Quality and Reliability Engineering International

WILEY

RESEARCH ARTICLE OPEN ACCESS

Simple Heterogeneous Step-Stress Accelerated Life Testing Model Under Type II Censoring

Yao Lu | Maria Kateri

Institute of Statistics, RWTH Aachen University, Aachen, Germany

Correspondence: Maria Kateri (maria.kateri@rwth-aachen.de)

Received: 29 October 2024 | **Revised:** 12 February 2025 | **Accepted:** 13 February 2025

Funding: The work is funded by the German Ministry of Education and Research, within the project BALd in the Cluster BattNutzung, Grant No. 03XP0320A.

Keywords: bootstrap | cumulative exposure model | EM algorithm | exponential lifetime | log-link function | mixture model | model misspecification

ABSTRACT

Common step-stress accelerated life testing (SSALT) models assume that all testing items are sampled from a homogeneous population. However, this is often not the case in practice. Practitioners observe inhomogeneous aging patterns among items of the same production batch. This work proposes a simple SSALT model with exponentially distributed, Type II censored lifetimes that accounts for underlying heterogeneity in aging. To capture the inhomogeneity, a mixture model is introduced and an expectation–maximization (EM) algorithm for censored data is constructed for approximating the maximum likelihood estimates of the model's parameters. The validity of the suggested model and its advantage over the SSALT model in the presence of heterogeneity are demonstrated via simulation studies. Additionally, the log-link function, used to extrapolate the inferential results to normal operating condition (NOC), is adjusted to accommodate the heterogeneous setup. For log-link models, it is demonstrated that in presence of heterogeneity, the common model always overestimates the lifetime under NOC. In contrast, the proposed model, accounting for heterogeneity, reduces the bias in estimation and extrapolation.

1 | Introduction

It is challenging for engineers to evaluate products' lifetime within a constrained test time and expense due to their high reliability. Accelerated life testing (ALT) was developed and widely implemented in the industry as a solution to this problem. Failures are induced far sooner than under normal operating condition (NOC) when higher-than-usual levels of stress are present. Examples of such stresses include temperature, voltage, and pressure. Data obtained from ATL are analyzed and then extrapolated at the NOC. Some key references in the area of ALT are [1–6], where optimal designs of ALTs, statistical models, and applications are discussed. When performing an ALT experiment, different types of load, for example, constant stress, step stress, or linearly increasing stress can be imposed on the testing units. In particular, step-stress ALT (SSALT) setups allow stress levels to change during the experiment, often from low to high, exposing thus the same testing units to more than one stress levels. Statistical references and optimal designs for a homogeneous population with various life distributions under Type I or II censoring in SSALT have been extensively discussed, particularly the simple SSALT model with only two stress levels, see [7–13].

ALT experimental data may exhibit heterogeneity in aging of the testing items. For instance, Figure 1 in [14] presented a circuit-board ALT test with relative humidity as the stress factor, where the failure times of 72 test units at 75.4% relative humidity can be separated into two clusters. This divergence in aging behavior often becomes more apparent during the later stages of an ALT test. For example, a late-stage heterogeneity is evident in the laser lifetime data presented in Example 13.10 (see Figure 13.14) in [6]. Another notable example is battery cells

performance. Cells tend to display nearly identical performance during the initial testing period but diverge significantly after a certain number of cycles (see, e.g., [15]). Furthermore, in ALT experiments, it is often observed that the variability in product performance tends to increase at higher stress levels, as evidenced in Example 21.1 and the corresponding Figure 21.1 in [6], where a more pronounced separation into two groups is observed at higher temperatures. Similarly, Figure 2 in [16] demonstrated this behavior for metal-film resistors.

Differences in aging behaviors and the presence of groups with similar aging patterns are expected when the tested products come from different subsidiary corporations or mass production lines, as variations in manufacturing processes may account for the differences in aging. In this case, aging groups coincide with, for example, production lines. However, it may happen in life testing experiments that items from the same population (e.g., the same batch or production line), subjected to identical testing conditions, age differently at later stages, leading to the formation of distinct groups, as seen in the previously mentioned examples. In such cases, group membership is unknown beforehand, since the sample is assumed to be homogeneous before testing. However, subtle differences, possibly at the material level, may lead to heterogeneity emerging in later stages. Detecting possible aging patterns and clustering the items accordingly, especially when group information is unavailable, is crucial for improving lifetime predictions and better understanding the variability within the tested population.

There exists some literature handling heterogeneity under the ALT setting, considering scenarios in which group information is either available or not. Al-Hussaini and Abdel-Hamid [17] considered finite mixture models under ALT and derived the maximum likelihood estimates (MLEs) of the parameters of a mixture of two Weibull components with unknown group membership, each representing a different failure mechanism, by solving nonlinear equations. Later, they extended this idea to the progressive stress ALT and step partially ALT models [18, 19]. Leon et al. [20] developed a Bayesian method to make credible estimations and predictions from ALT data when the test units originate from known heterogeneous groups. Lv et al. [21] incorporated both group effects and different failure mechanisms using a regression model with random effects and nonconstant shape parameters of the Weibull distribution in ALT, utilizing the available group membership. Lin et al. [22] proposed a flexible two-step reliability assessment model to deal with the underlying heterogeneous problem, using Weibull mixtures with unknown mixture proportions. Seo and Pan [23] described a generalized linear mixed model approach for right-censored ALT data to encompass inhomogeneity caused by the random group effect, where group membership information is available. Zhuang et al. [24] introduced a modeling framework and developed the statistical inference for progressive stress ALT with group effects under progressive censoring, using known group membership.

Nevertheless, modeling heterogeneity under the SSALT setup has not been widely contemplated. For example, Seo and Pan [25] extended their generalized linear mixed model to accommodate an SSALT setup. Wang [26] built a nonlinear mixed-effect model with Weibull-distributed lifetime data to analyze the random group effects in SSALT. However, all SSALT models considered so far refer to mixtures of groups with items of known group membership. Thus, there is a lack of treating heterogeneity in SSALT modeling when the group membership is not known ahead, which is often the case in real applications, as evidenced by the examples above. In this work, instead of considering the heterogeneous group effect introduced by subsampling or random blocks in [25] and [26] where the group information is already available, we develop an SSALT model based on a mixture model that depicts the inhomogeneous aging patterns of the test units without any prior knowledge about their aging group membership.

To the best of our knowledge, the point and interval inference in SSALT without prior information about inhomogeneity has not been addressed so far in the literature. In this paper, we propose such a heterogeneous SSALT (hSSALT) model based on a mixture of lifetime distributions to capture the heterogeneity in mean lifetimes, considering exponential lifetimes under Type II censoring and assuming the cumulative exposure (CE) model. The expectation–maximization (EM) algorithm is adapted to analyze failure time data from the proposed model. Specifically, we consider a simple SSALT experiment under which the test units behave homogeneously at the first stress level but diverge into two subgroups under the second stress level.

The remainder of the paper is organized as follows. In Section 2, the proposed hSSALT model, which incorporates a mixture model, is introduced and described in detail, providing the corresponding distribution function, the density function and the likelihood function under Type II censoring. The point estimation of its parameters is discussed in Section 3. In particular, the associated likelihood equations are provided, along with the proof of the identifiability of the model's parameters. The MLEs of the parameters in the mixture are approximated by an EM algorithm developed for our setup. Furthermore, the relationship between the estimates from the (homogeneous) SSALT and the hSSALT models is commented. Ways to construct asymptotic and bootstrap confidence intervals (CIs) are discussed in Section 4. A simulation study for comparing the coverage probabilities of the considered CIs is carried out in Section 5. Next, two illustrative examples are presented in Section 6, where the results of hSSALT and SSALT are compared in both heterogeneous and homogeneous scenarios. Besides, CIs are also computed for hSSALT model. In Section 7, the link function for hSSALT to extrapolate lifetime under NOC is described. Finally, a summary of the results and some concluding remarks are provided in Section 8.

2 | Model Description

Let us consider the following simple SSALT experiment: n identical units are placed on a life test at stress level x_1, where x_1 is higher than the NOC level x_0. Let τ denote the prespecified time point at which the stress level is increased to x_2, $x_2 > x_1$. Let n_1 and n_2 be the number of failures that occur before and after τ, respectively. The experiment continues until r (usually $r < n$) failures are observed (Type II censoring), with $n_1 + n_2 = r$. When $r = n$, no censoring occurs and the entire sample is observed. This case can be treated as a special case of Type II censoring.

Let $t_{1:n} < t_{2:n} < \cdots < t_{r:n}$ denote the ordered observed failure times. We further suppose that the data come from a CE model; see [2, 5]. Additionally, we assume that the tested units behave quite similarly under x_1, but differentiate into two groups under x_2. An extension to more than two groups is also possible. The assumption of heterogeneity at the second stress level aligns with the discussion in the introduction observed in real ALT experiments, where it is common for heterogeneity to emerge during later stages of testing, particularly at higher stress levels. However, the model can be adjusted to account for heterogeneity occurring at the first stress level as well, depending on the specific experimental conditions and assumptions of the model.

In the sequel, we focus on the case of heterogeneity arising under the second stress level x_2 and being captured by two subgroups of testing items. Thus, the corresponding cumulative distribution function (CDF) for the random lifetimes T_1 and T_2 of restricted two components in a CE setup are given, respectively, by the following:

$$G_j(t) = \begin{cases} F_1(t), & t < \tau \\ F_{2j}(s_j + t - \tau), & t \geqslant \tau \end{cases}, \; j = 1, 2,$$

where F_1, F_{21}, and F_{22} are CDFs of exponential distributions with means θ_1, θ_{21}, and θ_{22}, respectively, that is, $F_1(t) = 1 - \exp\left\{-\frac{t}{\theta_1}\right\}$, for $t \geqslant 0$, with F_{2j}, $j = 1, 2$, being defined analogously. In the CE setup, it is obvious that $F_{21}(s_1) = F_{22}(s_2) = F_1(\tau)$, leading to $s_j = \frac{\theta_{2j}}{\theta_1}\tau$ and hence to $s_j + t - \tau = t - \tau + \frac{\theta_{2j}}{\theta_1}\tau$, for $j = 1, 2$.

The whole test sample is a mixture of two subsamples with proportion p for the first group and $1 - p$ for the second, with CDF $G^*(t) = pG_1(t) + (1 - p)G_2(t)$. Then, the hSSALT model is defined by the following:

$$G^*(t) = \begin{cases} 1 - \exp\left\{-\frac{t}{\theta_1}\right\}, & t < \tau \\ 1 - p\exp\left\{-\frac{t-\tau}{\theta_{21}} - \frac{\tau}{\theta_1}\right\} & \\ \quad -(1-p)\exp\left\{-\frac{t-\tau}{\theta_{22}} - \frac{\tau}{\theta_1}\right\}, & t \geqslant \tau \end{cases}. \tag{1}$$

The corresponding probability density function (PDF) is given by the following:

$$g^*(t) = \begin{cases} g_1^*(t), & t < \tau \\ g_2^*(t), & t \geqslant \tau \end{cases} = \begin{cases} \frac{1}{\theta_1}\exp\left\{-\frac{t}{\theta_1}\right\}, & t < \tau \\ p\frac{1}{\theta_{21}}\exp\left\{-\frac{t-\tau}{\theta_{21}} - \frac{\tau}{\theta_1}\right\} & \\ \quad + (1-p)\frac{1}{\theta_{22}}\exp\left\{-\frac{t-\tau}{\theta_{22}} - \frac{\tau}{\theta_1}\right\}, & t \geqslant \tau \end{cases}. \tag{2}$$

Given observed data $\boldsymbol{t} = (t_{1:n}, t_{2:n}, \ldots, t_{r:n})$, the likelihood is as follows:

$$L(\theta_1, \theta_{21}, \theta_{22}, p \mid \boldsymbol{t}) \propto \left(\prod_{i=1}^{n_1} g_1^*(t_{i:n})\right)\left(\prod_{i=n_1+1}^{r} g_2^*(t_{i:n})\right) \times [1 - G^*(t_{r:n})]^{n-r}, \tag{3}$$

where $\prod_{i=1}^{0}(\cdot) = 1$ and $\prod_{i=r+1}^{r}(\cdot) = 1$. It is evident that, once the data are observed, one of the following three cases will occur: (i) all failures are observed at level x_1, that is, $\mathbb{I}(t_{r:n} \leqslant \tau) = 1$ and $n_1 = r$; (ii) all failures happen under level x_2, that is $\mathbb{I}(t_{1:n} > \tau) = 1$ and $n_1 = 0$; or (iii) failures occur under both stress levels. Of special interest is case (iii) with $1 \leqslant n_1 \leqslant r - 3$, since under this condition, the MLEs of all parameters exist. However, these estimates are conditional MLEs. It has to be pointed out that this condition is stricter than the corresponding one for the common simple SSALT model. Recall that the estimability of the parameters of the simple SSALT model is ensured if at least one failure is observed per stress level, while for an hSSALT model, at least one observation under x_1 and at least three observations under x_2 are required. For this, further analysis refers only to the case with $1 \leqslant n_1 \leqslant r - 3$.

