

A stochastic expectation-maximization algorithm for the analysis of system lifetime data with known signature

Yandan Yang¹ · Hon Keung Tony Ng¹ ·
Narayanaswamy Balakrishnan²

Received: 28 August 2014 / Accepted: 8 May 2015 / Published online: 22 May 2015
© Springer-Verlag Berlin Heidelberg 2015

Abstract Statistical estimation of the model parameters of component lifetime distribution based on system lifetime data with known system structure is discussed here. We propose the use of stochastic expectation-maximization (SEM) algorithm for obtaining the maximum likelihood estimates of model parameters based on complete and censored system lifetimes. Different ways of implementing the SEM algorithm are also studied. We have shown that the proposed methods are feasible and are easy to implement for various families of component lifetime distributions. The proposed methodologies are then illustrated with two popular lifetime models—the Weibull and Birnbaum-Saunders distributions. Monte Carlo simulation is then used to compare the performance of the proposed methods with the corresponding estimation by direct maximization. Finally, two illustrative examples are presented along with some concluding remarks.

Keywords Expectation-maximization algorithm · Maximum likelihood estimation · Monte Carlo simulation · Reliability data · Type-II censoring

Electronic supplementary material The online version of this article (doi:[10.1007/s00180-015-0586-6](https://doi.org/10.1007/s00180-015-0586-6)) contains supplementary material, which is available to authorized users.

✉ 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, Ontario L8S 4K1, Canada

1 Introduction

In system reliability engineering, systems are made up of different components and these systems can be complex in their form and structure. For various purposes, engineers and statisticians are often interested in the lifetime distribution of the system as well that of the components that form the system. In many cases, the lifetimes of an n -component system can be observed, but not the lifetime of its components. For example, we will encounter this problem when n -component systems are placed on the field to work and our interest is then in monitoring the reliability of the components in the system as well as the reliability of the system. Therefore, novel statistical methods are required for the analysis of these system-level lifetime data; the development of statistical inference for the lifetime distribution of components based on these system lifetimes then becomes of interest. To develop statistical inference based on system-level data, information about the system structure of the n -component system become essential. System signature (Samaniego 2007), which is an effective way to express the system structure of a n -component system, is considered here. Through the use of system signature, parametric and nonparametric inferential procedures have been developed for component lifetime characteristics based on system-level data (see, for example, Balakrishnan et al. 2011a, b; Bhattacharya and Samaniego 2010; Chahkandi et al. 2014; Ng et al. 2012).

In industrial engineering, censoring is common in life-testing experiments due to time or budget constraints. One of the commonly used censoring schemes is Type-II right censoring scheme, in which testing of systems gets terminated when a pre-fixed number of system failures have been observed. Complete sample can be considered as a special case of Type-II censoring. Based on Type-II right censored system-level data with known system signature, best linear unbiased estimation (BLUE) (Balakrishnan et al. 2011b), maximum likelihood estimation (MLE) (Balakrishnan et al. 2011b; Ng et al. 2012), and regression-based method (Zhang et al. 2015) have all been developed for different statistical distributions for the estimation of model parameters of the component lifetime distribution. However, the development of these estimation methods are specific for certain families of distributions and moreover, the theoretical derivations of the estimators are tedious. Therefore, it is of interest to develop an easy-to-use method for the estimation of model parameters that can be applied generally to any component lifetime distribution.

In principle, the MLE of the unknown parameters can be obtained by directly maximizing the likelihood function. Due to the complexity of the likelihood function based on system lifetime data, the MLEs can only be calculated through numerical approaches such as the Newton–Raphson algorithm and the Nelder–Mead algorithm. But, these numerical methods require heavy computation and/or complicated second derivatives of the log-likelihood function.

Moreover, the convergence of these numerical approaches cannot be guaranteed. Even when these algorithms converge, they may not converge to the global maximum if inappropriate initial values are used. Furthermore, when the rate of missing data is high and numerical approximations are involved in the likelihood function, these direct maximization methods may not be stable (see, for example, Ye and Ng 2014). In practice, therefore, engineers and practitioners may prefer to use

some simple tools that do not require heavy computational efforts. For this reason, we propose the use of stochastic expectation-maximization (SEM) algorithm, proposed by [Celeux and Diebolt \(1985\)](#), as an alternative method to approximate the MLEs of the component lifetime distribution parameters based on complete or censored system-level data. In this paper, we discuss the implementations as well as some merits of the SEM algorithm and then exam the performance of the proposed methods.

The rest of this article is organized as follows. In Sect. 2, we introduce the notation and the definitions of system signature and ordered system signature. We also describe here the maximum likelihood estimation method for the parameter estimation. Then, in Sect. 3, the SEM algorithm is discussed and different approaches of implementing the SEM algorithm are proposed based on complete and Type-II censored system-level data. An extensive Monte Carlo simulation study is used to evaluate the performances of the proposed approaches and the obtained results are presented in Sect. 4. Two commonly used lifetime models—Weibull and Birnbaum–Saunders (BS) distributions—are considered in the Monte Carlo simulation study. Finally, in Sect. 5, the proposed methodologies are illustrated with two numerical examples, and some concluding remarks are made at the end.

2 System signature and likelihood inference based on system-level data

Consider an n -component system with the component lifetimes, denoted by X_1, X_2, \dots, X_n , being independent and identically distributed (i.i.d.) with probability density function (p.d.f.) $f_X(x; \theta)$, cumulative distribution function (c.d.f.) $F_X(x; \theta)$ and survival function (s.f.) $\bar{F}_X(x; \theta)$, where θ is the vector of parameters. We further denote the lifetime of the n -component system by T , with p.d.f. $f_T(t; \theta)$, c.d.f. $F_T(t; \theta)$ and s.f. $\bar{F}_T(t; \theta)$. Suppose m such n -component systems are placed on a life-testing experiment and only the first r ($r \leq m$) system failure times are observed. The remaining $(m - r)$ system failure times are censored resulting in a Type-II censored sample with ordered system lifetimes, denoted by $T_{1:m} < T_{2:m} < \dots < T_{r:m}$. Our interest then is to estimate the parameter θ based on the observed ordered system lifetimes $T_{1:m} < T_{2:m} < \dots < T_{r:m}$.

In system reliability study, a system is said to be coherent if every component is relevant and if its structure function is monotone ([Barlow and Proschan 1965](#)). We consider coherent and mixed reliability systems with n components having i.i.d. lifetimes, where a mixed reliability system is defined as a stochastic mixture of coherent systems ([Boland and Samaniego 2004](#)). In the study of coherent systems, system signature, introduced by [Samaniego \(1985\)](#) and discussed further by [Kochar et al. \(1999\)](#), is an index that characterizes a system with i.i.d. components in a simple and elegant probabilistic way. The system signature is an n -dimensional probability vector $\mathbf{s} = (s_1, s_2, \dots, s_n)$, where the i -th element s_i is the probability that the i -th ordered component failure causes the failure of the system, i.e.,

$$s_i = \Pr(T = X_{i:n}), \quad i = 1, 2, \dots, n,$$

Fig. 1 A 3-component series-parallel system

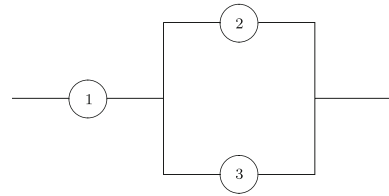


Table 1 Six possible arrangements of the component lifetimes in a 3-component series-parallel system

Arrangements	System lifetime T
$X_1 \leq X_2 \leq X_3$	$X_{1:3}$
$X_1 \leq X_3 \leq X_2$	$X_{1:3}$
$X_2 \leq X_1 \leq X_3$	$X_{2:3}$
$X_2 \leq X_3 \leq X_1$	$X_{2:3}$
$X_3 \leq X_1 \leq X_2$	$X_{2:3}$
$X_3 \leq X_2 \leq X_1$	$X_{2:3}$

and $\sum_{i=1}^n s_i = 1$. For example, let us consider a 3-component series-parallel system, as shown in Fig. 1. Let the lifetimes of these three components be X_1, X_2, X_3 . There are then six possible arrangements of the component lifetimes (see Table 1).

In the first two cases, the system lifetime $T = X_{1:3}$, and therefore $s_1 = \Pr(T = X_{1:3}) = 2/6$. In the last four cases, the system lifetime $T = X_{2:3}$, so that $s_2 = \Pr(T = X_{2:3}) = 4/6$, and similarly $s_3 = \Pr(T = X_{3:3}) = 0$. The signature of this system is, therefore, $(2/6, 4/6, 0) = (1/3, 2/3, 0)$.

Note that many of the systems used in industry can be described by system signature and hence the methodology developed here can readily be used in these cases. For instance, Frenkel and Khvatskin (2006) considered a real-life prototype for n -component phosphor acid filter system used in Rotem/Deshanim Chemical Processing Facility, Arad, Israel. The filter system has n identical turning rollers and a system failure occurs when two adjacent rollers stop working according to the technical specification. The system described in Frenkel and Khvatskin (2006) is a consecutive 2-out-of- n : F system with i.i.d. components, whereas the system structure can be described by using a system signature. For example, the system signatures of consecutive 2-out-of- n : F system with $n = 4, 6$ and 8 components are $s = (0, 1/2, 1/2, 0)$, $s = (0, 1/3, 7/15, 1/5, 0, 0)$, and $s = (0, 1/4, 11/28, 2/7, 1/14, 0, 0, 0)$, respectively. For more information on the signature-based analysis of consecutive k -out-of- n : F systems, one can refer to Eryilmaz (2010), Eryilmaz et al. (2010) and the references cited therein.

System signature is a distribution-free representation of the system structure meaning that it does not depend on the distribution of the component lifetimes. For an n component system with i.i.d. components, the p.d.f. and s.f. of the system lifetime T can be expressed as (Kochar et al. 1999, p. 512)

$$f_T(t; \theta) = \sum_{i=1}^n s_i \binom{n}{i} i f_X(t; \theta) [F_X(t; \theta)]^{i-1} [\bar{F}_X(t; \theta)]^{n-i} \tag{1}$$

and

$$\bar{F}_T(t, \theta) = \sum_{i=1}^n s_i \sum_{j=0}^{i-1} \binom{n}{j} [F_X(t; \theta)]^j [\bar{F}_X(t; \theta)]^{n-j}, \tag{2}$$

respectively.

The log-likelihood function for θ based on Type-II censored system lifetime data is

$$l(\theta; t_{1:m}, t_{2:m}, \dots, t_{r:m}) = \sum_{k=1}^r \ln f_T(t_{k:m}; \theta) + (m - r) \ln \bar{F}_T(t_{r:m}; \theta), \tag{3}$$

where f_T and \bar{F}_T can be expressed in terms of f_X and F_X as given in Eqs. (1) and (2), respectively. For a specific parametric component lifetime density f_X (or equivalently F_X or \bar{F}_X), the MLE of θ can be obtained by maximizing the log-likelihood function in Eq. (3).

