The structure function for system reliability as predictive (imprecise) probability

Frank P.A. Coolen[1a], Tahani Coolen-Maturi[2]

[1]*Department of Mathematical Sciences, Durham University, UK*
[2]*Durham University Business School, Durham University, UK*

Abstract

In system reliability, the structure function models functioning of a system for given states of its components. As such, it is typically a straightforward binary function which plays an essential role in reliability assessment, yet it has received remarkably little attention in its own right. We explore the structure function in more depth, considering in particular whether its generalization as a, possibly imprecise, probability can provide useful further tools for reliability assessment in case of uncertainty. In particular, we consider the structure function as a predictive (imprecise) probability, which enables uncertainty and indeterminacy about the next task the system has to perform to be taken into account. The recently introduced concept of 'survival signature' provides a useful summary of the structure function to simplify reliability assessment for systems with many components of multiple types. We also consider how the (imprecise) probabilistic structure function can be linked to the survival signature. We briefly discuss some related research topics towards implementation for large practical systems and networks, and we outline further possible generalizations.

Keywords: Lower and upper probabilities, structure function, survival signature, system reliability

1. INTRODUCTION

In the mathematical theory of reliability, the main focus is on the functioning of a system, given the structure of the system and the functioning,

[a]Corresponding author: frank.coolen@durham.ac.uk

or not, of its components. The mathematical concept which is central to this theory is the *structure function*. For a system with m components, let state vector $\underline{x} = (x_1, x_2, \ldots, x_m) \in \{0,1\}^m$, 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 $\underline{x}$. In this paper, as in most of the literature, attention is restricted to coherent systems, for which $\phi(\underline{x})$ 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 (generalization to allow incoherent systems is possible but would make concepts and notation later in the paper more complex). Coherence of a system is further usually assumed to imply that all the system's components are relevant, meaning that the functioning or not of each component makes a difference to the functioning of the system for at least one set of states for the other components.

The structure function is a powerful tool for reliability quantification, but in practice there may be uncertain or unknown aspects related to a system's functioning which can be taken into account by a generalization of the structure function to a probabilistic structure function. A main motivation for this generalization is that the system may have to deal with a variety of tasks of different types, which put different requirements on the system. We focus then on a specific future task to be performed, calling it the 'next task', and take uncertainty about the type of this task into account by using probabilities over the different types of tasks, and by generalizing this to imprecise probabilities which enables uncertainty and indeterminacy to be included in the modelling. This approach is very flexible; it can even be used to include the possibility of a fully unknown type of task, which might for example be suitable to reflect possible unknown threats to the system. A further motivation comes from the fact that there may simply be too many uncertainties affecting the system's functioning, which cannot be modelled in detail due to lack of meaningful information or limited time for detailed analysis. Throughout, it should be kept in mind that the proposed (imprecise) probabilistic structure function generalized the classical structure function, and as such provides a more flexible tool for reliability quantification. If one strongly feels that one can always model scenarios in full detail then one can argue that this generalization is not required, in which case perhaps the interest in the probabilistic structure function would be merely from the perspective of a mathematical exercise. However, we believe that there are plenty of real-world scenarios that will benefit from the flexibility provided by the (imprecise) probabilistic structure function when compared to the special case of the deterministic structure function, ensuring that the contribution of this paper goes far beyond merely a mathematical exercise.

Section 2 presents the structure function as a, possibly imprecise, predictive probability. This section also presents a motivating example. Section 3 explains how the (imprecise) probabilistic structure function can be incorporated into (lower and upper) survival signatures for efficient system reliability quantification [1]. Uncertainty with regard to the type of the next task is considered in Section 4, and illustrated in an extensive example in Section 5. The paper concludes with a discussion of some related aspects in Section 6.

2. THE STRUCTURE FUNCTION AS (IMPRECISE) PROBABILITY

A simple way to reflect uncertainty about the system's functioning given the state vector $\underline{x}$ is by defining the structure function as a probability, so $\phi : \{0,1\}^m \to [0,1]$. We define $\phi(\underline{x})$ as the probability that the system functions for a specific state vector $\underline{x}$ and for the next task the system is required to perform. It should be emphasized that explicit focus on the next task is not necessary when generalizing the structure function in this way, but it provides a natural tool for further uncertain aspects which we discuss later. We will simply refer to this generalized structure function as *probabilistic structure function*, and for simplicity we keep using the same notation $\phi(\underline{x}$ which is reasonable as the classical deterministic structure function is a special case of the probabilistic structure functioning, only using probabilities 0 and 1. Note that one could similarly define a probabilistic structure function for a system that has to perform multiple future tasks, this is left as a topic for future consideration.

We wish to emphasize that considering the structure function as a probability differs essentially from the classical use of the structure function with randomness on whether or not the individual components function. The corresponding probability that the system functions with random functioning of the components is usually called the 'reliability function' [2]. The important novelty in this paper is that the system functioning can be uncertain for given states of the components, which can occur due to a range of practical circumstances. Combining the structure function as a probability with randomness for the state of the components is a straigthforward further step, simplified by the use of the survival signature, as also presented in this paper.

Let S denote the event that the system functions as required for the next task it has to perform, then

$$\phi(\underline{x}) = P(S|\underline{x}) \tag{1}$$

We have kept the same notation for the structure function, as a probability, as in Section 1, which should not cause problems and is justified as the earlier definition of structure function can be regarded as a special case of this generalized definition with all probabilities either 0 or 1.

