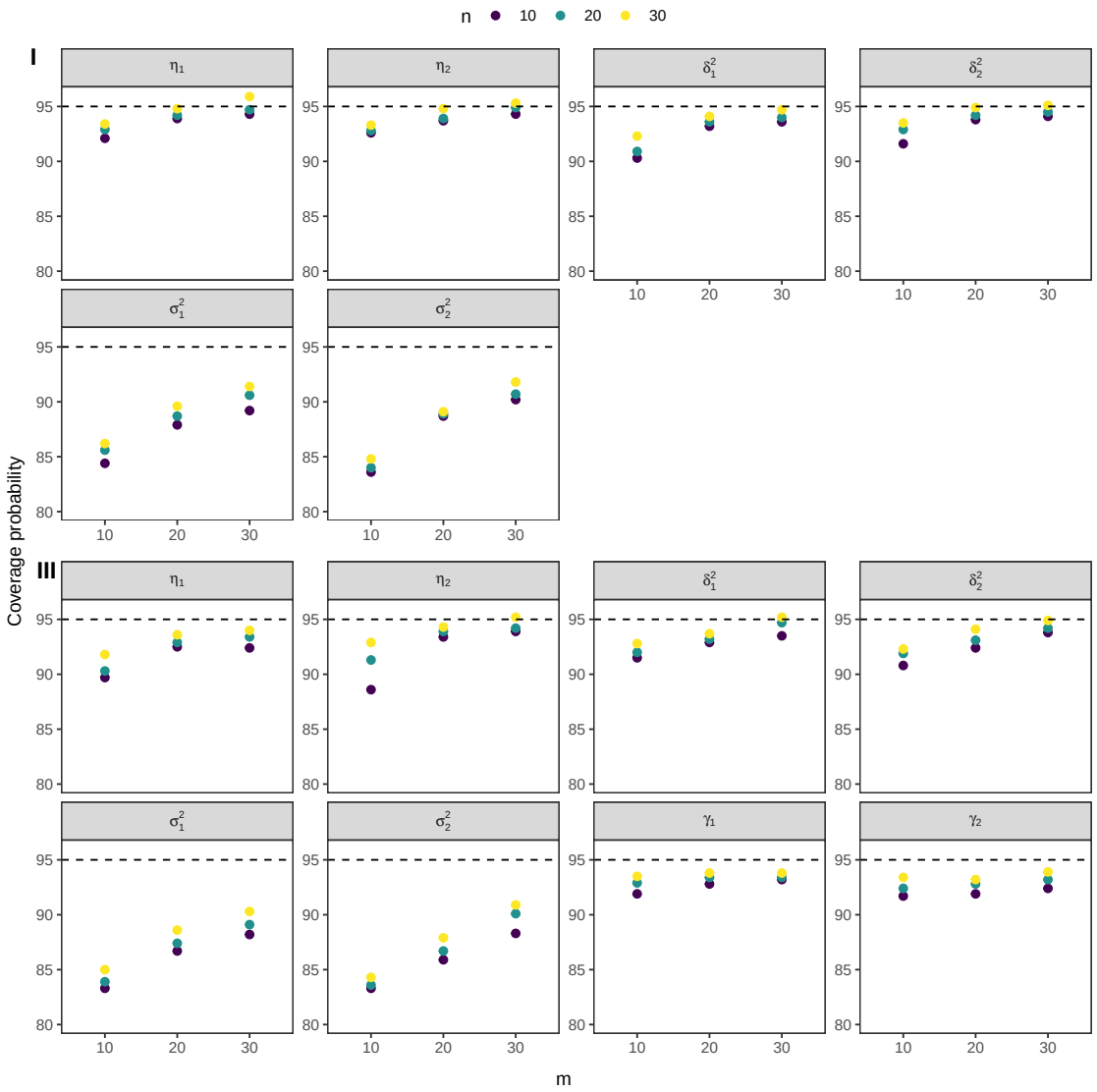


Figure 5: CP ($\times 100\%$) for parameter estimators in scenarios I and III.

The failure thresholds for three PCs are assumed to be $150, 300$, and 400, respectively. Figure 7 presents boxplots of RMSE for MTTF estimators under various sample sizes in scenario IV, with green points indicating the average RMSE. The results show that the model estimates have a consistently small RMSE, decreasing as the sample size increases. In comparison to the multivariate Wiener process, the proposed model demonstrates a smaller average RMSE. Therefore, accurately characterizing the dependent degradation data with heavy-tailed characteristics of multivariate PCs is crucial for assessing system reliability. These precise reliability estimates aid in predicting system lifespan effectively, minimizing

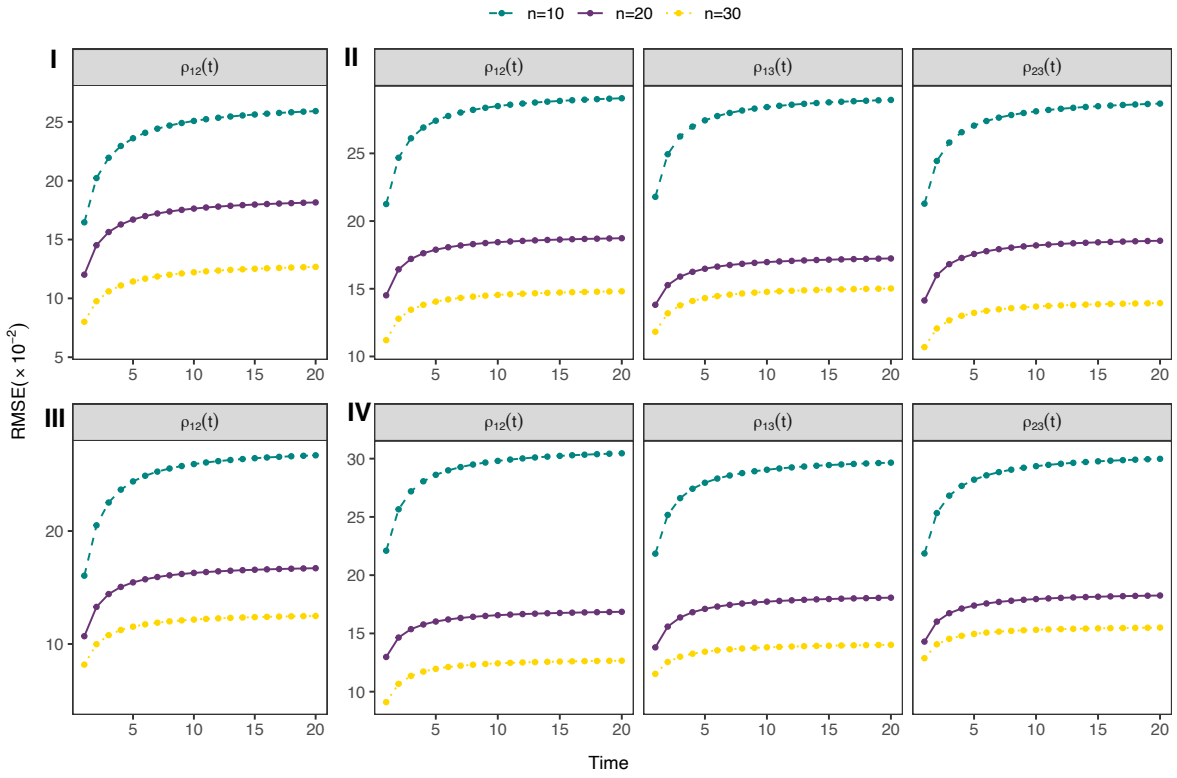


Figure 6: RMSE ($\times 10^{-2}$) of correlation estimators across different sample sizes and scenarios.

unnecessary maintenance costs and the risk of production disruptions.

