

Expectation–maximization algorithm for system-based lifetime data with unknown system structure

Yandan Yang¹ · Hon Keung Tony Ng¹  ·
Narayanaswamy Balakrishnan²

Received: 19 June 2017 / Accepted: 28 February 2018 / Published online: 9 March 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract In science and engineering, we are often interested in learning about the lifetime characteristics of the system as well as those of the components that made up the system. However, in many cases, the system lifetimes can be observed but not the component lifetimes, and so we may not also have any knowledge on the structure of the system. Statistical procedures for estimating the parameters of the component lifetime distribution and for identifying the system structure based on system-level lifetime data are developed here using expectation–maximization (EM) algorithm. Different implementations of the EM algorithm based on system-level or component-level likelihood functions are proposed. A special case that the system is known to be a coherent system with unknown structure is considered. The methodologies are then illustrated by considering the component lifetimes to follow a two-parameter Weibull distribution. A numerical example and a Monte Carlo simulation study are used to evaluate the performance and related merits of the proposed implementations of the EM algorithm. Lognormally distributed component lifetimes and a real data example are used to illustrate how the methodologies can be applied to other lifetime models in addition to the Weibull model. Finally, some recommendations along with concluding remarks are provided.

Keywords Latent variable · Likelihood inference · Lognormal distribution · Truncated distribution · Weibull distribution

✉ Hon Keung Tony Ng
ngh@mail.smu.edu

¹ Department of Statistical Science, Southern Methodist University, Dallas, TX 75275-0332, USA

² Department of Mathematics and Statistics, McMaster University, Hamilton, ON L8S 4K1, Canada

1 Introduction

In science and engineering, systems are made up of different components and these systems can be complex in form and structure. Reliability engineers and researchers are often interested in the lifetime distribution of the system as well as the lifetime distribution of the components that form the system. In many cases, the lifetimes of a n -component system can be observed, but not the lifetimes of the components themselves. This problem may arise when it is not possible to put the individual components on a life test after the n -component system is already built. A system with specified performance characteristics but unknown or unspecified constituents and means of operation is also known as a black-box system. For example, an integrated circuit (IC) is made up of different electrical components, such as transistors, resistors, capacitors and diodes, that are connected to each other in different ways. Manufacturers of IC are not only interested in the lifetime of the IC, but also the lifetime distributions of the electrical components that form the IC. Once the IC is built, it may not be possible to test these electronic components inside the IC individually. Therefore, in this situation, the development of statistical inference for the component reliability characteristics based on system lifetime data becomes necessary.

In recent years, parametric and nonparametric inferential methods for the lifetime distribution of components based on system lifetimes have been developed by many authors. Coherent systems (i.e., each component contributes to the functioning/failure of the system and the reliability of the system is monotone) are commonly considered in these studies. For nonparametric inference, Bhattacharya and Samaniego (2010) considered the nonparametric estimation of component lifetime distributions from system failure-time data, while Balakrishnan et al. (2011a) developed exact nonparametric confidence intervals for population quantiles as well as tolerance intervals. Al-Nefaiee and Coolen (2013) discussed nonparametric inference for system lifetimes with exchangeable components. For parametric inference, Eryilmaz et al. (2011) considered the analysis of consecutive k -out-of- n systems with exchangeable components. Ng et al. (2012) discussed the estimation of parameters in the component lifetime distribution under a proportional hazard rate model. Balakrishnan et al. (2011b) developed the best linear unbiased estimation method for the location and scale parameters of the lifetime distribution of components based on complete and censored system lifetime data. Zhang et al. (2015) discussed different estimation methods for the parameters when the component lifetime distribution is in the location-scale family of distributions or log-location-scale family of distributions. In these existing works, the structure of the systems under investigation was assumed to be known. However, in some practical situations, the system of interest is a black box, in which case there is a lack of knowledge on the system's internal structure. System identification is the subject that tries to estimate a black box based on observed experimental system lifetime data. There is a rich literature on system identification based only on data (data driven) without having previous knowledge of the system (see, for example, Eykhoff 1974; Ljung 1999; Tangirala 2015). In this paper, we are interested in developing statistical inferential procedures for estimating the parameters of the component lifetime distribution as well as for identifying the system structure based on the system-level lifetime data.

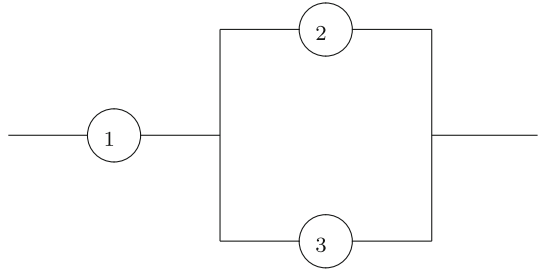
Based on the special features of the system lifetime data, we treat the system-level data as incomplete data and apply the expectation–maximization (EM) algorithm (Dempster et al. 1977; McLachlan and Krishnan 2008) to obtain the maximum likelihood estimates (MLEs). EM algorithm is an iterative method for finding the MLEs of the parameters in statistical models that involve some unobserved latent variables. It is based on the idea of replacing a difficult likelihood maximization with a sequence of easier maximization steps which will ultimately converge to the value obtainable by direct maximization of the likelihood. It is particularly useful for solving missing data problems. EM algorithm has been widely used in lifetime data analysis, for example, in the analysis of progressively censored data (Ng et al. 2002), for hybrid censored data (Dube et al. 2011), and for left- truncated and right-censored data (Balakrishnan and Mitra 2012). Based on the representations of the system structure, system signature (Samaniego 2007) and minimal signature (Navarro et al. 2007), we propose here different ways of implementing the EM algorithm and evaluate relative merits of these implementations.

The rest of this paper is organized as follows. Section 2 provides the details of the form of system lifetime data and the notion of system signature. In Sect. 3, we discuss the statistical inference based on maximum likelihood method and present the framework of the EM algorithm when considering the system lifetime distribution as a mixture distribution. Different implementations of the EM algorithm based on system-level and component-level likelihood functions are presented in Sects. 3.1–3.3. Then, in Sect. 4, we discuss the applications of the EM algorithm when the system is known to be a coherent system with an unknown system structure. In Sect. 5, we illustrate the methodologies by considering the lifetimes of the components in the system to follow a two-parameter Weibull distribution. An illustrative example is presented in Sect. 5.2, and a Monte Carlo simulation study is carried out in Sect. 5.3 to evaluate the performance of the proposed implementations of the EM algorithm when the component lifetimes follow the Weibull distribution. In Sect. 6, we consider the case when the component lifetimes follow a lognormal distribution to further illustrate how the methodologies can be applied to different lifetime models in addition to the Weibull model. A real data example of water-reservoir control system is used to verify the performance of the proposed estimation procedures. Finally, some concluding remarks are made in Sect. 7.

2 System lifetime data and system signature

Let the random variable T be the lifetime of a coherent system with independent and identically distributed (i.i.d.) component lifetimes X_1, X_2, \dots, X_n from a common absolutely continuous cumulative distribution function (c.d.f.) $F_X(\cdot)$, probability density function (p.d.f.) $f_X(\cdot)$ and survival function (s.f.) $\bar{F}_X(\cdot) = 1 - F_X(\cdot)$. We denote the order statistics corresponding to the n component lifetimes by $X_{1:n} < X_{2:n} < \dots < X_{n:n}$. We shall further denote the c.d.f., p.d.f. and s.f. of the i -th order statistic by $F_{i:n}(\cdot)$, $f_{i:n}(\cdot)$ and $\bar{F}_{i:n}(\cdot)$, and the c.d.f., p.d.f. and s.f. of the lifetime of the system T by $F_T(\cdot)$, $f_T(\cdot)$ and $\bar{F}_T(\cdot)$, respectively. Now, let us assume m independent such n -component coherent systems are placed on a life-testing experiment and

Fig. 1 A 3-component parallel-series system



their lifetimes are observed, with the corresponding ordered system lifetimes being $T_{1:m} < T_{2:m} < \dots < T_{m:m}$.

System signature (Samaniego 2007) is an index that characterizes a system without using complex structure function. The system signature of a coherent n -component system is a n -dimensional probability vector, $s = (s_1, s_2, \dots, s_n)$, defined by

$$\begin{aligned} s_i &= \Pr(\text{system fails upon the failure of the } i\text{-th component}) \\ &= \Pr(T = X_{i:n}), \end{aligned}$$

where $0 \leq s_i \leq 1$, $i = 1, 2, \dots, n$, and $\sum_{i=1}^n s_i = 1$. System signature is a distribution-free representation (Kochar et al. 1999) of the system structure meaning that s does not depend on the distribution of the component lifetimes. For example, let us consider a 3-component series-parallel system, as shown in Fig. 1. The signature for this system is $(1/3, 2/3, 0)$. Once we establish a system signature s , the distribution function of the system lifetime T can be expressed readily in terms of s and F_X alone. For a n -component system with i.i.d. components, the p.d.f. and s.f. of the system lifetime T can be expressed as

$$f_T(t|\boldsymbol{\theta}, s) = \sum_{i=1}^n s_i f_{i:n}(t|\boldsymbol{\theta}) = \sum_{i=1}^n s_i \binom{n}{i} i f_X(t|\boldsymbol{\theta}) [F_X(t|\boldsymbol{\theta})]^{i-1} [\bar{F}_X(t|\boldsymbol{\theta})]^{n-i} \quad (1)$$

and

$$\bar{F}_T(t|\boldsymbol{\theta}, s) = \sum_{i=1}^n s_i \bar{F}_{i:n}(t|\boldsymbol{\theta}) = \sum_{i=1}^n s_i \sum_{j=0}^{i-1} \binom{n}{j} [F_X(t|\boldsymbol{\theta})]^j [\bar{F}_X(t|\boldsymbol{\theta})]^{n-j}, \quad (2)$$

respectively, where $\boldsymbol{\theta}$ is the vector of model parameters. For more discussions on the system signatures and their applications, one may refer to Samaniego (2007) and the references therein.

Suppose m independent n -component systems with the same but unknown structure are placed on a life-testing experiment and the censoring time for the k -th system is τ_k . That is, the experiment is terminated at $T_k^* = \min\{T_k, \tau_k\}$ for the k -th system. Let d_k be the censoring indicator, i.e., $d_k = 1$ if $T_k < \tau_k$ and $d_k = 0$ otherwise. Then, the observation can be expressed as $T_k^* = d_k T_k + (1 - d_k) \tau_k$, $k = 1, 2, \dots, m$, and we denote the observed values as $t_k^* = d_k t_k + (1 - d_k) \tau_k$, $k = 1, 2, \dots, m$. In this paper,

we consider the system structure to be unknown but unchanged over time, i.e., s is considered as a vector of parameters. Our aim is to estimate the system signature s and the model parameter θ simultaneously based on the censored system-based lifetime data $(t^*, d) = ((t_1^*, d_1), \dots, (t_m^*, d_m))$.

3 Maximum likelihood estimation and EM algorithm

Suppose \mathbf{U} is the observed data which are generated from the distribution $f(u|\theta)$, and \mathbf{V} is the missing data. Then, the complete data are (\mathbf{U}, \mathbf{V}) with a joint distribution $f(\mathbf{u}, \mathbf{v}|\theta) = f(\mathbf{v}|\mathbf{u}, \theta)f(\mathbf{u}|\theta)$, where θ is the vector of model parameters. The problem of interest is to estimate the parameter vector θ . The log-likelihood based on the observed data \mathbf{u} is

$$\ell(\theta) = \ln f(\mathbf{u}|\theta). \tag{3}$$

The MLE of θ can be obtained by direct maximization of the log-likelihood function in Eq. (3). However, in some cases, the direct maximization of the log-likelihood function based on the observed data may not be straightforward. An alternative method to obtain the MLE of θ is by applying the EM algorithm. Consider the log-likelihood function based on the complete data (\mathbf{U}, \mathbf{V}) given by $\ln f(\mathbf{u}, \mathbf{v}|\theta)$. The log-likelihood function of (\mathbf{U}, \mathbf{V}) is in fact a random variable because \mathbf{V} is unobserved. The EM algorithm first takes the conditional expectation of the complete data log-likelihood with respect to the unobserved data \mathbf{V} , given the observed data \mathbf{U} and the parameter estimate in the h -th iteration $\theta^{(h)}$. Let us define

$$Q(\theta, \theta^{(h)}) = E_{\mathbf{V}} \left(\ln f(\mathbf{u}, \mathbf{V}|\theta) | \mathbf{u}, \theta^{(h)} \right) = \int_{\mathbf{V}} \ln f(\mathbf{u}, \mathbf{v}|\theta) f(\mathbf{v}|\mathbf{u}, \theta^{(h)}) d\mathbf{v},$$

where $f(\mathbf{v}|\mathbf{u}, \theta^{(h)})$ is the conditional marginal distribution of the unobserved data \mathbf{V} , given the observed data \mathbf{U} and the current parameter estimate $\theta^{(h)}$. The evaluation of the expectation $Q(\theta, \theta^{(h)})$ is the E-step of the EM algorithm. The first argument θ in $Q(\theta, \theta^{(h)})$ corresponds to the variable that maximizes the function $Q(\theta, \theta^{(h)})$, and the second argument $\theta^{(h)}$ is treated as a constant in each iteration.