Given that $1 \leqslant n_1 \leqslant r - 3$, Equation (3) with Equation (2) becomes

$$\begin{aligned} L(\theta_1, \theta_{21}, \theta_{22}, p \mid \boldsymbol{t}) \propto & \frac{1}{\theta_1^{n_1}}\exp\left\{-\frac{\sum_{i=1}^{n_1} t_{i:n}}{\theta_1}\right\}\prod_{i=n_1+1}^{r} \\ & \times \left(p\frac{1}{\theta_{21}}\exp\left\{-\frac{t_{i:n}-\tau}{\theta_{21}} - \frac{\tau}{\theta_1}\right\}\right. \\ & \left. + (1-p)\frac{1}{\theta_{22}}\exp\left\{-\frac{t_{i:n}-\tau}{\theta_{22}} - \frac{\tau}{\theta_1}\right\}\right) \\ & \times \left(p\exp\left\{-\frac{t_{r:n}-\tau}{\theta_{21}} - \frac{\tau}{\theta_1}\right\}\right. \\ & \left. + (1-p)\exp\left\{-\frac{t_{r:n}-\tau}{\theta_{22}} - \frac{\tau}{\theta_1}\right\}\right)^{n-r}, \end{aligned}$$

leading to the following expression for the log-likelihood (up to a constant)

$$\begin{aligned} \ell(\theta_1, \theta_{21}, \theta_{22}, p \mid \boldsymbol{t}) = & -n_1\log\theta_1 - \frac{\sum_{i=1}^{n_1} t_{i:n}}{\theta_1} - \frac{(n-n_1)\tau}{\theta_1} \\ & + \sum_{i=n_1+1}^{r}\log\left(p\frac{1}{\theta_{21}}\exp\left\{-\frac{t_{i:n}-\tau}{\theta_{21}}\right\}\right. \\ & \left. + (1-p)\frac{1}{\theta_{22}}\exp\left\{-\frac{t_{i:n}-\tau}{\theta_{22}}\right\}\right) \\ & + (n-r)\log\left(p\exp\left\{-\frac{t_{r:n}-\tau}{\theta_{21}}\right\}\right. \\ & \left. + (1-p)\exp\left\{-\frac{t_{r:n}-\tau}{\theta_{22}}\right\}\right). \end{aligned} \tag{4}$$

Notice that the log-likelihood can be written as a sum of two components

$$\ell(\theta_1, \theta_{21}, \theta_{22}, p \mid \boldsymbol{t}) = \ell^{(1)}(\theta_1 \mid \boldsymbol{t}) + \ell^{(2)}(\theta_{21}, \theta_{22}, p \mid \boldsymbol{t}),$$

with $\ell^{(1)}$ depending only on θ_1 and

$$\begin{aligned}\ell^{(2)}(\theta_{21},\theta_{22},p\mid \boldsymbol{t}) = &\sum_{i=n_1+1}^{r}\log\left(p\frac{1}{\theta_{21}}\exp\left\{-\frac{t_{i:n}-\tau}{\theta_{21}}\right\}\right.\\ &\left.+(1-p)\frac{1}{\theta_{22}}\exp\left\{-\frac{t_{i:n}-\tau}{\theta_{22}}\right\}\right)\\ &+(n-r)\log\left(p\exp\left\{-\frac{t_{r:n}-\tau}{\theta_{21}}\right\}\right.\\ &\left.+(1-p)\exp\left\{-\frac{t_{r:n}-\tau}{\theta_{22}}\right\}\right).\end{aligned} \tag{5}$$

depending only on the parameters under stress level x_2, that is, on p and $\boldsymbol{\theta}=(\theta_{21},\theta_{22})$. Furthermore, notice that the mixture components in the sum terms of Equation (5) correspond to the PDF's of exponential distributions, left truncated at τ, that is, of the two-parameter exponential distributions $\mathrm{Exp}(\theta_{2j},\tau)$, $j=1,2$.

Remark 1. For the uncensored case, obtained when $r=n$, that is, when the complete sample is observed, the log-likelihood is simplified, since the last term in Equation (4) is eliminated. All results presented next for Type II censored samples apply also to complete samples.

3 | Maximum Likelihood Estimation

Given a Type II censored sample observed under a simple hSSALT experiment and setting the partial derivatives of the log-likelihood function (4) with respect to the parameters of model (2) to 0, the corresponding likelihood equations are derived, which are provided in Appendix A.

From Equation (A1), the MLE of θ_1 is given in closed-form expression

$$\hat{\theta}_1=\frac{\sum_{i=1}^{n_1}t_{i:n}+(n-n_1)\tau}{n_1}. \tag{6}$$

Not surprisingly, Equation (6) coincides with the MLE expression for θ_1 of the simple SSALT model and is a biased estimator (see [8] for properties of $\hat{\theta}_1$ and their proofs). However, the MLEs for θ_{21}, θ_{22}, and p cannot be derived in closed-form expressions from (A2) to (A4), which have to be solved numerically. The hSSALT model under x_2 is a mixture of two exponential distributions, left truncated at τ, with unknown mixture proportion p. Hence an EM algorithm can be utilized for solving (A2)–(A4), discussed in the next section along with the existence of the MLEs.

Before proceeding to the estimation of θ_{21}, θ_{22}, and p, their identifiability is discussed next.

3.1 | Identifiability of Parameters

As a precondition for clustering or classification of objects, the identifiability of a finite mixture is important for parameter estimation. The identifiability of mixtures has been discussed and proved for different classes of finite mixing distributions. For example, [27] showed that the class of all mixtures of a one-parameter additively closed family of distributions is identifiable, while [28] proved the identifiability of all finite mixtures of normal distributions and of Gamma distributions by implementing the Laplace transform and the lexicographic order. Chandra [29] developed a general theory for the identifiability of a class of mixing distribution. He adjusted the sufficient condition proposed in [28] by using the moment-generating function (MGF) and simplifying thus the proofs. Barndorff-Nielsen [30] proved the identifiability of mixtures of exponential families. Yakowitz and Spragins [31] proved an identifiability characterization theorem and exhibited distribution families that generate identifiable mixtures, for example, the multivariate Gaussian family.

According to [29], a class of finite mixing distributions is defined as identifiable if and only if $\sum_{i=1}^{m}p_i f(\theta_i)=\sum_{j=1}^{n}p'_j f(\theta'_j)$ implies that $m=n$ and for all i, $(p_i,f(\theta_i))$ is a permutation of $(p'_j,f(\theta'_j))$.

Under the model assumption in Section 2 and focusing on the two-component exponential mixture under stress level x_2, that is, Equation (1) for $t>\tau$, we can prove that the class of all finite mixtures of left-truncated exponential distributions is identifiable using Theorem 2.4 in [29]. In particular, it is sufficient to show that (i) the lexicographic order of our two left-truncated exponential components of the CDF, when considered in terms of their MGFs, implies a subset relationship between the domains of these MGFs, and (ii) there exists a point t_1 such that the limit of the ratio of these two MGFs for $t\to t_1$ equals 0. For this, let X_j be continuous random variables with left-truncated density functions given by the following:

$$f_{X_j}(x)=\begin{cases}0, & x<\tau\\ \frac{f_{2j}(x)}{1-F_1(\tau)}, & x\geqslant\tau\end{cases},\ j=1,2, \tag{7}$$

where F_1 is the CDF of an exponential distribution with mean lifetime θ_1, and f_{2j} is the PDF of an exponential with θ_{2j} (recall that $F_1(\tau)=F_{2j}(s_j)$, $j=1,2$). The associated MGFs are given by the following:

$$\begin{aligned}\phi_j(t)=M_{X_j}(t)&=\int_{\tau}^{\infty}e^{tx}\frac{1}{\theta_{2j}}e^{(x-\tau)/\theta_{2j}}\,dx\\ &=\frac{e^{\tau t}}{1-\theta_{2j}t},\ t<\frac{1}{\theta_{2j}},\ j=1,2,\end{aligned} \tag{8}$$

with domain $D_{\phi_j}=(0,\frac{1}{\theta_{2j}})$, $j=1,2$. The CDFs F_{X_j} are lexicographical ordered by $F_{X_2}(x;\theta_{22})\leqslant F_{X_1}(x;\theta_{21})$ if $\frac{1}{\theta_{22}}<\frac{1}{\theta_{21}}$, that is, if $\theta_{21}<\theta_{22}$. Thus, $F_{X_2}\leqslant F_{X_1}$ implies $D_{\phi_2}\subseteq D_{\phi_1}$. Furthermore, for $t_1=\frac{1}{\theta_{22}}$, it holds $\lim_{t\to t_1}\frac{\phi_1(t)}{\phi_2(t)}=\lim_{t\to t_1}\frac{1-\theta_{22}t}{1-\theta_{21}t}=0$. Therefore, the class of all finite mixtures of the left-truncated exponential distributions is identifiable.

Thus, in the hSSALT model setup, the identifiability of the parameters θ_{21}, θ_{22}, and p requires the constraint $\theta_{21}<\theta_{22}$ to be imposed. Alternatively, the constraint $\theta_{21}>\theta_{22}$ can be imposed, adjusting the proof outlined above accordingly. This constraint ensures the parameters identifiability.

3.2 | EM Algorithm

The EM algorithm, introduced by [32], is one of the most well-known iterative algorithms to approach the MLEs of mixture models with an intricate nonlinear system of likelihood equations, where the observed data can be viewed as incomplete data with latent state information. Each iteration involves estimating a probability distribution for complete data given the current estimates (E-step) and maximizing the likelihood based on the previous step (M-step). EM algorithms have also been widely implemented in survival and reliability analysis, some of them handling also the issue of censoring. For example, [33] presented an extension of the stochastic EM algorithm for a censored mixture, while [34] determined the MLEs when the data are progressively Type II censored. Later, [35] fitted Weibull lifetime data with right-censored and left-truncated data using an adapted EM algorithm. Lee and Scott [36] presented EM algorithms with multivariate Gaussian mixture models where data are truncated, censored, or truncated and censored. Ateya [37] introduced the MLEs of generalized exponential distributions with censored mixtures. Verbelen et al. [38] discussed fitting mixtures of Erlangs to censored and truncated data with EM algorithms.

In our setting, we develop an EM algorithm for deriving the MLEs of the parameters $\theta = (\theta_{21}, \theta_{22})$ and p of the hSSALT model with PDF (2), that is, for solving the likelihood equations (A2)–(A4), which is equivalent to estimating the parameters of a mixture of two left-truncated exponential distributions (7) under Type II censoring, working along the lines of, for example, [36]. A variable that indicates the censoring time needs to be introduced to handle the right-censoring. In our Type II censoring setup, the censoring time is random and hence the censoring process is described by a random variable C with $C = T_{r:n}$. Then, the ordered censored lifetimes are as follows:

$$T^*_{i:n} = \min(T_{i:n}, C),\ D_i = \mathbb{I}(T_{i:n} \leqslant C),\ i = n_1 + 1, \dots, n,$$

where the binary random variable D_i indicates whether the i-th lifetime $T^*_{i:n}$ is observed ($D_i = 1$) or censored ($D_i = 0$). The indicator D_i can also be explained as whether the order of i-th failure is larger than prefixed r. Thus, the observed censored lifetimes consist of two vectors $\boldsymbol{t}^* = (t^*_{n_1+1:n}, \dots, t^*_{n:n})$ and $\boldsymbol{d} = (d_{n_1+1}, \dots, d_n)$, where $t^*_{i:n}$ and d_i are realizations of $T^*_{i:n}$ and D_i, respectively. Further, we consider a binary vector $Z_i = (Z_{i1}, Z_{i2})$ indicating the assigned subgroup for i-th failure with components

$$Z_{ij} = \begin{cases} 1 & t^*_{i:n} \text{ belongs to subgroup } j \\ 0 & \text{otherwise} \end{cases},\ i = n_1 + 1, \dots, n,\ j = 1, 2,$$

and $Z_{i1} + Z_{i2} = 1$, for $i = n_1 + 1, \dots, n$. The component indicator Z is regarded as the missing data. Thus $(\boldsymbol{t}^*, \boldsymbol{d})$ is the observed data (with unknown group membership for failures observed after the time point τ) and $(\boldsymbol{t}^*, \boldsymbol{d}, \boldsymbol{z})$ is the associated complete data.

For the i-th observed censored failure data $(t^*_{i:n}, d_i)$, with corresponding missing (group membership) data $z_i = (z_{i1}, z_{i2})$, and under the assumption that censoring is independent of the group membership, the contribution of the i-th complete data point to the likelihood is

$$L_i(\theta, p \mid t^*_{i:n}, d_i, z_i) \propto \prod_{j=1}^{2} \left\{ p_j f_{X_j}\left(t_{i:n} \mid z_i, \theta_{2j}\right)^{d_i} \overline{F}_{X_j}\left(t_{r:n} \mid z_i, \theta_{2j}\right)^{1-d_i} \right\}^{z_{ij}}, \tag{9}$$

with f_{X_j} being the PDF of the truncated exponential distribution given in Equation (7), F_{X_j} is the CDF corresponding to Equation (7), $\overline{F}_{X_j} = 1 - F_{X_j}$, $p = p_1 = \mathbb{P}(Z_{i1} = 1)$, and $p_2 = \mathbb{P}(Z_{i2} = 1) = 1 - p$. Hence, the p_js provide the group membership information.