Based on the ordered system lifetimes, [Balakrishnan and Volterman \(2014\)](#) recently provided an extension of the system signature, called the ordered system signature. Consider the lifetimes of the n components in the k -th system to be $X_{k1}, X_{k2}, \dots, X_{kn}$, $k = 1, 2, \dots, m$, and denote the corresponding ordered component lifetimes as $X_{k,1:n} < X_{k,2:n} < \dots < X_{k,n:n}$. Intuitively, for an early failed system among the m systems in the life-testing experiment, the failure would be more likely be due to more critical components. In other words, if the information about the ranks of the system lifetimes is given, then the system signature vectors of a system with lifetime $T_{a:m}$ and of a system with lifetime $T_{b:m}$ ($a < b$) should be different. More specifically, [Balakrishnan and Volterman \(2014\)](#) denoted the system signature vector of the k -th ordered system lifetime, $T_{k:m}$, as $\mathbf{s}^{(k:m)} = (s_1^{(k:m)}, \dots, s_n^{(k:m)})$, where

$$s_i^{(k:m)} = \sum_{j=1}^m \Pr(T_{k:m} = X_{j,i:n}), \quad i = 1, \dots, n,$$

is the probability that the k -th ordered system failure time correspond to its i -th ordered component failure. Let ℓ_i denote the number of systems that failed due to their i -th ordered component failure, $i = 1, \dots, n$, and $\mathcal{L} = \{\ell = (\ell_1, \dots, \ell_n) : \ell_1 + \dots + \ell_n = m\}$. Then, $s_i^{(k:m)}$ could be represented as ([Balakrishnan and Volterman 2014](#))

$$s_i^{(k:m)} = \sum_{\mathcal{L}} \binom{m}{\ell_1, \dots, \ell_n} \left\{ \prod_{j=1}^n s_j^{\ell_j} \right\} p_{i|\ell}^{(k:m)}, \tag{4}$$

where $p_{i|\ell}^{(k:m)}$ is the conditional probability that the k th ordered system failed due to the i -th ordered component failure, given that ℓ_j systems failed due to the j -th ordered component failure, for $j = 1, \dots, n$. If the system lifetimes are i.i.d., then $s_i^{(k:m)}$ can also be represented as

Table 2 Elements used to calculate the ordered system signature for $m = 2, n = 3$ and $s = (1/3, 2/3, 0)$

ℓ_1	ℓ_2	ℓ_3	$\binom{2}{\ell_1, \ell_2, \ell_3}$	$s_1^{\ell_1} s_2^{\ell_2} s_3^{\ell_3}$	$p_{1 \ell}^{(1:2)}$	$p_{2 \ell}^{(1:2)}$	$p_{3 \ell}^{(1:2)}$	$p_{1 \ell}^{(2:2)}$	$p_{2 \ell}^{(2:2)}$	$p_{3 \ell}^{(1:2)}$
2	0	0	1	1/9	1	0	0	1	0	0
1	1	0	2	2/9	4/5	1/5	0	1/5	4/5	0
0	2	0	1	4/9	0	1	0	0	1	0

$$s_i^{(k:m)} = \Pr(T_j = X_{j,i:n} | T_{k:m} = T_j), \quad i = 1, 2, \dots, n.$$

For the 3-component series-parallel system presented in Fig. 1, suppose we have two systems, i.e., $m = 2, n = 3$ and $s = (s_1, s_2, s_3) = (1/3, 2/3, 0)$. Since $s_3 = 0$, only the cases $\ell_3 = 0$ will contribute to the ordered system signature. All the possible arrangements of \mathcal{L} and the corresponding $p_{i|\ell}^{(k:m)}$ values are presented in Table 2. The elements in the ordered system signatures could be calculated as

$$\begin{aligned} s_1^{(1:2)} &= 1 \times \frac{1}{9} \times 1 + 2 \times \frac{2}{9} \times \frac{4}{5} = \frac{7}{15}, \\ s_2^{(1:2)} &= 2 \times \frac{2}{9} \times \frac{1}{5} + 1 \times \frac{4}{9} \times 1 = \frac{8}{15}, \\ s_3^{(1:2)} &= 0, \\ s_1^{(2:2)} &= 1 \times \frac{1}{9} \times 1 + 2 \times \frac{2}{9} \times \frac{1}{5} = \frac{1}{5}, \\ s_2^{(2:2)} &= 2 \times \frac{2}{9} \times \frac{4}{5} + 1 \times \frac{4}{9} \times 1 = \frac{4}{5}, \\ s_3^{(2:2)} &= 0. \end{aligned}$$

Hence, the ordered system signatures are

$$\begin{aligned} \mathbf{s}^{(1:2)} &= (s_1^{(1:2)}, s_2^{(1:2)}, s_3^{(1:2)}) = (7/15, 8/15, 0), \\ \mathbf{s}^{(2:2)} &= (s_1^{(2:2)}, s_2^{(2:2)}, s_3^{(2:2)}) = (1/5, 4/5, 0). \end{aligned}$$

Since the system signature is a distribution-free representation of the system structure, as mentioned earlier, Eq. (4) shows that the ordered system signatures also do not depend on the underlying component lifetime distribution. In general, the ordered system signatures are not available in simple form, but their distribution-free nature enables a simple approximation by Monte Carlo methods. In the following section, we will utilize the ordered system signature to develop different approaches for implementing the SEM algorithm.

3 Stochastic expectation-maximization algorithm for system-level data

For the estimation of parameters of the component lifetime distribution, since the system-level lifetime data can be viewed as a missing data problem, the well-known

expectation-maximization (EM) algorithm (Dempster et al. 1977) (see, also McLachlan and Krishnan 2008) can be utilized for determining the MLE of the vector of parameters. The EM algorithm is an iterative procedure that repeatedly fills the missing data in the complete-data log-likelihood with their conditional expected values (E-step) and then maximizes the expected complete data log-likelihood to update the parameter estimates (M-step). Specifically, suppose \mathbf{T} denotes the observed data and \mathbf{Z} denotes the missing data, the complete data is $\mathbf{X} = (\mathbf{T}, \mathbf{Z})$, and the likelihood function based on complete data is $L(\boldsymbol{\theta}; \mathbf{X})$. The E-step of the EM algorithm in the h -th iteration requires the computation of the conditional expectation of the log-likelihood function, with respect to the conditional distribution of \mathbf{Z} , given $\mathbf{T} = \mathbf{t}$, at the current estimate of the parameter vector $\boldsymbol{\theta}^{(h)}$:

$$Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(h)}) = E[\ln L(\boldsymbol{\theta}; \mathbf{T}, \mathbf{Z})|\mathbf{T} = \mathbf{t}, \boldsymbol{\theta}^{(h)}], \quad (5)$$

and the M-step involves maximizing $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(h)})$ with respect to $\boldsymbol{\theta}$, and the revised estimate becomes

$$\boldsymbol{\theta}^{(h+1)} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}}\{Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(h)})\}. \quad (6)$$

Therefore, in order to facilitate the EM algorithm, we need to determine the conditional distribution of \mathbf{Z} , given $\mathbf{T} = \mathbf{t}$ and the current value of the parameter. In some cases, the EM algorithm might be hard to implement, if not impossible, due to the complex forms of the conditional expectations involved in the E-step. SEM algorithm, as long as the missing data is easy to impute, simply replaces the E-step with a stochastic step (S-step):

S-step: Given the values of $\boldsymbol{\theta}^{(h)}$ and $\mathbf{T} = \mathbf{t}$, simulate values of \mathbf{Z} from the conditional distribution of \mathbf{Z} , given $\mathbf{T} = \mathbf{t}$.

The S-step simulates a pseudo-complete data set, and then the M-step involves maximizing the likelihood function based on a complete sample. S-step might be much easier to implement in many cases as compared to the E-step since only a single draw from the conditional distribution is needed.

Let H be the total number of iterations of the SEM algorithm. The sequence of estimates $\boldsymbol{\theta}^{(h)}$, $h = 1, 2, \dots, H$, do not converge to a single point, but this sequence of estimates is a Markov chain that rapidly converges to a stationary distribution under some regularity conditions (Diebolt and Celeux 1993; Diebolt and Ip 1996). This stationary distribution could be obtained after a burn-in period, and thus a point estimate, $\tilde{\boldsymbol{\theta}}$, could be obtained by averaging the sequence of estimates after this burn-in period. One of the advantages of the SEM algorithm is that the random perturbations of the Markov chains prevent the sequence of the estimates being trapped in a local maximum or saddle point. Other commonly used methods for computing the MLE, such as the EM algorithm and the Newton–Raphson method, do not guarantee the convergence to a global maximum or even a local maximum. These methods might converge to a stationary point close to the starting value and that point could be a saddle point (Wu 1983; Redner and Walker 1984). An example in Ip (1994) has displayed

that the SEM algorithm might lead to an improvement over the EM algorithm in terms of convergence to a “good” stationary point.

To further observe the advantage of the SEM algorithm over other commonly used numerical methods for maximizing the log-likelihood function in Eq. (3), we substitute the p.d.f. and c.d.f. of the system lifetimes in Eqs. (1) and (2) into the log-likelihood function in Eq. (3) to obtain

$$\begin{aligned}
 l(\boldsymbol{\theta}; t_{1:m}, t_{2:m}, \dots, t_{r:m}) &= \sum_{k=1}^r \ln \left\{ \sum_{i=1}^n s_i \binom{n}{i} i f_X(t_{k:m}; \boldsymbol{\theta}) [F_X(t_{k:m}; \boldsymbol{\theta})]^{i-1} [\bar{F}_X(t_{k:m}; \boldsymbol{\theta})]^{n-i} \right\} \\
 &+ (m-r) \ln \left\{ \sum_{i=1}^n s_i \sum_{j=0}^{i-1} \binom{n}{j} [F_X(t_{r:m}; \boldsymbol{\theta})]^j [\bar{F}_X(t_{r:m}; \boldsymbol{\theta})]^{n-j} \right\}.
 \end{aligned}$$

This log-likelihood can become computational involved when n is large. Moreover, the c.d.f. and s.f. of the component lifetime, F_X and \bar{F}_X , involved in the log-likelihood function might not be in closed-form for some commonly used component lifetime distributions, such as gamma distribution and log-normal distribution and so numerical approximations are required in the evaluation of these functions. Therefore, the evaluation of the log-likelihood function and any numerical methods that maximize the log-likelihood function directly can become computationally quite intensive. In addition, since the c.d.f. and s.f. of the component lifetime are powers of F_X up to n in the log-likelihood function, any numerical error in approximating these functions will get amplified. The SEM algorithm proposed in this paper can effectively avoid these issues. It is noteworthy that instead of taking the average of the sequence of SEM estimates after burn-in as the parameter estimate, one can use the SEM estimate encountered during the SEM iterations that maximizes the observed data log-likelihood as the parameter estimate. However, this will involve the evaluation of the log-likelihood function in Eq. (3) and hence the issue of error accumulation described above will remain and consequently the advantage of the SEM may get reduced. Therefore, we only consider the use of taking the average of the sequence of SEM estimates after burn-in as the final parameter estimate.

The asymptotic properties of the SEM estimate $\tilde{\boldsymbol{\theta}}$ has been studied by Ip (1994), Diebolt and Ip (1996) and Nielsen (2000). These authors have pointed out that SEM estimate is asymptotically equivalent to the MLE $\hat{\boldsymbol{\theta}}$. For exponential family of distributions, Ip (1994) has shown that the mean of the stationary distribution differs from the MLE by $O(1/m)$ under some appropriate assumptions, where m is the sample size. In addition, Nielsen (2000) has shown that the distribution of SEM estimates converges to a normal distribution with mean equal to the MLE and variance of order $O(1/m)$.

For these reasons, we consider here the SEM algorithm for the analysis of system-level data. We also make use of the ordered system signature (Balakrishnan and Volterman 2014) to develop different ways of implementing the SEM algorithm.

3.1 SEM algorithm for complete system lifetime data

Under the setting described in Sect. 2, suppose m i.i.d. n -component systems are placed on a life-testing experiment and the observed data is $\mathbf{t} = (t_{1:m} < t_{2:m} < \dots < t_{m:m})$. Since our interest is on the component lifetime distribution, the complete data is taken to be $\mathbf{X}_k = (X_{k1}, X_{k2}, \dots, X_{kn})$, $k = 1, 2, \dots, m$, containing $m \times n$ component lifetimes. The log-likelihood function based on the complete data \mathbf{X}_k , $k = 1, 2, \dots, m$, can then be expressed as

$$\ln L(\boldsymbol{\theta}; \mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_m) = \sum_{k=1}^m \sum_{i=1}^n \ln f_X(x_{ki}; \boldsymbol{\theta}).$$

For most commonly used statistical distributions in lifetime data analysis, the maximum likelihood estimation based on complete data is well developed and can be readily computed by the use of standard statistical software. So, the maximization involved in the M-step can be done easily.