Then, the M-step of the EM algorithm maximizes the expected log-likelihood function obtained from the E-step, i.e.,

$$\theta^{(h+1)} = \arg \max_{\theta} Q(\theta, \theta^{(h)}).$$

The MLE of the parameter vector θ is obtained by iterating the E-step and the M-step successively until convergence. Each iteration in the EM algorithm is guaranteed to increase the log-likelihood. Detailed discussions on the convergence of the EM algorithm can be found in Dempster et al. (1977) and Wu (1983). An estimate of the asymptotic variance–covariance matrix of the MLE of parameter vector θ , denoted by $\hat{\theta}$, can be obtained by inverting the observed (local) Fisher information matrix (see, for example, Zhang et al. 2015)

$$\mathbf{I}_U(\hat{\theta}) = \left(-\frac{\partial^2 \ell(\theta)}{\partial \theta \partial \theta^T} \right) \Big|_{\theta=\hat{\theta}}.$$

While the EM algorithm is used to find the MLE of θ , an alternative way to obtain the observed Fisher information matrix is through the use of the missing information principle (Louis 1982; Tanner 1993):

$$\text{Observed information} = \text{Complete information} - \text{Missing information};$$

that is,

$$\mathbf{I}_U(\hat{\theta}) = \mathbf{I}_{(U,V)}(\hat{\theta}) - \mathbf{I}_{V|U}(\hat{\theta}),$$

where $\mathbf{I}_{(U,V)}(\hat{\theta}) = E_V \left(-\frac{\partial^2 \ln f(\mathbf{u}, \mathbf{V}|\theta)}{\partial \theta \partial \theta^T} \right) \Big|_{\theta=\hat{\theta}}$ is the complete information and $\mathbf{I}_{V|U}(\hat{\theta}) = E_V \left(-\frac{\partial^2 \ln f(\mathbf{V}|\mathbf{u}, \theta)}{\partial \theta \partial \theta^T} \right) \Big|_{\theta=\hat{\theta}}$ is the missing information. In the following subsections, we describe different approaches of implementing the EM algorithm based on the censored system-level lifetime data.

3.1 EM algorithm based on system-level likelihood function

Based on the censored system lifetime data $(\mathbf{t}^*, \mathbf{d}) = ((t_1^*, d_1), \dots, (t_m^*, d_m))$, the log-likelihood function is

$$\ell(\theta, \mathbf{s}) = \sum_{k=1}^m d_k \ln \left(\sum_{i=1}^n s_i f_{i:n}(t_k|\theta) \right) + \sum_{k=1}^m (1 - d_k) \ln \left(\sum_{i=1}^n s_i \bar{F}_{i:n}(\tau_k|\theta) \right). \quad (4)$$

The MLEs of θ and \mathbf{s} can be obtained by maximizing the log-likelihood function in (4) with respect to θ and \mathbf{s} . However, direct maximization of (4) with respect to θ and \mathbf{s} simultaneously is not an easy task, especially when the number of components n is large. For this reason, we consider an alternative iterative procedure in which we fix one parameter vector at a time and maximize the log-likelihood function with respect to the other parameter vector.

First, we consider the parameter vector θ to be fixed (i.e., assume that θ is known) and then maximize the log-likelihood function with respect to \mathbf{s} . Here, maximizing Eq. (4) with respect to \mathbf{s} with the constraint $\sum_{i=1}^n s_i = 1$ can be done by maximizing

$$\begin{aligned} & \sum_{k=1}^m d_k \ln \left(\sum_{i=1}^n s_i f_{i:n}(t_k|\theta) \right) + \sum_{k=1}^m (1 - d_k) \ln \left(\sum_{i=1}^n s_i \bar{F}_{i:n}(\tau_k|\theta) \right) \\ & - \nu \left(\sum_{i=1}^n s_i - 1 \right), \end{aligned} \quad (5)$$

where ν is the Lagrange multiplier. Taking the derivative of Eq. (5) with respect to s_i , and setting the derivative to 0, we obtain

$$\sum_{k=1}^m d_k \frac{f_{i:n}(t_k|\boldsymbol{\theta})}{\sum_{l=1}^n s_l f_{l:n}(t_k|\boldsymbol{\theta})} + \sum_{k=1}^m (1 - d_k) \frac{\bar{F}_{i:n}(\tau_k|\boldsymbol{\theta})}{\sum_{l=1}^n s_l \bar{F}_{l:n}(\tau_k|\boldsymbol{\theta})} - \nu = 0, \tag{6}$$

for $i = 1, \dots, n$. Multiplying both sides of Eq. (6) by s_i , and summing over all s_i , for $i = 1, \dots, n$, we can obtain $\nu = m$ and

$$s_i = \frac{1}{m} \left[\sum_{k=1}^m d_k \frac{s_i f_{i:n}(t_k|\boldsymbol{\theta})}{\sum_{l=1}^n s_l f_{l:n}(t_k|\boldsymbol{\theta})} + \sum_{k=1}^m (1 - d_k) \frac{s_i \bar{F}_{i:n}(\tau_k|\boldsymbol{\theta})}{\sum_{l=1}^n s_l \bar{F}_{l:n}(\tau_k|\boldsymbol{\theta})} \right], \tag{7}$$

for $i = 1, \dots, n$.

Second, we consider the system signature \mathbf{s} to be fixed (i.e., assume that \mathbf{s} is known) and then maximize the log-likelihood function with respect to $\boldsymbol{\theta}$. Taking the first derivative of $\ell(\boldsymbol{\theta}, \mathbf{s})$ in Eq. (4) with respect to the parameter $\boldsymbol{\theta}$, we obtain

$$\begin{aligned} \frac{\partial \ell(\boldsymbol{\theta}, \mathbf{s})}{\partial \boldsymbol{\theta}} &= \sum_{k=1}^m d_k \left(\sum_{i=1}^n \frac{s_i f_{i:n}(t_k|\boldsymbol{\theta})}{\sum_{l=1}^n s_l f_{l:n}(t_k|\boldsymbol{\theta})} \frac{\partial \ln f_{i:n}(t_k|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right) \\ &\quad + \sum_{k=1}^m (1 - d_k) \left(\sum_{i=1}^n \frac{s_i \bar{F}_{i:n}(\tau_k|\boldsymbol{\theta})}{\sum_{l=1}^n s_l \bar{F}_{l:n}(\tau_k|\boldsymbol{\theta})} \frac{\partial \ln \bar{F}_{i:n}(\tau_k|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right). \end{aligned} \tag{8}$$

Defining

$$\delta_{ki}(\boldsymbol{\theta}, \mathbf{s}) = \frac{s_i f_{i:n}(t_k|\boldsymbol{\theta})}{\sum_{l=1}^n s_l f_{l:n}(t_k|\boldsymbol{\theta})} \quad \text{and} \quad \tilde{\delta}_{ki}(\boldsymbol{\theta}, \mathbf{s}) = \frac{s_i \bar{F}_{i:n}(\tau_k|\boldsymbol{\theta})}{\sum_{l=1}^n s_l \bar{F}_{l:n}(\tau_k|\boldsymbol{\theta})},$$

for $k = 1, \dots, m$, $i = 1, \dots, n$, Eq. (8) can be rewritten as

$$\begin{aligned} \frac{\partial \ell(\boldsymbol{\theta}, \mathbf{s})}{\partial \boldsymbol{\theta}} &= \sum_{k=1}^m d_k \left(\sum_{i=1}^n \delta_{ki}(\boldsymbol{\theta}, \mathbf{s}) \frac{\partial \ln f_{i:n}(t_k|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right) \\ &\quad + \sum_{k=1}^m (1 - d_k) \left(\sum_{i=1}^n \tilde{\delta}_{ki}(\boldsymbol{\theta}, \mathbf{s}) \frac{\partial \ln \bar{F}_{i:n}(\tau_k|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right). \end{aligned}$$

Here, the observed data \mathbf{U} is the system lifetimes or censoring times $\mathbf{t}^* = (t_1^*, t_2^*, \dots, t_m^*)$ and the censoring indicators $\mathbf{d} = (d_1, d_2, \dots, d_m)$, and the missing (unobserved) data \mathbf{V} is the number of failed components at the time of the system failure for each of the m systems. Let the number of failed components at the time of the system failure for system k be I_k and $\mathbf{I} = (I_1, I_2, \dots, I_m)$. Here, \mathbf{I} can be considered as the latent variable and the complete sample for the system-level EM algorithm is $(\mathbf{U}, \mathbf{V}) = (\mathbf{t}^*, \mathbf{d}, \mathbf{I})$.

With an initial estimate of the parameter vector θ and an initial estimate of the system signature s , say $\theta^{(0)}$ and $s^{(0)}$, respectively, and the estimates at the h -th iteration $(\theta^{(h)}, s^{(h)})$, the $(h + 1)$ -th iteration of the EM algorithm based on the system-level likelihood function can be described as follows:

E-step Evaluate the conditional probability of latent variable I as

$$\begin{aligned} \Pr(I_k = i | T_k = t_k, \theta^{(h)}, s^{(h)}) &= \delta_{ki}(\theta^{(h)}, s^{(h)}) \\ &= \frac{s_i^{(h)} f_{i:n}(t_k | \theta^{(h)})}{\sum_{l=1}^n s_l^{(h)} f_{l:n}(t_k | \theta^{(h)})} \text{ for } d_k = 1, \end{aligned} \quad (9)$$

$$\begin{aligned} \Pr(I_k = i | T_k > \tau_k, \theta^{(h)}, s^{(h)}) &= \tilde{\delta}_{ki}(\theta^{(h)}, s^{(h)}) \\ &= \frac{s_i^{(h)} \bar{F}_{i:n}(\tau_k | \theta^{(h)})}{\sum_{l=1}^n s_l^{(h)} \bar{F}_{l:n}(\tau_k | \theta^{(h)})} \text{ for } d_k = 0, \end{aligned} \quad (10)$$

$k = 1, 2, \dots, m, i = 1, 2, \dots, n$. Then, we obtain the expectation of the complete data on system-level log-likelihood as

$$\begin{aligned} Q(\theta, (\theta^{(h)}, s^{(h)})) &= \sum_{k=1}^m \sum_{i=1}^n d_k \delta_{ki}(\theta^{(h)}, s^{(h)}) \ln f_{i:n}(T_k = t_k | \theta) \\ &\quad + \sum_{k=1}^m \sum_{i=1}^n (1 - d_k) \tilde{\delta}_{ki}(\theta^{(h)}, s^{(h)}) \ln \bar{F}_{i:n}(\tau_k | \theta). \end{aligned} \quad (11)$$

M-step Step 1. Maximize $Q(\theta, (\theta^{(h)}, s^{(h)}))$ in Eq. (11) with respect to θ to obtain $\theta^{(h+1)}$;

Step 2. The updated estimate of s can be obtained as

$$s_i^{(h+1)} = \frac{1}{m} \sum_{k=1}^m \left[d_k \delta_{ki}(\theta^{(h)}, s^{(h)}) + (1 - d_k) \tilde{\delta}_{ki}(\theta^{(h)}, s^{(h)}) \right], \quad (12)$$

for $i = 1, \dots, n$.

The above E-step and M-step are repeated until convergence occurs to the desired level of accuracy. The MLEs of the model parameter θ and the system signature s based on the EM algorithm are the convergent values of the sequence $\theta^{(h)}$ and $s^{(h)}$, respectively. In the following comparative study, we denote this implementation of the EM algorithm based on system-level likelihood function by *EM-sys*.

3.2 EM algorithm based on component-level likelihood function

Due to the complexity of the distribution of order statistics, maximizing the function in Eq. (11) usually does not yield an explicit solution. The EM algorithm will lose its advantage if the M-step in each iteration is not straightforward. Therefore, we consider here a different way of applying the EM algorithm by considering the component-level likelihood function.

When all the n component lifetimes in a system are observed, finding the MLE is simply reduced to maximizing the likelihood function based on a complete sample from the component lifetime distribution f_X , which is a much simpler problem because computational algorithms to compute MLEs of parameters of commonly used lifetime distributions are well-known.

Consider a single n -component system, say system k , with system lifetime $T_k = t_k$ and the system fails due to the failure of the I_k -th ordered component failure. Let $\mathbf{x}_k = (x_{k1}, x_{k2}, \dots, x_{kn})$ be the vector of the n component lifetimes in system k . If all the n component lifetimes in system k are observed, the log-likelihood function can be expressed as

$$\ell_k(\boldsymbol{\theta}) = \sum_{j=1}^n \ln f_X(x_{kj}|\boldsymbol{\theta}). \tag{13}$$

If the lifetime of system k is observed as $T_k = t_k$, then the system lifetime must correspond to one of the n component lifetimes, i.e., $t_k = x_{k(I_k:n)}$. In this case, the number of component failures at the time of the system failure (I_k) and the lifetimes of the other $n - 1$ component lifetimes ($x_{k(1:n)}, x_{k(2:n)}, \dots, x_{k(I_k-1:n)}, x_{k(I_k+1:n)}, \dots, x_{k(n:n)}$) can be considered as missing data and the EM algorithm can then be applied in this framework. On the other hand, if the lifetime of system k is right censored at τ_k , then all the n component lifetimes, $\mathbf{x}_k = (x_{k(1:n)}, x_{k(2:n)}, \dots, x_{k(n:n)})$, can be considered as missing data and the EM algorithm can then be applied. We will describe now the ways of handling these two situations in the following subsections.

