

## Imprecise reliability analysis of complex interconnected networks

J. Behrendorf & M. Broggi

*Institute for Risk and Reliability, Leibniz Universitt Hannover, Hannover, Germany*

M. Beer

*Institute for Risk and Reliability, Leibniz Universitt Hannover, Hannover, Germany*

*Institute for Risk and Uncertainty, University of Liverpool, Liverpool, UK*

*International Joint Research Center for Engineering Reliability and Stochastic Mechanics (ERSM),  
Tongji University, Shanghai, China*

**ABSTRACT:** The effect of natural and man made disasters on critical infrastructures are substantial, as evident from recent history. Break downs of critical systems such as electrical power grids, water supply networks, communication networks or transportation can have dire consequences on the availability of aid in such a crisis. That is why, reliability analyses of these networks are of paramount importance. Two important factors must taken into consideration during reliability analysis. First, the networks are subject to complex interdependencies and must not be treated as individual units. Second, the reliability analysis is typically based on some form of data and or expert knowledge. However, this information is rarely precise or even available. Therefore, it is important to account for different kinds of uncertainties, namely aleatory uncertainty and epistemic uncertainty. Aleatory uncertainty represents the natural randomness in a process, while epistemic uncertainty represents vagueness or lack of knowledge in the model. In this work we present an approach to the numerical reliability analysis of complex networks and systems extending a previously developed method based on Monte Carlo simulation and survival signature. The extended method treats both kinds of uncertainties, thus, yielding better results. We show how Monte Carlo simulation controls aleatory uncertainty and apply sets of distributions (probability boxes) to treat epistemic uncertainties in component failures. In this framework, dependencies are modelled using copulas. Copulas possess the unique property of decoupling the modelling of the univariate margins from the modelling of the dependence structure for continuous multivariate distributions. Analogous to the p-boxes we use sets of copulas to include imprecision in the dependencies. Finally, the method is applied to an example system of coupled networks.

### 1 INTRODUCTION

Modern infrastructure systems are highly complex and subject to a multitude of different dependencies. Disasters in recent years have shown how critical the impact of these dependencies can be. Failures in one network such as a power outage will surely impact other dependent systems. In worst case scenarios these dependencies can lead to cascading effect ultimately breaking down entire networks (Buldyrev et al. 2010). This highlights the need for methods of reliability analysis that can deal with these complexities.

Recently, the survival signature (Coolen and Coolen-Maturi 2013) has gained in popularity as a tool to aid with this task. The survival signature allows to decouple the structural evaluation from the probabilistic analysis, allowing for highly efficient simulation. In Behrendorf et al. 2017 we introduced a method for the reliability analysis of complex interdependent networks. Other efficient algorithms can be found for example in Patelli et al. 2017.

A secondary task during the reliability analysis is the accurate modelling of component failures and dependencies. Typically, this is done based on data or expert assessments. However, both are subject to two kinds of imprecisions, namely, aleatory and epistemic uncertainty (Beer et al. 2013). Aleatory uncertainty represents the randomness inherent in a process, such as component degradation and external forces affecting the system (natural hazards, earthquakes, etc.), while epistemic uncertainty describes the uncertainty in the model due to a lack of or vagueness of knowledge about the system. The latter is usually regarded as reducible through acquiring of additional data and information.

In this work we expand our previously developed technique by inclusion of imprecision. The method is based on Monte Carlo simulation and as such already deals with aleatory uncertainty. In this extension the modelling of component failures is refined by applying probability-boxes (p-boxes) to account for epistemic uncertainty. Feng et al. 2016 have shown the advantages of using p-boxes

in reliability analysis. Additionally, the rather simple dependency modelling of the initially developed method is replaced by imprecise copulas (Montes et al. 2015). Copulas split continuous multivariate distributions in a dependence structure and univariate marginals, which in turn allows for separate flexible modelling of the two (Joe 2014).

This paper is outlined as follows. First we introduce the previously developed method for the reliability analysis of networks. Then, after presenting basic theory and notation on copulas, we discuss how to model dependencies with copulas and how to translate these methodologies into an imprecise setting. Finally, we apply the developed techniques to a simple example. The paper closes with some concluding remarks and an insight into future work.

## 2 RELIABILITY ANALYSIS

This section presents the survival signature based method to calculate the reliability of a system, as first presented in Behrendorf et al. 2017.

