JIRSS (2021)
Vol. 20, No. 01, pp 123-152
DOI:10.52547/jirss.20.1.123

Stage Life Testing with Missing Stage Information - an EM-Algorithm Approach

Erhard Cramer [1], **and Benjamin Laumen** [1]

[1] Institute of Statistics, RWTH Aachen University, D-52062 Aachen, GERMANY.

Received: 23/11/2020, Revision received: 07/02/2021, Published online: 03/04/2021

Abstract. We consider a stage life testing model and assume that the information at which levels the failures occurred is not available. In order to find estimates for the lifetime distribution parameters, we propose an EM-algorithm approach which interprets the lack of knowledge about the stages as missing information. Furthermore, we illustrate the implementation difficulties caused by an increasing number of stages. The study is supplemented by a data example as well as simulations.

Keywords. EM-Algorithm, Exponential Distribution, Missing Information, Progressive Censoring, Stage Life Testing, Weibull Distribution

MSC: 62N05, 62F10.

1 Introduction

The notion of stage life testing (SLT) has been proposed in Laumen (2017) and Laumen and Cramer (2019b, 2021a) as an extension of progressive Type-I censoring (for a version with random stage changing times, see Laumen and Cramer (2021b)). The approach provides models that allow to incorporate additional life time information

Corresponding Author: Erhard Cramer (erhard.cramer@rwth-aachen.de)
Benjamin Laumen (benjamin.laumen@rwth-aachen.de)

of progressively censored objects by performing additional testing of the removed items (for comments in this direction, cf., e.g., Balakrishnan and Aggarwala 2000, p. 3, Balakrishnan et al. 2011, p. 336, Balakrishnan and Cramer (2014, 2021), Cramer (2017)). In fact, it is assumed that the progressively censored objects are further tested but under different conditions (called stages) whereas the remaining items are continued to be monitored under the initial conditions. An illustration of this concept is depicted in Figure 1 for $k-1$ stage changing times $\tau_1 < \cdots < \tau_{k-1}$ and an effectively applied stage changing plan $(r_1^\star, \dots, r_{k-1}^\star)$. The experimental design of such a life test requires that $r_j^\star$ objects are randomly withdrawn from the life test at time τ_j, $1 \leq j \leq k-1$. Notice that the stage changing plan may be specified in different ways. The life span test is terminated when either the last remaining object in the life test fails or the last item is removed from the test. Notice that we do not assume that the experiment is stopped at time τ_{k-1} as is commonly done in progressive Type-I censoring (cf. progressive Type-I censoring with fixed censoring times discussed in Laumen and Cramer 2019a).

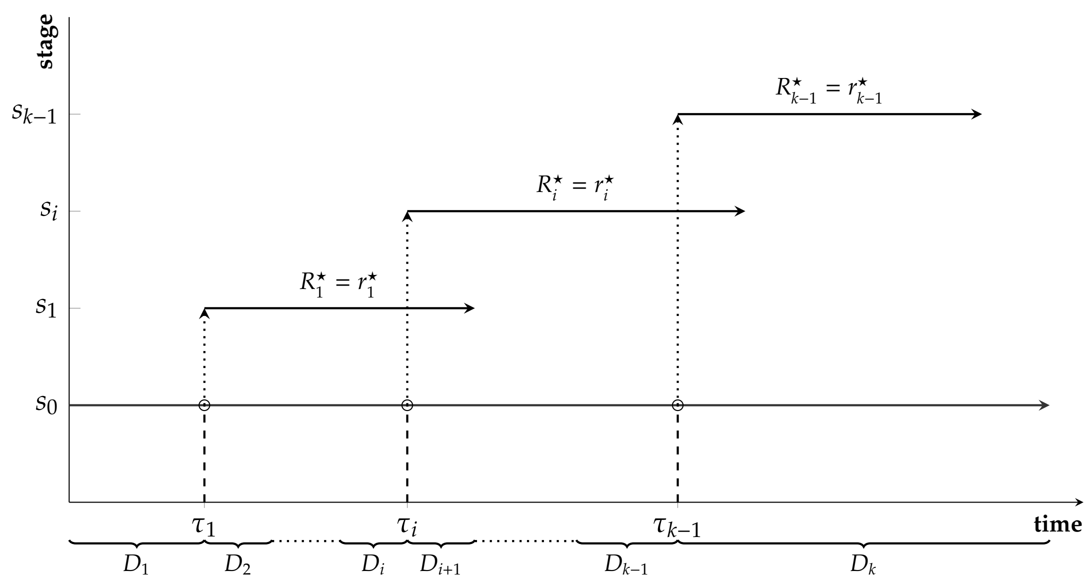

Figure 1: Illustration of k-step SLT with stage-changing times $\tau_1 < \cdots < \tau_{k-1}$ and an effectively applied stage changing plan $(r_1^\star, \dots, r_{k-1}^\star)$.

As can be seen from Figure 1, the sample is split at change-time τ_1 in these items which are tested under the initial conditions (stage s_0), and those which are tested on stage s_1 with a possibly different load (which, of course, may be higher or lower). This process is continued for the items remaining on stage s_0 at the following stage changing times $\tau_2, \ldots, \tau_{k-1}$. The adaption of the load is modelled by the cumulative exposure model approach (see, e.g., Kundu and Ganguly (2017)). Furthermore, it should be mentioned that SLT can also be interpreted as a modified simple step-stress model (see Balakrishnan (2009), Kundu and Ganguly (2017)) where only a proportion of the items tested under initial conditions is selected for testing under other stress conditions. In this regard, SLT can also be interpreted as a model of accelerated life testing. For details, we refer to Laumen and Cramer (2019b, 2021a).

This paper is organized as follows. In Section 2, we introduce briefly the SLT model and recall some results presented in Laumen and Cramer (2019b, 2021a). In Section 3, we address maximum likelihood estimation in the SLT model under missing stage information and present the likelihood for k-step SLT. Afterwards, we illustrate the EM-algorithm approach for the 2- and 3-step SLT. In particular, we start with an exponential distribution on both stages in Section 3.2.1. In Section 3.2.2, we consider the combination of a Weibull and an exponential distribution for the stages of the SLT model. In Section 4.1, we provide an illustrative example. Finally, we present the results of a simulation study in Section 4.2.

2 SLT Model

Our discussion is based on the notation of k-step SLT order statistics as introduced in Laumen and Cramer (2021a). Assume that n identical objects with iid lifetimes $X_1, \ldots, X_n$ are placed on a life test at the initial stage s_0. At the jth prefixed stage-change time τ_j, $R_j^\star \geq 0$ of the surviving items are randomly withdrawn (if possible) from the sample and further tested on the changed stage s_j, the remaining objects are further tested under the conditions of stage s_0, $1 \leq j \leq k-1$. The life test is terminated when all n objects have failed.

The (random) numbers of failures observed on the initial stage s_0 in the intervals

$$(-\infty, \tau_1],\ (\tau_1, \tau_2],\ \ldots,\ (\tau_{k-2}, \tau_{k-1}],\ (\tau_{k-1}, \infty),$$

are denoted by $D_1, D_2, \ldots, D_{k-1}, D_k$, where $\tau_0 = -\infty$, $\tau_k = \infty$ (see Figure 1).

Let $M = D_{\bullet k} = \sum_{j=1}^{k} D_j$ be the total number of observations failed on level s_0 where $D_{\bullet 0} = 0$. Arranging the data according to the stage levels $s_0, s_1, \ldots, s_{k-1}$, the observed failure times on these levels are denoted by

- $\boldsymbol{Y}_{h,D_h} = \left(Y_{D_{\bullet h-1}+1:M:n}, \ldots, Y_{D_{\bullet h}:M:n}\right)$ denote the ordered failure times observed in the interval $(\tau_{h-1}, \tau_h]$ on stage s_0, $h = 1, \ldots, k$. Notice that $\boldsymbol{Y}_{1,D_1}, \ldots, \boldsymbol{Y}_{k,D_k}$ forms a progressively Type-I censored sample with fixed censoring times as discussed in Laumen and Cramer (2019a). This connection is reflected by the notation used;
- $\boldsymbol{Z}_{j,R_j^\star} = \left(Z_{j,1:R_j^\star}, \ldots, Z_{j,R_j^\star:R_j^\star}\right)$ with $Y_{D_{\bullet j}:M:n} \le \tau_j < Z_{j,1:R_j^\star}$ denote the ordered failure times observed on stage s_j, $j = 1, \ldots, k-1$.

The order statistics on stage s_0 and the order statistics on the stages $s_1, \ldots, s_{k-1}$ are represented by the vectors $\boldsymbol{Y} = \left(\boldsymbol{Y}_{1,D_1}, \ldots, \boldsymbol{Y}_{k,D_k}\right)$ and $\boldsymbol{Z} = \left(\boldsymbol{Z}_{1,R_1^\star}, \ldots, \boldsymbol{Z}_{k-1,R_{k-1}^\star}\right)$, respectively. The complete sample is given by $(\boldsymbol{Y}, \boldsymbol{Z})$. Notice that the partitioning of the sample induces the assignment of failures to stages. But, in the following, we assume that the information about this assignment is not available. In order to describe the situation, we introduce random indicators $\Sigma_1, \ldots, \Sigma_n$ which provide the information about the stage, that is,

$$\Sigma_i = \begin{cases} 0, & \text{object } i \text{ has failed on stage } s_0 \\ 1, & \text{object } i \text{ has failed on stage } s_1 \\ & \vdots \\ k-1, & \text{object } i \text{ has failed on stage } s_{k-1} \end{cases}, \quad i = 1, \ldots, n;$$

indicates whether an object has failed on stage $s_0, \ldots, s_{k-1}$. Thus, each observation X_i is accompanied by an indicator Σ_i so that we observe a pair (X_i, Σ_i) where Σ_i provides the information about the stage. By considering the order statistics $\boldsymbol{X}^* = (X_{1:n}, \ldots, X_{n:n})$ of the sample $\boldsymbol{X}$, the stage indicators can be interpreted as a concomitant (see David and Nagaraja (1998), Bairamov and Eryılmaz (2006), Izadi and Khaledi (2007), Balakrishnan and Cramer (2014)), that is, we get the (bivariate) 'ordered' sample

$$(X_{1:n}, \Sigma_{[1:n]}), \ldots, (X_{n:n}, \Sigma_{[n:n]}).$$

Due to the construction of the sample, we know that $X_{h:n} = Y_{h:M:n}$ and $\Sigma_{[h:n]} = 0$, $h = 1, \ldots, D_1$. For brevity, we subsequently write $X_i^* = X_{i:n}, \Sigma_i^* = \Sigma_{[i:n]}$, $1 \le i \le n$. In the present discussion, we use the following notation and assumptions:

- Throughout the manuscript, we use the notation $\boldsymbol{w}_j = (w_1, \ldots, w_j)$ for the vector of j components $w_1, \ldots, w_j$ as well as $w_{\bullet j} = \sum_{i=1}^{j} w_i$ for their partial sum.

- The random censoring number $R_j^\star$ is generated by a (deterministic) function ϱ_j of the failures observed before τ_j, that is, $R_j^\star = \varrho_j(\boldsymbol{D}_j)$, $1 \le j \le k-1$. In the following, these functions may be chosen according to the needs of the experimenter. In Laumen and Cramer (2019b, 2021a) two options to generate the withdrawal number $R_j^\star$, $j = 1, \ldots, k-1$, have been proposed:

$$\varrho_j(\boldsymbol{d}_j) = \begin{cases} \left\lfloor \pi_j \cdot \left(n - d_{\bullet j} - \sum_{i=1}^{j-1} \varrho_i(\boldsymbol{d}_i)\right)\right\rfloor, & \text{Type-P}, \\ \min\left\{R_j^0, \max\{n - d_{\bullet j} - \sum_{i=1}^{j-1} \varrho_i(\boldsymbol{d}_i), 0\}\right\}, & \text{Type-M}, \end{cases} \tag{2.1}$$

where $\boldsymbol{d}_j = (d_1, \ldots, d_j)$, $1 \le j \le k$, and $\lfloor t \rfloor$ is defined as the largest integer not exceeding $t \in \mathbb{R}$. The proportions $\pi_j \in (0,1)$ as well as the numbers $R_j^0 \in \mathbb{N}$, $1 \le j \le k-1$, are specified in advance, respectively. These choices can be interpreted as follows:

  - *Type-P*: At τ_j, a (prefixed) proportion π_j of the surviving objects is selected for testing on stage s_j, $j = 1, \ldots, k-1$.
  - *Type-M*: The second way to generate $R_j^\star$, $j = 1, \ldots, k-1$, is similar to the censoring procedure of progressive censoring with fixed failure times (see Laumen and Cramer (2019a)). Given a prefixed number R_j^0, it is intended to select at τ_j as many items as possible (at most R_j^0) for testing on stage s_j, $j = 1, \ldots, k-1$.