3.1.1 SEM algorithm based on ordinary system signature

To develop the SEM algorithm based on system signature, we consider the observed system lifetime of the k -th system (not necessarily in order) among the m systems in the experiment, t_k . Assume that the λ -th component failure in the k -th system caused the failure of the system, i.e., $t_k = x_{k,\lambda:n}$. Then, the conditional distributions of the other $(n - 1)$ components are random variables with either left-truncated or right-truncated distributions. Specifically, the conditional density of the first $(\lambda - 1)$ ordered component lifetimes, $X_{k,1:n}, X_{k,2:n}, \dots, X_{k,(\lambda-1):n}$, given $t_k = x_{k,\lambda:n}$, is a right-truncated density

$$g_{RX}(x; \boldsymbol{\theta} | t_k, \lambda) = \frac{f_X(x; \boldsymbol{\theta})}{F_X(t_k; \boldsymbol{\theta})}, \quad x < t_k, \tag{7}$$

and similarly the conditional density of the last $(n - \lambda)$ ordered component lifetimes, $X_{k,(\lambda+1):n}, X_{k,(\lambda+2):n}, \dots, X_{k,n:n}$, given $t_k = x_{k,\lambda:n}$, is a left-truncated density

$$g_{LX}(x; \boldsymbol{\theta} | t_k, \lambda) = \frac{f_X(x; \boldsymbol{\theta})}{1 - F_X(t_k; \boldsymbol{\theta})}, \quad x > t_k. \tag{8}$$

In the $(h + 1)$ -th SEM iteration, given the current value of the parameter vector $\boldsymbol{\theta}^{(h)}$ and the observed system lifetime t_k , the S-step and M-step then proceed as follows:

- S-step* 1. For the k -th system, generate a discrete random variable Λ based on the system signature of the n -component system with probability mass function $\Pr(\Lambda = \lambda) = s_\lambda$, $\lambda = 1, 2, \dots, n$, and we denote the realization as λ ;
- 2. Generate $\lambda - 1$ random variates from the conditional distribution in Eq. (7) with $\boldsymbol{\theta} = \boldsymbol{\theta}^{(h)}$, say $\tilde{x}_{k1}, \tilde{x}_{k2}, \dots, \tilde{x}_{k(\lambda-1)}$;
- 3. Generate $n - \lambda$ random variates from the conditional distribution in Eq. (8) with $\boldsymbol{\theta} = \boldsymbol{\theta}^{(h)}$, say $\tilde{x}_{k(\lambda+1)}, \tilde{x}_{k(\lambda+2)}, \dots, \tilde{x}_{kn}$;

4. The pseudo-complete sample for system k is then

$$\tilde{\mathbf{x}}_k = (\tilde{x}_{k1}, \tilde{x}_{k2}, \dots, \tilde{x}_{k(\lambda-1)}, \tilde{x}_{k\lambda}, \tilde{x}_{k(\lambda+1)}, \tilde{x}_{k(\lambda+2)}, \dots, \tilde{x}_{kn}),$$

where $\tilde{x}_{k\lambda} = t_k$. Repeat Steps 1–3 for $k = 1, \dots, m$ to obtain the pseudo-complete sample $\tilde{\mathbf{x}}_k$, $k = 1, 2, \dots, m$.

M-step Maximize the log-likelihood function,

$$\ln L(\boldsymbol{\theta}; \tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2, \dots, \tilde{\mathbf{x}}_m) = \sum_{k=1}^m \sum_{i=1}^n \ln f_X(\tilde{x}_{ki}; \boldsymbol{\theta}),$$

with respect to $\boldsymbol{\theta}$ to obtain $\boldsymbol{\theta}^{(h+1)}$ for the next cycle.

To obtain an estimate of $\boldsymbol{\theta}$, we run the SEM algorithm to obtain a sequence of $\boldsymbol{\theta}^{(h)}$, $h = 1, 2, \dots, B$, discard the first few iterations for burn-in, and average over the estimates from the remaining iterations to get an estimate of $\boldsymbol{\theta}$ (say, $\hat{\boldsymbol{\theta}}$).

3.1.2 SEM algorithm based on the ordered system signature

Instead of considering the unordered observed system lifetimes (t_1, t_2, \dots, t_m) and developing the SEM algorithm as described in Sect. 3.1.1, we now incorporate the ordered system signature (Balakrishnan and Volterman 2014) together with the ordered system lifetimes $\mathbf{t} = (t_{1:m} < t_{2:m} < \dots < t_{m:m})$. The ordered system signature for the k -th ordered system lifetime $T_{k:m}$, $\mathbf{s}^{(k:m)} = (s_1^{(k:m)}, \dots, s_n^{(k:m)})$, can be computed explicitly by using the formulas provided in Balakrishnan and Volterman (2014) for some special cases. However, to keep the main advantages of SEM, namely, being simple and easy to implement, we use Monte Carlo method to approximate the ordered system signatures $\mathbf{s}^{(k:m)}$, $k = 1, 2, \dots, m$, based on the fact that the ordered system signature is free of the underlying component lifetime distribution. In particular, for a given system signature, we first generate m system lifetimes and keep track of which component failure led to the system failure (from any distribution F_X) and repeat this process N times. The elements in the ordered system signatures can then be approximated by

$$\hat{s}_i^{(k:m)} = \frac{\text{No. of times that event } A \text{ occurred}}{N},$$

where $A = \{\text{the } k\text{-th ordered system failed due to its } i\text{-th ordered component failure}\}$, for $i = 1, 2, \dots, n$, $k = 1, 2, \dots, m$. In our simulation study, we used $N = 100,000$. Under these considerations, the Step 1 in the **S-Step** described in Sect. 3.1.1 can then be modified as follows:

S-Step 1.* For the k -th system (ordered based on the system lifetimes), generate a discrete random variable Λ based on the ordered system signature with probability mass function $\Pr(\Lambda = \lambda) = \hat{s}_\lambda^{(k:m)}$, $\lambda = 1, 2, \dots, n$, and we denote the realization as λ .

The remaining steps involved in the **S-Step** and the **M-Step** are exactly the same as described earlier in Sect. 3.1.1.

3.2 SEM algorithm for Type-II censored system lifetimes

Suppose a Type-II censored system lifetime data, $\mathbf{t} = (t_{1:m} < t_{2:m} < \dots < t_{r:m})$, $r \leq m$, is observed. For a Type-II censored sample, the early system failures will tend to fail due to the first few ordered component failures. Hence, the r observed system lifetimes are not exchangeable and the ordinary system signature cannot be used in this case. So, ordered system signature will be quite useful to develop the SEM algorithm for Type-II censored system lifetime data. Under Type-II censoring, we can either impute the censored system lifetimes first and then the unobserved component lifetimes in those censored systems, or directly impute the component lifetimes in the censored systems. Based on these two approaches, we propose here two ways of implementing the SEM algorithm for Type-II censored system lifetime data.

3.2.1 Impute component lifetimes directly (SEM-I)

For those failed systems, the unobserved component lifetimes could be imputed by using the methods described in Sect. 3.1.2. For those censored (un-failed) systems, we first consider the conditional probability mass function of the number of failed components in a system, given that the system is still working at time τ . Suppose an n -component system with system signature $\mathbf{s} = (s_1, s_2, \dots, s_n)$ is working at time τ and the system lifetime is ranked j among the m systems ($j = r + 1, \dots, m$) for those censored systems. We then consider the conditional probability vector $\mathbf{p}^{(j:m)} = (p_0^{(j:m)}, p_1^{(j:m)}, \dots, p_{n-1}^{(j:m)})$, where

$$\begin{aligned}
 p_l^{(j:m)} &= \Pr(\text{lout of } n \text{ components failed} | T_{j:m} > \tau) \\
 &= \frac{(1 - \sum_{k=1}^l s_k^{(j:m)}) \binom{n}{l} [F_X(\tau; \boldsymbol{\theta})]^l [\bar{F}_X(\tau; \boldsymbol{\theta})]^{n-l}}{\bar{F}_T(\tau; \boldsymbol{\theta})}, l = 0, 1, \dots, n - 1.
 \end{aligned}$$

Note that it is not possible to have all n components failed for a working coherent system. For the considered Type-II censored system lifetime data, the censoring time is $\tau = t_{r:m}$. So, the h -th iteration of the S-step in the SEM algorithm can be implemented as follows:

- S-step* 1. For the r failed systems with \mathbf{s} , the component lifetimes can be imputed based on ordered system signatures described in Sect. 3.1.2. Hence, we have $r \times n$ imputed component lifetime $\tilde{\mathbf{x}}_k, k = 1, 2, \dots, r$;
2. For the censored system with rank j among the m systems, $j = r + 1, r + 2, \dots, m$, the component lifetimes can be imputed based on the conditional probability vector $\mathbf{p}^{(j:m)}$ as follows:
- i. Generate a discrete random variate Δ based on the conditional probability vector $\mathbf{p}^{(j:m)}$ given $\tau = t_{r:m}$ with probability mass function

$$\Pr(\Delta = \delta) = p_\delta^{(j:m)}, \delta = 0, 1, \dots, n - 1;$$

- ii. Generate δ random variates from the conditional distribution in Eq. (7) with the point of truncation $t_k = t_{r:m}$ and $\boldsymbol{\theta} = \boldsymbol{\theta}^{(h)}$, say $\tilde{x}_{j1}, \tilde{x}_{j2}, \dots, \tilde{x}_{j\delta}$;

- iii. Generate $n - \delta$ random variates from the conditional distribution in Eq. (8) with the point of truncation $t_k = t_{r:m}$ and $\theta = \theta^{(h)}$, say $\tilde{x}_{j(\delta+1)}, \tilde{x}_{j(\delta+2)}, \dots, \tilde{x}_{jn}$;
- iv. The pseudo-complete sample for j -th ranked censored system is then

$$\tilde{\mathbf{x}}_j = (\tilde{x}_{j1}, \tilde{x}_{j2}, \dots, \tilde{x}_{jn});$$

- v. Repeat Steps i – iv until we have the imputed samples $\tilde{\mathbf{x}}_j, j = r + 1, r + 2, \dots, m$.

Then, the **M-Step** proceeds exactly the same way as detailed earlier in Sect. 3.1.1.

3.2.2 Impute the censored system lifetimes (SEM-II)

Another way of implementing the SEM algorithm is by implementing the unobserved component lifetimes based on a two-stage procedure. In the first stage, we generate the $(m - r)$ censored system lifetimes from a left-truncated distribution with c.d.f.

$$H_T(y; \theta) = \frac{F_T(y; \theta) - F_T(\tau; \theta)}{1 - F_T(\tau; \theta)}, y > \tau. \tag{9}$$

Here, $\tau = t_{r:m}$ is the censoring time of the experiment. Then, we order these $(m - r)$ samples to obtain $y_{r+1:m} < y_{r+2:m} < \dots < y_{m:m}$. In the second stage, we treat the sample $(t_{1:m} < t_{2:m} < \dots < t_{r:m} < y_{r+1:m} < y_{r+2:m} < \dots < y_{m:m})$ as a pseudo-complete system lifetime data. Since these pseudo-complete system lifetimes are naturally ordered, we consider the implementation of the SEM algorithm based on the ordered system signature.

4 Monte Carlo simulation study

In this section, Monte Carlo simulation is used to evaluate the performance of the SEM algorithm in approximating the MLEs of the parameters in the component lifetime distribution based on complete and censored system-level lifetime data. The performance of different implementations of the SEM algorithm are also compared. Two popular statistical models in lifetime analysis—the Weibull and BS distributions—are considered here for the Monte Carlo evaluation process.

Assume that the component lifetime X follows a Weibull distribution, with its c.d.f. as

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

where α is the shape parameter and β is the scale parameter. Since it is convenient to work with location-scale model, we consider the log-lifetimes of the components $W = \ln X$ in the analysis and W then follows the smallest extreme-value (SEV) distribution with c.d.f.