Focusing only on the second stress level, the relevant part of the log-likelihood is $\ell^{(2)}$, provided in Equation (5). By applying Equation (9), $\ell^{(2)}$ can equivalently be expressed based on the complete data as follows:

$$\begin{aligned} \ell^{(2)}_c(\theta, p) &= \log\left[\prod_{i=n_1+1}^{n} L_i(\theta, p \mid t^*_{i:n}, d_i, z_i)\right] \\ &= \sum_{i=n_1+1}^{n}\sum_{j=1}^{2} z_{ij} \log\left[p_j f_{X_j}(t_{i:n} \mid z_i, \theta)^{d_i} \overline{F}_{X_j}(t_{r:n} \mid z_i, \theta)^{1-d_i}\right]. \end{aligned} \tag{10}$$

Let $\vartheta = (\theta, p)$ denote the parameter vector, for simplicity, and $\vartheta^{(k)} = (\theta^{(k)}_{21}, \theta^{(k)}_{22}, p^{(k)})$ the estimate vector at the kth iteration of the EM algorithm. The last can be equivalently be expressed as $\vartheta^{(k)} = (\theta^{(k)}_{21}, \theta^{(k)}_{22}, p^{(k)}_1, p^{(k)}_2)$ with $p^{(k)}_1 + p^{(k)}_2 = 1$. Given $\vartheta^{(k)}$, the expectation of the complete log-likelihood function (10) with respect to the missing data, conditioned on the observed data, is defined by the following:

$$\begin{aligned} Q\left(\vartheta \mid \vartheta^{(k)}\right) &= \mathbb{E}\left[\ell^{(2)}_c(\vartheta) \mid \boldsymbol{t}^*, \boldsymbol{d}, \vartheta^{(k)}\right] \\ &= \sum_{i=n_1+1}^{n}\sum_{j=1}^{2} \mathbb{E}\left(Z_{ij} \mid t^*_{i:n}, d_i, \vartheta^{(k)}\right) \\ &\quad \times \log\left[p_j f_{X_j}(t_{i:n} \mid z_i, \theta)^{d_i} \overline{F}_{X_j}(t_{r:n} \mid z_i, \theta)^{1-d_i}\right]. \end{aligned} \tag{11}$$

In the E-step of the EM algorithm, the expectation of the missing variable Z_i given the current estimate $\vartheta^{(k)}$ can be computed by considering only the conditional probability of $Z_{ij} = 1$ (since Z_{ij} is binary), which is equivalent to the posterior probability conditional on the data and $\vartheta^{(k)}$. This posterior probability for the i-th observed censored failure datum with the jth group indicator is expressed by the following:

$$\begin{aligned} h^{(k)}_{ij} &:= \mathbb{P}\left(Z_{ij} = 1 \mid t^*_{i:n}, d_i, \vartheta^{(k)}\right) \\ &= \left(\frac{p^{(k)}_j f_{X_j}\left(t_{i:n} \mid \theta^{(k)}_{2j}\right)}{\sum_{j=1}^{2} p^{(k)}_j f_{X_j}\left(t_{i:n} \mid \theta^{(k)}_{2j}\right)}\right)^{d_i} \left(\frac{p^{(k)}_j \overline{F}_{X_j}\left(t_{r:n} \mid \theta^{(k)}_{2j}\right)}{\sum_{j=1}^{2} p^{(k)}_j \overline{F}_{X_j}\left(t_{r:n} \mid \theta^{(k)}_{2j}\right)}\right)^{1-d_i}. \end{aligned} \tag{12}$$

The first or second term in Equation (12) represents the associated conditional probability for the observed or censored lifetimes, respectively.

After substituting Equation (12) for the conditional expectation of Z_{ij} and applying the PDF and CDF of the left-truncated exponential distribution from Equation (7) in Equation (11), the parameter vector in the Q function to be maximized is derived as follows:

$$\begin{aligned}\boldsymbol{\vartheta}^{(k+1)} = \arg\max_{\boldsymbol{\vartheta}} & \sum_{i=n_1+1}^{n}\sum_{j=1}^{2} h_{ij}^{(k)} \\ & \times \left[\log p_j + d_i \log\left(\frac{1}{\theta_{2j}}\exp\left\{-\frac{t_{i:n}-\tau}{\theta_{2j}}\right\}\right)\right. \\ & \left. + (1-d_i)\log\left(\exp\left\{-\frac{t_{r:n}-\tau}{\theta_{2j}}\right\}\right)\right].\end{aligned} \tag{13}$$

To simplify the formulation, the censored part in Equation (13) is not written individually in the subsequent derivation but is absorbed in the introduced t^* notation. In particular, let $t^*_{i:n} = t_{i:n}\mathbb{I}(n_1+1 \leqslant i \leqslant r) + t_{r:n}\mathbb{I}(r+1 \leqslant i \leqslant n)$. Moreover, the Lagrange multiplier ϕ is introduced due to the constraint $\sum_{j=1}^{2} p_j = 1$. Thus, Equation (13) can also be written as follows:

$$\begin{aligned}\boldsymbol{\vartheta}^{(k+1)} = \arg\min_{\boldsymbol{\vartheta}} & \sum_{i=n_1+1}^{n}\sum_{j=1}^{2} h_{ij}^{(k)}\left[d_i \log\theta_{2j} + \frac{t^*_{i:n}-\tau}{\theta_{2j}} - \log p_j\right] \\ & + \phi\left(\sum_{j=1}^{2} p_j - 1\right),\end{aligned} \tag{14}$$

since the negative sign in the exponent of the exponential distribution alters the direction of the optimization problem in Equation (13), leading to setting the estimate as the point that minimizes Equation (14) in the next iteration. The updated estimate $\boldsymbol{\vartheta}^{(k+1)}$ can be obtained by taking the first-order derivative with respect to the mixture proportion p_j and the mean time to failure θ_{2j} from Equation (14) and setting them to 0, that is, $\frac{\partial Q(\boldsymbol{\vartheta};\boldsymbol{\vartheta}^{(k)})}{\partial p_j} = 0$ and $\frac{\partial Q(\boldsymbol{\vartheta};\boldsymbol{\vartheta}^{(k)})}{\partial \theta_{2j}} = 0$, respectively. This defines the M-step of the EM algorithm. The update rule is in closed-form expressions with

$$\begin{aligned}\theta_{2j}^{(k+1)} &= \frac{\sum_{i=n_1+1}^{n} h_{ij}^{(k)}(t^*_{i:n}-\tau)}{\sum_{i=n_1+1}^{n} h_{ij}^{(k)} d_i}, \\ p_j^{(k+1)} &= \frac{\sum_{i=n_1+1}^{n} h_{ij}^{(k)}}{\sum_{i=n_1+1}^{n}\sum_{j=1}^{2} h_{ij}^{(k)}} = \frac{\sum_{i=n_1+1}^{n} h_{ij}^{(k)}}{n-n_1},\ j=1,2.\end{aligned} \tag{15}$$

The MLEs of parameters are therefore obtained by updating equations in Equation (15) until the EM algorithm converges.

When analyzing a complete sample in which all objects fail during the experiment, the aforementioned procedure reduces to a standard EM algorithm for a mixture, as all d_is are equal to 1. Accordingly, Equation (15) simplifies to the following:

$$\theta_{2j}^{(k+1)} = \frac{\sum_{i=n_1+1}^{n} h_{ij}^{(k)}(t_{i:n}-\tau)}{\sum_{i=n_1+1}^{n} h_{ij}^{(k)}},\ p_j^{(k+1)} = \frac{1}{n-n_1}\sum_{i=n_1+1}^{n} h_{ij}^{(k)},\ j=1,2. \tag{16}$$

3.3 | Convergence of the EM Algorithm

As noted by [39], the convergence of the EM algorithm is achieved via a sequence of likelihood values that are bounded above. Wu [40] proposed regularity conditions for the convergence of an EM sequence to a stationary value. These presumptions are satisfied in our hSSALT setup and are detailed in Appendix B. As a consequence, the sequence $\{L(\boldsymbol{\vartheta}^{(k)})\}$ is bounded above for any $\boldsymbol{\vartheta}^{(0)} \in \boldsymbol{\Omega}$. Further, $Q(\boldsymbol{\vartheta} \mid \boldsymbol{\vartheta}')$ is continuous in both $\boldsymbol{\vartheta}$ and $\boldsymbol{\vartheta}'$. According to Theorem 2 in [40], this continuity ensures that as the EM algorithm progresses, the likelihood values approach a maximum value corresponding to a stationary point.

However, the implementation of the EM algorithm may not guarantee convergence to the global maximum in case the likelihood equation has multiple roots, corresponding to multiple local maxima or even saddle points, which highly depends upon the initial value (see [39] and references therein). One approach, suggested by [39], identifies all roots of the likelihood equation using diverse initial values from the partitioned parameter space. Then the largest observed local maximum $\hat{\boldsymbol{\vartheta}}$ is considered to be the MLE. Nevertheless, this approach is impractical since it is impossible to ensure that all roots have been found. In practice, a common strategy is to run the EM algorithm from various starting values and select the local maximizer that corresponds to the highest local maximum as the MLE of the parameter vector $\boldsymbol{\vartheta}$, as discussed in [41]. Besides, extensive studies have discussed the optimal choice of starting values for the EM algorithm, but we will not elaborate on them here. In our setup, we explored a range of parameter combinations to reach the MLE. The maximum likelihood asymptotic theory, including consistency, asymptotic normality, and asymptotic efficiency, applies within the EM framework. Furthermore, [42] demonstrated the asymptotic properties of MLE based on Type II censored data. Therefore, asymptotic results can be applied to the MLEs obtained from Equation (15) in our hSSALT setup and will be discussed in the subsequent sections.

3.4 | Estimation of θ_2 Under the SSALT and hSSALT Models

We conduct next a comparative analysis between the MLEs derived for the SSALT and the hSSALT models, with a primary focus on the inhomogeneous situation (hSSALT) and the implications when using one or the other model.

In the presence of heterogeneity, the mean lifetime under stress level x_2 for the hSSALT model is formulated as the weighted average of two distinct subgroups, denoted by $\theta_2^* = p\theta_{21} + (1-p)\theta_{22}$. The MLE of θ_2^* is then

$$\hat{\theta}_2^* = \hat{p}\hat{\theta}_{21} + (1-\hat{p})\hat{\theta}_{22}, \tag{17}$$

with $\hat{p}$, $\hat{\theta}_{21}$, and $\hat{\theta}_{22}$ approximated by the EM algorithm described above. Conversely, if one applies the common simple SSALT model ignoring the existence of heterogeneity, the MLE of the single parameter θ_2 under stress level x_2 is given by the following:

$$\hat{\theta}_2 = \frac{\sum_{i=n_1+1}^{r}(t_{i:n}-\tau) + (n-r)(t_{r:n}-\tau)}{n_2}, \tag{18}$$

and is an unbiased estimator of θ_2 in the homogeneous SSALT model (see [8]).

In the case of a complete sample, that is, for $r = n$ and $n_2 = n - n_1$, it is interesting to notice that the estimates of the expected lifetime under stress level x_2 for hSSALT and SSALT, $\hat{\theta}_2^*$ and $\hat{\theta}_2$, respectively, coincide, as demonstrated in the sequel. Assuming that the EM algorithm for the fit of the hSSALT model has converged after $k+1$ iterations, we set $\hat{p}_j = \hat{p}_j^{(k+1)}$ and $\hat{\theta}_{2j} = \hat{\theta}_{2j}^{(k+1)}$, for $j = 1, 2$. From Equations (17) and (18), we then have

$$\begin{aligned}\hat{\theta}_2^* &= \hat{p}_1^{(k+1)}\hat{\theta}_{21}^{(k+1)} + \hat{p}_2^{(k+1)}\hat{\theta}_{22}^{(k+1)} \\ &= \frac{1}{n_2}\sum_{i=n_1+1}^{n} h_{i1}^{(k)}(t_{i:n} - \tau) + \frac{1}{n_2}\sum_{i=n_1+1}^{n} h_{i2}^{(k)}(t_{i:n} - \tau) \\ &= \frac{1}{n_2}\sum_{i=n_1+1}^{n}\sum_{j=1}^{2} h_{ij}^{(k)}(t_{i:n} - \tau) = \frac{1}{n_2}\sum_{i=n_1+1}^{n}(t_{i:n} - \tau) = \hat{\theta}_2.\end{aligned}$$

In contrast, in the case of a censored sample, where $r < n$, it can be shown that $\hat{\theta}_2^* > \hat{\theta}_2$ always holds (see the proof in Appendix C). Therefore, though the adoption of the SSALT model yields a correct estimation of the average lifetime under stress level x_2 when all failures are observed, even in the presence of heterogeneity, it leads to a biased estimation of the expected lifetime when the experiment is Type II censored with $r < n$. In particular, the misspecified model underestimates the mean lifetime at x_2.

The comparison between the SSALT and hSSALT models is illustrated by a simulation study, considering three different sample sizes ($n = 35$, 50, and 1000). The parameter values used for simulating the hSSALT model (2) are $\theta_1 = e^{3.5} = 33.12$, $\theta_{21} = e^{-0.2} = 0.82$, $\theta_{22} = e^{2.0} = 7.39$, $p = 0.4$, and $\tau = 8$, while the same Type II censoring proportion of 14% is assumed for all sample sizes, leading to $r = 30, 86, 860$ for $n = 35, 100, 1000$, respectively. We focus on point estimation under x_2.

The box plots of the parameters' estimates of θ_2^* and θ_2 along with the kernel density estimation plots, based on 1000 replications, are provided in Figure 1. The left, middle, and right panels refer to sample sizes of 35, 50, and 1000, respectively. The vertical gray lines depict the weighted mean of the true parameters θ_{21} and θ_{22}, which is $\theta_2^* = 4.761$. The means of the $\hat{\theta}_2^*$ and $\hat{\theta}_2$ estimates are provided in the kernel densities plots by the green and red dashed lines, respectively.

The box plot reveals the result proved above across various sample sizes that the point estimates derived from the misspecified homogeneous model consistently tend to underestimate the true weighted mean. Even with smaller sample sizes, a significant portion (third quantile) of $\hat{\theta}_2$s falls below the actual value θ_2^*. This discrepancy becomes more pronounced as the sample size increases, indicating a worsening performance of the homogeneous model with larger datasets. For instance, although not appropriate for an SSALT experiment, only for illustrative purposes, in repeated simulations with a sample size of 5000, it is always observed that $\hat{\theta}_2 < \theta_2^*$ (not shown here). This discrepancy is further highlighted by the significant difference between the true value θ_2^* and the mean of the 1000 estimates of $\hat{\theta}_2$ derived from the simple SSALT model, denoted as $\bar{\hat{\theta}}_2$, as depicted in the kernel density plot. The mean values $\bar{\hat{\theta}}_2$ for sample sizes 35, 100, and 1000 are 4.153, 4.196, and 4.188, respectively. The deviation from the true value θ_2^* amounts to approximately 10%, even in a sample size of 1000. In contrast, the weighted estimates $\hat{\theta}_2^*$ derived from the hSSALT model consistently converge around the true value θ_2^*, despite displaying relatively high volatility due to the increased complexity of the mixture model. However, this fluctuation stabilizes as the sample size increases. Specifically, the average values of 1000-time θ_2^* from the simple hSSALT model are equal to $\bar{\hat{\theta}}_2^* = 4.816, 4.824, 4.764$ for sample sizes of 35, 100, and 1000 respectively. These values gradually approach the actual weighted mean of parameters θ_2^*.

4 | Confidence Intervals

In this section, we discuss the CIs for the parameters. The exact distributions and the associated CIs cannot be obtained since the MLEs of the model parameters under x_2 cannot be derived in closed form. As a result, the asymptotic CIs and the bootstrap CIs are presented.