This generalization already enables an important range of real-world scenarios to be modelled in a straightforward way. Scenarios where the flexibility of the structure function as a probability might be useful are, of course, situations where even with known status of the components, it is not certain whether or not the system functions, that is performs its task as required. This may be due to varying circumstances or requirements which may not be modelled explicitly, or may not even be fully known. For example, one could consider a collection of wind turbines as one system, with the task to contribute to overall generation of a level of energy required to provide a specific area with sufficient electricity for a specific period of time (we can consider this to be the 'next task'). One could consider each wind turbine as a component (with several other types of components in the system, that is irrelevant for now). Even if one knows the number of functioning components at a particular time, factors such as the weather, the availability of other electricity generating resources for the network, and the specific electricity demand, can lead to uncertainty about whether or not the system meets the actual requirements. To fit with the established deterministic definition of the structure function one can define system functioning in far more detail, but this may be hard to do in practice. As another example, one could think about a network of computers which together form a system for complex computations, where its actual success in dealing with required tasks might be achieved with some computers not functioning, but with some lack of knowledge about the exact number of computers required to complete tasks of different types.

The generalization to consider the structure function as a probability, although mathematically straightforward, requires substantial information in order to assess the probabilities of system functioning for all possible state vectors $\underline{x}$. While this modelling might explicitly take co-variates into account, thus possibly benefitting from a large variety of statistical models, it may be difficult to actually formulate the important co-variates and one might not know their specific values. This leads to two further topics we wish to discuss, namely what precisely is meant when we say that the system functions, and a generalization of probability to allow lack of knowledge to be reflected.

Whether or not a real-world system performs its task well may depend on many circumstances beyond the states of the system components. It may be too daunting to specify system functioning for all possible circumstances, and it may even be impossible to know all possible circumstances. Hence, speaking of 'system functioning' in the traditional theoretic way seems rather restricted. One suggestion would be to only define system functioning for one (or a specified number of) application(s), e.g. whether or not a system functions at its next required use. This will not be sufficient for all real-world scenarios, but it will enable important aspects of uncertainty on factors such as different tasks and circumstances to be taken into account. We believe that this is a topic that requires further attention, it links to many system dependability concepts including flexibility and resilience.

The generalization of the structure function as a probability provides substantial enhanced modelling opportunities for system reliability and dependability. However, the use of single-valued probabilities for events does not enable the strength or lack of information to be taken into account, with most obvious limitation the inability to reflect if 'no information at all' is available about an event of interest. In recent decades, theory of *imprecise probability* [3] has gained increasing attention from the research community, including contributions to reliability and risk [4]. It generalizes classical, precise, probability theory by assigning to each event two values, a lower probability and an upper probability, denoted by $\underline{P}$ and $\overline{P}$, respectively, with $0 \leq \underline{P} \leq \overline{P} \leq 1$. These can be interpreted in several ways, for the current discussion it suffices to regard them as the sharpest bounds for a probability based on the information available, where the lower probability typically reflects the information available in support of the event of interest and the corresponding upper probability reflects the information available against this event. The case of no information at all can be reflected by $[\underline{P}, \overline{P}] = [0, 1]$ while equality $\underline{P} = \overline{P}$ reflects perfect knowledge about the probability and results in classical precise probability as a special case of imprecise probability.

We propose the further generalization of the structure function within imprecise probability theory by introducing the *lower structure function*

$$\underline{\phi}(\underline{x}) = \underline{P}(S|\underline{x}) \tag{2}$$

and the *upper structure function*

$$\overline{\phi}(\underline{x}) = \overline{P}(S|\underline{x}) \tag{3}$$

This provides substantial flexibility for practical application of methods to quantify system reliability and other dependability concepts. For example, it may be known historically that, under different external circumstances, a system with a certain subset $\underline{x}$ of its components functioning manages a task well in 85 to 95 percent of all cases. While it might be possible to go into further detail and e.g. describe beliefs within this range by a probability distribution, or assume this for mathematical convenience, this may not be required or it may actually be impossible in a meaningful way, and one can use lower probability 0.85 and upper probability 0.95 to accurately reflect this information. If one has to rely on expert judgements to assign the values of the structure function, then time may often be too limited to meaningfully assign precise probabilities for system functioning for all possible component state vectors. In such cases, the use of imprecise probabilities also offers suitable flexibility. Assigning a subset of probabilities for some events (or bounds for these) will imply bounds for all other related events under suitable assumptions [3] where particularly assumed coherence of the system, which implies that any additional component failure can never improve system functioning, is useful and practically justifiable in many applications.

3. INCORPORATING THE PROBABILISTIC STRUCTURE FUNCTION INTO THE SURVIVAL SIGNATURE

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. Recently, we introduced such a summary, called the *survival signature*, to facilitate reliability analyses for systems with multiple types of components [1]. In case of just a single type of components, the survival signature is closely related to the system signature [5], 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. We briefly review the survival signature and some recent advances, then link it to the above suggested generalization of the structure function.

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 [6], while full independence is assumed for the random failure times of components of different types. 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 sub-vector 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 $k \in \{1, \ldots, K\}$ [1].