5. Case studies

5.1. PMB degradation data

In this section, we illustrate the implementation of the proposed methodology by analyzing the PMB degradation data as shown in Figure 1(a). We apply the proposed model to fit the data, providing two forms of the time-scale transformation function: i) linear form $\Lambda(t) = t$; ii) power form $\Lambda(t) = t^{\gamma}$. These are referred to as models M_l and M_p, respectively. To assess the superiority of the proposed model, we consider two corresponding multivariate Wiener process models, denoted as M_l^W and M_p^W, respectively. These multivariate Wiener processes are regarded as special cases within the framework described by Fang and Pan (2023), where the impact of random initial values is not considered. The EM algorithm is employed for parameter estimation in both models, and the convergence of model inference is evaluated through trace plots of parameter estimates, as shown in the Supplementary Section S2.

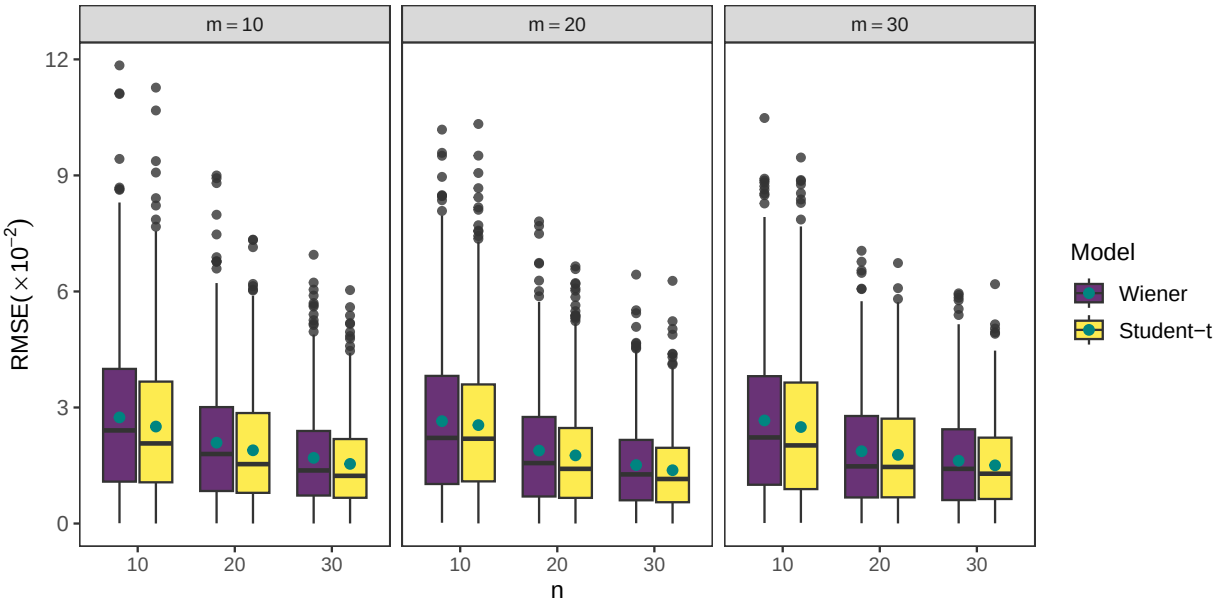


Figure 7: RMSE ($\times 10^{-2}$) for MTTF estimators under various sample sizes in scenario IV (green points mean the average RMSEs).

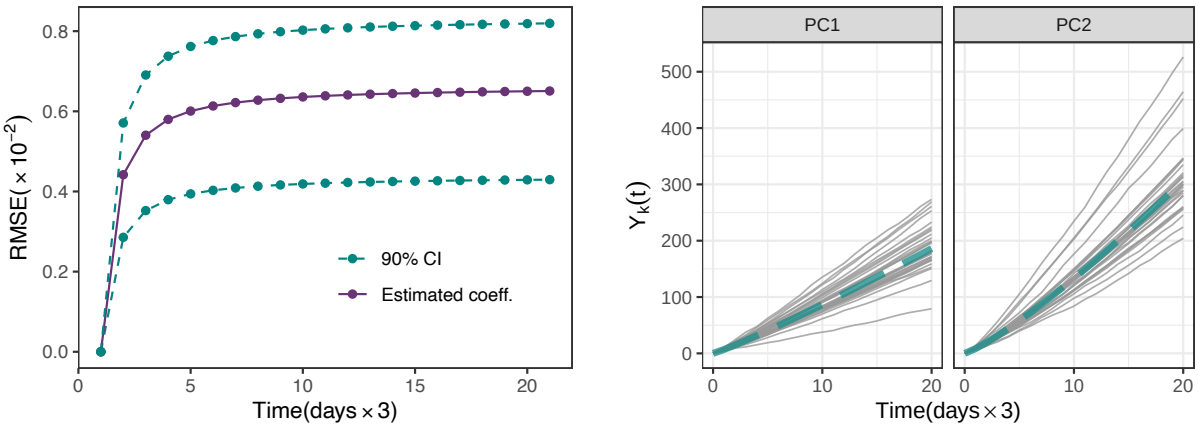
Table 2 provides parameter estimates, 90% CIs, and Akaike Information Criterion (AIC) values for different models. From these results, it is evident that the AIC values based on the tail-weighted multivariate model consistently remain lower than the corresponding AIC values for the Wiener model. Among the four models, the power form of the proposed model (i.e., M_p) yields the lowest AIC, leading us to consider it as the optimal model for the PMB data. The estimation results for γ reveal a faster degradation rate for PC2 compared to PC1, consistent with observations in Figure 1(a). Additionally, the estimation results for ν indicate that the proposed model effectively captures the heavy-tail characteristics in the data. Figure 8(a) displays the estimated correlation between $Y_1(t)$ and $Y_2(t)$ at various times, based on the proposed model, accompanied by 90% bootstrap CIs. The estimated correlation coefficients demonstrate an increase over time. Specifically, at $t = 20$, the estimated correlation coefficient is $\rho_{12}(20) = 0.642$, with a 90% CI of $(0.429, 0.811)$, indicating a moderate correlation between the two PCs. Figure 8(b) illustrates the average degradation fitting performance of model M_p on the PMB data, demonstrating a satisfactory fit to the degradation paths. This provides further validation of the effectiveness of the proposed model.

For reliability analysis, we substitute the estimated parameters into Algorithm 1 and employ MC methods with $Q = 5000$ to obtain an approximate reliability function for system

Table 2: Parameter point estimation and 90% CI for the PMB data.

