Failure Analysis for Truncated and Fully Censored Lifetime Data With a Hierarchical Grid Algorithm

QI LI ©, (Member, IEEE), DONGDONG LI, BING HUANG, ZENGQIANG JIANG ©, MINGCHENG E. ©, AND JING MA ©
Department of Mechanical Engineering, Beijing Jiaotong University, Beijing 100044, China
Corresponding author: Zengqiang Jiang (zqjiang@bjtu.edu.cn)

This work was supported in part by the Fundamental Research Funds for the Central Universities of China under Grant 2019JB053, and in part by the Beijing Natural Science Foundation under Grant 9184030.

ABSTRACT In the field of failure analysis and reliability evaluation, truncated and censored lifetime data due to observational constraints and periodic inspection schemes are highly common. The expectation-maximization algorithm is a widely employed approach for the parameter estimations of truncated and censored lifetime data. In this study, a new hierarchical grid algorithm is proposed to estimate the model parameters based on left-truncated and fully-censored data. An improved expectation-maximization algorithm is also adapted under incomplete information. These two methods are compared using Monte Carlo simulations. The confidence intervals corresponding to different quantile methods of the nonparametric bootstrap are compared in terms of coverage probabilities. In a case study, the failure analysis and prediction intervals for individual coupler knuckles of rail wagons are discussed. This simulation study and case study verify the effectiveness of the proposed framework.

INDEX TERMS Lifetime estimation, algorithm design and analysis, maximum likelihood estimation, nonparametric bootstrap method, left-truncation.

I. INTRODUCTION
Lifetime data has been used to evaluate the time-to-failure distribution and to predict remaining useful life (RUL) for a long time in the field of reliability. The quality of the lifetime data collection strongly affects further research. However, incomplete data containing truncation and censored information are highly common due to long-interval inspections, especially during the process of establishing the condition monitoring system.

Censoring is generated by the scenario that researchers cannot fully observe the occurrence of a research event during the observation time window due to realistic limitations. This characteristic is an attribute of individuals in the sampling group, which means that under the same observation conditions, the entry information of a sample is known, but the start-stop information of individuals is (partly) unknown. Unlike censoring, truncation is an attribute of the population. Due to the limitations of observation conditions, the subjects can only be observed when they appear within a particular range, and the existence of individuals who fall outside the certain range cannot be known. In other words, the entry information of the population is incomplete. The books by Cohen [1], Meeker and Escobar [2] provide detail discussions on truncation and censoring.

Many researchers have developed statistical inference methods based on truncated and censored data. Among these methods, the left-truncated and right-censored structure is the most widely studied data scenarios. Hong et al. [3] developed a statistically based analysis and predicted the RUL of high-voltage power transformers for maintenance planning and capital expenditures purposes. In their study, the collected power transformers lifetime data were left-truncated at the starting date of record keeping and right-censored at the ending date of the study. Following studies by these researchers, Balakrishnan and Mitra [4]–[6], Emura and Shiu [7], and Kundu and Mitra [8], [9] worked on likelihood-based inference for left-truncated and right-censored data. Zhao et al. [10] and Bai [11] also studied left-truncated and right-censored data in other areas, such as clinical trials.

Right censoring could be viewed as a special form of interval censoring. Interval-censored observations consist of the
upper and lower bounds of a failure time and often occur in, for example, epidemiological and medical follow-up studies. Shen [12], [13], Wang et al. [14], Shen and Li [15], and Shen et al. [16] proposed a series on studies of the conditional likelihood approach for statistical inference based on left-truncated and interval-censored data in a AIDS cohort study. Balakrishnan and Mitra [17] proposed to replace the continuous checkup with a periodic checkup of power transformers to obtain left-truncated and interval-censored data.

The expectation-maximization (EM) algorithm [18]–[20] has been used by numerous researchers, especially in the context of incomplete data. The EM algorithm consists of two steps: the expectation step (E-step) and the maximization step (M-step). In the E-step, the conditional expectation of the complete data likelihood is obtained, given the observed incomplete data and the current value of the parameter. In the M-step, this expected complete likelihood is maximized by adjusting the parameters. The E-step and M-step are then iterated until reaching the convergence criterion [21].

Balakrishnan and Mitra [4] developed an EM algorithm for fitting a lognormal distribution to the left-truncated and right-censored data. Similar EM algorithms have been developed for Weibull distribution [5] and generalized gamma distribution [6]. Recently, Shen et al. [16] proposed an EM algorithm based on left-truncated and interval-censored data for obtaining the conditional maximum likelihood estimator (cMLE). Balakrishnan and Mitra [17] developed a stochastic expectation-maximization (St-EM) algorithm to estimate the model parameters.

However, the inference of parameter estimation methods and their applications to left-truncated and fully censored data have not been studied extensively. Table 1 summarizes a detailed comparison of the proposed method and the related literature.

## A. REALISTIC LEFT-TRUNCATED AND FULLY CENSORED DATA SCENARIO

The lifetime data of coupler knuckles is collected from a rail wagon maintenance factory, which was established in 2015. The lifetime data is observable only when the unit failed after the year 2015, as careful archival record keeping of knuckle couplers’ lifetime data started in that year. In other words, the lifetime data has been left-truncated at 2015. The recording time-window is from 2015 to 2019, a four-year period. Clearly, the interval is significantly shorter than the average lifetime of a coupler knuckle of 8-9 years under normal operating conditions. However, the customers of the maintenance factory maintain numerous rail wagons, which have been running for decades.

Fig. 1 shows an illustration of the truncation and censoring mechanism based on the observation time-window. For “Subject a”, the service starting time and the failure time of the coupler knuckle are both before 2015. There is no information available (including its existence) for Subject a, as this coupler knuckle is completely outside the observation.

### TABLE 1. Comparison of parameter estimation methods for selected literature.

| Reference | Year | Data type | Lifetime distribution | Parameter estimation methods | Case study |
|-----------|------|-----------|-----------------------|-----------------------------|------------|
| [3]       | 2009 | Left-truncated and right-censored | Weibull | NR method | Power transformers |
| [4]       | 2011 | Left-truncated and right-censored | Lognormal | EM algorithm | Power transformers |
| [17]      | 2019 | Left-truncated and right-censored | Lognormal, Weibull | NR method and EM algorithm | Power transformers |
| [15]      | 2019 | Left-truncated and interval-censored | Lognormal, Weibull | Stochastic EM algorithm | Cancer clinical trials |
| This paper | 2020 | Left-truncated and fully censored | Weibull | HG algorithm and improved EM algorithm | Coupler knuckles |
time-window. For “Subject b”, the coupler knuckle started service before 2015, and it failed between 2015 and 2019. Hence, Subject b is from a left-truncated distribution. For “Subject c”, the service starting time is after 2015 and the failure time is before 2019, making it a complete observation data entry. For “Subject d”, the coupler knuckle was installed before 2015, and it was still working as of 2019. Hence, Subject d is both left-truncated and right-censored.

Fig. 2 shows the mechanism for generating different censored data based on the maintenance policy. All the observable units are pulled back to the same starting point (the beginning of their service time). It is known that coupler knuckles are inspected every two years. “Subject e” failed at the first inspection, and its failure should have happened sometime before the first inspection. Hence, Subject e is left-censored at the first inspection. The “Subject f” passed the first inspection and failed at the second inspection. The failure of this subject should have happened sometime between the first inspection and the second inspection. Hence, Subject f is interval-censored. “Subject g” remained in service after the third inspection, and its failure should be sometime after the third inspection. Hence, Subject g is right-censored at the third inspection. It is worth noting that these collected data are also used in the case study section to illustrate the proposed models.