4.1 | Asymptotic Confidence Intervals

For a large sample size n, we can construct the approximate CIs using the asymptotic normality property of MLEs. The pivotal quantities needed for $100(1-\alpha)\%$ CIs of θ_1, θ_{21}, θ_{22}, and p are given by $[\hat{\eta}_l - E(\hat{\eta}_l)]/\sqrt{V_{\hat{\eta}_l}}$, $l = 1, \ldots, 4$, where $\boldsymbol{\eta} = (\eta_1, \eta_2, \eta_3, \eta_4) = (\theta_1, \theta_{21}, \theta_{22}, p)$, and with the variances being estimated through the inverse of the observed Fisher information matrix. The associated observed Fisher information matrix is presented in Appendix D.

Due to the bias of $\hat{\theta}_1$, the two-sided $100(1-\alpha)\%$ CI for θ_1, $\hat{\theta}_1 \pm z_{1-\frac{\alpha}{2}}\sqrt{V_{\hat{\theta}_1}}$, is corrected by replacing $\hat{\theta}_1$ with $\hat{\theta}_1 - (\sum_{j=1}^{r-1}\sum_{k=0}^{j} c_{j,k}\tau_{j,k})$, where $\tau_{j,k} = (\tau/j)(n-j+k)$, $c_{j,k} = \frac{(-1)^k}{\sum_{i=1}^{r-1} q_i}\binom{n}{j}\binom{j}{k}e^{-(\tau/\theta_1)(n-j+k)}$, $q_i = \binom{n}{i}(1-e^{-(\tau/\theta_1)})^i e^{-(\tau/\theta_1)(n-i)}$, see [8]. The two-sided $100(1-\alpha)\%$ CIs for the remaining three parameters are then given by $\hat{\eta}_l \pm z_{1-\frac{\alpha}{2}}\sqrt{V_{\hat{\eta}_l}}$, $l = 2, 3, 4$.

4.2 | Bootstrap Confidence Intervals

In this subsection, we discuss two bootstrap methods for constructing CIs for parameters: the percentile interval and the bias-corrected accelerated (BCa) percentile interval. We begin by describing the procedure outlined in [43], adjusted in our hSSALT context.

i. Given the original sample $\boldsymbol{t} = (t_{1:n}, t_{2:n}, \ldots, t_{r:n})$, obtain the MLE $\hat{\theta}_1$, $\hat{\theta}_{21}$, $\hat{\theta}_{22}$, and $\hat{p}$ from Equations (6) and (15), respectively.

ii. Simulate the first r order statistics from a data sample of size n from a Uniform distribution $U(0,1)$, denoted as $U_{1:n}, \ldots, U_{r:n}$.

iii. Find n_1 such that $U_{n_1:n} \leqslant 1 - \exp\{-\frac{\tau}{\hat{\theta}_1}\} < U_{n_1+1:n}$. Set $T_{i:n} = -\hat{\theta}_1 ln(1 - U_{i:n})$ for $i \leqslant n_1$.

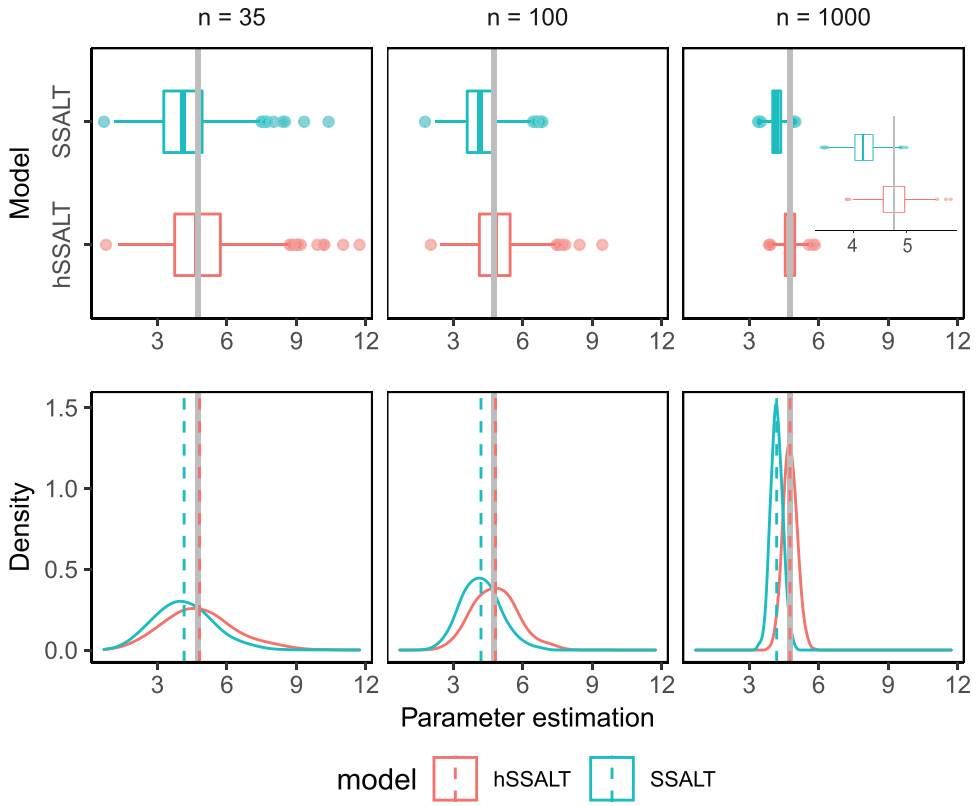

FIGURE 1 | Box plots (top panel) and kernel density estimation plots (bottom panel) of the estimators $\hat{\theta}_2$ and $\hat{\theta}_2^*$, based on 1000 parameter estimations for a simple (homogeneous) SSALT model and a simple hSSALT model with two stress levels ($m = 2$) and stress level change at $\tau = 8$. The columns from left to right show the results for a sample size of 35, 100, and 1000. The gray solid lines represent θ_2^*, while the dashed lines represent the mean values of the estimates from the two models.

iv. Generate n_{21} from a binomial distribution with sample size $n - n_1$ and probability parameter $\hat{p}$. Set $n_{22} = n - n_1 - n_{21}$ as n_{22} and generate data samples of size n_{21} and n_{22} from $U(0,1)$, respectively. Set the sorted $-\hat{\theta}_{2j} \ln(1 - U_{k:n_{2j}}) + \tau$ to $T_{i:n}$, where $j = 1, 2$, $k = 1, \ldots, n_{2j}$, and $i > n_1$. Combine the mixture and select the first $r - n_1$ sorted values as the censored sample under x_2.

v. Estimate the parameters $\boldsymbol{\eta} = (\eta_1, \eta_2, \eta_3, \eta_4) = (\theta_1, \theta_{21}, \theta_{22}, p)$ based on the simulated bootstrap sample, denote by $\hat{\boldsymbol{\eta}}^{(1)}$.

vi. Repeat Steps (ii)–(v) B times to obtain $\hat{\boldsymbol{\eta}}^{(1)}, \ldots, \hat{\boldsymbol{\eta}}^{(B)}$.

vii. Arrange each component of $\boldsymbol{\eta}$ in ascending order and obtain $\hat{\eta}_l^{[1]}, \ldots, \hat{\eta}_l^{[B]}$ $(l = 1, \ldots, 4)$.

A two-sided $100(1-\alpha)\%$ percentile bootstrap CI for a specific parameter is obtained by $(\hat{\eta}_l^{[\alpha B/2]}, \hat{\eta}_l^{[(1-\alpha)B/2]})$. An associated two-sided $100(1-\alpha)\%$ BCa CI of η_l is given by (η_l^L, η_l^U), where $\eta_l^L = \hat{\eta}_l^{[B\alpha_{1l}]}$, $\eta_l^U = \hat{\eta}_l^{[B\alpha_{2l}]}$ with

$$\alpha_{1l} = \Phi\left(\hat{z}_{0l} + \frac{\hat{z}_{0l} + z_{\alpha/2}}{1 - \hat{a}_l\left(\hat{z}_{0l} + z_{\alpha/2}\right)}\right),$$

and

$$\alpha_{2l} = \Phi\left(\hat{z}_{0l} + \frac{\hat{z}_{0l} + z_{1-\alpha/2}}{1 - \hat{a}_l\left(\hat{z}_{0l} + z_{1-\alpha/2}\right)}\right).$$

Here, $\Phi(\cdot)$ denotes the standard normal CDF, and the value of the bias correction $\hat{z}_{0l}$ is computed by

$$\hat{z}_{0l} = \Phi^{-1}\left(\frac{\#(\hat{\eta}_l^{(b)} < \hat{\eta}_l)}{B}\right), \quad l = 1, \ldots, 4, \; b = 1, \ldots, B.$$

Further, choose the good estimate of the acceleration $\hat{a}_l$ with

$$\hat{a}_l = \frac{\sum_{i=1}^{r}\left(\hat{\eta}_{l(\cdot)} - \hat{\eta}_{l(i)}\right)^3}{6\left[\sum_{i=1}^{r}\left(\hat{\eta}_{l(\cdot)} - \hat{\eta}_{l(i)}\right)^2\right]^{\frac{3}{2}}}, \quad l = 1, \ldots, 4,$$

where $\hat{\eta}_{l(i)}$ represents the jackknife estimate for η_l based on the original sample, i.e., the MLE of parameter η_l with the i-th observation deleted, and $\hat{\eta}_{l(\cdot)}$ denotes the mean of the r jackknife estimates for l-th parameter. This method can only be applied when there are at least two observations at each stress level in the initial sample, as explained in detail in Remark 4.2 of [44].

5 | Simulation Study

To evaluate the performance of the interval inference methods described in the preceding section, we conduct a simulation study with varying sample sizes (n = 35, 50, and 100). The same censoring proportion of 14% as in Section 3.4 is used, resulting in r = 30, 43, 86 for the three scenarios, respectively.

Additionally, we consider the complete sample case where $r = n$. Given that the EM algorithm may not always provide valid estimates for both scenarios using the same sample, especially after applying the BCa adjustment, it is important to note that the censored scenarios are sampled independently from the complete scenarios with the same sample size n. We choose $\tau = 8$, $\theta_1 = e^{3.5} = 33.12$, $\theta_{21} = e^{-0.2} = 0.82$, and $\theta_{22} = e^{2.0} = 7.39$.

We conduct 1000 Monte Carlo simulations to obtain estimated coverage probabilities, with true coverage probability levels set at 90%, 95%, and 99%. Furthermore, 1000 valid bootstrap replications are carried out for each simulation. Since the jackknife estimate in the BCa method requires at least two observations under each stress level, to ensure comparability, this requirement is considered for all methodologies.

To focus only on the heterogeneous cases, homogeneous estimates are excluded from the bootstrap replications. An estimate is considered as homogeneous if $\hat{p} \notin [0.01, 0.99]$ or if $\hat{\theta}_{22} - \hat{\theta}_{21} \leqslant 0.05\hat{\theta}_{22}$. In addition, results that fail to converge within 1000 iterations in the EM algorithm are discarded. Consequently, to ensure valid bootstrap results with $B = 1000$, the requisite number of replications is higher than 1000. For instance, when $n = 35$, the median (mean) number of simulations for complete and censored cases are 1014 (1045) and 1032 (1079), respectively. The means are higher due to randomly simulated samples resulting in some extreme values, particularly when the θ_{2j}s are closely located. These adjustments are also applied in the subsequent section where interval estimations are provided.

Tables 1 and 2 present the coverage probabilities and average interval lengths for 90%, 95%, and 99% CIs across various scenarios, respectively. The results in Table 1 indicate that the percentile approach consistently provides the highest coverage probabilities relative to the nominal values, while the approximate method consistently yields the lowest. As expected, the approximate method outperforms in large sample sizes and generally provides the narrowest average interval lengths. However, for $n = 35$, the accuracy of the point estimates associated with the approximate method is unsatisfactory. This can be attributed to the limited number of observations available at x_2 (approximately 25), which are insufficient for estimating three parameters in a mixture model. As expected, as the sample size increases, better coverage probabilities are observed. Furthermore, the coverage probabilities for complete samples tend to be closer to the nominal values compared to censored cases, which is a natural observation, since censoring introduces uncertainty and reduces the amount of available information.

Regarding the two bootstrap methods for constructing CIs, when the sample size is small ($n = 35$), the percentile approach tends to yield conservative coverage probabilities for the parameters under x_2, particularly for p and θ_{22}. However, with a larger sample size, this method does improve its coverage probabilities. On the other hand, the BCa method demonstrates somewhat satisfactory levels of coverage probability relative to the nominal values for all parameters, even with a small sample size. Compared to the percentile method, bias correction is considerably beneficial in maintaining a decent coverage probability. However, for θ_{21}, the average interval lengths are slightly narrower with the percentile method, while the coverage probabilities are also satisfactory. Hence, percentile CIs are recommended for θ_{21}. For θ_{22}, we opt for the BCa approach, as it provides coverage probabilities closer to the nominal values, despite having wider average interval lengths. Kirby and Gerlanc [45] explained that the bounds of CIs in the bootstrap approach can be influenced by a few extreme values when the sample size is small. Moreover, the BCa method can aggravate the discrepancy, as the adjusted cutoffs may lead to more extreme values in the tail of the distribution. Thus, the BCa CIs are substantially broader than the percentile CIs.

Based on this simulation study, for small-to-moderate sample sizes and for θ_{21}, we recommend the parametric percentile method, as it achieves shorter interval lengths while maintaining satisfactory coverage probabilities. For the other three parameters, we recommend the parametric BCa bootstrap method, as it consistently yields coverage probabilities closest to the predetermined nominal values. In the case of large sample sizes, the approximate method can be employed instead due to its computational simplicity and the reasonable closeness of the coverage probabilities to the nominal levels.

6 | Examples

In this section, we present two examples to demonstrate the performance of both the simple SSALT model and our proposed hSSALT model in heterogeneous and homogeneous populations, respectively. Our analysis focuses on the estimates under x_2, as the estimates of θ_1 coincide for both models.

