Imprecise system reliability using the survival signature

Coolen, F.P.A.; Coolen-Maturi, T.; Aslett, L.; Walter, G.M.

Published in:
ICAMER'16 : Proceedings of the 1st International Conference on Applied Mathematics in Engineering and Reliability. 4-6 May 2016, Ho Chi Minh City, Vietnam

Document license:
Unspecified

Published: 26/01/2016

Document Version
Accepted manuscript including changes made at the peer-review stage

Imprecise system reliability using the survival signature

Frank P.A. Coolen
Department of Mathematical Sciences
Durham University, Durham, United Kingdom

Tahani Coolen-Maturi
Durham University Business School
Durham University, Durham, United Kingdom

Louis J.M. Aslett
Department of Statistics
Oxford University, Oxford, United Kingdom

Gero Walter
School of Industrial Engineering
Eindhoven University of Technology, Eindhoven, The Netherlands

ABSTRACT: The survival signature has been introduced to simplify quantification of reliability of systems which consist of components of different types, with multiple components of at least one of these types. The survival signature generalizes the system signature, which has attracted much interest in the theoretical reliability literature but has limited practical value as it can only be used for systems with a single type of components. The key property for uncertainty quantification of the survival signature, in line with the signature, is full separation of aspects of the system structure and failure times of the system components. This is particularly useful for statistical inference on the system reliability based on component failure times.
This paper provides a brief overview of the survival signature and its use for statistical inference for system reliability. We show the application of generalized Bayesian methods and nonparametric predictive inference, both these inference methods use imprecise probabilities to quantify uncertainty, where imprecision reflects the amount of information available. The paper ends with a discussion of related research challenges.

1 INTRODUCTION

In mathematical theory of reliability the main focus is on the functioning of a system given the functioning, or not, of its components and the structure of the system. The mathematical concept which is central to this theory is the *structure function* (Barlow & Proschan 1975). For a system with m components, the state vector is $\underline{x} = (x_1, x_2, \ldots, x_m) \in \{0,1\}^m$, with $x_i = 1$ if the ith component functions and $x_i = 0$ if not. The labelling of the components is arbitrary but must be fixed to define $\underline{x}$. The structure function $\phi : \{0,1\}^m \rightarrow \{0,1\}$, defined for all possible $\underline{x}$, takes the value 1 if the system functions and 0 if the system does not function for state vector $\underline{x}$. Most practical systems are coherent, which means that $\phi(\underline{x})$ is not decreasing in any of the components of $\underline{x}$, so system functioning cannot be improved by worse performance of one or more of its components. The assumption of coherent systems is made throughout this paper and is convenient from the perspective of uncertainty quantification for system reliability. It is further logical to assume that $\phi(\underline{0}) = 0$ and $\phi(\underline{1}) = 1$, so the system fails if all its components fail and it functions if all its components function.

For larger systems, working with the full structure function may be complicated, and one may particularly only need a summary of the structure function in case the system has exchangeable components of one or more types. We use the term 'exchangeable components' to indicate that the failure times of the components in the system are exchangeable (De Finetti 1974). Recently, we introduced such a summary, called the *survival signature*, to facilitate reliability analyses for systems with multiple types of components (Coolen & Coolen-Maturi 2012). In

case of just a single type of components, the survival signature is closely related to the system signature (Samaniego 2007), which is well-established and the topic of many research papers during the last decade. However, generalization of the signature to systems with multiple types of components is extremely complicated (as it involves ordering order statistics of different distributions), so much so that it cannot be applied to most practical systems. In addition to the possible use for such systems, where the benefit only occurs if there are multiple components of the same types, the survival signature is arguably also easier to interpret than the signature.

Consider a system with $K \geq 1$ types of components, with m_k components of type $k \in \{1, \ldots, K\}$ and $\sum_{k=1}^{K} m_k = m$. Assume that the random failure times of components of the same type are exchangeable (De Finetti 1974). Due to the arbitrary ordering of the components in the state vector, components of the same type can be grouped together, leading to a state vector that can be written as $\underline{x} = (\underline{x}^1, \underline{x}^2, \ldots, \underline{x}^K)$, with $\underline{x}^k = (x_1^k, x_2^k, \ldots, x_{m_k}^k)$ the subvector representing the states of the components of type k.