### B. LIKELIHOOD CONSTRUCTION

Likelihood represents the probability of data generation. From this perspective, truncation could be viewed as a conditional sampling distribution, and censoring could be viewed as the cumulative likelihood of an interval. The likelihood construction of truncated and censored data has been well established in the literature [22], [23]. In this section, a concise summary of related concepts is given to avoid any misunderstanding.

The lifetime of a unit is denoted by the random variable $t$. Let $i$ be the index of the unit and $j$ be the index of inspection. Let $t_{ij}$ denote the lifetime of the $i^{th}$ unit in the $j^{th}$ inspection. For a unit installed before 2015 (the truncation point), let $\tau_i^L$ denote the time between the year of installation and 2015. For a unit that is still working after 2019 (the right-censored point), let $c_i^R$ denote the time between the year of installation and 2019. Let $\alpha_i$ denote the truncation indicator, that is,

$$\alpha_i = \begin{cases} 0, & \text{if the observation is truncated.} \\ 1, & \text{if the observation is not truncated.} \end{cases} \quad (1)$$

Let $\beta_i$ denote the censoring indicator, such that:

$$\beta_i = \begin{cases} 0, & \text{if the observation is interval censored.} \\ 1, & \text{if the observation is not interval censored.} \end{cases} \quad (2)$$

Further, let $S_1$ and $S_2$ be two sets of units, where $S_1$ is the set of units installed on or after 2015 and $S_2$ is the set of units installed before 2015. The likelihood function for the lifetime data is represented as

$$L(\theta \mid \text{DATA}) = \prod_{S_1} \left[1 - F(c_i^R; \theta)\right]^{\beta_i} \times \left[F(t_i^j; \theta) - F(t_i^{j-1}; \theta)\right]^{1 - \beta_i}$$

$$\times \prod_{S_2} \left[1 - F(\tau_i^L; \theta)\right]^{\beta_i} \times \left[F(t_i^j; \theta) - F(t_i^{j-1}; \theta)\right]^{1 - \beta_i} \quad (3)$$

After simplification, the log-likelihood function for lifetime data is given as

$$\log L(\theta \mid \text{DATA}) = \sum_{i=1}^{n} \beta_i \log \left[1 - F(c_i^R; \theta)\right]$$

$$+ (1 - \beta_i) \left[F(t_i^j; \theta) - F(t_i^{j-1}; \theta)\right]$$

$$- \sum_{i=1}^{n} \alpha_i \log \left[1 - F(\tau_i^L; \theta)\right] \quad (4)$$

### III. PARAMETER ESTIMATION METHODS

In this section, the idea and procedure of the proposed hierarchical grid algorithm are introduced. An improved version of the EM algorithm that can handle fully censored data is also presented for comparison purposes.

#### A. PROPOSED HIERARCHICAL GRID ALGORITHM

For right-censored lifetime data, the log-likelihood function is relatively simple, and the analytical formula can be obtained by EM algorithm with Newton-Raphson method. For right-censored and left-truncated lifetime data, approximate solutions could be obtained by EM algorithm with the missing information principle [19]. However, the data structure studied in this paper contains fully censored data with a high truncation ratio, and it is highly difficult to obtain the derivative of likelihood functions. Since the proposed log-likelihood function model is convex—that is, it has a unique peak value, and the function value changes gently near the peak value—this paper proposes a hierarchical grid algorithm inspired by the idea of the gradient descent algorithm.
The gradient descent algorithm is one of the simplest and oldest methods for solving unconstrained optimization problems. Based on this approach, many effective algorithms have been improved. The gradient descent method uses the negative gradient direction as the search direction. Thus, in the proposed HG algorithm, the log-likelihood function is converted to its negative value, and then the maximum of the likelihood function is converted to obtain the minimum of the likelihood function. The multivariate function $-Q(\theta)$ is defined to be equivalent to the log likelihood function $\log L(\theta | DATA)$. Then, solving the optimal parameters of the log-likelihood function is equivalent to solving the extremum of this multivariate function. Suppose the parameters to be solved follow normal distributions. Then the multivariate function to be solved is a binary function.

In view of the monotonicity of the likelihood function, if the maximum of the current-log-likelihood function is located in the center of a rectangular grid lattice, the global maximum exists in the four-unit domains adjacent to the current maximum. In short, the HG algorithm first searches the extremum region. As shown in Fig. 3, the HG is roughly divided into the following two stages.

1) STAGE 1. FIND THE REGION WHERE THE EXTRENUM EXISTS

**Step 1:** Set the current minimum point $\theta_i = (\mu_i, \sigma_i)$. Set the current interval length $\Delta \theta_i = (\Delta \mu_i, \Delta \sigma_i)$, $0 < \Delta \theta_i \ll \theta^*, i = 0, 1, 2, \ldots$, where $\theta^*$ denote the global optimal value. Construct an equally spaced initial collection of points $\Theta = \{\theta_{i,1 \times 1}, \theta_{i,1 \times 2}, \ldots, \theta_{i,1 \times n}, \theta_{i,2 \times 1}, \ldots, \theta_{i,n \times (n-1)}, \theta_{i,n \times n}\}$. Divide $\Theta$ into two complementary sets $A$ and $A'$, where $A = \{\theta_{i,1 \times s} : t mod n > a, s mod n > a\}$, where $a = 0, 1$ or 2.

Define the diagonal length of the unit grid as the iteration step size of the corresponding parameter, and its size is $\sqrt{\Delta \mu_i^2 + \Delta \sigma_i^2}$. The description of the symbols in Fig. 4 is as follows:

$$
\mu_{i1} = \mu_i - [f_i] \Delta \mu_i; \quad \mu_{in} = \mu_i + [f_i] \Delta \mu_i
$$
$$
\sigma_{i1} = \sigma_i - [f_i] \Delta \sigma_i; \quad \sigma_{in} = \sigma_i + [f_i] \Delta \sigma_i
$$
$$
f_i = \frac{C}{1 + e^{-i}} \times \frac{m - 1}{2}, \quad C \geq 2, \quad m \geq 3, \quad i = 0, 1, 2, \ldots,
$$

where $\sigma_{i1}$ and $\sigma_{in}$ are the minimum and maximum values of the collection of points $\Theta$, in the $x$ direction, $\mu_{i1}$ and $\mu_{in}$ are the minimum and maximum values of collection of points $\Theta$, in the $\mu$ direction, $f_i$ is the step coefficient, and $C$ is a constant. The larger $C$ is, the larger the coverage of the collection of points $\Theta$. The constant $m$ determines the coverage of the initial point set. Let $\theta_{i,1 \times 1}, \theta_{i,1 \times n}, \theta_{i,n \times 1}$, and $\theta_{i,n \times n}$ be the lower-left endpoint, the upper-left endpoint, the lower-right endpoint, and the upper-right endpoint of the collection of points, respectively.

