

# PARTIAL MONITORING OF MULTISTATE SYSTEMS

Kristina Skutlaberg, Arne Bang Huseby, Bent Natvig  
University of Oslo, Norway

## Abstract

For large multicomponent systems it is typically too costly to monitor the entire system constantly. In the present paper we consider a case where a component is unobserved in a time interval  $[0, T]$ . Here  $T$  is a stochastic variable with a distribution which depends on the structure of the system and the lifetime distribution of the other components. Thus, different systems will result in different distributions of  $T$ , the main focus of the paper is on how the unobserved period of time affects what we learn about the unobserved component during this period. We analyse this by considering three different cases. In the first case we consider both  $T$  as well as the state of the unobserved component at time  $T$  as given. In the second case we allow the state of the unobserved component at time  $T$  to be stochastic, while in the third case both  $T$  and the state are treated as stochastic variable. In all cases we study the problem using preposterior analysis. That is, we investigate how much information we can expect to get by the end of the time interval  $[0, T]$ . The methodology is also illustrated on a more complete example.

## 1 Introduction

Multistate systems were introduced in the mid 1970's as a natural generalisation of binary systems. For an extensive introduction and key results on such systems, see [Natvig 2011]. See also [Lisnianski et. al. 2010]. As more powerful computers have become available, advanced analysis of multistate systems is now becoming possible. Such analysis includes evaluating the system reliability and availability, as well as calculating advanced multistate importance measures. See [Natvig et. al. 2009]. A key method in relation to this is *discrete event simulation*. This is described in [Huseby et. al. 2010]. Improved algorithms also allows various optimization problems related to multistate systems to be solved. See e.g., [Liu 2010] and [Huang and Wang 2016]. In [Skutlaberg and Natvig 2016] a new approach to importance measures based on decision analysis and optimization is introduced.

Continuously monitoring individual components in a complex system is typically very costly. Thus, when gathering operational data from such systems, some or all of the components will typically be monitored only at certain points of time e.g., when component degradation or errors are displayed through changes in the system state. As a result of such partial monitoring one ends up with various types of censored observations. In [Natvig and Gåsemyr 1998] a binary, monotone system whose component states are dependent through the possible occurrence of independent common shocks was considered. The system is assumed to be observed until it fails. At this instant, the set of failed components and the failure time of the system are noted. The failure times of the individual components, however, are assumed to be unknown. Using a Bayesian approach the paper shows how to handle such autopsy data. A similar situation is considered in [Gåsemyr and Natvig 2001]. This paper also discusses applications to preventive system maintenance. In [Gåsemyr and Natvig 2005] a marked point process linked to partial monitoring of some components is considered. Furthermore, the paper demonstrates how to arrive at the posterior distribution for the relevant parameter vector by using a data augmentation method. The more recent paper [Curcurù et. al. 2010] also considers predictive maintenance policies and proposes a procedure for the computation of the maintenance time that minimizes the global maintenance cost. By adopting a stochastic model for the degradation process and by hypothesizing the use of imperfect monitoring, the paper shows how to carry

out Bayesian updating of prior knowledge.

Partial monitoring of components is particularly relevant for systems where both the components as well as the system itself can be in multiple states. In the present paper we provide a new methodology for analyzing data from such systems. In particular, we show how the effects of the imperfect monitoring can be quantified by using *preposterior analysis*.

We consider a multistate system with component set  $E = \{1, \dots, n\}$ . For all  $i \in E$  we let  $S_i = \{0, 1, \dots, s_i\}$  denote the possible states of component  $i$ . The states are ordered in the natural way such that  $s_i$  is the best state while 0 denotes the state where the component is completely failed. The state of component  $i$  at time  $t$  is denoted by  $X_i(t)$ ,  $t \geq 0$ ,  $i \in E$ . At time 0 each component starts out in its best state. That is, we assume that  $X_i(0) = s_i$ ,  $i \in E$ . As time goes by the components go through their respective states until reaching state 0. At this point of time the components will be replaced or repaired back to its best state again. However, in this context we consider only the first life cycles of the components, and introduce for all  $i \in E$  and  $j \in S_i$ :

- $T_{ij}$  = The time component  $i$  spends in state  $j$  during its first life cycle,
- $U_{ij}$  = The point of time when component  $i$  enters state  $j$  in its first life cycle.

Thus,  $U_{i,s_i} = 0$ ,  $U_{i,s_i-1} = T_{i,s_i}$ ,  $U_{i,s_i-2} = T_{i,s_i} + T_{i,s_i-1}$  and so on. Moreover, for all  $i \in E$  and  $j \in S_i$  we have:

$$X_i(t) = j, \quad t \in [U_{i,j}, U_{i,j-1}).$$

For all  $i \in E$  we also let:

$$\begin{aligned} \mathbf{T}_i &= (T_{i0}, T_{i1}, \dots, T_{i,s_i}), \\ \boldsymbol{\lambda}_i &= (\lambda_{i0}, \lambda_{i1}, \dots, \lambda_{i,s_i}). \end{aligned}$$

Regarding distributions we assume that  $T_{ij} | \lambda_{ij} \sim \text{Exp}(\lambda_{ij})$  and that  $\lambda_{ij} \sim \text{Gamma}(\alpha_{ij}, \beta_{ij})$ , for all  $i \in E$  and  $j \in S_i$ . We refer to the  $\lambda_{ij}$ 's as the *transition rates* of the  $i$ th component. We also assume that the components are stochastically independent.

Throughout this paper the chosen parametrisation of a  $\text{Gamma}(\alpha, \beta)$ -distribution is such that if  $\lambda \sim \text{Gamma}(\alpha, \beta)$  a priori, then the prior density of  $\lambda$  is given by:

$$\pi(\lambda) = \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} \exp(-\beta\lambda), \quad \lambda > 0.$$

The resulting expectation and variance are thus  $E[\lambda] = \alpha/\beta$  and  $\text{Var}[\lambda] = \alpha/\beta^2$ . If  $T$  is a stochastic variable where  $T | \lambda \sim \text{Exp}(\lambda)$ , it follows from Bayes' theorem that  $\lambda | T = t \sim \text{Gamma}(\alpha + 1, \beta + t)$ , and that  $\lambda | T > t \sim \text{Gamma}(\alpha, \beta + t)$ .

We now focus on a specific component  $i \in E$ . In order to simplify the notation we define  $s$  to be the best state of this component, i.e.,  $s_i = s$ . The main idea in this article is to study a situation where this component is unmonitored in an interval  $[0, T]$ . Here  $T$  denotes a stochastic variable with a probability distribution that may depend on the structure of the system as well as on the lifetime distributions of the other components.  $T$  may e.g., be the point of time where the system reaches a state where it needs to be fully inspected, and we refer to  $T$  as the *censoring point of time*.

In Section 3 both the censoring point of time and the state of component  $i$  at this time are considered to be fixed. In Section 4 we still assume that censoring point of time is fixed, but that the state of component  $i$  is assumed to be stochastic, while in Section 5 both the censoring point of time and the state are stochastic. Finally, we apply the methodology to multistate systems in Section 6.

## 2 Preposterior analysis

Before we proceed with the main part of the paper, we review the basic ideas of preposterior analysis. For a more extensive introduction to this see [Raiffa and Schlaifer 1961], [Barlow and Wu 1981] or [Berger 1985]. We will now describe this principle.

Assume that  $\theta \in \Theta$  is an unknown parameter with prior distribution  $\pi(\theta)$ . By observing  $X \in \mathcal{X}$  our knowledge of  $\theta$  will increase. With  $\theta$  given, then  $X$  follows the distribution  $p(x|\theta)$ . From Bayes Theorem we find that the posterior distribution for  $\theta$  is

$$\pi(\theta|x) = \frac{p(x|\theta)\pi(\theta)}{p(x)},$$

where the unconditional distribution for  $X$  is given by:

$$p(x) = \int_{\Theta} p(x|\theta)\pi(\theta)d\theta.$$

A decision  $d \in D$  is made based on  $X$ . Let  $L(\theta, d)$  be the loss connected to this decision. The optimal decision is found by minimizing the expected loss with respect to  $X$  given the decision  $d$ :

$$E[L(\theta, d)|X] = \int_{\Theta} L(\theta, d)\pi(\theta|X)d\theta.$$

Let  $\delta(X)$  denote the optimal decision. We now have that

$$E[L(\theta, \delta(X))|X] \leq E[L(\theta, d)|X], \quad \forall d \in D.$$

When using preposterior analysis one wishes to estimate the expected loss *before*  $X$  is observed. This means that prior to the observation of  $X$  we try to find out how well we are able to make the decision after we observe  $X$ . We find the following:

$$E[L(\theta, \delta(X))] = \int_{\mathcal{X}} E[L(\theta, \delta(x))|x]p(x)dx$$

Assume that we are interested in estimating  $\theta$  with quadratic loss, that is  $L(\theta, d) = (\theta - d)^2$ . It is well known that the optimal estimate for  $\theta$  is  $\delta(X) = E[\theta|X]$ . See for instance [Berger 1985]. We get:

$$E[L(\theta, \delta(X))|X] = E[(\theta - E[\theta|X])^2] = \text{Var}[\theta|X].$$

The expected loss based on the observation  $X$ , but calculated *before* the actual observation, can be expressed as:

$$E[L(\theta, \delta(X))] = \int_{\mathcal{X}} E[L(\theta, \delta(x))|x]p(x)dx = \int_{\mathcal{X}} \text{Var}[\theta|x]p(x)dx = E[\text{Var}[\theta|X]]. \quad (1)$$

The expectation  $E[L(\theta, \delta(X))]$  is referred to as the *preposterior loss* based on the observation  $X$ . Note that the preposterior loss differs from what we would get if we were to estimate  $\theta$  *without* knowledge of  $X$ . If we were unable to observe  $X$ , the best estimate of  $\theta$  would be the expectation in the prior distribution. Let  $\delta_0$  denote this estimate. Thus, we have  $\delta_0 = E[\theta]$ . The expected loss for this decision is therefore:

$$E[L(\theta, \delta_0)] = E[(\theta - E[\theta])^2] = \text{Var}[\theta],$$

which is the prior variance. This can be expressed as:

$$\text{Var}[\theta] = E[\text{Var}[\theta|X]] + \text{Var}[E[\theta|X]].$$

Since  $\text{Var}[E[\theta|X]] \geq 0$ , it follows that:

$$E[L(\theta, \delta_0)] \geq E[L(\theta, \delta(X))].$$

Furthermore, the value of the observation  $X$  is the reduction in the expected loss as a consequence of the fact that we can observe  $X$ :

$$E[L(\theta, \delta_0)] - E[L(\theta, \delta(X))] = \text{Var}[E[\theta|X]].$$

We will use this expression as a quantification of the value of a future observation.

### 3 Fixed censoring point of time and state

In this section we compute the posterior distributions for the transition rates of the unmonitored component  $i$  given the censoring point of time  $T = t$  as well as the state  $X_i(t)$ .

Since component  $i$  is a multistate component, the posterior distributions for the transition rates will depend on the state of this component at the censoring point of time. Due to the somewhat complex censoring patterns, it is not always possible to calculate these posterior distributions analytically. When analytical calculations are not possible, we will use rejection sampling to estimate the posterior distributions for the transition rates. It is also possible to use data augmentation and Gibbs-sampling, as described in [Gåsemyr and Natvig 2005].