The *survival signature* for such a system, denoted by $\Phi(l_1, \ldots, l_K)$, with $l_k = 0, 1, \ldots, m_k$ for $k = 1, \ldots, K$, is defined as the probability for the event that the system functions given that *precisely* l_k of its m_k components of type k function, for each $k \in \{1, \ldots, K\}$ (Coolen & Coolen-Maturi 2012). There are $\binom{m_k}{l_k}$ state vectors $\underline{x}^k$ with $\sum_{i=1}^{m_k} x_i^k = l_k$. Let $S_{l_k}^k$ denote the set of these state vectors for components of type k and let $S_{l_1,\ldots,l_K}$ denote the set of all state vectors for the whole system for which $\sum_{i=1}^{m_k} x_i^k = l_k$, $k = 1, \ldots, K$. We also introduce the notation $\underline{l} = (l_1, \ldots, l_K)$. Due to the exchangeability assumption for the failure times of the m_k components of type k, all the state vectors $\underline{x}^k \in S_{l_k}^k$ are equally likely to occur, hence (Coolen & Coolen-Maturi 2012)

$$\Phi(\underline{l}) = \left[\prod_{k=1}^{K} \binom{m_k}{l_k}^{-1} \right] \times \sum_{\underline{x} \in S_{l_1,\ldots,l_K}} \phi(\underline{x})$$

Let $C_t^k \in \{0, 1, \ldots, m_k\}$ denote the number of components of type k in the system that function at time $t > 0$. Then, for system failure time T_S,

$$P(T_S > t) = \sum_{l_1=0}^{m_1} \cdots \sum_{l_K=0}^{m_K} \Phi(\underline{l}) P(\bigcap_{k=1}^{K} \{C_t^k = l_k\})$$

There are no restrictions on dependence of the failure times of components of different types, as the probability $P(\bigcap_{k=1}^{K} \{C_t^k = l_k\})$ can take any form of dependence into account, for example one can include common-cause failures quite straightforwardly into this approach (Coolen & Coolen-Maturi 2015). However, there is a substantial simplification if one assumes that the failure times of components of different types are independent, and even more so if one assumes that the failure times of components of type k are conditionally independent and identically distributed with CDF $F_k(t)$. With these assumptions, we get

$$P(T_S > t) = \sum_{l_1=0}^{m_1} \cdots \sum_{l_K=0}^{m_K} \Phi(\underline{l}) \times$$

$$\prod_{k=1}^{K} \binom{m_k}{l_k} [F_k(t)]^{m_k - l_k} [1 - F_k(t)]^{l_k} \qquad (1)$$

The main advantage of the survival signature, in line with this property of the signature for systems with a single type of components (Samaniego 2007), is that the information about the system structure is fully separated from the information about functioning of the components, which simplifies related statistical inference as well as considerations of optimal system design. In particular for study of system reliability over time, with the structure of the system, and hence the survival signature, not changing, this separation also enables relatively straightforward statistical inferences where even the use of imprecise probabilistic methods (Augustin, Coolen, de Cooman, & Troffaes 2014, Coolen & Utkin 2011) is quite straightforward. Such methods have the advantage that imprecision for the system survival function reflects the amount of information available. The next two sections briefly discuss such methods of statistical inference for the system failure time. First we show an application of generalized Bayesian methods, with a set of prior distributions instead of a single prior distribution. This is followed by a brief discussion and application of nonparametric predictive inference (Coolen 2011), a frequentist statistical method which is based on relatively few assumptions, enabled through the use of imprecise probabilities, and which does not require the use of prior distributions. The paper ends with a brief discussion of research challenges, particularly with regard to upscaling the survival signature methodology for application to large-scale real-world systems and networks.

2 IMPRECISE BAYESIAN INFERENCE

The reliability of a system, for which the survival signature is available, is quite straightforwardly quantified through its survival function, as shown in the previous section. We briefly consider a scenario where we have test data that enable learning about the reliability of the components of different types in the system, where we assume independence of the failure times of components of different types. The numbers of components in the system, of each type, that are functioning at time t, denoted by C_t^k for $k = 1, \ldots, K$, are the random quantities of main interest. One attractive statistical method to learn about these random quantities from test data is provided by the Bayesian

framework of statistics, which can be applied with the assumption of a parametric distribution for the component failure times (Walter, Graham, & Coolen 2015) or in a nonparametric manner (Aslett, Coolen, & Wilson 2015). We briefly illustrate the latter approach.