The first example refers to the inhomogeneous situation. The complete sample is generated based on the simulated dataset presented by [10], where $n = 35$ with $\tau = 8$, $\theta_1 = e^{3.5} = 33.12$, $\theta_2 = e^{2.0} = 7.39$. Particularly, θ_2 in the common SSALT corresponds to θ_{22} in our setting. Additionally, we set $p = 0.4$ as the mixture proportion for the component with a smaller mean, that is, θ_{21}. To create a heterogeneous sample, we randomly select 60% of original data from x_2. Subsequently, the remaining 40% of the original data set in x_2 is substituted with the exponential distribution with $\theta_{21} = e^{-0.2} = 0.82$, representing another component. The combined sample is presented in Table 3.

A grid of starting values comprising 540 combinations is utilized to verify the convergence and stability of the EM algorithm. The mixture proportion p varies from 0.1 to 0.9 with a step size of 0.1, θ_{21} ranges from 0.5 to 3 with a step of 0.5, and θ_{22} spans from 6 to 15 with a step of 1. Remarkably, all combinations within this dataset yield the same outcome, rounded to three decimal places. The MLE of θ_1, obtained from Equation (6), is found to be $\hat{\theta}_1 = 31.449$, while the MLEs of θ_{21}, θ_{22}, and p, derived from Equation (16), are $\hat{\theta}_{21} = 0.798$, $\hat{\theta}_{22} = 6.612$, and $\hat{p} = 0.379$, respectively. These estimates closely approximate the true parameters.

The MLE $\hat{\theta}_2 = 4.410$, based on Equation (18), is obtained by (wrongly) modeling this dataset by the simple SSALT for a homogeneous population. The estimated weighted mean lifetime of the hSSALT model, $\hat{\theta}_2^* = \hat{p}\hat{\theta}_{21} + (1 - \hat{p})\hat{\theta}_{22}$, also equals 4.410. Although $\hat{\theta}_2 = \hat{\theta}_2^*$, Figure 2 shows that the hSSALT model outperforms the corresponding homogeneous model in fitting the empirical data. The presented point-wise CIs are constructed based on the quantiles of the respective CDF.

TABLE 1 | Estimated coverage probabilities (in percentage) based on 1000 simulations with $\theta_1 = 33.12$, $p = 0.4$, $\theta_{21} = 0.82$, $\theta_{22} = 7.39$, $\tau = 8$, and $B = 1000$.

		90% CI			95% CI			99% CI		
n	***r***	**Approx**	**Percentile**	**BCa**	**Approx**	**Percentile**	**BCa**	**Approx**	**Percentile**	**BCa**
						CI of θ_1				
35	35	84.4	86.7	88.5	88.0	94.0	94.6	93.9	97.6	98.2
	30	86.5	86.8	90.0	89.6	95.2	95.7	94.3	98.7	99.0
50	50	88.8	88.8	90.7	91.5	94.4	95.1	95.6	98.5	99.1
	43	88.1	88.1	89.3	90.9	93.9	94.2	95.1	98.7	98.7
100	100	89.9	90.3	90.3	93.9	94.6	95.8	98.3	98.6	99.5
	86	89.9	89.6	91.1	93.5	95.6	96.4	97.8	98.7	99.1
						CI of p				
35	35	86.0	95.0	86.1	91.2	98.9	91.0	95.3	100.0	98.0
	30	90.1	98.3	82.0	93.5	99.6	88.4	97.0	100.0	96.2
50	50	89.4	93.5	88.3	92.9	96.3	92.6	96.9	99.8	98.3
	43	90.5	98.3	86.9	94.0	99.5	93.1	97.3	99.9	97.9
100	100	91.3	89.8	89.4	95.3	93.8	92.8	98.2	98.1	98.1
	86	91.3	97.3	90.7	93.9	99.3	95.5	98.1	99.9	99.2
						CI of θ_{21}				
35	35	82.3	91.2	86.8	86.0	95.5	93.1	89.5	98.8	98.7
	30	81.1	89.9	84.9	84.1	94.3	92.8	88.1	98.7	98.4
50	50	84.5	91.3	88.5	87.7	95.2	93.6	91.9	98.8	98.1
	43	84.0	91.1	87.7	86.8	95.2	93.6	91.6	98.8	98.1
100	100	91.2	89.8	89.4	94.0	95.8	94.9	97.1	99.1	98.7
	86	86.4	93.1	90.2	90.1	96.7	95.6	95.8	99.4	98.9
						CI of θ_{22}				
35	35	88.2	95.2	88.3	91.8	98.1	94.9	95.2	99.9	99.1
	30	82.6	93.9	85.1	86.7	97.4	92.1	90.9	99.3	97.1
50	50	90.2	96.0	91.6	94.3	98.5	96.7	97.4	99.8	99.2
	43	87.1	93.8	89.6	89.7	98.0	93.8	93.3	99.9	98.8
100	100	91.0	92.1	90.7	94.4	97.6	96.4	97.9	99.9	99.4
	86	88.4	94.1	92.0	91.5	97.6	96.7	94.7	99.7	99.5

Comparatively, let us consider a Type II censored dataset based on Table 3 where $r = 30$ is supposed as the prefixed number of failures, as shown in Table 4.

In this case, we obtain the same MLE of θ_1 to be $\hat{\theta}_1 = 31.449$. The point estimates from Equation (15) are $\hat{\theta}_{21} = 0.791$, $\hat{\theta}_{22} = 6.447$, and $\hat{p} = 0.372$. Clearly, the complete sample produces more accurate estimates with less bias due to a higher number of observed failures, particularly for the subgroup with a larger mean lifetime parameter. Furthermore, for the censored case, we find $\hat{\theta}_2 = 3.864$ and $\hat{\theta}_2^* = 4.342$, illustrating the established property $\hat{\theta}_2 < \hat{\theta}_2^*$ as discussed in Section 3.4.

In the second example, we examine a homogeneous Type II censored dataset previously presented by [7] in Table 5, where $r = 16, n = 20$, with $\tau = 5, \theta_1 = e^{2.5} = 12.18$, and $\theta_{22} = e^{1.5} = 4.48$. From the SSALT model, we calculate $\hat{\theta}_1 = 23.518$ and $\hat{\theta}_2 = 5.056$. The EM algorithm, introduced in Section 3, is applied using the same 540 sets of starting values as in the first example. All runs converge to the same log-likelihood, rounded to five decimal places, with the corresponding estimates for θ_{21} varying from 5.01 to 5.05, and for θ_{22} being equal to 5.06 or 5.07. The counts of each estimated combination for the parameters θ_{21} and θ_{22} under stress level x_2 are detailed in Table 6, evidencing that the two subgroups' mean lifetime estimates are very close, fulfilling the homogeneity criteria defined in Section 5. This suggests that the sample is relatively homogeneous and does not support the presence of two distinct groups.

It can be concluded from the above two examples that the hSSALT model with the EM algorithm is capable of both identifying heterogeneity and remaining valid in homogeneous scenarios.

TABLE 2 | Average lengths of CIs based on 1000 simulations with $\theta_1 = 33.12$, $p = 0.4$, $\theta_{21} = 0.82$, $\theta_{22} = 7.39$, $\tau = 8$, and $B = 1000$.

		90% CI			95% CI			99% CI		
n	***r***	**Approx**	**Percentile**	**BCa**	**Approx**	**Percentile**	**BCa**	**Approx**	**Percentile**	**BCa**
						CI of θ_1				
35	35	50.40	72.94	61.12	58.11	93.91	77.90	70.68	143.97	116.19
	30	49.39	74.11	59.90	56.89	96.38	77.95	69.12	147.68	117.80
50	50	40.05	51.63	45.28	47.29	68.01	57.00	60.18	111.49	87.41
	43	40.22	51.45	45.53	47.50	68.40	58.10	60.39	112.76	87.30
100	100	25.58	28.12	26.95	30.48	34.84	32.97	40.06	50.22	46.12
	86	25.65	28.34	26.97	30.57	34.96	32.96	40.18	50.26	46.70
						CI of p				
35	35	0.55	0.54	0.51	0.64	0.65	0.60	0.79	0.80	0.77
	30	0.59	0.51	0.45	0.68	0.59	0.53	0.84	0.71	0.66
50	50	0.48	0.46	0.45	0.56	0.57	0.55	0.71	0.76	0.74
	43	0.51	0.46	0.43	0.60	0.54	0.51	0.75	0.67	0.64
100	100	0.37	0.33	0.32	0.44	0.40	0.40	0.57	0.60	0.58
	86	0.37	0.36	0.35	0.47	0.43	0.42	0.61	0.56	0.55
						CI of θ_{21}				
35	35	1.47	1.72	1.86	1.69	2.21	2.37	2.05	3.23	3.36
	30	1.53	1.51	1.78	1.74	1.86	2.17	2.10	2.58	2.83
50	50	1.29	1.47	1.54	1.50	1.92	2.02	1.85	3.01	3.10
	43	1.30	1.32	1.48	1.51	1.63	1.81	1.84	2.28	2.45
100	100	1.06	1.03	1.05	1.25	1.35	1.38	1.59	2.28	2.31
	86	1.05	1.04	1.10	1.24	1.29	1.37	1.57	1.83	1.90
						CI of θ_{22}				
35	35	9.10	9.09	10.52	10.73	11.62	13.87	13.74	18.31	21.02
	30	13.72	11.68	12.94	15.88	14.87	15.85	19.67	22.55	21.76
50	50	7.52	7.52	8.51	8.90	9.64	11.48	11.53	15.85	18.99
	43	11.11	10.66	11.59	12.97	13.61	14.36	16.27	20.99	20.47
100	100	5.15	4.94	5.24	6.13	6.16	6.77	8.02	10.13	12.38
	86	7.78	8.10	8.43	9.18	10.35	10.59	11.82	16.18	15.68

TABLE 3 | Simulated sample based on [10].

Stress level	**Mean failure time**	**Times to failure**							
x_1	$\theta_1 = e^{3.5}$	1.46	2.22	3.92	4.24	5.47	5.60	6.12	6.56
x_2	$\theta_{22} = e^{2.0}$	8.30	8.98	9.43	9.87	11.14	11.85	12.14	13.49
		14.19	14.33	15.28	16.58	17.80	21.09	26.34	28.66
	$\theta_{21} = e^{-0.2}$	8.01	8.10	8.22	8.59	8.77	8.80	8.80	8.84
		8.90	8.97	9.62					

Further, we assess the performance of the asymptotic, the bootstrap percentile, and the BCa CIs using the data illustrated in Tables 3 and 4. The associated point estimates are $\hat{\theta}_1 = 31.449$, $\hat{\theta}_{21} = 0.798$, $\hat{\theta}_{22} = 6.612$, and $\hat{p} = 0.379$ from Table 3 and $\hat{\theta}_1 = 31.449$, $\hat{\theta}_{21} = 0.791$, $\hat{\theta}_{22} = 6.447$, and $\hat{p} = 0.372$ from Table 4.

The point estimates obtained from these tables are used along with the asymptotic variances to construct the approximate CIs, while the bootstrap CIs are simulated based on the condition where $1 \leqslant n_1 \leqslant r - 3$. The results in Table 7 indicate that the approximate method, which is more suitable for larger sample

TABLE 4 | Type II censored sample from the simulated data in Table 3.

Stress level	Mean failure time	Times to failure							
x_1	$\theta_1 = e^{3.5}$	1.46	2.22	3.92	4.24	5.47	5.60	6.12	6.56
x_2	$\theta_{22} = e^{2.0}$	8.30	8.98	9.43	9.87	11.14	11.85	12.14	13.49
		14.19	14.33	15.28					
	$\theta_{21} = e^{-0.2}$	8.01	8.10	8.22	8.59	8.77	8.80	8.80	8.84
		8.90	8.97	9.62					

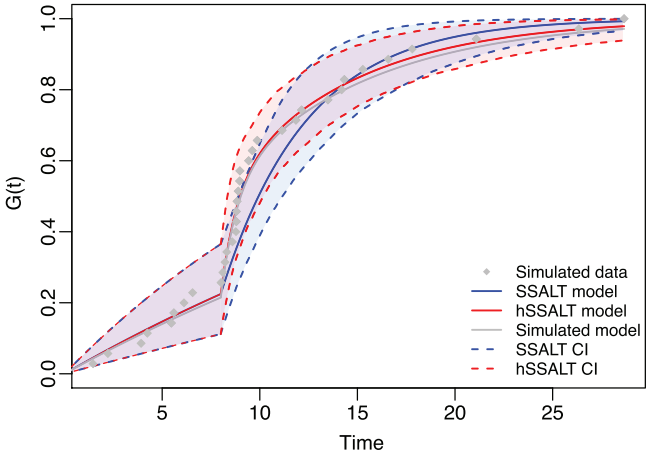

FIGURE 2 | Estimated CDFs (G) for the homogeneous SSALT model (blue) and hSSALT model (red) with two stress levels ($m = 2$) and stress level changes at $\tau = 8$, based on the data of Table 3. The corresponding dashed lines represent the calculated lower and upper bounds of CDF for pointwise 95% CIs. The gray line is the CDF from which the data are simulated. The gray points correspond to the empirical CDF of the simulated data.

TABLE 5 | Simulated sample of [7].

Stress level	Mean failure time	Times-to-failure					
x_1	$\theta_1 = e^{2.5}$	2.01	3.60	4.12	4.34		
x_2	$\theta_2 = e^{1.5}$	5.04	5.94	6.68	7.09	7.17	7.49
		7.60	8.23	8.24	8.25	8.69	12.05

TABLE 6 | Summary of the MLEs of θ_{21} and θ_{22} for the data of Table 5, computed by the EM algorithm based on 540 initial values.