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 $\sum_{i=1}^{m_k} x_i^k = l_k$, $k = 1, \ldots, K$. 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 [1]

$$\Phi(l_1, \ldots, l_K) = \left[\prod_{k=1}^{K} \binom{m_k}{l_k}^{-1}\right] \times \sum_{\underline{x} \in S_{l_1,\ldots,l_K}} \phi(\underline{x}) \tag{4}$$

We now consider the survival signature with the generalized structure function as discussed above, using the lower structure function (2) and upper structure function (3). The survival signature can straightforwardly be adapted to include these, due to its monotone dependence on the structure function. This leads to the following definitions of the *lower survival signature*

$$\underline{\Phi}(l_1, \ldots, l_K) = \left[\prod_{k=1}^{K} \binom{m_k}{l_k}^{-1}\right] \times \sum_{\underline{x} \in S_{l_1,\ldots,l_K}} \underline{\phi}(\underline{x}) \tag{5}$$

and the corresponding *upper survival signature*

$$\overline{\Phi}(l_1, \ldots, l_K) = \left[\prod_{k=1}^{K} \binom{m_k}{l_k}^{-1}\right] \times \sum_{\underline{x} \in S_{l_1,\ldots,l_K}} \overline{\phi}(\underline{x}) \tag{6}$$

These are the sharpest possible bounds for the survival signature corresponding to the lower and upper structure functions, and as such indeed the lower and upper probabilities for the event that the system functions given that precisely l_k of its m_k components of type k function, for each $k \in \{1, \ldots, K\}$.

These lower and upper survival signatures can be used for imprecise reliability quantifications. Particularly if chosen quantifications are monotone functions of the survival signature, this is again a straightforward generalization of the precise approach [1]. Let us consider the event that the system functions for the next task it has to perform, denoted by S. Let $C_k \in \{0, 1, \ldots, m_k\}$ denote the number of components of type k in the system which function when required for the next task. The probability for the event S is [1]

$$P(S) = \sum_{l_1=0}^{m_1} \cdots \sum_{l_K=0}^{m_K} \Phi(l_1, \ldots, l_K) P(\bigcap_{k=1}^{K} \{C_k = l_k\}) \tag{7}$$

With the generalization of the survival signature, we get the lower probability for the event that the systems functions for the next task

$$\underline{P}(S) = \sum_{l_1=0}^{m_1} \cdots \sum_{l_K=0}^{m_K} \underline{\Phi}(l_1, \ldots, l_K) P(\bigcap_{k=1}^{K} \{C_k = l_k\}) \tag{8}$$

and the corresponding upper probability

$$\overline{P}(S) = \sum_{l_1=0}^{m_1} \cdots \sum_{l_K=0}^{m_K} \overline{\Phi}(l_1, \ldots, l_K) P(\bigcap_{k=1}^{K} \{C_k = l_k\}) \tag{9}$$

For this imprecise case, just as for the precise case [1], assuming independence of the functioning of components of different types leads to simplification of (8) and (9) by using, for $l_k \in \{0, 1, \ldots, m_k\}$ for each $k \in \{1, \ldots, K\}$,

$$P(\bigcap_{k=1}^{K} \{C_k = l_k\}) = \prod_{k=1}^{K} P(C_k = l_k)$$

If in addition it is assumed that functioning of components of the same type is conditionally independent given probability $p_k \in [0,1]$ that a component of type k functions for the next task, then further simplification is achieved by using

$$P(\bigcap_{k=1}^{K}\{C_k = l_k\}) = \prod_{k=1}^{K} \binom{m_k}{l_k} p_k^{l_k}[1-p_k]^{m_k-l_k}$$

This leads to relatively straightforward computations for reliability metrics, which we do not discuss further in this paper. It is important though to emphasize that exactly the same approach can be followed when interest is in processes over time, where instead of focussing on functioning of the system for the next task one can consider the probability that the system functions at a given time [1].

The probabilities for the numbers of functioning components can also be generalized to lower and upper probabilities, as e.g. done by Coolen et al. [7] within the nonparametric predictive inference framework of statistics [8], where lower and upper probabilities for the events $C_k = l_k$ are inferred from test data on components of the same types as those in the system. This step is less trivial as one must ensure to have probability distributions for these events, thus summing to one over $l_k = 0, 1, \ldots, m_k$ for each type k. For coherent systems this is not very complicated due to the monotonicity of the (lower or upper) survival signature, see [7] where the method for dealing with imprecision on the random quantities C_k is presented.

The main advantage of the survival signature, in line with this property of the signature for systems with a single type of components [5], as shown by Equation (7), 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. This property clearly also holds for the lower and upper survival signatures as is shown by Equations (8) and (9).

4. MULTIPLE TYPES OF TASKS

If a system may need to deal with different tasks, the (lower or upper) structure function should, ideally, be defined for each specific type of task. Let there be $R \geq 1$ types of tasks. The (lower or upper) structure function for a specific type of task $r \in \{1, \ldots, R\}$ is the (lower or upper) probability for the event that the system functions for component states $\underline{x}$ and for known type of task r, we denote these as before with an additional subscript r (we generalize earlier notation in this way throughout this section without explicit introduction), so

$$\phi_r(\underline{x}) = P(S|\underline{x}, r) \quad \underline{\phi}_r(\underline{x}) = \underline{P}(S|\underline{x}, r) \quad \overline{\phi}_r(\underline{x}) = \overline{P}(S|\underline{x}, r)$$