Given that  $T = t$  and  $X_i(t) = j$ , the rejection sampling method can be described as follows:

**Algorithm 3.1** Repeat the following steps until a sufficient amount of samples (not counting the rejected ones) has been generated:

- STEP 1. Generate  $\lambda_i$  from the given priors.
- STEP 2. Generate  $T_i$  from the conditional distributions given  $\lambda_i$ .
- STEP 3. Calculate the resulting value of  $X_i(t)$ .
- STEP 4. If  $X_i(t) = j$ , we keep  $\lambda_i$ . Otherwise the vector is rejected.

The calculation of  $X_i(t)$  in Step 3 is done as follows: If  $T_{is} > t$ , no transitions have occurred in  $[0, t]$ . Thus, in this case  $X_i(t) = s$ . If  $T_{is} \leq t$  and  $T_{is} + T_{i,s-1} > t$ , a transition from state  $s$  to state  $s - 1$  has occurred in  $[0, t]$ . Thus, in this case  $X_i(t) = s - 1$ . If  $T_{is} + T_{i,s-1} \leq t$  and  $T_{is} + T_{i,s-1} + T_{i,s-2} > t$ , a transition from state  $s$  to state  $s - 1$  and a transition from state  $s - 1$  to state  $s - 2$  have occurred, implying that  $X_i(t) = s - 2$ . By proceeding in this fashion the state  $X_i(t)$  is eventually identified.

If  $X_i(t) = s$ , this implies that  $T_{is} > t$ . This means that the state of component  $i$  has been the same throughout the interval  $[0, t]$ . This implies that we have no information beyond the prior belief concerning  $\lambda_{i1}, \dots, \lambda_{i,s-1}$ . Moreover, in this case the posterior distribution of  $\lambda_{is}$  can be calculated analytically, and we get that  $\lambda_{is} | T_{is} > t \sim \text{Gamma}(\alpha_{is}, \beta_{is} + t)$ . For all other transition rates the posterior distribution equals the prior distribution.

If  $X_i(t) = s - 1$ , then  $T_{is} \leq t$  while  $T_{is} + T_{i,s-1} > t$ . In this case, we use rejection sampling to estimate the posterior distributions for both  $\lambda_{is}$  and  $\lambda_{i,s-1}$ . For the other failure rates, the posterior distribution equals the prior distribution.

If  $X_i(t) = s - 2$ , then  $T_{is} + T_{i,s-1} \leq t$ , while  $T_{is} + T_{i,s-1} + T_{i,s-2} > t$ . Estimation of the posterior distribution for  $\lambda_{is}$ ,  $\lambda_{i,s-1}$  and  $\lambda_{i,s-2}$  are done using rejection sampling. The posterior distribution equals the prior distribution for  $\lambda_{i,s-3}, \dots, \lambda_{i1}$ .

This scheme continues until  $X_i(t) = 0$  where  $T_{is} + \dots + T_{i1} \leq t$ . In this case, rejection sampling is used to estimate the posterior distributions for the failure rates  $\lambda_{is}, \dots, \lambda_{i1}$ .

In order to illustrate this we consider a specific example where  $S_i = \{0, 1, 2, 3\}$ . In this example we consider three different censoring points:  $t = 5$ ,  $t = 10$  and  $t = 15$ . Since we do not include repairs in the model, we focus on the transition rates  $\lambda_{i1}$ ,  $\lambda_{i2}$ , and  $\lambda_{i3}$ . We assume that the parameters in the prior distributions for these transition rates are  $\alpha_{ij} = 1.0$  and that  $\beta_{ij} = 5.0$  for  $j = 1, 2, 3$ . Thus, we have the following prior expectations and variances:

$$\begin{aligned} E[\lambda_{ij}] &= \frac{\alpha_{ij}}{\beta_{ij}} = \frac{1}{5} = 0.2, \quad j = 1, 2, 3. \\ \text{Var}[\lambda_{ij}] &= \frac{\alpha_{ij}}{\beta_{ij}^2} = \frac{1}{5^2} = 0.04, \quad j = 1, 2, 3. \end{aligned}$$

In order to estimate the posterior distributions Algorithm 3.1 was used. After all rejected samples were eliminated, we ended up with 500000 samples.

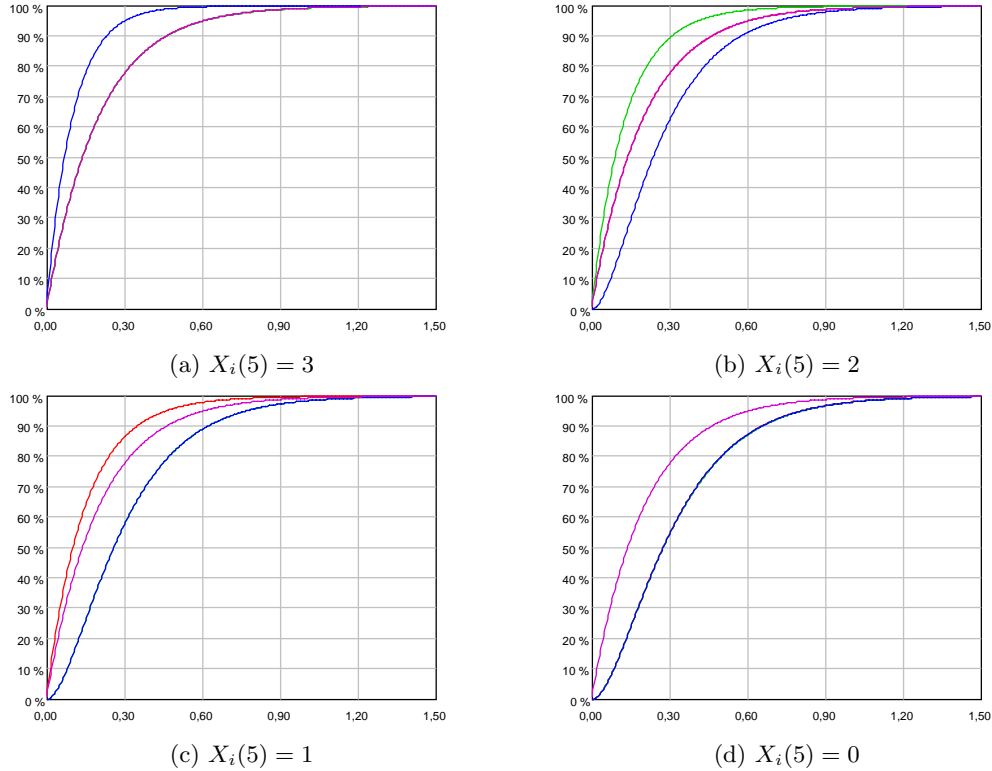


Figure 1: Posterior distributions for  $t = 5$  for  $\lambda_{ij}$ ,  $j = 1, 2, 3$  (red, green and blue curves respectively) along with the common prior distribution,  $\text{Gamma}(1, 5)$  (violet curve).

In Figure 1, Figure 2 and Figure 3 we have shown the estimated posterior cumulative distributions for  $\lambda_{ij}$ ,  $j = 1, 2, 3$  (red, green and blue curves respectively) given that  $X_i(t) = 0, 1, 2, 3$  along with the common prior distribution,  $\text{Gamma}(1, 5)$  (violet curve).

When  $X_i(t) = 0$ , all the posterior distributions are identical. Thus, the red, green and blue curves are displayed on top of each other. We observe that in this case the posterior distributions are located to the right of the prior distribution (violet curve). This indicates that all the transition rates are higher than the prior distribution implies since in this case all possible transitions have occurred in  $[0, t]$ . This effect is most visible when  $t$  is small since in this case all transitions occur within a short interval. When  $t$  is larger, the effect is weaker since in this case all transitions occur within a longer interval.

When  $X_i(t) = 1$ , the posterior distributions of  $\lambda_{i2}$  (green curve) and  $\lambda_{i3}$  (blue curve) are identical and located to the right of the prior distribution (violet curve). This indicates that the transition rates  $\lambda_{i2}$  and  $\lambda_{i3}$  are higher than the prior distribution implies. This reflects that in this case both transitions from state 3 to state 2 and from state 2 to state 1 have occurred in  $[0, t]$ . Again this effect is most visible when  $t = 5$ , and less dramatic as  $t$  increases. The posterior distribution of  $\lambda_{i1}$  (red curve), on the other hand, is located to the left of the prior distribution. The reason for this is that when  $X_i(t) = 1$ , no transition from state 1 to state 0 has occurred in  $[0, t]$ . This indicates that the transition rate  $\lambda_{i1}$  is lower than implied by the prior distribution. This effect is stronger when  $t$  is large than when  $t$  is small.

When  $X_i(t) = 2$ , the posterior distribution of  $\lambda_{i3}$  (blue curve) is located to the right of the prior distribution (violet curve). This indicates that  $\lambda_{i3}$  is higher than implied by the prior. This reflects that in this case a transition from state 3 to state 2 has occurred in  $[0, t]$ . This effect is stronger when  $t$  is small, where the transition from state 3 to state 2 must occur early, and weaker when  $t$  is large, where the transitions from state 3 to state 2 may occur later. The posterior distribution of  $\lambda_{i2}$  (green curve) is located to the left of the prior distribution (violet curve). This indicates that

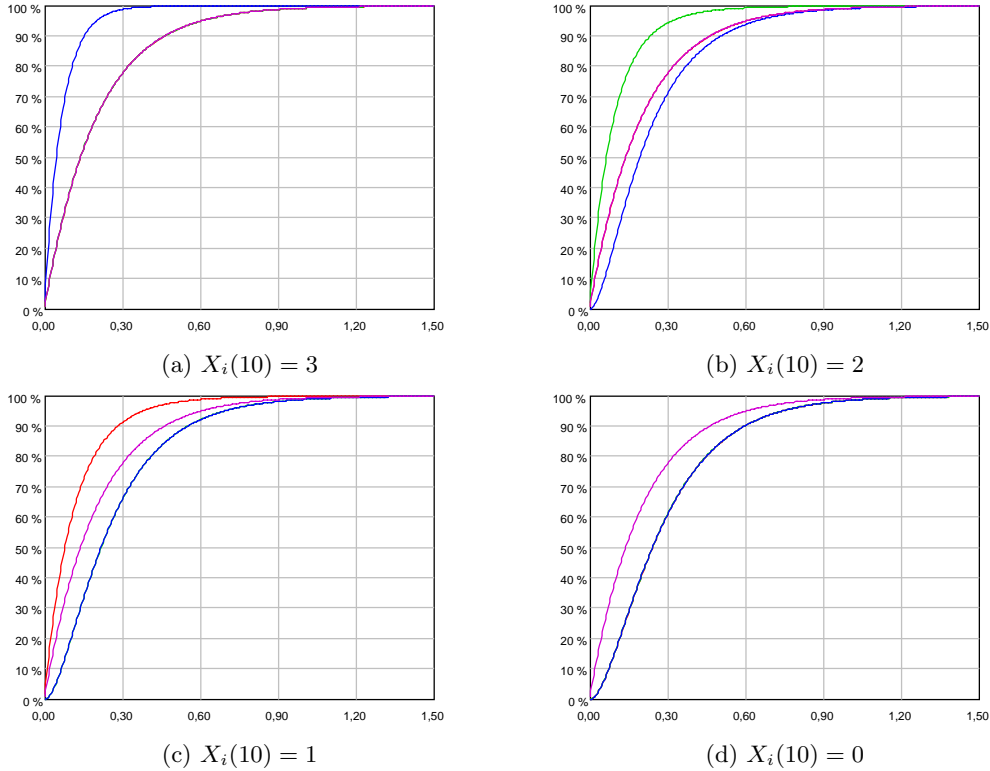


Figure 2: Posterior distributions for  $t = 10$  for  $\lambda_{ij}$ ,  $j = 1, 2, 3$  (red, green and blue curves respectively) along with the common prior distribution,  $\text{Gamma}(1, 5)$  (violet curve).