Parameter	M_l	M_p	M_l^W	M_p^W
η_1	9.166	6.883	9.369	6.998
	(8.593, 9.685)	(6.465, 7.297)	(8.789, 9.916)	(6.585, 7.428)
η_2	15.095	8.311	15.412	8.432
	(14.318, 15.970)	(7.772, 8.940)	(14.554, 16.369)	(7.955, 8.934)
δ_1^2	1.077	0.409	1.133	0.572
	(0.999, 1.160)	(0.346, 0.498)	(1.074, 1.198)	(0.542, 0.603)
δ_2^2	2.874	0.613	2.951	0.837
	(2.660, 3.096)	(0.523, 0.741)	(2.772, 3.108)	(0.798, 0.886)
σ_1^2	1.759	0.914	1.915	1.285
	(1.358, 2.156)	(0.694, 1.174)	(1.511, 2.335)	(1.037, 1.581)
σ_2^2	2.654	1.163	2.967	1.535
	(2.019, 3.329)	(0.889, 1.490)	(2.377, 3.692)	(1.219, 1.871)
ρ_{12}	0.652	0.642	0.668	0.631
	(0.440, 0.793)	(0.429, 0.811)	(0.449, 0.819)	(0.412, 0.787)
ν	15.986	2.596	-	-
	(9.726, 30.000)	(1.792, 3.812)		
γ_1	-	1.098	-	1.098
		(1.085, 1.114)		(1.092, 1.105)
γ_2	-	1.202	-	1.202
		(1.185, 1.221)		(1.194, 1.210)
AIC	6302.373	**2488.158**	6386.120	3088.899

failure time. For ease of presentation, we assume that the threshold values for the two PCs are $\omega_1 = 400$ and $\omega_2 = 600$, respectively. Figure 9 illustrates the reliability curves of system failure time based on models M_p and M_p^W. Furthermore, we compute the pseudo failure time (PFT) of the system by determining the minimum time at which the fitted curve for each PC reaches the threshold value. This can be achieved by simply fitting the data for each degradation path using ordinary least squares (Peng and Cheng, 2020). Utilizing the Anderson-Darling (AD) test for model validation, we assess the goodness of fit for models M_p and M_p^W. The p-values for the two models are 0.703 and 0.289, respectively, indicating strong performance at a significance level of 0.05. However, the proposed model M_p exhibits a higher p-value, suggesting its superior performance in fitting the PMB data. This commendable performance can be attributed to the robustness of the proposed model,



(a) Correlation coefficient estimation (b) Average degradation path estimation

Figure 8: Summary of the PMB data analysis results: Correlation coefficient estimation ($\hat{\rho}_{12}$) and average degradation path estimation under M_p model.

enabling it to handle potential outliers or noise more effectively.

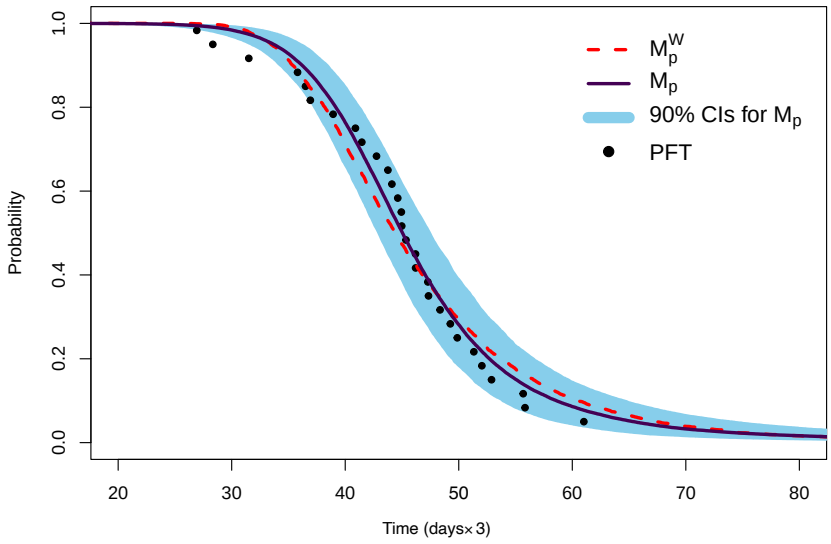


Figure 9: The estimated reliability of the PMB data.

While the tail-weighted multivariate degradation process excels in handling data with heavy-tail characteristics, its performance becomes uncertain when modeling data without heavy-tail features. Therefore, we use the following example to demonstrate the robustness and flexibility of the proposed model.

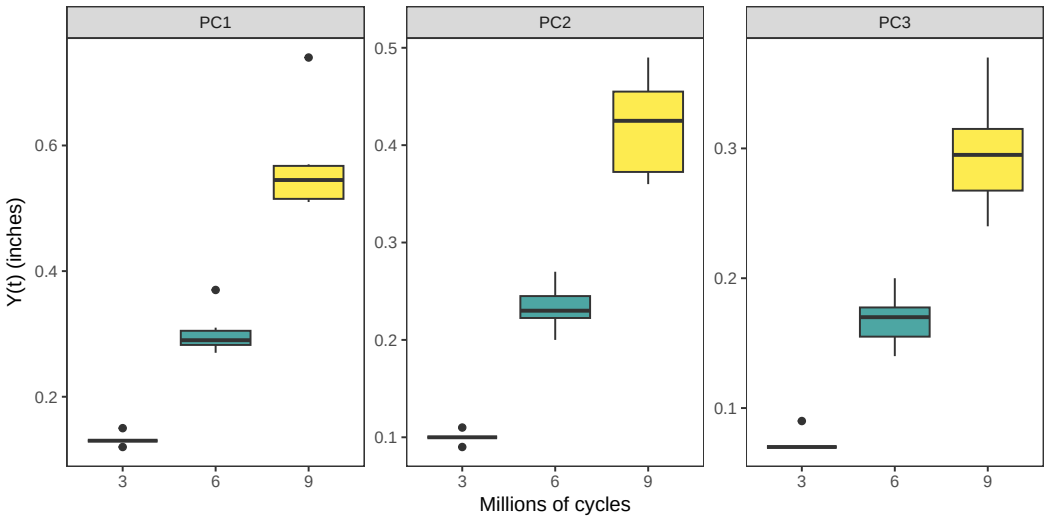


Figure 10: Degradation paths of the FCS data.

5.2. Fatigue crack-size data

The alloy FCS data is obtained from Meeker et al. (2022), and we divide the dataset into three segments, creating a three-dimensional degradation process (Fang et al., 2022). For simplicity, we subtract the initial crack size of 0.9, and the resulting degradation paths over some observation times are depicted in Figure 10. The boxplot indicates the absence of outliers in the degradation values for the majority of PCs. For such data without strong heavy-tail features, we utilize the proposed model, fitting it with the exponential time scale transformation function $\lambda(t) = \exp(\gamma t) - 1$, denoted as M_e. Additionally, the corresponding multivariate Wiener process model, denoted as M_e^W, is considered. Parameters for each model are estimated using the EM algorithm, and their convergence trace plots can be found in Supplementary Section S2. Table 3 lists the estimated parameters along with their respective AIC values for each model. Due to suboptimal fitting results and page constraints, the outcomes for the M_l and M_l^W models are omitted. The estimated parameter ν for model M_e is 27.827, suggesting that the data doesn't show significant heavy-tail behavior, consistent with the results in Figure 10. Furthermore, we compare the multivariate IG process (Fang et al., 2022), where the model's AIC value of -1074.186 is much higher than our proposed model. In the AIC model selection, M_e performs comparably to M_e^W.

Figure 11 displays the estimated correlation coefficient under M_e model, accompanied by 90% bootstrap CIs. At $t = 10$, $\rho_{12}(10) = 0.601$, $\rho_{13}(10) = 0.608$, and $\rho_{23}(10) = 0.498$, with 90% CIs of $(0.484, 0.768)$, $(0.522, 0.768)$, and $(0.384, 0.706)$, respectively. These values indicate a moderate correlation among the three PCs, increasing over time. Assuming

Table 3: Parameter point estimation and 90% CI for the FCS data.