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

with location parameter $\mu = \ln \beta$ and scale parameter $\sigma = 1/\alpha$. Without loss of generality, we set the location parameter $\mu = 0$ and the scale parameter $\sigma = 1$ in the simulation study. For details on the estimation of parameters in Weibull and SEV distributions, one may refer to [Meeker and Escobar \(1998\)](#) and [Nelson \(2005\)](#).

The BS distribution was originally proposed to model failures due to cracks, and is known as the fatigue life distribution. The BS distribution can be used to model failure times of components in general as well. If the component lifetime X follows a BS distribution, the c.d.f. of X is

$$F_X(x; \gamma, \kappa) = \Phi \left\{ \frac{1}{\gamma} \left[\left(\frac{x}{\kappa} \right)^{1/2} - \left(\frac{\kappa}{x} \right)^{1/2} \right] \right\}, \quad x > 0, \gamma, \kappa > 0,$$

where $\Phi(\cdot)$ is the standard normal c.d.f., γ is the shape parameter and κ is the scale parameter. We set the shape parameter $\gamma = 0.2, 0.5, 1$ and scale parameter $\kappa = 1$ in the simulation study. The MLE of the parameters in BS distribution have been discussed extensively in the literature, see, for example, [Balakrishnan and Zhu \(2014\)](#), [Ng et al. \(2003\)](#) and the references therein.

Three different systems with different system signatures are considered in the simulation study:

- sys1: 4-component series-parallel III system with system signature (1/4, 1/4, 1/2, 0);
- sys2: 4-component mixed parallel I system with system signature (0, 1/2, 1/4, 1/4);
- sys3: 3-component parallel-series system with system signature (0, 2/3, 1/3).

These systems were studied in detail and named as such by [Navarro et al. \(2007\)](#). For each setting, if a complete sample is generated, the estimates were obtained by SEM algorithm based on the ordinary system signature (SEM-SS), SEM algorithm based on ordered system signature (SEM-OSS) and the direct numerical MLEs. For censored sample, the estimates were obtained by SEM algorithm with unobserved component lifetimes imputed directly (SEM-I) and SEM algorithm with unobserved component lifetimes generated indirectly (SEM-II), as well as the direct numerical MLEs. Initial values were set as $\mu^{(0)} = 2$ and $\sigma^{(0)} = 2$ for SEV distribution and $\gamma^{(0)} = 3$ and $\kappa^{(0)} = 3$ for BS distribution, which were intentionally chosen to be far away from the true values in order to examine the sensitivity of SEM algorithm to the initial values. According to some reports (see, for example, [Ye and Ng 2014](#)), a burn-in period of 100 cycles is long enough, while an additional 900 iterations is sufficient to estimate θ well. [Ye and Ng \(2014\)](#) also suggested to use a trace plot of the $\{\theta^{(h)}\}$ sequence versus the iterations for checking the sufficiency of the burn-in, and determining a more appropriate burn-in duration, if necessary. Therefore, in each SEM cycle, the length of the parameter chain was set to be 1000. Bias and mean squared error (MSE) for each estimation method were estimated based on $Y = 10,000$ realizations. Let $\tilde{\theta}_{(y)}$ be the SEM estimate in the y -th realization. Then the estimated bias is calculated as $\widehat{\text{Bias}} = \sum_{y=1}^Y (\tilde{\theta}_{(y)} - \theta) / Y$, and the estimated MSE is calculated as $\widehat{\text{MSE}} = \sum_{y=1}^Y (\tilde{\theta}_{(y)} - \theta)^2 / Y$. In addition, for evaluating the performance of SEM algorithm in approximating the MLEs, the differences between the estimators based

on SEM algorithm and MLEs were calculated in each simulation. Let $\hat{\theta}_{(y)}$ be the MLE in the y -th realization. Then, the estimated difference between SEM estimate and the MLE was calculated as $\widehat{\text{Diff.}} = \sum_{y=1}^Y (\tilde{\theta}_{(y)} - \hat{\theta}_{(y)})/Y$, and the estimated mean squared differences (MSD) is calculated as $\widehat{\text{MSD}} = \sum_{y=1}^Y (\tilde{\theta}_{(y)} - \hat{\theta}_{(y)})^2/Y$. Note that the numerical MLEs here were computed by using direct optimization via Nelder–Mead algorithm (Lange 2001) with the estimate obtained from SEM as the initial value.

4.1 Complete system lifetime data

For the case when complete system lifetime data are observed, we consider the sample sizes $m = 10, 20, 30, 50, 100$. The estimated bias and MSEs of the estimates obtained from two different SEM algorithms and the direct numerical MLEs of the location and scale parameters in SEV distribution are all summarized in Table 3. Also the differences between the estimates obtained from SEM algorithms and the direct numerical MLEs are presented in Table 5. The results for the estimation of the shape parameter $\gamma = 1$ and $\kappa = 1$ in the BS distribution are summarized in Tables 4 and 6. Similar results were also obtained when the shape parameter $\gamma = 0.2, 0.5$ and $\kappa = 1$ in BS distribution. For the sake of brevity, we only present the simulation results for the case $\gamma = 1$ and $\kappa = 1$ here.

From Tables 3 and 4, the performance of SEM-OSS and MLE are seen to be quite close, while the bias and MSEs of SEM-SS tend to be larger than those of the numerical MLEs. For SEM-OSS, the bias and MSE decrease as the sample size increases. From Tables 5 and 6, we can further observe that the SEM algorithm based on the ordered system signature provides a close approximation to the direct numerical MLEs. As the sample size increases, the SEM-OSS estimates become quite close to the MLEs. Although the ordered system signature is calculated based on the ordinary system signature only, it provides more information. A possible reason for the SEM-OSS performing better than SEM-SS is that the imputations of unobserved component lifetimes from the ordered system signature give values distributed more evenly over the whole distribution compared to the imputations based on the ordinary system signature.

4.2 Type-II censored system lifetime data

For Type-II censored system lifetime data, we consider the sample sizes $m = 10, 30$ and 50, with censoring proportion $q = 10\%, 20\%, 30\%, 40\%, 50\%$ for $m = 10$; $q = 20\%, 30\%, 40\%, 50\%$ for $m = 30$; and $q = 20\%, 30\%$ for $m = 50$. Other simulation settings are exactly the same as in the case of complete system lifetime data. The estimated bias and MSEs of the estimates obtained from the two different approaches of implementing the SEM algorithm (SEM-I and SEM-II) and the direct numerical MLEs are presented in Tables 9 and 10 for the SEV and BS distributions, respectively. To study the effectiveness of the SEM algorithm in approximating the MLE, the differences between the SEM estimates and the numerical MLEs in each iteration were computed, and summarized in Tables 7 and 8 for the SEV and BS distributions, respectively.

Table 3 Estimated bias and MSEs for complete sample based on SEM-SS, SEM-OSS, direct numerical MLEs, with different signatures and the underlying component lifetime distribution is SEV distribution with $\mu = 0$ and $\sigma = 1$

<i>s</i>	<i>m</i>	μ						σ												
		SEM-SS			SEM-OSS			MLE			SEM-SS			SEM-OSS			MLE			
		Bias	MSE	MSE	Bias	MSE	MSE	Bias	MSE	MSE	Bias	MSE	MSE	Bias	MSE	MSE	Bias	MSE	MSE	
sys1	10	0.067	0.078	0.076	-0.046	0.076	0.076	0.210	0.170	0.076	0.076	0.076	0.076	0.076	0.076	0.076	-0.073	0.074	0.074	0.074
	20	0.107	0.047	0.036	-0.017	0.036	0.036	0.292	0.153	0.036	0.036	0.036	0.036	0.036	0.036	0.036	-0.042	0.036	0.036	0.036
	30	0.117	0.037	0.024	-0.013	0.024	0.024	0.330	0.156	0.024	0.024	0.024	0.024	0.024	0.024	0.024	0.026	0.024	0.024	0.024
	50	0.123	0.030	0.015	-0.010	0.015	0.015	0.360	0.158	0.015	0.015	0.015	0.015	0.015	0.015	0.015	-0.013	0.014	0.014	0.014
	100	0.132	0.025	0.007	-0.003	0.007	0.007	0.378	0.157	0.007	0.007	0.007	0.007	0.007	0.007	0.007	-0.008	0.007	0.007	0.007
sys2	10	0.057	0.078	0.075	0.031	0.075	0.075	0.191	0.121	0.075	0.075	0.075	0.075	0.075	0.075	0.075	-0.080	0.057	0.057	0.057
	20	0.071	0.044	0.040	-0.014	0.040	0.040	0.283	0.125	0.040	0.040	0.040	0.040	0.040	0.040	0.040	-0.041	0.027	0.027	0.027
	30	0.073	0.031	0.026	-0.011	0.026	0.026	0.319	0.133	0.026	0.026	0.026	0.026	0.026	0.026	0.026	-0.025	0.018	0.018	0.018
	50	0.076	0.021	0.015	-0.006	0.015	0.015	0.344	0.138	0.015	0.015	0.015	0.015	0.015	0.015	0.015	-0.016	0.011	0.011	0.011
	100	0.079	0.014	0.008	-0.003	0.008	0.008	0.364	0.142	0.008	0.008	0.008	0.008	0.008	0.008	0.008	-0.009	0.005	0.005	0.005
sys3	10	0.021	0.070	0.070	-0.009	0.070	0.070	0.029	0.069	0.070	0.070	0.070	0.070	0.070	0.070	0.070	-0.080	0.062	0.062	0.062
	20	0.022	0.035	0.000	-0.006	0.035	0.035	0.105	0.047	0.035	0.035	0.035	0.035	0.035	0.035	0.035	-0.037	0.029	0.029	0.029
	30	0.022	0.023	-0.001	0.023	0.023	0.023	0.131	0.042	0.023	0.023	0.023	0.023	0.023	0.023	0.023	-0.024	0.019	0.019	0.019
	50	0.024	0.014	0.001	-0.002	0.014	0.014	0.148	0.037	0.014	0.014	0.014	0.014	0.014	0.014	0.014	-0.016	0.011	0.011	0.011
	100	0.024	0.007	0.000	-0.001	0.007	0.007	0.166	0.035	0.007	0.007	0.007	0.007	0.007	0.007	0.007	-0.006	0.005	0.005	0.005

Table 4 Estimated bias and MSEs for complete sample based on SEM-SS, SEM-OSS, direct numerical MLEs, with different signatures and the underlying component lifetime distribution is BS distribution with $\gamma = 1$ and $\kappa = 1$