If interest is in the next task that the system has to perform, and it is known of which type this task is, then we are back to the setting discussed before. If the type of task is not known with certainty, then there are several possible scenarios. First, suppose that one can assign a precise probability for the event that the next task is of type r, denoted by p_r, for each $r \in \{1, \ldots, R\}$. Then the system structure function for the next task can be derived via the theorem of total probability, which also applies straightforwardly to the corresponding lower and upper structure functions in the generalized case. This leads to

$$\phi(\underline{x}) = \sum_{r=1}^{R} \phi_r(\underline{x}) p_r \quad \underline{\phi}(\underline{x}) = \sum_{r=1}^{R} \underline{\phi}_r(\underline{x}) p_r \quad \overline{\phi}(\underline{x}) = \sum_{r=1}^{R} \overline{\phi}_r(\underline{x}) p_r$$

For this scenario the corresponding lower and upper survival signatures that apply for the next task, of random type, are easily derived and given by

$$\begin{aligned}
\underline{\Phi}(l_1, \ldots, l_K) &= \left[\prod_{k=1}^{K} \binom{m_k}{l_k}^{-1}\right] \times \sum_{\underline{x} \in S_{l_1, \ldots, l_K}} \sum_{r=1}^{R} \underline{\phi}_r(\underline{x}) p_r \\
&= \sum_{r=1}^{R} \underline{\Phi}_r(l_1, \ldots, l_K) p_r \\
\overline{\Phi}(l_1, \ldots, l_K) &= \left[\prod_{k=1}^{K} \binom{m_k}{l_k}^{-1}\right] \times \sum_{\underline{x} \in S_{l_1, \ldots, l_K}} \sum_{r=1}^{R} \overline{\phi}_r(\underline{x}) p_r \\
&= \sum_{r=1}^{R} \overline{\Phi}_r(l_1, \ldots, l_K) p_r
\end{aligned}$$

These results hold as all sums involved are finite, hence the order of summations can be changed, which can also be applied to derive

$$\underline{P}(S) = \sum_{r=1}^{R} \underline{P}_r(S) p_r$$
$$\overline{P}(S) = \sum_{r=1}^{R} \overline{P}_r(S) p_r$$

Secondly, one may only be able to assign bounds for the probabilities p_r, where the sharpest bounds one can assign are lower and upper probabilities, denoted by $\underline{p}_r$ and $\overline{p}_r$. Let p denote any probability vector of dimension R, so $p = (p_1, \ldots, p_R)$ with all $p_r \geq 0$ and $\sum_{r=1}^{R} p_r = 1$, and let $\mathcal{P}$ denote the set of all such probability vectors with $\underline{p}_r \leq p_r \leq \overline{p}_r$ for all $r \in \{1, \ldots, R\}$. In this situation, deriving the lower and upper structure functions for the next task is less straigthforward, as they require optimisation over the set $\mathcal{P}$ of probability vectors

$$\underline{\phi}(\underline{x}) = \min_{p \in \mathcal{P}} \sum_{r=1}^{R} \underline{\phi}_r(\underline{x}) p_r \quad \overline{\phi}(\underline{x}) = \max_{p \in \mathcal{P}} \sum_{r=1}^{R} \overline{\phi}_r(\underline{x}) p_r \tag{10}$$

In case of a precise structure function, the lower and upper structure functions on the right-hand sides of these equations are just equal to the precise structure function, with imprecision still resulting from the set $\mathcal{P}$ of probability vectors. While these optima are not available in closed-form, their computation is quite straightforward, solutions are obtained by setting all p_r equal to either $\underline{p}_r$ or $\overline{p}_r$ apart from one which will take on a value within its corresponding range $[\underline{p}_r, \overline{p}_r]$ such that the individual probabilities sum up to one.

For this scenario, deriving the corresponding lower and upper survival signatures is less straightforward than for the first scenario above. These lower and upper survival signatures are

$$\underline{\Phi}(l_1, \ldots, l_K) = \min_{p \in \mathcal{P}} \left(\left[\prod_{k=1}^{K} \binom{m_k}{l_k}^{-1} \right] \times \sum_{\underline{x} \in S_{l_1, \ldots, l_K}} \sum_{r=1}^{R} \underline{\phi}_r(\underline{x}) p_r \right) \tag{11}$$

$$\overline{\Phi}(l_1, \ldots, l_K) = \max_{p \in \mathcal{P}} \left(\left[\prod_{k=1}^{K} \binom{m_k}{l_k}^{-1} \right] \times \sum_{\underline{x} \in S_{l_1, \ldots, l_K}} \sum_{r=1}^{R} \overline{\phi}_r(\underline{x}) p_r \right) \tag{12}$$

which generally requires solving complex optimisation problems. From computational perspective it is far easier to calculate

$$\left[\prod_{k=1}^{K}\binom{m_k}{l_k}^{-1}\right] \times \sum_{\underline{x}\in S_{l_1,\ldots,l_K}} \left(\min_{p\in\mathcal{P}} \sum_{r=1}^{R} \underline{\phi}_r(\underline{x}) p_r\right) \tag{13}$$

and

$$\left[\prod_{k=1}^{K}\binom{m_k}{l_k}^{-1}\right] \times \sum_{\underline{x}\in S_{l_1,\ldots,l_K}} \left(\max_{p\in\mathcal{P}} \sum_{r=1}^{R} \overline{\phi}_r(\underline{x}) p_r\right) \tag{14}$$

These expressions follow by inserting the lower and upper structure functions (10) into the equations for the lower and upper survival signatures, and require many optimisations to be performed, but as just mentioned these are all quite straightforward. Generally, the lower survival signature (11) is greater than or equal to expression (13) and the upper survival signature (12) is less than or equal to expression (14). If all optimisations in expression (13) have the same probability vector within $\mathcal{P}$ as solution, then the lower survival signature (11) is equal to expression (13), and similarly for the upper survival signature (12) with regard to the optimisations in expression (14). While this may appear to be unlikely, we will illustrate a case were it applies in the example in the following section. Further investigation into the optimisation problems for general situations is left as an important challenge for future research.