$\lambda_{i2}$  is lower than implied by the prior. This reflects that in this case a transition from state 2 to state 1 has not occurred in  $[0, t]$ . This effect is stronger when  $t$  is large than when  $t$  is small which again is as expected. Finally, since no transition from state 1 to state 0 can occur in this case, the posterior distribution of  $\lambda_{i1}$  (red curve) is identical to the prior distribution (violet curve).

When  $X_i(t) = 3$ , the posterior distribution of  $\lambda_{i3}$  (blue curve) is located to the left of the prior distribution (violet curve). Again this indicates that  $\lambda_{i3}$  is lower than implied by the prior since in this case no transitions from state 3 to state 2 occurred in  $[0, t]$ . As one can expect, this effect is stronger when  $t$  is large than when  $t$  is small. Since no transitions from state 2 to state 1 or from state 1 to state 0 can occur in this case, the posterior distributions of  $\lambda_{i2}$  (green curve) and  $\lambda_{i1}$  (red curve) are identical to the prior distribution (violet curve).

As a step towards the preposterior analysis, we have estimated conditional variances:

$$\text{Var}[\lambda_{ij}|X_i(t) = x], \quad x = 0, 1, 2, 3, \quad j = 1, 2, 3.$$

The results are listed in Table 17, Table 19, and Table 21 in Appendix A. As a comparison we have also estimated the corresponding conditional variances given that component  $i$  is monitored continuously throughout the interval  $[0, t]$ :

$$\text{Var}[\lambda_{ij}|X(u) : \text{for all } u \in [0, t], X(t) = x], \quad x = 0, 1, 2, 3, \quad j = 1, 2, 3.$$

The results are listed in Table 18, Table 20, and Table 22 in Appendix A.

## 4 Fixed censoring point of time, stochastic state

In this section we proceed with the preposterior analysis. In this section, we will assume, as we did in Section 3, that  $t$  is fixed. However, now the state of component  $i$  at time  $t$ ,  $X_i(t)$ , is considered

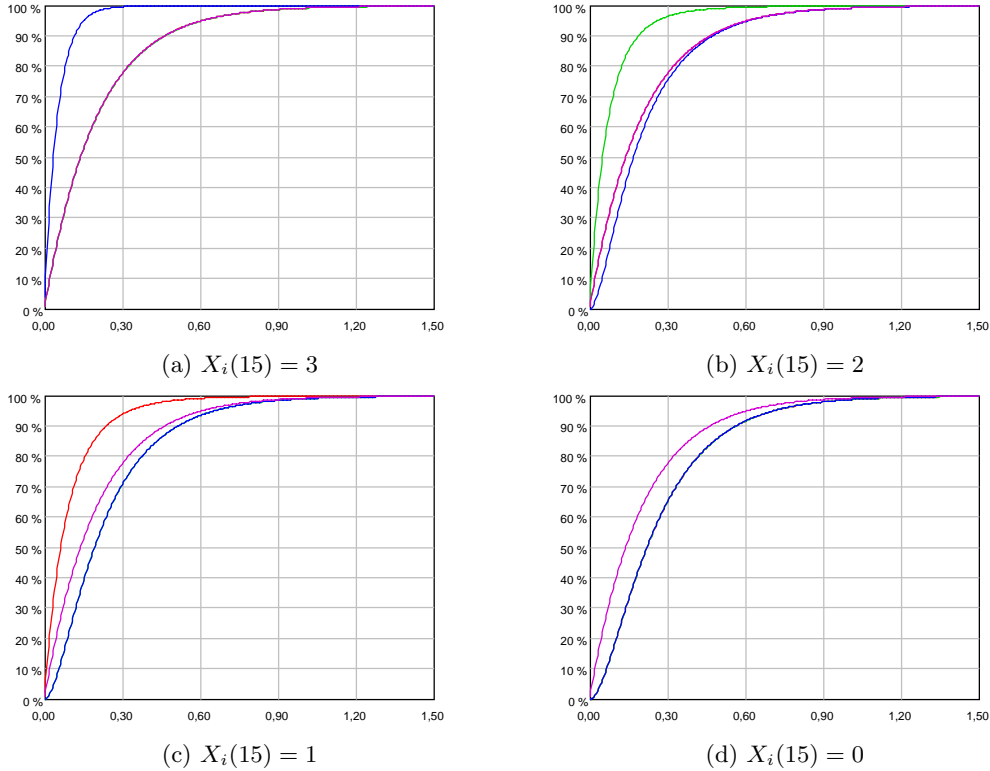


Figure 3: Posterior distributions for  $t = 15$  for  $\lambda_{ij}$ ,  $j = 1, 2, 3$  (red, green and blue curves respectively) along with the common prior distribution,  $\text{Gamma}(1, 5)$  (violet curve).

to be stochastic. The goal is to describe the value of observing the state  $X_i(t)$ . It is also of interest to compare this to the value of observing component  $i$  in the entire interval  $[0, t]$ .

As described previously, we consider the case where the failure rates  $\lambda_{i1}, \dots, \lambda_{is}$  are estimated with quadratic loss. This implies that the resulting preposterior losses are given by:

$$E[\text{Var}[\lambda_{ij}|X_i(t)]], \quad j = 1, \dots, s.$$

Since there are only a finite number of possible states for component  $i$ , the preposterior losses can easily be calculated as:

$$E[\text{Var}[\lambda_{ij}|X_i(t)]] = \sum_{x=0}^s \text{Var}[\lambda_{ij}|X_i(t) = x] \Pr(X_i(t) = x), \quad j = 1, \dots, s. \quad (2)$$

In (2) the posterior variances  $\text{Var}[\lambda_{ij}|X_i(t) = x]$ ,  $x = 0, 1, \dots, s$  are estimated using the posterior distributions we found using Algorithm 3.1. In fact it is easy to include this estimation as a part of the rejection sampling. All the posterior variances needed for these calculations are listed in Appendix A.

The only remaining problem is to estimate the probability distribution of  $X_i(t)$ . This can be done using the first three steps of Algorithm 3.1. However, instead of rejecting samples, we use *all* samples, and thus obtain an estimate the distribution of  $X_i(t)$  by calculating the fractions of these samples where  $X_i(t) = x$ ,  $x = 0, 1, 2, 3$ .

In order to illustrate this we consider the example introduced in Section 3, and calculate the estimates for the distribution of  $X_i(t)$  for  $t = 5, 10, 15$ . The results are shown in Table 1. We observe that for  $t = 5$  the probability distribution is skewed towards the higher values. For such a short time interval there is a high probability that the number of transitions is either 0 or 1, i.e., that

$X_i(t)$  is either 3 or 2. On the other hand when  $t = 15$ , the probability distribution of  $X_i(t)$  is close to uniform. This reflects that when the time interval  $[0, t]$  is longer, there is a higher probability that more transitions have occurred. By combining the probability distributions in Table 1 with the

	$x = 0$	$x = 1$	$x = 2$	$x = 3$
$P(X_i(5) = x)$	0.0490	0.1302	0.3207	0.5000
$P(X_i(10) = x)$	0.1632	0.1996	0.3038	0.3333
$P(X_i(15) = x)$	0.2769	0.2120	0.2612	0.2498

Table 1: Estimated probability distribution of  $X_i(t)$  for  $t = 5, 10, 15$ .

corresponding variance estimates in the Appendix, using (2), we arrive at the estimated preposterior losses. The results are given in Table 2.

	$\lambda_{i1}$	$\lambda_{i2}$	$\lambda_{i3}$
$E[\text{Var}[\lambda_{ij} X_i(5)]]$	0.0387	0.0361	0.0300
$E[\text{Var}[\lambda_{ij} X_i(10)]]$	0.0371	0.0344	0.0309
$E[\text{Var}[\lambda_{ij} X_i(15)]]$	0.0362	0.0342	0.0320

Table 2: Estimated preposterior losses when component  $i$  is *not* monitored in  $[0, t]$ ,  $t = 5, 10, 15$ .

As a comparison we have also estimated the preposterior losses given that component  $i$  is monitored continuously throughout the interval  $[0, t]$ . In this case the exact points of times when the state of component  $i$  changes would be known, and thus, the resulting posterior variances can be calculated analytically using the following formulas:

CASE 0.  $X_i(t) = 0$ .

$$\begin{aligned}\text{Var}[\lambda_{i1}|X_i(u) : u \in [0, t]] &= \frac{\alpha_1 + 1}{(\beta_1 + T_{i1})^2}, \\ \text{Var}[\lambda_{i2}|X_i(u) : u \in [0, t]] &= \frac{\alpha_2 + 1}{(\beta_2 + T_{i2})^2}, \\ \text{Var}[\lambda_{i3}|X_i(u) : u \in [0, t]] &= \frac{\alpha_3 + 1}{(\beta_3 + T_{i3})^2},\end{aligned}\tag{3}$$

CASE 1.  $X_i(t) = 1$ .

$$\begin{aligned}\text{Var}[\lambda_{i1}|X_i(u) : u \in [0, t]] &= \frac{\alpha_1}{(\beta_1 + t - T_{i2} - T_{i3})^2}, \\ \text{Var}[\lambda_{i2}|X_i(u) : u \in [0, t]] &= \frac{\alpha_2 + 1}{(\beta_2 + T_{i2})^2}, \\ \text{Var}[\lambda_{i3}|X_i(u) : u \in [0, t]] &= \frac{\alpha_3 + 1}{(\beta_3 + T_{i3})^2},\end{aligned}\tag{4}$$

CASE 2.  $X_i(t) = 2$ .

$$\begin{aligned}\text{Var}[\lambda_{i1}|X_i(u) : u \in [0, t]] &= \frac{\alpha_1}{\beta_1^2}, \\ \text{Var}[\lambda_{i2}|X_i(u) : u \in [0, t]] &= \frac{\alpha_2}{(\beta_2 + t - T_{i3})^2}, \\ \text{Var}[\lambda_{i3}|X_i(u) : u \in [0, t]] &= \frac{\alpha_3 + 1}{(\beta_3 + T_{i3})^2},\end{aligned}\tag{5}$$



CASE 3.  $X_i(t) = 3$ .

$$\begin{aligned}\text{Var}[\lambda_{i1}|X_i(u) : u \in [0, t]] &= \frac{\alpha_1}{\beta_1^2}, \\ \text{Var}[\lambda_{i2}|X_i(u) : u \in [0, t]] &= \frac{\alpha_2}{\beta_2^2}, \\ \text{Var}[\lambda_{i3}|X_i(u) : u \in [0, t]] &= \frac{\alpha_3}{(\beta_3 + t)^2},\end{aligned}\tag{6}$$

The relevant formula, i.e., either (3), (4), (5) or (6) chosen according to the value of  $X_i(t)$ , is applied in each of the simulations, and the average values of the resulting variances are calculated. The resulting preposterior losses are obtained by combining the average variances with the estimated probability distribution for  $X_i(t)$ . The results for  $t = 5, 10, 15$  are listed in Table 3, while in Table 4 we have listed the differences in preposterior losses.