By construction, D_k is a (deterministic) function of $D_1, \ldots, D_{k-1}$ $\left(\text{and } (R_1^\star, \ldots, R_{k-1}^\star)\right)$, that is,

$$D_k = n - D_{\bullet k-1} - R_{\bullet k-1}^\star = n - D_{\bullet k-1} - \sum_{i=1}^{k-1} \varrho_i(\boldsymbol{D}_i).$$

- The support of $(D_1, \ldots, D_k)$ is represented by the set

$$\mathfrak{D}_{(k)} = \Big\{\boldsymbol{a}_k \in \mathbb{N}_0^k \,\Big|\, a_i \le \max\{n - a_{\bullet i-1} - r_{\bullet i-1}^\star(\boldsymbol{a}_{i-1}), 0\},\ i = 1, \ldots, k-1, \\ a_k = \max\{n - a_{\bullet k-1} - r_{\bullet k-1}^\star(\boldsymbol{a}_{k-1}), 0\}\Big\};$$

- $\mathscr{P}_* = (\pi_1, \ldots, \pi_{k-1})$ denotes the proportional stage-changing plan (see Type-P);
- $\mathscr{R}_*^0 = (R_1^0, \ldots, R_{k-1}^0)$ denotes the initially planned stage-changing plan (see Type-M);

- ◇ F_j denotes the absolutely continuous cumulative distribution function with density function f_j on stage s_j, $j \in \{0, \ldots, k-1\}$;
- ◇ The cumulative exposure model is supposed to hold, that is, the distribution function F_j on stage j is connected to the baseline distribution function F_0 as follows: For $1 \leq j \leq k-1$, we have values $v_1, \ldots, v_{k-1}$ such that v_j is the solution of the equation

  $$F_0(\tau_j) = F_j(v_j).$$

  Hence, given the stage-changing time τ_j, the cumulative distribution function and the corresponding probability density functions of a test unit on stage s_j are given by

  $$F_{0,j}(t) = \begin{cases} F_0(t), & t \leq \tau_j \\ F_j(t + v_j - \tau_j), & \tau_j < t \end{cases}, \qquad f_{0,j}(t) = \begin{cases} f_0(t), & t \leq \tau_j \\ f_j(t + v_j - \tau_j), & \tau_j < t \end{cases}. \tag{2.2}$$

  Details on the cumulative exposure model can be found in Kundu and Ganguly (2017, Chapter 2).

Laumen and Cramer (2021a) have obtained the joint density function of k-step SLT order statistics as given in Theorem 2.1.

Theorem 2.1. *Let $\boldsymbol{Y}_{1,D_1}, \ldots, \boldsymbol{Y}_{k,D_k}$ and $\boldsymbol{Z}_{1,R_1^\star}, \ldots, \boldsymbol{Z}_{k-1,R_{k-1}^\star}$ be k-step SLTOSs and let F_i be an absolutely continuous cumulative distribution function with density function f_i, $i \in \{0, \ldots, k-1\}$. Further, let $-\infty = \tau_0 < \tau_1 < \cdots < \tau_{k-1} < \tau_k = \infty$.*

Then, the joint density function $f_{1\ldots n}^{\boldsymbol{Y},\boldsymbol{Z},\boldsymbol{D}_k}$ of $\boldsymbol{Y} = \left(\boldsymbol{Y}_{1,D_1}, \ldots, \boldsymbol{Y}_{k,D_k}\right)$, $\boldsymbol{Z} = \left(\boldsymbol{Z}_{1,R_1^\star}, \ldots, \boldsymbol{Z}_{k-1,R_{k-1}^\star}\right)$ and $\boldsymbol{D}_k = (D_1, \ldots, D_k)$ (w.r.t. the product of the n dimensional Lebesgue measure and the k dimensional counting measure) is given by

$$f^{\boldsymbol{Y},\boldsymbol{Z},\boldsymbol{D}_k}(\boldsymbol{y}, \boldsymbol{z}, \boldsymbol{d}_k) = \prod_{j=1}^{k} \binom{n - d_{\bullet j-1} - r_{\bullet j-1}^\star}{d_j} d_j! \left\{ \prod_{i=d_{\bullet j-1}+1}^{d_{\bullet j}} f_0(y_{i:m:n}) \mathbb{1}_{(\tau_{j-1},\tau_j]}(y_{i:m:n}) \right\}$$
$$\times \left\{ \prod_{a=1}^{k-1} r_a^\star! \prod_{b=1}^{r_a^\star} f_a(z_{a,b:r_a^\star} + v_a - \tau_a) \mathbb{1}_{(\tau_a,\infty)}(z_{a,b:r_a^\star}) \right\}, \tag{2.3}$$

for $\boldsymbol{d}_k \in \mathfrak{D}_{(k)}$,

$$\boldsymbol{y} = \left(\boldsymbol{y}_{1,d_1}, \ldots, \boldsymbol{y}_{k,d_k}\right) \quad \text{with} \quad \boldsymbol{y}_{j,d_j} = \left(y_{d_{\bullet j-1}+1:m:n}, \ldots, y_{d_{\bullet j}:m:n}\right), \quad 1 \leq j \leq k,$$
$$\boldsymbol{z} = \left(\boldsymbol{z}_{1,r_1^\star}, \ldots, \boldsymbol{z}_{k-1,r_{k-1}^\star}\right) \quad \text{with} \quad \boldsymbol{z}_{a,r_a^\star} = \left(z_{a,1:r_a^\star}, \ldots, z_{a,r_a^\star:r_a^\star}\right), \quad 1 \leq a \leq k-1.$$

Notice that $(\boldsymbol{Y}, \boldsymbol{Z}, \boldsymbol{D}_k)$ are determined by $(\boldsymbol{X}, \Sigma)$ and vice versa. Therefore, using the above notation, the density function in (2.3) can equivalently be written as

$$f^{\boldsymbol{X},\Sigma}(\boldsymbol{x}, \boldsymbol{\sigma}) = c \prod_{i:\sigma_i=0} f_0(x_i) \prod_{j=1}^{k-1} \prod_{i:\sigma_i=j} f_j(x_i + v_j - \tau_j), \tag{2.4}$$

with $d_j = \sum_{i:\sigma_i=0} \mathbb{1}_{(\tau_{j-1},\tau_j]}(x_i)$, $j = 1, \ldots, k$, and normalizing constant $c = c(\boldsymbol{d}_k)$. The corresponding density function of the ordered data is given by

$$f^{\boldsymbol{X}^*,\Sigma^*}(\boldsymbol{x}^*, \boldsymbol{\sigma}^*) = c^* \prod_{i:\sigma_i^*=0} f_0(x_i^*) \prod_{j=1}^{k-1} \prod_{i:\sigma_i^*=j} f_j(x_i^* + v_j - \tau_j), \tag{2.5}$$

with $d_j = \sum_{i:\sigma_i^*=0} \mathbb{1}_{(\tau_{j-1},\tau_j]}(x_i^*)$, $j = 1, \ldots, k$, and normalizing constant $c^* = c^*(\boldsymbol{d}_k)$.

3 Maximum Likelihood Estimation in SLT under Missing Stage Information

3.1 k-step SLT

Given a statistical model with parameter vector $\boldsymbol{\theta} = (\vartheta_j)_{j=0,\ldots,k-1} \in \Theta = \bigtimes_{j=0}^{k-1} \Theta_j$, the likelihood function is obtained from (2.5) as

$$\mathsf{L}(\vartheta_0, \ldots, \vartheta_{k-1} \mid \boldsymbol{x}^*, \boldsymbol{\sigma}^*) = f^{\boldsymbol{X}^*,\Sigma^*}_{\vartheta_0,\ldots,\vartheta_{k-1}}(\boldsymbol{x}^*, \boldsymbol{\sigma}^*) \propto \prod_{i:\sigma_i^*=0} f_{0,\vartheta_0}(x_i^*) \prod_{j=1}^{k-1} \prod_{i:\sigma_i^*=j} f_{j,\vartheta_j}(x_i^* + v_j - \tau_j), \tag{3.1}$$

with density functions f_{j,ϑ_j}, $j = 0, \ldots, k-1$. Assuming $\boldsymbol{\sigma}^*$ as known, that is, we know which failure occurred on which stage, likelihood inference has been discussed for exponential and Weibull distributions in Laumen and Cramer (2019b, 2021a). However, if the stage information $\boldsymbol{\Sigma}^* = \boldsymbol{\sigma}$ is not available, the respective likelihood is obtained from (3.1) by summing over all possible values of $\boldsymbol{\sigma}$ so that the corresponding likelihood is obtained as marginal density function of $\boldsymbol{X}^*$. It reads

$$\mathsf{L}_{\mathrm{MI}}(\vartheta_0, \ldots, \vartheta_{k-1} \mid \boldsymbol{x}^*) = \sum_{\boldsymbol{\sigma}} f^{\boldsymbol{X}^*,\Sigma^*}_{\vartheta_0,\ldots,\vartheta_{k-1}}(\boldsymbol{x}^*, \boldsymbol{\sigma}) \propto \sum_{\boldsymbol{\sigma}} \prod_{i:\sigma_i=0} f_{0,\vartheta_0}(x_i^*) \prod_{j=1}^{k-1} \prod_{i:\sigma_i=j} f_{j,\vartheta_j}(x_i^* + v_j - \tau_j). \tag{3.2}$$

Notice that the normalizing constant does not depend on the failure assignment $\boldsymbol{\sigma}$. Due to the sum representation, direct optimization of the likelihood (3.2) may be quite hard.

However, treating the values of $\boldsymbol{\Sigma}^* = \boldsymbol{\sigma}$ as missing information, allows to address the maximization problem by an EM-algorithm type approach. In case of an exponential distribution $\mathsf{Exp}(\vartheta)$ with mean $\vartheta > 0$, that is, the probability density function and the cumulative distribution function are given by

$$f(x) = \frac{1}{\vartheta}\mathrm{e}^{-x/\vartheta}\mathbb{1}_{(0,\infty)}(x), \quad F(x) = \left(1 - \mathrm{e}^{-x/\vartheta}\right)\mathbb{1}_{(0,\infty)}(x), \quad \vartheta > 0,\ x \in \mathbb{R}.$$

this approach is particularly useful since the resulting MLEs (under complete information) are available in a closed form representation. Suppose the lifetime distributions on stage s_j are exponential with mean ϑ_j, $j = 0, \ldots, k-1$. Then, one gets $v_j = \tau_1 \frac{\vartheta_j}{\vartheta_0}$. Furthermore, from (3.1), the likelihood function is given by

$$\begin{aligned}
&\mathsf{L}(\vartheta_0, \ldots, \vartheta_{k-1} \mid \boldsymbol{x}^*, \boldsymbol{\sigma}) \\
&\quad = c^*\vartheta_0^{-m}\prod_{a=1}^{k-1}\vartheta_a^{-r_j^\star}\exp\left\{-\sum_{a=1}^{k-1} r_a^\star \tau_a\left(\frac{1}{\vartheta_0} - \frac{1}{\vartheta_a}\right) - \frac{1}{\vartheta_0}\sum_{j=1}^{n} x_j^*\mathbb{1}_{\{0\}}(\sigma_j) - \sum_{a=1}^{k-1}\frac{1}{\vartheta_a}\sum_{j=1}^{n} x_j^*\mathbb{1}_{\{a\}}(\sigma_j)\right\}.
\end{aligned} \tag{3.3}$$

Using results of Laumen and Cramer (2021a), the corresponding MLEs are given by

$$\widehat{\vartheta}_0 = \frac{1}{M}\left(\sum_{i=1}^{n} X_i^*\mathbb{1}_{\{0\}}(\Sigma_i) + \sum_{j=1}^{k-1} R_j^\star\,\tau_j\right), \tag{3.4}$$

and, provided $R_h^\star > 0$,

$$\widehat{\vartheta}_h = \frac{1}{R_h^\star}\sum_{i=1}^{n}(X_i^* - \tau_h)\mathbb{1}_{\{h\}}(\Sigma_i), \quad h \in \{1, \ldots, k-1\}. \tag{3.5}$$

3.2 EM-Algorithm for Two-Step SLT

3.2.1 Exponential-Exponential Case

Let the sample $(X_1^*, \Sigma_1^*), \ldots, (X_n^*, \Sigma_n^*)$ be incomplete in the sense that the information on which stage the observed failures occurred is missing, that is, the values of the stage indicators Σ^*'s are not available. The design of the life test is still known, i.e., we know n, τ_1, π_1, and R_1^0, respectively. Furthermore, the ordered failure times $x_1^* \leq \ldots \leq x_n^*$ are observed. This situation is depicted in Figure 2. The question marks indicate that the assignment of the failure to the stage is not known. Notice that the failure times $x_{d_1+1}^*, \ldots, x_n^*$ could have been observed on stages s_0 or s_1. In order to estimate the unknown parameters ϑ_0 and ϑ_1, we utilize an EM-algorithm (see, e.g., Dempster et al. 1977).