Assume that there are m_k components of type k in the system, and we are interested in the probability distribution of C_t^k. Suppose that n_k components of the same type k were tested, these are not the components that are in the system but their failure times are assumed to be exchangeable with those in the system. We assume that for all tested components the failure time has been observed, let s_t^k denote the number of these components that still functioned at time t. A convenient and basic model for C_t^k is the Binomial distribution, where the probability of 'success', that is a component still to be functioning at time t, can, in the Bayesian framework, be conveniently modelled as a random quantity with a Beta prior distribution. Different to the standard parameterization for the Beta distribution, we define a Beta prior distribution through parameters $n_{k,t}^{(0)}$ and $y_{k,t}^{(0)}$ with as interpretations a pseudocount of components and the expected value of the success probability, respectively. Hence, these parameters can be interpreted in the sense that the prior distribution represents beliefs reflecting the same information as would result from observing $n_{k,t}^{(0)}$ components of which $n_{k,t}^{(0)}y_{k,t}^{(0)}$ still function at time t (Walter 2013). Doing this leads to straightforward updating, using the test information consisting of observations of n_k components of which $s_{k,t}$ were still functioning at time t. The updating results in a similar Beta distribution as the prior, but now with parameter values $n_{k,t}^{(n)} = n_{k,t}^{(0)} + n_k$ and $y_{k,t}^{(n)}$ the weighted average of $y_{k,t}^{(0)}$ and $s_{k,t}/n_{k,t}$, with weights proportional to $n_{k,t}^{(0)}$ and n_k, respectively. This leads to the posterior predictive distribution (Walter, Aslett, & Coolen 2016)

$$P(C_t^k = l_k \mid s_t^k) = \binom{m_k}{l_k} \times$$

$$\frac{B(l_k + n_{k,t}^{(n)}y_{k,t}^{(n)}, m_k - l_k + n_{k,t}^{(n)}(1 - y_{k,t}^{(n)}))}{B(n_{k,t}^{(n)}y_{k,t}^{(n)}, n_{k,t}^{(n)}(1 - y_{k,t}^{(n)}))}$$

This model can also relatively straightforwardly be used with a set of Beta prior distributions rather than a single one, a generalization fitting in the theory of imprecise probability (Augustin, Coolen, de Cooman, & Troffaes 2014). At each value of t one calculates the infimum and supremum of the probability $P(C_t^k = l_k \mid s_t^k)$ over the set of prior parameters, with $n_{k,t}^{(0)} \in [\underline{n}_{k,t}^{(0)}, \overline{n}_{k,t}^{(0)}]$ and $y_{k,t}^{(0)} \in [\underline{y}_{k,t}^{(0)}, \overline{y}_{k,t}^{(0)}]$, with the bounds of these intervals chosen to reflect a priori available knowledge and its limitations. The use of such prior sets, with only an interval of possible values specified for each parameter, provides much flexibility for modelling prior beliefs and indeterminacy, together with interesting ways in which the corresponding sets of posterior (predictive) distributions and related inferences can vary. Most noticeably, this model enables conflict between prior beliefs and data to be shown through increased imprecision, that is difference between upper and lower probabilities for an event of interest (Walter 2013). We illustrate the use of this model, together with the survival signature, for a small system in Example 1, without attention to such prior-data conflict, further details on this will be presented elsewhere (Walter, Aslett, & Coolen 2016).

Example 1

As a small example, consider the system with three types of components presented in Figure 1. The survival signature of this system is given in Table 1, where all cases with $l_3 = 0$ have been omitted as the system cannot function if the component of Type 3 does not function, hence $\Phi(l_1, l_2, 0) = 0$ for all (l_1, l_2).

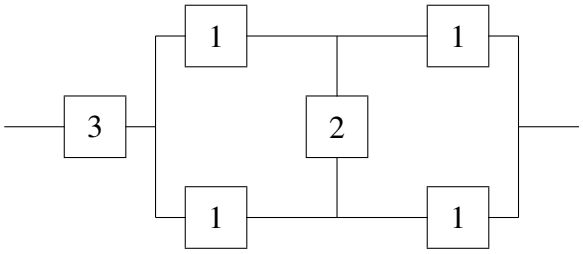

Figure 1: System with 3 types of components

l_1	l_2	$\Phi(l_1, l_2, 1)$	l_1	l_2	$\Phi(l_1, l_2, 1)$
0	0	0	0	1	0
1	0	0	1	1	0
2	0	1/3	2	1	2/3
3	0	1	3	1	1
4	0	1	4	1	1

Table 1: Survival signature for the system in Figure 1 for cases with $l_3 = 1$.

For component types 1 and 2, we consider a near-noninformative set of prior survival functions. For components of type 3, we consider an informative set of prior survival functions as given in Table 2. This set could result from eliciting prior survival probabilities at times $t = 0, 1, 2, 3, 4, 5$ only, and using those values to deduce such prior probabilities for all other values of t without further assumptions. These prior assumptions, together with sets of posterior survival functions, are illustrated in Figure 3 (presented at the end of the paper); test data for components of type 1 and 2 are taken as $\{2.2, 2.4, 2.6, 2.8\}$ and

t	$[0,1)$	$[1,2)$	$[2,3)$	$[3,4)$	$[4,5)$
$\underline{y}_{3,t}^{(0)}$	0.625	0.375	0.250	0.125	0.010
$\overline{y}_{3,t}^{(0)}$	0.999	0.875	0.500	0.375	0.250

Table 2: Lower and upper prior functioning probability bounds for component type 3 in the system of Figure 1.