	$\lambda_{i1}$	$\lambda_{i2}$	$\lambda_{i3}$
$E[\text{Var}[\lambda_{ij} X_i(u) : u \in [0, 5]]]$	0.0386	0.0353	0.0283
$E[\text{Var}[\lambda_{ij} X_i(u) : u \in [0, 10]]]$	0.0362	0.0323	0.0271
$E[\text{Var}[\lambda_{ij} X_i(u) : u \in [0, 15]]]$	0.0343	0.0307	0.0269

Table 3: Estimated preposterior losses when component  $i$  is monitored in  $[0, t]$ ,  $t = 5, 10, 15$ .

	$\lambda_{i1}$	$\lambda_{i2}$	$\lambda_{i3}$
Loss difference, $t = 5$	0.0001	0.0008	0.0017
Loss difference, $t = 10$	0.0009	0.0021	0.0038
Loss difference, $t = 15$	0.0019	0.0035	0.0051

Table 4: Differences between estimated preposterior losses when component  $i$  is monitored and when the component is *not* monitored.

As one could expect, the preposterior losses when component  $i$  is monitored in  $[0, t]$  is somewhat smaller than the corresponding preposterior losses when component  $i$  is *not* monitored in  $[0, t]$ . The difference is greatest when  $t$  is large since in this case more information is lost by not monitoring component  $i$ . We also note that the difference is a bit greater for  $\lambda_{i3}$  than for the other parameters. In fact for  $\lambda_{i3}$  the preposterior loss when component  $i$  is not monitored increases as  $t$  increases. Since component  $i$  starts out in state 3, we get the most precise information about the transition rate  $\lambda_{i3}$  in the beginning. As time goes by, however, the exact point of time when the transition from state 3 to state 2 occurs, becomes more and more uncertain. As a result the preposterior loss increases. On the other hand if component  $i$  is monitored continuously, the preposterior losses are always decreasing by increasing  $t$  since no information is lost.

For all the transition rates the prior loss, i.e., the prior variance is  $\alpha_{ij}/\beta_{ij}^2 = 0.04$ . Hence, the value of observing  $X_i(t)$ , i.e., the difference between the prior loss and the preposterior loss can easily be calculated. The results are shown in Table 5.

	$\lambda_{i1}$	$\lambda_{i2}$	$\lambda_{i3}$
Value of observing $X_i(5)$	0.0013	0.0039	0.0100
Value of observing $X_i(10)$	0.0029	0.0056	0.0091
Value of observing $X_i(15)$	0.0038	0.0058	0.0080

Table 5: The value of observing  $X_i(t)$  expressed as the difference between the prior loss and the preposterior loss.

We observe that the value of observing  $X_i(t)$  increases as  $t$  increases for  $\lambda_{i1}$  and  $\lambda_{i2}$ , and decreases for  $\lambda_{i3}$ . These effects are explained above.

## 5 Stochastic censoring point of time and state

While the rejection algorithm ensures that  $\text{Var}[\lambda_{ij}|X_i(t) = x]$  is estimated very precisely for each value of  $x$ , the algorithm throws away huge amount of samples where the calculated value of  $x$  does not match the desired one. Since we want to estimate  $\text{Var}[\lambda_{ij}|X_i(t) = x]$  for all values of  $x$ , all these rejected samples could in fact be used as well. However, if we do this, we lose control over how many samples we get for each value of  $x$ . In particular for values of  $x$  where  $P(X_i(t) = x)$  is small we get few samples, while for values of  $x$  where  $P(X_i(t) = x)$  is large we get many samples. Still, when calculating the preposterior loss  $E[\text{Var}[\lambda_{ij}|X_i(t)]]$  using (2), it is the terms where  $P(X_i(t) = x)$  is large where the precision is most important.

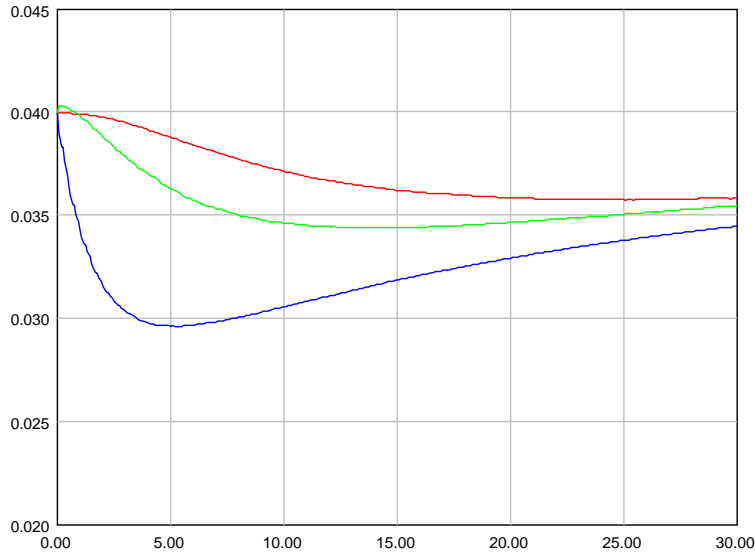


Figure 4: Preposterior losses for  $\lambda_{i1}$  (red curve),  $\lambda_{i2}$  (green curve) and  $\lambda_{i3}$  (blue curve) as a function of the censoring point of time  $t$  when component  $i$  is *not* monitored in  $[0, t]$ . 500000 simulations are used in order to estimate each of the points on the curves.

This issue becomes more relevant when censoring point of time,  $t$ , varies over a longer interval, or when this quantity is stochastic. In particular, when  $T$  is small, the distribution of  $X_i(t)$  becomes strongly concentrated on the higher states, while the probability of getting  $X_i(t) = 0$  is approximately zero. In such cases the rejection algorithm takes very long time to converge, as a huge amount of samples are rejected. Thus, we instead use the following algorithm:

**Algorithm 5.1** Repeat the following steps until a sufficient amount of samples has been generated:

- STEP 1. Generate  $\lambda_i$  from the given priors.
- STEP 2. Generate  $T_i$  from the conditional distributions given  $\lambda_i$ .
- STEP 3. Calculate the resulting value of  $X_i(t)$ .
- STEP 4. Add  $\lambda_i$  to the sampled vectors for this particular value of  $X_i(t)$ , and keep track of the number of times the different states  $X_i(t)$  occur.

Note that by counting the number of times the different states  $X_i(t)$  occur in Step 4 of Algorithm 5.1, allows the probability distribution of  $X_i(t)$  to be estimated.

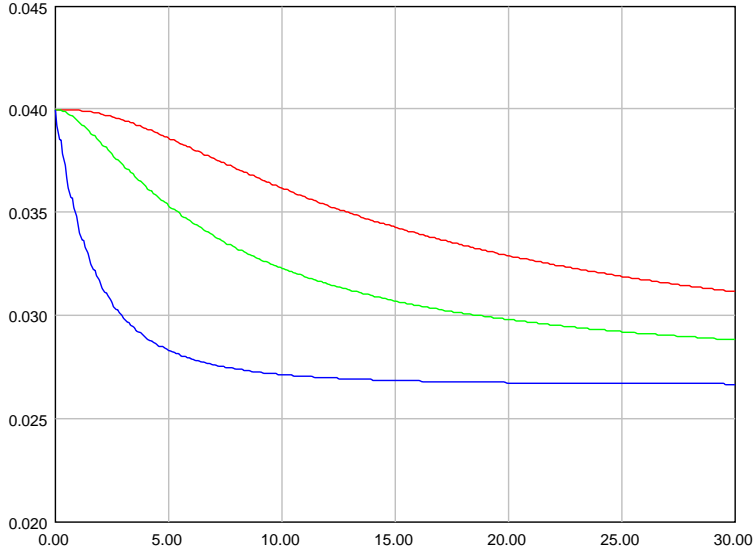


Figure 5: Preposterior losses for  $\lambda_{i1}$  (red curve),  $\lambda_{i2}$  (green curve) and  $\lambda_{i3}$  (blue curve) as a function of the censoring point of time  $t$  when component  $i$  is monitored in  $[0, t]$ . 500000 simulations are used in order to estimate each of the points on the curves.

Extensive tests have shown that Algorithm 5.1, where all samples are used, actually works very well with respect to the precision of the preposterior loss. In order to illustrate this we have run this algorithm where we have varied  $t$  in the range  $[0, 30]$ . The results for the unmonitored case are shown in Figure 4, while the corresponding results for the monitored case is shown in Figure 5. While the curves for the monitored case are decreasing all the way, the curves for the unmonitored case are decreasing in the beginning and then increasing again as  $t$  increases.

In order to take a closer look at the asymptotic behaviour of the preposterior loss curves, we start by considering the probability distribution of  $X_i(t)$ . The event that  $X_i(t) = 0$  is equivalent with the event that component  $i$  has gone through all states from 3 down to 0 at time  $t$ , i.e., that  $T_{i3} + T_{i2} + T_{i1} < t$ . Since the sum  $T_{i3} + T_{i2} + T_{i1}$  is finite with probability one, we have

$$\lim_{t \rightarrow \infty} P(X_i(t) = 0) = \lim_{t \rightarrow \infty} P(T_{i3} + T_{i2} + T_{i1} < t) = 1.$$

Hence, by (2) we get that:

$$\lim_{t \rightarrow \infty} E[\text{Var}[\lambda_{ij}|X_i(t)]] = \lim_{t \rightarrow \infty} \text{Var}[\lambda_{ij}|X_i(t) = 0], \quad j = 1, 2, 3.$$

We consider first the unmonitored case, and let  $A \subseteq [0, \infty)$  be an arbitrary measurable set. We then have:

$$\begin{aligned} P(\lambda_{ij} \in A | X_i(t) = 0) &= \frac{P(\lambda_{ij} \in A \cap X_i(t) = 0)}{P(X_i(t) = 0)} \\ &= \frac{P(\lambda_{ij} \in A) \cdot P(X_i(t) = 0 | \lambda_{ij} \in A)}{P(X_i(t) = 0)} \end{aligned}$$

Thus, since obviously  $\lim_{t \rightarrow \infty} P(X_i(t) = 0) = \lim_{t \rightarrow \infty} P(X_i(t) = 0 | \lambda_{ij} \in A) = 1$ , it follows that:

$$\lim_{t \rightarrow \infty} P(\lambda_{ij} \in A | X_i(t) = 0) = P(\lambda_{ij} \in A).$$

Thus, the conditional distribution of  $\lambda_{ij}$  given  $X_i(t) = 0$  converges to the prior distribution of  $\lambda_{ij}$ ,

$j = 1, 2, 3$  as  $t$  goes to infinity. Hence, we get:

$$\lim_{t \rightarrow \infty} E[\text{Var}[\lambda_{ij}|X_i(t)]] = \text{Var}[\lambda_{ij}] = \frac{\alpha_{ij}}{\beta_{ij}^2} = 0.04, \quad j = 1, 2, 3.$$

That is, the preposterior loss converges to the prior loss as  $t$  goes to infinity. This explains why the curves in Figure 4 increases for larger values of  $t$ .