Finally, one may wish to use statistical inference for the p_r in case one has relevant data. There is a variety of options, including Bayesian methods, which might be generalized through the use of sets of prior distributions as in the imprecise Dirichlet model for multinomial data [3] and nonparametric predictive inference [9, 10]. The latter approach may be of specific interest as it provides the possibility to take unobserved or even undefined tasks into consideration [11].

5. Example

We present an extensive example to illustrate the new concepts and methods presented in this paper. The uncertainty about system functioning given the states of its components is assumed to result from two cases, requiring different numbers of components to function for successful functioning of the system. Five different types of task are then considered, each varying with

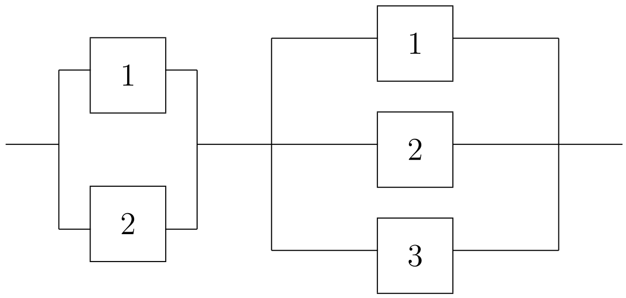

Figure 1: System with three types of components

regard to the (imprecise) probability that they belong to either of the two cases. We then explore predictive (imprecise) probabilities of system functioning for 4 situations with different information about the next task that the system has to perform. Crucially, the fourth situation includes the possibility that the next task is totally unknown, which is a specific advantage of the presented method with the structure function as predictive imprecise probability.

Consider the system presented in Figure 1, consisting of two subsystems in series configuration, with three types of component as indicated by the numbers in the figure. Consider the following variation for the second subsystem consisting of three components: for some tasks to be performed according to the requirements it is sufficient for one of the three components to function, but for other tasks (or under other circumstances) it is necessary to have at least two components functioning. We will refer to these as Case 1 and Case 2, respectively. The survival signatures for this system corresponding to these two cases are presented in Table 1, denoted by Φ_1 and Φ_2, where the quite trivial entries for which both survival signatures are equal to 0 or 1 are not included.

Suppose that five different possible tasks have been identified which this system may have to deal with. This may actually be different tasks, or just due to different circumstances under which the tasks may need to be performed. For Task A Case 1 applies, so only one functioning component in the second subsystem is required. For Task B Case 2 applies. For Task C there is uncertainty about whether one or two components need to function in the second subsystem, with either case having probability 1/2. For Task D the same uncertainty occurs, but the probabilities that either case applies are

l_1	l_2	l_3	$\Phi_1(l_1, l_2, l_3)$	$\Phi_2(l_1, l_2, l_3)$
0	1	1	1/2	0
0	2	0	1	0
1	0	1	1/2	0
1	1	0	1/2	0
1	1	1	3/4	1/2
1	2	0	1	1/2
2	0	0	1	0
2	1	0	1	1/2

Table 1: Survival signatures for system in Figure 1, two cases

not precisely known, with lower and upper probability for Case 1 equal to 0.4 and 0.8, respectively, which by the conjugacy property for lower and upper probabilities [3] implies lower and upper probability 0.2 and 0.6 for Case 2. Finally, for Task E the same uncertainty occurs but there is no knowledge at all about the probability with which each case applies, represented by lower and upper probabilities 0 and 1, respectively, for both cases.

The survival signatures for Tasks A and B are just $\Phi_A = \Phi_1$ and $\Phi_B = \Phi_2$. For Tasks C-E, the generalized structure functions are easily derived and lead to the (lower and upper) survival signatures given in Table 2, where for completeness also Φ_A and Φ_B are given and entries which are either equal to 0 or 1 for all these functions have been left out.

For these (lower and upper) survival signatures, the following ordering holds for all (l_1, l_2, l_3),

$$\Phi_B = \underline{\Phi}_E \leq \underline{\Phi}_D \leq \Phi_C \leq \overline{\Phi}_D \leq \overline{\Phi}_E = \Phi_A$$

This means that in this example the special case applies in which expressions (13) and (14) give the lower and upper survival signatures, as the minimisations to derive the following lower survival signatures are all solved by the same probability vector in $\mathcal{P}$, and similar for the maximisations to derive the upper survival signatures. While this special case does not illustrate the full modelling ability of the concepts presented in this paper, it is of practical interest in situations such as discussed in this example, where there are a number of basic tasks which differ with regard to their demands on

l_1	l_2	l_3	Φ_A	Φ_B	Φ_C	$[\underline{\Phi}_D, \overline{\Phi}_D]$	$[\underline{\Phi}_E, \overline{\Phi}_E]$
0	1	1	0.5	0	0.25	$[0.2, 0.4]$	$[0, 0.5]$
0	2	0	1	0	0.5	$[0.4, 0.8]$	$[0, 1]$
1	0	1	0.5	0	0.25	$[0.2, 0.4]$	$[0, 0.5]$
1	1	0	0.5	0	0.25	$[0.2, 0.4]$	$[0, 0.5]$
1	1	1	0.75	0.5	0.625	$[0.6, 0.7]$	$[0.5, 0.75]$
1	2	0	1	0.5	0.75	$[0.7, 0.9]$	$[0.5, 1]$
2	0	0	1	0	0.5	$[0.4, 0.8]$	$[0, 1]$
2	1	0	1	0.5	0.75	$[0.7, 0.9]$	$[0.5, 1]$