Model	Parameters					
M_e	η_1	η_2	η_3	δ_1^2	δ_2^2	δ_3^2
	0.302	0.510	0.379	0.019	0.030	0.026
	(0.215, 0.482)	(0.295, 2.003)	(0.185, 0.806)	(0.016, 0.030)	(0.022, 0.061)	(0.017, 0.039)
	σ_1^2	σ_2^2	σ_3^2	σ_{12}	σ_{13}	σ_{23}
	0.035	0.051	0.048	0.042	0.041	0.050
	(0.011, 0.076)	(0.017, 0.166)	(0.011, 0.106)	(0.014, 0.060)	(0.020, 0.067)	(0.017, 0.074)
	γ_1	γ_2	γ_3	ν	AIC	
	0.116	0.066	0.063	27.827	-1336.660	
	(0.089, 0.139)	(0.021, 0.097)	(0.037, 0.101)	(8.614, 59.871)		
M_e^W	η_1	η_2	η_3	δ_1^2	δ_2^2	δ_3^2
	0.305	0.515	0.384	0.020	0.031	0.027
	(0.205, 0.494)	(0.204, 0.995)	(0.165, 0.757)	(0.016, 0.028)	(0.023, 0.220)	(0.018, 0.673)
	σ_1^2	σ_2^2	σ_3^2	σ_{12}	σ_{13}	σ_{23}
	0.038	0.053	0.050	0.045	0.044	0.051
	(0.014, 0.069)	(0.011, 0.098)	(0.012, 0.108)	(0.004,0.050)	(0.024, 0.062)	(0.013, 0.073)
	γ_1	γ_2	γ_3	ν	AIC	
	0.116	0.066	0.063	-	-1339.663	
	(0.081, 0.145)	(0.017, 0.122)	(0.002, 0.116)			

threshold values for three PCs are $\omega_1 = 0.9$, $\omega_2 = 0.5$, and $\omega_3 = 0.4$. Figure 12 illustrates reliability estimates of system lifetime distribution under models M_e and M_e^W, along with the 90% CI for the reliability of the M_e model. The fitting performance of the proposed model closely aligns with the results of the multivariate Wiener process, highlighting the model's adaptability and flexibility. It adjusts tail heaviness through the degree of freedom parameter, enabling its suitability for data without prominent heavy-tail features.

6. Conclusion

In this study, we introduce a novel class of tail-weighted multivariate degradation models. Our model proficiently accounts for both within-unit variability and dependencies among PCs while allowing flexible tuning of the tail heaviness through the parameter of the degree of freedom. Based on the model, we derive the system reliability and provide an efficient MC

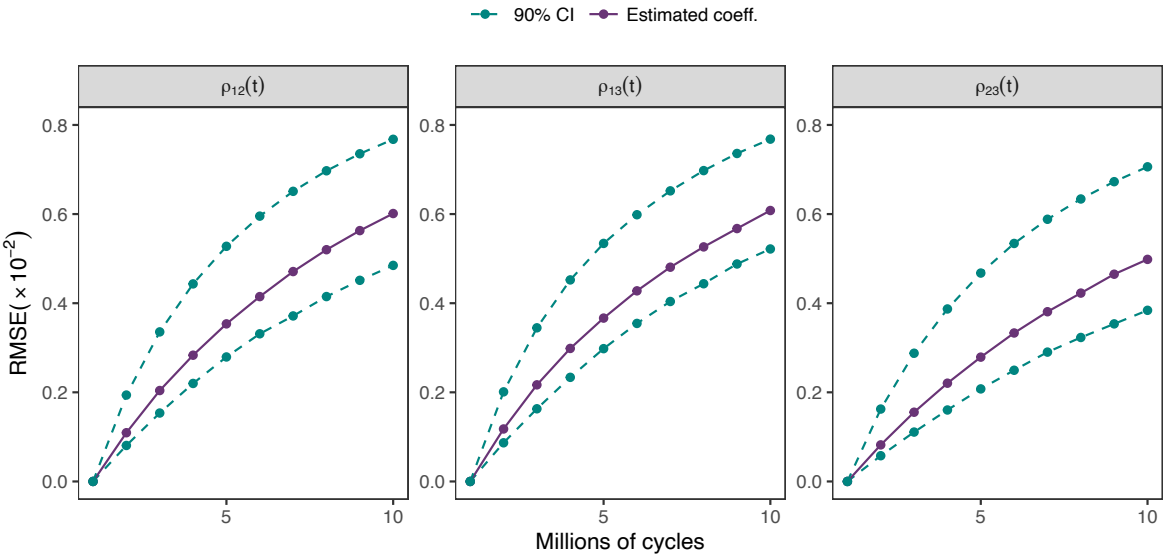


Figure 11: Correlation coefficient estimation for the FCS data under M_e model.

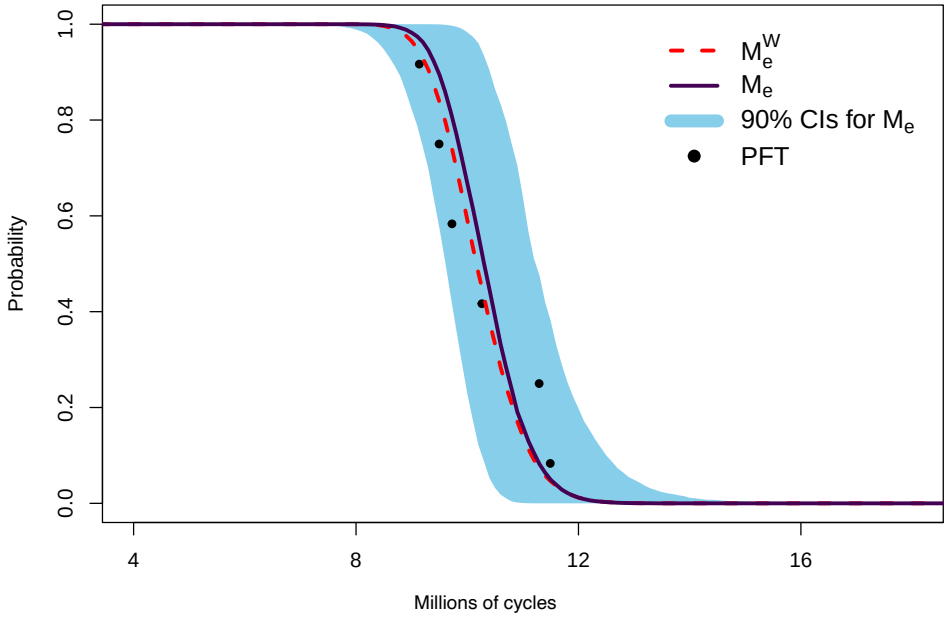


Figure 12: The estimated reliability of the FCS data.

method for reliability assessment. Considering the complexities of the likelihood function and the non-linear time scale transformation, we design an innovative two-stage parameter estimation method, which seamlessly combines NLS and EM methods. Furthermore, a bootstrap method is utilized to construct interval estimates for these parameters. Comprehensive simulation studies are conducted to validate the effectiveness of our inference method. The findings indicate that: i) Larger sample sizes improve the accuracy of point estimates; ii) The bootstrap method for interval estimation offers reasonable CPs; iii) Ignoring heavy-tail characteristics in the model results in significant biases when estimating the MTTF. Finally, we demonstrate the applicability of our proposed methodology through two case studies. Compared to other existing models, the proposed model is versatile to accommodate multi-dimensional degradation data, both with and without heavy-tail characteristics, by adjusting the degree of freedom parameters. This not only highlights the model's adaptability but also indicates its potential for diverse applications.