![FIGURE 3. Algorithm flowchart of the proposed method.](image)

In Stage 1, $\theta_0$ assumes the prior parameter, such as the mean and standard deviation of the sample. Let $\Delta \theta_i = \Delta \theta_{i-1}, i = 1, 2, 3, \ldots$.

**Step 2:** As shown in Fig. 4, we obtain the current minimum point $\theta_{i+1}$ in $\Theta$. If $\theta_{i+1} \in A$, repeat Stage 1. If $\theta_{i+1} \in A'$, start Stage 2. In addition, record the current minimum point $\theta_{i+1}$ as $\theta_k, k = i + 1$.

2) STAGE 2. OBTAIN THE EXTRENUM IN THE REGION WHERE THE EXTRENUM EXISTS

**Step 3:** Set the current minimum point $\theta_j = (\mu_j, \sigma_j), j = k, k + 1, \ldots$. Set the current interval length $\Delta \theta_j = (\Delta \mu_j, \Delta \sigma_j)$, $0 < \Delta \theta_j \ll \theta^*, i = 1, 2, \ldots$. Let parameter $\theta^*$ denote the global optimal value. Construct the equally spaced initial collection of points $\Theta_j = \{\theta_{j,1 \times 1}, \theta_{j,1 \times 2}, \ldots, \theta_{j,1 \times n}, \theta_{j,2 \times 1}, \ldots, \theta_{j,n \times (n-1)}, \theta_{j,n \times n}\}$. The description of the symbols in Fig. 5 is as follows:

$$
\mu_{j1} = \mu_j - [f_j] \Delta \mu_j; \quad \mu_{jn} = \mu_j + [f_j] \Delta \mu_j
$$
$$
\sigma_{j1} = \sigma_j - [f_j] \Delta \sigma_j; \quad \sigma_{jn} = \sigma_j + [f_j] \Delta \sigma_j
$$
$$
f_j = \frac{m - 1}{2}, \quad m \geq 3; \quad \Delta \theta_j = \frac{1}{C} \Delta \theta_{j-1}, 2 \leq C' \leq n - 1
$$

(6)
where $\sigma_{j1}$ and $\sigma_{jn}$ are the minimum and maximum values of the collection of points $\Theta_j$ in the $\sigma$ direction, $\mu_{j1}$ and $\mu_{jn}$ are the minimum and maximum values of the collection of points $\Theta_j$ in the $\mu$ direction, $f_j$ is the step coefficient, and $C'$ is a constant. The larger $C'$ is, the larger the coverage of the collection of points $\Theta_j$. Let $\theta_{j,1\times1}, \theta_{j,1\times n}, \theta_{j,n\times1}$, and $\theta_{j,n\times n}$ be the lower-left endpoint, the upper-left endpoint, the lower-right endpoint, and the upper-right endpoint of the collection of points, respectively.

**Step 4:** As shown in Fig. 5, we obtain the current minimum point $\theta_{i+1}$ in $\Theta_j$. Repeat Stage 2 when $|\theta_{i+1} - \theta_i| > \varepsilon$. The iterations continue until $|\theta_{i+1} - \theta_i| \leq \varepsilon$. It is worth noting that this algorithm requires that the precision of the two parameters both meet the requirements, rather than only one parameter.

In summary, the proposed HG algorithm shares three basic elements with the gradient descent method, namely, the initial point, the descending direction and the descending step size. Table 2 compares the three elements of the proposed HG algorithm and the gradient descent method.

### B. IMPROVED EM ALGORITHM

The EM algorithm is a generalized iterative algorithm that computes a maximum likelihood estimation of a parameter in the case of incomplete data. The approach corrects the censored data using the parameter estimation from the previous step and then updates the parameter estimation with the maximum likelihood criterion. The EM algorithms proposed in the literature were mainly focused on the case of left-truncated and right-censored data. Therefore, EM algorithms need to be improved to handle interval-censored data. In this section, an improved EM algorithm is proposed to analyze left-truncated and fully censored data with different censoring types. In general, the partial derivative of the function can be calculated to obtain the general expression of the extreme point of the

---

**TABLE 2. Feature comparison for the gradient descent algorithm and the hierarchical grid algorithm.**

| Feature               | The gradient descent algorithm [24-25]                                      | The hierarchical grid algorithm                      |
|-----------------------|----------------------------------------------------------------------------|------------------------------------------------------|
| Initial point         | Random or all zero vectors                                                 | Take the basic statistic of the left endpoint of the lifetime as the prior parameter. |
| Descending direction  | Take the partial derivative of the likelihood function to obtain the gradient vector, and take the fastest descending direction as the descending direction. | Calculate the difference between the maximum value point and the central point of the current likelihood function in the parameter domain, and connect the two points as the descending direction. When the domain of the valley point is not found, the step size is gradually increased, otherwise the step size is decreased gradually. |
| Descending step size  | The smaller the step size and the slower the advance, the closer to the target value. |                                                     |
function. If the multivariate function is complex and it is difficult to derive the partial derivative of the function, then the numerical solution can be used to solve the approximate optimal solution of the function.

Let $\mathcal{Y} = \{y_1, \ldots, y_{k+1}, \ldots, y_n\}$ be the complete log-transformed lifetime data. It is assumed that the lifetime of the units follows the normal distribution with a probability density function (PDF) as:

$$ f(t) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(t-\mu)^2}{2\sigma^2}\right] $$  \hspace{1cm} (7)

Since certain lifetime data points are not observed exactly, let $\mathcal{Z} = \{z_1, \ldots, z_k, z_{k+1}^+, \ldots, z_n^+\}$ be the observed function of $\mathcal{Y}$, and among the points, $z_{k+1}^+, \ldots, z_n^+$ are the censored data. Define $z_{k+1}^+ = (z_{k+1}, u_{k+1})$, where

$$ u_{k+1} \begin{cases} < \text{inf}, & \text{if the observation is interval censored;} \\ = \text{inf}, & \text{if the observation is right censored.} \end{cases} $$  \hspace{1cm} (8)

Based on the left-truncated and full-censored dataset $\mathcal{Z}$, this paper obtains the model parameters $\theta = (\mu, \sigma)$ by the following improved EM algorithm.

**E-step:** Calculate the conditional expectation of a log-likelihood function.

$$ Q(\theta | \theta^{(i)}) = -n \ln \sqrt{2\pi} - n \ln \sigma - \frac{1}{2\sigma^2} \sum_{j=1}^{n} E[(y_j - \mu^2 | z, \mu, \sigma)] $$

$$ - \sum_{S_1} \ln \left(1 - \Phi \left(\frac{\tau_j - \mu}{\sigma}\right)\right), $$

where $\mu, \sigma$ are the parameter values at the beginning of the $i + 1^{th}$ iteration and $S_1$ is the left truncated data sets. To simplify the function, the following expressions are defined:

$$ a_{ij} = \frac{\int_{z_j}^{y_j} x f(y; \mu, \sigma, \sigma^{(i)}) \, dy}{\int_{z_j}^{y_j} f(y; \mu, \sigma, \sigma^{(i)}) \, dy} $$  \hspace{1cm} (10)

$$ b_{ij} = \frac{\int_{z_j}^{y_j} x^2 f(y; \mu, \sigma, \sigma^{(i)}) \, dy}{\int_{z_j}^{y_j} f(y; \mu, \sigma, \sigma^{(i)}) \, dy} $$  \hspace{1cm} (11)

where $i = 1, 2, \ldots$ and $j = k+1, \ldots$. Then, the summation part of the log likelihood function can be expressed as

$$ E[\ln f(y)] = \sum_{j=1}^{k} (z_j - \mu)^2 $$

$$ + \sum_{j=k+1}^{n} \int_{z_j}^{y_j} (y - \mu)^2 f(y; \mu, \sigma, \sigma^{(i)}) \, dy $$

$$ = \sum_{j=1}^{k} (z_j^2 - \mu^2) $$

$$ + \sum_{j=k+1}^{n} (b_{ij} - 2\mu a_{ij} + \mu^2) $$  \hspace{1cm} (12)