Index	$\hat{\theta}_{21}$	$\hat{\theta}_{22}$	Counts
1	5.01	5.06	2
2	5.02	5.06	3
3	5.03	5.06	9
4	5.04	5.06	88
5	5.05	5.06	391
6	5.05	5.07	47

TABLE 7 | Interval estimation for the simulated sample in Table 3 and its censored variant in Table 3 with $\theta_1 = 33.12$, $p = 0.4$, $\theta_{21} = 0.82$, $\theta_{22} = 7.39$, and $B = 1000$.

r		90%	95%	99%
			CI for θ_1	
n	Approx CI	(8.41, 44.98)	(4.90, 48.49)	(0.00, 55.34)
	Percentile CI	(18.86, 66.22)	(17.48, 68.28)	(14.90, 92.63)
	BCa CI	(18.77, 64.65)	(17.37, 66.83)	(12.94, 90.53)
30	Approx CI	(8.40, 44.98)	(4.90, 48.49)	(0.00, 55.33)
	Percentile CI	(18.71, 65.42)	(16.97, 68.46)	(14.71, 136.36)
	BCa CI	(18.18, 54.63)	(16.28, 66.84)	(13.76, 91.83)
			CI for p	
n	Approx CI	(0.11, 0.65)	(0.06, 0.70)	(0.00, 0.80)
	Percentile CI	(0.14, 0.69)	(0.09, 0.80)	(0.06, 0.91)
	BCa CI	(0.14, 0.73)	(0.11, 0.82)	(0.06, 0.91)
30	Approx CI	(0.07, 0.67)	(0.02, 0.73)	(0.00, 0.84)
	Percentile CI	(0.13, 0.63)	(0.09, 0.66)	(0.05, 0.73)
	BCa CI	(0.19, 0.68)	(0.15, 0.72)	(0.08, 0.78)
			CI for θ_{21}	
n	Approx CI	(0.16, 1.43)	(0.04, 1.55)	(0.00, 1.79)
	Percentile CI	(0.19, 1.84)	(0.10, 2.26)	(0.01, 3.65)
	BCa CI	(0.24, 2.02)	(0.14, 2.53)	(0.01, 3.78)
30	Approx CI	(0.14, 1.44)	(0.02, 1.56)	(0.00, 1.80)
	Percentile CI	(0.19, 1.64)	(0.11, 1.89)	(0.01, 2.29)
	BCa CI	(0.32, 2.01)	(0.24, 2.26)	(0.06, 2.49)
			CI for θ_{22}	
n	Approx CI	(3.34, 9.89)	(2.71, 10.51)	(1.49, 11.74)
	Percentile CI	(4.10, 11.52)	(3.72, 14.29)	(2.81, 18.71)
	BCa CI	(4.48, 14.30)	(4.11, 17.23)	(3.61, 24.76)
30	Approx CI	(2.16, 10.74)	(1.34, 11.56)	(0.00, 13.16)
	Percentile CI	(3.15, 12.87)	(2.77, 14.56)	(2.25, 21.64)
	BCa CI	(3.62, 14.47)	(3.11, 16.96)	(2.38, 22.47)

sizes, generally produces the narrowest CIs. This method tends to yield the lowest coverage probabilities, as observed in Table 1. Inversely, with such a small sample size $n = 35$ as well, the percentile CIs are wider, which aligns with their conservative coverage probabilities in Table 1. The BCa CIs are wider than the percentile CIs in several setups with small sample sizes, as they

TABLE 8 | Interval estimation for large sample size based on simulated data with $\theta_1 = 33.12$, $p = 0.4$, $\theta_{21} = 0.82$, $\theta_{22} = 7.39$, and $B = 1000$.

			90%	95%	99%
n	***r***	**CI**		**CI for θ_1**	
50	43	Approx CI	(12.84, 40.68)	(10.17, 43.34)	(4.96, 48.56)
		Percentile CI	(19.82, 51.92)	(18.56, 53.81)	(16.35, 75.61)
		BCa CI	(18.75, 46.45)	(17.50, 52.47)	(13.45, 63.48)
100	86	Approx CI	(20.92, 47.32)	(18.39, 49.85)	(13.45, 54.79)
		Percentile CI	(25.68, 52.69)	(24.36, 57.09)	(23.01, 63.52)
		BCa CI	(25.05, 50.07)	(23.91, 53.40)	(19.99, 62.66)
				CI for p	
50	43	Approx CI	(0.16, 0.68)	(0.11, 0.73)	(0.01, 0.83)
		Percentile CI	(0.21, 0.64)	(0.17, 0.68)	(0.11, 0.73)
		BCa CI	(0.21, 0.64)	(0.17, 0.68)	(0.09, 0.72)
100	86	Approx CI	(0.16, 0.56)	(0.12, 0.60)	(0.04, 0.67)
		Percentile CI	(0.21, 0.53)	(0.17, 0.56)	(0.14, 0.64)
		BCa CI	(0.20, 0.51)	(0.16, 0.54)	(0.12, 0.61)
				CI for θ_{21}	
50	43	Approx CI	(0.10, 1.20)	(0.00, 1.31)	(0.00, 1.51)
		Percentile CI	(0.24, 1.19)	(0.18, 1.33)	(0.08, 1.80)
		BCa CI	(0.28, 1.27)	(0.22, 1.51)	(0.11, 1.85)
100	86	Approx CI	(0.13, 1.25)	(0.03, 1.36)	(0.00, 1.57)
		Percentile CI	(0.37, 1.20)	(0.31, 1.35)	(0.21, 1.67)
		BCa CI	(0.39, 1.31)	(0.33, 1.44)	(0.22, 1.92)
				CI for θ_{22}	
50	43	Approx CI	(2.54, 10.88)	(1.74, 11.68)	(0.18, 13.24)
		Percentile CI	(3.62, 12.61)	(3.04, 14.76)	(2.39, 20.95)
		BCa CI	(3.72, 12.87)	(3.05, 15.18)	(2.18, 20.69)
100	86	Approx CI	(4.24, 9.97)	(3.69, 10.51)	(2.62, 11.59)
		Percentile CI	(5.00, 10.72)	(4.77, 11.33)	(4.18, 13.77)
		BCa CI	(5.02, 10.78)	(4.77, 11.33)	(3.93, 13.62)

highly depend on the sampled data and the randomly generated bootstrap samples. Additionally, [45] emphasized the impact of small sample size on BCa CIs. Furthermore, it is worth noticing that the bootstrap CIs for the complete samples are sometimes wider than those for the censored samples. This is likely due to the robustness of censoring, which effectively reduces the impact of extreme simulated values. This phenomenon persists even when the sample size is increased, see Table 8.

It is worth mentioning that the widths of the percentile bootstrap CIs for θ_1 always exceed those of the BCa intervals, a trend not observed for the other three parameters. Additionally, the BCa CIs for θ_1 are left-shifted in comparison to the percentile CIs. This can be attributed to the effectiveness of the BCa method, especially in handling known biases in the estimator $\hat{\theta}_1$. This suggests that the BCa approach might be particularly suitable for situations where bias correction is essential.

The estimated CIs, particularly for p, are unsatisfactory due to the limitation of a relatively small sample size. To evaluate the sample size effect, two further simulations are conducted with sample size $n = 50$ and 100, and with r set to 43, 86 in accordance with the proportionate relationship 14% in the original sample. The associated point estimates are as follows:

$$n = 50: \ \hat{\theta}_1 = 29.312, \ \hat{\theta}_{21} = 0.652, \ \hat{\theta}_{22} = 6.712, \ \hat{p} = 0.422,$$

$$n = 100: \ \hat{\theta}_1 = 35.893, \ \hat{\theta}_{21} = 0.693, \ \hat{\theta}_{22} = 7.103, \ \hat{p} = 0.357.$$

We observe a significant reduction in the width of CIs in Table 8 compared to Table 7. As expected, the corresponding CIs get narrower as the sample size increases. Note that, for large sample sizes, asymptotic CIs for some parameters are slightly wider than the bootstrap intervals, while asymptotic intervals are narrower when the sample size is moderate to small. This phenomenon aligns with observations in Table IX of [9]. Additionally, the three different types of CIs become closer to each other as the sample size grows.

7 | Link Function

To extrapolate the lifetime under NOC x_0, a link function is introduced to quantify the relationship between stress levels and lifetimes. The log-link function described below is a common choice in literature

$$\log \theta_s = \alpha + \beta x_s, \quad s = 0, \ldots, m, \tag{19}$$

see [3]. This link model is adapted to incorporate the heterogeneous assumption.

To model the life–stress relationship group-wise at x_2, the following two equations are considered, respectively:

$$A: \ \log \theta_{s1} = \alpha_A + \beta_A x_s, \ s = 1, 2,$$

$$B: \ \log \theta_{s2} = \alpha_B + \beta_B x_s, \ s = 1, 2,$$

with $\theta_{11} = \theta_{12} = \theta_1$ and thus $\alpha_B = \alpha_A + (\beta_A - \beta_B)x_1$, where the subscripts A and B denote the two subgroups with lifetime θ_{21} and θ_{22}, respectively. Analogous to the simple SSALT, the parameter vector $(\alpha_A, \beta_A, \beta_B)$ in the case of a simple hSSALT is also just a reparameterization of $(\theta_1, \theta_{21}, \theta_{22})$, where

$$\alpha_A = \frac{x_2 \log \theta_1 - x_1 \log \theta_{21}}{x_2 - x_1}, \ \beta_A = \frac{\log \theta_{21} - \log \theta_1}{x_2 - x_1},$$

$$\beta_B = \frac{\log \theta_{22} - \log \theta_1}{x_2 - x_1}.$$

A segmented linear function is considered under the assumption in Section 2 that units perform commonly before τ and diverge subsequently. The extrapolation to the expected lifetime under NOC for a randomly selected item from the heterogeneous population with unknown group membership is facilitated through the additional parameter p, representing the mixing fraction for θ_{21} (subgroup A) and the corresponding weighted average θ_2^*. For a simple hSSALT model, the log-link equation has the same form as Equation (19) with $m = 2$, but $\alpha =$

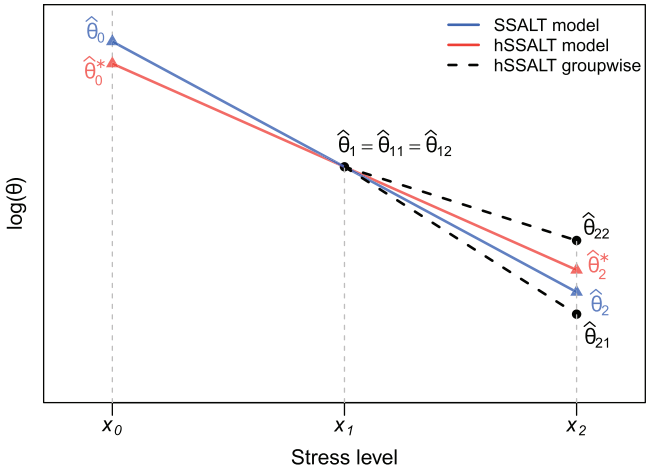

FIGURE 3 | Extrapolation of log-link function for the SSALT model (blue) and the hSSALT model (red) with two stress levels ($m = 2$). The dashed black lines represent the group-wise relation under stress level x_2 due to the existence of heterogeneity. The labeled $\hat{\theta}_i$, $i = 0, 1, 2$, $\hat{\theta}_{ij}$, $i, j = 1, 2$ and $\hat{\theta}_i^*$, $i = 0, 2$ points in the plot are in logarithmic scale (log is omitted for brevity).

$p\alpha_A + (1-p)\alpha_B$ and $\beta = p\beta_A + (1-p)\beta_B$. If $\hat{\alpha}_A$, $\hat{\beta}_A$, $\hat{\beta}_B$, and $\hat{p}$ are estimated, the parameter of interest θ_0 is then derived by $\hat{\theta}_0^* = \exp\left\{\hat{\alpha}_A + (1-\hat{p})(\hat{\beta}_A - \hat{\beta}_B)x_1 + \left(\hat{p}\hat{\beta}_A + (1-\hat{p})\hat{\beta}_B\right)x_0\right\}$. Once $\hat{\theta}_0^*$ is obtained, the lifetime distribution under the NOC can be estimated, along with the associated failure rate, survival function, and other reliability metrics to assess the behavior of the test items. The CI for θ_0^* can be then obtained, for example, through Monte Carlo simulations.

The comparison between the log-link relationships with and without heterogeneity, as estimated based on a Type II censored sample, is shown in Figure 3. The blue line illustrates the log-linear relationship between stress levels and the associated expected lifetimes, as estimated by fitting the simple SSALT model on the dataset, with $\hat{\theta}_2$ representing the estimated lifetime parameter under x_2 and $\hat{\theta}_0$ the corresponding extrapolation under x_0. The two dashed black lines correspond to the log-linear relationships connecting the expected lifetimes between x_1 and x_2 for each group formed under x_2. The red line segment between x_1 and x_2 represents the log-linear relationship between θ_1 (under x_1) and the expected lifetime θ_2^* for the whole population under x_2, which is proposed for the extrapolation to NOC in case of the hSSALT model.

Figure 3 demonstrated that for a Type II censored sample, it holds $\hat{\theta}_2^* > \hat{\theta}_2$, as shown in Section 3.4. Thus, if heterogeneity under x_2 is present, the SSALT model tends to overestimate the lifetime at x_0 (indicated by the blue line), leading to prediction bias. In contrast, the hSSALT model adjusts for heterogeneity, providing a more reliable extrapolation.

8 | Discussion

In this paper, we have proposed a simple hSSALT model with exponentially distributed lifetimes, focusing on scenarios where heterogeneity exists at second stress level x_2 under Type II censoring. Our approach utilizes a mixture model under x_2 to capture heterogeneity without prior knowledge of group membership. To obtain the estimates in such a mixture model, the EM algorithm has been adapted in a censored scenario to fit our setup. We have investigated the relationship between the estimates under x_2 in both heterogeneous and homogeneous settings. Simulation studies have been carried out to compare different approaches for interval inference in terms of their coverage probabilities. Additionally, we have provided two examples and demonstrated the validity and performance of our introduced model using three types of CIs. Furthermore, the extrapolation of the lifetime at the user level x_0 for the hSSALT model has been considered and elaborated using a segmented log-link function.