We then consider the monitored case. In this case the event  $X_i(t) = 0$  implies that the exact values of  $T_{i3}$ ,  $T_{i2}$  and  $T_{i1}$  are known at time  $t$ . The posterior distributions of  $\lambda_{i3}$ ,  $\lambda_{i2}$  and  $\lambda_{i1}$  can then easily be calculated using Bayes' theorem, and we get that:

$$\lambda_{ij}|T_{ij} \sim \text{Gamma}(\alpha_{ij} + 1, \beta_{ij} + T_{ij}), \quad j = 1, 2, 3.$$

Hence, as in (3), we have:

$$\text{Var}[\lambda_{ij}|X_i(t) = 0] = \text{Var}[\lambda_{ij}|T_{ij}] = \frac{\alpha_{ij} + 1}{(\beta_{ij} + T_{ij})^2}, \quad j = 1, 2, 3.$$

The preposterior loss is then given by:

$$E[\text{Var}[\lambda_{ij}|T_{ij}]] = E\left[\frac{\alpha_{ij} + 1}{(\beta_{ij} + T_{ij})^2}\right] = \frac{\alpha_{ij}(\alpha_{ij} + 1)}{\beta_{ij}^2(\alpha_{ij} + 2)} = \frac{2}{75} = 0.0267, \quad j = 1, 2, 3.$$

In Figure 5 we see that the curve for the preposterior loss of  $\lambda_{i3}$  appears to converge rapidly to the value 0.0267 as  $t$  increases. For  $\lambda_{i2}$  and  $\lambda_{i1}$  the convergence is slower, but as  $t$  increases these curves will converge towards the same value as well.

By using a similar sampling method, estimates of the preposterior losses for the case where  $T$  is stochastic can be obtained. We now make the important assumption that the censoring point of time,  $T$ , is independent of all the variables affecting the state of component  $i$ , i.e.,  $\lambda_i$  and  $T_i$ . It then follows that the preposterior losses can be calculated using the following extended version of (2):

$$E[\text{Var}[\lambda_{ij}|X_i(T)]] = \int_0^\infty \left[ \sum_{x=0}^s \text{Var}[\lambda_{ij}|X_i(t) = x] \Pr(X_i(t) = x) \right] f_T(t) dt, \quad j = 1, \dots, s, \quad (7)$$

where  $f_T$  denotes the probability density of  $T$ . In order to calculate this integral analytically we would need to estimate both the posterior variances as well as the probability distribution of  $X_i(t)$  for all  $t \geq 0$ . However, this integral can easily be estimated using Monte Carlo simulation. In order to do this the simulations must be run within a double loop. In the outer loop  $T$  is sampled from its distribution. This loop is run 10000 times. In the inner loop the modified sampling procedure with no rejections is run in order to estimate the preposterior losses for each given value  $t$  of  $T$  as well as the probability distribution of  $X_i(t)$ . This inner loop is also run 10000 times. A formal description of this algorithm is as follows:

**Algorithm 5.2** Repeat the following steps until a sufficient amount of samples has been generated:

STEP 1. Generate  $T$  from the given distribution.

For each value value of  $T$  repeat the following steps until a sufficient amount of samples has been generated:

STEP 2. Generate  $\lambda_i$  from the given priors.

STEP 3. Generate  $T_i$  from the conditional distributions given  $\lambda_i$ .

STEP 4. Calculate the resulting value of  $X_i(T)$ .

STEP 5. Add  $\lambda_i$  to the sampled vectors for this particular value of  $X_i(T)$ , and keep track of the number of times the different states  $X_i(T)$  occur.

	E[T]	Var[T]
Gamma(4, 0.4)	10	25
Gamma(8, 0.4)	20	50
Gamma(12, 0.4)	30	75

Table 6: Distributions for  $T$ .

As before counting the number of times the different states  $X_i(T)$  occur, enables us to estimate the probability distribution of  $X_i(T)$   $i \in E$ .

Three different Gamma-distributions for  $T$  were considered in the simulations. The parameters as well as the corresponding expectations and variances are listed in Table 6.

Averaging the estimated preposterior losses with respect to the sampled values of  $T$  we arrived at the preposterior losses for the three different distributions for  $T$  given respectively in Table 7, Table 8 and Table 9. In these table we have listed the preposterior losses for the unmonitored case, the monitored case as well as the difference in preposterior losses.

	$\lambda_{i1}$	$\lambda_{i2}$	$\lambda_{i3}$
E[Var[ $\lambda_{ij} X_i(T)$ ]]	0.0374	0.0349	0.0310
E[Var[ $\lambda_{ij} X_i(u) : u \in [0, T]$ ]]	0.0364	0.0329	0.0276
Difference	0.0010	0.0020	0.0034

Table 7: Estimated preposterior losses when  $T \sim \text{Gamma}(4, 0.4)$ .

	$\lambda_{i1}$	$\lambda_{i2}$	$\lambda_{i3}$
E[Var[ $\lambda_{ij} X_i(T)$ ]]	0.0362	0.0345	0.0330
E[Var[ $\lambda_{ij} X_i(u) : u \in [0, T]$ ]]	0.0333	0.0301	0.0268
Difference	0.0029	0.0044	0.0062

Table 8: Estimated preposterior losses when  $T \sim \text{Gamma}(8, 0.4)$ .

	$\lambda_{i1}$	$\lambda_{i2}$	$\lambda_{i3}$
E[Var[ $\lambda_{ij} X_i(T)$ ]]	0.0360	0.0351	0.0345
E[Var[ $\lambda_{ij} X_i(u) : u \in [0, T]$ ]]	0.0314	0.0290	0.0267
Difference	0.0046	0.0061	0.0078

Table 9: Estimated preposterior losses when  $T \sim \text{Gamma}(12, 0.4)$ .

We observe that when  $T \sim \text{Gamma}(4, 0.4)$ , the differences in preposterior losses between the unmonitored and the monitored case are not very large. The main reason for this is that the expected value of  $T$  is only 10. As we saw in Figure 4 and Figure 5 as well as in the analysis of the asymptotic behaviour of the two cases, the preposterior losses in the unmonitored case increase and converge to the prior loss when  $t$  increases, while the corresponding preposterior losses in the monitored case converges to a significantly lower value when  $t$  increases. Exactly the same effects are seen in the stochastic cases. Thus, the differences in preposterior losses increases considerably when  $T$  is sampled from the Gamma(12, 0.4)-distribution. These effects are intuitively very reasonable since the consequence of not monitoring a component obviously is worse if the length of the time interval with no information about the state of the component is increased.

## 6 Applications to multistate systems

In this section we show how the methodology introduced in the previous sections can be applied to multistate systems. As in the previous sections we let  $T_{ij}$  denote the time component  $i$  spends in state  $j$  during its first life cycle,  $j \in S_i$ ,  $i \in E$ . We assume that  $T_{ij}|\lambda_{ij} \sim \text{Exp}(\lambda_{ij})$  and that  $\lambda_{ij} \sim \text{Gamma}(\alpha_{ij}, \beta_{ij})$ ,  $j \in S_i$ ,  $i \in E$ .

The states of the individual components are not monitored continuously. Instead the system is left unmonitored until its state reaches a given critical level  $C$ . We denote the point of time when this happens by  $T$ . At that point of time the system is inspected and the states of the components are recorded. Thus, the situation is similar to the case considered in [Natvig and Gåsemyr 1998]. However, since we are considering multistate systems, the critical level  $C$  does not need to be zero.

Given the censoring point of time,  $T$ , and the states of the components at time  $T$ ,  $X_i(T)$ ,  $i \in E$ , the posterior distributions for the transition rates can in principle be estimated. However, in this case  $T$  is not independent of the other variables. In fact  $T$  is a function of the vectors  $\mathbf{T}_i$ ,  $i \in E$ . While it is not difficult to obtain an estimate of the marginal distribution of  $T$  using Monte Carlo simulation, it is more difficult to obtain the necessary conditional distributions needed for the inner loop in order to use Algorithm 5.2. Still this is at least possible by using e.g., the Metropolis-Hastings algorithm.

In this study we have chosen a simpler solution where we instead estimate the posterior distributions of the  $\lambda_{ij}$ 's given only the states  $X_i$ ,  $i \in E$  at the censoring point of time. This essentially implies that the exact value of  $T$  is ignored. It should be noted, however, that since the  $X_i$ 's obviously depend strongly on the value of  $T$ , the loss of information is not as great as one could suspect. In fact, as we shall see, the increases in preposterior losses even compared to the monitored case turn out to be small.

With this simplification the necessary data can be obtained by applying the following simulation algorithm:

**Algorithm 6.1** Repeat the following steps until a sufficient amount of samples has been generated:

STEP 2. Generate  $\lambda_i$  from the given priors for all  $i \in E$ .

STEP 3. Generate  $\mathbf{T}_i$  from the conditional distributions given  $\lambda_i$  for all  $i \in E$ .

STEP 4. Calculate the states of the components,  $X_i$  at the censoring point of time for all  $i \in E$ .

STEP 5. Add  $\lambda_i$  to the sampled vectors for this particular value of  $X_i$  for all  $i \in E$ , and keep track of the number of times the different states  $X_i$  occur.

In order to calculate the states of the components in Step 4 of Algorithm 6.1, a flexible and convenient approach is to use a discrete event model. An *event* in this context is a state transition for one of the components in the system. All such events are processed in chronological order using an *event queue*. As an event is processed, we also update the system state. When the system state reaches the critical level  $C$ , the processing stops and the relevant quantities are stored. This approach is easy to implement even for large complex multistate systems, and allows fast and efficient calculation of the states of the components. For further details see [Huseby et. al. 2010].

In Step 5 of Algorithm 6.1 we keep track of the number of times the different states  $X_i$  occur. As in the previous sections these counts are used to estimate the probability distribution of  $X_i$   $i \in E$ . The resulting preposterior losses can be calculated by the following simplified formula:

$$E[\text{Var}[\lambda_{ij}|X_i]] = \sum_{x=0}^s \text{Var}[\lambda_{ij}|X_i = x] \Pr(X_i = x), \quad j = 1, \dots, s. \quad (8)$$

While it is easy to apply this approach to arbitrary complex multistate systems, we have chosen to illustrate the methodology by considering a very simple multistate parallel system consisting of just two components. We assume that  $S_i = \{0, 1, 2, 3\}$ ,  $i = 1, 2$ , and that the state of the system at time  $t$  is given by the structure function  $\phi(t) = \phi(X_1(t), X_2(t)) = X_1(t) + X_2(t)$ . Thus,  $\phi(t) \in \{0, 1, \dots, 6\}$  for all  $t \geq 0$ . We assume that  $X_1(0) = X_2(0) = 3$ , implying that  $\phi(0) = 3 + 3 = 6$ . As time goes by the components go through their respective states until reaching the complete failure state 0. The parameter values of the prior distributions are listed in Table 10.

i	$\alpha_{i1}$	$\beta_{i1}$	$\alpha_{i2}$	$\beta_{i2}$	$\alpha_{i3}$	$\beta_{i3}$
1	1.0	4.0	1.0	4.0	1.0	4.0
2	1.0	6.0	1.0	6.0	1.0	6.0

Table 10: Parameters in the prior distributions

In the simulation study three different critical levels for the system are considered:  $C = 2$ ,  $C = 3$  and  $C = 4$ . As in the previous sections we start out by estimating the posterior losses given the state of the components. These results are given in Appendix B. Moreover, we estimate the probability distributions of  $X_i(T)$ ,  $i = 1, 2$ . The results for the three different critical values  $C = 2$ ,  $C = 3$  and  $C = 4$  are given in Table 11, Table 12 and Table 13 respectively.