$\{3.2, 3.4, 3.6, 3.8\}$, respectively. For components of type 3 test data are taken as $\{0.5, 1.5, 2.5, 3.5\}$, which are well in line with expectations according to the set of prior distributions. The posterior sets of survival functions for each component type and for the whole system show considerably smaller imprecision than the corresponding prior sets, which is mainly due to the low prior strength intervals we chose for this example, namely $[\underline{n}_{1,t}^{(0)}, \overline{n}_{1,t}^{(0)}] = [\underline{n}_{2,t}^{(0)}, \overline{n}_{2,t}^{(0)}] = [1, 2]$, $[\underline{n}_{3,t}^{(0)}, \overline{n}_{3,t}^{(0)}] = [1, 4]$, for all t. We see that posterior lower and upper survival functions drop at those times t when there is a failure time in the test data, or a drop in the prior survival probability bounds. Note that the lower bound for prior system survival function is zero for all t due to the prior lower bound of zero for type 1 components, and for the system to function at least two components of type 1 must function. A further reason why the imprecision reduces substantially in this example is that the data do not conflict with the prior beliefs. With these sets of prior distributions such prior-data conflict can only really occur for components of type 3, as such conflict logically requires at least reasonably strong prior beliefs to be taken into account through the set of prior distributions. If test failure times for the components of type 3 were unexpectedly small or large, the imprecision in the lower and upper posterior survival functions for this component would increase, with a similar effect on the corresponding overall lower and upper system survival functions. A detailed analysis illustrating this effect will be presented elsewhere (Walter, Aslett, & Coolen 2016).

3 NONPARAMETRIC PREDICTIVE INFERENCE

Nonparametric predictive inference (NPI) (Coolen 2011) is a frequentist statistical framework based on relatively few assumptions and considering events of interest which are explicitly in terms of one or more future observations. NPI can be considered suitable if there is hardly any knowledge about the random quantity of interest, other than the data which we assume consist of n observations, or if one does not want to use such further information, e.g. to study effects of additional assumptions underlying other statistical methods. NPI uses lower and upper probabilities, also known as imprecise probabilities, to quantify uncertainty (Augustin, Coolen, de Cooman, & Troffaes 2014) and has strong consistency properties from frequentist statistics perspective (Augustin & Coolen 2004, Coolen 2011). NPI provides a solution to some explicit goals formulated for objective (Bayesian) inference, which cannot be obtained when using precise probabilities (Coolen 2006), and it never leads to results that are in conflict with inferences based on empirical probabilities. NPI for system survival functions, using the survival signature, was recently presented (Coolen, Coolen-Maturi, & Al-nefaiee 2014) and is briefly summarized here.

We now present NPI lower and upper survival functions for the failure time T_S of a system consisting of multiple types of components, using the system signature combined with NPI for Bernoulli data (Coolen 1998). This enables the NPI method to be applied to, in principle, all systems. The failure times of components of different types are assumed to be independent. It must be emphasized that the NPI framework does not assume an underlying population distribution in relation to random quantities, and therefore also not that these are conditionally independent given some probability distribution. In fact, NPI explicitly takes the inter-dependence of multiple future observations into account. This requires a somewhat different approach for dealing with imprecise probabilities to that presented for the imprecise Bayesian approach in the previous section.

NPI will be used for learning about the components of a specific type in the system, from data consisting of failure times for components that are exchangeable with these. We assume therefore that such data are available, for example resulting from testing or previous use of such components. It is assumed that failure times are available for all tested components. As in the previous section, let n_k, for $k \in \{1, \ldots, K\}$, denote the number of components of type k for which test failure data are available, and let s_t^k denote the number of these components which still function at time t.

The NPI lower survival function is derived as follows. Remember that C_t^k denotes the number of components of type k in the system which function at time t, where it is assumed that failure ends the functioning of a component. Under the assumptions for the NPI approach (Coolen 1998), we derive the following lower bound for the survival function

$$P(T_S > t) \geq \sum_{l_1=0}^{m_1} \cdots \sum_{l_K=0}^{m_K} \Phi(\underline{l}) \prod_{k=1}^{K} \overline{D}(C_t^k = l_k)$$

where

$$\overline{D}(C_t^k = l_k) = \overline{P}(C_t^k \leq l_k) - \overline{P}(C_t^k \leq l_k - 1) =$$

$$\binom{n_k + m_k}{n_k}^{-1} \binom{s_t^k - 1 + l_k}{s_t^k - 1} \times$$

$$\binom{n_k - s_t^k + m_k - l_k}{n_k - s_t^k}$$

In this expression, $\overline{P}$ denotes the NPI upper probability for Bernoulli data (Coolen 1998). For each