From our study, we recommend using our hSSALT model when dealing with scenarios involving heterogeneity during the testing procedure. The homogeneous SSALT model may lead to model misspecification as it cannot effectively identify and model heterogeneity. Furthermore, in case of heterogeneity, the common SSALT assuming a homogeneous population throughout the test, tends to overestimate the lifetime under NOC. In contrast, our adapted hSSALT model provides more accurate lifetime predictions by accounting for heterogeneity. When it comes to CIs, the BCa method outperforms the approximate and percentile approaches for all parameters under both stress levels except θ_{21}, particularly in small-to-moderate sample sizes. Therefore, we recommend using the BCa method for interval reference in such scenarios, with the option of switching to the percentile bootstrap method for θ_{21}. For large sample sizes, the approximate method is suggested due to its computational simplicity and relatively close results compared to nominal values.

This is a first work on SSALT models with heterogeneity, captured by a mixture for the heterogeneous groups, while considering unknown group membership. There are further topics that deserve to be investigated. These can be related to the estimation procedure, considering alternative estimation approaches, for example, Bayesian. Optimal design for SSALT experiments under heterogeneity is also an interesting topic. One potential direction for future research could involve exploring the hSSALT model with different lifetime distributions and censoring schemes. Investigation of the model's performance under various scenarios could provide valuable insights into its applicability and robustness. With respect to robustness, alternative estimators to MLE can be exploited, as for example, minimum density power divergence estimators [46] or minimum ϕ-divergence estimators [47]. Another interesting topic we are working on at the moment is the development of a test procedure to assess the necessity of employing an hSSALT model. This work offers significant practical value, as highlighted in Section 7. Overlooking the presence of heterogeneity under the second stress level and applying the simple SSALT model can result in overestimated lifetime predictions under NOC. This, in turn, may lead to excessively long warranty periods, which, while appealing to consumers, could result in financial losses for manufacturers due to higher warranty claims. On the other hand, implementing the hSSALT model when heterogeneity exists provides a more realistic and conservative lifetime estimation under NOC. This allows manufacturers to set warranty periods appropriately, thereby reducing risks of overcommitment. Furthermore, the use

of the hSSALT model enhances manufacturers' understanding of product reliability in the presence of heterogeneity, enabling them to optimize maintenance schedules, reduce unexpected failures, and strengthen brand reputation by well-calibrated expectations.

Finally, the heterogeneity concept can be extended to multiple SSALT with $m > 2$ stress levels. A realistic model assumes the population to be homogeneous up to level k ($1 \leqslant k \leqslant m-1$) and heterogeneous afterward. Further, the mixture proportion p for component A is assumed to remain fixed across all heterogeneous stress levels. Under this framework, a link model involving four parameters can be constructed in the same manner as in Section 7 for the simple hSSALT model. Hence, estimating parameters in a link model is more parsimonious compared to estimating all θ_is to extrapolate to θ_0.

Acknowledgments

The authors sincerely thank the reviewers for their constructive and useful comments on an earlier version of the manuscript. The work is funded by the German Ministry of Education and Research, within the project BALd in the Cluster BattNutzung, Grant No. 03XP0320A.

Ethics Statement

The author has nothing to report.

Consent

The author has nothing to report.

Data Availability Statement

The data that support the findings of this study are openly available in *Annals of the Institute of Statistical Mathematics* at https://link.springer.com/article/10.1007/s10463-007-0135-3.

References

1. W. B. Nelson and W. Q. Meeker, "Theory for Optimum Accelerated Censored Life Tests for Weibull and Extreme Value Distributions," *Technometrics* 20, no. 2 (1978): 171–177.

2. W. B. Nelson, "Accelerated Life Testing-Step-Stress Models and Data Analyses," *IEEE Transactions on Reliability* 29, no. 2 (1980): 103–108.

3. R. Miller and W. B. Nelson, "Optimum Simple Step-Stress Plans for Accelerated Life Testing," *IEEE Transactions on Reliability* 32, no. 1 (1983): 59–65.

4. D. S. Bai, M. Kim, and S. Lee, "Optimum Simple Step-Stress Accelerated Life Tests With Censoring," *IEEE Transactions on Reliability* 38, no. 5 (1989): 528–532.

5. W. B. Nelson, *Accelerated Life Testing: Statistical Models, Test Plans and Data Analysis* (Wiley, 1990).

6. W. Q. Meeker and L. A. Escobar, *Statistical Methods for Reliability Data* (Wiley, 1998).

7. C. Xiong, "Inferences on a Simple Step-Stress Model With Type-II Censored Exponential Data," *IEEE Transactions on Reliability* 47, no. 2 (1998): 142–146.

8. N. Balakrishnan, D. Kundu, K. T. Ng, and N. Kannan, "Point and Interval Estimation for a Simple Step-Stress Model With Type-II Censoring," *Journal of Quality Technology* 39, no. 1 (2007): 35–47.

9. M. Kateri and N. Balakrishnan, "Inference for a Simple Step-Stress Model With Type-II Censoring, and Weibull Distributed Lifetimes," *IEEE Transactions on Reliability* 57, no. 4 (2008): 616–626.

10. N. Balakrishnan, Q. Xie, and D. Kundu, "Exact Inference for a Simple Step-Stress Model From the Exponential Distribution Under Time Constraint," *Annals of the Institute of Statistical Mathematics* 61 (2009): 251–274.

11. N. Balakrishnan, L. Zhang, and Q. Xie, "Inference for a Simple Step-Stress Model With Type-I Censoring and Lognormally Distributed Lifetimes," *Communications in Statistics - Theory and Methods* 38, no. 10 (2009): 1690–1709.

12. M. S. Kotb and M. M. Mohie El-Din, "Parametric Inference for Step-Stress Accelerated Life Testing From Rayleigh Distribution Under Ordered Ranked Set Sampling," *IEEE Transactions on Reliability* 71, no. 1 (2022): 16–27.

13. M. Kateri and N. I. Nikolov, "Product of Spacings Estimation in Step-Stress Accelerated Life Testing: An Alternative to Maximum Likelihood," *IEEE Transactions on Reliability* 73 (2024): 1433–1445.

14. W. Q. Meeker and M. J. LuValle, "An Accelerated Life Test Model Based on Reliability Kinetics," *Technometrics* 37, no. 2 (1995): 133–146.

15. S. J. Harris, D. J. Harris, and C. Li, "Failure Statistics for Commercial Lithium-Ion Batteries: A Study of 24 Pouch Cells," *Journal of Power Sources* 342 (2017): 589–597.

16. B. Zheng, C. Chen, and Y. Lin, et al., "Optimal Design of Step-Stress Accelerated Degradation Test Oriented by Nonlinear and Distributed Degradation Process," *Reliability Engineering & System Safety* 217 (2022): 108087.

17. E. K. Al-Hussaini and A. H. Abdel-Hamid, "Accelerated Life Tests Under Finite Mixture Models," *Journal of Statistical Computation and Simulation* 76, no. 8 (2006): 673–690.

18. A. H. Abdel-Hamid and E. K. Al-Hussaini, "Progressive Stress Accelerated Life Tests Under Finite Mixture Models," *Metrika* 66 (2007): 213–231.

19. A. H. Abdel-Hamid and E. K. Al-Hussaini, "Step Partially Accelerated Life Tests Under Finite Mixture Models," *Journal of Statistical Computation and Simulation* 78, no. 10 (2008): 911–924.

20. R. V. León, R. Ramachandran, A. J. Ashby, and J. Thyagarajan, "Bayesian Modeling of Accelerated Life Tests With Random Effects," *Journal of Quality Technology* 39, no. 1 (2007): 3–16.

21. S. Lv, Z. Niu, L. Qu, S. He, and Z. He, "Reliability Modeling of Accelerated Life Tests With Both Random Effects and Nonconstant Shape Parameters," *Quality Engineering* 27, no. 3 (2015): 329–340.

22. K. Lin, Y. Chen, and D. Xu, "Reliability Assessment Model Considering Heterogeneous Population in a Multiple Stresses Accelerated Test," *Reliability Engineering & System Safety* 165 (2017): 134–143.

23. K. Seo and R. Pan, "Data Analysis for Accelerated Life Tests With Constrained Randomization," in *2016 Annual Reliability and Maintainability Symposium (RAMS)* (IEEE, 2016), 1–7.

24. L. Zhuang, A. Xu, B. Wang, Y. Xue, and S. Zhang, "Data Analysis of Progressive-Stress Accelerated Life Tests With Group Effects," *Quality Technology & Quantitative Management* 20, no. 6 (2023): 763–783.

25. K. Seo and R. Pan, "Data Analysis of Step-Stress Accelerated Life Tests With Heterogeneous Group Effects," *IISE Transactions* 49, no. 9 (2017): 885–898.

26. J. Wang, "Data Analysis of Step-Stress Accelerated Life Test With Random Group Effects Under Weibull Distribution," *Mathematical Problems in Engineering* 2020 (2020): 1–11.

27. H. Teicher, "Identifiability of Mixtures," *Annals of Mathematical Statistics* 32, no. 1 (1961): 244–248.

28. H. Teicher, "Identifiability of Finite Mixtures," *Annals of Mathematical Statistics* 34, no. 4 (1963): 1265–1269.

29. S. Chandra, "On the Mixtures of Probability Distributions," *Scandinavian Journal of Statistics* 4, no. 3 (1977): 105–112.

30. O. Barndorff-Nielsen, "Identifiability of Mixtures of Exponential Families," *Journal of Mathematical Analysis and Applications* 12, no. 1 (1965): 115–121.

31. S. J. Yakowitz and J. D. Spragins, "On the Identifiability of Finite Mixtures," *Annals of Mathematical Statistics* 39, no. 1 (1968): 209–214.

32. A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum Likelihood From Incomplete Data via the EM Algorithm," *Journal of the Royal Statistical Society Series B: Statistical Methodology* 39, no. 1 (1977): 1–22.

33. D. Chauveau, "A Stochastic EM Algorithm for Mixtures With Censored Data," *Journal of Statistical Planning and Inference* 46, no. 1 (1995): 1–25.

34. H. Ng, P. Chan, and N. Balakrishnan, "Estimation of Parameters From Progressively Censored Data Using EM Algorithm," *Computational Statistics & Data Analysis* 39, no. 4 (2002): 371–386.

35. N. Balakrishnan, D. Mitra, "Likelihood Inference for Lognormal Data With Left Truncation and Right Censoring With an Illustration," *Journal of Statistical Planning and Inference* 141, no. 11 (2011): 3536–3553.

36. G. Lee and C. Scott, "EM Algorithms for Multivariate Gaussian Mixture Models With Truncated and Censored Data," *Computational Statistics & Data Analysis* 56, no. 9 (2012): 2816–2829.

37. S. F. Ateya, "Maximum Likelihood Estimation Under a Finite Mixture of Generalized Exponential Distributions Based on Censored Data," *Statistical Papers* 55 (2014): 311–325.

38. R. Verbelen, L. Gong, K. Antonio, A. Badescu, and S. Lin, "Fitting Mixtures of Erlangs to Censored and Truncated Data Using the EM Algorithm," *ASTIN Bulletin: The Journal of the IAA* 45, no. 3 (2015): 729–758.

39. G. J. McLachlan and T. Krishnan, *The EM Algorithm and Extensions* (John Wiley & Sons, 2007).

40. C. J. Wu, "On the Convergence Properties of the EM Algorithm," *Annals of Statistics* 11, no. 1 (1983): 95–103.

41. G. J. McLachlan and D. Peel, *Finite Mixture Models* (John Wiley & Sons, 2000).

42. G. K. Bhattacharyya, "The Asymptotics of Maximum Likelihood and Related Estimators Based on Type-II Censored Data," *Journal of the American Statistical Association* 80, no. 390 (1985): 398–404.

43. B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap* (CRC Press, 1994).

44. P. Bobotas and M. Kateri, "The Step-Stress Tampered Failure Rate Model Under Interval Monitoring," *Statistical Methodology* 27 (2015): 100–122.

45. K. N. Kirby and D. Gerlanc, "BootES: An R Package for Bootstrap Confidence Intervals on Effect Sizes," *Behavior Research Methods* 45 (2013): 905–927.

46. N. Balakrishnan, E. Castilla, M. Jaenada, and L. Pardo, "Robust Inference for Nondestructive One-Shot Device Testing Under Step-Stress Model With Exponential Lifetimes," *Quality and Reliability Engineering International* 39, no. 4 (2023): 1192–1222.

47. E. Castilla, "A New Estimation Approach Based on Phi-Divergence Measures for One-Shot Device Accelerated Life Testing," *Quality and Reliability Engineering International* 40, no. 4 (2024): 2048–2066.