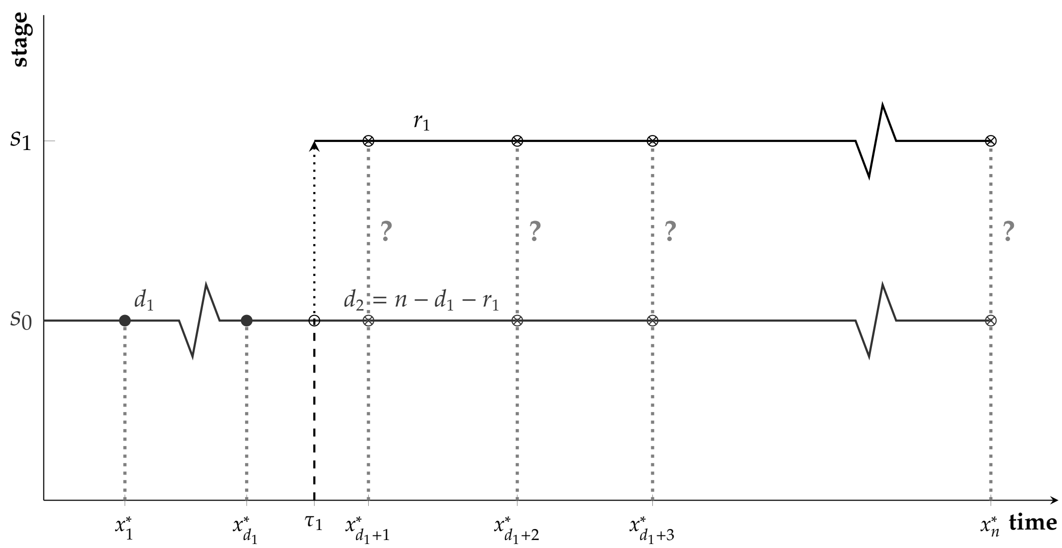

Figure 2: 2-step SLTOSs with missing information. The observations $x_{d_1+1}^*, \ldots, x_n^*$ could have been observed on stages s_0 or s_1.

Using the observations $x_1^*, \ldots, x_n^*$ and the design of the life test, the quantities d_1, d_2, and $r_1^\star$ are obtained as

$$d_1 = \sum_{i=1}^{n} \mathbb{1}_{(-\infty,\tau_1]}(x_i^*), \quad r_1^\star = \varrho(d_1) = \begin{cases} \lfloor \pi_1 \cdot (n - d_1) \rfloor, & \text{Type-P} \\ \min\{n - d_1, R_1^0\}, & \text{Type-M} \end{cases}, \quad d_2 = n - d_1 - r_1^\star.$$

Remark 1. (i) When $d_1 = n$, we know that all observed failures occurred on stage s_0. Hence, $\widehat{\vartheta}_0$ can be determined from equation (3.4) and the MLE $\widehat{\vartheta}_1$ does not exist.

(ii) When $d_1 \geq 0$, $d_2 > 0$, and $r_1^\star = 0$, we know that all observed failures after τ_1 occurred on stage s_0. Thus, the MLE $\widehat{\vartheta}_1$ does not exist and $\widehat{\vartheta}_0$ can be determined from equation (3.4).

(iii) When $d_1 > 0$, $d_2 = 0$, and $r_1^\star > 0$, we know that all observed failures after τ_1 occurred on stage s_1. Therefore, $\widehat{\vartheta}_0$ and $\widehat{\vartheta}_1$ can be obtained from (3.4) and (3.5), respectively.

In order to define the EM-algorithm in the present situation, we consider first the scenario of available stage information $(\boldsymbol{X}^*, \boldsymbol{\Sigma}^*)$. Let $\boldsymbol{\theta} = (\vartheta_0, \vartheta_1)$ and $\boldsymbol{\theta}^{(t)} = \left(\vartheta_0^{(t)}, \vartheta_1^{(t)}\right)$, $t \in \mathbb{N}_0$. Using equation (3.3), the likelihood function for the complete data $(\boldsymbol{X}^*, \boldsymbol{\Sigma}^*) = (\boldsymbol{x}^*, \boldsymbol{\sigma}^*)$ can be written in the form

$$\begin{aligned}
\mathsf{L}(\boldsymbol{\theta} \mid \boldsymbol{x}^*, \boldsymbol{\sigma}^*) &= c^* \prod_{i=1}^{d_1} f_0(x_i^*) \prod_{j=d_1+1}^{n} \Big[\mathbb{1}_{\{0\}}(\sigma_j^*)\, f_0(x_j^*) + \mathbb{1}_{\{1\}}(\sigma_j^*)\, f_1(x_j^* + v_1 - \tau_1)\Big] \\
&= \frac{c^*}{\vartheta_0^{d_1}} \exp\left\{-\frac{1}{\vartheta_0} \sum_{i=1}^{d_1} x_i^*\right\} \exp\left\{ \sum_{j=d_1+1}^{n} \mathbb{1}_{\{0\}}(\sigma_j^*) \left[-\frac{x_j^*}{\vartheta_0} - \log(\vartheta_0)\right]\right\} \\
&\quad \times \exp\left\{ \sum_{j=d_1+1}^{n} \mathbb{1}_{\{1\}}(\sigma_j^*) \left[-\frac{1}{\vartheta_1}\left(x_j^* + \tau_1 \frac{\vartheta_1}{\vartheta_0} - \tau_1\right) - \log(\vartheta_1)\right]\right\},
\end{aligned}$$

where $\sum_{j=d_1+1}^{n} \sigma_j^* = r_1^\star$. Notice that this simplification is possible since $\sigma_j^* \in \{0, 1\}$ for $k = 2$. This yields the log-likelihood function for the complete data $(\boldsymbol{X}^*, \boldsymbol{\Sigma}^*) = (\boldsymbol{x}^*, \boldsymbol{\sigma}^*)$ given by

$$\begin{aligned}
\ell(\boldsymbol{\theta} \mid \boldsymbol{x}^*, \boldsymbol{\sigma}^*) = {} & \log(c^*) - d_1 \log(\vartheta_0) - \frac{1}{\vartheta_0} \sum_{i=1}^{d_1} x_i^* + \sum_{j=d_1+1}^{n} \mathbb{1}_{\{0\}}(\sigma_j^*) \left[-\frac{x_j^*}{\vartheta_0} - \log(\vartheta_0)\right] \\
& + \sum_{j=d_1+1}^{n} \mathbb{1}_{\{1\}}(\sigma_j^*) \left[-\frac{1}{\vartheta_1}\left(x_j^* + \tau_1 \frac{\vartheta_1}{\vartheta_0} - \tau_1\right) - \log(\vartheta_1)\right]. \qquad (3.6)
\end{aligned}$$

In order to perform the **E-step** of the EM-algorithm, we have to calculate the expectation of $\ell(\boldsymbol{\theta} \mid \boldsymbol{X}^*, \boldsymbol{\Sigma}^*)$ w.r.t. $P_{\boldsymbol{\theta}^{(t)}}^{\boldsymbol{\Sigma}^* \mid \boldsymbol{X}^* = \boldsymbol{x}^*}$ for the current estimate $\boldsymbol{\theta}^{(t)}$. Thus, we get with (3.6)

$$Q\left(\boldsymbol{\theta}\,; \boldsymbol{\theta}^{(t)}\right) = E_{\boldsymbol{\theta}^{(t)}}\Big[\ell\left(\boldsymbol{\theta} \mid \boldsymbol{X}^*, \boldsymbol{\Sigma}^*\right) \Big| \boldsymbol{X}^* = \boldsymbol{x}^*\Big]$$

$$= \log(c^*) - d_1 \log(\vartheta_0) - \frac{1}{\vartheta_0} \sum_{i=1}^{d_1} x_i^* + \sum_{j=d_1+1}^{n} P_{0,j}^{(t)} \Big[-\frac{x_j^*}{\vartheta_0} - \log(\vartheta_0) \Big]$$
$$+ \sum_{j=d_1+1}^{n} P_{1,j}^{(t)} \Big[-\frac{1}{\vartheta_1} \Big(x_j^* + \tau_1 \frac{\vartheta_1}{\vartheta_0} - \tau_1 \Big) - \log(\vartheta_1) \Big],$$

where

$$P_{s,j}^{(t)} = E_{\boldsymbol{\theta}^{(t)}} \Big[\mathbb{1}_{\{s\}}(\Sigma_j^*) \,\Big|\, \boldsymbol{X}^* = \boldsymbol{x}^* \Big] = P_{\boldsymbol{\theta}^{(t)}} \Big(\Sigma_j^* = s \,\Big|\, \boldsymbol{X}^* = \boldsymbol{x}^* \Big) = \frac{f_{\boldsymbol{\theta}^{(t)}}^{\Sigma_j^*, \boldsymbol{X}^*}(s, \boldsymbol{x}^*)}{f_{\boldsymbol{\theta}^{(t)}}^{\boldsymbol{X}^*}(\boldsymbol{x}^*)}, \quad s \in \{0, 1\},$$

$j \in \{d_1 + 1, \dots, n\}$, does not depend on $\boldsymbol{\theta}$. To proceed further, we introduce the sets

$$\mathfrak{S}_n = \Big\{ \boldsymbol{\sigma}^* \in \{0,1\}^n \,\Big|\, \sigma_1^* = 0, \dots, \sigma_{d_1}^* = 0, \sum_{j=d_1+1}^{n} \sigma_j^* = r_1^\star \Big\},$$
$$\mathfrak{S}_n^{j,0} = \mathfrak{S}_n \cap \Big\{ \sigma_j^* = 0 \Big\} \quad \text{and} \quad \mathfrak{S}_n^{j,1} = \mathfrak{S}_n \cap \Big\{ \sigma_j^* = 1 \Big\}, \quad j \in \{d_1 + 1, \dots, n\}.$$

Hence, we have

$$f_{\boldsymbol{\theta}^{(t)}}^{\Sigma_j^*, \boldsymbol{X}^*}(0, \boldsymbol{x}^*) = \frac{1}{\vartheta_0^{(t)}} \exp\Big\{ -\frac{x_j^*}{\vartheta_0^{(t)}} \Big\} \prod_{h=1}^{d_1} \frac{1}{\vartheta_0^{(t)}} \exp\Big\{ -\frac{x_h^*}{\vartheta_0^{(t)}} \Big\}$$
$$\times \sum_{\boldsymbol{\sigma}_n^* \in \mathfrak{S}_n^{j,0}} \Bigg(\prod_{\substack{i=d_1+1 \\ i \neq j}}^{n} f_{\boldsymbol{\theta}^{(t)}}^{0,1}(x_i^*, \sigma_i^*) \Bigg), \tag{3.7a}$$

$$f_{\boldsymbol{\theta}^{(t)}}^{\Sigma_j^*, \boldsymbol{X}^*}(1, \boldsymbol{x}^*) = \frac{1}{\vartheta_1^{(t)}} \exp\Big\{ -\frac{1}{\vartheta_1^{(t)}} \Big(x_j^* + \tau_1 \frac{\vartheta_1^{(t)}}{\vartheta_0^{(t)}} - \tau_1 \Big) \Big\} \prod_{h=1}^{d_1} \frac{1}{\vartheta_0^{(t)}} \exp\Big\{ -\frac{x_h^*}{\vartheta_0^{(t)}} \Big\}$$
$$\times \sum_{\boldsymbol{\sigma}_n^* \in \mathfrak{S}_n^{j,1}} \Bigg(\prod_{\substack{i=d_1+1 \\ i \neq j}}^{n} f_{\boldsymbol{\theta}^{(t)}}^{0,1}(x_i^*, \sigma_i^*) \Bigg), \tag{3.7b}$$

$j \in \{d_1 + 1, \dots, n\}$, and

$$f_{\boldsymbol{\theta}^{(t)}}^{\boldsymbol{X}^*}(\boldsymbol{x}^*) = \prod_{h=1}^{d_1} \frac{1}{\vartheta_0^{(t)}} \exp\Big\{ -\frac{x_h^*}{\vartheta_0^{(t)}} \Big\} \sum_{\boldsymbol{\sigma}_n^* \in \mathfrak{S}_n} \Bigg(\prod_{i=d_1+1}^{n} f_{\boldsymbol{\theta}^{(t)}}^{0,1}(x_i^*, \sigma_i^*) \Bigg), \tag{3.8}$$

with

$$f^{0,1}_{\boldsymbol{\theta}^{(t)}}(x_i^*, \sigma_i^*) = \mathbb{1}_{\{0\}}(\sigma_i^*)\,\frac{1}{\vartheta_0^{(t)}}\exp\Big\{-\frac{x_i^*}{\vartheta_0^{(t)}}\Big\} + \mathbb{1}_{\{1\}}(\sigma_i^*)\,\frac{1}{\vartheta_1^{(t)}}\exp\Big\{-\frac{1}{\vartheta_1^{(t)}}\Big(x_i^* + \tau_1\,\frac{\vartheta_1^{(t)}}{\vartheta_0^{(t)}} - \tau_1\Big)\Big\}.$$