component type k, the function $\overline{D}$ ensures that maximum possible probability, corresponding to NPI for Bernoulli data (Coolen 1998), is assigned to the event $C_t^k = 0$, so $\overline{D}(C_t^k = 0) = \overline{P}(C_t^k = 0)$. Then, $\overline{D}(C_t^k = 1)$ is defined by putting the maximum possible remaining probability mass, from the total probability mass available for the event $C_t^k \leq 1$, to the event $C_t^k = 1$. This is achieved by $\overline{D}(C_t^k = 1) = \overline{P}(C_t^k \leq 1) - \overline{P}(C_t^k = 0)$. This argument is continued, by assigning for increasing l_k the maximum possible remaining probability mass $\overline{D}(C_t^k = l_k)$. As the survival signature is increasing in l_k for coherent systems, as assumed in this paper, and the resulting $\overline{D}$ is a precise probability distribution, the right-hand side of the inequality above is indeed a lower bound and it is the maximum possible lower bound. As such, it is the NPI lower probability for the event $T_S > t$, giving the NPI lower survival function for the system failure time (for $t > 0$)

$$\underline{P}(T_S > t) = \sum_{l_1=0}^{m_1} \cdots \sum_{l_K=0}^{m_K} \Phi(\underline{l}) \prod_{k=1}^{K} \overline{D}(C_t^k = l_k)$$

The corresponding NPI upper survival function for T_S is similarly derived, using the upper bound

$$P(T_S > t) \leq \sum_{l_1=0}^{m_1} \cdots \sum_{l_K=0}^{m_K} \Phi(\underline{l}) \prod_{k=1}^{K} \underline{D}(C_t^k = l_k)$$

where

$$\underline{D}(C_t^k = l_k) = \underline{P}(C_t^k \leq l_k) - \underline{P}(C_t^k \leq l_k - 1) =$$

$$\binom{n_k + m_k}{n_k}^{-1} \binom{s_t^k + l_k}{s_t^k} \times$$

$$\binom{n_k - s_t^k + m_k - l_k - 1}{n_k - s_t^k}$$

In this expression, $\underline{P}$ denotes the NPI lower probability for Bernoulli data (Coolen 1998). This construction ensures that minimum possible weight is given to small values of C_t^k, resulting in the NPI upper survival function for the system failure time

$$\overline{P}(T_S > t) = \sum_{l_1=0}^{m_1} \cdots \sum_{l_K=0}^{m_K} \Phi(\underline{l}) \prod_{k=1}^{K} \underline{D}(C_t^k = l_k)$$

We illustrate this NPI method for system reliability using the survival signature in Example 2 (Coolen, Coolen-Maturi, & Al-nefaiee 2014).

Example 2.
Consider the system with $K = 2$ types of components as presented in Figure 2. The survival signature for this system is presented in Table 3, it is easily verified by checking all possible combinations of the specific components of each type which function or not.

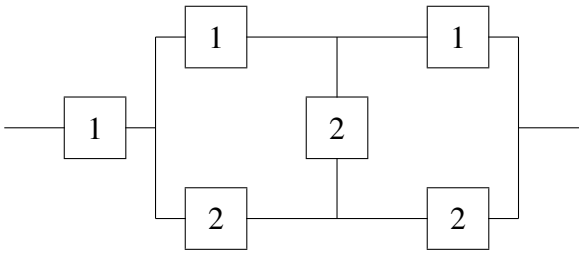

Figure 2: System with 2 types of components

l_1	l_2	$\Phi(l_1, l_2)$	l_1	l_2	$\Phi(l_1, l_2)$
0	0	0	2	0	0
0	1	0	2	1	0
0	2	0	2	2	4/9
0	3	0	2	3	6/9
1	0	0	3	0	1
1	1	0	3	1	1
1	2	1/9	3	2	1
1	3	3/9	3	3	1

Table 3: Survival signature of the system in Figure 2

To illustrate NPI for the system survival time, suppose that $n_1 = 2$ components exchangeable with those of type 1 and $n_2 = 2$ components exchangeable with those of type 2 were tested. First suppose that failure times $t_1^2 < t_1^1 < t_2^2 < t_2^1$ were observed, with t_j^k the j-th ordered failure time of a component of type k. The resulting NPI lower and upper survival functions for the system failure time T_S are specified in Table 4, together with the results for the case with the test failure times ordered as $t_1^1 < t_1^2 < t_2^1 < t_2^2$.