<i>s</i>	<i>m</i>	γ						κ												
		SEM-SS			SEM-OSS			MLE			SEM-SS			SEM-OSS			MLE			
		Bias	MSE	MSE	Bias	MSE	MSE	Bias	MSE	MSE	Bias	MSE	MSE	Bias	MSE	MSE	Bias	MSE	MSE	
sys1	10	0.212	0.133	0.056	-0.067	0.056	0.057	-0.086	0.057	0.033	0.062	0.062	0.017	0.060	0.060	0.022	0.060	0.060	0.022	0.060
	20	0.301	0.139	0.027	-0.030	0.027	0.028	-0.042	0.028	0.026	0.029	0.028	0.012	0.028	0.028	0.014	0.028	0.028	0.014	0.028
	30	0.330	0.142	0.018	-0.020	0.018	0.018	-0.029	0.018	0.018	0.020	0.018	0.005	0.020	0.019	0.007	0.020	0.019	0.007	0.020
	50	0.357	0.147	0.011	-0.011	0.011	0.011	-0.017	0.011	0.017	0.012	0.011	0.004	0.012	0.011	0.005	0.012	0.011	0.005	0.011
	100	0.376	0.152	0.005	-0.005	0.005	0.005	-0.008	0.005	0.013	0.006	0.006	0.001	0.006	0.006	0.002	0.006	0.006	0.002	0.006
sys2	10	0.218	0.134	0.055	-0.061	0.055	0.056	-0.080	0.056	0.050	0.069	0.054	0.054	0.070	0.065	0.033	0.070	0.065	0.033	0.070
	20	0.303	0.139	0.027	-0.028	0.027	0.027	-0.041	0.027	0.017	0.030	0.025	0.025	0.031	0.030	0.014	0.031	0.030	0.014	0.030
	30	0.334	0.145	0.018	-0.017	0.018	0.018	-0.026	0.018	0.010	0.019	0.019	0.019	0.020	0.019	0.011	0.020	0.019	0.011	0.020
	50	0.356	0.147	0.011	-0.012	0.011	0.011	-0.017	0.011	-0.001	0.011	0.011	0.009	0.012	0.011	0.005	0.012	0.011	0.005	0.011
	100	0.376	0.152	0.005	-0.006	0.005	0.005	-0.008	0.005	-0.004	0.006	0.006	0.007	0.006	0.006	0.005	0.006	0.006	0.005	0.006
sys3	10	0.044	0.069	0.058	-0.085	0.058	0.058	-0.080	0.058	0.059	0.067	0.067	0.067	0.069	0.066	0.051	0.069	0.066	0.051	0.066
	20	0.111	0.047	0.028	-0.041	0.028	0.028	-0.041	0.028	0.025	0.030	0.036	0.036	0.032	0.031	0.027	0.032	0.031	0.027	0.031
	30	0.134	0.042	0.018	-0.028	0.018	0.018	-0.028	0.018	0.010	0.019	0.021	0.021	0.020	0.020	0.016	0.020	0.020	0.016	0.020
	50	0.155	0.039	0.011	-0.015	0.011	0.011	-0.015	0.011	-0.002	0.011	0.010	0.010	0.011	0.011	0.007	0.011	0.011	0.007	0.011
	100	0.169	0.036	0.006	-0.007	0.006	0.006	-0.007	0.006	-0.006	0.005	0.007	0.007	0.006	0.006	0.005	0.006	0.006	0.005	0.006

Table 5 Estimated distance between SEM estimates and direct numerical MLEs, with different signatures and the underlying component lifetime distribution is SEV distribution with $\mu = 0$ and $\sigma = 1$

<i>s</i>	<i>m</i>	μ		σ					
		SEM-SS versus MLE		SEM-SS versus MLE					
		Difference	MSE	Difference	MSE				
sys1	10	0.1135	0.0155	0.0009	0.0004	0.2829	0.0954	-0.0015	0.0010
	20	0.1246	0.0168	0.0006	0.0002	0.3342	0.1219	0.0003	0.0006
	30	0.1293	0.0176	0.0003	0.0001	0.3562	0.1344	0.0009	0.0004
	50	0.1330	0.0182	0.0002	0.0001	0.3730	0.1440	0.0006	0.0003
	100	0.1353	0.0186	0.0001	0.0000	0.3854	0.1511	0.0004	0.0001
sys2	10	-0.0399	0.0020	-0.0145	0.0004	-0.0142	0.0007	-0.0327	0.0015
	20	0.0851	0.0077	0.0123	0.0006	0.3244	0.1084	0.0047	0.0003
	30	0.0840	0.0074	0.0080	0.0004	0.3447	0.1211	0.0040	0.0002
	50	0.0829	0.0071	0.0050	0.0002	0.3603	0.1313	0.0028	0.0001
	100	0.0820	0.0068	0.0025	0.0001	0.3724	0.1395	0.0016	0.0000
sys3	10	0.0308	0.0012	0.0112	0.0002	0.1092	0.0131	-0.0150	0.0006
	20	0.0281	0.0009	0.0063	0.0001	0.1425	0.0211	-0.0061	0.0002
	30	0.0271	0.0008	0.0041	0.0001	0.1547	0.0245	-0.0034	0.0001
	50	0.0261	0.0007	0.0026	0.0000	0.1642	0.0273	-0.0017	0.0000
	100	0.0255	0.0007	0.0012	0.0000	0.1723	0.0299	-0.0007	0.0000

Table 6 Estimated distance between SEM estimates and direct numerical MLEs, with different signatures and the underlying component lifetime distribution is BS distribution with $\gamma = 1$ and $\kappa = 1$

<i>s</i>	<i>m</i>	γ	κ						
			SEM-SS versus MLE			SEM-OSS versus MLE			
			Difference	MSD	MSS	Difference	MSD	MSS	
sys1	10	0.2976	0.0956	0.0193	0.0009	0.0107	0.0006	-0.0046	0.0010
	20	0.3434	0.1224	0.0121	0.0004	0.0116	0.0004	-0.0026	0.0006
	30	0.3594	0.1324	0.0090	0.0003	0.0117	0.0003	-0.0019	0.0004
	50	0.3735	0.1415	0.0057	0.0002	0.0118	0.0003	-0.0012	0.0003
	100	0.3843	0.1488	0.0031	0.0001	0.0119	0.0002	-0.0006	0.0001
sys2	10	0.2986	0.0961	0.0193	0.0009	0.0171	0.0009	0.0215	0.0015
	20	0.3437	0.1226	0.0123	0.0004	0.0036	0.0003	0.0112	0.0007
	30	0.3604	0.1331	0.0092	0.0003	-0.0012	0.0002	0.0077	0.0005
	50	0.3736	0.1416	0.0058	0.0002	-0.0056	0.0002	0.0044	0.0003
	100	0.3844	0.1488	0.0030	0.0001	-0.0088	0.0001	0.0022	0.0002
sys3	10	0.1248	0.0175	-0.0041	0.0003	0.0073	0.0002	0.0157	0.0005
	20	0.1520	0.0244	-0.0002	0.0001	-0.0025	0.0001	0.0084	0.0002
	30	0.1618	0.0271	0.0002	0.0001	-0.0061	0.0001	0.0056	0.0001
	50	0.1701	0.0295	0.0006	0.0001	-0.0089	0.0001	0.0034	0.0001
	100	0.1761	0.0313	0.0003	0.0000	-0.0112	0.0001	0.0018	0.0000

Table 7 Estimated distance between SEM and direct numerical MLEs, for censored sample with different signatures and the underlining component lifetime distribution is SEV distribution with $\mu = 0$ and $\sigma = 1$

s	m	q (%)	μ	σ						
				SEM-I versus MLE			SEM-II versus MLE			
				Diff.	MSD	MSD	Diff.	MSD	MSD	
sys1	10	10	-0.0187	0.0008	-0.0069	0.0005	0.0098	0.0013	0.0147	0.0013
			-0.0312	0.0017	-0.0077	0.0006	-0.0007	0.0011	0.0014	0.0009
		20	-0.0496	0.0038	-0.0173	0.0011	-0.0135	0.0013	-0.0231	0.0014
			-0.0753	0.0082	-0.0412	0.0033	-0.0307	0.0023	-0.0607	0.0050
		30	-0.1119	0.0176	-0.0869	0.0168	-0.0540	0.0050	-0.1130	0.0174
			-0.0232	0.0008	-0.0015	0.0002	-0.0064	0.0005	-0.0065	0.0004
	30	20	-0.0399	0.0020	-0.0145	0.0004	-0.0142	0.0007	-0.0327	0.0015
			-0.0632	0.0048	-0.0467	0.0025	-0.0251	0.0013	-0.0768	0.0065
		40	-0.0941	0.0104	-0.1134	0.0136	-0.0383	0.0024	-0.1439	0.0220
			-0.0214	0.0006	-0.0001	0.0001	-0.0075	0.0004	-0.0080	0.0003
		50	-0.0379	0.0017	-0.0139	0.0003	-0.0144	0.0005	-0.0347	0.0015
			-0.0100	0.0004	0.0169	0.0005	-0.0111	0.0008	0.0140	0.0012
sys2	10	20	-0.0451	0.0024	0.0192	0.0007	-0.0419	0.0024	0.0023	0.0012
			-0.0842	0.0081	0.0175	0.0011	-0.0690	0.0057	-0.0176	0.0020
		30	-0.1204	0.0168	0.0098	0.0029	-0.0928	0.0104	-0.0474	0.0055
			-0.1505	0.0272	-0.0030	0.0225	-0.1138	0.0161	-0.0838	0.0233

Table 7 continued

s	m	q (%)	μ	σ						
				SEM-I versus MLE		SEM-II versus MLE				
				Diff.	MSD	Diff.	MSD			
30	20	30	-0.0442	0.0021	0.0157	0.0003	-0.0361	0.0015	0.0039	0.0004
			-0.0804	0.0068	0.0158	0.0005	-0.0560	0.0034	0.0008	0.0008
			-0.1126	0.0134	0.0066	0.0006	-0.0699	0.0054	0.0032	0.0032
			-0.1356	0.0196	-0.0230	0.0074	-0.0792	0.0070	0.0165	0.0165
			-0.0436	0.0020	0.0146	0.0003	-0.0347	0.0013	0.0003	0.0003
50	30	10	-0.0786	0.0064	0.0152	0.0004	-0.0524	0.0029	-0.0164	0.0006
			-0.0185	0.0004	0.0043	0.0001	-0.0261	0.0010	0.0004	0.0004
			-0.0435	0.0021	0.0057	0.0001	-0.0491	0.0029	0.0009	0.0009
			-0.0673	0.0051	0.0033	0.0002	-0.0680	0.0054	0.0027	0.0027
			-0.0875	0.0089	-0.0076	0.0006	-0.0860	0.0089	0.0078	0.0078
30	20	40	-0.1061	0.0136	-0.0327	0.0045	-0.1065	0.0141	-0.1322	0.0217
			-0.0414	0.0018	0.0087	0.0001	-0.0370	0.0015	0.0003	0.0003
			-0.0644	0.0043	0.0074	0.0001	-0.0491	0.0026	0.0013	0.0013
			-0.0813	0.0070	-0.0035	0.0002	-0.0569	0.0036	0.0053	0.0053
			-0.0931	0.0093	-0.0335	0.0019	-0.0649	0.0048	0.0184	0.0184
50	20	30	-0.0407	0.0017	0.0095	0.0001	-0.0341	0.0012	-0.0057	0.0002
			-0.0635	0.0041	0.0082	0.0001	-0.0442	0.0021	0.0010	0.0010

Table 8 Estimated distance between SEM and direct numerical MLEs, for censored sample with different signatures and the underlying component lifetime distribution is BS distribution with $\gamma = 1$ and $\kappa = 1$

s	m	q (%)	γ			κ										
			SEM-I versus MLE			SEM-II versus MLE			SEM-I versus MLE			SEM-II versus MLE				
			Diff.	MSD		Diff.	MSD		Diff.	MSD		Diff.	MSD			
sys1	10	10	0.0138	0.0008	0.0180	0.0008	0.0008	-0.0089	0.0011	-0.0026	0.0010					
			0.0070	0.0008	0.0139	0.0007	0.0015	0.0012	-0.0151	0.0015	-0.0003	0.0012				
			-0.0017	0.0009	0.0078	0.0007	0.0021	0.0015	-0.0226	0.0021	0.0041	0.0015				
			-0.0129	0.0014	-0.0019	0.0012	0.0033	0.0026	-0.0330	0.0033	0.0088	0.0026				
			-0.0265	0.0029	-0.0078	0.0071	0.0056	0.0278	-0.0445	0.0056	0.0313	0.0278				
	30	20	20	-0.0032	0.0003	0.0004	0.0002	0.0007	-0.0137	0.0007	-0.0016	0.0004				
				-0.0118	0.0005	-0.0101	0.0003	0.0012	0.0005	-0.0244	0.0012	-0.0034	0.0005			
				-0.0220	0.0010	-0.0256	0.0010	0.0022	0.0006	-0.0365	0.0022	-0.0051	0.0006			
				-0.0344	0.0020	-0.0478	0.0032	0.0041	0.0016	-0.0524	0.0041	-0.0083	0.0016			
				-0.0067	0.0002	-0.0038	0.0001	0.0005	0.0003	-0.0134	0.0005	-0.0018	0.0003			
	50	20	30	-0.0152	0.0004	-0.0157	0.0004	0.0010	-0.0241	0.0010	-0.0046	0.0003				
				-0.0091	0.0005	0.0156	0.0009	0.0005	0.0010	0.0040	0.0005	0.0221	0.0010			
				-0.0398	0.0021	0.0116	0.0009	0.0008	0.0009	-0.0197	0.0008	0.0244	0.0009			
				-0.0668	0.0054	0.0111	0.0010	0.0028	0.0014	-0.0456	0.0028	0.0302	0.0014			
				-0.0875	0.0095	0.0149	0.0021	0.0059	0.0037	-0.0665	0.0059	0.0434	0.0037			
sys2	10	50	-0.1026	0.0139	0.0345	0.0133	0.0093	-0.0806	0.0093	0.0826	0.0256					