**M-step:** Obtain the maximum point of the function $Q(\theta | \theta^{(i)})$.

First, the following equations are defined to further simplify the likelihood function:

$$ \psi(t) = \Phi \left(\frac{t - \mu^{(i)}}{\sigma^{(i)}}\right) $$  \hspace{1cm} (13)

$$ \Gamma(t) = \frac{1}{1 - \psi(t) \sqrt{2\pi}} \exp\left[-\frac{(t - \mu^{(i)})^2}{2\sigma^{(i)^2}}\right] $$  \hspace{1cm} (14)

$$ a'(t) = \frac{\int_{t}^{\infty} yf(y; \mu^{(i)}, \sigma^{(i)}) \, dy}{\int_{t}^{\infty} f(y; \mu^{(i)}, \sigma^{(i)}) \, dy} $$  \hspace{1cm} (15)

$$ b'(t) = \frac{\int_{t}^{\infty} y^2 f(y; \mu^{(i)}, \sigma^{(i)}) \, dy}{\int_{t}^{\infty} f(y; \mu^{(i)}, \sigma^{(i)}) \, dy} $$  \hspace{1cm} (16)

Then, it is easy to verify that $a'(t) = \Gamma(t) + \mu^{(i)}$ and $b'(t) = (t + \mu^{(i)}) \Gamma(t) + \mu^{(i)^2} + \sigma^{(i)^2}$. Thus, (9) and (10) can be updated as:

$$ a_{ij} = \left[ a'(z_j) (1 - \psi(z_j)) - a'(u_j) (1 - \psi(u_j)) \right] $$

$$ b_{ij} = \left[ b'(z_j) (1 - \psi(z_j)) - b'(u_j) (1 - \psi(u_j)) \right] $$  \hspace{1cm} (17)

If the incomplete data does not contain truncated data, then the likelihood function can be partially derived to obtain the general formula of the parameter estimation as:

$$ \mu^{(i+1)} = \frac{1}{n} \left( \sum_{j=1}^{k} z_j + \sum_{j=k+1}^{n} a_{ij} \right) $$  \hspace{1cm} (19)

$$ \sigma^{(i+1)} = \frac{1}{n} \left( \sum_{j=1}^{k} z_j^2 + \sum_{j=k+1}^{n} b_{ij} - n\mu^{(i+1)^2} \right) $$  \hspace{1cm} (20)

If the incomplete data contains truncated data, then it is difficult to solve the partial derivative of the likelihood function. Therefore, the Nelder-Mead simplex algorithm is used to obtain the approximate optimal solution. Through improvement, the improved EM algorithm can solve the parameters based on left-truncated and full-censored data.

**IV. NUMERICAL VALIDATION**

In this section, the validity, superiority, and robustness of the proposed HG algorithm are discussed. The comparisons of different bootstrap confidence interval estimating methods are also discussed based on a simulation study.

**A. NUMERICAL VALIDATION BASED ON LEFT-TRUNCATED AND RIGHT-CENSORED DATA**

To verify the proposed HG algorithm in the case of left-truncated and right-censored data, the EM algorithm of log-normal distribution introduced by Balakrishnan and Mitra [4] is employed for comparison. The sample mean $\hat{\mu} = 3.03$ and standard deviation $\hat{\sigma} = 0.41$ are taken as the initial values to obtain the target parameter estimations as $\hat{\mu} = 3.481$ and $\hat{\sigma} = 0.456$. The maximization step of the EM algorithm involves an approximation using a Taylor series expansion.
The tolerance is indicated by the Cartesian distance between the corresponding parameter point of the current iteration and that of the previous iteration. The iterative process and results of the two algorithms are shown in Table 3.

As shown in Table 3, the parameter estimation results of the HG algorithm and the improved EM algorithm are almost the same, which indicates that the HG algorithm is also suitable for dealing with left-truncation and right-censored data. The HG algorithm takes fewer steps to reach convergence than does the improved EM algorithm. The HG algorithm only needs four iterations to reach a tolerance of less than 0.001. These results reveal that the HG algorithm converges more rapidly than does the improved EM algorithm.

**B. SIMULATION COMPARISONS FOR FULLY CENSORED DATA**

To verify the applicability of the proposed HG algorithm to left-truncated and fully censored data, this section adds interval censoring to the simulation data from Balakrishnan and Mitra [4]. The improved EM algorithm is used for verification and comparison. Per Hong et al. [3], the truncation time is fixed as 1980, and the censoring time has been fixed as 2008. The specific simulation steps are summarized as follows.

Step 1: To analyze the influences of the percentage of truncation on the algorithm, the truncation percentages are set to be 30% or 60%.

Step 2: The index of the installation year is split into two parts: (1960-1979) and (1980-1995). Then, the installation years are simulated according to the sampling probabilities shown in Table 4.

Step 3: Set the sample size n. In this study, three levels of sample size are considered: n = 75, 100, and 200.

Step 4: Simulate the lifetime from a lognormal distribution with two different settings of the parameter vector (μ, σ) as (3.5, 0.5) or (3.0, 0.2).

Step 5: Combine the lifetime with the corresponding installation year, and then, the year of failure of the machine is obtained. If the year of failure is before 1980, we discard the unobserved machine and generate a new observed one with the new installation year and lifetime. If the year of failure is between 1980 and 2008, then this entry is interval censored. If the year of failure is after 2008, this entry is right-censored.

Step 6: Define two parameters following uniform distribution, i.e., \( \delta_1, \delta_2 \sim U(0.05, 0.1) \). If the machine is interval-censored, then add these two parameters to the endpoints of the lifetime, i.e., \( (L_i, R_i] = (X_i - \delta_1, X_i + \delta_2] \). If the machine is right-censored, then add the first parameter to the left endpoint of the lifetime, i.e., \( (L_i, R_i] = (X_i - \delta_1, +\infty] \).

Since the generated sample data is fully censored data, the sample mean and standard deviation of the left endpoint of the lifetime can be taken as the initial value. The results including the bias and mean square error of the parameter estimation, average number of iterations needed, and average tolerance limit based on a 1000-time Monte Carlo simulation are summarized in Table 5. EM corresponds to the improved EM algorithm for adapting left-truncated and fully censored data, while HG corresponds to the hierarchical grid algorithm.

The simulation results of Table 5 show that the bias and mean square error values of the model parameters of the two methods are approximately the same. When the parameter has a large true value, the average number of iterations needed and the average tolerance limit of the HG algorithm are significantly lower than those of the EM algorithm, which means that the iteration step size of the HG algorithm is larger than those of the EM algorithm.