For the ordering $t_1^2 < t_1^1 < t_2^2 < t_2^1$, in the first interval in Table 4 we have not yet seen a failure in the test data, so the NPI upper probability that the system will function is equal to one, which is logical as we base the inferences on the data with few additional assumptions. In the second interval, one failure of type 2 has occurred but we do not have any evidence from the data against the possibility that a component of type 1 will certainly function at times in this interval, so the NPI upper survival function remains one. In the fourth interval, both type 2 components have failed but only one component of type 1 has failed. In this interval, to consider the lower survival function the system is effectively reduced to a series system consisting of three components of type 1, with one 'success' and one 'failure' as data, denoted by $(2, 1)$. As such a series system only functions if all three components function, the NPI lower survival function within this fourth interval is equal to $\underline{S}_{T_S}(t) = \frac{1}{3} \times \frac{2}{4} \times \frac{3}{5}$

= 0.100, which follows by sequential reasoning, using that, based on n observations consisting of s successes and $n - s$ failures, denoted as data (n, s), the NPI lower probability for the next observation to be a success is equal to $s/(n + 1)$ (Coolen 1998). The NPI lower probability for the first component to function, given test data $(2, 1)$, is equal to $1/3$. Then the second component is considered, conditional on the first component functioning, which combines with the test data to two out of three components observed (or assumed) to be functioning, so combined data $(3, 2)$, hence this second component will also function with NPI lower probability $2/4$. Similarly, the NPI lower probability for the third component to function, conditional on functioning of the first two components in the system, so with combined data $(4, 3)$, is equal to $3/5$. In the final interval, we are beyond the failure times of all the tested components, so we no longer have evidence in favour of the system to function, so $\underline{S}_{T_S}(t) = 0$, but the system might of course still function as reflected by $\overline{S}_{T_S}(t) = 0.148$.

For the second case in Table 4, with data ordering $t_1^1 < t_1^2 < t_2^1 < t_2^2$, we have $\overline{S}_{T_S}(t) = 0.667$ in the second interval, where one failure of type 1 has occurred in the test data. In the fourth interval, both tested components of type 1 have failed, leading to $\underline{S}_{T_S}(t) = 0$. Both of these values are directly related to the required functioning of the left-most component in Figure 2.

$$t_1^2 < t_1^1 < t_2^2 < t_2^1$$

$t \in$	$\underline{P}(T_S > t)$	$\overline{P}(T_S > t)$
$(0, t_1^2)$	0.553	1
(t_1^2, t_1^1)	0.458	1
(t_1^1, t_2^2)	0.148	0.553
(t_2^2, t_2^1)	0.100	0.458
(t_2^1, ∞)	0	0.148

$$t_1^1 < t_1^2 < t_2^1 < t_2^2$$

$t \in$	$\underline{P}(T_S > t)$	$\overline{P}(T_S > t)$
$(0, t_1^1)$	0.553	1
(t_1^1, t_1^2)	0.230	0.667
(t_1^2, t_2^1)	0.148	0.553
(t_2^1, t_2^2)	0	0.230
(t_2^2, ∞)	0	0.148

Table 4: Lower and upper survival functions for the system in Figure 2 and two data orderings

4 DISCUSSION

The survival signature is a powerful and quite basic concept. As such, further generalizations are conceptually easy, for example one can straightforwardly generalize the survival signature to multi-state systems such that it again summarizes the structure function in a manner that is sufficient for a range of uncertainty quantifications for the system reliability. The survival signature can also be used with a generalization of the system structure function where the latter is a probability instead of a binary function, or even an imprecise probability. This enables uncertainty of system functioning for given states of its components to be taken into account, which may be convenient, for example, to take uncertain demands or environments for the system into consideration. In this paper, we only considered test data with observed failure times for all tested components. If test data also contain right-censored observations, this can also be dealt with, both in the imprecise Bayesian and NPI approaches (Walter, Graham, & Coolen 2015, Coolen & Yan 2004, Maturi 2010) (more information about NPI is available from `www.npi-statistics.com`). This generalization is further relevant as, instead of assuming availability of test data, it allows us to take process data for the actual components in a system into account while this system is operating, hence enabling inference on the remaining time until system failure.

Upscaling the survival signature to large real-world systems and networks, consisting of thousands of components, is a major challenge. However, even for such systems the fact that one only needs to derive the survival signature once for a system is an advantage, and also the monotonicity of the survival signature for coherent systems is very useful if one can only derive it partially. For small to medium-sized systems and networks, the survival signature is particularly easy to compute using the `ReliabilityTheory` R package (Aslett 2016b), available from `www.louisaslett.com`. Using this package it is straightforward to express your system in terms of an undirected graphical structure, after which a single call to the function `computeSystemSurvivalSignature` suffices. The function will compute all of the cut sets of the system and perform the combinatorial analysis, returning a table which contains the survival signature just as in Table 2 and 3. For example, computation of the survival signature for the system in Figure 1 is achieved with 3 simple commands

```
s <- graph.formula(s-1-2-3-t,
                   s-1-4-5-t,
                   2:4-6-3:5)
setCompTypes(s,
             list("T1"=c(2,4,3,5),
                  "T2"=6,
                  "T3"=1))
computeSystemSurvivalSignature(s)
```

Full instructions and some worked examples are available within the package. There are numerous other functions in the package, enabling computation of: the legacy system signature (Samaniego 2007); the

continuous-time Markov chain representation of repairable systems; as well as numerous inference algorithms for Bayesian inference on the system signature using only system-level data (Aslett 2013).