Table 8 continued

s	m	q (%)	γ				κ			
			SEM-I versus MLE		SEM-II versus MLE		SEM-I versus MLE		SEM-II versus MLE	
			Diff.	MSD	Diff.	MSD	Diff.	MSD	Diff.	MSD
sys3	30	20	-0.0379	0.0016	0.0057	0.0003	-0.0264	0.0009	0.0107	0.0002
			-0.0614	0.0041	0.0065	0.0004	-0.0501	0.0028	0.0167	0.0004
			-0.0781	0.0067	0.0109	0.0009	-0.0706	0.0055	0.0283	0.0011
			-0.0865	0.0084	0.0233	0.0041	-0.0830	0.0078	0.0531	0.0054
	50	20	-0.0375	0.0015	0.0028	0.0002	-0.0272	0.0008	0.0077	0.0001
			-0.0596	0.0037	0.0040	0.0002	-0.0505	0.0027	0.0136	0.0002
			-0.0291	0.0011	-0.0100	0.0004	0.0000	0.0002	0.0153	0.0004
			-0.0509	0.0030	-0.0160	0.0006	-0.0176	0.0005	0.0164	0.0004
	10	30	-0.0667	0.0052	-0.0217	0.0008	-0.0331	0.0014	0.0183	0.0005
			-0.0789	0.0076	-0.0297	0.0013	-0.0439	0.0025	0.0220	0.0009
			-0.0870	0.0098	-0.0391	0.0026	-0.0495	0.0033	0.0284	0.0024
			-0.0408	0.0018	-0.0078	0.0002	-0.0243	0.0007	0.0073	0.0001
30	20	-0.0547	0.0032	-0.0121	0.0003	-0.0402	0.0018	0.0102	0.0001	
		-0.0625	0.0043	-0.0186	0.0006	-0.0509	0.0028	0.0146	0.0003	
		-0.0658	0.0049	-0.0290	0.0014	-0.0552	0.0034	0.0213	0.0008	
		-0.0379	0.0015	-0.0064	0.0001	-0.0255	0.0007	0.0050	0.0000	
50	20	-0.0513	0.0027	-0.0107	0.0002	-0.0414	0.0018	0.0082	0.0001	

Table 9 Estimated bias and MSEs for censored sample based on SEM-I, SEM-II, MLEs, with different signatures and the underlying component lifetime distribution is SEV distribution with $\mu = 0$ and $\sigma = 1$

s	m	q (%)	μ	σ													
				μ						σ							
				SEM-I		SEM-II		MLE		SEM-I		SEM-II		MLE			
				Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
sys1	10	10	-0.077	0.089	0.087	-0.065	0.089	-0.059	0.085	-0.079	0.090	-0.074	0.092	-0.089	0.089		
				0.107	0.103	-0.092	0.101	-0.084	-0.111	0.106	0.110	-0.109	0.110	-0.110	0.105		
				0.140	0.134	-0.131	0.128	-0.114	-0.141	0.128	0.136	-0.151	0.136	-0.128	0.127		
				0.203	0.199	-0.206	0.180	-0.164	-0.186	0.155	0.174	-0.216	0.174	-0.155	0.153		
	30	20	30	-0.351	0.329	0.354	-0.326	0.285	-0.240	-0.244	0.195	-0.303	0.241	-0.190	0.191		
					0.033	0.031	-0.033	0.031	-0.031	-0.042	0.035	0.037	-0.043	0.037	-0.036	0.035	
					0.042	0.039	-0.050	0.038	-0.036	-0.057	0.042	0.046	-0.075	0.046	-0.042	0.041	
					0.061	0.060	-0.101	0.052	-0.054	-0.075	0.051	0.061	-0.127	0.061	-0.050	0.050	
	50	20	50	-0.167	0.094	0.105	-0.186	0.077	-0.073	-0.101	0.063	-0.206	0.091	-0.063	0.061		
					0.019	0.018	-0.017	0.018	-0.017	-0.028	0.021	0.022	-0.028	0.022	-0.020	0.021	
					0.025	0.023	-0.035	0.022	-0.022	-0.038	0.026	0.029	-0.058	0.029	-0.024	0.025	
					0.085	0.083	-0.022	0.084	-0.039	-0.100	0.070	0.070	-0.075	0.070	-0.089	0.070	
sys2	10	20	-0.094	0.096	0.089	-0.030	0.090	-0.049	-0.144	0.088	-0.099	0.088	-0.102	0.084			
				0.118	0.105	-0.045	0.103	-0.063	-0.186	0.111	0.115	-0.135	0.115	-0.117	0.103		
				0.152	0.134	-0.080	0.125	-0.090	-0.233	0.139	0.154	-0.188	0.154	-0.140	0.125		
				0.218	0.224	-0.145	0.172	-0.142	-0.291	0.179	0.241	-0.261	0.241	-0.177	0.157		

Table 9 continued

<i>s</i>	<i>m</i>	<i>q</i> (%)	μ						σ					
			SEM-I Bias	MSE	SEM-II Bias	MSE	MLE Bias	MSE	SEM-I Bias	MSE	SEM-II Bias	MSE	MLE Bias	MSE
30	20	30	-0.060	0.032	0.000	0.029	-0.015	0.029	-0.066	0.029	-0.026	0.029	-0.030	0.027
			-0.102	0.043	-0.006	0.034	-0.022	0.034	-0.091	0.036	-0.051	0.036	-0.035	0.033
			-0.146	0.058	-0.026	0.042	-0.033	0.040	-0.115	0.046	-0.094	0.050	-0.045	0.040
			-0.179	0.076	-0.067	0.067	-0.044	0.052	-0.139	0.059	-0.165	0.083	-0.059	0.083
50	20	30	-0.055	0.021	0.003	0.018	-0.011	0.018	-0.057	0.018	-0.019	0.017	-0.022	0.016
			-0.089	0.027	0.004	0.020	-0.011	0.020	-0.074	0.023	-0.038	0.022	-0.021	0.020
			-0.034	0.072	-0.011	0.071	-0.015	0.071	-0.112	0.073	-0.092	0.073	-0.086	0.072
			-0.072	0.079	-0.023	0.074	-0.029	0.074	-0.150	0.091	-0.122	0.091	-0.101	0.086
sys3	10	20	-0.113	0.094	-0.043	0.084	-0.046	0.084	-0.184	0.111	-0.161	-0.116	0.104	
			-0.160	0.114	-0.080	0.100	-0.072	0.097	-0.226	0.139	-0.221	0.153	-0.140	0.128
			-0.208	0.150	-0.135	0.141	-0.102	0.126	-0.274	0.175	-0.300	0.214	-0.167	0.159
			-0.050	0.027	0.000	0.024	0.009	0.025	-0.070	0.029	-0.042	0.029	-0.033	0.027
30	20	30	-0.082	0.033	-0.010	0.027	-0.017	0.027	-0.092	0.037	-0.074	0.037	-0.043	0.033
			-0.104	0.040	-0.026	0.031	-0.022	0.031	-0.108	0.046	-0.122	0.053	-0.051	0.042
			-0.131	0.052	-0.071	0.044	-0.038	0.039	-0.123	0.057	-0.190	0.081	-0.059	0.051
			-0.048	0.017	0.002	0.014	-0.007	0.015	-0.055	0.018	-0.026	0.017	-0.020	0.017
50	20	30	-0.073	0.021	-0.001	0.015	-0.009	0.016	-0.069	0.022	-0.054	0.022	-0.024	0.020

Table 10 Estimated bias and MSEs for censored sample based on SEM-I, SEM-II, MLEs, with different signatures and the underlying component lifetime distribution is BS with $\gamma = 1$ and $\kappa = 1$

s	m	q (%)	γ	κ																			
				SEM-I			SEM-II			MLE			SEM-I			SEM-II			MLE				
				Bias	MSE	MSE	Bias	MSE	MSE	Bias	MSE	MSE	Bias	MSE	MSE	Bias	MSE	MSE	Bias	MSE	MSE		
sys1	10	10	10	-0.083	0.066	0.066	-0.079	0.066	0.067	0.067	0.067	0.007	0.062	0.062	0.014	0.063	0.063	0.016	0.063	0.063			
				-0.110	0.079	0.080	-0.103	0.080	0.081	0.081	-0.013	0.064	0.064	0.002	0.066	0.066	0.002	0.066	0.066	0.002	0.065	0.065	
				-0.139	0.098	0.101	-0.130	0.101	0.100	0.100	-0.031	0.071	0.071	-0.004	0.075	0.075	-0.008	0.075	0.075	-0.008	0.073	0.073	
				-0.179	0.123	0.134	-0.168	0.134	0.126	0.126	-0.058	0.080	0.080	-0.017	0.091	0.091	-0.025	0.091	0.091	-0.025	0.084	0.084	
				-0.229	0.161	0.204	-0.210	0.204	0.166	0.166	-0.107	0.100	0.100	-0.031	0.166	0.166	-0.063	0.166	0.166	-0.063	0.107	0.107	
	30	20	20	30	-0.043	0.025	0.026	-0.039	0.026	0.025	0.025	-0.012	0.022	0.022	0.000	0.022	0.022	0.001	0.022	0.022			
					-0.058	0.032	0.033	-0.057	0.033	0.031	0.031	-0.030	0.024	0.024	-0.009	0.024	0.024	-0.006	0.024	0.024	-0.006	0.024	0.024
					-0.079	0.041	0.045	-0.083	0.045	0.040	0.040	-0.046	0.029	0.029	-0.015	0.031	0.031	-0.009	0.031	0.031	-0.009	0.030	0.030
					0.899	0.851	0.837	0.885	0.837	0.919	0.919	-0.070	0.038	0.038	-0.026	0.044	0.044	-0.018	0.044	0.044	-0.018	0.039	0.039
					-0.032	0.015	0.015	-0.029	0.015	0.015	0.015	-0.012	0.013	0.013	0.000	0.013	0.013	0.002	0.013	0.013	0.002	0.013	0.013
	50	20	20	40	-0.046	0.019	0.020	-0.046	0.020	0.018	0.018	-0.025	0.015	0.015	-0.006	0.015	0.015	-0.001	0.015	0.015			
					-0.104	0.069	0.068	-0.079	0.068	0.068	0.068	0.037	0.070	0.070	0.056	0.074	0.074	0.033	0.074	0.074	0.033	0.068	0.068
					-0.143	0.082	0.079	-0.092	0.079	0.079	0.079	0.005	0.067	0.067	0.049	0.076	0.076	0.025	0.076	0.076	0.025	0.070	0.070
					-0.184	0.101	0.098	-0.106	0.098	0.094	0.094	-0.019	0.066	0.066	0.057	0.083	0.083	0.027	0.083	0.083	0.027	0.073	0.073
					-0.230	0.127	0.129	-0.127	0.129	0.116	0.116	-0.055	0.068	0.068	0.055	0.096	0.096	0.012	0.096	0.096	0.012	0.078	0.078
sys2	10	10	50	-0.275	0.160	0.204	-0.138	0.204	0.147	0.147	-0.086	0.077	0.077	0.078	0.154	0.154	-0.005	0.154	0.154				
				-0.172	0.142	0.142	-0.142	0.142	0.142	0.142	-0.055	0.068	0.068	0.055	0.096	0.096	0.012	0.096	0.096	0.012	0.078	0.078	
				-0.172	0.142	0.142	-0.142	0.142	0.142	0.142	-0.055	0.068	0.068	0.055	0.096	0.096	0.012	0.096	0.096	0.012	0.078	0.078	
				-0.172	0.142	0.142	-0.142	0.142	0.142	0.142	-0.055	0.068	0.068	0.055	0.096	0.096	0.012	0.096	0.096	0.012	0.078	0.078	
				-0.172	0.142	0.142	-0.142	0.142	0.142	0.142	-0.055	0.068	0.068	0.055	0.096	0.096	0.012	0.096	0.096	0.012	0.078	0.078	