3.2.1 Observed system failures

Suppose the lifetime of system k is observed as $T_k = t_k$. For notational simplicity, we use $x_{k1} < x_{k2} < \dots < x_{k(I_k-1)} < t_k = x_{kI_k} < x_{k(I_k+1)} < \dots < x_{kn}$ to denote the ordered component lifetimes in system k . Let us further denote the observed system lifetime by U_k , i.e., $U_k = T_k$, the missing data by $\mathbf{V}_k = (V_{k1}, \mathbf{V}_{k2})$, where $V_{k1} = I_k$ is the number of component failures at the time of the system failure and $\mathbf{V}_{k2} = (X_{k1}, X_{k2}, \dots, X_{k(I_k-1)}, X_{k(I_k+1)}, \dots, X_{kn})$ is the vector of unobserved component lifetimes. From Eq. (13), the log-likelihood function of the complete data $(U_k, V_{k1}, \mathbf{V}_{k2})$ for system k is

$$\begin{aligned} \ell_k(\boldsymbol{\theta}|U_k, V_{k1}, \mathbf{V}_{k2}) &= \ell_k(\boldsymbol{\theta}|T_k = x_{kI_k}, I_k, (x_{k1}, \dots, x_{k(I_k-1)}, x_{k(I_k+1)}, \dots, x_{kn})) \\ &= \sum_{j=1}^{I_k-1} \ln f_X(x_{kj}|\boldsymbol{\theta}) + \ln f_X(t_k|\boldsymbol{\theta}) + \sum_{j=I_k+1}^n \ln f_X(x_{kj}|\boldsymbol{\theta}). \end{aligned} \tag{14}$$

As mentioned in Sect. 2, the EM algorithm first takes the conditional expectation of the log-likelihood function based on complete data with respect to the missing data $(V_{k1}, \mathbf{V}_{k2})$. Suppose the estimates of $\boldsymbol{\theta}$ and s after the h -th iteration are $\boldsymbol{\theta}^{(h)}$ and $s^{(h)}$, respectively. We then need to evaluate the function

$$\begin{aligned}
Q_k(\boldsymbol{\theta}, (\boldsymbol{\theta}^{(h)}, \mathbf{s}^{(h)})) &= E_{(V_{k1}, \mathbf{V}_{k2})|U_k} \left(\ell_k(\boldsymbol{\theta}|U_k, V_{k1}, \mathbf{V}_{k2})|u_k, \boldsymbol{\theta}^{(h)}, \mathbf{s}^{(h)} \right) \\
&= E_{V_{k2}|U_k} \left[E_{V_{k1}|(V_{k2}, U_k)} \left(\ell_k(\boldsymbol{\theta}|U_k, V_{k1}, \mathbf{V}_{k2})|u_k, \boldsymbol{\theta}^{(h)}, \mathbf{s}^{(h)} \right) \right]. \quad (15)
\end{aligned}$$

We first consider the inner layer of the expectation

$$\begin{aligned}
E_{V_{k1}|(V_{k2}, U_k)} \left(\ell_k(\boldsymbol{\theta}|U_k, V_{k1}, \mathbf{V}_{k2})|u_k, \boldsymbol{\theta}^{(h)}, \mathbf{s}^{(h)} \right) \\
= \sum_{i=1}^n \Pr(I_k = i | T_k = t_k, \boldsymbol{\theta}^{(h)}, \mathbf{s}^{(h)}) \ell_k(\boldsymbol{\theta}|U_k, V_{k1} = I_k = i, \mathbf{V}_{k2}). \quad (16)
\end{aligned}$$

Next, we consider the outer layer of the expectation in Eq. (15). The conditional distribution of the first $(I_k - 1)$ ordered component lifetimes, given t_k , is i.i.d. random variables from a right-truncated distribution

$$g_{RX}(x_{kj}|\boldsymbol{\theta}, t_k) = \frac{f_X(x_{kj}|\boldsymbol{\theta})}{F_X(t_k|\boldsymbol{\theta})}, \quad x_{kj} < t_k,$$

for $j = 1, \dots, I_k - 1$; see Arnold et al. (1992). The outer layer of the expectation of the log-likelihood in Eq. (15) with respect to the components failed before the system failure (i.e., $X < t$) can be written as

$$E_{X|X < t}[\ln f_X(x|\boldsymbol{\theta})] = \int_{-\infty}^t \ln f_X(x|\boldsymbol{\theta}) g_{RX}(x|\boldsymbol{\theta}, t) dx.$$

Similarly, given the lifetime of the system k , t_k , the last $(n - I_k)$ ordered component lifetimes are i.i.d. random variables from a left-truncated distribution

$$g_{LX}(x_{kj}|\boldsymbol{\theta}, t_k) = \frac{f_X(x_{kj}|\boldsymbol{\theta})}{\bar{F}_X(t_k|\boldsymbol{\theta})}, \quad x_{kj} > t_k,$$

for $j = I_k + 1, \dots, n$; see Arnold et al. (1992). The outer layer of the expectation of the log-likelihood in Eq. (15) with respect to the components survived at the time of system failure (i.e., $X > t$) can be written as

$$E_{X|X > t}[\ln f_X(x|\boldsymbol{\theta})] = \int_t^{\infty} \ln f_X(x|\boldsymbol{\theta}) g_{LX}(x|\boldsymbol{\theta}, t) dx.$$

Then, Eq. (15) can be expressed as

$$\begin{aligned}
Q_k(\boldsymbol{\theta}, (\boldsymbol{\theta}^{(h)}, \mathbf{s}^{(h)})) \\
= \sum_{i=1}^n \Pr \left(I_k = i | T_k = t_k, \boldsymbol{\theta}^{(h)}, \mathbf{s}^{(h)} \right)
\end{aligned}$$

$$\left[(i - 1) \int_{-\infty}^{t_k} \ln f_X(x|\boldsymbol{\theta})g_{RX}(x|\boldsymbol{\theta}^{(h)}, t_k)dx + f_X(t_k|\boldsymbol{\theta}) + (n - i) \int_{t_k}^{\infty} \ln f_X(x|\boldsymbol{\theta})g_{LX}(x|\boldsymbol{\theta}^{(h)}, t_k)dx \right]. \tag{17}$$

3.2.2 Censored system lifetimes

Suppose the component lifetimes in system k are $\mathbf{x}_k = (x_{k1}, \dots, x_{kn})$, and the system lifetime T_k is right censored at time τ_k . Let Λ_k be the random variable that denotes the number of failed components in system k before time τ_k . Λ_k is a discrete random variable with support on $\{0, 1, \dots, n - 1\}$. In this case, the observed data are $U_k = \tau_k$ and the missing data are $\mathbf{V}_k = (\mathbf{V}_{k1}, \mathbf{V}_{k2})$, where $\mathbf{V}_{k1} = \Lambda_k$ and $\mathbf{V}_{k2} = (X_{k1}, \dots, X_{k\Lambda_k}, X_{k(\Lambda_k+1)}, \dots, X_{kn})$. From Eq. (13), the log-likelihood function of the complete data $(U_k, \mathbf{V}_{k1}, \mathbf{V}_{k2})$ can be written as

$$\ell_k(\boldsymbol{\theta}|U_k, \mathbf{V}_{k1}, \mathbf{V}_{k2}) = \sum_{j=1}^{\Lambda_k} \ln f_X(x_{k(j:n)}|\boldsymbol{\theta}) + \sum_{j=\Lambda_k+1}^n \ln f_X(x_{k(j:n)}|\boldsymbol{\theta}). \tag{18}$$

As in the case when the system failure is observed, the conditional distribution of the first Λ_k ordered component lifetimes, given the censoring time τ_k , are i.i.d. random variables from a right-truncated distribution

$$g_{RX}(x_{kj}|\boldsymbol{\theta}, \tau_k) = \frac{f_X(x_{kj}|\boldsymbol{\theta})}{F_X(\tau_k|\boldsymbol{\theta})}, \quad x_{kj} < \tau_k,$$

for $j = 1, \dots, \Lambda_k$. Given the censoring time of system k , τ_k , the last $(n - \Lambda_k)$ ordered component lifetimes are i.i.d. random variables from a left-truncated distribution

$$g_{LX}(x_{kj}|\boldsymbol{\theta}, \tau_k) = \frac{f_X(x_{kj}|\boldsymbol{\theta})}{\bar{F}_X(\tau_k|\boldsymbol{\theta})}, \quad x_{kj} > \tau_k,$$

for $j = \Lambda_k + 1, \dots, n$. Suppose the estimates of $\boldsymbol{\theta}$ and \mathbf{s} after the h -th iteration are $\boldsymbol{\theta}^{(h)}$ and $\mathbf{s}^{(h)}$, respectively. We then need to evaluate the function

$$\begin{aligned} Q_k(\boldsymbol{\theta}, (\boldsymbol{\theta}^{(h)}, \mathbf{s}^{(h)})) &= \sum_{\lambda=0}^{n-1} \Pr(\Lambda_k = \lambda|T_k > \tau_k, \boldsymbol{\theta}^{(h)}, \mathbf{s}^{(h)}) \left[\lambda \int_{-\infty}^{\tau_k} \ln f_X(x|\boldsymbol{\theta})g_{RX}(x|\boldsymbol{\theta}^{(h)}, \tau_k)dx \right. \\ &\quad \left. + (n - \lambda) \int_{\tau_k}^{\infty} \ln f_X(x|\boldsymbol{\theta})g_{LX}(x|\boldsymbol{\theta}^{(h)}, \tau_k)dx \right], \end{aligned} \tag{19}$$

where

$$\begin{aligned} \Pr(\Lambda_k = \lambda | T_k > \tau_k, \boldsymbol{\theta}^{(h)}, \mathbf{s}^{(h)}) &= \frac{\Pr(\Lambda_k = \lambda, T_k > \tau_k | \boldsymbol{\theta}^{(h)}, \mathbf{s}^{(h)})}{\Pr(T_k > \tau_k | \boldsymbol{\theta}^{(h)}, \mathbf{s}^{(h)})} \\ &= \frac{(1 - \sum_{l=1}^{\lambda} s_l^{(h)}) \binom{n}{\lambda} [F_X(\tau_k | \boldsymbol{\theta}^{(h)})]^\lambda [\bar{F}_X(\tau_k | \boldsymbol{\theta}^{(h)})]^{n-\lambda}}{\bar{F}_T(\tau_k | \boldsymbol{\theta}^{(h)}, \mathbf{s}^{(h)})} \stackrel{\text{def}}{=} \delta_{k\lambda}^c(\boldsymbol{\theta}^{(h)}, \mathbf{s}^{(h)}), \end{aligned} \quad (20)$$

$(1 - \sum_{l=1}^{\lambda} s_l^{(h)})$ is the probability that the system did not fail due to the i -th ordered component failure, $i = 1, 2, \dots, \lambda$, and $\binom{n}{\lambda} [F_X(\tau_k | \boldsymbol{\theta}^{(h)})]^\lambda [\bar{F}_X(\tau_k | \boldsymbol{\theta}^{(h)})]^{n-\lambda}$ is the probability that λ -out-of- n components are failed at time τ_k .

3.2.3 Implementation of the EM algorithm

Based on the results presented in Sects. 3.2.1 and 3.2.2, when the system signature s is unknown and given the estimates of $\boldsymbol{\theta}$ and \mathbf{s} after the h -th iteration are $\boldsymbol{\theta}^{(h)}$ and $\mathbf{s}^{(h)}$, respectively, the $(h + 1)$ -th iteration of the EM algorithm on component-level can be described as follows:

E-step Compute the conditional probabilities $\delta_{ki}(\boldsymbol{\theta}^{(h)}, \mathbf{s}^{(h)})$ in Eq. (9) for $d_k = 1$, $i = 1, \dots, n$, and $\delta_{k\lambda}^c(\boldsymbol{\theta}^{(h)}, \mathbf{s}^{(h)})$ in Eq. (20) for $d_k = 0$, $\lambda = 0, \dots, n - 1$, $k = 1, \dots, m$. We then have

$$\begin{aligned} Q(\boldsymbol{\theta}, (\boldsymbol{\theta}^{(h)}, \mathbf{s}^{(h)})) &= \sum_{k=1}^m d_k \sum_{i=1}^n \delta_{ki}(\boldsymbol{\theta}^{(h)}, \mathbf{s}^{(h)}) \\ &\quad \times \left[(i - 1) \int_{-\infty}^{\tau_k} \ln f_X(x | \boldsymbol{\theta}) g_{RX}(x | \boldsymbol{\theta}^{(h)}, \tau_k) dx + f_X(\tau_k | \boldsymbol{\theta}) \right. \\ &\quad \left. + (n - i) \int_{\tau_k}^{\infty} \ln f_X(x | \boldsymbol{\theta}) g_{LX}(x | \boldsymbol{\theta}^{(h)}, \tau_k) dx \right] \\ &\quad + \sum_{k=1}^m (1 - d_k) \sum_{\lambda=0}^{n-1} \delta_{k\lambda}^c(\boldsymbol{\theta}^{(h)}, \mathbf{s}^{(h)}) \\ &\quad \times \left[\lambda \int_{-\infty}^{\tau_k} \ln f_X(x | \boldsymbol{\theta}) g_{RX}(x | \boldsymbol{\theta}^{(h)}, \tau_k) dx \right. \\ &\quad \left. + (n - \lambda) \int_{\tau_k}^{\infty} \ln f_X(x | \boldsymbol{\theta}) g_{LX}(x | \boldsymbol{\theta}^{(h)}, \tau_k) dx \right]. \end{aligned} \quad (21)$$

M-step Step 1. Maximize $Q(\boldsymbol{\theta}, (\boldsymbol{\theta}^{(h)}, \mathbf{s}^{(h)}))$ in Eq. (21) with respect to $\boldsymbol{\theta}$ to obtain $\boldsymbol{\theta}^{(h+1)}$;

Step 2. The updated estimate of \mathbf{s} can be obtained by Eq. (12) for $i = 1, \dots, n$.