In the future, we can further enhance the model's predictive accuracy and generalizability by incorporating more complex dependencies and external dynamic environmental factors (Hajiha et al., 2021). For instance, integrating variables such as operational conditions, environmental stressors, and usage patterns can provide a more comprehensive understanding of degradation processes. Moreover, applying our model to maintenance scheduling and reliability management is a promising direction (Zhao et al., 2022). By accurately predicting degradation trajectories, our model can help optimize maintenance schedules, thereby reducing downtime and maintenance costs. This may involve developing decision-support systems that leverage our model's predictions to recommend proactive maintenance actions.

Acknowledgments

The authors sincerely appreciate the editor and anonymous reviewers for their detailed comments and valuable suggestions.

Funding

The research is supported by Zhejiang Provincial Natural Science Foundation of China (LZ24A010002, LQ22G010003), Natural Science Foundation of China (12171432, 72201242), the China Scholarship Council program (202308330508), the Fundamental Research Funds for the Provincial Universities of Zhejiang, the Summit Advancement Disciplines of Zhejiang

Province (Zhejiang Gongshang University—Statistics), and Collaborative Innovation Center of Statistical Data Engineering Technology & Application.

Notes on contributors

Ancha Xu received the B.S. degree in Mathematics from the Central China Normal University, Wuhan, China, in 2006, and the Ph.D. degree in Probability Theory and Mathematical Statistics from East China Normal University, Shanghai, China, in 2011. From 2011 to 2020, he was an Associate Professor of Statistics at Wenzhou University, Wenzhou, China. Since 2021, he has been a Professor at the Department of Statistics, Zhejiang Gongshang University, Hangzhou, China. He is the author of more than 50 articles. His research interests include Bayesian computation, data-driven degradation modeling, and industrial statistics.

Guanqi (Kevin) Fang is an Associate Professor in School of Statistics and Mathematics at Zhejiang Gongshang University, Hangzhou, China. He obtained his Ph.D. degree in Industrial Engineering and a concurrent M.S. degree in Statistics from Arizona State University. His research interests include reliability modeling, statistical inference, and multivariate data analysis, among others.

Liangliang Zhuang received his B.S. and M.S. degrees in Applied Statistics from Wenzhou University, China, in 2019 and 2022, respectively. He is currently pursuing a Ph.D. in Statistics at Zhejiang Gongshang University, China, and is also a visiting Ph.D. student at the National University of Singapore, Singapore. His research interests include reliability analysis and optimization, prognostics, condition monitoring, and maintenance optimization.

Cheng Gu received the B.S. degree in Information and Computing Science from the Zhejiang University of Technology, Hangzhou, China, in 2021. He is currently working toward the M.S. degree in statistics with Zhejiang Gongshang University, Hangzhou, China. His research interests mainly focus on degradation modeling.

Appendix

Appendix A. Proof of Theorem 1

According to (8) and Bayes' theorem, the joint conditional distribution of $\boldsymbol{\Theta}_i$ and τ_i

$$
\begin{aligned}
f(\boldsymbol{\Theta}_i, \tau_i|\Delta\boldsymbol{Y}_i) &\propto f(\Delta\boldsymbol{Y}_i|\boldsymbol{\Theta}_i, \tau_i) \times f(\boldsymbol{\Theta}_i|\tau_i) \times f_{\mathcal{G}}(\tau_i) \\
&\propto \prod_{k=1}^{m_i} \tau_i^{p/2} \exp\left\{-\frac{(\Delta\boldsymbol{Y}_{i,k} - \Delta\boldsymbol{\Sigma}(t_{i,k})\boldsymbol{\Theta}_i)' \left(\boldsymbol{\Omega}_\delta \Delta\boldsymbol{\Sigma}(t_{i,k})/\tau_i\right)^{-1} (\Delta\boldsymbol{Y}_{i,k} - \Delta\boldsymbol{\Sigma}(t_{i,k})\,\boldsymbol{\Theta}_i)}{2}\right\} \\
&\quad \times \tau_i^{p/2} \exp\left\{-\frac{(\boldsymbol{\Theta}_i - \boldsymbol{\eta})^\top (\boldsymbol{\Sigma}_0/\tau_i)^{-1} (\boldsymbol{\Theta}_i - \boldsymbol{\eta})}{2}\right\} \times \tau_i^{\mu/2-1} \exp(-\tau_i\nu/2) \\
&\propto \prod_{k=1}^{m_i} \exp\left\{-\frac{\left(\boldsymbol{\Theta}_i - \Delta\boldsymbol{\Sigma}^{-1}(t_{i,k})\,\Delta\boldsymbol{Y}_{i,k}\right)'(\boldsymbol{\Omega}_\delta \Delta\boldsymbol{\Sigma}^{-1}(t_{i,k})/\tau_i)^{-1}(\boldsymbol{\Theta}_i - \Delta\boldsymbol{\Sigma}^{-1}(t_{i,k})\Delta\boldsymbol{Y}_{i,k})}{2}\right\} \\
&\quad \times \tau_i^{(m_i+1)p/2+\nu/2-1} \exp\left\{-\frac{(\boldsymbol{\Theta}_i - \boldsymbol{\eta})^\top (\boldsymbol{\Sigma}_0/\tau_i)^{-1} (\boldsymbol{\Theta}_i - \boldsymbol{\eta})}{2}\right\} \\
&\propto \tau_i^{p/2} \exp\left\{-\frac{(\boldsymbol{\Theta}_i - \boldsymbol{\mu}_i)' (\boldsymbol{\Sigma}_{\boldsymbol{\Theta}_i}/\tau_i)^{-1}(\boldsymbol{\Theta}_i - \boldsymbol{\mu}_i)}{2}\right\} \tau_i^{m_i p/2+\nu/2-1} \exp\left\{-\frac{\tau_i\,(K_{i,1} - K_{i,2} + \nu)}{2}\right\},
\end{aligned}
$$

where

$$
\begin{aligned}
&\boldsymbol{\Sigma}_{\boldsymbol{\Theta}_i} = \left[\boldsymbol{\Sigma}_0^{-1} + \boldsymbol{\Omega}_\delta^{-1}\boldsymbol{\Sigma}(t_{im_i})\right]^{-1}, \quad \boldsymbol{\mu}_i = \boldsymbol{\Sigma}_{\boldsymbol{\Theta}_i}(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\eta} + \boldsymbol{\Omega}_\delta^{-1}\boldsymbol{Y}_{i,m_i}), \\
&K_{i,1} = \sum_{k=1}^{m_i} \Delta\boldsymbol{Y}_{i,k}'\boldsymbol{\Omega}_\delta^{-1}\Delta\boldsymbol{\Sigma}^{-1}(t_{i,k})\Delta\boldsymbol{Y}_{i,k} + \boldsymbol{\eta}^\top\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\eta}, \quad K_{i,2} = \boldsymbol{\mu}_i'\boldsymbol{\Sigma}_{\boldsymbol{\Theta}_i}^{-1}\boldsymbol{\mu}_i.
\end{aligned}
$$

Then we know that given $\Delta\boldsymbol{Y}_i$ and τ_i, the conditional distribution of $\boldsymbol{\Theta}_i$ is proportional to

$$
\tau_i^{p/2} \exp\left\{-\frac{(\boldsymbol{\Theta}_i - \boldsymbol{\mu}_i)' (\boldsymbol{\Sigma}_{\boldsymbol{\Theta}_i}/\tau_i)^{-1}(\boldsymbol{\Theta}_i - \boldsymbol{\mu}_i)}{2}\right\}.
$$

Thus, we have that $\boldsymbol{\Theta}_i|\Delta\boldsymbol{Y}_i, \tau_i \sim \mathcal{N}_p\left(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_{\boldsymbol{\Theta}_i}/\tau_i\right)$. Given $\Delta\boldsymbol{Y}_i$, the distribution of τ_i

$$
f(\tau_i|\Delta\boldsymbol{Y}_i) = \int f(\boldsymbol{\Theta}_i, \tau_i|\Delta\boldsymbol{Y}_i)\mathrm{d}\boldsymbol{\Theta}_i \propto \tau_i^{m_i p/2+\nu/2-1} \exp\left\{-\frac{\tau_i\,(K_{i,1} - K_{i,2} + \nu)}{2}\right\},
$$

which is the gamma distribution $\mathcal{G}\left(\frac{m_i p+\nu}{2}, \frac{2}{K_{i,1}-K_{i,2}+\nu}\right)$.