Table 10 continued

s	m	q (%)	γ						κ					
			SEM-I			SEM-II			SEM-I			SEM-II		
			Bias	MSE	MLE	Bias	MSE	MLE	Bias	MSE	MLE	Bias	MSE	MLE
sys3	30	20	-0.072	0.027	-0.028	0.025	-0.034	0.024	-0.016	0.021	0.023	0.011	0.022	
			-0.100	0.035	-0.032	0.032	-0.038	0.030	-0.042	0.022	0.024	0.008	0.023	
			-0.123	0.044	-0.034	0.044	-0.044	0.037	-0.066	0.025	0.030	0.004	0.025	
			-0.144	0.055	-0.034	0.067	-0.057	0.046	0.086	0.029	0.043	-0.003	0.029	
	50	20	-0.062	0.017	-0.022	0.016	-0.024	0.015	-0.023	0.012	0.013	0.004	0.013	
			-0.085	0.022	-0.021	0.018	-0.025	0.017	-0.046	0.014	0.015	0.005	0.014	
			-0.117	0.068	-0.098	0.067	-0.088	0.067	0.050	0.067	0.070	0.050	0.067	
			-0.154	0.081	-0.120	0.077	-0.104	0.076	0.030	0.066	0.074	0.047	0.070	
	10	30	-0.192	0.101	-0.147	0.096	-0.125	0.092	0.005	0.065	0.075	0.038	0.071	
			-0.227	0.124	-0.178	0.123	-0.148	0.113	-0.019	0.064	0.076	0.025	0.071	
			-0.276	0.160	-0.228	0.170	-0.189	0.146	-0.046	0.067	0.085	0.003	0.074	
			-0.074	0.027	-0.041	0.025	-0.033	0.025	-0.009	0.019	0.021	0.015	0.020	
30	40	-0.092	0.032	-0.049	0.030	-0.037	0.028	-0.026	0.020	0.022	0.014	0.021		
		-0.111	0.040	-0.067	0.039	-0.049	0.035	-0.042	0.022	0.024	0.009	0.022		
		-0.127	0.050	-0.090	0.055	-0.061	0.044	-0.055	0.024	0.027	0.000	0.024		
		-0.062	0.017	-0.030	0.015	-0.024	0.015	-0.017	0.012	0.012	0.009	0.012		
50	30	-0.075	0.021	-0.034	0.018	-0.023	0.017	-0.034	0.012	0.012	0.008	0.012		

Based on these results, it can be seen that the estimates based on SEM-I and SEM-II perform well compared to the numerical MLEs. In general, the SEM-I estimates and SEM-II estimates approximate MLEs very well when the censoring proportion is moderate (say, $30 \leq q < 50\%$). The bias and MSEs of the estimates increase when the censoring proportion increases. When the censoring proportion is small (say, $q \leq 20\%$), the estimates obtained based on SEM-II are seen to be as good as the direct MLEs. SEM-II performs slightly better than SEM-I when the censoring proportion is small in terms of bias, MSE and the distance to numerical MLEs. However, further study shows that when the censoring proportion is large (say, $q \geq 70\%$), the bias of SEM-II indeed becomes larger than that of SEM-I.

5 Illustrative examples

5.1 4-component series-parallel III system with Weibull distributed components

A system lifetime sample generated by Balakrishnan et al. (2011b) is used here to illustrate the proposed SEM algorithm. The sample system lifetime data were generated from a 4-component series-parallel III system (Navarro et al. 2007) with system signature $\mathbf{s} = (1/4, 1/4, 1/2, 0)$ and sample size $m = 10$, and with component lifetimes following a Weibull distribution with scale parameter $\beta = 3$ and shape parameter $\alpha = 2$ (equivalent to the SEV distribution with location parameter $\mu = 1.0986$ and scale parameter $\sigma = 0.5$). These data are presented in Table 11.

We first present the performance of estimates based on SEM-SS and SEM-OSS, MLEs and BLUE when there is no censoring. We implemented the SEM algorithm with starting values $\mu^{(0)} = 0$ and $\sigma^{(0)} = 1$. The first 100 iterations were used as burn-in, and additional 900 iterations were then used to obtain the final parameter estimates. The unobserved component lifetimes were imputed based on the system signature $\mathbf{s} = (1/4, 1/4, 1/2, 0)$ in the S-step of SEM-SS, while the unobserved component lifetimes were imputed based on the ordered system signatures for SEM-OSS. Monte Carlo method was employed to approximate the ordered system signatures. For instance, in this example, the approximated second ordered system signature is $\hat{\mathbf{s}}^{(2:10)} = (0.5577, 0.2865, 0.1558, 0)$ and the approximated 9-th ordered system signature is $\hat{\mathbf{s}}^{(9:10)} = (0.0268, 0.1579, 0.8153, 0)$. The Fisher information matrix and the asymp-

Table 11 Simulated system lifetimes with signature $(1/4, 1/4, 1/2, 0)$, and component distribution Weibull(3,2), taken from Balakrishnan et al. (2011b)

j	1	2	3	4	5
$T_{j:10}$	0.72717	1.02050	1.38633	1.61244	1.70590
$X_{j:10} = \ln T_{j:10}$	-0.13836	0.00881	0.14187	0.20748	0.23195
j	6	7	8	9	10
$T_{j:10}$	1.76789	2.67863	3.02676	3.25943	3.78497
$X_{j:10} = \ln T_{j:10}$	0.24746	0.42791	0.48098	0.51314	0.57806

Table 12 Estimation results from the data in Table 11

q (%)	Method	μ			σ		
		Estimate	Variance	95 % CI	Estimate	Variance	95 % CI
0	SEM-SS	1.0327	0.0312	(0.6862, 1.3792)	0.6360	0.0591	(0.3006, 1.3457)
	SEM-OSS	0.9774	0.0189	(0.7083, 1.2465)	0.5072	0.0192	(0.2970, 0.8660)
	MLE	0.9912	0.0184	(0.7254, 1.2572)	0.4990	0.0181	(0.2940, 0.8468)
	BLUE	1.0141	0.0216	(0.7262, 1.3020)	0.5366	0.0254	(0.2999, 0.9601)
30	SEM-I	0.9462	0.0224	(0.6531, 1.2393)	0.4943	0.0246	(0.2652, 0.9211)
	SEM-II	0.9774	0.0259	(0.6622, 1.2927)	0.5072	0.0276	(0.2668, 0.9642)
	MLE	0.9949	0.0297	(0.6569, 1.3328)	0.5292	0.0331	(0.2699, 1.0377)
	BLUE	1.0623	0.0499	(0.6243, 1.4798)	0.5985	0.0584	(0.2713, 1.3202)

otic variance-covariance matrix of the estimates were obtained by substituting the SEM estimates into the corresponding formulas presented in Zhang et al. (2015). 95 % confidence intervals of the parameters were calculated according to Equations (5) and (8) in Zhang et al. (2015). The point and interval estimates obtained by different methods are presented in Table 12. Again, this example shows that the estimates obtained from SEM-OSS are very close to the direct numerical MLEs. The variances of the estimates based on SEM-OSS are seen to be the smallest among the four methods.

Second, we consider the case of Type-II censoring with censoring proportion $q = 30\%$. Balakrishnan et al. (2011b) and Zhang et al. (2015) considered this setting and obtained the BLUEs and MLEs of the model parameters. The evolution paths of SEM estimates based on the direct imputation of component lifetimes (SEM-I) and based on the two-stage procedure by first imputing the censored system lifetimes (SEM-II) are shown in Figs. 2 and 3, respectively. The SEM cycles oscillate around the horizontal lines which indicate the MLEs but do not show any upward or downward trend. This reveals that chains $\theta^{(h)}$ have converged to a stationary distribution, similar to the results obtained in Nielsen (2000). It is sufficient to use the average of the chains to estimate the parameter, and the estimate will be close to the MLE.

The estimates obtained from the proposed SEM algorithms are presented in Table 12. In this case, it is clear that estimates based on SEM-I and SEM-II approximate the MLEs very well and the variance of the estimates based on SEM algorithms are also smaller.

5.2 Consecutive 2-out-of-8: F system with Birnbaum–Saunders distributed components

The phosphor acid filter system (Frenkel and Khvatskin 2006) as described in Sect. 2 is a real life prototype for n -component system. It is a consecutive 2-out-of- n : F system, in which, the failure of any 2 adjacent components failure will lead to the system failure. For illustrative purpose, we provide here a numerical example based on the consecutive 2-out-of-8: F system. If the components in the system have equal chance of failure, then the system signature is $s = (0, 1/4, 11/28, 2/7, 1/14, 0, 0, 0)$.

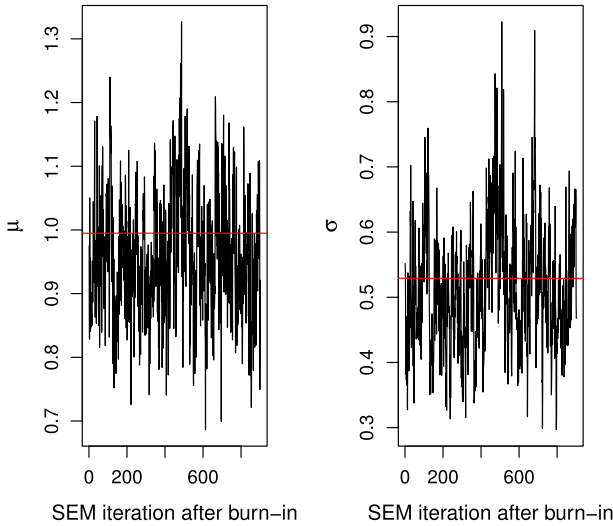


Fig. 2 Evolution paths of the parameter values via the SEM-I

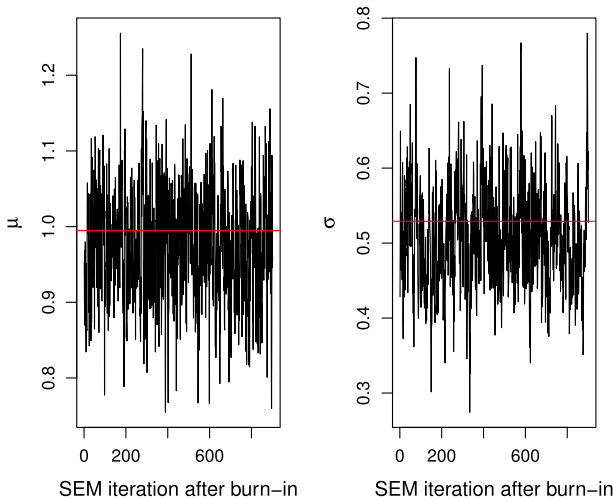


Fig. 3 Evolution paths of the parameter values via the SEM-II

We simulated the lifetimes of $m = 20$ consecutive 2-out-of-8: F systems in which the component lifetimes follow the BS distribution with shape parameter $\gamma = 1$ and scale parameter $\kappa = 1$. The data are presented in Table 13. As in the previous example, SEM-SS algorithm and SEM-OSS algorithm were employed to compare with the MLEs. The first 100 iterations were used as burn-in, and additional 900 iterations were used to obtain the final parameter estimates. Then, the case of Type-II censoring with censoring proportion $q = 20\%$ was considered. The estimates based on SEM-I algorithm and SEM-II algorithm were obtained. The required ordered system signature was obtained by Monte Carlo simulation. The variance of the parameter estimates for both cases