The MLEs of the model parameter $\boldsymbol{\theta}$ and the system signature \mathbf{s} can be obtained by repeating the E-step and M-step until convergence occurs to the desired level of accuracy. In the following comparative study, we denote this implementation of the EM algorithm based on component-level likelihood function by *EM-comp*.

3.3 Alternate implementation based on component-level likelihood function

Suppose a Type II censored system lifetime data, $t = (t_{1:m} < t_{2:m} < \dots < t_{r:m})$, $r \leq m$, are observed. According to the results in order statistics (Arnold et al. 1992, pp. 23), the censored system lifetimes $T_{r+1:m}, \dots, T_{m:m}$ have the same distribution as order statistics $Z_{1:m-r}, \dots, Z_{m-r:m-r}$, where Z_1, \dots, Z_{m-r} are generated from

$$g_L(z|t_{r:m}) = \frac{f_T(z)}{1 - F_T(t_{r:m})}, \quad z > t_{r:m}.$$

Then, the distribution of order statistic $T_{r+j:m}$ is simply

$$g_{(r+j):m}(z|t_{r:m}) = \frac{(m-r)!}{(j-1)!(m-r-j)!} \times g_L(z|t_{r:m})G_L(z|t_{r:m})^{j-1}(1 - G_L(z|t_{r:m}))^{m-r-j}, \quad z \geq t_{r:m},$$

$j = 1, 2, \dots, (m-r)$, where $G_L(z|t_{r:m}) = \int_{t_{r:m}}^z g_L(x|t_{r:m})dx$. The expectation of $T_{r+j:m}$ can then be expressed as

$$E(T_{(r+j):m}) = \int_{t_{r:m}}^{\infty} x g_{(r+j):m}(x|t_{r:m})dx, \quad j = 1, \dots, m-r.$$

Once we have the expectation of the censored system lifetimes, those expectations could be used to replace the censored system lifetimes, i.e., assuming the observed system lifetimes to be $T_{1:m}, \dots, T_{r:m}, E(T_{r+1:m}), \dots, E(T_{m:m})$ as an approximation. The theoretical basis of this approximation is approximating $E[\ln f_T(T)|T > t_{r:m}]$ by $\ln f_T[E(T|T > t_{r:m})]$ based on the first-order Taylor expansion of $\ln f_T(T)$ at $E(T|T > t_{r:m})$. Then, the EM algorithm discussed in Sects. 3.1 and 3.2 can be applied with $d_1 = 1, \dots, d_m = 1$ on the system lifetimes. We denote this implementation of the EM algorithm by *EM-exp*.

4 Coherent systems with unknown structure

In the preceding sections, when the system signature is unknown, we have developed different implementations of the EM algorithm for the estimation of the model parameter θ and the system signature s , but we did not put any restriction on the estimate of the system signature except $\sum_{i=1}^n \hat{s}_i = 1$. However, in practice, if we have the information that the observed system lifetimes are coming from a coherent system with n components, there are a limited number of possible system signatures that need to be considered. For example, if the system is a 3-component coherent system, then there are only 5 possible arrangements, as shown in Table 1.

If the system is a 4-component coherent system, then there are 20 possible arrangements (Navarro et al. 2007, p. 184). Therefore, we should take into account the restriction on the limited possible system signatures in the parameter estimation procedures. This can be achieved by applying the proposed EM algorithms with known

Table 1 All possible 3-component coherent system

System	System lifetime	System signature
Series	$\min(X_1, X_2, X_3) = X_{1:3}$	$s_1 = (1, 0, 0)$
Series-parallel	$(X_1, \max(X_2, X_3))$	$s_2 = (1/3, 2/3, 0)$
2-out-of-3	$\max_{1 \leq i < j \leq 3} (\min(X_i, X_j)) = X_{2:3}$	$s_3 = (0, 1, 0)$
Parallel-series	$\max(X_1, \min(X_2, X_3))$	$s_4 = (0, 2/3, 1/3)$
Parallel	$\max(X_1, X_2, X_3) = X_{3:3}$	$s_5 = (0, 0, 1)$

system signature for each of the possible system signature and then choose the system signature that results in the largest likelihood value. Specifically, if there are L possible system signatures, $s_l, l = 1, 2, \dots, L$, we can apply the proposed implementations of EM algorithms by fixing the system signature as s_l to obtain the MLE of θ as $\hat{\theta}_l, l = 1, 2, \dots, L$, and then computing the value of the log-likelihood in Eq. (4) as $\ell(\hat{\theta}_l, s_l)$. Then, suppose

$$\ell(\hat{\theta}_{l^*}, s_{l^*}) = \max_{l=1,2,\dots,L} \{\ell(\hat{\theta}_l, s_l)\},$$

the final estimates of the parameter vector and the system signature will be $\hat{\theta}_{l^*}$ and s_{l^*} , respectively.

5 Weibull distributed component lifetime data

To illustrate the methodologies developed in the preceding sections, we consider the lifetimes of the components in an n -component system to follow a two-parameter Weibull distribution. Then, Monte Carlo simulations and a numerical example are used here to evaluate the performance of different implementations of the EM algorithm in obtaining the MLEs of the parameters in the subsequent sections.

5.1 Implementations of the EM algorithm

Suppose that the component lifetimes follow a two-parameter Weibull distribution (see, for example, Weibull 1951; Murthy et al. 2004; Rinne 2010) with c.d.f. and p.d.f.

$$F_X(x|\alpha, \beta) = 1 - \exp\left[-\left(\frac{x}{\beta}\right)^\alpha\right], x > 0$$

and $f_X(x|\alpha, \beta) = \frac{\alpha}{x} \left(\frac{x}{\beta}\right)^{\alpha-1} e^{-\left(\frac{x}{\beta}\right)^\alpha}, x > 0,$

respectively, where $\alpha > 0$ is the shape parameter and $\beta > 0$ is the scale parameter. The log-transformation of the component lifetimes follows the smallest extreme value

(SEV) distribution. Let $W = \ln X$ be the log-lifetime of the component. Then, the c.d.f. and p.d.f. of the log-transformed component lifetimes are

$$F_W(w|\mu, \sigma) = 1 - \exp \left\{ - \exp \left[\frac{w - \mu}{\sigma} \right] \right\}, -\infty < w < \infty$$

$$\text{and } f_W(w|\mu, \sigma) = \frac{1}{\sigma} e^{\frac{w-\mu}{\sigma}} \exp \left[-e^{\frac{w-\mu}{\sigma}} \right], -\infty < w < \infty,$$

respectively, with location parameter $-\infty < \mu = \ln \beta < \infty$ and scale parameter $\sigma = 1/\alpha > 0$.

The p.d.f. and c.d.f. of the i -th order statistic $W_{i:n}$ are, respectively,

$$f_{i:n}(w|\mu, \sigma) = \binom{n}{i} i (F_W(w|\mu, \sigma))^{i-1} f_W(w|\mu, \sigma) (1 - F_W(w|\mu, \sigma))^{n-i}$$

$$= \frac{1}{\sigma} \binom{n}{i} i \exp \left(\frac{w - \mu}{\sigma} \right) \left\{ \exp \left[- \exp \left(\frac{w - \mu}{\sigma} \right) \right] \right\}^{n-i+1}$$

$$\times \left\{ 1 - \exp \left[- \exp \left(\frac{w - \mu}{\sigma} \right) \right] \right\}^{i-1}, -\infty < w < \infty \quad (22)$$

$$\text{and } F_{i:n}(w|\mu, \sigma) = \sum_{l=i}^n \binom{n}{l} [F_W(w|\mu, \sigma)]^l [1 - F_W(w|\mu, \sigma)]^{n-l}$$

$$= \sum_{l=i}^n \binom{n}{l} \left\{ 1 - \exp \left[- \exp \left(\frac{w - \mu}{\sigma} \right) \right] \right\}^l$$

$$\times \left\{ \exp \left[- \exp \left(\frac{w - \mu}{\sigma} \right) \right] \right\}^{n-l}, -\infty < w < \infty, \quad (23)$$

where $-\infty < \mu < \infty$ and $\sigma > 0$. Here after, we consider the log-transformed system and component lifetimes instead of the original lifetimes as it is more convenient. For notational simplicity, we still use X to denote the log-transformed component lifetime and T to denote the log-transformed system lifetime, unless stated otherwise.

In the following subsections, we illustrate different implementations of the EM algorithm described earlier in Sect. 3 when the component lifetime distribution is Weibull. The algorithms and the required computational formulae for the conditional expectations are presented.

5.1.1 EM-sys algorithm

The EM-sys algorithm is realized by directly applying Eqs. (22) and (23) to the algorithm described earlier in Sect. 3.1. Suppose the estimates of the parameter vector and the system signature in the h -th iteration are, respectively, $\theta^{(h)} = (\mu^{(h)}, \sigma^{(h)})$ and $\mathbf{s}^{(h)}$. Then, the $(h + 1)$ -th iteration of the EM-sys algorithm can be described as follows:

- E-step* Evaluate the conditional probability of \mathbf{I} in Eq. (9) by plugging in the p.d.f. in Eq. (22) and the c.d.f. in Eq. (23);

M-step Step 1. Maximize $Q(\theta, (\theta^{(h)}, s^{(h)}))$ in Eq. (11) with respect to μ and σ to obtain $\mu^{(h+1)}$ and $\sigma^{(h+1)}$;
 Step 2. Obtain the updated $s^{(h)}$ by Eq. (12).

5.1.2 EM-comp algorithm

Suppose a sample $w = (w_1, w_2, \dots, w_n)$ follows SEV distribution with location parameter μ and scale parameter σ . Then, the MLEs of μ and σ can be obtained by solving the following equation (Lawless 2011):

$$\hat{\sigma} = \frac{\sum_{i=1}^n w_i \exp(w_i/\hat{\sigma})}{\sum_{i=1}^n \exp(w_i/\hat{\sigma})} - \frac{1}{n} \sum_{i=1}^n w_i, \tag{24}$$

and then substituting the solution $\hat{\sigma}$ into

$$\hat{\mu} = \hat{\sigma} \ln \left[\frac{1}{n} \sum_{i=1}^n \exp \left(\frac{w_i}{\hat{\sigma}} \right) \right]. \tag{25}$$

The M-step of the EM algorithm based on component-level likelihood function can be simplified to solve Eq. (24).

In order to compute the conditional expectations in the E-step, we consider a random variable Z_k that follows a left-truncated SEV distribution with truncation at time t_k (see, for example, Ng et al. 2002). The p.d.f. of Z_k is then given by

$$\begin{aligned} f_{Z_k}(z_k | Z_k > t_k, \mu, \sigma) &= \frac{f_X(z_k | \mu, \sigma)}{1 - F_X(t_k | \mu, \sigma)} \\ &= \frac{\exp[\exp(\xi_k)]}{\sigma} \exp \left[\frac{z_k - \mu}{\sigma} - \exp \left(\frac{z_k - \mu}{\sigma} \right) \right], \quad z_k > t_k, \\ &\quad -\infty < \mu < \infty, \sigma > 0, \end{aligned}$$

where $\xi_k = \frac{t_k - \mu}{\sigma}$.