Both algorithms use conditional probability to correct the deviation caused by the truncation so that the estimation results should have less deviation from the true value. As the number of data point increases, the deviation between the estimation result and the true value decreases. In such cases, the truncation ratio has effects on the parameter estimation.

Based on the above analysis, the stability of both methods is good, but the HG algorithm has a faster convergence speed. For sets of simulation data with small true values of model parameters, both algorithms show better stability and convergence performance.

**C. DISCUSSION ON INITIAL VALUE AND STEP SIZE**

The above discussions verify the effectiveness of the HG algorithm. Based on the derivation and the simulation results, it is easy to conclude that the EM algorithm can find the general formula of the parameter estimate without considering the truncated data. In this section, the influence of initial value and step size is discussed.

The EM algorithm depends on the choice of initial values and shows a problem of convergence when a good choice is not given. Table 6 presents the average number of iterations needed and the running time of the above two algorithms. The running time is counted when the algorithm ends iteratively, needed and the running time of the above two algorithms. The above discussions verify the effectiveness of the HG algorithm. Based on the derivation and the simulation results, it is easy to conclude that the EM algorithm can find the general formula of the parameter estimate without considering the truncated data. In this section, the influence of initial value and step size is discussed.

The EM algorithm depends on the choice of initial values and shows a problem of convergence when a good choice is not given. Table 6 presents the average number of iterations needed and the running time of the above two algorithms. The running time is counted when the algorithm ends iteratively, that is, when the parameter accuracy requirement is met. This discussion is based on the simulation data in the above section with \( \mu = 3.5, \sigma = 0.5 \), and \( n = 200 \). This study selects...
TABLE 5. Bias ($B$), mean square error ($MSE$), average number of iterations ($AI$), and average tolerance limit ($AT$) for the lognormal distribution using two different methods (EM$_I$ and HG).

| $(\mu, \sigma)$ | Truncation (%) | Method | $B(\mu)$ | $B(\sigma)$ | $MSE(\mu)$ | $MSE(\sigma)$ | $AI$ | $AT$ |
|-----------------|----------------|--------|----------|------------|-------------|--------------|------|------|
| $(3.5, 0.5)$    | 30             | EM$_I$ | 4.27E-03 | -4.88E-03 | 6.50E-03    | 5.19E-03    | 24.9 | 6.01E-05 |
|                 |                | HG     | 4.42E-03 | -4.72E-03 | 6.52E-03    | 5.20E-03    | 8.01 | 7.62E-05 |
|                 | 60             | EM$_I$ | 2.18E-03 | -4.63E-03 | 5.15E-03    | 4.33E-03    | 17.4 | 5.40E-05 |
|                 |                | HG     | 2.24E-03 | -4.52E-03 | 5.15E-03    | 4.53E-03    | 7.73 | 1.02E-04 |
| $(3.0, 0.2)$    | 30             | EM$_I$ | 1.27E-03 | -4.33E-03 | 6.40E-04    | 3.69E-04    | 6.38 | 1.03E-05 |
|                 |                | HG     | 1.27E-03 | -4.33E-03 | 6.40E-04    | 3.69E-04    | 4.77 | 1.15E-04 |
|                 | 60             | EM$_I$ | -4.40E-05| -5.37E-03 | 7.05E-04    | 3.84E-04    | 5.38 | 1.44E-04 |
|                 |                | HG     | -4.10E-05| -5.37E-03 | 7.05E-04    | 3.84E-04    | 4.96 | 9.62E-05 |
| $(3.5, 0.5)$    | 30             | EM$_I$ | 1.99E-03 | -7.28E-03 | 5.17E-03    | 3.90E-03    | 24.6 | 5.91E-05 |
|                 |                | HG     | 2.13E-03 | -7.12E-03 | 5.18E-03    | 3.91E-03    | 8.03 | 7.72E-05 |
|                 | 60             | EM$_I$ | 7.93E-03 | -2.13E-03 | 3.91E-03    | 3.47E-03    | 17.6 | 5.62E-05 |
|                 |                | HG     | 7.99E-03 | -2.02E-03 | 3.91E-03    | 3.48E-03    | 7.77 | 1.03E-04 |
| $(3.0, 0.2)$    | 30             | EM$_I$ | 1.48E-03 | -4.85E-03 | 4.75E-04    | 3.00E-04    | 6.36 | 9.03E-06 |
|                 |                | HG     | 1.48E-03 | -4.84E-03 | 4.75E-04    | 3.00E-04    | 4.75 | 1.17E-04 |
|                 | 60             | EM$_I$ | 7.43E-04 | -5.38E-03 | 5.18E-04    | 3.01E-04    | 5.37 | 1.03E-05 |
|                 |                | HG     | 7.48E-04 | -5.38E-03 | 5.18E-04    | 3.01E-04    | 4.94 | 9.72E-05 |
| $(3.5, 0.5)$    | 30             | EM$_I$ | 2.33E-03 | -1.32E-03 | 2.35E-03    | 1.96E-03    | 24.4 | 6.11E-05 |
|                 |                | HG     | 2.46E-03 | -1.17E-03 | 2.36E-03    | 1.96E-03    | 8.00 | 8.12E-05 |
|                 | 60             | EM$_I$ | 4.27E-03 | -1.33E-03 | 2.02E-03    | 1.94E-03    | 17.3 | 6.01E-05 |
|                 |                | HG     | 4.33E-03 | -1.22E-03 | 2.02E-03    | 1.94E-03    | 7.73 | 1.06E-04 |
| $(3.0, 0.2)$    | 30             | EM$_I$ | 1.18E-03 | -4.63E-03 | 2.46E-04    | 1.50E-04    | 6.35 | 7.02E-06 |
|                 |                | HG     | 1.18E-03 | -4.63E-03 | 2.46E-04    | 1.50E-04    | 4.76 | 1.16E-04 |
|                 | 60             | EM$_I$ | 9.90E-05 | -4.53E-03 | 2.96E-04    | 1.38E-04    | 5.36 | 1.12E-05 |
|                 |                | HG     | 1.02E-04 | -4.52E-03 | 2.96E-04    | 1.38E-04    | 4.99 | 9.34E-05 |

TABLE 6. Number of iterations needed and the running time of the EM$_I$ algorithm and the HG algorithm.

| \(\mu_0\) | \(\sigma_0\) | The number of iterations | The running time (s) |
|----------|-------------|--------------------------|---------------------|
|          |             | EM$_I$ | HG | EM$_I$ | HG |                      |
| 10       | 0.01        | -     | 10 | -     | 27.99                   |
|          | 5.00        | 41    | 10 | 3.66  | 0.57                     |
| 0.1      | 0.01        | -     | 8  | -     | 0.47                     |
|          | 5.00        | 40    | 10 | 3.43  | 0.61                     |

As shown in Table 6, the EM$_I$ algorithm takes more steps to converge, and it even cannot converge in some cases. In contrast, the HG algorithm requires fewer iterations, has a shorter running time, and exhibits a good robustness.

Considering that the parameter value of the first solution is in the edge region, setting the size of the new parameter domain directly determines whether the global minimum point can be found. This section also compares and analyzes the effect of the step size on seeking the global minimum point.