	$x = 0$	$x = 1$	$x = 2$	$x = 3$
$P(X_1(T) = x)$	0.4484	0.2705	0.2811	0.0000
$P(X_2(T) = x)$	0.2811	0.2705	0.4484	0.0000

Table 11: Estimated probability distribution of  $X_i(t)$ ,  $i = 1, 2$  for  $C = 2$ .

	$x = 0$	$x = 1$	$x = 2$	$x = 3$
$P(X_1(T) = x)$	0.2444	0.3393	0.2755	0.1408
$P(X_2(T) = x)$	0.1408	0.2755	0.3393	0.2444

Table 12: Estimated probability distribution of  $X_i(t)$ ,  $i = 1, 2$  for  $C = 3$ .

	$x = 0$	$x = 1$	$x = 2$	$x = 3$
$P(X_1(T) = x)$	0.0000	0.3582	0.4134	0.2284
$P(X_2(T) = x)$	0.0000	0.2284	0.4134	0.3582

Table 13: Estimated probability distribution of  $X_i(t)$ ,  $i = 1, 2$  for  $C = 4$ .

Note that since the system is inspected when its state reaches level  $C$ , we always have:

$$P(\phi(T) = C) = P(X_1(T) + X_2(T) = C) = 1.$$

This implies that  $X_1(T) = x$  if and only if  $X_2(T) = C - x$ . Thus, the events  $X_1(T) = x$  and  $X_2(T) = C - x$  must have equal probability, for  $\max(0, C - 3) \leq x \leq \min(3, C)$ . By examining Table 11, Table 12 and Table 13 we see that this property is satisfied.

Also note that if  $C = 2$ , then  $X_i(T) \leq 2$ ,  $i = 1, 2$ . Thus, in this case  $P(X_i(T) = 3) = 0$ . Similarly, if  $C = 4$ , then  $X_i(T) \geq 1$ ,  $i = 1, 2$ . Thus, in this case  $P(X_i(T) = 0) = 0$ .

The preposterior losses can now be calculated using (8) by combining the estimated variances listed in Appendix B with the estimated probability distributions. The results are shown in Table 14 ( $C = 2$ ), Table 15 ( $C = 3$ ) and Table 16 ( $C = 4$ ). In these tables we have also included preposterior losses for the corresponding cases where the components are monitored continuously. As before these are calculated using (3), (4), (5) or (6).

The prior losses for component 1 are  $\text{Var}(\lambda_{1j}) = \alpha_{1j}/\beta_{1j}^2 = 1/4^2 = 0.0625$  for  $j = 1, 2, 3$ . The corresponding prior losses for component 2 are  $\text{Var}(\lambda_{2j}) = \alpha_{2j}/\beta_{2j}^2 = 1/6^2 = 0.0278$  for  $j = 1, 2, 3$  for component 2. That is, the prior losses are larger for component 1 than for component 2. As could be expected, these differences also affect the preposterior losses which are larger for component 1 than for component 2.

The differences between the unmonitored cases and the monitored cases are largest when  $C = 2$  and smallest when  $C = 4$ . This is intuitively reasonable since when  $C = 2$  the system is left unmonitored for a longer period of time compared to when  $C = 4$ . We also note that the differences in preposterior losses are larger for  $\lambda_{1,3}$  and  $\lambda_{2,3}$  compared to e.g.,  $\lambda_{1,1}$  and  $\lambda_{2,1}$ . The reason for this is that both the components start out in state 3. Thus, when monitoring the components one would with a very high probability observe the exact value of  $T_{1,3}$  and  $T_{2,3}$ . On the other hand, when the components are not monitored, this information will be lost.

Finally, we note that when  $C = 4$ , we have  $X_1(T) + X_2(T) = 4$ . From this it follows that we must have  $X_i(T) \geq 1$ ,  $i = 1, 2$ . Thus, there will never be a transition from state 1 to state 0 for any component within the interval  $[0, T]$ . In fact, at time  $T$  when the system reaches state  $C = 4$ , we would not have any new information about  $\lambda_{1,1}$  or  $\lambda_{2,1}$ . Hence, the preposterior losses for these parameters are equal to the corresponding prior losses, respectively 0.0625 for component 1 and 0.0278 for component 2, both in the unmonitored case and the monitored case.

	$\lambda_{i1}$	$\lambda_{i2}$	$\lambda_{i3}$
$E[\text{Var}[\lambda_{1j} X_1(T)]]$	0.0602	0.0596	0.0597
$E[\text{Var}[\lambda_{1j} X_1(u) : u \in [0, T]]]$	0.0511	0.0451	0.0416
Difference	0.0091	0.0145	0.0181
$E[\text{Var}[\lambda_{2j} X_2(T)]]$	0.0270	0.0265	0.0263
$E[\text{Var}[\lambda_{2j} X_2(u) : u \in [0, T]]]$	0.0244	0.0213	0.0185
Difference	0.0026	0.0052	0.0078

Table 14: Estimated preposterior losses when  $C = 2$ .

	$\lambda_{i1}$	$\lambda_{i2}$	$\lambda_{i3}$
$E[\text{Var}[\lambda_{1j} X_1(T)]]$	0.0613	0.0584	0.0565
$E[\text{Var}[\lambda_{1j} X_1(u) : u \in [0, T]]]$	0.0559	0.0475	0.0419
Difference	0.0054	0.0109	0.0146
$E[\text{Var}[\lambda_{2j} X_2(T)]]$	0.0274	0.0263	0.0242
$E[\text{Var}[\lambda_{2j} X_2(u) : u \in [0, T]]]$	0.0260	0.0226	0.0189
Difference	0.0014	0.0037	0.0053

Table 15: Estimated preposterior losses when  $C = 3$ .

	$\lambda_{i1}$	$\lambda_{i2}$	$\lambda_{i3}$
$E[\text{Var}[\lambda_{1j} X_1(T)]]$	0.0625	0.0600	0.0554
$E[\text{Var}[\lambda_{1j} X_1(u) : u \in [0, T]]]$	0.0625	0.0524	0.0427
Difference	0.0000	0.0076	0.0127
$E[\text{Var}[\lambda_{2j} X_2(T)]]$	0.0278	0.0268	0.0242
$E[\text{Var}[\lambda_{2j} X_2(u) : u \in [0, T]]]$	0.0278	0.0247	0.0195
Difference	0.0000	0.0021	0.0047

Table 16: Estimated preposterior losses when  $C = 4$ .

## 7 Conclusions

In the present paper we have shown how to quantify the effect of partial monitoring of multistate components in a multistate system. The quantification is done by calculating *preposterior loss*.



As part of the process of calculating preposterior losses we also show how to compute posterior distributions under various assumptions. This methodology provides a framework for determining an optimal monitoring plan balancing the cost of monitoring with the loss in information. It is e.g., possible to identify which of the components that should be given priority with respect to monitoring.

In cases where all components in the system are monitored continuously, it is often possible to calculate the posterior loss analytically. In the unmonitored case, however, the calculations must typically be based on Monte Carlo simulation. Fortunately, the simulation methods can be done very efficiently even for large complex multistate systems by using discrete event simulation.

In all our calculations we have assumed that the components are stochastically independent of each other. A more realistic model should allow for various types of dependence, both physical dependence as well as dependence due to similar components sharing the same transition rates. The model should ideally also allow for analysis of more than one life cycle, and include perfect or imperfect repairs.

## References

- [Barlow and Wu 1981] Barlow, R. E. and Wu, A. S. Preposterior Analysis of Bayes Estimators of Mean Life *Biometrika* (68): 403–410, 1981.
- [Berger 1985] Berger, J. O. *Statistical Decision Theory and Bayesian Analysis* Springer-Verlag, 1985.
- [Curcurù et. al. 2010] Curcurù, G., Galante, G. and Lombardo, A. A predictive maintenance policy with imperfect monitoring. *Reliability Engineering & System Safety*, 95 (9): 989–997, 2010.
- [Gåsemyr and Natvig 2001] Gåsemyr, J. and Natvig, B. Bayesian inference based on partial monitoring of components with applications to preventive system maintenance. *Naval Research Logistics*, (48): 551–577, 2001.
- [Gåsemyr and Natvig 2005] Gåsemyr, J. and Natvig, B. Probabilistic modelling of monitoring and maintenance of multistate monotone systems with dependent components. *Methodology and Computing in Applied Probability* (7): 63–78, 2005.
- [Huang and Wang 2016] Huang, C.H. and Wang C.H. Optimization of preventive maintenance for a multi-state degraded system by monitoring component performance. *Journal of Intelligent Manufacturing*, 27 (6): 1151–1170, 2016.
- [Huseby et. al. 2010] Huseby, A., Natvig, B., Gåsemyr, J., Skutlaberg, K. and Isaksen, S. Advanced discrete event simulation methods with application to importance measure estimation in reliability. In: *Discrete Event Simulation*, INTECH, 205–222, 2010.
- [Lisnianski et. al. 2010] Lisnianski, A., Frenkel, I. and Ding, Y. *Multi-state system reliability analysis and optimization for engineers and industrial managers*. London: Springer, 2010.
- [Liu 2010] Liu, Y. Optimal Replacement Policy for MultiState System Under Imperfect Maintenance. *IEEE Transactions on Reliability*, 59(3): 483–495, 2010.
- [Natvig and Gåsemyr 1998] Natvig, B. and Gåsemyr, J. The posterior distribution of the parameters of component lifetimes based on autopsy data in a shock model. *Scandinavian Journal of Statistics*, 25.(2): 271–292, 1998.
- [Natvig et. al. 2009] Natvig, B., Eide, K. A., Gåsemyr, J., Huseby, A. and Isaksen, S. Simulation based analysis and an application to an offshore oil and gas production system of the Natvig measures of component importance in repairable systems. *Reliability Engineering & System Safety*, 94.(10): 1629–1638, 2009.

- [Natvig 2011] Natvig, B. *Multistate Systems Reliability Theory with Applications*. John Wiley & Sons, Wiley Series in Probability and Statistics, 2011.
- [Raïffa and Schlaifer 1961] Raïffa, H. and Schlaifer, R. *Applied Statistical Decision Theory*. Harvard University and MIT Press, 1961.
- [Skutlaberg and Natvig 2016] Skutlaberg, K. and Natvig, B. Minimization of the Expected Total Net Loss in a Stationary Multistate Flow Network System. *Applied Mathematics*, 7.(8): 793–817, 2016.