Let $Z_k^* = \frac{Z_k - \mu}{\sigma}$ and $\eta_k = \exp[\exp(\xi_k)]$. Then, the p.d.f. of Z_k^* could be expressed as

$$f_{Z_k^*}(z_k^* | Z_k^* > \xi_k, \mu, \sigma) = \eta_k \exp [z_k^* - \exp(z_k^*)], \quad z_k^* > \xi_k, \quad -\infty < \mu < \infty, \sigma > 0.$$

The moment generating function of Z_k^* , conditional on $Z_k^* > \xi_k$, is given by

$$M_{Z_k^*}(s) = \eta_k \Gamma(s + 1, e^{\xi_k}) = \Gamma(s + 1) \left[\eta_k - \sum_{p=0}^{+\infty} \frac{e^{(s+p+1)\xi_k}}{\Gamma(s + p + 2)} \right],$$

where $\Gamma(\alpha) = \int_0^\infty u^{\alpha-1} e^{-u} du$ is the gamma function and $\Gamma(\alpha, x) = \int_x^\infty u^{\alpha-1} e^{-u} du$ is the incomplete gamma function. Then, we have the following expectations:

$$\begin{aligned}
 E[Z_k|\xi_k, \mu, \sigma] &= \sigma E_{R,1,k} + \mu, \\
 E[e^{Z_k/\sigma}|\xi_k, \mu, \sigma] &= e^{\mu/\sigma} E_{R,2,k} \\
 \text{and } E[Z_k e^{Z_k/\sigma}|\xi_k, \mu, \sigma] &= e^{\mu/\sigma} [\sigma E_{R,3,k} + \mu E_{R,2,k}],
 \end{aligned}
 \tag{26}$$

where

$$\begin{aligned}
 E_{R,1,k} &= E(Z_k^*|\xi_k, \mu, \sigma) \\
 &= \psi(1)\eta_k + \sum_{p=0}^{+\infty} \frac{e^{(p+1)\xi_k} \psi(p+2)}{\Gamma(p+2)} - [\xi_k + \psi(1)] \sum_{p=0}^{+\infty} \frac{e^{(p+1)\xi_k}}{\Gamma(p+2)}, \\
 E_{R,2,k} &= E(e^{Z_k^*}|\xi_k, \mu, \sigma) \\
 &= e^{\xi_k} + 1 \\
 \text{and } E_{R,3,k} &= E(Z_k^* e^{Z_k^*}|\xi_k, \mu, \sigma) \\
 &= \psi(2)\eta_k + \sum_{p=0}^{+\infty} \frac{e^{(p+2)\xi_k} \psi(p+3)}{\Gamma(p+3)} - [\xi_k + \psi(2)] \sum_{p=0}^{+\infty} \frac{e^{(p+2)\xi_k}}{\Gamma(p+3)},
 \end{aligned}$$

$\psi(\cdot) = \frac{d}{dx} \ln \Gamma(x)$ is the digamma function.

Similarly, let Y_k be a random variable with a right-truncated SEV distribution at time t_k . The p.d.f. of Y_k is then

$$\begin{aligned}
 f_{Y_k}(y_k|Y_k < t_k, \mu, \sigma) &= \frac{f_X(y_k|\mu, \sigma)}{F_X(t_k|\mu, \sigma)} \\
 &= \frac{1}{\sigma[1 - \exp(\exp(\xi_k))]} \exp\left[\frac{y_k - \mu}{\sigma} - \exp\left(\frac{y_k - \mu}{\sigma}\right)\right], \\
 & \quad y_k < t_k, -\infty < \mu < \infty, \sigma > 0.
 \end{aligned}$$

If $Y_k^* = \frac{Y_k - \mu}{\sigma}$, the p.d.f. of Y_k^* can be expressed as

$$\begin{aligned}
 f_{Y_k^*}(y_k^*|Y_k^* < \xi_k, \mu, \sigma) &= \frac{1}{[1 - \exp(\exp(\xi_k))]} \exp[y_k^* - \exp(y_k^*)], \\
 & \quad y_k^* > \xi_k, -\infty < \mu < \infty, \sigma > 0.
 \end{aligned}$$

The moment generating function of Y_k^* , conditional on $Y_k^* < \xi_k$, is given by

$$M_{Y_k^*}(s) = (\eta_k - 1)^{-1} [\eta_k \Gamma(s + 1) - M_{Z_k^*}(s)].$$

The conditional expectations of interest are then given by

$$\begin{aligned}
 E[Y_k|\xi_k, \mu, \sigma] &= \sigma E_{L,1,k} + \mu, \\
 E[e^{Y_k/\sigma}|\xi_k, \mu, \sigma] &= e^{\mu/\sigma} E_{L,2,k} \\
 \text{and } E[Y_k e^{Y_k/\sigma}|\xi_k, \mu, \sigma] &= e^{\mu/\sigma} [\sigma E_{L,3,k} + \mu E_{L,2,k}],
 \end{aligned}
 \tag{27}$$

where

$$\begin{aligned}
 E_{L,1,k} &= E(Y_k^*|\xi_k, \mu, \sigma) \\
 &= (\eta_k - 1)^{-1} [\eta_k \psi(1) - E(Z_k^*|\xi_k, \mu, \sigma)], \\
 E_{L,2,k} &= E(e^{Y_k^*}|\xi_k, \mu, \sigma) \\
 &= (\eta_k - 1)^{-1} [\eta_k - E(e^{Z_k^*}|\xi_k, \mu, \sigma)] \\
 \text{and } E_{L,3,k} &= E(Y_k^* e^{Y_k^*}|\xi_k, \mu, \sigma) \\
 &= (\eta_k - 1)^{-1} [\eta_k \psi(2) - E(Z_k^* e^{Z_k^*}|\xi_k, \mu, \sigma)].
 \end{aligned}$$

As mentioned in Sect. 3.2, suppose system k failed at time t_k . Then, the components still working at time t_k have the same distribution as the random variable Z_k , and the components failed before time t_k have the same distribution as the random variable Y_k . If system k is censored at time τ_k , with $\xi_k^* = \frac{\tau_k - \mu}{\sigma}$, the above formulations are still applicable.

Following the notation and setting described earlier in Sect. 3.1, the $(h + 1)$ -th iteration of the *EM-comp* is as follows:

E-step Evaluate the conditional probabilities $\delta_{ki}(\mu^{(h)}, \sigma^{(h)}, \mathbf{s}^{(h)})$ in Eq. (9) for $i = 1, \dots, n$, and $\delta_{k\lambda}^c(\mu^{(h)}, \sigma^{(h)}, \mathbf{s}^{(h)})$ in Eq. (20), for $\lambda = 0, \dots, n - 1, k = 1, \dots, m$;

M-step Step 1. Update the estimates of the parameters by Eqs. (24) and (25) as

$$\begin{aligned}
 \sigma^{(h+1)} &= \frac{\sum_{k=1}^m (\rho_3(k|\mu^{(h)}, \sigma^{(h)}, \mathbf{s}^{(h)}))}{\sum_{k=1}^m (\rho_2(k|\mu^{(h)}, \sigma^{(h)}, \mathbf{s}^{(h)}))} - \frac{\sum_{k=1}^m (\rho_1(k|\mu^{(h)}, \sigma^{(h)}, \mathbf{s}^{(h)}))}{mn} \\
 \text{and } \mu^{(h+1)} &= \sigma^{(h+1)} \ln \left[\frac{\sum_{k=1}^m (\rho_2(k|\mu^{(h)}, \sigma^{(h)}, \mathbf{s}^{(h)}))}{mn} \right],
 \end{aligned}$$

where

$$\begin{aligned}
 \rho_l(k|\mu^{(h)}, \sigma^{(h)}, \mathbf{s}^{(h)}) &= d_k \sum_{i=1}^n E_{o,l}(k, i|\mu^{(h)}, \sigma^{(h)}, \mathbf{s}^{(h)}) \\
 &\quad + (1 - d_k) \sum_{\lambda=0}^{n-1} E_{c,l}(k, \lambda|\mu^{(h)}, \sigma^{(h)}, \mathbf{s}^{(h)}),
 \end{aligned}$$

for $l = 1, 2, 3$, and

$$\begin{aligned}
 E_{o,1}(k, i|\mu^{(h)}, \sigma^{(h)}, \mathbf{s}^{(h)}) &= \delta_{ki}(\mu^{(h)}, \sigma^{(h)}, \mathbf{s}^{(h)}) \left[(i - 1)E(Y_k|\xi_k, \mu^{(h)}, \sigma^{(h)}) \right. \\
 &\quad \left. + t_k + (n - i)E(Z_k|\xi_k, \mu^{(h)}, \sigma^{(h)}) \right], \\
 E_{c,1}(k, \lambda|\mu^{(h)}, \sigma^{(h)}, \mathbf{s}^{(h)}) &= \delta_{k\lambda}^c(\mu^{(h)}, \sigma^{(h)}, \mathbf{s}^{(h)}) \left[\lambda E(Y_k|\xi_k^*, \mu^{(h)}, \sigma^{(h)}) \right. \\
 &\quad \left. + (n - \lambda)E(Z_k|\xi_k^*, \mu^{(h)}, \sigma^{(h)}) \right],
 \end{aligned}$$

$$\begin{aligned}
 E_{o,2}(k, i | \mu^{(h)}, \sigma^{(h)}, \mathbf{s}^{(h)}) &= \delta_{ki}(\mu^{(h)}, \sigma^{(h)}, \mathbf{s}^{(h)}) \left[(i - 1)E(e^{Y_k/\sigma^{(h)}} | \xi_k, \mu^{(h)}, \sigma^{(h)}) \right. \\
 &\quad \left. + e^{t_k/\sigma^{(h)}} + (n - i)E(e^{Z_k/\sigma^{(h)}} | \xi_k, \mu^{(h)}, \sigma^{(h)}) \right], \\
 E_{c,2}(k, \lambda | \mu^{(h)}, \sigma^{(h)}, \mathbf{s}^{(h)}) &= \delta_{k\lambda}^c(\mu^{(h)}, \sigma^{(h)}, \mathbf{s}^{(h)}) \left[\lambda E(e^{Y_k/\sigma^{(h)}} | \xi_k^*, \mu^{(h)}, \sigma^{(h)}) \right. \\
 &\quad \left. + (n - \lambda)E(e^{Z_k/\sigma^{(h)}} | \xi_k^*, \mu^{(h)}, \sigma^{(h)}) \right], \\
 E_{o,3}(k, i | \mu^{(h)}, \sigma^{(h)}, \mathbf{s}^{(h)}) &= \delta_{ki}(\mu^{(h)}, \sigma^{(h)}, \mathbf{s}^{(h)}) \left[(i - 1)E(Y_k e^{Y_k/\sigma^{(h)}} | \xi_k, \mu^{(h)}, \sigma^{(h)}) \right. \\
 &\quad \left. + t_k e^{t_k/\sigma^{(h)}} + (n - i)E(Z_k e^{Z_k/\sigma^{(h)}} | \xi_k, \mu^{(h)}, \sigma^{(h)}) \right] \\
 \text{and } E_{c,3}(k, \lambda | \mu^{(h)}, \sigma^{(h)}, \mathbf{s}^{(h)}) &= \delta_{k\lambda}^c(\mu^{(h)}, \sigma^{(h)}, \mathbf{s}^{(h)}) \left[\lambda E(Y_k e^{Y_k/\sigma^{(h)}} | \xi_k^*, \mu^{(h)}, \sigma^{(h)}) \right. \\
 &\quad \left. + (n - \lambda)E(Z_k e^{Z_k/\sigma^{(h)}} | \xi_k^*, \mu^{(h)}, \sigma^{(h)}) \right];
 \end{aligned}$$

Step 2. Update $\mathbf{s}^{(h+1)}$ by Eq. (12).

5.1.3 EM-exp algorithm

The algorithm of *EM-exp* based on Weibull distributed component lifetimes can be described as follows:

E-step Step 1. Evaluate the expectations if the system lifetimes t_k are censored, for $k = r + 1, \dots, m$, obtaining the pseudo-complete system lifetimes $t_1, t_2, \dots, t_r, E(T_{r+1}), \dots, E(T_m)$;

Step 2. Evaluate the conditional probabilities $\delta_{ki}(\mu^{(h)}, \sigma^{(h)}, \mathbf{s}^{(h)})$ in Eq. (9) for $i = 1, \dots, n, k = 1, \dots, m$;

M-step The same as the M-step in Sect. 5.1.2 with $d_k = 1$ for $k = 1, \dots, m$.

5.2 Illustrative example

A numerical example is given here to illustrate the three different implementations of the EM algorithm: *EM-sys*, *EM-comp* and *EM-exp*, proposed in Sects. 3.1, 3.2 and 3.3 by assuming the component lifetime distribution to be Weibull. A sample of $m = 30$ system lifetimes was simulated from a 4-component system with system signature $\mathbf{s} = (1/4, 1/4, 1/2, 0)$, and the component lifetimes follow a Weibull distribution with scale parameter $\beta = 1$ and shape parameter $\alpha = 1$. In other words, the logarithm of the component lifetimes is assumed to follow a SEV distribution with location parameter $\mu = 0$ and scale parameter $\sigma = 1$. The ordered log-transformed system lifetimes are presented in Table 2. We consider the estimation of parameters and the system signature when the system structure is unknown with different censoring proportions in the following subsections.

We are interested in estimating both the system signature \mathbf{s} and the model parameters μ and σ . We apply the three different implementations of the EM algorithm proposed in Sects. 3.1, 3.2 and 3.3, and the numerical maximization to obtain the MLEs based on the dataset presented in Table 2 with censoring proportions $q = 0, 10$ and 30% . We consider four different combinations of initial values for these algorithms:

Table 2 Logarithm of $m = 30$ simulated system lifetimes of a 4-component system with system signature $\mathbf{s} = (1/4, 1/4, 1/2, 0)$ when the component lifetimes are from *Weibull*(1, 1)

-4.5469	-4.4799	-3.9642	-3.2333	-2.8369	-2.6229	-1.7921	-1.4315	-1.2964	-1.2523
-1.1188	-1.1021	-1.0527	-0.8610	-0.5840	-0.5731	-0.3609	-0.2885	-0.2457	0.0230
0.1509	0.2665	0.3275	0.3735	0.3877	0.4422	0.4823	0.5025	0.6734	1.1459

- (i) $\theta^{(0)} = (\hat{\mu}^{(0)}, \hat{\sigma}^{(0)}) = (0, 1), \mathbf{s}^{(0)} = (0.25, 0.25, 0.50, 0.00)$;
- (ii) $\theta^{(0)} = (\hat{\mu}^{(0)}, \hat{\sigma}^{(0)}) = (2, 3), \mathbf{s}^{(0)} = (0.25, 0.25, 0.50, 0.00)$;
- (iii) $\theta^{(0)} = (\hat{\mu}^{(0)}, \hat{\sigma}^{(0)}) = (0, 1), \mathbf{s}^{(0)} = (0.40, 0.30, 0.20, 0.10)$;
- (iv) $\theta^{(0)} = (\hat{\mu}^{(0)}, \hat{\sigma}^{(0)}) = (2, 3), \mathbf{s}^{(0)} = (0.40, 0.30, 0.20, 0.10)$.