Table 13 Consecutive 2-out-of-8 system lifetimes with signature (0, 1/4, 11/28, 2/7, 1/14, 0, 0, 0), and component distribution BS(1,1)

j	1	2	3	4	5	6	7	8	9	10
$T_{j:20}$	0.2598	0.2803	0.3329	0.4172	0.4532	0.459	0.5541	0.5769	0.5842	0.7784
j	11	12	13	14	15	16	17	18	19	20
$T_{j:20}$	0.7917	0.8565	0.8895	0.9186	0.9348	1.1130	1.2049	1.3938	1.4406	1.6351

Table 14 Estimation results from the data in Table 13

q	Method	γ			κ		
		Estimate	Variance	95 % C.I	Estimate	Variance	95 % C.I
0 %	SEM-SS	1.2263	0.0795	(0.7814, 1.9246)	1.1277	0.0396	(0.7981, 1.5934)
	SEM-OSS	1.0137	0.0284	(0.7319, 1.4039)	1.0586	0.0201	(0.8142, 1.3763)
	MLE	1.0190	0.0290	(0.7342, 1.4141)	1.0559	0.0200	(0.8124, 1.3725)
20 %	SEM-I	1.0575	0.0414	(0.7253, 1.5420)	1.0741	0.0270	(0.7958, 1.4496)
	SEM-II	1.0137	0.0338	(0.7103, 1.4467)	1.0586	0.0235	(0.7970, 1.4061)
	MLE	1.0949	0.0500	(0.7337, 1.6340)	1.1090	0.0346	(0.7982, 1.5407)

were obtained by inverting the observed Fisher information matrix, according to the formulas given in Zhang et al. (2015). 95 % confidence intervals of the parameters were then calculated according to Equation (8) in Zhang et al. (2015).

The estimates obtained from the SEM algorithms are presented in Table 14. For complete sample case, the estimates obtained from SEM algorithm based on ordered system signature are very close to the direct numerical MLEs, both in the case of point and interval estimation. For the Type-II censoring case with censoring proportion 20 %, the two approaches of the SEM algorithm do provide good approximation to the direct MLEs. This numerical example demonstrates that the proposed SEM algorithm works very well for different values of n and different lifetime distributions.

5.3 Software

All the computations were conducted in R version 3.1.1 (R Core Team 2014). Code in R for different approaches to implement the SEM algorithm and the computation of the MLEs are available in the supplementary material. Detailed descriptions of these R programs and examples based on the numerical examples in Sects. 5.1 and 5.2 are also provided in the supplementary material. The results in Tables 12 and 14 were generated using a specific random seed value (see supplementary material). It should be noted that a change in the random seed value will result in a change in the estimates and confidence intervals. However, since the results are based on 1000 SEM iterations (with the first 100 as burn-in), we can expect the results obtained with different seed values to yield results that are reasonably close.

6 Concluding remarks

In this paper, the estimation of parameters of the component lifetime distribution based on complete and censored system lifetime data have been studied. Different ways have been suggested for implementing the SEM algorithm and their performances have been evaluated by Monte Carlo simulations. It has been shown that the SEM algorithm based on ordered system signature approximates the directly numerical MLE very closely.

Since the use of the ordered system signature does not add complexity to the computation and that the exact ordered system signatures are not required, the SEM algorithm with the ordered system signature is the one we would recommended to use. The SEM algorithm is easy to program and implement as long as the truncated distributions of the underlying component lifetime distributions are easy to simulate from. In the case when several statistical models are under consideration for model fitting, instead of going through all the cumbersome derivations of the MLEs and writing computer programs to obtain the MLEs for every single distribution under consideration, the SEM algorithm provides an attractive alternative since we only need the algorithms to generate from the component lifetime distribution g_{RX} in Eq. (7) and g_{LX} in Eq. (8).

SEM algorithm might not be as efficient computationally as the MLEs obtained by direct maximization when the censoring proportion is large. SEM-II algorithm runs slower than SEM-I algorithm as it needs to impute two layers of the missing information, both the missing system lifetime and component lifetime. Moreover, the SEM-II algorithm requires to simulate the missing system lifetime which might introduce large bias when the censoring proportion is large.

SEM algorithm based on ordered signature is a highly competitive algorithm for the analysis of system-level lifetime data. For the complete data case, the SEM-OSS algorithm results in estimates very close the direct numerical MLEs, and better than those resulting from the SEM-SS algorithm. For Type-II censored data case, SEM-I algorithm and SEM-II algorithm perform similarly in the sense that both approximate the MLEs very well. When censoring proportion is small to moderate (say, $q \leq 40\%$), SEM-II algorithm is the one to be recommended for use. When censoring proportion is large (say, $q > 40\%$), SEM-I algorithm is recommended for use.

The SEM algorithms presented in this work could be extended to Type-I censored data with some straightforward modifications. Specifically, suppose a Type-I censored life testing experiment with m systems gets terminated at a pre-fixed time τ_0 and r system failures are observed before time τ_0 while $(m - r)$ system lifetimes are censored. Here, r is a random variable. Conditional on the observed value of r , the SEM algorithm (SEM-I) presented in Sect. 3.2.1 can be used by setting $\tau = \tau_0$ and the point of truncation $t_k = \tau_0$ in Eqs. (7) and (8). Similarly, the SEM algorithm (SEM-II) presented in Sect. 3.2.2 can be used by setting $\tau = \tau_0$. Besides Type-I censoring, the SEM algorithm proposed in the preceding sections could be applied to other generalized censoring schemes such as hybrid censoring (Epstein 1954; Balakrishnan and Kundu 2013) and progressive censoring (Balakrishnan and Cramer 2014). Since the results developed in this work are for coherent and mixed systems with i.i.d. components, it will be of interest to generalize these results to reliability systems with non-i.i.d. components and/or dynamic systems. Research in this direction is currently under progress, and we hope to report these findings in a future paper.

Acknowledgments The authors thank two anonymous referees and the associate editor for their useful comments and suggestions on an earlier version of this manuscript which resulted in this improved version. H.K.T. Ng's work was supported by a grant from the Simons Foundation (#280601).

References

- Balakrishnan N, Ng HKT, Navarro J (2011) Exact nonparametric inference for component lifetime distribution based on lifetime data from systems with known signatures. *J Nonparametric Stat* 23:741–752
- Balakrishnan N, Ng HKT, Navarro J (2011) Linear inference for type-II censored lifetime data of reliability systems with known signatures. *IEEE Trans Reliab* 60:426–440
- Balakrishnan N, Cramer E (2014) *The art of progressive censoring: applications to reliability and quality*. Birkhäuser, Boston
- Balakrishnan N, Kundu D (2013) Hybrid censoring: models, inferential results and applications. *Comput Stat Data Anal* 57:166–209 (with discussions)
- Balakrishnan N, Volterman W (2014) On the signatures of ordered system lifetimes. *J Appl Probab* 51:82–91
- Balakrishnan N, Zhu X (2014) On the existence and uniqueness of the maximum likelihood estimates of the parameters of Birnbaum–Saunders distribution based on type-I, type-II and hybrid censored samples. *Statistics* 48:1013–1032
- Barlow RE, Proschan F (1965) *Mathematical theory of reliability*. Wiley, New York
- Bhattacharya D, Samaniego FJ (2010) Estimating component characteristics from system failure-time data. *Nav Res Logist* 57:380–389
- Boland PJ, Samaniego FJ (2004) The signature of a coherent system and its applications in reliability. In: Soyer R, Mazzuchi T, Singpurwalla N (eds) *Mathematical reliability: an expository perspective*. Kluwer, Boston, pp 1–29
- Celeux G, Chauveau D, Diebolt J (1996) Stochastic versions of the EM algorithm: an experimental study in the mixture case. *J Stat Comput Simul* 55:287–314
- Celeux G, Diebolt J (1985) The SEM algorithm: a probabilistic teacher algorithm derived from the EM algorithm for the mixture problem. *Comput Stat Q* 2:73–82
- Celeux G, Diebolt J (1992) A stochastic approximation type EM algorithm for the mixture problem. *Stochastics* 41:119–134
- Chahkandi M, Ahmadi J, Baratpour S (2014) Non-parametric prediction intervals for the lifetime of coherent systems. *Stat Pap* 55:1019–1034
- Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the EM algorithm. *J R Stat Soc Ser B* 39:1–38
- Diebolt J, Celeux G (1993) Asymptotic properties of a stochastic EM algorithm for estimating mixing proportions. *Stoch Models* 9:599–613
- Diebolt J, Ip E (1996) *Markov Chain Monte Carlo in practice*. Springer, New York
- Epstein B (1954) Truncated life tests in the exponential case. *Ann Math Stat* 25:555–564
- Eryılmaz S (2010) Conditional lifetimes of consecutive k -out-of- n systems. *IEEE Trans Reliab* 59:178–182
- Eryılmaz S, Koutras MV, Triantafyllou IS (2010) Signature based analysis of m -consecutive- k -out-of- n : F systems with exchangeable components. *Nav Res Logist* 58:344–354
- Frenkel I, Khvatskin L (2006) Cost-effective maintenance with preventive: replacement of oldest component. *Maint Reliab* 2:37–39
- Ip E (1994) *A stochastic EM estimator in the presence of missing data: theory and applications*. PhD Thesis, Department of Statistics, Stanford University
- Kocher S, Mukerjee H, Samaniego FJ (1999) The signature of a coherent system and its application to comparisons among systems. *Nav Res Logist* 46:507–523
- Lange K (2001) *Numerical analysis for statisticians*. Springer, New York
- McLachlan G, Krishnan T (2008) *The EM algorithm and extensions*, 2nd edn. Wiley, Hoboken
- Meeker WQ, Escobar LA (1998) *Statistical methods for reliability data*. Wiley, New York
- Navarro J, Ruiz JM, Sandoval CJ (2007) Properties of coherent systems with dependent components. *Commun Statist Theory Methods* 36:175–191
- Nelson W (2005) *Applied life data analysis*. Wiley, Hoboken
- Ng HKT, Kundu D, Balakrishnan N (2003) Modified moment estimation for the two-parameter Birnbaum–Saunders distribution. *Comput Stat Data Anal* 43:283–298

- Ng HKT, Navarro J, Balakrishnan N (2012) Parametric inference from system lifetime data under a proportional hazard rate model. *Metrika* 75:367–388
- Nielsen SF (2000) The stochastic EM algorithm: estimation and asymptotic results. *Bernoulli* 6:457–489
- R Core Team (2014) R: a language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. ISBN: 3-900051-07-0
- Redner RA, Walker HF (1984) Mixture densities, maximum likelihood and the EM algorithm. *SIAM Rev* 26:195–239
- Samaniego FJ (1985) On closure of the IFR class under formation of coherent systems. *IEEE Trans Reliab* 34:69–72
- Samaniego FJ (2007) System signatures and their applications in engineering reliability. Springer, New York
- Wu CJ (1983) On the convergence properties of the EM algorithm. *Ann Stat* 11:95–103
- Ye ZS, Ng HKT (2014) On analysis of incomplete field failure data. *Ann Appl Stat* 8:1713–1727
- Zhang J, Ng HKT, Balakrishnan N (2015) Statistical inference of component lifetimes with location-scale distributions from censored system failure data with known signature. *IEEE Trans Reliab* (to appear)