Table 2: Lower and upper survival signatures for Tasks A-E

the system, and a variety of scenarios for the next possible task to be performed, each of these being represented by a different (imprecise) probability distribution over those basic tasks. For all such scenarios, the optimisations involved in deriving the lower and upper survival signatures for the next task to be performed by the system are straightforward, as in this example. We now consider several scenarios with different levels of knowledge about the type of the next task, the lower and upper survival signatures are presented in Table 3 (again leaving out those which are trivially equal to 0 or 1).

Suppose first, Scenario I, that the next task can be of any of the five types $A - E$, each with probability 0.2. The lower survival signature for the next task in this scenario, denoted by $\underline{\Phi}_I$, is derived as the average of the (lower) survival signatures for tasks A-E, and similar for the upper survival signature.

For Scenario II, suppose that the next task can again be of types A, B or C with probability 0.2 each, but there is uncertainty ('indeterminacy') with regard to the probability that this task may be of types D or E, reflected through lower and upper probabilities of 0.1 and 0.3, respectively, for both these types. To derive the lower survival signature for the next task in this scenario, we assign maximum probability 0.3 to $\underline{\Phi}_E$ for all (l_1, l_2, l_3), as this is never greater than $\underline{\Phi}_D$, which of course is assigned the minimum possible probability 0.1 to remain within the set of probability vectors $\mathcal{P}$. Similarly, due to $\overline{\Phi}_E \geq \overline{\Phi}_D$ for all (l_1, l_2, l_3), the corresponding upper survival signature is derived by assigning probability 0.3 to $\overline{\Phi}_E$ and 0.1 to $\overline{\Phi}_D$.

l_1	l_2	l_3	$[\underline{\Phi}_I, \overline{\Phi}_I]$	$[\underline{\Phi}_{II}, \overline{\Phi}_{II}]$	$[\underline{\Phi}_{III}, \overline{\Phi}_{III}]$	$[\underline{\Phi}_{IV}, \overline{\Phi}_{IV}]$
0	1	1	$[0.19, 0.33]$	$[0.17, 0.34]$	$[0.095, 0.415]$	$[0.17, 0.39]$
0	2	0	$[0.38, 0.66]$	$[0.34, 0.68]$	$[0.19, 0.83]$	$[0.34, 0.68]$
1	0	1	$[0.19, 0.33]$	$[0.17, 0.34]$	$[0.095, 0.415]$	$[0.17, 0.39]$
1	1	0	$[0.19, 0.33]$	$[0.17, 0.34]$	$[0.095, 0.415]$	$[0.17, 0.39]$
1	1	1	$[0.595, 0.665]$	$[0.585, 0.67]$	$[0.5475, 0.7075]$	$[0.535, 0.695]$
1	2	0	$[0.69, 0.83]$	$[0.67, 0.84]$	$[0.595, 0.915]$	$[0.62, 0.84]$
2	0	0	$[0.38, 0.66]$	$[0.34, 0.68]$	$[0.19, 0.83]$	$[0.34, 0.68]$
2	1	0	$[0.69, 0.83]$	$[0.67, 0.84]$	$[0.595, 0.915]$	$[0.62, 0.84]$

Table 3: Lower and upper survival signatures for Scenarios *I*-*IV*

To illustrate a greater level of indeterminacy with regard to the next task, Scenario III considers that it may be of each of the five identified types with lower probability 0.1 and upper probability 0.5. With the ordering of the (lower and upper) survival signatures for the five types, it is easy to verify that the lower survival signature over this set of probability vectors $\mathcal{P}$ is derived by assigning probability 0.5 to Φ_B, 0.2 to $\underline{\Phi}_E$ (these two values can be chosen differently as long as they sum up to 0.7 and are both between 0.1 and 0.5) and 0.1 to each of $\underline{\Phi}_D$, Φ_C and Φ_A. Similarly, the upper survival signature is derived by assigning probability 0.5 to Φ_A, 0.2 to $\overline{\Phi}_E$ and 0.1 to each of $\overline{\Phi}_D$, Φ_C and Φ_B.

Finally, we return to the setting of Scenario II, but with an important addition. For Scenario IV, suppose that it is judged that the next task the system needs to perform could actually also be a totally unknown task, for which it is not known at all whether or not the system can deal with it. This goes beyond the two basic tasks discussed throughout this example, for which the structure functions were given in Table 1. To reflect total lack of knowledge of such an unknown ('unidentified', 'unforeseen') task, which we indicate by index U, we can assign lower structure function $\underline{\phi}_U(l_1, l_2, l_3) = 0$ and upper structure function $\overline{\phi}_U(l_1, l_2, l_3) = 1$ for all (l_1, l_2, l_3), reflecting that even with all components functioning we do not know if the system can deal with this task, and that even with no components functioning it might be possible that this task can be satisfactorily dealt with. While these values may appear to be extreme, it covers all possibilities for unknown tasks, including e.g. targeted attacks on the system. It should be emphasized that such lack of knowledge cannot be taken into account adequately when restricted to the use of precise probabilities, and thus illustrates one of the major advantages of the use of imprecise probabilities. Let us assume that the next task can be of type U with lower probability 0 and upper probability 0.1, so the set of probability vectors over the six types $A - E$ and U consists of all probability vectors with $p_A = p_B = p_C = 0.2$, $p_D, p_E \in [0.1, 0.3]$ and $p_U \in [0, 0.1]$. To derive the lower survival signature for the next task in this scenario, we assign, in addition to the fixed probabilities 0.2 to types A, B, C, probability 0.1 to $\underline{\Phi}_U$, 0.2 to $\underline{\Phi}_E$ and 0.1 to $\underline{\Phi}_D$. To derive the corresponding upper survival signature, we similarly assign probability 0.1 to $\overline{\Phi}_U$, 0.2 to $\overline{\Phi}_E$ and 0.1 to $\overline{\Phi}_D$.