Set the step size coefficients of the HG algorithm to be \(m = 25\) and \(c = 10\). Fig. 6 shows the process of obtaining the global minimum value according to the step size coefficient function \(f_i = \frac{c}{1 + e^{-i}} \times \frac{m-1}{2}\) and constant, i.e., \(f_i = \frac{m-1}{2}\). Evidently, the change in the step length effectively reduces the number of iterations of the algorithm and helps the algorithm to converge quickly. The authors suggest to further study the effect of the step size coefficient function on the iterative step.

The proposed HG algorithm is similar to the Nelder-Mead method [26], since both of them are numerical optimization methods. The HG algorithm combines the ideas of simplex method and gradient descent method. In the first phase of the HG algorithm, the number of grid points is relatively large, which consumes more computing resources. The HG algorithm searches the grid points to get the optimal solution. The values of grid points are enumerated and compared to determine the peak domain, and the global optimal solution is obtained within the peak domain. It is worth noting the feature of global convergence of the HG algorithm is under the assumption that the log-likelihood function is strictly convex.
convex near the peak with a flat unimodal area. In summary, the number of iterations of the HG algorithm is significantly less than that of the improved EM algorithm, and the proposed method converges more quickly when given a far-off initial value. The HG algorithm has stronger applicability and can obtain a global optimal solution.

**D. COMPARISONS AMONG BOOTSTRAP CONFIDENCE INTERVALS**

The bootstrap method was originally developed by B. Efron as a method for deriving the standard error of arbitrary estimates, and it can be divided into the nonparametric bootstrap approach and the parametric bootstrap approach [27]. The parametric bootstrap method makes assumptions about the underlying distribution. However, when dealing with data involving censoring or truncation, the parametric bootstrap approach requires specifications as to how the censored and truncated data were generated. Since the distribution assumptions on the truncation time and censoring time are not available for most cases, the sample is fully censored, and the sample size is sufficiently large, this study employed the nonparametric bootstrap method to obtain the parameter interval estimations.

The classical approach for constructing the nonparametric bootstrap confidence intervals of a parameter of interest is to use the “appropriate” quantiles of the empirical bootstrap distribution of that parameter estimator. In this study, four approaches of quantile selection are considered, namely, the simple percentile method (SP), the BCa percentile method, the BC percentile method, and the nonparametric basic bootstrap method (BM). Table 7 and Table 8 present the coverage probabilities of the four confidence interval methods based on the nonparametric bootstrap for parameters $\mu$ and $\sigma$ with different percentages of truncation and different nominal confidence levels.

From the results presented in Table 7, it is observed that for the parameter $\mu$, the coverage probability of BM is closest to the nominal confidence level when the true value of the parameter is large. The coverage probability of SP is closer to the nominal confidence level when the true value is small. When the truncation ratio is higher, the coverage probability of BM is closer to the nominal confidence interval, and when the truncation ratio is lower, SP is closer to the nominal confidence interval. The choice of confidence level has an impact on the coverage probability. As the sample size increases, the coverage probability of SP for different combinations is closer to the nominal confidence interval. It is worth mentioning that the difference among the four methods exists in the thousandth position for the parameter $\mu$, which can be ignored for most of the cases. From the results presented in Table 8, it is observed that for the parameter $\sigma$, the coverage probability of the BCa bootstrap method is closest to the
TABLE 8. Coverage probabilities for the two asymptotic confidence intervals for $\sigma$ with different nominal confidence levels and different simulation settings.

| $\sigma$ | Truncation (%) | Nominal CL (%) | Coverage probability |
|----------|----------------|----------------|---------------------|
|          |                |                | SP                  | BCa                  | BC                  | BM                  |
| $n=75$   |                |                |                     |                      |                     |                     |
| 0.5      | 30             | 90             | 0.836               | 0.865               | 0.859               | 0.839               |
|          | 95             |                | 0.890               | 0.925               | 0.920               | 0.882               |
|          | 60             | 90             | 0.846               | 0.879               | 0.869               | 0.840               |
|          |                | 95             | 0.907               | 0.914               | 0.912               | 0.889               |
| 0.2      | 30             | 90             | 0.845               | 0.867               | 0.864               | 0.849               |
|          |                | 95             | 0.887               | 0.930               | 0.920               | 0.890               |
|          | 60             | 90             | 0.823               | 0.851               | 0.844               | 0.822               |
|          |                | 95             | 0.874               | 0.911               | 0.902               | 0.885               |
| $n=100$  |                |                |                     |                      |                     |                     |
| 0.5      | 30             | 90             | 0.862               | 0.880               | 0.872               | 0.849               |
|          | 95             |                | 0.905               | 0.932               | 0.931               | 0.890               |
|          | 60             | 90             | 0.869               | 0.879               | 0.876               | 0.856               |
|          |                | 95             | 0.917               | 0.927               | 0.927               | 0.906               |
| 0.2      | 30             | 90             | 0.845               | 0.870               | 0.869               | 0.844               |
|          |                | 95             | 0.895               | 0.931               | 0.923               | 0.895               |
|          | 60             | 90             | 0.848               | 0.871               | 0.867               | 0.850               |
|          |                | 95             | 0.906               | 0.908               | 0.923               | 0.905               |
| $n=200$  |                |                |                     |                      |                     |                     |
| 0.5      | 30             | 90             | 0.871               | 0.872               | 0.869               | 0.859               |
|          |                | 95             | 0.922               | 0.930               | 0.926               | 0.909               |
|          | 60             | 90             | 0.861               | 0.890               | 0.883               | 0.858               |
|          |                | 95             | 0.927               | 0.930               | 0.929               | 0.921               |
| 0.2      | 30             | 90             | 0.836               | 0.860               | 0.858               | 0.844               |
|          |                | 95             | 0.887               | 0.908               | 0.901               | 0.890               |
|          | 60             | 90             | 0.826               | 0.877               | 0.864               | 0.832               |
|          |                | 95             | 0.906               | 0.922               | 0.922               | 0.895               |

nominal confidence level. In addition, the coverage probabilities of the different methods present significant differences in the interval estimation of the parameter $\sigma$. In summary, the BCa bootstrap method is the preferred approach.

Based on the simulation study, the proposed HG algorithm can handle left-truncated and fully censored data. The parameter estimation accuracy of the algorithm is at the same level as that of the improved EM algorithm. The convergence speed of the algorithm is generally higher than that of the EM algorithm. The dependence of the algorithm on the initial value is more robust than that of the EM algorithm.

V. CASE STUDY

The lifetime data of coupler knuckles mentioned in the previous section is employed as a case study to illustrate the proposed method and its applications. The coupler knuckle is a component used to achieve the connection between a locomotive and a car or between cars in a rail wagon. This component transmits traction and impact forces and maintains a certain distance between the vehicles. The failure of a coupler knuckle threatens the safety of rail transport.

In this case, randomly selected maintenance data by a fixed observation interval are collected. Based on the collected data and the methodological framework above, this section provides the lifetime analysis model for the coupler knuckle data and the predictions for the RUL of individuals. The cumulative number of failures for the overall population of coupler knuckles in service is also analyzed as a function of time.

The dataset used in this study contains 6724 observations including 3872 failures. Table 9 gives a summary of four types of lifetime data from different manufacturers. Except for the MB manufacturer group, most of the lifetime data of coupler knuckles are truncated. The overall truncation accounts for 52.62%.