The survival signature enables some interesting applications which would otherwise be intractably difficult. For example, often a system designer may consider the design (structure) of their system to be a trade secret and so be unwilling to release it to component manufacturers, while at the same time component manufacturers are frequently unwilling to release anything more than summary figures for components, e.g. mean-time-between-failures. These two opposing goals lead to a situation in which it would seem unrealistic to achieve a full probabilistic reliability assessment and to honour the privacy requirements of all parties. However, recent work (Aslett 2016a) makes use of the survival signature to allow cryptographically secure evaluation of the system reliability function, where the functional form resulting from the survival signature decomposition in Equation (1) is crucial to enabling encrypted computation using so-called homomorphic encryption schemes (Aslett, Esperança, & Holmes 2015). The equivalent decomposition in terms of the structure function leads to difficulties in encrypted computation, so that this application may be intractable without use of the survival signature.

ACKNOWLEDGEMENTS

The authors wish to thank Professor Radim Bris for his kind invitation to present this work at the ICAMER 2016 conference. Louis Aslett was supported by the i-like project (EPSRC grant reference number EP/K014463/1). Gero Walter was supported by the DINALOG project 'Coordinated Advanced Maintenance and Logistics Planning for the Process Industries' (CAMPI).

REFERENCES

Aslett, L. (2013). *MCMC for Inference on Phase-type and Masked System Lifetime Models*. Ph. D. thesis, Trinity College Dublin.

Aslett, L. (2016a). Cryptographically secure multiparty evaluation of system reliability. *Pending journal submission*.

Aslett, L. (2016b). *ReliabilityTheory: Tools for structural reliability analysis.* R package.

Aslett, L., F. Coolen, & S. Wilson (2015). Bayesian inference for reliability of systems and networks using the survival signature. *Risk Analysis 35*, 1640–1651.

Aslett, L., P. Esperança, & C. Holmes (2015). A review of homomorphic encryption and software tools for encrypted statistical machine learning. Technical report, University of Oxford.

Augustin, T. & F. Coolen (2004). Nonparametric predictive inference and interval probability. *Journal of Statistical Planning and Inference 124*, 251–272.

Augustin, T., F. Coolen, G. de Cooman, & M. Troffaes (2014). *Introduction to Imprecise Probabilities*. Chichester: Wiley.

Barlow, R. & F. Proschan (1975). *Statistical Theory of Reliability and Life Testing*. New York: Holt, Rinehart and Winston.

Coolen, F. (1998). Low structure imprecise predictive inference for Bayes' problem. *Statistics & Probability Letters 36*, 349–357.

Coolen, F. (2006). On nonparametric predictive inference and objective Bayesianism. *Journal of Logic, Language and Information 15*, 21–47.

Coolen, F. (2011). Nonparametric predictive inference. In M. Lovric (Ed.), *International Encyclopedia of Statistical Science*, pp. 968–970. Springer.

Coolen, F. & T. Coolen-Maturi (2012). On generalizing the signature to systems with multiple types of components. In W. Zamojski, J. Mazurkiewicz, J. Sugier, T. Walkowiak, and J. Kacprzyk (Eds.), *Complex Systems and Dependability*, pp. 115–130. Springer.

Coolen, F. & T. Coolen-Maturi (2015). Predictive inference for system reliability after common-cause component failures. *Reliability Engineering and System Safety 135*, 27–33.

Coolen, F., T. Coolen-Maturi, & A. Al-nefaiee (2014). Nonparametric predictive inference for system reliability using the survival signature. *Journal of Risk and Reliability 228*, 437–448.

Coolen, F. & L. Utkin (2011). Imprecise reliability. In M. Lovric (Ed.), *International Encyclopedia of Statistical Science*, pp. 649–650. Springer.

Coolen, F. & K. Yan (2004). Nonparametric predictive inference with right-censored data. *Journal of Statistical Planning and Inference 126*, 25–54.

De Finetti, B. (1974). *Theory of Probability*. Chichester: Wiley.

Maturi, T. (2010). *Nonparametric Predictive Inference for Multiple Comparisons*. Ph. D. thesis, Durham University.

Samaniego, F. (2007). *System Signatures and their Applications in Engineering Reliability*. New York: Springer.