The algorithm is considered to have converged when

$$\max(|\hat{\mu}^{(h+1)} - \hat{\mu}^{(h)}|, |\hat{\sigma}^{(h+1)} - \hat{\sigma}^{(h)}|, |\hat{s}_i^{(h+1)} - \hat{s}_i^{(h)}|, i = 1, 2, 3, 4) < 10^{-6}.$$

The parameter estimates (say, $\hat{\theta} = (\hat{\mu}, \hat{\sigma})$) and the estimate of system signature $\hat{\mathbf{s}}$ of the three different implementations of the EM algorithm and the parameter estimates obtained from direct maximization of the log-likelihood function are presented in Table 3. In addition, to assess which method results in estimates with the largest likelihood, the values of the log-likelihood in Eq. (4) are evaluated at the parameter estimates and are denoted by $\ell(\hat{\theta}, \hat{\mathbf{s}})$. The largest value of $\ell(\hat{\theta}, \hat{\mathbf{s}})$ for each setting is marked with an asterisk *. We can observe from Table 3 that different initial values $(\theta^{(0)}, \mathbf{s}^{(0)})$ result in different estimates of the parameters and the system signature. The initial value of the system signature, $\mathbf{s}^{(0)}$, plays an important role in the resulting estimates, especially in the estimation of the system signature. It is noteworthy that the elements set to be zero in the initial values of the system signature $\mathbf{s}^{(0)}$ will never become nonzero in the algorithm, and hence, the final estimate of those elements is always equal to zero. From Table 3, the estimate of θ is very close to the true value when the censoring proportion is small, regardless of the initial values. We will further investigate the effect of the initial values on the final estimates by using a Monte Carlo simulation study in the following section.

For Type II censored sample, the first r -out-of- m system failures are observed, and thus, we define the censoring proportion by $q = (m - r)/m$. For Type II censored sample with $r = 27$ (i.e., $q = 10\%$) in Table 2 with initial values $(\hat{\mu}^{(0)}, \hat{\sigma}^{(0)}) = (2, 3)$ and $\mathbf{s}^{(0)} = (0.25, 0.25, 0.50, 0.00)$, the implementations *EM-sys*, *EM-comp* and *EM-exp* take 78 iterations (in 17.54 s), 176 iterations (in 1.06 s) and 205 iterations (in 1.22 s), respectively, to converge to the final estimates. In this example, we can see that the implementations of EM algorithm based on component-level likelihood function, *EM-comp* and *EM-exp*, take less time to converge to the final estimates. This is because the M-step of *EM-sys* requires a two-dimensional numerical optimization, while the M-step in *EM-comp* and *EM-exp* only requires solving one nonlinear equation.

Since these three proposed implementations of the EM algorithm converge to different values, to evaluate the correctness of the estimates obtained by these implementations, we compute the values of the log-likelihood function in Eq. (4), $\ell(\hat{\theta}, \hat{\mathbf{s}})$

Table 3 MLEs obtained by three different implementations of the EM algorithm with unknown system signature

q (%)	Implementation	$\hat{\mu}$	$\hat{\sigma}$	\hat{s}	$\ell(\hat{\theta}, \hat{s})$
$\theta^{(0)} = (0, 1), s^{(0)} = (.25, .25, .5, 0)$					
0	<i>EM-sys</i>	0.0025	1.1675	(0.3719, 0.0054, 0.6226, 0.0000)	− 50.5854
	<i>EM-comp/EM-exp</i>	− 0.0004	1.1678	(0.3729, 0.0000, 0.6271, 0.0000)	− 50.5779*
10	<i>EM-sys</i>	0.0081	1.1915	(0.3660, 0.0064, 0.6277, 0.0000)	− 51.2285
	<i>EM-comp</i>	0.0046	1.1913	(0.3673, 0.0000, 0.6327, 0.0000)	− 51.2209*
	<i>EM-exp</i>	0.0232	1.1535	(0.3853, 0.0000, 0.6147, 0.0000)	− 51.2289
30	<i>EM-sys</i>	0.1366	1.3421	(0.3554, 0.0317, 0.6129, 0.0000)	− 50.6604*
	<i>EM-comp</i>	0.1075	1.3400	(0.3589, 0.0002, 0.6409, 0.0000)	− 50.6653
	<i>EM-exp</i>	0.1779	1.2146	(0.4240, 0.0001, 0.5759, 0.0000)	− 50.7007
$\theta^{(0)} = (2, 3), s^{(0)} = (.25, .25, .5, 0)$					
0	<i>EM-sys</i>	0.0034	1.1676	(0.3714, 0.0076, 0.6211, 0.0000)	− 50.5884
	<i>EM-comp/EM-exp</i>	− 0.0004	1.1678	(0.3729, 0.0000, 0.6271, 0.0000)	− 50.5779*
10	<i>EM-sys</i>	0.0084	1.1916	(0.3659, 0.0069, 0.6272, 0.0000)	− 51.2292
	<i>EM-comp</i>	0.0046	1.1913	(0.3673, 0.0000, 0.6327, 0.0000)	− 51.2209*
	<i>EM-exp</i>	0.0232	1.1535	(0.3853, 0.0000, 0.6147, 0.0000)	− 51.2289
30	<i>EM-sys</i>	0.1315	1.3419	(0.3558, 0.0266, 0.6176, 0.0000)	− 50.6611*
	<i>EM-comp</i>	− 0.237	2.3526	(0.0000, 0.1914, 0.8086, 0.0000)	− 50.7275
	<i>EM-exp</i>	0.1786	1.2148	(0.4238, 0.0010, 0.5752, 0.0000)	− 50.7004
$\theta^{(0)} = (0, 1), s^{(0)} = (.4, .3, .2, .1)$					
0	<i>EM-sys</i>	0.0162	1.1644	(0.3656, 0.0385, 0.5925, 0.0034)	− 50.5835
	<i>EM-comp/EM-exp</i>	− 0.0004	1.1678	(0.3729, 0.0000, 0.6271, 0.0000)	− 50.5779*
10	<i>EM-sys</i>	− 0.7414	1.1930	(0.2598, 0.0068, 0.1432, 0.5902)	− 51.1351
	<i>EM-comp</i>	− 0.7456	1.1949	(0.2609, 0.0014, 0.1462, 0.5915)	− 51.1337*
	<i>EM-exp</i>	0.0233	1.1535	(0.3852, 0.0002, 0.6146, 0.0000)	− 51.2291
30	<i>EM-sys</i>	− 0.6633	1.3368	(0.2183, 0.0797, 0.1659, 0.5362)	− 50.6486
	<i>EM-comp</i>	− 0.6617	1.3350	(0.2222, 0.0695, 0.1757, 0.5326)	− 50.6443
	<i>EM-exp</i>	− 0.5370	1.2351	(0.2740, 0.0164, 0.2558, 0.4538)	− 50.6405*
$\theta^{(0)} = (2, 3), s^{(0)} = (.4, .3, .2, .1)$					
0	<i>EM-sys</i>	0.0025	1.1675	(0.3719, 0.0055, 0.6226, 0.0000)	− 50.5854*
	<i>EM-comp/EM-exp</i>	0.7165	1.1846	(0.4840, 0.5160, 0.0000, 0.0000)	− 51.3797
10	<i>EM-sys</i>	− 0.7391	1.1925	(0.2619, 0.0010, 0.1505, 0.5866)	− 51.1328*
	<i>EM-comp</i>	− 0.7453	1.1949	(0.2613, 0.0000, 0.1478, 0.5909)	− 51.1331
	<i>EM-exp</i>	0.7590	1.2039	(0.4913, 0.5087, 0.0000, 0.0000)	− 51.8952
30	<i>EM-sys</i>	− 0.6619	1.3365	(0.2175, 0.0823, 0.1640, 0.5361)	− 50.6492
	<i>EM-comp</i>	− 0.6617	1.3350	(0.2222, 0.0695, 0.1757, 0.5326)	− 50.6443*
	<i>EM-exp</i>	0.1786	1.2148	(0.4238, 0.0010, 0.5752, 0.0000)	− 50.7004

and compare these values. These values of the log-likelihood function are also presented in Table 3 with “*” to indicate the largest log-likelihood among the three

implementations. We do observe that these values of log-likelihood based on different implementations are quite close to each other.

5.3 Monte Carlo simulation study

In this subsection, we evaluate the performance of the proposed implementations of the EM algorithm using Monte Carlo simulation. Once again, the component lifetime distribution is assumed to be Weibull, and hence, the logarithm of the component lifetimes follows SEV distribution. Without loss of generality, we set $\mu = 0$ and $\sigma = 1$ in the following simulation studies.

5.3.1 Coherent systems with unknown system signature

In Sect. 4, we described the procedure to estimate the parameter vector and the system signature when we have the information that the systems are coherent systems. A Monte Carlo simulation study is conducted here to evaluate the performance of the proposed implementations of the EM algorithm in correctly identifying the system signature.

We consider the 3-component coherent systems with sample sizes $m = 10$ and 100. The five possible coherent systems and their corresponding system signatures are presented in Table 1. We generate samples of m system lifetimes with system signature s_l , $l = 1, 2, 3, 4, 5$, presented in Table 1. Then, we apply the procedure described in Sect. 4 to obtain the estimate of θ . Since all three implementations of the EM algorithm converge to the same value when the system signature is assumed to be known, we present the results based only on *EM-comp* in this simulation study. The simulation results were based on 10,000 Monte Carlo runs.

The obtained results are summarized in Tables 4 and 5, where the sample sizes are $m = 10$ and 100, respectively. The values in Tables 4 and 5 represent the percentages that the data sampled from the row category have the largest likelihood value assuming it is sampled from the column category. For example, the cell in row 1 (Series) column 2 (Series-parallel) in Table 4 is 0.292, which means of the 10,000 iterations that the data are generated from a series system, the value of the likelihood of a series-parallel system is the largest 2920 times compared to the other four likelihoods (series, 2-out-of-3, parallel-series and parallel systems).

From Tables 4 and 5, we observe that the EM algorithm could not correctly detect which system the data are generated from when the sample size is small (say $m = 10$). When $m = 100$, the percentage of correctly detecting the system structure increases. It may also be seen that when the sample size is small, the most unlikely choices are parallel-series and 2-out-of-3 systems.

To investigate this problem further, a sample of size 30 was generated from a series-parallel system with signature $(1/3, 2/3, 0)$, and with log-transformed components following a SEV distribution with location $\mu = 0$ and scale $\sigma = 1$. The EM algorithm is then used to estimate the parameters μ and σ for each of the five possible system signatures separately. Figure 2 presents the p.d.f. of the log-system lifetime for each of the possible system, when log-component lifetimes follow SEV distribution with

Table 4 The proportion of correct detection of system signature when the sample size is 10

Generated from	Identified as				
	Series	Series–parallel	2-out-of-3	Parallel–series	Parallel
Series	0.295	0.292	0.076	0.033	0.304
Series–parallel	0.285	0.313	0.075	0.037	0.290
2-out-of-3	0.271	0.195	0.081	0.044	0.409
Parallel–series	0.239	0.171	0.082	0.045	0.462
Parallel	0.122	0.227	0.123	0.044	0.484

Table 5 The proportion of correct detection of system signature when the sample size is 100

Generated from	Identified as				
	Series	Series–parallel	2-out-of-3	Parallel–series	Parallel
Series	0.392	0.349	0.193	0.027	0.039
Series–parallel	0.119	0.525	0.131	0.030	0.195
2-out-of-3	0.211	0.079	0.329	0.135	0.247
Parallel–series	0.093	0.030	0.258	0.188	0.431
Parallel	0.057	0.021	0.227	0.163	0.532

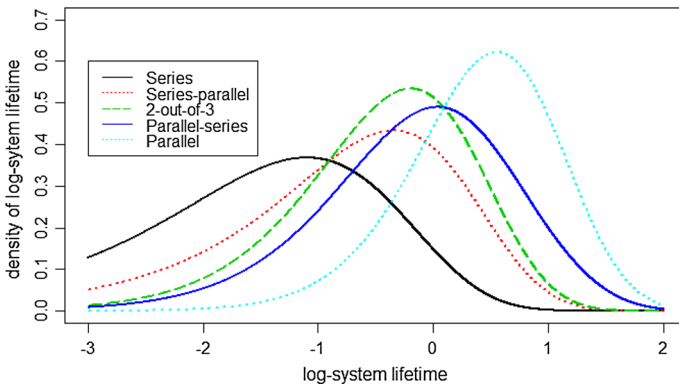


Fig. 2 Distributions of log-system lifetimes, when log-component lifetimes follow SEV distribution with $\mu = 0$ and $\sigma = 1$, based on different system signatures

$\mu = 0$ and $\sigma = 1$. Figure 3 presents the estimated system lifetime distribution by assuming different underlying system structures and the parameter estimates are also presented. From Fig. 3, it can be seen that the estimated distribution of the log-transformed system lifetime under different assumptions is very similar, even when the parameter estimates vary.