The survival signature was developed as an extension to the system signature (Samaniago 2007), overcoming the limitations that restrict the system signature to systems of one single component type (Coolen and Coolen-Maturi 2013). The main function of the survival signature is to separate the structural information of a network from its probabilistic characteristics.

### 2.1 Survival signature

Considering a system with  $m$  components, the survival signature for  $l$  out of  $m$  components working is defined as

$$\Phi(l) = \binom{m}{l}^{-1} \sum_{\underline{x} \in S_l} \varphi(\underline{x}), \quad (1)$$

where  $\underline{x} = (x_1, \dots, x_m)$  denotes the state vector of the system with  $x_i = 1$  and  $x_i = 0$  representing a working or failed component respectively and  $\varphi(\underline{x})$  is the structure function returning the state of the full system with  $\varphi(\underline{x}) = 1$  indicating a working system and  $\varphi(\underline{x}) = 0$  indicating a failed system.

Extending the survival signature to systems with  $K$  component types and  $m_k$  components for each type  $k (k=1, \dots, K)$  and  $l_k$  out of  $m_k$  components working results in

$$\Phi(l_1, \dots, l_K) = \left[ \prod_{k=1}^K \binom{m_k}{l_k}^{-1} \right] \times \sum_{\underline{x} \in S_{l_1, \dots, l_K}} \varphi(\underline{x}). \quad (2)$$

An efficient algorithm to compute the survival signature can be found in Aslett 2012.

### 2.2 Survival function

The next step in calculating the reliability of a system is the definition of the survival function. This function uses the survival signature to calculate the probability that a system is working at time  $t$  and as such calculates the reliability. The survival function is defined as

$$P(T_s > t) = \sum_{l_1=0}^{m_1} \dots \sum_{l_k=0}^{m_k} \Phi(l_1, \dots, l_k) \cdot P\left(\bigcap_{k=1}^K \{C_t^k = l_k\}\right). \quad (3)$$

Note especially the separation of structural information (left) and probabilistic information (right). This means, that the structural evaluation of the system must occur only once for the entire reliability analysis. In the last remaining step, the probabilistic part of the survival function is approximated using Monte Carlo Simulation in order to be able to include imprecisions and interdependencies in the analysis.

### 2.3 Monte Carlo simulation

The simulation starts by selecting a sufficient number of samples  $N_{MC}$  and small time step followed by the sampling of component failure times from the assumed copula (see section 3). Next, in two nested loops over all combinations  $l_1, \dots, l_k$  where  $\Phi(l_1, \dots, l_k) > 0$  and all time steps  $t$  the number of samples in the same configuration are counted as  $N_{l_1, \dots, l_k}(t)$ . Then, the probabilistic part of the survival function is approximated by

$$P\left(\bigcap_{k=1}^K \{C_t^k = l_k\}\right) \approx \frac{N_{l_1, \dots, l_k}(t)}{N_{MC}} \quad (4)$$

Finally, the partial reliabilities obtained in the previous step are multiplied by their probability from the survival signature and summed up, yielding the full reliability of the network.

## 3 MODELLING DEPENDENCIES

This section introduces the necessary notation of copulas and how to apply them to model dependencies in and between networks. In more detail, section 3.4 presents how to use copulas to model common causes of failure while section 3.5 shows how to model interdependencies. This is a very basic introduction, for a thorough discussion of copulas see Nelsen 2006.

### 3.1 Copulas

The basis of copulas is built by what is today known as Sklar's theorem (Sklar 1959). The theorem states that any multivariate distribution  $H$  (in dimensions  $d \geq 2$ ) can be separated into its univariate marginal distributions  $F_i$  and a copula function  $C : [0,1]^d \rightarrow [0,1]$ .

**Theorem 1. Sklar's theorem** *Let  $H$  be a  $d$ -dimensional distribution function with margins  $F_1, \dots, F_d$ . There exists an  $n$ -dimensional copula  $C$  such that for all  $\mathbf{x}$  in  $\mathbb{R}^d$*

$$H(\mathbf{x}) = C(F_1(x_1), \dots, F_d(x_d)). \quad (5)$$

*If the marginals  $F_1, \dots, F_d$  are continuous, then  $C$  is unique; otherwise,  $C$  is unique on  $\text{Range}(F_1) \times \dots \times \text{Range}(F_d)$ .*

*Conversely, if  $C$  is a  $d$ -copula and  $F_1, \dots, F_d$  are distribution functions, then the function  $H$  defined by Eq. 5 is an  $d$ -dimensional distribution function with margins  $F_1, \dots, F_d$ .*

This facilitates separate modelling of the marginal distributions from modelling of the dependence structure, in turn allowing for effective treatment of imprecisions in both parts (see section 4).