As is clear from Table 3, increase in indeterminacy, reflected through increased imprecision in the assigned lower and upper probabilities, leads to more imprecise lower and upper survival signatures in a logically nested way. From the perspective of risk management, the lower survival signatures are likely to be of most interest, as they reflect the most pessimistic scenario for system functioning corresponding to the information and assumptions made. As this example shows, the lower survival signature is derived by assigning the maximum possible probabilities to the possible types of task for which the system is least likely to function well.

In Scenario IV, we illustrated the possibility to include a totally unknown type of task by assigning lower and upper probabilities of 0 and 0.1 for the event that the next task is of such nature. In most risk scenarios, it would make sense to have lower probability 0 for such an event. The upper probability is, of course, more important for risk management as, combined with the lower probability for the system functioning well for such a task, it relates to the most pessimistic scenario. To illustrate our method we just chose the value 0.1 for this upper probability, yet it is worth mentioning that the nonparametric predictive inference (NPI) approach can actually provide a meaningful numerical value for the upper probability for the event that an as yet unobserved or even undefined event occurs [9, 10, 11, 12]. This NPI upper probability, which we do not discuss further in this paper, is based on relatively weak assumptions and is decreasing as function of the number of events considered in the data yet increasing as function of the number of different types of tasks the system had to deal with thus far. Of course, the presented upper and lower survival signatures for different scenarios, in this example, would not be the ultimate inference in most reliability investigations. Typically, they are combined, as illustrated before, with probabilities about the functioning of the components in order to get lower and upper probabilities of system functioning. This in turn can be used to support decisions on e.g. design of the system or inspection and maintenance strategies, this is beyond the scope of the current paper.

6. DISCUSSION

Traditional theory of system reliability tends to be based on quite strong assumptions with regard to knowledge about systems and their practical use. As shown in this paper, rather straightforward generalization of the structure function to consider it as a probability increases modelling opportunities substantially. Beyond that, the use of imprecise probabilities enables us to reflect indeterminacy, which is particularly important in risk scenarios where one may have limited knowledge and experience of the system functioning, or where the system may need to be resilient in case of unforeseen tasks. In this paper we have illustrated the approach mainly by considering different types of tasks, which in the example were related to two basic ways a given system could need to function, namely with one subsystem either requiring only one or at least two of its three components to function. The main advantage of the survival signature, as shown in this paper, is that this generalization of the structure function is straigthforwardly embedded in its definition, leading to lower and upper survival signatures. These are formulated for a single future task, which is important if one wishes to use statistical methods to infer system reliability and to reflect the amount of information available. Developing such statistical methods related to the lower and upper survival signatures is an interesting challenge for future research.

One could argue that using imprecise probability to reflect indeterminacy is an easy way out, as one effectively considers both the most optimistic and pessimistic scenarios which correspond to the information available, and reports the bounds based on these as the results of the inferences. The importance of this generalization of probability should, however, not be underestimated, as it avoids choosing precise values even in cases where there is no justification for doing so. Seeing the quality of the available information reflected explicitly in the reliability quantification, without lack of detailed information being hidden due to stronger assumptions or precise input values chosen for convenience, provides useful information for managing risks. If one does have quite detailed information it can be included in the inferences, and indeed doing so will normally lead to less imprecision, so it is certainly worth aiming to use all available information. In addition, one can also explore the influence of further assumptions or information on the imprecise results, which can be helpful if one wishes to explore what to focus on in order to derive the most useful information for a specific problem.

Following the first steps presented in this paper, there are many research challenges in order to develop a methodology that is applicable to large scale systems. It is important for such research challenges to be taken on with direct relation to real world applications, in order to discover the real problems and to see how results can be implemented. Part of such challenges will be in computation, as deriving the survival signature involves complex calculations, the number of which increases exponentially with the size of the system. Aslett [13] has developed a function in the statistical software R which can compute the survival signature for small to medium sized systems, but for practical systems and networks more research is required. For monotone systems, working with bounds for the survival signature, if it is only calculated for a subset of all combinations of numbers of functioning components of different types, is quite straightforward [7], this is also easily generalized to the lower and upper survival signatures as presented in this paper.