For the **M-step**, we have to maximize $Q\big(\boldsymbol{\theta}\,;\boldsymbol{\theta}^{(t)}\big)$ w.r.t. $\boldsymbol{\theta}$. First, we have

$$Q\big(\boldsymbol{\theta}\,;\boldsymbol{\theta}^{(t)}\big) = \widetilde{Q}_0(\vartheta_0) + \widetilde{Q}_1(\vartheta_1),$$

where

$$\widetilde{Q}_0(\vartheta_0) = \log(c^*) - \big(d_1 + d_2^{(t)}\big)\log(\vartheta_0) - \frac{1}{\vartheta_0}\sum_{i=1}^{d_1} x_i^* - \frac{1}{\vartheta_0}\sum_{j=d_1+1}^{n} P_{0,j}^{(t)}\,x_j^* - \frac{r_1^{\star\,(t)}\tau_1}{\vartheta_0},$$

$$\widetilde{Q}_1(\vartheta_1) = -\,r_1^{\star\,(t)}\log(\vartheta_1) - \frac{1}{\vartheta_1}\sum_{j=d_1+1}^{n} P_{1,j}^{(t)}\,x_j^* + \frac{r_1^{\star\,(t)}\tau_1}{\vartheta_1},$$

with

$$d_2^{(t)} = \sum_{j=d_1+1}^{n} P_{0,j}^{(t)} = E_{\boldsymbol{\theta}^{(t)}}\Big[\sum_{j=d_1+1}^{n} \mathbb{1}_{\{0\}}\big(\Sigma_j^*\big)\,\Big|\,\boldsymbol{X}^* = \boldsymbol{x}^*\Big] = n - d_1 - r_1^\star \quad\text{and}\quad r_1^{\star\,(t)} = \sum_{j=d_1+1}^{n} P_{1,j}^{(t)} = r_1^\star.$$

Thus, the updated estimates $\vartheta_0^{(t+1)}$ and $\vartheta_1^{(t+1)}$ in the $(t+1)$th iteration step are given by

$$\vartheta_0^{(t+1)} = \frac{1}{n - r_1^\star}\Big(\sum_{i=1}^{d_1} x_i^* + \sum_{j=d_1+1}^{n} P_{0,j}^{(t)}\,x_j^* + r_1^\star\tau_1\Big) \quad\text{and}\quad \vartheta_1^{(t+1)} = \frac{1}{r_1^\star}\Big(\sum_{j=d_1+1}^{n} P_{1,j}^{(t)}\,x_j^* - r_1^\star\tau_1\Big),$$

respectively. Possible initial values for the EM-algorithm are given by

$$\vartheta_0^{(0)} = \frac{1}{d_1}\Big(\sum_{i=1}^{d_1} x_i^* + (n - d_1)\tau_1\Big) \quad\text{and}\quad \vartheta_1^{(0)} = \frac{1}{n - d_1}\Big(\sum_{j=d_1+1}^{n} x_j^* - (n - d_1)\tau_1\Big),$$

when $d_1 > 0$, and by

$$\vartheta_0^{(0)} = \vartheta_1^{(0)} = \frac{1}{n}\Big(\sum_{j=1}^{n} x_j^* - n\tau_1\Big),$$

when $d_1 = 0$.

Alternatively, the MLEs under missing stage information (IMLE) can be computed by direct maximization of $f^{\boldsymbol{X}^*_n}_{\boldsymbol{\theta}^{(t)}}$ w.r.t. $\boldsymbol{\theta}^{(t)}$ (see equations (3.2) and (3.8)). Notice that this function is the marginal density function of $\boldsymbol{X}^*$. In Section 4.2, we compare the results of both approaches showing that they lead almost to the same estimates.

3.2.2 Exponential-Weibull Case

We assume the same situation of missing information as in Section 3.2.1 but with Weibull lifetimes on stage s_0. The probability density function and the cumulative distribution function of the Weibull distribution $\mathsf{Wei}(\vartheta, \beta)$ are given by

$$f(x) = \frac{\beta}{\vartheta} x^{\beta-1} \mathrm{e}^{-x^\beta/\vartheta} \mathbb{1}_{(0,\infty)}(x), \quad F(x) = \left(1 - \mathrm{e}^{-x^\beta/\vartheta}\right) \mathbb{1}_{(0,\infty)}(x), \quad \vartheta > 0,\ \beta > 0,\ x \in \mathbb{R}.$$

We assume that the lifetimes on stage s_0 are $\mathsf{Wei}(\vartheta_0, \beta)$-distributes whereas they are $\mathsf{Exp}(\vartheta_1)$-distributed on stage s_1. Therefore, $v_1 = \tau_1^\beta \frac{\vartheta_1}{\vartheta_0}$ and, with $\boldsymbol{\theta} = (\vartheta_0, \beta, \vartheta_1)$ and $\boldsymbol{\theta}^{(t)} = \left(\vartheta_0^{(t)}, \beta^{(t)}, \vartheta_1^{(t)}\right)$, $t \in \mathbb{N}_0$, the log-likelihood function for the complete data $(\boldsymbol{X}^*, \boldsymbol{\Sigma}^*) = (\boldsymbol{x}^*, \boldsymbol{\sigma}^*)$ is given by

$$\begin{aligned}
\ell(\boldsymbol{\theta} \mid \boldsymbol{x}^*, \boldsymbol{\sigma}^*) = \log(c^*) + d_1 \log(\beta) - d_1 \log(\vartheta_0) - \frac{1}{\vartheta_0} \sum_{i=1}^{d_1} (x_i^*)^\beta + (\beta - 1) \sum_{i=1}^{d_1} \log(x_i^*) \\
+ \sum_{j=d_1+1}^{n} \mathbb{1}_{\{0\}}(\sigma_j^*) \left[-\frac{(x_j^*)^\beta}{\vartheta_0} + \log(\beta) - \log(\vartheta_0) + (\beta - 1) \log(x_j^*) \right] \\
+ \sum_{j=d_1+1}^{n} \mathbb{1}_{\{1\}}(\sigma_j^*) \left[-\frac{1}{\vartheta_1} \left(x_j^* + \tau_1^\beta \frac{\vartheta_1}{\vartheta_0} - \tau_1 \right) - \log(\vartheta_1) \right].
\end{aligned}$$

Hence, we get

$$\begin{aligned}
Q\left(\boldsymbol{\theta}\,; \boldsymbol{\theta}^{(t)}\right) = \log(c^*) + d_1 \log(\beta) - d_1 \log(\vartheta_0) - \frac{1}{\vartheta_0} \sum_{i=1}^{d_1} (x_i^*)^\beta + (\beta - 1) \sum_{i=1}^{d_1} \log(x_i^*) \\
+ \sum_{j=d_1+1}^{n} P_{\beta,0,j}^{(t)} \left[-\frac{(x_j^*)^\beta}{\vartheta_0} + \log(\beta) - \log(\vartheta_0) + (\beta - 1) \log(x_j^*) \right]
\end{aligned}$$

$$+\sum_{j=d_1+1}^{n} P_{\beta,1,j}^{(t)} \left[-\frac{1}{\vartheta_1}\Big(x_j^* + \tau_1^{\beta}\frac{\vartheta_1}{\vartheta_0} - \tau_1 \Big) - \log(\vartheta_1) \right],$$

where the weights

$$P_{\beta,s,j}^{(t)} = \frac{f_{\boldsymbol{\theta}^{(t)}}^{\Sigma_j^*, \boldsymbol{X}^*}(s, \boldsymbol{x}^*)}{f_{\boldsymbol{\theta}^{(t)}}^{\boldsymbol{X}^*}(\boldsymbol{x}^*)}, \quad s \in \{0,1\}, \quad j \in \{d_1+1, \ldots, n\},$$

do not depend on $\boldsymbol{\theta}$. Further, with obvious changes, the functions $f_{\boldsymbol{\theta}^{(t)}}^{\Sigma_j^*, \boldsymbol{X}^*}$ and $f_{\boldsymbol{\theta}^{(t)}}^{\boldsymbol{X}^*}$ are as defined in (3.7) and (3.8), respectively. Moreover, we have to maximize $Q\big(\boldsymbol{\theta}\,;\boldsymbol{\theta}^{(t)}\big)$ w.r.t. $\boldsymbol{\theta}$. First, we have

$$Q\big(\boldsymbol{\theta}\,;\boldsymbol{\theta}^{(t)}\big) = \widetilde{Q}_0(\vartheta_0, \beta) + \widetilde{Q}_1(\vartheta_1),$$

where

$$\begin{aligned}
\widetilde{Q}_0(\vartheta_0, \beta) = &\ \log(c_1) + (n - r_1^{\star}) \log(\beta) \\
&+ (\beta - 1) \sum_{i=1}^{d_1} \log(x_i^*) + (\beta - 1) \sum_{j=d_1+1}^{n} P_{\beta,0,j}^{(t)} \log(x_j^*) \\
&- (n - r_1^{\star}) \log(\vartheta_0) - \frac{1}{\vartheta_0} \sum_{i=1}^{d_1} (x_i^*)^{\beta} - \frac{1}{\vartheta_0} \sum_{j=d_1+1}^{n} P_{\beta,0,j}^{(t)} (x_j^*)^{\beta} - \frac{r_1^{\star}\tau_1^{\beta}}{\vartheta_0}, \\
\widetilde{Q}_1(\vartheta_1) = &\ - r_1^{\star} \log(\vartheta_1) - \frac{1}{\vartheta_1} \sum_{j=d_1+1}^{n} P_{\beta,1,j}^{(t)} x_j^* + \frac{r_1^{\star}\tau_1}{\vartheta_1}.
\end{aligned}$$

The updated estimates of ϑ_0 and ϑ_1 in the $(t+1)$th iteration step are given by

$$\vartheta_0^{(t+1)}\big(\beta^{(t+1)}\big) = \frac{1}{n - r_1^{\star}} \underbrace{\Bigg(\sum_{i=1}^{d_1} (x_i^*)^{\beta^{(t+1)}} + \sum_{j=d_1+1}^{n} P_{\beta,0,j}^{(t)} (x_j^*)^{\beta^{(t+1)}} + r_1^{\star}\tau_1^{\beta^{(t+1)}} \Bigg)}_{=A_{\beta^{(t+1)}},\ \text{say}}, \tag{3.9}$$

and

$$\vartheta_1^{(t+1)} = \frac{1}{r_1^{\star}} \Bigg(\sum_{j=d_1+1}^{n} P_{\beta,1,j}^{(t)} x_j^* - r_1^{\star}\tau_1 \Bigg),$$

respectively. The partial derivative of $Q\big(\boldsymbol{\theta}\,;\boldsymbol{\theta}^{(t)}\big)$ w.r.t. β is given by

$$\frac{\partial Q\big(\boldsymbol{\theta}\,;\boldsymbol{\theta}^{(t)}\big)}{\partial\beta} = \frac{d_1}{\beta} - \frac{1}{\vartheta_0}\underbrace{\sum_{i=1}^{d_1}(x_i^*)^\beta\log(x_i^*)}_{=B_\beta,\text{ say,}} + \sum_{i=1}^{d_1}\log(x_i^*) - \frac{1}{\vartheta_0}r_1^\star\tau_1^\beta\log(\tau_1)$$
$$- \frac{1}{\vartheta_0}\underbrace{\sum_{j=d_1+1}^{n}P_{\beta,0,j}^{(t)}\Big[(x_j^*)^\beta\log(x_j^*)\Big]}_{=C_\beta,\text{ say,}} + \sum_{j=d_1+1}^{n}P_{\beta,0,j}^{(t)}\Big[\frac{1}{\beta}+\log(x_j^*)\Big] = 0.$$

Equation (3.9) plugged into the above equation yields

$$\frac{1}{\beta^{(t+1)}} - \frac{B_{\beta^{(t+1)}} + C_{\beta^{(t+1)}} + r_1^\star\tau_1^{\beta^{(t+1)}}\log(\tau_1)}{A_{\beta^{(t+1)}}} + \frac{1}{n-r_1^\star}\left(\sum_{i=1}^{d_1}\log(x_i^*) + \sum_{j=d_1+1}^{n}P_{\beta,0,j}^{(t)}\log(x_j^*)\right) = 0. \tag{3.10}$$