5.3.2 Unknown system signature

Simulation studies are used in this subsection to evaluate the performance of the proposed methodologies. The log-component lifetimes in the system with system

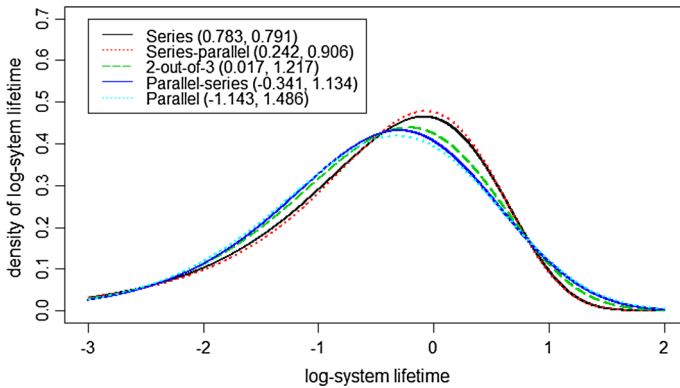


Fig. 3 Estimated $f_T(t)$ with parameters estimated from different system signatures

signature $s = (.25, .25, .5, 0)$ are generated from SEV distribution with $\mu = 0$ and $\sigma = 1$ with sample size $m = 20, 30$, and proportion of censoring $q = 0, 10, 30, 50\%$. Different initial values are considered. The stopping criterion used is

$$\max(|\hat{\mu}^{(h+1)} - \hat{\mu}^{(h)}|, |\hat{\sigma}^{(h+1)} - \hat{\sigma}^{(h)}|, |\hat{s}^{(h+1)} - \hat{s}^{(h)}|, i = 1, 2, 3, 4) < 10^{-4}.$$

The simulation results are based on 10,000 Monte Carlo runs, and these results are presented in Tables 6 and 7. We have also considered different system signatures in the simulation study, and since the results demonstrate similar behavior, for the sake of brevity, we only present here the results for $s = (.25, .25, .5, 0)$. The average of parameter estimates and the mean square errors (MSEs) of the estimates were obtained. Moreover, the values of the likelihood function in Eq. (4) were evaluated with the estimates of the component lifetime distribution parameters and system signature, i.e., $\ell(\hat{\theta}, \hat{s})$. The proportions of cases in which the implementation results in estimates with the largest likelihood function among the three proposed implementations, denoted as “Prop”, are also presented in Tables 6 and 7. The obtained results are presented in Tables 6 and 7 in which the starting values of (μ, σ) and s are $\{(0, 1), (.25, .25, .5, 0)\}$ and $\{(0, 1), (.1, .4, .3, .2)\}$, respectively. The results for starting values $\{(1, 2), (.25, .25, .5, 0)\}, \{(1, 2), (.1, .4, .3, .2)\}$ are not presented here for the sake of conciseness.

It is observed that those estimates obtained from the three implementations of the EM algorithm depend on the choice of the initial values, especially the initial value of the system signature, when the system lifetimes are Type II censored. When the censoring proportions are small (say, $q < 30\%$), the effect of the initial values to the final estimates is relatively small, and the three proposed implementations *EM-sys*, *EM-comp* and *EM-exp* all give quite similar estimates which are indeed close to the true value.

For the cases when the true value of the unknown system signature is used as the starting value of the EM algorithm (see Table 6), *EM-sys* has the largest proportion of times that the likelihood function is the largest in most of the cases, especially for small censoring proportion $q \leq 10\%$. For the estimation of the parameters of the component lifetime distribution, *EM-exp* gives the smallest MSEs in most of the cases.

Table 6 Estimates of the parameters and the proportions of $\ell(\hat{\theta}, \hat{s})$ being the largest based on *EM-sys*, *EM-comp*, *EM-exp* when the starting values for (μ, σ) and s are (0, 1) and (.25, .25, .5, 0), respectively

<i>q</i> (%)	Implementation	$\hat{\mu}$		$\hat{\sigma}$		Ave. of \hat{s}	Prop
		Ave.	MSE	Ave.	MSE		
<i>m</i> = 20							
0	<i>EM-sys</i>	−0.0004	0.0661	0.9345	0.0543	(0.3293, 0.1533, 0.5173, 0.0000)	0.582*
	<i>EM-comp/EM-exp</i>	−0.0031	0.0637	0.9376	0.0499	(0.3246, 0.1545, 0.5210, 0.0000)	0.418
10	<i>EM-sys</i>	−0.0750	0.0756	0.9484	0.0774	(0.3252, 0.0624, 0.6124, 0.0000)	0.485*
	<i>EM-comp</i>	−0.0904	0.0783	0.9639	0.0679	(0.3061, 0.0597, 0.6341, 0.0000)	0.207
	<i>EM-exp</i>	−0.0247	0.0620	0.8950	0.0621	(0.3465, 0.1050, 0.5483, 0.0000)	0.308
30	<i>EM-sys</i>	−0.1497	0.0977	0.9213	0.0979	(0.3210, 0.0099, 0.6690, 0.0000)	0.462*
	<i>EM-comp</i>	−0.1824	0.1176	0.9705	0.0798	(0.2819, 0.0037, 0.7144, 0.0000)	0.275
	<i>EM-exp</i>	−0.1204	0.0819	0.8192	0.0858	(0.3526, 0.0495, 0.5978, 0.0000)	0.263
50	<i>EM-sys</i>	−0.2813	0.1885	0.8792	0.1187	(0.2799, 0.0045, 0.7155, 0.0000)	0.436*
	<i>EM-comp</i>	−0.3131	0.2158	0.9715	0.0848	(0.2237, 0.0005, 0.7758, 0.0000)	0.267
	<i>EM-exp</i>	−0.3002	0.1664	0.7450	0.1225	(0.3117, 0.0618, 0.6263, 0.0000)	0.297
<i>m</i> = 30							
0	<i>EM-sys</i>	−0.0067	0.0597	0.9652	0.0484	(0.3200, 0.1441, 0.5360, 0.0000)	0.555*
	<i>EM-comp/EM-exp</i>	−0.0100	0.0568	0.9678	0.0445	(0.3159, 0.1457, 0.5384, 0.0000)	0.445
10	<i>EM-sys</i>	−0.0768	0.0654	0.9793	0.0679	(0.3136, 0.0693, 0.6171, 0.0000)	0.516*
	<i>EM-comp</i>	−0.0819	0.0661	0.9806	0.0589	(0.3057, 0.0599, 0.6344, 0.0000)	0.205
	<i>EM-exp</i>	−0.0221	0.0547	0.9212	0.0535	(0.3464, 0.0970, 0.5565, 0.0000)	0.279
30	<i>EM-sys</i>	−0.1349	0.0798	0.9614	0.0866	(0.3142, 0.0190, 0.6668, 0.0000)	0.463*
	<i>EM-comp</i>	−0.1598	0.0927	0.9943	0.0763	(0.2884, 0.0102, 0.7014, 0.0000)	0.294
	<i>EM-exp</i>	−0.1005	0.0643	0.8564	0.0702	(0.3569, 0.0372, 0.6058, 0.0000)	0.243
50	<i>EM-sys</i>	−0.2434	0.1366	0.9328	0.0992	(0.2789, 0.0076, 0.7135, 0.0000)	0.458*
	<i>EM-comp</i>	−0.2692	0.1562	0.9943	0.0802	(0.2420, 0.0011, 0.7569, 0.0000)	0.269
	<i>EM-exp</i>	−0.2643	0.1300	0.7987	0.0937	(0.3142, 0.0285, 0.6572, 0.0000)	0.273

For the cases when the starting value of the unknown system signature is different from the true value (see Table 7), *EM-exp* has the largest proportion of times that the likelihood function is the largest when the censoring proportion is moderate to large ($q \geq 30\%$). *EM-exp* also gives the smallest MSEs for the estimation of the parameters of the component lifetime distribution in most of the cases. If the estimation of the parameters of the component lifetime distribution is the major focus, we would therefore recommend the use of *EM-exp* method.

6 Lognormal distributed component lifetime data

In order to illustrate how the proposed methodologies can be applied to different lifetime models in addition to the Weibull model, we consider the case when the component lifetimes follow lognormal distribution in this section. The component lifetime X is assumed to follow a lognormal distribution with c.d.f. and p.d.f.

Table 7 Estimates of the parameters and the proportions of $\ell(\hat{\theta}, \hat{s})$ being the largest based on *EM-sys*, *EM-comp*, *EM-exp* when the starting values for (μ, σ) and s are (0, 1) and (.1, .4, .3, .2), respectively

q (%)	Implementation	$\hat{\mu}$		$\hat{\sigma}$		Ave. of \hat{s}	Prop
		Ave.	MSE	Ave.	MSE		
<i>m</i> = 20							
0	<i>EM-sys</i>	-0.2851	0.1655	0.8557	0.0709	(0.2837, 0.1778, 0.2406, 0.2979)	0.547*
	<i>EM-comp/EM-exp</i>	-0.1733	0.0881	0.8902	0.0562	(0.2837, 0.1871, 0.3633, 0.1659)	0.453
10	<i>EM-sys</i>	-0.5180	0.4093	0.8568	0.0856	(0.2486, 0.1779, 0.0301, 0.5434)	0.386
	<i>EM-comp</i>	-0.5056	0.3942	0.8521	0.0795	(0.2507, 0.1820, 0.0258, 0.5414)	0.156
	<i>EM-exp</i>	-0.0895	0.0885	0.8334	0.0753	(0.2241, 0.2722, 0.1521, 0.3515)	0.458*
30	<i>EM-sys</i>	-0.6846	0.5942	0.8613	0.1064	(0.2135, 0.1326, 0.0096, 0.6443)	0.356
	<i>EM-comp</i>	-0.6833	0.5935	0.8593	0.0902	(0.2090, 0.1383, 0.0031, 0.6496)	0.109
	<i>EM-exp</i>	-0.2068	0.0979	0.7480	0.1118	(0.2382, 0.1706, 0.2247, 0.3664)	0.535*
50	<i>EM-sys</i>	-0.8829	0.9016	0.8683	0.1343	(0.1671, 0.0930, 0.0033, 0.7367)	0.352*
	<i>EM-comp</i>	-0.8965	0.9337	0.8871	0.0936	(0.1464, 0.1084, 0.0001, 0.7451)	0.323
	<i>EM-exp</i>	-0.3230	0.1418	0.6689	0.1550	(0.2222, 0.1371, 0.3570, 0.2835)	0.325
<i>m</i> = 30							
0	<i>EM-sys</i>	-0.2835	0.1555	0.8780	0.0585	(0.2830, 0.1607, 0.2704, 0.2859)	0.504*
	<i>EM-comp/EM-exp</i>	-0.1625	0.0796	0.9052	0.0482	(0.2954, 0.1598, 0.3920, 0.1528)	0.496
10	<i>EM-sys</i>	-0.5265	0.4142	0.8887	0.0729	(0.2435, 0.1638, 0.0601, 0.5326)	0.427
	<i>EM-comp</i>	-0.5069	0.3932	0.8814	0.0697	(0.2483, 0.1678, 0.0580, 0.5259)	0.155
	<i>EM-exp</i>	-0.0791	0.0871	0.8423	0.0684	(0.2336, 0.2736, 0.1343, 0.3584)	0.418*
30	<i>EM-sys</i>	-0.6896	0.5992	0.9155	0.0923	(0.1999, 0.1399, 0.0199, 0.6403)	0.382
	<i>EM-comp</i>	-0.6870	0.5972	0.9105	0.0816	(0.1991, 0.1447, 0.0082, 0.6480)	0.098
	<i>EM-exp</i>	-0.2131	0.1000	0.7632	0.1031	(0.2473, 0.1616, 0.1898, 0.4012)	0.520*
50	<i>EM-sys</i>	-0.8783	0.8790	0.9283	0.1176	(0.1608, 0.0995, 0.0073, 0.7324)	0.344
	<i>EM-comp</i>	-0.8855	0.8982	0.9307	0.0900	(0.1544, 0.1045, 0.0003, 0.7408)	0.061
	<i>EM-exp</i>	-0.3356	0.1489	0.6977	0.1341	(0.2276, 0.1038, 0.3277, 0.3408)	0.595*

$$F_X(x|\mu, \sigma) = \Phi \left[\frac{\ln x - \mu}{\sigma} \right] \quad \text{and} \quad f_X(x|\mu, \sigma) = \frac{1}{\sigma x} \phi \left(\frac{\ln x - \mu}{\sigma} \right), \quad x > 0,$$

respectively, where $\phi(z) = e^{-(z^2/2)}/\sqrt{2\pi}$ is the p.d.f. and $\Phi(z) = \int_{-\infty}^z \phi(u)du$ is the c.d.f. of the standard normal distribution, $\exp(\mu)$ is a scale parameter and $\sigma > 0$ is a shape parameter. The log-transformation of the component lifetimes, $W = \ln X$, follows the normal distribution with location parameter $\mu \in (-\infty, \infty)$ and scale parameter $\sigma > 0$.