Appendix B. Proof of Theorem 2

Firstly, we list the following lemma that will be used in the proof. The details of the lemma can be found on page 107 of Rencher and Schaalje (2008).

Lemma 1. *If $\boldsymbol{x}$ is a random vector with a mean vector of $\boldsymbol{\beta}$ and a covariance matrix of $\boldsymbol{B}$, and if $\boldsymbol{A}$ is a symmetric matrix composed of constants, then* $\mathrm{E}_x(\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}) = \mathrm{tr}(\boldsymbol{A}\boldsymbol{B}) + \boldsymbol{\beta}'\boldsymbol{A}\boldsymbol{\beta}$.

From Lemma 1, we can obtain the following result.

Lemma 2. *Under the conditions of Lemma 1, for any constant vector $\boldsymbol{\beta}_0$,*

$$\mathrm{E}_x\left[(\boldsymbol{x}-\boldsymbol{\beta}_0)'\boldsymbol{A}(\boldsymbol{x}-\boldsymbol{\beta}_0)\right] = \mathrm{tr}(\boldsymbol{A}\boldsymbol{B}) + (\boldsymbol{\beta}-\boldsymbol{\beta}_0)'\boldsymbol{A}(\boldsymbol{\beta}-\boldsymbol{\beta}_0).$$

Proof: Firstly, we can expand $(\boldsymbol{x}-\boldsymbol{\beta}_0)'\boldsymbol{A}(\boldsymbol{x}-\boldsymbol{\beta}_0)$ as follows:

$$(\boldsymbol{x}-\boldsymbol{\beta}_0)'\boldsymbol{A}(\boldsymbol{x}-\boldsymbol{\beta}_0) = (\boldsymbol{x}-\boldsymbol{\beta})'\boldsymbol{A}(\boldsymbol{x}-\boldsymbol{\beta}) - 2(\boldsymbol{x}-\boldsymbol{\beta})'\boldsymbol{A}(\boldsymbol{\beta}-\boldsymbol{\beta}_0) + (\boldsymbol{\beta}-\boldsymbol{\beta}_0)'\boldsymbol{A}(\boldsymbol{\beta}-\boldsymbol{\beta}_0).$$

According to Lemma 1, we have

$$\mathrm{E}_x\left[(\boldsymbol{x}-\boldsymbol{\beta})'\boldsymbol{A}(\boldsymbol{x}-\boldsymbol{\beta})\right] = \mathrm{tr}(\boldsymbol{A}\boldsymbol{B}),$$

because $\mathrm{E}_x[\boldsymbol{x}-\boldsymbol{\beta}] = 0$ and $\mathrm{Cov}_x[\boldsymbol{x}-\boldsymbol{\beta}] = \boldsymbol{B}$. Thus, the result holds.

Now we proceed to prove Theorem 2. Recalling that $\ell_{i,0} = (\boldsymbol{\Theta}_i-\boldsymbol{\eta})'\boldsymbol{\Sigma}_0^{-1}(\boldsymbol{\Theta}_i-\boldsymbol{\eta})$ and $\boldsymbol{\Theta}_i|\Delta\boldsymbol{Y}_i,\tau_i,\boldsymbol{\Psi}^{(s)} \sim \mathcal{N}_p\left(\boldsymbol{\mu}_i^{(s)}, \boldsymbol{\Sigma}_{\boldsymbol{\Theta}_i}^{(s)}/\tau_i\right)$. According to Lemma 2, it is straightforward to obtain

$$\mathrm{E}\left(\ell_{i,0}|\Delta\boldsymbol{Y}_i,\tau_i,\boldsymbol{\Psi}^{(s)}\right) = \mathrm{tr}\left(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_{\boldsymbol{\Theta}_i}^{(s)}/\tau_i\right) + \left(\boldsymbol{\mu}_i^{(s)}-\boldsymbol{\eta}\right)'\boldsymbol{\Sigma}_0^{-1}\left(\boldsymbol{\mu}_i^{(s)}-\boldsymbol{\eta}\right).$$

Then, we can obtain

$$\begin{aligned}\mathrm{E}\left(\tau_i\ell_{i,0}|\Delta\boldsymbol{Y}_i,\boldsymbol{\Psi}^{(s)}\right) &= \mathrm{E}_{\tau_i}\left[\mathrm{E}(\tau_i\ell_{i,0}|\Delta\boldsymbol{Y}_i,\tau_i,\boldsymbol{\Psi}^{(s)})\right] = \mathrm{E}_{\tau_i}\left[\tau_i\mathrm{E}\left(\ell_{i,0}|\Delta\boldsymbol{Y}_i,\tau_i,\boldsymbol{\Psi}^{(s)}\right)\right]\\ &= \mathrm{tr}\left(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_{\boldsymbol{\Theta}_i}^{(s)}\right) + \mathrm{E}\left(\tau_i|\Delta\boldsymbol{Y}_i,\boldsymbol{\Psi}^{(s)}\right)\left(\boldsymbol{\mu}_i^{(s)}-\boldsymbol{\eta}\right)'\boldsymbol{\Sigma}_0^{-1}\left(\boldsymbol{\mu}_i^{(s)}-\boldsymbol{\eta}\right).\end{aligned}$$

Notice that

$$\begin{aligned}\ell_{i,k} &= (\Delta\boldsymbol{Y}_{i,k}-\Delta\boldsymbol{\Sigma}(t_{i,k})\boldsymbol{\Theta}_i)'\left(\boldsymbol{\Omega}_\delta\Delta\boldsymbol{\Sigma}(t_{i,k})\right)^{-1}(\Delta\boldsymbol{Y}_{i,k}-\Delta\boldsymbol{\Sigma}(t_{i,k})\boldsymbol{\Theta}_i)\\ &= \left(\boldsymbol{\Theta}_i-\Delta\boldsymbol{\Sigma}^{-1}(t_{i,k})\,\Delta\boldsymbol{Y}_{i,k}\right)'(\boldsymbol{\Omega}_\delta\Delta\boldsymbol{\Sigma}^{-1}(t_{i,k}))^{-1}(\boldsymbol{\Theta}_i-\Delta\boldsymbol{\Sigma}^{-1}(t_{i,k})\Delta\boldsymbol{Y}_{i,k}).\end{aligned}$$

Similarly, we can obtain that

$$\begin{aligned}\mathrm{E}\left(\tau_i\ell_{i,k}|\Delta\boldsymbol{Y}_i,\boldsymbol{\Psi}^{(s)}\right) &= \mathrm{tr}\left(\Delta\boldsymbol{\Sigma}(t_{i,k})\boldsymbol{\Omega}_\delta^{-1}\boldsymbol{\Sigma}_{\boldsymbol{\Theta}_i}^{(s)}\right) + \mathrm{E}\left(\tau_i|\Delta\boldsymbol{Y}_i,\boldsymbol{\Psi}^{(s)}\right)\\ &\quad\times\left(\boldsymbol{\mu}_i^{(s)}-\Delta\boldsymbol{\Sigma}^{-1}(t_{i,k})\Delta\boldsymbol{Y}_{i,k}\right)'(\boldsymbol{\Omega}_\delta\Delta\boldsymbol{\Sigma}^{-1}(t_{i,k}))^{-1}\left(\boldsymbol{\mu}_i^{(s)}-\Delta\boldsymbol{\Sigma}^{-1}(t_{i,k})\Delta\boldsymbol{Y}_{i,k}\right).\end{aligned}$$