Therefore, we get the update variable $\beta^{(t+1)}$ of β for the $(t+1)$th iteration step by solving equation (3.10) numerically for $\beta^{(t+1)}$. As shown in Section A.1, equation (3.10) has a unique solution for $\beta^{(t+1)} > 0$. Possible initial values for the EM-algorithm are given by

$$\vartheta_0^{(0)}\big(\beta^{(0)}\big) = \frac{1}{d_1}\left(\sum_{i=1}^{d_1}(x_i^*)^{\beta^{(0)}} + (n-d_1)\tau_1^{\beta^{(0)}}\right) \quad\text{and}\quad \vartheta_1^{(0)} = \frac{1}{n-d_1}\left(\sum_{j=d_1+1}^{n}x_j^* - (n-d_1)\tau_1\right),$$

when $d_1 > 0$, where $\beta^{(0)}$ is the (unique) numerical solution of

$$\frac{1}{\beta^{(0)}} - \frac{(n-d_1)\tau_1^{\beta^{(0)}}\log(\tau_1) + \sum_{i=1}^{d_1}(x_i^*)^{\beta^{(0)}}\log(x_i^*)}{(n-d_1)\tau_1^{\beta^{(0)}} + \sum_{i=1}^{d_1}(x_i^*)^{\beta^{(0)}}} + \frac{1}{d_1}\sum_{i=1}^{d_1}\log(x_i^*) = 0.$$

For $d_1 = 0$, one may choose the initial values

$$\beta^{(0)} = 1 \quad\text{and}\quad \vartheta_0^{(0)} = \vartheta_1^{(0)} = \frac{1}{n}\left(\sum_{j=1}^{n}x_j^* - n\tau_1\right).$$

3.3 EM-Algorithm for Three-Step SLT

For $k > 2$, the situation is more involved. In order to illustrate the additional difficulties caused by more stages, we discuss the case of two fixed stage-change times, i.e., the case $k = 3$. An extension to more stages may be developed in the same manner. In particular, the EM-algorithm proposed in Section 3.2.1 is extended to a second stage-change time. As above, we assume that the design of the life test is still known, that is, $n, \tau_1, \tau_2, \mathscr{P}_* = (\pi_1, \pi_2)$, and $\mathscr{R}^0_* = (R^0_1, R^0_2)$ are given. The sample is given by the ordered failure times $x^*_1 \leq \cdots \leq x^*_n$ only. Since the estimators for ϑ_1 and ϑ_2 do not exist if no failures have been observed, our calculations are conditional on $R^\star_1 > 0$ and $R^\star_2 > 0$. For $R^\star_1 = 0$ and/or $R^\star_2 = 0$, the following EM-algorithm can be used with the necessary adjustments.

Using the sample $x^*_1, \ldots, x^*_n$ and the design of the life test, we have

$$d_1 = \sum_{i=1}^{n} \mathbb{1}_{(-\infty,\tau_1]}(x^*_i) \quad \text{and} \quad r^\star_1 = \varrho_1(d_1) = \begin{cases} \lfloor \pi_1 \cdot (n - d_1) \rfloor, & \text{Type-P,} \\ \min\big\{n - d_1, R^0_1\big\}, & \text{Type-M.} \end{cases}$$

Further, we introduce the following counters:

- ⋄ Number of observations in the interval $(\tau_1, \tau_2]$: $b_1 = \sum_{i=1}^{n} \mathbb{1}_{(\tau_1,\tau_2]}(x^*_i)$;
- ⋄ Number of observations in the interval (τ_2, ∞): $b_2 = \sum_{i=1}^{n} \mathbb{1}_{(\tau_2,\infty)}(x^*_i)$;
- ⋄ Number of observations on stage s_1 in the interval (τ_2, ∞): $r_1 = r^\star_1 + d_2 - b_1$.

Note that we can generally not determine d_2 from the available information. Therefore, we have to consider all possible values $d_2 \in \{\max\{b_1 - r^\star_1, 0\}, \ldots, b_1\}$ in our calculations so that

$$r^\star_2 = \varrho_2(\boldsymbol{d}_2) = \left\{ \begin{array}{ll} \lfloor \pi_2 \cdot (n - d_1 - d_2 - r^\star_1) \rfloor, & \text{Type-P} \\ \min\big\{n - d_1 - d_2 - r^\star_1, R^0_2\big\}, & \text{Type-M} \end{array} \right\} > 0.$$

The situation is illustrated in Figure 2. Note that the failure times $x^*_{d_1+1}, \ldots, x^*_{d_1+b_1}$ could have been observed on stages s_0 or s_1, whereas the failure times $x^*_{d_1+b_1+1}, \ldots, x^*_n$ could have been observed on stages s_0, s_1, or s_2. Thus, in the intervals $(\tau_1, \tau_2]$ and (τ_2, ∞), we have two and three options for each observation to allocate the observed data, respectively.

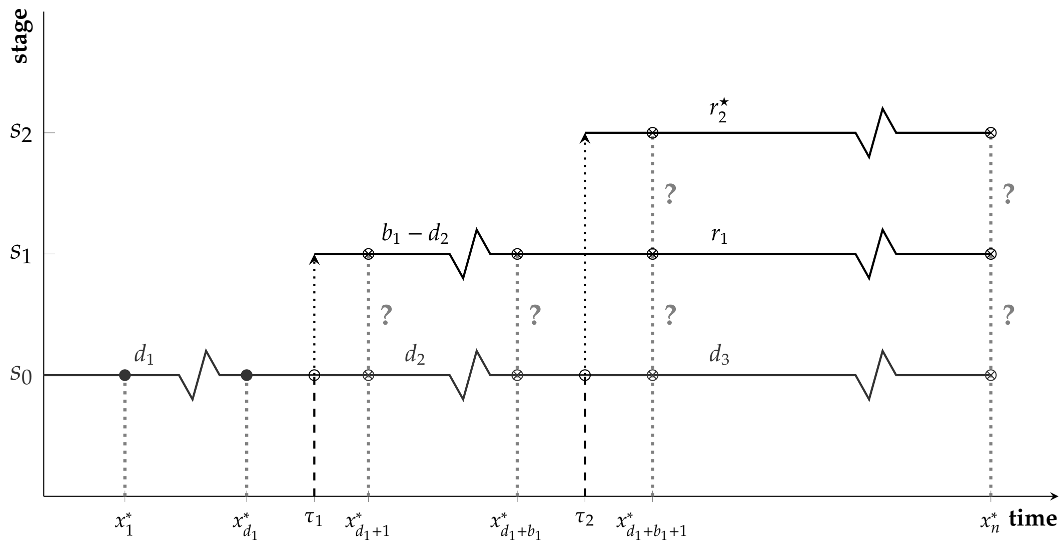

Figure 3: 3-step SLTOSs with missing information. The observations $x^*_{d_1+1}, \dots, x^*_{d_1+b_1}$ could have been observed on stages s_0 or s_1. The observations $x^*_{d_1+b_1+1}, \dots, x^*_n$ could have been observed on stages s_0, s_1, or s_2.

Additional and supplementary computations as well as the corresponding notation are presented in the appendix. Applying an EM-procedure, these computations yield the updated estimates in the $(t+1)$th iteration step given by

$$\vartheta_0^{(t+1)} = \frac{1}{d_1 + d_2^{(t)} + d_3^{(t)}} \times \left(\sum_{a=1}^{d_1} x_a^* + \sum_{h=d_1+1}^{d_1+b_1} P_{2,0,h}^{(t)} x_h^* + \sum_{i=d_1+b_1+1}^{n} P_{3,0,i}^{(t)} x_i^* + \left(r_1^{\star\,(t)} + r_1^{(t)} \right) \tau_1 + r_2^{\star\,(t)} \tau_2 \right), \tag{3.11a}$$

$$\vartheta_1^{(t+1)} = \frac{1}{r_1^{\star\,(t)} + r_1^{(t)}} \left(\sum_{h=d_1+1}^{d_1+b_1} P_{2,1,h}^{(t)} x_h^* + \sum_{i=d_1+b_1+1}^{n} P_{3,1,i}^{(t)} x_i^* - \left(r_1^{\star\,(t)} + r_1^{(t)} \right) \tau_1 \right), \tag{3.11b}$$

$$\vartheta_2^{(t+1)} = \frac{1}{r_2^{\star\,(t)}} \left(\sum_{i=d_1+b_1+1}^{n} P_{3,2,i}^{(t)} x_i^* - r_2^{\star\,(t)} \tau_2 \right). \tag{3.11c}$$

Possible initial values for the EM-algorithm are given by

$$\vartheta_0^{(0)} = \frac{1}{d_1}\Big(\sum_{a=1}^{d_1} x_a^* + \sum_{j=1}^{2} b_j\,\tau_j\Big), \quad \vartheta_1^{(0)} = \frac{1}{b_1}\sum_{h=d_1+1}^{d_1+b_1} x_h^*, \quad \vartheta_2^{(0)} = \frac{1}{b_2}\sum_{i=d_1+b_1+1}^{n} x_i^*,$$

when $d_1 > 0$, $b_1 > 0$ and $b_2 > 0$.

Note that during the calculations of the conditional probabilities (cf. (A.2), (A.3)) for all possible values $d_2 \in \Big\{\max\{b_1 - r_1^\star, 0\}, \dots, b_1\Big\}$ in each iteration, the following combinatorial counts result:

- ⋄ Number of possibilities to allocate the data in the interval $(\tau_1, \tau_2]$: $\binom{b_1}{b_1-d_2}$;
- ⋄ Number of possibilities to allocate the data in the interval (τ_2, ∞): $\binom{b_2}{r_2^\star} \cdot \binom{b_2-r_2^\star}{r_1}$.

As above, the MLEs under missing stage information (IMLE) may be computed by direct maximization of the likelihood $f_{\boldsymbol{\theta}^{(t)}}^{\boldsymbol{X}^*}$ w.r.t. $\boldsymbol{\theta}^{(t)}$ (see equation (A.4)).

4 Illustrative Example and Simulations

4.1 Illustrative Example

In order to illustrate the approach, we consider the data given in Laumen and Cramer (2019b), Section 5, where a cumulative exposure model with exponential distributions is discussed. The respective parameters are given by $n = 16$, $\vartheta_0 = 40$, $\vartheta_1 = 20$, and $\tau_1 = 15$. The resulting SLT order statistics for the two scenarios Type-P and Type-M are given in Tables 1 and 2, respectively. Notice that the stage information is available.

Table 1: Stage life testing sample with $p_1 = 0.5$ and $\tau_1 = 15$ (Type-P).

$s_0, \le \tau_1$	2.55	7.77	9.01	9.23	12.34	
$s_0, > \tau_1$	18.61	19.39	27.54	39.67	43.25	91.46
$s_1, > \tau_1$	24.51	30.73	36.88	55.26	65.39	

Table 2: Stage life testing sample with $R_1^0 = 8$ and $\tau_1 = 15$ (Type-M).

$s_0, \leq \tau_1$	2.55	7.77	9.01	9.23	12.34			
$s_0, > \tau_1$	39.67	43.25	95.53					
$s_1, > \tau_1$	16.81	17.19	21.27	24.51	30.73	36.88	53.23	65.39

In particular, we get for both options Type-P and Type-M the number of observations on each stage:

$$\begin{aligned} \text{Type-P:} \quad & d_1 = 5,\ d_2 = 6,\ \text{and } r_1^\star = 5, \\ \text{Type-M:} \quad & d_1 = 5,\ d_2 = 3,\ \text{and } r_1^\star = 8. \end{aligned}$$

Based on the values in Tables 1 and 2, the MLEs (with stage information) are given by $\widehat{\vartheta}_0 = 32.35, \widehat{\vartheta}_1 = 27.55$ (Type-P) and $\widehat{\vartheta}_0 = 42.42, \widehat{\vartheta}_1 = 18.25$ (Type-M), respectively. Using the initial values suggested in Section 3.2.1, we find the following values for the MLEs under missing stage information applying both the proposed EM-algorithm (EME) as well as the direct optimization (IMLE):

$$\begin{aligned} \text{Type-P:} \quad & \text{initial values: } \vartheta_0^{(0)} = 41.18 \quad \text{and} \quad \vartheta_1^{(0)} = 26.16 \\ & \widehat{\vartheta}_{0,\text{EME}} = 34.93, \quad \widehat{\vartheta}_{0,\text{IMLE}} = 34.92 \quad \text{and} \quad \widehat{\vartheta}_{1,\text{EME}} = 21.88, \quad \widehat{\vartheta}_{1,\text{IMLE}} = 21.92, \\ \text{Type-M:} \quad & \text{initial values: } \vartheta_0^{(0)} = 41.18 \quad \text{and} \quad \vartheta_1^{(0)} = 25.41 \\ & \widehat{\vartheta}_{0,\text{EME}} = 38.40, \quad \widehat{\vartheta}_{0,\text{IMLE}} = 38.48 \quad \text{and} \quad \widehat{\vartheta}_{1,\text{EME}} = 22.27, \quad \widehat{\vartheta}_{1,\text{IMLE}} = 22.25. \end{aligned}$$

It turns out that the estimates obtained by the EM-algorithm as well as by the direct optimization are quite close.