Walter, G. (2013). *Generalized Bayesian Inference under Prior-Data Conflict*. Ph. D. thesis, Ludwig Maximilian University of Munich.

Walter, G., L. Aslett, & F. Coolen (2016). Bayesian nonparametric system reliability with sets of priors. *In submission*.

Walter, G., A. Graham, & F. Coolen (2015). Robust Bayesian estimation of system reliability for scarce and surprising data. In L. Podofillini, B. Sudret, B. Stojadinović, E. Zio, and W. Kröger (Eds.), *Safety and Reliability of Complex Engineered Systems: ESREL 2015*, pp. 1991–1998. CRC Press.

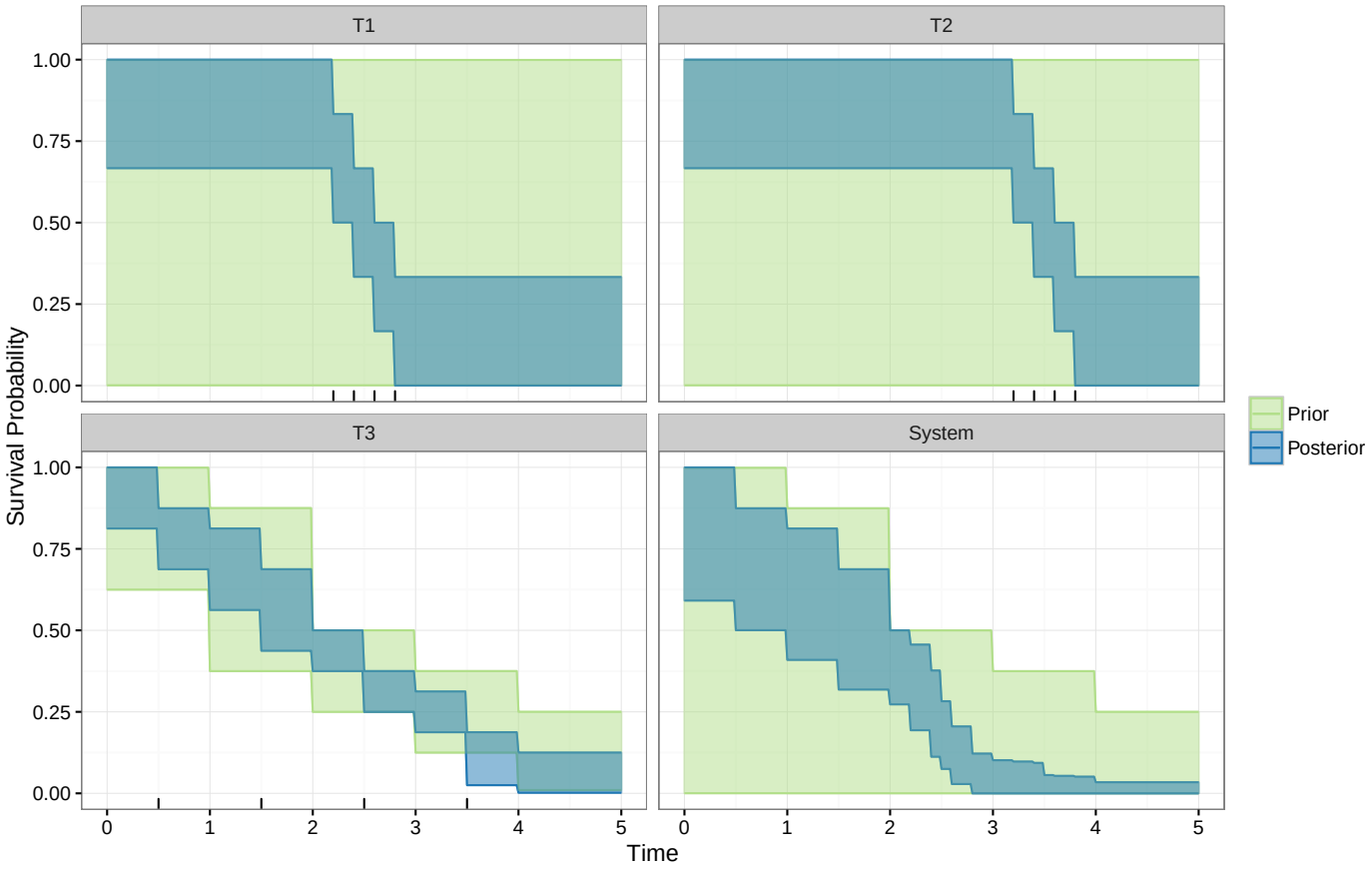

Figure 3: Prior and posterior sets of survival functions for the system in Figure 1 and its three component types. The component failure times, that form the test data, are denoted with tick marks near the time axis.