Appendix C. Proof of Theorem 3

Taking the partial derivative of $Q\left(\boldsymbol{\Psi} \mid \boldsymbol{\Psi}^{(s)}\right)$ with respect to $\boldsymbol{\eta}$ and setting it to zero

$$\frac{\partial Q\left(\boldsymbol{\Psi} \mid \boldsymbol{\Psi}^{(s)}\right)}{\partial \boldsymbol{\eta}} = \sum_{i=1}^{n} \mathrm{E}\left(\tau_i | \Delta \boldsymbol{Y}_i, \boldsymbol{\Psi}^{(s)}\right) \boldsymbol{\Sigma}_0^{-1}\left(\boldsymbol{\mu}_i^{(s)} - \boldsymbol{\eta}\right) = 0.$$

Solving the above equation can get the solution $\boldsymbol{\eta}^{(s+1)}$:

$$\boldsymbol{\eta}^{(s+1)} = \frac{\sum_{i=1}^{n} \boldsymbol{\mu}_i^{(s)} \mathrm{E}\left(\tau_i | \boldsymbol{\Delta Y}_i, \boldsymbol{\Psi}^{(s)}\right)}{\sum_{i=1}^{n} \mathrm{E}\left(\tau_i | \boldsymbol{\Delta Y}_i, \boldsymbol{\Psi}^{(s)}\right)}.$$

Taking the partial derivative of $Q\left(\boldsymbol{\Psi} \mid \boldsymbol{\Psi}^{(s)}\right)$ with respect to $\boldsymbol{\eta}$ and setting it to zero

$$\frac{\partial Q\left(\boldsymbol{\Psi} \mid \boldsymbol{\Psi}^{(s)}\right)}{\partial \boldsymbol{\Sigma}_0} = -\frac{n}{2}\boldsymbol{\Sigma}_0^{-1} + \frac{1}{2}\sum_{i=1}^{n}\left[\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_{\boldsymbol{\Theta}_i}^{(s)}\boldsymbol{\Sigma}_0^{-1} + \boldsymbol{\Sigma}_0^{-1}\left(\boldsymbol{\mu}_i^{(s)} - \boldsymbol{\eta}\right)\left(\boldsymbol{\mu}_i^{(s)} - \boldsymbol{\eta}\right)'\boldsymbol{\Sigma}_0^{-1}\right] = 0.$$

Solving the above equation and substituting the optimal solution $\boldsymbol{\eta}^{(s+1)}$ into the result, we have

$$\boldsymbol{\Sigma}_0^{(s+1)} = \frac{\sum_{i=1}^{n}\left[\boldsymbol{\Sigma}_{\boldsymbol{\Theta}_i}^{(s)} + \mathrm{E}(\tau_i | \Delta \boldsymbol{Y}_i, \boldsymbol{\Psi}^{(s)})(\boldsymbol{\mu}_i^{(s)} - \boldsymbol{\eta}^{(s+1)})(\boldsymbol{\mu}_i^{(s)} - \boldsymbol{\eta}^{(s+1)})'\right]}{n}.$$

Similarly, we take the partial derivative of $Q\left(\boldsymbol{\Psi} \mid \boldsymbol{\Psi}^{(s)}\right)$ with respect to $\boldsymbol{\Omega}_\delta$ and set it to zero.

$$\begin{aligned}\frac{\partial Q\left(\boldsymbol{\Psi} \mid \boldsymbol{\Psi}^{(s)}\right)}{\partial \boldsymbol{\Omega}_\delta} = -\frac{1}{2}\sum_{i=1}^{n}\boldsymbol{\Omega}_\delta^{-1} + \frac{1}{2}\sum_{i=1}^{n}\sum_{k=1}^{m_i}\Big[&\boldsymbol{\Omega}_\delta^{-1}\Delta\boldsymbol{\Sigma}(t_{i,k})\boldsymbol{\Sigma}_{\boldsymbol{\Theta}_i}^{(s)}\boldsymbol{\Omega}_\delta^{-1} + \mathrm{E}\left(\tau_i | \Delta \boldsymbol{Y}_i, \boldsymbol{\Psi}^{(s)}\right) \\ &\times \boldsymbol{\Omega}_\delta^{-1}\Delta\boldsymbol{\Sigma}(t_{i,k})\left(\boldsymbol{\mu}_i^{(s)} - \Delta\boldsymbol{\Sigma}^{-1}(t_{i,k})\Delta\boldsymbol{Y}_{i,k}\right)\left(\boldsymbol{\mu}_i^{(s)} - \Delta\boldsymbol{\Sigma}^{-1}(t_{i,k})\Delta\boldsymbol{Y}_{i,k}\right)'\boldsymbol{\Omega}_\delta^{-1}\Big] = 0.\end{aligned}$$

Then we can get

$$\begin{aligned}\boldsymbol{\Omega}_\delta^{(s+1)} = \frac{1}{\sum_{i=1}^{n} m_i}\sum_{i=1}^{n}\sum_{k=1}^{m_i}\Big[&\Delta\boldsymbol{\Sigma}(t_{i,k})\boldsymbol{\Sigma}_{\boldsymbol{\Theta}_i}^{(s)} + \mathrm{E}(\tau_i | \Delta \boldsymbol{Y}_i, \boldsymbol{\Psi}^{(s)}) \\ &\times (\boldsymbol{\mu}_i^{(s)} - \Delta\boldsymbol{\Sigma}^{-1}(t_{i,k})\Delta\boldsymbol{Y}_{i,k})(\boldsymbol{\mu}_i^{(s)} - \Delta\boldsymbol{\Sigma}^{-1}(t_{i,k})\Delta\boldsymbol{Y}_{i,k})'\Big].\end{aligned}$$

References

Bae, S.J., Kvam, P.H., 2004. A nonlinear random-coefficients model for degradation testing. Technometrics 46, 460–469. doi:10.1198/004017004000000464.

Chen, P., Ye, Z.S., 2018. Uncertainty quantification for monotone stochastic degradation models. Journal of Quality Technology 50, 207–219. doi:10.1080/00224065.2018.1436839.

Chen, P., Ye, Z.S., Xiao, X., 2019. Pairwise model discrimination with applications in lifetime distributions and degradation processes. Naval Research Logistics 66, 675–686. doi:10.1002/nav.21875.

Dempster, A.P., Laird, N.M., Rubin, D.B., 1977. Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society: Series B 39, 1–22. doi:10.1111/j.2517-6161.1977.tb01600.x.

Di Nardo, E., Nobile, A.G., Pirozzi, E., Ricciardi, L., 2001. A computational approach to first-passage-time problems for Gauss–Markov processes. Advances in Applied Probability 33, 453–482. doi:10.1239/aap/999188324.

Fan, J., Yung, K.C., Pecht, M., 2012. Lifetime estimation of high-power white LED using degradation-data-driven method. IEEE Transactions on Device and Materials Reliability 12, 470–477. doi:10.1109/TDMR.2012.2190415.

Fang, G., Pan, R., 2023. A class of hierarchical multivariate Wiener processes for modeling dependent degradation data. Technometrics 66, 141–156. doi:10.1080/00401706.2023.2242413.