4.2 Simulation Study: Missing Stage Information

4.2.1 Exponential Distributions

In order to illustrate the EM-algorithm, we present the results of a simulation study in Table 3 for a 2-step SLT. The results are based on $N = 1000$ samples from the exponential distribution with $\vartheta_0 = 1.0$ on stage s_0 and $\vartheta_1 = 2.0$ on stage s_1 as well as sample size $n = 12$. We computed the EMEs and the IMLEs as mentioned above. We used the package "maxLik" from the statistical software R to maximize the log-likelihood function numerically. The implemented procedure for the Newton-Raphson method is called "maxNR". This procedure is always used with the default options (i.e., grad = NULL, hess = NULL, tol = 10^{-8}, steptol = 10^{-10}, iterlim = 150). In addition to the average estimates, we computed the standard deviation $SD_{\widehat{\vartheta}} = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\left(\widehat{\vartheta}^{(i)} - \overline{\widehat{\vartheta}}\right)^2}$. Further, the counter $N_1^* \le N$ denotes the number of samples where at least one failure has been observed on stage s_1, that is, $r_1^\star > 0$.

Table 3: EMEs and IMLEs ("maxNR") for $\vartheta_0 = 1.0$ and $\vartheta_1 = 2.0$ with $\tau_1 = 0.5$, $n = 12$, $N = 10^3$, and $\varepsilon = 10^{-10}$ (SLT for exponential distributions).

Model	π_1	R_1^0	EME $\overline{\widehat{\vartheta}}_0$ $(SD_{\widehat{\vartheta}_0})$	EME $\overline{\widehat{\vartheta}}_1$ $(SD_{\widehat{\vartheta}_1})$	IMLE $\overline{\widehat{\vartheta}}_0$ $(SD_{\widehat{\vartheta}_0})$	IMLE $\overline{\widehat{\vartheta}}_1$ $(SD_{\widehat{\vartheta}_1})$	$\overline{d_1}$	$\overline{d_2}$	$\overline{r_1^\star}$	N_1^*
Type-P	0.25	–	1.08088914 (0.385702)	1.66217278 (1.750974)	1.08122775 (0.386620)	1.65280462 (1.728466)	4.73	5.83	1.44	984
	0.50		1.14755232 (0.519729)	1.81071981 (1.210331)	1.14755349 (0.519732)	1.80981357 (1.208123)	4.73	3.90	3.37	1000
	0.75		1.20856106 (0.683318)	1.89482235 (0.962973)	1.20856105 (0.683318)	1.89453765 (0.961835)	4.73	2.20	5.07	1000
Type-M	–	3	1.09765251 (0.443357)	1.63414105 (1.166364)	1.09783365 (0.444094)	1.63327092 (1.167054)	4.73	4.27	3.00	1000
		6	1.16272888 (0.584143)	1.83580770 (0.878168)	1.16272892 (0.584143)	1.83580772 (0.877168)	4.73	1.50	5.77	1000
		9	1.21788872 (0.750887)	1.96632143 (0.758704)	1.21788877 (0.750887)	1.96632142 (0.758704)	4.73	0.11	7.16	1000

For a 3-step SLT, we conducted a simulation study based on $N = 1000$ samples and sample size $n = 16$. The results are presented in Table 4. The value $N_2^* = 583$ indicates that, for the first case with option Type-P, the estimate of ϑ_2 does not exist in 417 cases of the 1000 samples. We make the following observations:

- ◇ The EMEs are close to the IMLEs;
- ◇ The higher the number of observations on a stage, the closer the corresponding estimate is to the true value.
- ◇ For the 3-step SLT, the estimates for ϑ_0 and ϑ_2 are overestimating the true values,

whereas the estimates for ϑ_1 are underestimating them. This reflects the influence of the data being allocated to the "wrong" stage.

Further, note that the computation of the IMLEs with "maxNR" (package "maxLik" from the statistical software R (cf. Henningsen and Toomet, 2011)) is faster than the computation of the EMEs, since the Newton-Raphson method needs less iterations than the EM-algorithm. However, further investigations of the simulated results indicate that, for a few samples, the "maxNR" procedure generates values that are significantly different from those computed with the EM-algorithm. This observation might be an explanation for the discrepancy of the computed means in Table 4.

Table 4: EMEs and IMLEs ("maxNR") for $\vartheta_0 = 1.0$, $\vartheta_1 = 3.0$ and $\vartheta_2 = 2.0$ with $\tau_1 = 0.5$, $\tau_2 = 1.0$, $n = 16$, $N = 10^3$ and $\varepsilon = 10^{-8}$ (3-step SLT for exponential distributions).

		EME			IMLE								
Model	$\mathscr{P}_*/\mathscr{R}_*^0$	$\overline{\widehat{\vartheta}_0}$ $(SD_{\widehat{\vartheta}_0})$	$\overline{\widehat{\vartheta}_1}$ $(SD_{\widehat{\vartheta}_1})$	$\overline{\widehat{\vartheta}_2}$ $(SD_{\widehat{\vartheta}_2})$	$\overline{\widehat{\vartheta}_0}$ $(SD_{\widehat{\vartheta}_0})$	$\overline{\widehat{\vartheta}_1}$ $(SD_{\widehat{\vartheta}_1})$	$\overline{\widehat{\vartheta}_2}$ $(SD_{\widehat{\vartheta}_2})$	$\overline{d_1}$	$\overline{r_1^\star}$	$\overline{b_1}$	$\overline{b_2}$	N_1^*	N_2^*
Type-P	(0.2, 0.2)	1.225756 (0.4203)	1.145561 (1.5705)	2.460058 (3.1364)	1.226914 (0.4152)	1.089701 (1.3850)	2.504700 (3.2546)	6.28	1.57	3.48	6.25	993	583
	(0.4, 0.4)	1.468713 (0.7119)	1.375882 (1.3933)	2.977349 (2.5674)	1.475549 (0.7094)	1.361291 (1.3905)	2.937543 (2.5725)	6.28	3.50	3.04	6.69	1000	809
	(0.6, 0.6)	1.436692 (0.8099)	2.114799 (1.3725)	3.394885 (2.8515)	1.445394 (0.8096)	2.087464 (1.3574)	3.389059 (2.8564)	6.28	5.45	2.55	7.18	1000	822
Type-M	(2, 2)	1.217683 (0.4179)	1.065040 (1.3832)	2.649664 (2.1505)	1.221743 (0.4231)	1.052376 (1.3510)	2.654893 (2.1651)	6.28	2.00	3.42	6.31	1000	996
	(4, 4)	1.394729 (0.6738)	1.479444 (1.3648)	2.870876 (1.8156)	1.398108 (0.6814)	1.480421 (1.3652)	2.862943 (1.8165)	6.28	4.00	2.94	6.79	1000	973
	(6, 6)	1.291120 (0.8376)	2.355805 (1.3080)	2.954240 (2.0452)	1.291120 (0.8376)	2.355804 (1.3080)	2.954241 (2.0452)	6.28	5.97	2.43	7.29	1000	880

4.2.2 Weibull-Exponential Case

To illustrate the presented EM-algorithm, we conducted a simulation study. We applied the Newton-Raphson method by using the procedure "nleqslv" with initial value $\widetilde{\beta} = 1$ to solve equation (3.10). The results and the model parameters of a simulation study for the EM-algorithm are given in Table 5. We used the conjugate gradient method (cf.

Atkinson, 1989, pp. 562–569) for the derivation of the IMLEs by applying the procedure "maxCG" with the default options (i.e., grad = NULL, hess = NULL, tol = 10^{-8}, iterlim = 500). The outcome of the simulations leads to similar conclusions as those in Section 4.2.1 (cf. Table 3). The presented EM-algorithm works well despite the efforts caused by the numerical computation of the updates for β during each iteration.

Table 5: EMEs ("nleqslv") and IMLEs ("maxCG") for $\vartheta_0 = 2.0, \beta = 4.0$ and $\vartheta_1 = 3.0$ with $\tau_1 = 1.1$, $n = 24$, $N = 10^3$, and $\varepsilon = 10^{-8}$ (SLT for Weibull and exponential distribution).

Model	π_1	R_1^0	EME $\overline{\widehat{\vartheta}_0}$ $(SD_{\widehat{\vartheta}_0})$	EME $\overline{\widehat{\beta}}$ $(SD_{\widehat{\beta}})$	EME $\overline{\widehat{\vartheta}_1}$ $(SD_{\widehat{\vartheta}_1})$	IMLE $\overline{\widehat{\vartheta}_0}$ $(SD_{\widehat{\vartheta}_0})$	IMLE $\overline{\widehat{\beta}}$ $(SD_{\widehat{\beta}})$	IMLE $\overline{\widehat{\vartheta}_1}$ $(SD_{\widehat{\vartheta}_1})$	$\overline{d_1}$	$\overline{d_2}$	$\overline{r_1^\star}$	N_1^*
Type-P	0.25	–	2.161087 (0.6731)	4.240794 (0.8078)	2.960759 (1.9616)	2.161612 (0.6736)	4.240911 (0.8075)	2.952323 (1.9297)	12.55	8.96	2.49	999
	0.50		2.159399 (0.7018)	4.276169 (0.9315)	2.954151 (1.2849)	2.160377 (0.7028)	4.274265 (0.9353)	2.952445 (1.2865)	12.55	5.99	5.46	1000
	0.75		2.162283 (0.6987)	4.315022 (1.1030)	2.986807 (1.0456)	2.162283 (0.6987)	4.315021 (1.1030)	2.986805 (1.0456)	12.55	3.25	8.20	1000
Type-M	–	6	2.150220 (0.6969)	4.340566 (0.9774)	2.955920 (1.2063)	2.150220 (0.6969)	4.340566 (0.9774)	2.955919 (1.2063)	12.55	5.46	5.99	1000
		12	2.144948 (0.6889)	4.390657 (1.2480)	2.999659 (0.9331)	2.144948 (0.6889)	4.390657 (1.2480)	3.002483 (0.9322)	12.55	0.69	10.76	1000
		18	2.157151 (0.7056)	4.362316 (1.2631)	3.004980 (0.9014)	2.157151 (0.7056)	4.362316 (1.2631)	3.004980 (0.9014)	12.55	0.00	11.45	1000

5 Conclusion and Outlook

We have presented an EM-algorithm approach under SLT to compute estimates for the distribution parameters when the stage information of the failures is not available. Furthermore, we have illustrated our approach by a data set as well as by some simulations for a 2- and 3-step SLT, respectively. In particular, an extension of the presented EM-algorithm to three or more fixed stage-change times results in an increase of relevant counting variables (cf. b_1, b_2, and r_1). Therefore, the set $\mathfrak{S}_n$ representing all possible outcomes of $\mathbf{\Sigma}^*$ becomes significantly more complex. Furthermore, additional possibilities for allocating the data to the "right" stage have to be taken into account. Therefore, an implementation of the EM-algorithm for more than two stage-change times is generally possible but leads to very complex expressions and has to incorporate combinatorial identities. In particular, the number of possible combinations will grow fast. Thus, a more efficient implementation of the algorithm should be subject of future research. Additionally, further lifetime distributions can be discussed. For instance, the case of Weibull distributions on each stage is under investigation. So far, it turns out that the computational complexity grows considerably leading to instabilities in the computations. Thus, further research is also necessary in this direction.

As mentioned in the introduction, stage life testing can be considered as an extension of simple step stress testing. In this regard, it seems to be possible that a similar approach can also be applied to simple step stress assuming that the stress changing time τ_1 may be not known for some reason. Therefore, it would be interesting to see whether the ideas of the present paper can be utilized in such a model.

Finally, it has to be mentioned that the SLT model discussed in the present paper can be subject to additional censoring. In particular, the present setting can be extended to a Type-I censoring of the data, that is, an additional termination time $\tau_k > \tau_{k-1}$ can be introduced where the SLT experiment is stopped. Of course, one may choose an overall stopping time which is applied to any stage. But, one may also consider a stage dependent termination time. Such models will also be subject of future research.

References

Atkinson, K. E. (1989). *An introduction to numerical analysis*. John Wiley & Sons, Inc., 2nd edition.

Bairamov, I., and Eryılmaz, S. (2006). Spacings, exceedances and concomitants in progressive Type II censoring scheme. *J. Statist. Plann. Inference*, **136**(3), 527-536.

Balakrishnan, N. (2009). A synthesis of exact inferential results for exponential step-stress models and associated optimal accelerated life-tests. *Metrika*, **69**, 351-396.

Balakrishnan, N., and Aggarwala, R. (2000). *Progressive censoring: theory, methods, and applications*. Birkhäuser, Boston.

Balakrishnan, N., and Cramer, E. (2014). *The Art of Progressive Censoring: Applications to Reliability and Quality*. Statistics for Industry and Technology. Birkhäuser, New York.

Balakrishnan, N., and Cramer, E. (2021). *Progressive censoring methodology: A review. In*

Pham, H., editor, Springer Handbook of Engineering Statistics, Springer, New York, 2. edition, to appear.