The theory presented in this paper is particularly useful for systems and networks with multiple types of components and with many components of the same type, as the survival signature is a sufficient summary of the system's structure which, in such cases, provides a substantial reduction compared to the complete structure function. One might encounter such systems and networks in many application areas, for example complex computer or communication systems with many parallel servers, energy networks, and transport infrastructure including rail networks. It may further be relevant for biology and medical research, exploring the opportunities for applications is an exciting challenge. In many modern applications emphasis is on real-time monitoring and online prediction. The setting presented in this paper may be suitable for such inference, in particular when combined with nonparametric predictive inference (NPI) [8] where inferences are in terms of the next event and take all data into account. The combined use of NPI and signatures has been presented for systems consisting of only a single type of components [14, 15]. Recently, NPI has also been applied together with the survival signature [7], this also requires a substantial research effort to become implementable to large scale practical problems. The use of

Bayesian methods in combination with the survival signature for quantification of system reliability has also recently been presented [16], and provides many further research challenges. In practice, components may not fail independently due to common failure causes. This has been investigated within the NPI framework, combined with the use of the survival signature [17], and can also be generalized by using the lower and upper survival signatures as presented in this paper.

There has, of course, been some research into the role of the structure function in system reliability, with consideration of generalizations, for example changing the system functioning, and hence the structure function, from binary to multi-state [18, 19] or even a continuous state space, for the latter situation a regression method has been proposed to derive the structure function [20]. Such generalizations can also be relevant together with the generalization proposed in this paper, and suggest further topics for research. Beyond this, one can also incorporate uncertainty about the functioning of the components, where again both precise and imprecise probabilistic methods may be explored. One may also generalize the structure function further to represent both a level (or 'quality') of system functioning and the probability that this level is attained, or, perhaps particularly relevant for networks, to reflect successful passing through the network with multiple possible start and end point; all such topics require careful consideration at the level of the structure function, but can be generalized by considering the structure function as a (possibly imprecise) probability.

We have advocated the explicit consideration of the 'next task' the system has to perform, but this is not a restriction on the proposed methodology. However, following this approach, one can also consider multiple future tasks, which may be interconnected as occurs in phased-missions. If there is dependence between the components' functioning or the system requirements for different phases, this must be modelled with care, but once more there is no conceptual difficulty in allowing the structure functions for each of such tasks to be generalized as a (imprecise) probability.

An important research topic from practical perspective is how such methods can deal with the effect of maintenance and inspection activities. Basically, if condition monitoring of components reveals information which leads to distinction between components which were initially deemed to be exchangeable, then this should be dealt with by creating a new type of component, and hence by changing the survival signature. The same may be required upon replacement of a component. The computational requirements for this have not yet been studied, where particularly methods that are applicable to large scale practical systems will provide interesting and important research challenges. As always, real-world application of the concepts introduced in this paper will be of major benefit to provide further research directions.

Acknowledgements

We thank two anonymous reviewers for the enthusiastic support of this work and suggestions to improve the presentation.

References

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

[2] Barlow, R.E., Proschan F. (1975). *Statistical Theory of Reliability and Life Testing.* Holt, Rinehart and Winston, Inc. New York.

[3] Augustin T., Coolen F.P.A., de Cooman G., Troffaes M.C.M. (2014). *Introduction to Imprecise Probabilities.* Wiley, Chichester.

[4] Coolen F.P.A., Utkin L.V. (2011). Imprecise reliability. In: *International Encyclopedia of Statistical Science*, Lovric M. (Ed.). Springer, Berlin, pp. 649–650.

[5] Samaniego F.J. (2007). *System Signatures and their Applications in Engineering Reliability.* Springer, New York.

[6] De Finetti B. (1974). *Theory of Probability.* Wiley, Chichester.

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

[8] Coolen F.P.A. (2011). Nonparametric predictive inference. In: *International Encyclopedia of Statistical Science*, Lovric M. (Ed.). Springer, Berlin, pp. 968–970.

[9] Coolen F.P.A., Augustin T. (2005). Learning from multinomial data: a nonparametric predictive alternative to the Imprecise Dirichlet Model. In: F. Cozman *et al* (Eds.), *Proceedings ISIPTA'05*, pp. 125–134.

[10] Coolen F.P.A., Augustin T. (2009). A nonparametric predictive alternative to the Imprecise Dirichlet Model: the case of a known number of categories. *International Journal of Approximate Reasoning* **50** 217–230.

[11] Coolen F.P.A. (2007). Nonparametric prediction of unobserved failure modes. *Journal of Risk and Reliability* **221** 207–216.

[12] Coolen-Maturi T., Coolen F.P.A. (2011). Unobserved, re-defined, unknown or removed failure modes in competing risks. *Journal of Risk and Reliability* **225** 461–474.

[13] Aslett L.J.M. (2012). ReliabilityTheory: Tools for structural reliability analysis. R package (*www.louisaslett.com*).

[14] Coolen F.P.A., Al-nefaiee A.H. (2012). Nonparametric predictive inference for failure times of systems with exchangeable components. *Journal of Risk and Reliability* **226** 262–273.

[15] Al-nefaiee A.H., Coolen F.P.A. (2013). Nonparametric predictive inference for system failure time based on bounds for the signature. *Journal of Risk and Reliability* **227** 513–522.

[16] Aslett L.J.M., Coolen F.P.A., Wilson S.P. (2015). Bayesian inference for reliability of systems and networks using the survival signature. *Risk Analysis* **35** 1640–1651.

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

[18] Bendell A., Ansell J. (1984). The incoherency of multistate coherent systems. *Reliability Engineering* **8** 165–178.

[19] Aven T. (1993). On performance measures for multistate monotone systems. *Reliability Engineering and System Safety* **41** 259–266.

[20] Gamiz M.L., Martinez Miranda M.D. (2010). Regression analysis of the structure function for reliability evaluation of continuous-state systems. *Reliability Engineering and System Safety* **95** 134–142.