Fang, G., Pan, R., Wang, Y., 2022. Inverse Gaussian processes with correlated random effects for multivariate degradation modeling. European Journal of Operational Research 300, 1177–1193. doi:10.1016/j.ejor.2021.10.049.

Hajiha, M., Liu, X., Hong, Y., 2021. Degradation under dynamic operating conditions: Modeling, competing processes and applications. Journal of Quality Technology 53, 347–368. doi:10.1080/00224065.2020.175739.

Hong, Y., Zhang, M., Meeker, W.Q., 2018. Big data and reliability applications: The complexity dimension. Journal of Quality Technology 50, 135–149. doi:10.1080/00224065.2018.1438007.

Jiang, D., Chen, T., Xie, J., Cui, W., Song, B., 2023. A mechanical system reliability degradation analysis and remaining life estimation method——With the example of an aircraft hatch lock mechanism. Reliability Engineering & System Safety 230, 108922. doi:10.1016/j.ress.2022.108922.

Kang, R., Gong, W., Chen, Y., 2020. Model-driven degradation modeling approaches: Investigation and review. Chinese Journal of Aeronautics 33, 1137–1153. doi:10.1016/j.cja.2019.12.006.

Kou, B., Chen, W., Jin, Y., 2021. A novel cage-secondary permanent magnet linear eddy current brake with wide speed range and its analytical model. IEEE Transactions on Industrial Electronics 69, 7130–7139. doi:10.1109/TIE.2021.3097603.

Liu, B., Pandey, M.D., Wang, X., Zhao, X., 2021. A finite-horizon condition-based maintenance policy for a two-unit system with dependent degradation processes. European Journal of Operational Research 295, 705–717. doi:10.1016/j.ejor.2021.03.010.

Lu, C.J., Meeker, W.O., 1993. Using degradation measures to estimate a time-to-failure distribution. Technometrics 35, 161–174. doi:10.1080/00401706.1993.10485038.

Lu, L., Wang, B., Hong, Y., Ye, Z., 2020. General path models for degradation data with multiple characteristics and covariates. Technometrics 63, 354–369. doi:10.1080/00401706.2020.1796814.

Luo, C., Tan, C.H., Liu, X., 2020. Maximum excess dominance: Identifying impractical solutions in linear problems with interval coefficients. European Journal of Operational Research 282, 660–676. doi:10.1016/j.ejor.2019.09.030.

Luo, F., Hu, L., Wang, Y., Yu, X., 2024. Statistical inference of reliability for a K-out-of-N: G system with switching failure under Poisson shocks. Statistical Theory and Related Fields doi:10.1080/24754269.2024.2341982.

Meeker, W.Q., Escobar, L.A., Pascual, F.G., 2022. Statistical Methods for Reliability Data. John Wiley & Sons.

Peng, C.Y., Cheng, Y.S., 2020. Student-t processes for degradation analysis. Technometrics 62, 223–235. doi:10.1080/00401706.2019.1630008.

Peng, W., Li, Y.F., Yang, Y.J., Huang, H.Z., Zuo, M.J., 2014. Inverse Gaussian process models for degradation analysis: A Bayesian perspective. Reliability Engineering & System Safety 130, 175–189. doi:10.1016/j.ress.2014.06.005.

Peng, W., Ye, Z.S., Chen, N., 2019. Joint online RUL prediction for multivariate deteriorating systems. IEEE Transactions on Industrial Informatics 15, 2870–2878. doi:10.1109/TII.2018.2869429.

Rencher, A., Schaalje, G., 2008. Linear Models in Statistics. John Wiley & Sons.

Si, X.S., Wang, W., Hu, C.H., Zhou, D.H., 2011. Remaining useful life estimation—A review on the statistical data driven approaches. European Journal of Operational Research 213, 1–14. doi:10.1016/j.ejor.2010.11.018.

Sun, F., Li, H., Cheng, Y., Liao, H., 2021. Reliability analysis for a system experiencing dependent degradation processes and random shocks based on a nonlinear Wiener process model. Reliability Engineering & System Safety 215, 107906. doi:10.1016/j.ress.2021.107906.

Wan, R., Bai, Y., 2024. Communication-efficient distributed statistical inference on zero-inflated poisson models. Statistical Theory and Related Fields 8, 81–106. doi:10.1080/24754269.2023.2263721.

Wang, X., Balakrishnan, N., Guo, B., 2015. Residual life estimation based on nonlinear-multivariate Wiener processes. Journal of Statistical Computation and Simulation 85, 1742–1764. doi:10.1080/00949655.2014.898765.

Wu, C.J., 1983. On the convergence properties of the EM algorithm. The Annals of Statistics 11, 95–103.

Xu, A., Shen, L., Wang, B., Tang, Y., 2018. On modeling bivariate Wiener degradation process. IEEE Transactions on Reliability 67, 897–906. doi:10.1109/TR.2018.2791616.

Yao, F., Hu, J., Li, B., Liu, H., Gong, F., 2024. Non-periodic inspection and replacement policy of system subject to non-homogeneous gamma degradation process. Quality and Reliability Engineering International 40, 1165–1181. doi:10.1002/qre.3473.

Ye, Z., Xie, M., 2014. Stochastic modelling and analysis of degradation for highly reliable products. Applied Stochastic Models in Business and Industry 31, 16–32. doi:10.1002/asmb.2063.

Zhai, Q., Xu, A., Yang, J., Zhou, Y., 2024. Statistical modeling and reliability analysis for degradation processes indexed by two scales,. IEEE Transactions on Industrial Informatics 20, 3675–3684. doi:10.1109/TII.2023.3313668.

Zhai, Q., Ye, Z.S., 2018. Degradation in common dynamic environments. Technometrics 60, 461–471. doi:10.1080/00401706.2017.1375994.

Zhai, Q., Ye, Z.S., 2023. A multivariate stochastic degradation model for dependent performance characteristics. Technometrics 65, 315–327. doi:10.1080/00401706.2022.2157881.

Zhang, S., Zhai, Q., Li, Y., 2023. Degradation modeling and RUL prediction with Wiener process considering measurable and unobservable external impacts. Reliability Engineering & System Safety 231, 109021. doi:10.1016/j.ress.2022.109021.

Zhao, H., Chen, Z., Shu, X., Shen, J., Lei, Z., Zhang, Y., 2023. State of health estimation for lithium-ion batteries based on hybrid attention and deep learning. Reliability Engineering & System Safety 232, 109066. doi:10.1016/j.ress.2022.109066.

Zhao, X., Chen, P., Gaudoin, O., Doyen, L., 2021. Accelerated degradation tests with inspection effects. European Journal of Operational Research 292, 1099–1114. doi:10.1016/j.ejor.2020.11.041.

Zhao, X., Liang, Z., Parlikad, A.K., Xie, M., 2022. Performance-oriented risk evaluation and maintenance for multi-asset systems: A Bayesian perspective. IISE Transactions 54, 251–270. doi:10.1080/24725854.2020.1869871.

Zheng, B., Chen, C., Lin, Y., Ye, X., Zhai, G., 2023. Reliability analysis based on a bivariate degradation model considering random initial state and its correlation with degradation rate. IEEE Transactions on Reliability 72, 37–48. doi:10.1109/TR.2022.3172416.

Zhuang, L., Xu, A., Wang, Y., Tang, Y., 2024. Remaining useful life prediction for two-phase degradation model based on reparameterized inverse Gaussian process. European Journal of Operational Research doi:10.1016/j.ejor.2024.06.032.