## A Estimated variances from Section 3

	$\lambda_{i1}$	$\lambda_{i2}$	$\lambda_{i3}$
$X_i(t) = 0$	0.0581	0.0588	0.0584
$X_i(t) = 1$	0.0238	0.0532	0.0533
$X_i(t) = 2$	0.0400	0.0192	0.0471
$X_i(t) = 3$	0.0400	0.0400	0.0100

Table 17:  $\text{Var}[\lambda_{ij}|X_i(t) = x]$  for  $t = 5$ .

	$\lambda_{i1}$	$\lambda_{i2}$	$\lambda_{i3}$
$X_i(t) = 0$	0.0563	0.0563	0.0563
$X_i(t) = 1$	0.0232	0.0507	0.0507
$X_i(t) = 2$	0.0400	0.0185	0.0435
$X_i(t) = 3$	0.0400	0.0400	0.0100

Table 18:  $\text{Var}[\lambda_{ij}|X_i(u) : \text{for all } u \in [0, t], X_i(t) = x]$  for  $t = 5$ .

	$\lambda_{i1}$	$\lambda_{i2}$	$\lambda_{i3}$
$X_i(t) = 0$	0.0506	0.0506	0.0506
$X_i(t) = 1$	0.0169	0.0450	0.0453
$X_i(t) = 2$	0.0400	0.0123	0.0397
$X_i(t) = 3$	0.0400	0.0400	0.0044

Table 19:  $\text{Var}[\lambda_{ij}|X_i(t) = x]$  for  $t = 10$ .

	$\lambda_{i1}$	$\lambda_{i2}$	$\lambda_{i3}$
$X_i(t) = 0$	0.0465	0.0465	0.0465
$X_i(t) = 1$	0.0156	0.0398	0.0397
$X_i(t) = 2$	0.0400	0.0112	0.0334
$X_i(t) = 3$	0.0400	0.0400	0.0044

Table 20:  $\text{Var}[\lambda_{ij}|X_i(u) : \text{for all } u \in [0, t], X_i(t) = x]$  for  $t = 10$ .

	$\lambda_{i1}$	$\lambda_{i2}$	$\lambda_{i3}$
$X_i(t) = 0$	0.0469	0.0471	0.0467
$X_i(t) = 1$	0.0132	0.0412	0.0413
$X_i(t) = 2$	0.0400	0.0090	0.0372
$X_i(t) = 3$	0.0400	0.0400	0.0025

Table 21:  $\text{Var}[\lambda_{ij}|X_i(t) = x]$  for  $t = 15$ .

	$\lambda_{i1}$	$\lambda_{i2}$	$\lambda_{i3}$
$X_i(t) = 0$	0.0411	0.0412	0.0412
$X_i(t) = 1$	0.0116	0.0342	0.0343
$X_i(t) = 2$	0.0400	0.0077	0.0291
$X_i(t) = 3$	0.0400	0.0400	0.0025

Table 22:  $\text{Var}[\lambda_{ij}|X_i(u)]$  : for all  $u \in [0, t]$ ,  $X_i(t) = x$  for  $t = 15$ .

## B Estimated variances and posterior distributions from Section 6

	$\lambda_{1,1}$	$\lambda_{1,2}$	$\lambda_{1,3}$
$X_1(T) = 0$	0.0676	0.0673	0.0670
$X_1(T) = 1$	0.0466	0.0622	0.0631
$X_1(T) = 2$	0.0616	0.0448	0.0448
$X_1(T) = 3$	n/a	n/a	n/a

Table 23:  $\text{Var}[\lambda_{1j}|X_1(T) = x]$ ,  $C = 2$ .

	$\lambda_{2,1}$	$\lambda_{2,2}$	$\lambda_{2,3}$
$X_2(t) = 0$	0.0314	0.0312	0.0313
$X_2(t) = 1$	0.0212	0.0300	0.0295
$X_2(t) = 2$	0.0277	0.0214	0.0217
$X_2(t) = 3$	n/a	n/a	n/a

Table 24:  $\text{Var}[\lambda_{2j}|X_2(T) = x]$ ,  $C = 2$ .

	$\lambda_{1,1}$	$\lambda_{1,2}$	$\lambda_{1,3}$
$X_1(T) = 0$	0.0699	0.0692	0.0700
$X_1(T) = 1$	0.0539	0.0664	0.0661
$X_1(T) = 2$	0.0623	0.0376	0.0578
$X_1(T) = 3$	0.0625	0.0613	0.0075

Table 25:  $\text{Var}[\lambda_{1j}|X_1(T) = x]$ ,  $C = 3$ .

	$\lambda_{2,1}$	$\lambda_{2,2}$	$\lambda_{2,3}$
$X_2(T) = 0$	0.0320	0.0320	0.0321
$X_2(T) = 1$	0.0243	0.0318	0.0317
$X_2(T) = 2$	0.0278	0.0182	0.0279
$X_2(T) = 3$	0.0276	0.0281	0.0060

Table 26:  $\text{Var}[\lambda_{2j}|X_2(T) = x]$ ,  $C = 3$ .

	$\lambda_{1,1}$	$\lambda_{1,2}$	$\lambda_{1,3}$
$X_1(T) = 0$	n/a	n/a	n/a
$X_1(T) = 1$	0.0620	0.0715	0.0712
$X_1(T) = 2$	0.0625	0.0485	0.0638
$X_1(T) = 3$	0.0619	0.0629	0.0152

Table 27:  $\text{Var}[\lambda_{1j}|X_1(T) = x]$ ,  $C = 4$ .

	$\lambda_{2,1}$	$\lambda_{2,2}$	$\lambda_{2,3}$
$X_2(T) = 0$	n/a	n/a	n/a
$X_2(T) = 1$	0.0275	0.0333	0.0337
$X_2(T) = 2$	0.0281	0.0226	0.0313
$X_2(T) = 3$	0.0279	0.0276	0.0100

Table 28:  $\text{Var}[\lambda_{2j}|X_2(T) = x]$ ,  $C = 4$ .

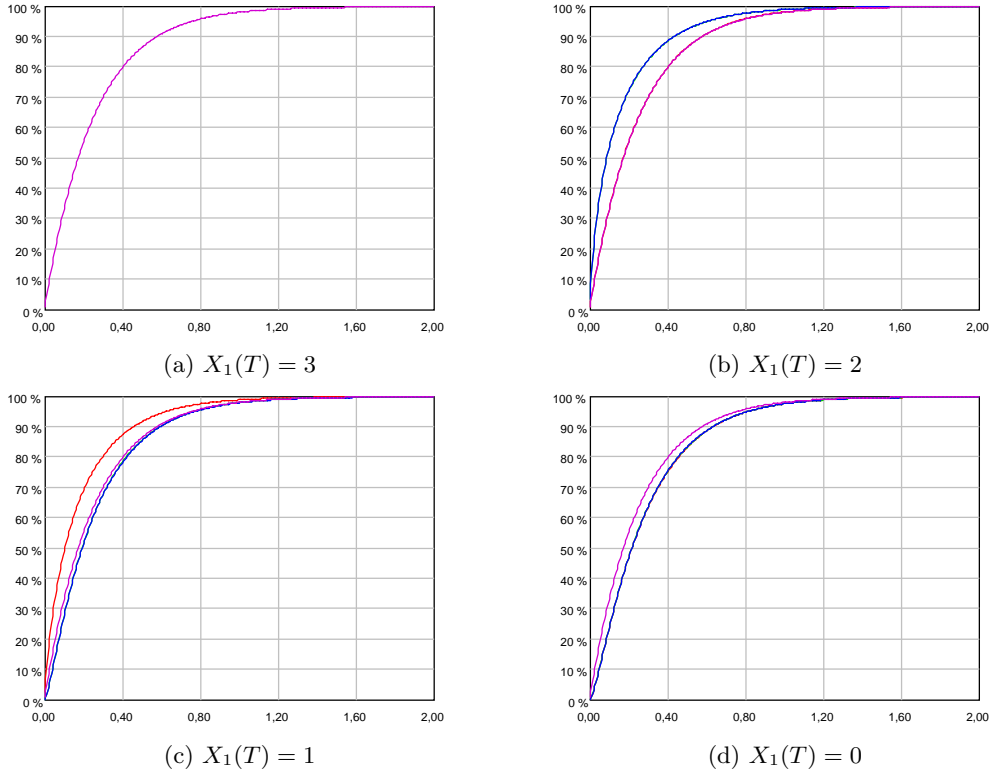


Figure 6: Posterior distributions for  $C = 2$  for  $\lambda_{1j}$ ,  $j = 1, 2, 3$ , (red, green and blue curves respectively) given that  $X_1(t) = 0, 1, 2, (3)$  along with the common prior distribution,  $\text{Gamma}(1, 4)$  (violet curve).

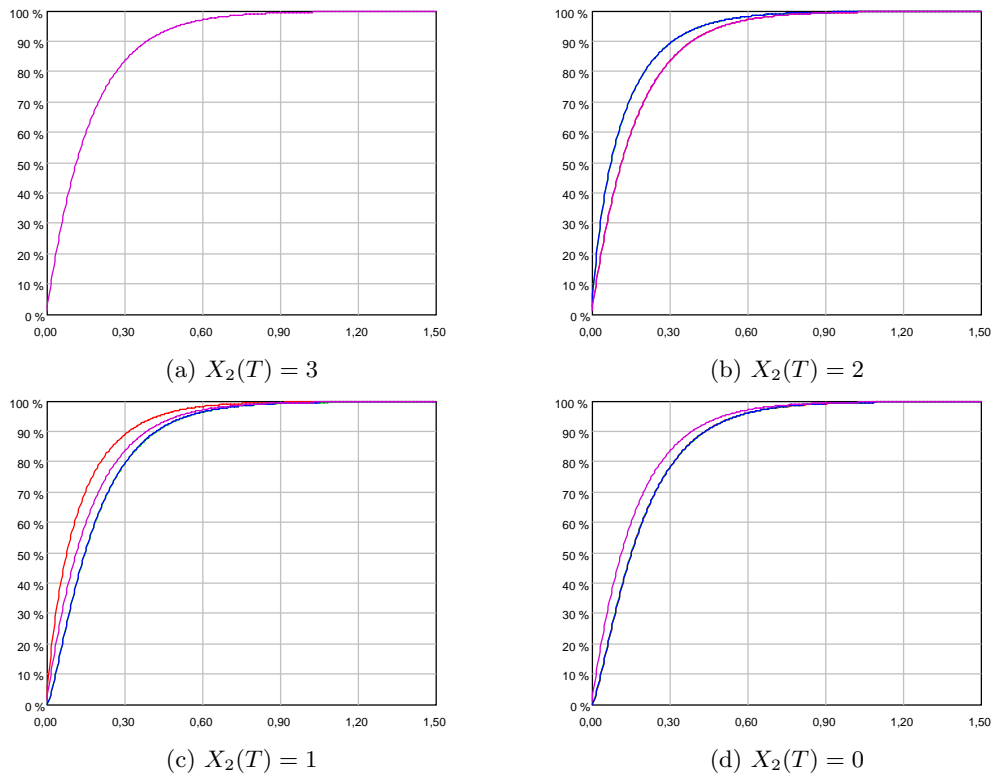


Figure 7: Posterior distributions for  $C = 2$  for  $\lambda_{2j}$ ,  $j = 1, 2, 3$ , (red, green and blue curves respectively) given that  $X_2(t) = 0, 1, 2, (3)$  along with the common prior distribution, Gamma(1, 6) (violet curve).