Based on a complete sample of the log-lifetimes (w_1, w_2, \dots, w_n) from a normal distribution, the MLEs of μ and σ can be simply obtained as

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n w_i \quad \text{and} \quad \hat{\sigma} = \left[\frac{1}{n} \sum_{i=1}^n (w_i - \hat{\mu})^2 \right]^{1/2}, \tag{28}$$

respectively.

Following the formulations presented in Sect. 5, to implement the *EM-comp* algorithm, the expectations of a truncated normally distributed random variable are needed. Let Z_k be the left-truncated normal random variable at time t_k , and Y_k be the right-truncated normal random variable. The required expectations could be obtained as

$$\begin{aligned} E(Z_k|\xi_k, \mu, \sigma) &= \mu + \sigma \lambda(\xi_k), \\ E(Z_k^2|\xi_k, \mu, \sigma) &= \sigma^2[1 + \xi_k \lambda(\xi_k)] + 2\mu\sigma \lambda(\xi_k) + \mu^2, \\ E(Y_k|\xi_k, \mu, \sigma) &= \mu - \sigma \lambda(-\xi_k) \\ \text{and } E(Y_k^2|\xi_k, \mu, \sigma) &= \sigma^2[1 - \xi_k \lambda(-\xi_k)] - 2\mu\sigma \lambda(-\xi_k) + \mu^2, \end{aligned}$$

where $\xi_k = \frac{t_k - \mu}{\sigma}$ and $\lambda(s) = \frac{\phi(s)}{1 - \Phi(s)}$ is the hazard function of the standard normal distribution.

The *Step 1* in M-step at the $(h + 1)$ -th iteration of the *EM-comp* can be described as follows:

M-step Step 1. Update the estimates of the parameters by Eq. (28) as

$$\begin{aligned} \mu^{(h+1)} &= \frac{1}{mn} \sum_{k=1}^m \left[\rho_1^*(k|\mu^{(h)}, \sigma^{(h)}, \mathbf{s}^{(h)}) \right] \\ \text{and } \sigma^{(h+1)} &= \left[\frac{1}{mn} \sum_{k=1}^m \rho_2^*(k|\mu^{(h)}, \sigma^{(h)}, \mathbf{s}^{(h)}) - (\mu^{(h+1)})^2 \right]^{1/2}, \end{aligned}$$

where

$$\begin{aligned} \rho_l^*(k|\mu^{(h)}, \sigma^{(h)}, \mathbf{s}^{(h)}) &= d_k \sum_{i=1}^n E_{o,l}^*(k, i|\mu^{(h)}, \sigma^{(h)}, \mathbf{s}^{(h)}) \\ &\quad + (1 - d_k) \sum_{\lambda=0}^{n-1} E_{c,l}^*(k, \lambda|\mu^{(h)}, \sigma^{(h)}, \mathbf{s}^{(h)}), \\ E_{o,l}^*(k, i|\mu^{(h)}, \sigma^{(h)}, \mathbf{s}^{(h)}) &= \delta_{ki}(\mu^{(h)}, \sigma^{(h)}, \mathbf{s}^{(h)}) \left[(i - 1)E(Y_k^l|\xi_k, \mu^{(h)}, \sigma^{(h)}) \right. \\ &\quad \left. + t_k + (n - i)E(Z_k^l|\xi_k, \mu^{(h)}, \sigma^{(h)}) \right] \\ \text{and } E_{c,l}^*(k, \lambda|\mu^{(h)}, \sigma^{(h)}, \mathbf{s}^{(h)}) &= \delta_{k\lambda}^c(\mu^{(h)}, \sigma^{(h)}, \mathbf{s}^{(h)}) \left[\lambda E(Y_k^l|\xi_k^*, \mu^{(h)}, \sigma^{(h)}) \right. \\ &\quad \left. + (n - \lambda)E(Z_k^l|\xi_k^*, \mu^{(h)}, \sigma^{(h)}) \right], \end{aligned}$$

for $l = 1, 2$

Here, we present a real data example to illustrate when the component lifetimes follow a lognormal distribution. The failure times of a simple software control logic for a water-reservoir control system presented in Pham and Pham (1991) and Teng and Pham (2002) are used here to illustrate the EM algorithm. Water is supplied via

Table 8 Failure times of water-reservoir control systems and the 2 components (software versions) in the system (Teng and Pham 2002)

Fault no.	1	2	3	4	5	6	7	8	9	10
Component 1	1.2	2.8	8.4	10.0	16.4	20.0	24.4	28.0	29.2	31.2
Component 2	3.6	8.4	12.8	14.4	17.2	18.0	20.0	23.2	25.2	28.0
System	3.6	8.4	12.8	14.4	17.2	20.0	24.4	28.0	29.2	31.2
Fault no.	11	12	13	14	15	16	17	18	19	20
Component 1	34.0	36.0	36.8	39.2	40.0	44.0	44.8	54.0	56.0	62.4
Component 2	28.4	30.8	31.2	34.8	36.4	36.8	38.0	39.2	41.6	42.0
System	34.0	36.0	36.8	39.2	40.0	44.0	44.8	54.0	56.0	62.4
Fault no.	21	22	23	24	25	26				
Component 1	80.0	92.0	99.6	100.0+	100.0+	100.0+				
Component 2	46.4	59.6	62.4	98.8	99.6	100.0				
System	80.0	92.0	99.6	100.0+	100.0+	100.0+				

+ The failure time is right censored

a source pipe controlled by a source valve and removed via a drain pipe controlled by a drain valve. There are two level sensors in the control system that maintain the water level between the high and low limits in the water reservoir. The water-reservoir control system achieves fault tolerance by using a two-version programming software control logic in which the system fails only when both of its components (software versions) fail. The control system can be treated as a two-component parallel system with system signature $s = (0, 1)$. The original dataset presented in Table IV of Teng and Pham (2002) contains the failure times of the two individual components. In order to verify the performance of the proposed estimation procedures discussed here, we assume that the system failure times are observed, but not the individual component lifetimes and the component lifetimes follow a lognormal distribution. The dataset is presented in Table 8.

Since the complete component lifetimes are available, we can obtain the estimates of the parameters based on the component lifetimes as $\hat{\mu} = 3.4608$ and $\hat{\sigma} = 0.9582$. The standard errors of $\hat{\mu}$ and $\hat{\sigma}$ obtained from the method using the missing information principle described in Sect. 3 are 0.1336 and 0.0981, respectively. In this example, we can observe that the parameter estimates based on system-level data (Table 9) are close to the parameter estimates based on component-level data, while the system-level data contain less information compared to the component-level data. As expected, the standard errors of the estimates based on system-level data are larger than those based on component-level data.

7 Concluding remarks

In this paper, we develop the EM algorithm for estimating the parameters of the component lifetime distribution and the system structure simultaneously based on censored system-level lifetime data. We propose three different implementations of

Table 9 MLEs obtained by direct maximization and the proposed EM algorithms with known system signature based on the data in Table 8

	$\hat{\mu}$ (SE)	$\hat{\sigma}$ (SE)	$l(\hat{\theta}, s = (0, 1))$
Max-II	3.0100 (0.1924)	1.1086 (0.1677)	– 33.5466
EM-sys	3.0103 (0.1924)	1.1086 (0.1676)	– 33.5466
EM-comp	3.0101 (0.1924)	1.1087 (0.1677)	– 33.5466
EM-exp	3.0133 (0.1910)	1.1022 (0.1653)	– 33.5474

the EM algorithm based on the system-level likelihood function and the component-level likelihood function. These implementations are illustrated when the component lifetimes follow a two-parameter Weibull distribution. The performances of these three implementations of the EM algorithm are evaluated by Monte Carlo simulations. As shown in the simulation results and the illustrative example given in Sects. 5.2 and 5.3, the final estimates depend on the initial value of the system signature. Among the three different implementations, *EM-comp* is the one that is most computationally efficient. Overall, we recommend the use of *EM-exp* method for estimating the parameters when we have no information on the system structure or when the censoring proportion is large. If one is more concerned about the computational efficiency, the *EM-comp* method is the best choice.

Since the estimate of the system structure is dependent on the initial estimate, if more information could be collected, for example, if one can use autopsy on some of the failed systems in order to observe the number of failed components, then a better estimate of the system signature can be obtained and used as an initial estimate. However, when little or no information is available on the system structure, we would suggest using different system signatures as the initial value for the EM algorithm and then comparing the resulting estimates. It will be of practical interest to see how a reliable starting value can be obtained for the implementation of the EM algorithm.

Acknowledgements The authors thank two anonymous referees for their careful review and useful comments and suggestions on an earlier version of this manuscript which resulted in this improved version. HKTNs work was supported by a Grant from the Simons Foundation (#280601).

References

- Al-Nefaiee, A.H., Coolen, F.: Nonparametric predictive inference for system failure time based on bounds for the signature. *Proc. Inst. Mech. Eng. O J. Risk Reliab.* **10**, 513–522 (2013)
- Arnold, B.C., Balakrishnan, N., Nagaraja, H.N.: *A First Course in Order Statistics*. Wiley, New York (1992)
- Balakrishnan, N., Mitra, D.: Left truncated and right censored weibull data and likelihood inference with an illustration. *Comput. Stat. Data Anal.* **56**, 4011–4025 (2012)
- Balakrishnan, N., Ng, H.K.T., Navarro, J.: Exact nonparametric inference for component lifetime distribution based on lifetime data from systems with known signatures. *J. Nonparametr. Stat.* **23**, 741–752 (2011a)
- Balakrishnan, N., Ng, H.K.T., Navarro, J.: Linear inference for type-II censored lifetime data of reliability systems with known signatures. *IEEE Trans. Reliab.* **60**, 426–440 (2011b)
- Bhattacharya, D., Samaniego, F.J.: Estimating component characteristics from system failure-time data. *Nav. Res. Logist.* **57**, 380–389 (2010)

- Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. Ser. B* **39**, 1–38 (1977)
- Dube, S., Pradhan, B., Kundu, D.: Parameter estimation of the hybrid censored log-normal distribution. *J. Stat. Comput. Simul.* **81**, 275–287 (2011)
- Eryilmaz, S., Koutras, M.V., Triantafyllou, I.S.: Signature based analysis of m -consecutive- k -out-of- n : f systems with exchangeable components. *Nav. Res. Logist.* **58**, 344–354 (2011)
- Eykhoff, P.: *System Identification—Parameter and System Estimation*. Wiley, New York (1974)
- Kochar, S., Mukerjee, H., Samaniego, F.J.: The signature of a coherent system and its application to comparisons among systems. *Nav. Res. Logist.* **46**, 507–523 (1999)
- Lawless, J.F.: *Statistical Models and Methods for Lifetime Data*, vol. 362, 2nd edn. Wiley, Hoboken (2011)
- Ljung, L.: *System Identification: Theory for the User*. Prentice-Hall Inc., Upper Saddle River (1999)
- Louis, T.A.: Finding the observed information matrix when using the EM algorithm. *J. R. Stat. Soc. Ser. B* **44**, 226–233 (1982)
- McLachlan, G., Krishnan, T.: *The EM Algorithm and Extensions*, 2nd edn. Wiley, Hoboken (2008)
- Murthy, D.P., Xie, M., Jiang, R.: *Weibull Models*, vol. 505. Wiley, Hoboken (2004)
- Navarro, J., Ruiz, J.M., Sandoval, C.J.: Properties of coherent systems with dependent components. *Commun. Stat. Theory Methods* **36**, 175–191 (2007)
- Ng, H.K.T., Chan, P.S., Balakrishnan, N.: Estimation of parameters from progressively censored data using EM algorithm. *Comput. Stat. Data Anal.* **39**, 371–386 (2002)
- Ng, H.K.T., Navarro, J., Balakrishnan, N.: Parametric inference from system lifetime data under a proportional hazard rate model. *Metrika* **75**, 367–388 (2012)
- Pham, H., Pham, M.: *Software reliability models for critical applications*. Technical Report EGG-2663, Idaho National Engineering Laboratory EG&G, Idaho Falls, ID (1991)
- Rinne, H.: *The Weibull Distribution: A Handbook*. CRC Press, Boca Raton (2010)
- Samaniego, F.: *System Signatures and Their Applications in Engineering Reliability*. Springer, New York (2007)
- Tangirala, A.: *Principles of System Identification: Theory and Practice*. CRC Press, Boca Raton (2015)
- Tanner, M.A.: *Tools for Statistical Inference: Observed Data and Data Augmentation Methods*, 2nd edn. Springer, New York (1993)
- Teng, X., Pham, H.: A software-reliability growth model for n -version programming systems. *IEEE Trans. Reliab.* **51**(3), 311–321 (2002)
- Weibull, W.: A statistical distribution function of wide applicability. *J. Appl. Mech.* **18**, 292–297 (1951)
- Wu, C.F.J.: On the convergence properties of the EM algorithm. *Ann. Stat.* **11**, 95–103 (1983)
- Zhang, J., Ng, H.K.T., Balakrishnan, N.: Statistical inference of component lifetimes with location-scale distributions from censored system failure data with known signature. *IEEE Trans. Reliab.* **64**, 613–626 (2015)