Appendix A

The likelihood equations for the parameters θ_1, θ_{21}, θ_{22}, and p, of model (2) follow

$$\frac{\partial \ell(\theta_1,\theta_{21},\theta_{22},p\mid \boldsymbol{t})}{\partial \theta_1} = -\frac{n_1}{\theta_1} + \frac{\sum_{i=1}^{n_1} t_{i:n}}{\theta_1^2} + \frac{(n-n_1)\tau}{\theta_1^2} = 0, \tag{A1}$$

$$\frac{\partial \ell(\theta_1,\theta_{21},\theta_{22},p\mid \boldsymbol{t})}{\partial \theta_{21}} = \sum_{i=n_1+1}^{r} \frac{p\frac{1}{\theta_{21}^2}\exp\left\{-\frac{t_{i:n}-\tau}{\theta_{21}}\right\}\left(\frac{t_{i:n}-\tau}{\theta_{21}}-1\right)}{p\frac{1}{\theta_{21}}\exp\left\{-\frac{t_{i:n}-\tau}{\theta_{21}}\right\}+(1-p)\frac{1}{\theta_{22}}\exp\left\{-\frac{t_{i:n}-\tau}{\theta_{22}}\right\}} + \frac{(n-r)(t_{r:n}-\tau)p\frac{1}{\theta_{21}^2}\exp\left\{-\frac{t_{r:n}-\tau}{\theta_{21}}\right\}}{p\exp\left\{-\frac{t_{r:n}-\tau}{\theta_{21}}\right\}+(1-p)\exp\left\{-\frac{t_{r:n}-\tau}{\theta_{22}}\right\}} = 0, \tag{A2}$$

$$\frac{\partial \ell(\theta_1,\theta_{21},\theta_{22},p\mid \boldsymbol{t})}{\partial \theta_{22}} = \sum_{i=n_1+1}^{r} \frac{(1-p)\frac{1}{\theta_{22}^2}\exp\left\{-\frac{t_{i:n}-\tau}{\theta_{22}}\right\}\left(\frac{t_{i:n}-\tau}{\theta_{22}}-1\right)}{p\frac{1}{\theta_{21}}\exp\left\{-\frac{t_{i:n}-\tau}{\theta_{21}}\right\}+(1-p)\frac{1}{\theta_{22}}\exp\left\{-\frac{t_{i:n}-\tau}{\theta_{22}}\right\}} + \frac{(n-r)(t_{r:n}-\tau)(1-p)\frac{1}{\theta_{22}^2}\exp\left\{-\frac{t_{r:n}-\tau}{\theta_{22}}\right\}}{p\exp\left\{-\frac{t_{r:n}-\tau}{\theta_{21}}\right\}+(1-p)\exp\left\{-\frac{t_{r:n}-\tau}{\theta_{22}}\right\}} = 0, \tag{A3}$$

$$\frac{\partial \ell(\theta_1,\theta_{21},\theta_{22},p\mid \boldsymbol{t})}{\partial p} = \sum_{i=n_1+1}^{r} \frac{\frac{1}{\theta_{21}}\exp\left\{-\frac{t_{i:n}-\tau}{\theta_{21}}\right\}-\frac{1}{\theta_{22}}\exp\left\{-\frac{t_{i:n}-\tau}{\theta_{22}}\right\}}{p\frac{1}{\theta_{21}}\exp\left\{-\frac{t_{i:n}-\tau}{\theta_{21}}\right\}+(1-p)\frac{1}{\theta_{22}}\exp\left\{-\frac{t_{i:n}-\tau}{\theta_{22}}\right\}} + \frac{(n-r)\left(\exp\left\{-\frac{t_{r:n}-\tau}{\theta_{21}}\right\}-\exp\left\{-\frac{t_{r:n}-\tau}{\theta_{22}}\right\}\right)}{p\exp\left\{-\frac{t_{r:n}-\tau}{\theta_{21}}\right\}+(1-p)\exp\left\{-\frac{t_{r:n}-\tau}{\theta_{22}}\right\}} = 0. \tag{A4}$$

Appendix B

Regularity conditions are as follows [40]:

i. $\boldsymbol{\Omega}$ is a subset in d-dimensional Euclidean space $\mathbb{R}^d$;

ii. $\boldsymbol{\Omega}_{\vartheta_0} = \{\vartheta \in \boldsymbol{\Omega} : L(\vartheta) \geqslant L(\vartheta_0)\}$ is compact for any $L(\vartheta_0) > -\infty$;

iii. $L(\vartheta)$ is continuous in $\boldsymbol{\Omega}$ and differentiable in the interior of $\boldsymbol{\Omega}$.

Proof. Focusing on the estimation of the parameters under stress level x_2, the parameter space is $\boldsymbol{\Omega} = \{\vartheta = (p,\theta_{21},\theta_{22}) : p \in [0,1],\ \theta_{21},\theta_{22} \in (0,\theta_1),\ \theta_{21} < \theta_{22}\}$. The restriction $\theta_{21} < \theta_{22}$ is imposed for the model's identifiability, while $\theta_{22} < \theta_1$ is a natural restriction since stress level x_2 is higher than x_1.

i. This condition is easily proved since the parameter space $\boldsymbol{\Omega}$ is a subset in $\mathbb{R}^3$.

ii. $L(\cdot) = L(\cdot|\boldsymbol{t})$ is the observed likelihood function, given the observed failure times $\boldsymbol{t}$, and ϑ_0 is some arbitrary value of $\vartheta \in \boldsymbol{\Omega}$. Notice that for all $\vartheta \in \boldsymbol{\Omega}$, $\theta_{21} > 0$ and converges to some $\epsilon_1 > 0$. Similarly, since $\theta_{21} < \theta_{22} < \theta_1$, there exists an ϵ_2, $\epsilon_1 < \epsilon_2 < \theta_1$, such that $\epsilon_1 \leqslant \theta_{21} < \theta_{22} \leqslant \epsilon_2$. The parameter space is further restricted to $\epsilon_1 \leqslant \theta_{21} \leqslant \epsilon_1^*$, $\epsilon_2^* \leqslant \theta_{22} \leqslant \epsilon_2$, for some $\epsilon_1^* < \epsilon_2^*$. Because $L(\vartheta)$ is continuous, the set $\boldsymbol{\Omega}_{\vartheta_0} = \{\vartheta \in \boldsymbol{\Omega} : L(\vartheta) \geqslant L(\vartheta_0)\}$ is a closed subset of a compact set and therefore is also compact.

iii. It can be easily proved according to the definition of the likelihood function L. □

Appendix C

Proof. According to Equation (12), $h_{ij}^{(k)} = \frac{p_j^{(k)} f\left(t_{i:n}|\theta_{2j}^{(k)}\right)}{\sum_{j=1}^{2} p_j^{(k)} f\left(t_{i:n}|\theta_{2j}^{(k)}\right)}$, for $n_1 + 1 \leqslant i \leqslant r$ and $h_{ij}^{(k)} = \frac{p_j^{(k)} \overline{F}\left(t_{r:n}|\theta_{2j}^{(k)}\right)}{\sum_{j=1}^{2} p_j^{(k)} \overline{F}\left(t_{r:n}|\theta_{2j}^{(k)}\right)}$ remains constant for $r+1 \leqslant i \leqslant n$. It can be verified that under the imposed assumption $\theta_{21} <$

θ_{22}, $h_{i1}^{(k)}$ is monotonically decreasing for $n_1+1 \leqslant i \leqslant r$. Since $h_{i1}^{(k)} + h_{i2}^{(k)} = 1$, $h_{i2}^{(k)}$ increases monotonically. Under the assumption between θ_{2j}s, $h_{r1}^{(k)} > h_{r+1,1}^{(k)}$ can be easily proved. Meanwhile, $h_{r2}^{(k)} < h_{r+1,2}^{(k)}$. $\hat{\theta}_2^*$ can be reformulated as $\hat{\theta}_2^* = \frac{1}{n-n_1}[\frac{(n-r)h_{r+1,1}^{(k)}}{\sum_{i=n_1+1}^{r} h_{i1}^{(k)}} \sum_{i=n_1+1}^{n} h_{i1}^{(k)}(t_{i:n}-\tau) + \frac{(n-r)h_{r+1,2}^{(k)}}{\sum_{i=n_1+1}^{r} h_{i2}^{(k)}} \sum_{i=n_1+1}^{n} h_{i2}^{(k)}(t_{i:n}-\tau) + \sum_{i=n_1+1}^{n}(t_{i:n}-\tau)]$. To prove $\hat{\theta}_2^* > \hat{\theta}_2$, we can equivalently prove the following:

$$\underbrace{\frac{h_{r+1,1}^{(k)}}{\sum_{i=n_1+1}^{r} h_{i1}^{(k)}} \sum_{i=n_1+1}^{n} h_{i1}^{(k)}(t_{i:n}-\tau) + \frac{h_{r+1,2}^{(k)}}{\sum_{i=n_1+1}^{r} h_{i2}^{(k)}} \sum_{i=n_1+1}^{n} h_{i2}^{(k)}(t_{i:n}-\tau)}_{(*)}$$

$$> \frac{1}{r-n_1} \sum_{i=n_1+1}^{n} (t_{i:n}-\tau).$$

The left side of the last inequality, $(*)$, can be rearranged as $\sum_{i=n_1+1}^{n}(t_{i:n}-\tau)(\frac{h_{r+1,1}^{(k)} h_{i1}^{(k)}}{\sum_{i=n_1+1}^{r} h_{i1}^{(k)}} + \frac{h_{r+1,2}^{(k)} h_{i2}^{(k)}}{\sum_{i=n_1+1}^{r} h_{i2}^{(k)}})$. Notice that the first multiplicative term of the summand above, $(t_{i:n}-\tau)$, is monotonically increasing for $i \geqslant n_1+1$. Furthermore, it can be proved that the second multiplicative term is also monotonically increasing. By applying Chebyshev's sum inequality, we get

$$(*) > \frac{1}{n-n_1} \sum_{i=n_1+1}^{n} (t_{i:n}-\tau)\left[1 + (n-r)\left(\frac{\left(h_{r+1,1}^{(k)}\right)^2}{\sum_{i=n_1+1}^{r} h_{i1}^{(k)}} + \frac{\left(h_{r+1,2}^{(k)}\right)^2}{\sum_{i=n_1+1}^{r} h_{i2}^{(k)}}\right)\right],$$

with the inequality being strict, since some sum terms are strictly increasing. Then the objective is to show $\frac{\left(h_{r+1,1}^{(k)}\right)^2}{\sum_{i=n_1+1}^{r} h_{i1}^{(k)}} + \frac{\left(h_{r+1,2}^{(k)}\right)^2}{\sum_{i=n_1+1}^{r} h_{i2}^{(k)}} \geqslant \frac{1}{r-n_1}$. However, this holds always due to $[(r-n_1)h_{r+1,1}^{(k)} - \sum_{i=n_1+1}^{r} h_{i1}^{(k)}]^2 \geq 0$, which can be easily verified. Therefore, we conclude that, under Type-II censoring, $\hat{\theta}_2^* > \hat{\theta}_2$ holds. □

Appendix D

The observed Fisher information matrix is a symmetric matrix:

$$\mathbf{I}_{obs}(\boldsymbol{\eta}) = \mathbf{I}_{obs}(\eta_1,\eta_2,\eta_3,\eta_4) = (O_{ij}) = \left(-\frac{\partial^2 l}{\partial\eta_i\partial\eta_j}\right),$$

with the elements

$$O_{11} = -\frac{n_1}{\theta_1^2} + \frac{2\left[\sum_{i=1}^{n_1} t_{i:n} + (n-n_1)\tau\right]}{\theta_1^3},$$

$$O_{22} = -\left[\sum_{i=n_1+1}^{r}\left(\frac{pF_{1i}(F_{2i}^2-4F_{2i}+2)}{\theta_{21}^3 A_i} - \left(\frac{pF_{1i}(F_{2i}-1)}{\theta_{21}^2 A_i}\right)^2\right) + (n-r)\left(\frac{(t_{r:n}-\tau)pF_{1r}(F_{2r}-2)}{\theta_{21}^3 B} - \left(\frac{(t_{r:n}-\tau)pF_{1r}}{\theta_{21}^2 B}\right)^2\right)\right],$$

$$O_{33} = -\left[\sum_{i=n_1+1}^{r}\left(\frac{(1-p)F_{3i}(F_{4i}^2-4F_{2i}+2)}{\theta_{22}^3 A_i} - \left(\frac{(1-p)F_{3i}(F_{4i}-1)}{\theta_{22}^2 A_i}\right)^2\right) + (n-r)\left(\frac{(t_{r:n}-\tau)(1-p)F_{3r}(F_{4r}-2)}{\theta_{22}^3 B} - \left(\frac{(t_{r:n}-\tau)(1-p)F_{3r}}{\theta_{22}^2 B}\right)^2\right)\right],$$

$$O_{44} = \left[\sum_{i=n_1+1}^{r}\left(\frac{\theta_{22}F_{1i}-\theta_{21}F_{3i}}{\theta_{21}\theta_{22}A_i}\right)^2 + (n-r)\left(\frac{F_{1r}-F_{3r}}{B}\right)^2\right],$$

$$O_{12} = O_{13} = O_{14} = 0,$$

$$O_{23} = \frac{p(1-p)}{(\theta_{21}\theta_{22})^2}\left[\sum_{i=n_1+1}^{r}\frac{F_{1i}F_{3i}(F_{2i}-1)(F_{4i}-1)}{A_i^2} + \frac{(n-r)(t_{r:n}-\tau)^2 F_{1r}F_{3r}}{B^2}\right],$$

$$O_{24} = -\left[\sum_{i=n_1+1}^{r}\left(\frac{F_{1i}(F_{2i}-1)}{\theta_{21}^2 A_i} - \frac{pF_{1i}(F_{2i}-1)(\theta_{22}F_{1i}-\theta_{21}F_{3i})}{\theta_{21}^3\theta_{22}A_i^2}\right) + (n-r)\frac{t_{r:n}-\tau}{\theta_{21}^2}\left(\frac{F_{1r}}{B} - \frac{pF_{1r}(F_{1r}-F_{3r})}{B^2}\right)\right],$$

$$O_{34} = \left[\sum_{i=n_1+1}^{r}\left(\frac{F_{3i}(F_{4i}-1)}{\theta_{22}^2 A_i} + \frac{(1-p)F_{3i}(F_{4i}-1)(\theta_{22}F_{1i}+\theta_{21}F_{3i})}{\theta_{21}\theta_{22}^3 A_i^2}\right) + (n-r)\frac{t_{r:n}-\tau}{\theta_{22}^2}\left(\frac{F_{3r}}{B} + \frac{(1-p)F_{3r}(F_{1r}-F_{3r})}{B^2}\right)\right],$$

where $F_{1i} = \exp\{-\frac{t_{i:n}-\tau}{\theta_{21}}\}$, $F_{2i} = \frac{t_{i:n}-\tau}{\theta_{21}}$, $F_{3i} = \exp\{-\frac{t_{i:n}-\tau}{\theta_{22}}\}$, $F_{4i} = \frac{t_{i:n}-\tau}{\theta_{22}}$,

$$A_i = p\frac{1}{\theta_{21}}\exp\left\{-\frac{t_{i:n}-\tau}{\theta_{21}}\right\} + (1-p)\frac{1}{\theta_{22}}\exp\left\{-\frac{t_{i:n}-\tau}{\theta_{22}}\right\},$$

$$i = n_1+1, \ldots, r,$$

$$B = p\exp\left\{-\frac{t_{r:n}-\tau}{\theta_{21}}\right\} + (1-p)\exp\left\{-\frac{t_{r:n}-\tau}{\theta_{22}}\right\}.$$