The Weibull distribution is the most widely used lifetime distribution for time-to-failure modeling. The lognormal distribution is suitable to describe the time to fracture fatigue crack growth in metals. In addition, Xue [28] proposed that the fatigue life of the fatigue zone of the hook body and that the coupler knuckle follows a normal distribution when considering the truncated and censored data. In this study, these three distributions are analyzed and discussed.

TABLE 9. Summary of the four types of lifetime data for different manufacturers.

| MFG | F&T | F&T' | C&T | C&T' | Total |
|-----|-----|------|-----|------|-------|
| MA  | 312 | 440  | 217 | 412  | 1381  |
| MB  | 1   | 167  | 4   | 113  | 1265  |
| MC  | 788 | 227  | 119 | 64   | 1198  |
| MD  | 550 | 120  | 210 | 309  | 1189  |
| ME  | 444 | 104  | 88  | 37   | 673   |
| MF  | 141 | 49   | 72  | 44   | 306   |
| Other | 480 | 49   | 112 | 51   | 692   |
TABLE 10. Parameter estimations of the lognormal, Weibull and normal distributions.

| Distribution      | Lognormal | Weibull  | Normal   |
|-------------------|-----------|----------|----------|
| Location parameter| 7.298     | 7.555    | 1671.054 |
| Scale parameter   | 0.654     | 0.492    | 850.546  |
| Log-likelihood value | -6941.653 | -6509.294 | -6490.529 |
| AIC               | 13887.306 | 13022.588 | 12985.058 |

It is straightforward to obtain the log-likelihood function of the normal distribution by substituting specific parameters, as shown in the following equation:

\[
\log L (\mu, \sigma | DATA) = \sum_{i=1}^{n} \beta_i \log \left[ 1 - \Phi_{nor} \left( \frac{c_i - \mu}{\sigma} \right) \right] + (1 - \beta_i) \left[ \Phi_{nor} \left( \frac{r_i - \mu}{\sigma} \right) - \Phi_{nor} \left( \frac{r_{i-1} - \mu}{\sigma} \right) \right] - \alpha_i \log \left[ 1 - \Phi_{nor} \left( \frac{\tau_i - \mu}{\sigma} \right) \right] \tag{21}
\]

Based on the logarithm of the lifetime data, the likelihood function in (21) can be used to represent the likelihood function of the lognormal distribution as well. The log-likelihood function of the Weibull distribution is shown in the following equation:

\[
\log L (\lambda, k | DATA) = \sum_{i=1}^{n} \beta_i \left[ - \left( \frac{c_i}{\lambda} \right)^k \right] + (1 - \beta_i) \log \left[ \exp \left( - \left( \frac{r_i}{\lambda} \right)^k \right) - \exp \left( - \left( \frac{r_{i-1}}{\lambda} \right)^k \right) \right] - \alpha_i \left( - \left( \frac{\tau_i}{\lambda} \right)^k \right) \tag{22}
\]

The Akaike information criterion (AIC) is selected as the model selection indicator in this study. Table 10 shows the parameter estimations based on the assumptions of the lognormal, Weibull, and normal distributions by the HG algorithm. The lognormal distribution exhibits the largest AIC among the three distributions, indicating that it is a relatively bad choice. However, the Weibull distribution and the normal distribution yield quite close results in all respects.

Prediction performance is also important in model selection. The Weibull distribution and the normal distribution are both used to predict the cumulative total failure number. Hong et al. [3] proposed the use of calibrated prediction intervals to obtain the RUL of the individual and the cumulative numbers of future failures in a population, as a function of time. This study adopts the same procedures to predict the cumulative failure number of coupler knuckles.

The cumulative number of failures are predicted for every three months of the next two years.

Fig. 7 shows the expected predictions and the 99% prediction bands for the cumulative number of future failures for three distributions. By partially amplifying the total number of failures after one year, it can be found that the number of failures of the normal distribution are significantly lower than that of the Weibull distribution. This study adopts a conservative strategy to predict the number of future failures. Therefore, the Weibull distribution is selected.

Fig. 8 shows expected predictions and the 90% and 95% pointwise prediction bands based on the Weibull distribution for the cumulative number of future failures.

It is known that there are 2852 coupler knuckles still in service within the sample data. As seen from Fig. 8, in the next year, the cumulative number of failures is approximately 664 with a 95% confidence band of [623,707].

Fig. 9 shows the 90% prediction intervals of the RUL for a subset of individual coupler knuckles that are at risk. For example, the current life of Subject 11 is 583 days with a 90% confidence probability that it will expire between 762 and 3350 days. The current life of Subject 22 is 2002 days, and it
will expire between 2050 and 3662 days with a confidence of 90%, which means it is at high risk of failure within a short period of time.

VI. CONCLUSION
In this paper, a unified framework for analyzing left-truncated and fully censored data was proposed. The hierarchical gradient algorithm was introduced and compared with the EM algorithm using the classical data set. To verify the validity of the HG algorithm for the left-truncated and fully censored data, the EM algorithm was improved, and the two algorithms were compared based on the simulation data. The HG algorithm presents fewer iterations and is more stable to the choice of the initial value. Within the framework of the HG algorithm, four quantile methods of nonparametric bootstrap were used to construct the confidence intervals and evaluate the performance in terms of the coverage probability. The coverage probability of the BCa percentile method is closest to the nominal confidence interval. Finally, the proposed algorithm framework was applied to an actual case study and predictions of the RULs of individual coupler knuckles and the cumulative number of failures were given.

It can be seen from the simulation results that the left truncation ratio has an influence on the parameter estimation. Furthermore, the magnitude of the true value of the parameters and the sample size also affect the parameter estimation. It is suggested to further study the effect of the truncation ratio on the parameter estimation. Another topic of interest may be extended from the outcomes of this study to develop a reasonable maintenance strategy. Predictions for the RULs of individual coupler knuckles and the cumulative number of failures in the future provide useful information for maintenance strategies and the inventory management of spare parts. Economic factors and security considerations should be included in the optimization of maintenance strategies.

ACKNOWLEDGMENT
The authors would like to thank the support from Shenhua Railway Freight Transport Co., Ltd.