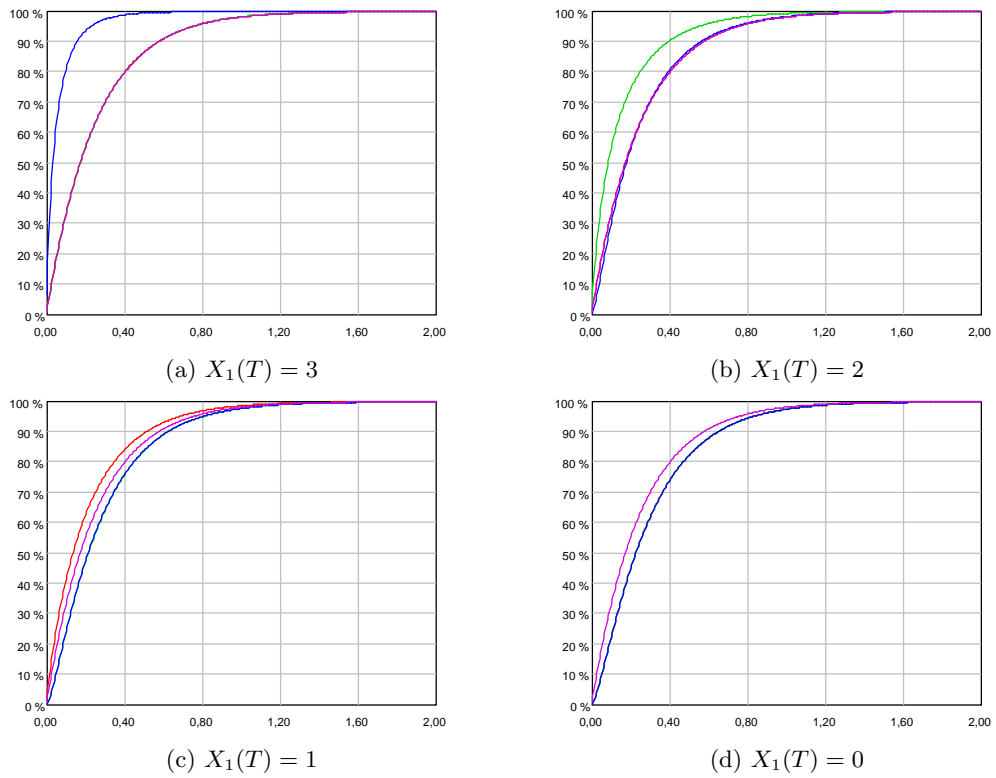


Figure 8: Posterior distributions for  $C = 3$  for  $\lambda_{1j}$ ,  $j = 1, 2, 3$ , (red, green and blue curves respectively) given that  $X_1(t) = 0, 1, 2, 3$  along with the common prior distribution,  $\text{Gamma}(1, 4)$  (violet curve).

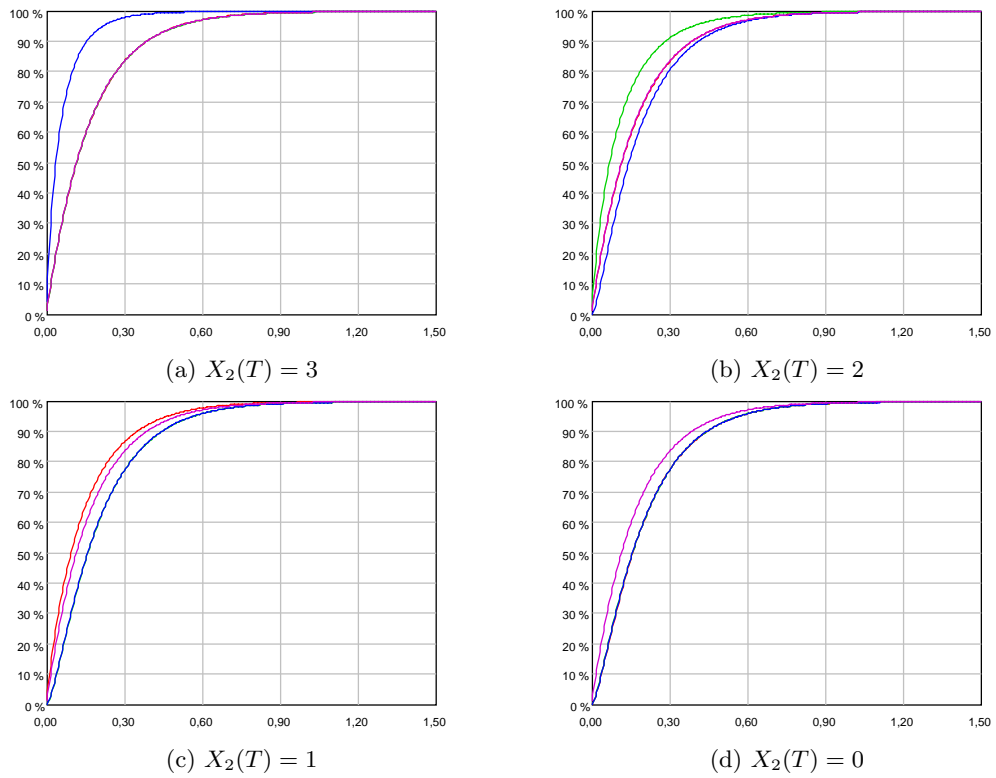


Figure 9: Posterior distributions for  $C = 3$  for  $\lambda_{2j}$ ,  $j = 1, 2, 3$ , (red, green and blue curves respectively) given that  $X_2(t) = 0, 1, 2, 3$  along with the common prior distribution,  $\text{Gamma}(1, 6)$  (violet curve).



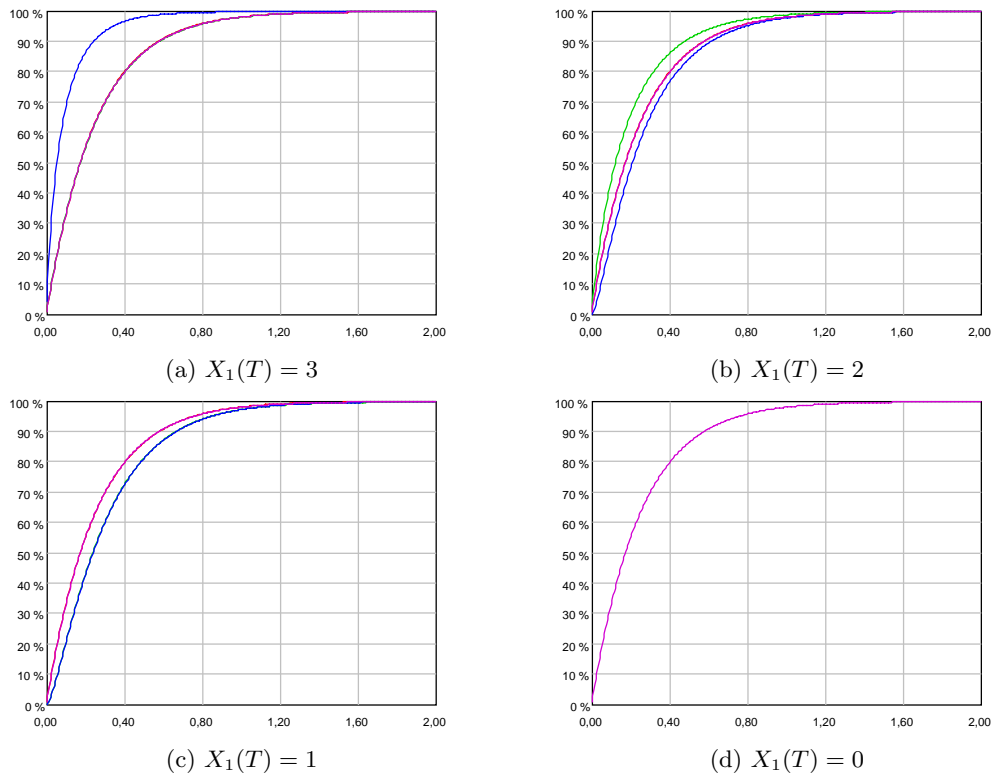


Figure 10: Posterior distributions for  $C = 4$  for  $\lambda_{1j}$ ,  $j = 1, 2, 3$ , (red, green and blue curves respectively) given that  $X_1(t) = (0), 1, 2, 3$  along with the common prior distribution, Gamma(1, 4) (violet curve).

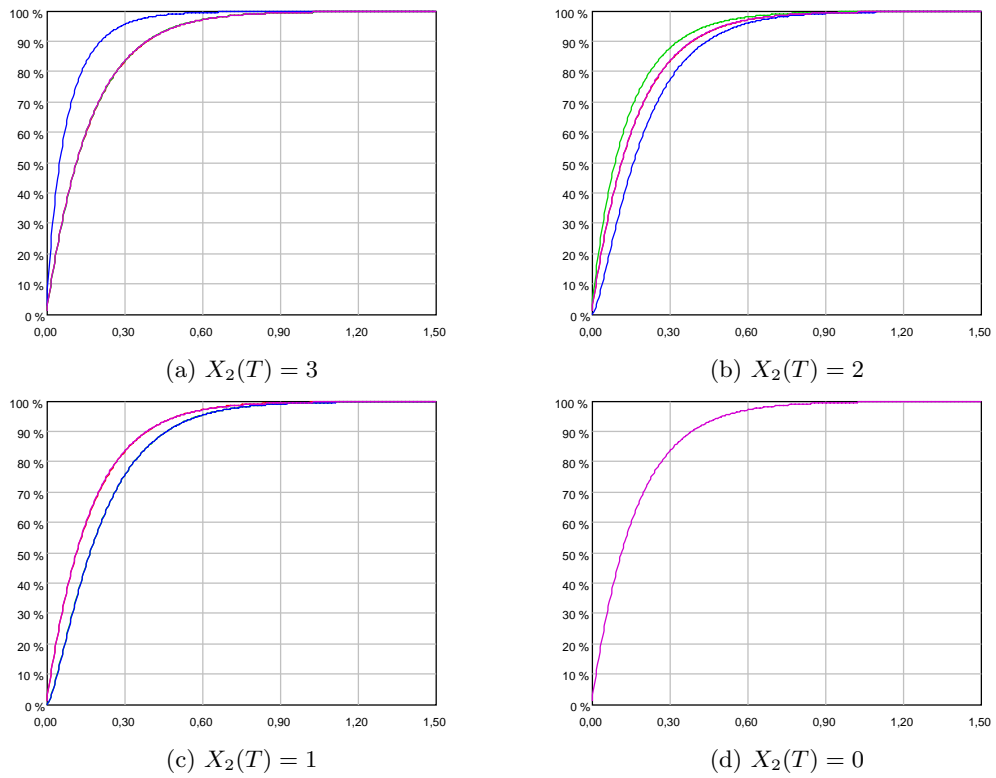


Figure 11: Posterior distributions for  $C = 4$  for  $\lambda_{2j}$ ,  $j = 1, 2, 3$ , (red, green and blue curves respectively) given that  $X_2(t) = (0), 1, 2, 3$  along with the common prior distribution, Gamma(1, 6) (violet curve).