Balakrishnan, N., Han, D., and Iliopoulos, G. (2011). Exact inference for progressively Type-I censored exponential failure data. *Metrika*, **73**(3), 335-358.

Balakrishnan, N., and Kateri, M. (2008). On the maximum likelihood estimation of parameters of Weibull distribution based on complete and censored data. *Statistics & Probability Letters*, **78**(17), 2971-2975.

Cramer, E. (2017). Progressive Censoring Schemes. In *Wiley StatsRef: Statistics Reference Online*, Hoboken, NJ. John Wiley & Sons, Ltd.

David, H. A., and Nagaraja, H. N. (1998). Concomitants of order statistics. In Balakrishnan, N. and Rao, C. R., editors, *Handbook of Statistics: Order Statistics: Theory & Methods*, **16**, 487-513. Elsevier, Amsterdam.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, **39**(1), 1–38.

Henningsen, A., and Toomet, O. (2011). maxLik: A package for maximum likelihood estimation in R. *Computational Statistics*, **26**(3), 443-458.

Izadi, M., and Khaledi, B.-E. (2007). Progressive Type II censored order statistics and their concomitants: some stochastic comparisons results. *J. Iran. Stat. Soc. (JIRSS)*, **6**, 111-124.

Kundu, D., and Ganguly, A. (2017). *Analysis of Step-Stress Models*. Academic Press Inc., London, UK.

Laumen, B. (2017). *Progressive Censoring and Stage Life Testing*. PhD thesis, RWTH Aachen University.

Laumen, B., and Cramer, E. (2019a). Progressive censoring with fixed censoring times. *Statistics*, **53**, 569-600.

Laumen, B., and Cramer, E. (2019b). Stage life testing. *Nav. Res. Logistics*, **53**, 632-647.

Laumen, B., and Cramer, E. (2021a). *k*-step stage life testing. *Statist. Neerlandica*, **75**, 203-223.

Laumen, B., and Cramer, E. (2021b). Stage life testing with random stage changing times. *Commun. Statist. Theory Meth.*, DOI:10.1080/03610926.2020.1805764, to appear.

A Appendix

A.1 Existence of a Unique Solution for Equation (3.10)

We prove the existence of a unique solution of equation (3.10) for $\beta^{(t+1)} > 0$ by analogy with Balakrishnan and Kateri (2008). First, we define the (continuous) function H for $\alpha > 0$ (cf. (3.10))

$$H(\alpha\,;\boldsymbol{x}^*) = \frac{\sum_{i=1}^{d_1}(x_i^*)^\alpha \log(x_i^*) + \sum_{j=d_1+1}^{n} P_{\beta,0,j}^{(t)}(x_j^*)^\alpha \log(x_j^*) + r_1^\star \tau_1^\alpha \log(\tau_1)}{r_1^\star \tau_1^\alpha + \sum_{i=1}^{d_1}(x_i^*)^\alpha + \sum_{j=d_1+1}^{n} P_{\beta,0,j}^{(t)}\,(x_j^*)^\alpha} - \frac{1}{n-r_1^\star}\left(\sum_{i=1}^{d_1}\log(x_i^*) + \sum_{j=d_1+1}^{n} P_{\beta,0,j}^{(t)}\,\log(x_j^*)\right).$$

Now, we can write equation (3.10) as

$$\alpha^{-1} = H(\alpha\,;\boldsymbol{x}^*).$$

We know that α^{-1} is a monotone decreasing function in α with a limit $+\infty$ at 0 and a limit 0 at $+\infty$.

In the following, we show that H is a monotone increasing function in α with a positive upper limit at $+\infty$. This ensures the existence and uniqueness of $\beta^{(t+1)}$. In order to verify the monotonicity of H, we have to ensure that

$$\frac{\partial H(\alpha\,;\boldsymbol{x}^*)}{\partial \alpha} = \frac{h(\alpha\,;\boldsymbol{x}^*)}{\left(r_1^\star \tau_1^\alpha + \sum_{i=1}^{d_1}(x_i^*)^\alpha + \sum_{j=d_1+1}^{n} P_{\beta,0,j}^{(t)}\,(x_j^*)^\alpha\right)^2} \geq 0,$$

or, equivalently, that

$$\begin{aligned} h(\alpha\,;\boldsymbol{x}^*) = &\left(\sum_{i=1}^{d_1}(x_i^*)^\alpha + \sum_{j=d_1+1}^{n} P_{\beta,0,j}^{(t)}\,(x_j^*)^\alpha + r_1^\star \tau_1^\alpha\right) \\ &\times\left(\sum_{i=1}^{d_1}(x_i^*)^\alpha \log^2(x_i^*) + \sum_{j=d_1+1}^{n} P_{\beta,0,j}^{(t)}(x_j^*)^\alpha \log^2(x_j^*) + r_1^\star \tau_1^\alpha \log^2(\tau_1)\right) \\ &- \left(\sum_{i=1}^{d_1}(x_i^*)^\alpha \log(x_i^*) + \sum_{j=d_1+1}^{n} P_{\beta,0,j}^{(t)}(x_j^*)^\alpha \log(x_j^*) + r_1^\star \tau_1^\alpha \log(\tau_1)\right)^2 \geq 0. \end{aligned} \quad \text{(A.1)}$$

Let $\boldsymbol{a}_{n+1} = (a_1, \ldots, a_{n+1})$ and $\boldsymbol{b}_{n+1} = (b_1, \ldots, b_{n+1})$ with

$$a_i = (x_i^*)^{\alpha/2}, \quad i = 1, \ldots, d_1, \quad a_j = \left(P_{\beta,0,j}^{(t)}\right)^{1/2} (x_j^*)^{\alpha/2}, \quad j = d_1 + 1, \ldots, n, \quad a_{n+1} = (r_1^\star)^{1/2} \tau_1^{\alpha/2},$$

and

$$b_i = a_i \log(x_i^*), \quad i = 1, \ldots, d_1, \quad b_j = a_j \log(x_j^*), \quad j = d_1 + 1, \ldots, n, \quad b_{n+1} = a_{n+1} \log(\tau_1).$$

Then, equation (A.1) reads

$$h(\alpha\,; \boldsymbol{x}^*) = \sum_{i=1}^{n+1} a_i^2 \sum_{i=1}^{n+1} b_i^2 - \left(\sum_{i=1}^{n+1} a_i\, b_i\right)^2.$$

By using the Cauchy-Schwarz inequality, we conclude that $h\left(\alpha\,; \boldsymbol{x}_n^*\right) \geq 0$. Therefore, the function H is a monotone increasing function in α with an upper limit

$$\begin{aligned}
\lim_{\alpha \to +\infty} H(\alpha; \boldsymbol{x}^*) &= \lim_{\alpha \to +\infty} \frac{r_1^\star \tau_1^\alpha \log(\tau_1) + \sum_{i=1}^n a_i^2 \log(x_i^*)}{\sum_{i=1}^{n+1} a_i^2} - \frac{1}{n - r_1^\star}\left(\sum_{i=1}^{d_1} \log(x_i^*) + \sum_{j=d_1+1}^{n} P_{\beta,0,j}^{(t)} \log(x_j^*)\right) \\
&= \log(x_n^*) - \frac{1}{n - r_1^\star}\left(\sum_{i=1}^{d_1} \log(x_i^*) + \sum_{j=d_1+1}^{n} P_{\beta,0,j}^{(t)} \log(x_j^*)\right) > 0.
\end{aligned}$$

The latter expression is positive since $x_1^* \leq \cdots \leq x_n^*$ and

$$\begin{aligned}
&\frac{1}{n - r_1^\star}\left(\sum_{i=1}^{d_1} \log(x_i^*) + \sum_{j=d_1+1}^{n} P_{\beta,0,j}^{(t)} \log(x_j^*)\right) \\
&\quad < \frac{1}{n - r_1^\star}\left(d_1 \cdot \log(x_n^*) + \log(x_n^*) \cdot \sum_{j=d_1+1}^{n} P_{\beta,0,j}^{(t)}\right) \\
&\quad = \frac{1}{n - r_1^\star}\left(d_1 \cdot \log(x_n^*) + \log(x_n^*) \cdot (n - d_1 - r_1^\star)\right) = \log(x_n^*).
\end{aligned}$$

A.2 Formulas for 3-Step SLT with Missing Stage Information

In the following, we present the formulas needed to implement of the EM-algorithm addressed in Section 3.3. We consider the complete information $(\boldsymbol{X}^*, \boldsymbol{\Sigma}^*)$ to define the

EM-algorithm in the present situation. Let $\boldsymbol{\theta} = (\vartheta_0, \vartheta_1, \vartheta_2)$ and $\boldsymbol{\theta}^{(t)} = \left(\vartheta_0^{(t)}, \vartheta_1^{(t)}, \vartheta_2^{(t)}\right)$, $t \in \mathbb{N}_0$. Then, the log-likelihood function for the complete data $(\boldsymbol{X}^*, \boldsymbol{\Sigma}^*) = (\boldsymbol{x}^*, \boldsymbol{\sigma}^*)$ is given by

$$\begin{aligned}
\ell(\boldsymbol{\theta} \mid \boldsymbol{x}^*, \boldsymbol{\sigma}^*) = {} & \log(c^*) - d_1 \log(\vartheta_0) - \frac{1}{\vartheta_0} \sum_{a=1}^{d_1} x_a^* \\
& + \sum_{h=d_1+1}^{d_1+b_1} \left(\mathbb{1}_{\{0\}}(\sigma_h^*) \left[-\frac{x_h^*}{\vartheta_0} - \log(\vartheta_0) \right] + \mathbb{1}_{\{1\}}(\sigma_h^*) \left[-\frac{1}{\vartheta_1} \left(x_h^* + \tau_1 \frac{\vartheta_1}{\vartheta_0} - \tau_1 \right) - \log(\vartheta_1) \right] \right) \\
& + \sum_{i=d_1+b_1+1}^{n} \left(\mathbb{1}_{\{0\}}(\sigma_i^*) \left[-\frac{x_i^*}{\vartheta_0} - \log(\vartheta_0) \right] + \mathbb{1}_{\{1\}}(\sigma_i^*) \left[-\frac{1}{\vartheta_1} \left(x_i^* + \tau_1 \frac{\vartheta_1}{\vartheta_0} - \tau_1 \right) - \log(\vartheta_1) \right] \right. \\
& \qquad \left. + \mathbb{1}_{\{2\}}(\sigma_i^*) \left[-\frac{1}{\vartheta_2} \left(x_i^* + \tau_2 \frac{\vartheta_2}{\vartheta_0} - \tau_2 \right) - \log(\vartheta_2) \right] \right).
\end{aligned}$$

We use the following sets

$$\begin{aligned}
\mathfrak{S}_n = \Bigg\{ \boldsymbol{\sigma}^* \in \{0,1,2\}^n \,\Bigg|\, & \sigma_1^* = 0, \ldots, \sigma_{d_1}^* = 0, \\
& \sigma_a^* \in \{0,1\},\ a \in \{d_1+1, \ldots, d_1+b_1\},\ \sum_{h=d_1+1}^{d_1+b_1} \sigma_h^* = b_1 - d_2, \\
& \sigma_b^* \in \{0,1,2\},\ b \in \{d_1+b_1+1, \ldots, n\}, \\
& \sum_{i=d_1+b_1+1}^{n} \mathbb{1}_{\{0\}}(\sigma_i^*) = d_3,\ \sum_{i=d_1+b_1+1}^{n} \mathbb{1}_{\{1\}}(\sigma_i^*) = r_1,\ \sum_{i=d_1+b_1+1}^{n} \mathbb{1}_{\{2\}}(\sigma_i^*) = r_2^\star \Bigg\},
\end{aligned}$$

$$\begin{aligned}
& \mathfrak{S}_n^{h,0} = \mathfrak{S}_n \cap \left\{ \sigma_h^* = 0 \right\} \quad \text{and} \quad \mathfrak{S}_n^{h,1} = \mathfrak{S}_n \cap \left\{ \sigma_h^* = 1 \right\}, \quad h \in \{d_1+1, \ldots, d_1+b_1\}, \\
& \mathfrak{S}_n^{i,0} = \mathfrak{S}_n \cap \left\{ \sigma_i^* = 0 \right\}, \quad \mathfrak{S}_n^{i,1} = \mathfrak{S}_n \cap \left\{ \sigma_i^* = 1 \right\} \quad \text{and} \quad \mathfrak{S}_n^{i,2} = \mathfrak{S}_n \cap \left\{ \sigma_i^* = 2 \right\}, \\
& i \in \{d_1+b_1+1, \ldots, n\}.
\end{aligned}$$