REFERENCES
[1] A. C. Cohen, Truncated and Censored Samples. New York, NY, USA: Marcel Dekker, 1991.
[2] W. Q. Meeker and L. A. Escobar, Statistical Methods for Reliability Data. Hoboken, NJ, USA: Wiley, 2014.
[3] Y. Hong, W. Q. Meeker, and J. D. McCalley, “Prediction of remaining life of power transformers based on left truncated and right censored lifetime data,” Ann. Appl. Stat., vol. 3, no. 2, pp. 857–879, Jun. 2009.
[4] N. Balakrishnan and D. Mitra, “Likelihood inference for lognormal data with left truncation and right censoring with an illustration,” J. Stat. Planning Inference, vol. 141, no. 11, pp. 3536–3553, Nov. 2011.
[5] N. Balakrishnan and D. Mitra, “Left truncated and right censored Weibull data and likelihood inference with an illustration,” Comput. Stat. Data Anal., vol. 56, no. 12, pp. 4011–4025, Dec. 2012.
[6] N. Balakrishnan and D. Mitra, “EM-based likelihood inference for some lifetime distributions based on left truncated and right censored data and associated model discrimination,” South Afr. Stat. J., vol. 48, no. 2, pp. 125–171, 2014.
[7] T. Emura and S.-K. Shiu, “Estimation and model selection for left-truncated and right-censored lifetime data with application to electric power transformers analysis,” Commun. Statist.-Simul. Comput., vol. 45, no. 9, pp. 3171–3189, Oct. 2016.
[8] D. Kundu and D. Mitra, “Bayesian inference of Weibull distribution based on left truncated and right censored data,” Comput. Stat. Data Anal., vol. 99, pp. 38–50, Jul. 2016.
[9] D. Kundu, D. Mitra, and A. Ganguly, “Analysis of left truncated and right censored competing risks data,” Comput. Stat. Data Anal., vol. 108, pp. 12–26, Apr. 2017.
[10] M. Zhao, H. Jiang, and X. Liu, “A note on estimation of the mean residual life function with left-truncated and right-censored data,” Stat. Probab. Lett., vol. 83, no. 10, pp. 2332–2336, Oct. 2013.
[11] Y. X. Bai, “Confidence interval construction for quantile residual lifetime under left-truncated and right-censored data,” M.S. thesis, Lanzhou Univ. Finance Econ., Lanzhou, China, 2017.
[12] P.-S. Shen, “Estimation in the Cox proportional hazards model with doubly censored and truncated data,” J. Stat. Comput. Simul., vol. 81, no. 11, pp. 1717–1725, Nov. 2011.
[13] P.-S. Shen, “Analysis of left-truncated right-censored or doubly censored data with linear transformation models,” Test, vol. 21, no. 3, pp. 584–603, Sep. 2012.
[14] P. Wang, X. Tong, S. Zhao, and J. Sun, “Efficient estimation for the additive hazards model in the presence of left-truncation and interval censoring,” Stat. Interface, vol. 8, no. 3, pp. 391–402, 2015.
[15] P.-S. Shen and L. N. Wang, “The Cox-Aalen model for left-truncated and mixed interval-censored data,” Statistics, vol. 53, no. 5, pp. 1152–1167, Sep. 2019.
[16] P.-S. Shen, H.-J. Chen, W.-H. Pan, and C.-M. Chen, “Semiparametric regression analysis for left-truncated and interval-censored data without or with a cure fraction,” Comput. Stat. Data Anal., vol. 140, pp. 74–87, Dec. 2019.
[17] D. Mitra and N. Balakrishnan, “Statistical inference based on left truncated and interval censored data from log-location-scale family of distributions,” Commun. Statist.-Simul. Comput., 2019, doi: 10.1080/03610918.2019.1577968.
[18] A. P. Dempster, N. M. Laird, and D. B. Rubin, “Maximum likelihood from incomplete data Via the EM algorithm,” J. Roy. Stat. Soc., B, to be published.
[19] O. G. Alma and R. Arabi Belaghi, “On the estimation of the extreme value and normal distribution parameters based on progressive type-II hybrid-censored data,” J. Stat. Comput. Simul., vol. 86, no. 3, pp. 569–596, Feb. 2016.
[20] R. Arabi Belaghi and M. Noori Asl, “Estimation based on progressively type-I hybrid censored data from the Burr XII distribution,” Stat. Papers, vol. 60, no. 3, pp. 761–803, Jun. 2019.
[21] G. J. McLachlan, T. Krishnan, The EM Algorithm and Extensions. Hoboken, NJ, USA: Wiley, 2008.
[22] U. Gürler and C. D. Yenigün, “Full and conditional likelihood approaches for hazard change-point estimation with truncated and censored data,” Comput. Stat. Data Anal., vol. 55, no. 10, pp. 2856–2870, Oct. 2011.
[23] J. F. Lawless, Statistical Models and Methods for Lifetime Data. Hoboken, NJ, USA: Wiley, 1982.
[24] A. Beck and M. Teboulle, “A fast iterative shrinkage-thresholding algorithm for linear inverse problems,” SIAM J. Imag. Sci., vol. 2, no. 1, pp. 183–202, Jan. 2009.
Q. Li et al.: Failure Analysis for Truncated and Fully Censored Lifetime Data With a Hierarchical Grid Algorithm

[25] C. Augustin, “Méthode générale pour la résolution des systemes d’équations simultanées,” *Comp. Rend. Sci. Paris*, vol. 25, no. 2, pp. 536–538, 1847.

[26] A. Bürmen, J. Puhan, and T. Tuma, “Grid restrained nelder-mead algorithm,” *Comput Optim. Appl.*, vol. 34, no. 3, pp. 359–375, Jul. 2006.

[27] A. C. Davison, D. V. Hinkley, *Bootstrap Methods and Their Application*. Cambridge, U.K.: Cambridge Univ. Press, 1997.

[28] H. Xue, “Study on fatigue reliability of heavy haul freight car couplings based on measured load spectrum,” Ph.D. dissertation, Beijing Jiaotong Univ., Beijing, China, 2017.

QI LI (Member, IEEE) was born in Qingdao, Shandong, China. He received the B.S. degree in industrial engineering from Beijing Jiaotong University, Beijing, China, in 2012, the M.S. degree in industrial engineering and the Ph.D. degree in industrial and manufacturing systems engineering and statistics from the Iowa State University, Iowa, USA, in 2014 and 2017, respectively. From 2012 to 2017, he was a Research Assistant with Iowa State University. Since 2017, he has been an Associate Professor with the Mechanical Engineering Department, Beijing Jiaotong University. His research interests include reliability, risk management, engineering statistics, and operations research.

DONGDONG LI was born in Wuhu, Anhui, China, in 1994. Since 2017, she has been a Postgraduate in industrial engineering with Beijing Jiaotong University, Beijing, China. Her research interests focus on reliability engineering, risk management, and operations research.

BING HUANG was born in Taizhou, Jiangsu, China, in 1996. Since 2018, he has been a Postgraduate in industrial engineering with Beijing Jiaotong University, Beijing, China. His research interests focus on optimal maintenance policy, reliability engineering, and operations research.

ZHENGQIANG JIANG was born in Yixing, Jiangsu, China. He received the B.S. and Ph.D. degrees in mechanical engineering from the Hefei University of Technology, Hefei, China, in 2002 and 2006, respectively. From 2006 to 2008, he was an Assistant Professor with the Hefei University of Technology, and became an Associate Professor, in 2008. Since 2011, he has been an Associate Professor with the Mechanical Engineering Department, Beijing Jiaotong University. His research interests include remaining useful life prediction and maintenance optimization, and intelligent manufacturing systems.

MINGCHENG E. was born in Dalian, Liaoning, China. He received the Ph.D. degree in mechanical engineering from Tianjin University, Tianjin, China, in 1997. From 1997 to 2000, he was an Assistant Professor with Beijing Jiaotong University. Since 2000, he has been an Associate Professor with the Mechanical Engineering Department, Beijing Jiaotong University. His research interests include remaining useful life estimation and optimal maintenance policy, and manufacturing system modeling and optimization.

JING MA was born in Guyuan, Ningxia, China. He received the B.S. and Ph.D. degrees in industrial engineering from the Hefei University of Technology, Hefei, China, in 2010 and 2015, respectively. From 2016 to 2018, he was a Post-doctoral Researcher with Tsinghua University. Since 2018, he has been an Assistant Professor with the Mechanical Engineering Department, Beijing Jiaotong University. His research interests include intelligent manufacturing systems planning, modeling, and optimization.

* * *