Further, we have the conditional probabilities

$$P_{2,s,h}^{(t)} = E_{\boldsymbol{\theta}^{(t)}} \left[\mathbb{1}_{\{s\}}(\Sigma_h^*) \,\middle|\, \boldsymbol{X}^* = \boldsymbol{x}^* \right] = \frac{f_{\boldsymbol{\theta}^{(t)}}^{\Sigma_h^*, \boldsymbol{X}^*}(s, \boldsymbol{x}^*)}{f_{\boldsymbol{\theta}^{(t)}}^{\boldsymbol{X}^*}(\boldsymbol{x}^*)}, \quad s \in \{0,1\}, \tag{A.2}$$

$h \in \{d_1+1,\dots,d_1+b_1\}$, and

$$P_{3,s,i}^{(t)} = E_{\boldsymbol{\theta}^{(t)}}\Big[\mathbb{1}_{\{s\}}(\Sigma_i^*)\,\Big|\,\boldsymbol{X}^* = \boldsymbol{x}^*\Big] = \frac{f_{\boldsymbol{\theta}^{(t)}}^{\Sigma_i^*,\boldsymbol{X}^*}(s,\boldsymbol{x}^*)}{f_{\boldsymbol{\theta}^{(t)}}^{\boldsymbol{X}^*}(\boldsymbol{x}^*)}, \quad s \in \{0,1,2\}, \tag{A.3}$$

$i \in \{d_1+b_1+1,\dots,n\}$. Moreover, we have the density functions

$$\begin{aligned}
f_{\boldsymbol{\theta}^{(t)}}^{\Sigma_h^*,\boldsymbol{X}^*}(0,\boldsymbol{x}^*) &= \frac{1}{\vartheta_0^{(t)}}\exp\Big\{-\frac{x_h^*}{\vartheta_0^{(t)}}\Big\}\prod_{j=1}^{d_1}\frac{1}{\vartheta_0^{(t)}}\exp\Big\{-\frac{x_j^*}{\vartheta_0^{(t)}}\Big\}\\
&\quad\times\sum_{\boldsymbol{\sigma}_n^*\in\mathfrak{S}_n^{h,0}}\Bigg(\prod_{\substack{a=d_1+1\\a\neq h}}^{d_1+b_1} f_{\boldsymbol{\theta}^{(t)}}^{0,1}(x_a^*,\sigma_a^*)\prod_{b=d_1+b_1+1}^{n} f_{\boldsymbol{\theta}^{(t)}}^{0,1,2}(x_b^*,\sigma_b^*)\Bigg),\\
f_{\boldsymbol{\theta}^{(t)}}^{\Sigma_h^*,\boldsymbol{X}^*}(1,\boldsymbol{x}^*) &= \frac{1}{\vartheta_1^{(t)}}\exp\Big\{-\frac{1}{\vartheta_1^{(t)}}\Big(x_h^*+\tau_1\frac{\vartheta_1^{(t)}}{\vartheta_0^{(t)}}-\tau_1\Big)\Big\}\prod_{j=1}^{d_1}\frac{1}{\vartheta_0^{(t)}}\exp\Big\{-\frac{x_j^*}{\vartheta_0^{(t)}}\Big\}\\
&\quad\times\sum_{\boldsymbol{\sigma}_n^*\in\mathfrak{S}_n^{h,1}}\Bigg(\prod_{\substack{a=d_1+1\\a\neq h}}^{d_1+b_1} f_{\boldsymbol{\theta}^{(t)}}^{0,1}(x_a^*,\sigma_a^*)\prod_{b=d_1+b_1+1}^{n} f_{\boldsymbol{\theta}^{(t)}}^{0,1,2}(x_b^*,\sigma_b^*)\Bigg),
\end{aligned}$$

$h \in \{d_1+1,\dots,d_1+b_1\}$,

$$\begin{aligned}
f_{\boldsymbol{\theta}^{(t)}}^{\Sigma_i^*,\boldsymbol{X}^*}(0,\boldsymbol{x}^*) &= \frac{1}{\vartheta_0^{(t)}}\exp\Big\{-\frac{x_i^*}{\vartheta_0^{(t)}}\Big\}\prod_{j=1}^{d_1}\frac{1}{\vartheta_0^{(t)}}\exp\Big\{-\frac{x_j^*}{\vartheta_0^{(t)}}\Big\}\\
&\quad\times\sum_{\boldsymbol{\sigma}_n^*\in\mathfrak{S}_n^{i,0}}\Bigg(\prod_{a=d_1+1}^{d_1+b_1} f_{\boldsymbol{\theta}^{(t)}}^{0,1}(x_a^*,\sigma_a^*)\prod_{\substack{b=d_1+b_1+1\\b\neq i}}^{n} f_{\boldsymbol{\theta}^{(t)}}^{0,1,2}(x_b^*,\sigma_b^*)\Bigg),\\
f_{\boldsymbol{\theta}^{(t)}}^{\Sigma_i^*,\boldsymbol{X}^*}(1,\boldsymbol{x}^*) &= \frac{1}{\vartheta_1^{(t)}}\exp\Big\{-\frac{1}{\vartheta_1^{(t)}}\Big(x_i^*+\tau_1\frac{\vartheta_1^{(t)}}{\vartheta_0^{(t)}}-\tau_1\Big)\Big\}\prod_{j=1}^{d_1}\frac{1}{\vartheta_0^{(t)}}\exp\Big\{-\frac{x_j^*}{\vartheta_0^{(t)}}\Big\}\\
&\quad\times\sum_{\boldsymbol{\sigma}_n^*\in\mathfrak{S}_n^{i,1}}\Bigg(\prod_{a=d_1+1}^{d_1+b_1} f_{\boldsymbol{\theta}^{(t)}}^{0,1}(x_a^*,\sigma_a^*)\prod_{\substack{b=d_1+b_1+1\\b\neq i}}^{n} f_{\boldsymbol{\theta}^{(t)}}^{0,1,2}(x_b^*,\sigma_b^*)\Bigg),\\
f_{\boldsymbol{\theta}^{(t)}}^{\Sigma_i^*,\boldsymbol{X}^*}(2,\boldsymbol{x}^*) &= \frac{1}{\vartheta_2^{(t)}}\exp\Big\{-\frac{1}{\vartheta_2^{(t)}}\Big(x_i^*+\tau_2\frac{\vartheta_2^{(t)}}{\vartheta_0^{(t)}}-\tau_2\Big)\Big\}\prod_{j=1}^{d_1}\frac{1}{\vartheta_0^{(t)}}\exp\Big\{-\frac{x_j^*}{\vartheta_0^{(t)}}\Big\}
\end{aligned}$$

$$\times \sum_{\boldsymbol{\sigma}_n^* \in \mathfrak{S}_n^{i,2}} \left(\prod_{a=d_1+1}^{d_1+b_1} f_{\boldsymbol{\theta}^{(t)}}^{0,1}(x_a^*, \sigma_a^*) \prod_{\substack{b=d_1+b_1+1 \\ b \neq i}}^{n} f_{\boldsymbol{\theta}^{(t)}}^{0,1,2}(x_b^*, \sigma_b^*) \right),$$

$i \in \{d_1 + b_1 + 1, \ldots, n\}$, and

$$\begin{aligned} f_{\boldsymbol{\theta}^{(t)}}^{\boldsymbol{X}^*}(\boldsymbol{x}^*) = & \prod_{j=1}^{d_1} \frac{1}{\vartheta_0^{(t)}} \exp\left\{-\frac{x_j^*}{\vartheta_0^{(t)}}\right\} \\ & \times \sum_{\boldsymbol{\sigma}_n^* \in \mathfrak{S}_n} \left(\prod_{a=d_1+1}^{n} f_{\boldsymbol{\theta}^{(t)}}^{0,1}(x_a^*, \sigma_a^*) \prod_{b=d_1+b_1+1}^{n} f_{\boldsymbol{\theta}^{(t)}}^{0,1,2}(x_b^*, \sigma_b^*) \right), \end{aligned} \tag{A.4}$$

with

$$\begin{aligned} f_{\boldsymbol{\theta}^{(t)}}^{0,1}(x_a^*, \sigma_a^*) = & \mathbb{1}_{\{0\}}(\sigma_a^*) \frac{1}{\vartheta_0^{(t)}} \exp\left\{-\frac{x_a^*}{\vartheta_0^{(t)}}\right\} \\ & + \mathbb{1}_{\{1\}}(\sigma_a^*) \frac{1}{\vartheta_1^{(t)}} \exp\left\{-\frac{1}{\vartheta_1^{(t)}}\left(x_a^* + \tau_1 \frac{\vartheta_1^{(t)}}{\vartheta_0^{(t)}} - \tau_1\right)\right\}, \\ f_{\boldsymbol{\theta}^{(t)}}^{0,1,2}(x_b^*, \sigma_b^*) = & \mathbb{1}_{\{0\}}(\sigma_b^*) \frac{1}{\vartheta_0^{(t)}} \exp\left\{-\frac{x_b^*}{\vartheta_0^{(t)}}\right\} \\ & + \mathbb{1}_{\{1\}}(\sigma_b^*) \frac{1}{\vartheta_1^{(t)}} \exp\left\{-\frac{1}{\vartheta_1^{(t)}}\left(x_b^* + \tau_1 \frac{\vartheta_1^{(t)}}{\vartheta_0^{(t)}} - \tau_1\right)\right\} \\ & + \mathbb{1}_{\{2\}}(\sigma_b^*) \frac{1}{\vartheta_2^{(t)}} \exp\left\{-\frac{1}{\vartheta_2^{(t)}}\left(x_b^* + \tau_2 \frac{\vartheta_2^{(t)}}{\vartheta_0^{(t)}} - \tau_2\right)\right\}. \end{aligned}$$

For the **E-step**, we have to calculate the expectation of $\ell(\boldsymbol{\theta} \mid \boldsymbol{X}^*, \boldsymbol{\Sigma}^*)$ w.r.t. $P_{\boldsymbol{\theta}^{(t)}}^{\boldsymbol{\Sigma}^* \mid \boldsymbol{X}^* = \boldsymbol{x}^*}$ for the current estimates $\boldsymbol{\theta}^{(t)}$. Therefore, we have

$$Q\left(\boldsymbol{\theta}\,; \boldsymbol{\theta}^{(t)}\right) = \widetilde{Q}_0(\vartheta_0) + \widetilde{Q}_1(\vartheta_1) + \widetilde{Q}_2(\vartheta_2),$$

where

$$\widetilde{Q}_0(\vartheta_0) = \log(C) - \left(d_1 + d_2^{(t)} + d_3^{(t)}\right) \log(\vartheta_0)$$

$$
\begin{aligned}
&\quad - \frac{1}{\vartheta_0}\left(\sum_{a=1}^{d_1} x_a^* + \sum_{h=d_1+1}^{d_1+b_1} P_{2,0,h}^{(t)}\, x_h^* + \sum_{i=d_1+b_1+1}^{n} P_{3,0,i}^{(t)}\, x_i^* + \left(r_1^{\star\,(t)} + r_1^{(t)}\right)\tau_1 + r_2^{\star\,(t)}\,\tau_2 \right),\\
\widetilde{Q}_1(\vartheta_1) = &- \left(r_1^{\star\,(t)} + r_1^{(t)}\right)\log(\vartheta_1)\\
&- \frac{1}{\vartheta_1}\left(\sum_{h=d_1+1}^{d_1+b_1} P_{2,1,h}^{(t)}\, x_h^* + \sum_{i=d_1+b_1+1}^{n} P_{3,1,i}^{(t)}\, x_i^* - \left(r_1^{\star\,(t)} + r_1^{(t)}\right)\tau_1 \right),
\end{aligned}
$$

and

$$
\widetilde{Q}_2(\vartheta_2) = - r_2^{\star\,(t)} \log(\vartheta_2) - \frac{1}{\vartheta_2}\left(\sum_{i=d_1+b_1+1}^{n} P_{3,2,i}^{(t)}\, x_i^* - r_2^{\star\,(t)}\,\tau_2 \right),
$$

with

$$
d_2^{(t)} = \sum_{h=d_1+1}^{d_1+b_1} P_{2,0,h}^{(t)} \quad \text{and} \quad d_3^{(t)} = \sum_{i=d_1+b_1+1}^{n} P_{3,0,i}^{(t)},
$$

and

$$
r_1^{\star\,(t)} = \sum_{h=d_1+1}^{d_1+b_1} P_{2,1,h}^{(t)}, \quad r_1^{(t)} = \sum_{i=d_1+b_1+1}^{n} P_{3,1,i}^{(t)} \quad \text{and} \quad r_2^{\star\,(t)} = \sum_{i=d_1+b_1+1}^{n} P_{3,2,i}^{(t)}.
$$

For the **M-step**, we have to maximize $Q(\boldsymbol{\theta}\,;\boldsymbol{\theta}^{(t)})$ w.r.t. $\boldsymbol{\theta}$. By using the same arguments as in Section 3.2.1, we get the updated estimates $\vartheta_0^{(t+1)}$, $\vartheta_1^{(t+1)}$, and $\vartheta_2^{(t+1)}$ in the $(t+1)$th iteration step given in (3.11).