In this work we apply three distinct copula families, namely the Gaussian copula, the Independence copula, and the Clayton copula. For an encompassing discussion of copula families the reader is referred to Nelsen 2006 and Joe 2014.

The  $d$ -dimensional Gaussian copula is defined as

$$C_R(u_1, \dots, u_d) = \Phi_d(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_d)), \quad (6)$$

where  $\mathbf{R} \in [-1,1]^{d \times d}$  is a positive definite correlation matrix and  $\Phi_d(\cdot; \mathbf{R})$  is the  $d$ -variate cumulative distribution of a  $\mathbb{N}_d(0, \mathbf{R})$  random vector.  $\Phi^{-1}$  denotes the inverse of the univariate standard Gaussian cdf (Joe 2014).

The latter two copulas belong to the class of Archimedean copulas. This family is particularly popular due to their easy construction and wide range of applications (Nelsen 2006). The Archimedean copulas used in this work are one parameter families, which allow for easy treatment of imprecision as seen in the subsequent section. Any  $d$ -dimensional Archimedean copula is constructed using a so called *generator* function  $\varphi : [0, \infty) \rightarrow [0,1]$  and its inverse  $\varphi^{-1}$  according to

$$C_\varphi(u_1, \dots, u_d) := \varphi(\varphi^{-1}(u_1) + \dots + \varphi^{-1}(u_d)), \quad (7)$$

where  $u_1, \dots, u_d \in [0,1]$  (Mai and Scherer 2012). Table 7 shows the generators and parameter ranges for the Independence and Clayton copula families.

### 3.2 Dependence measure

Studies have shown, that correlation is not a suitable measurement of dependence for copulas

(Schirma Schirmacher and Schirmacher 2008). Therefore, in this work, Kendall's tau is selected as the preferred dependence measure. Kendall's tau is based on *concordance*. A pair of random variables is said to be concordant if "large" values are associated with "large" values "small" values with "small". Formally, two observations  $(x_i, y_i)$  and  $(x_j, y_j)$  from a vector  $(X, Y)$  are *concordant* if  $(x_i - x_j)(y_i - y_j) > 0$  and *discordant* if  $(x_i - x_j)(y_i - y_j) < 0$ .

Then, Kendall's tau for a sample of  $n$  observations  $\{(x_i, y_i), \dots, (x_n, y_n)\}$  from a vector of continuous random variables  $(X, Y)$  can be defined as

$$t = \frac{c-d}{c+d} = (c-d) / \binom{n}{2}, \quad (8)$$

where  $c$  and  $d$  represent the number of concordant and discordant pairs among all possible pairs of observations. This value may also be interpreted as the probability of concordance minus the probability of discordance for a random pair of observations  $(x_i, y_i)$  and  $(x_j, y_j)$ . Based on this fact, Kendall's tau for the random variables  $X$  and  $Y$  can be defined by

$$\tau(X, Y) = P[(X - \tilde{X})(Y - \tilde{Y}) > 0] - P[(X - \tilde{X})(Y - \tilde{Y}) < 0], \quad (9)$$

where  $(\tilde{X}, \tilde{Y})$  is an independent copy of  $(X, Y)$  (Schirmacher and Schirmacher 2008).

Kendall's tau is not only used to measure dependence but also to find copula parameters representing a desired strength of dependence.

### 3.3 Vine copulas

In order to analyse the reliability of complex networks it one must build high dimensional copulas. However, in comparison to bivariate copulas, the available literature on multivariate copulas is scarce (Mai and Scherer 2012). For this reason, we employ a technique called *pair copula construction* to break down multivariate copulas into combinations of bivariate copulas. More accurately, we use vine copulas as a graphical tool to model pair cop-

Table 1. Generators, generator inverses and parameter ranges for the Clayton and Independence copula.

Name	Generator	Generator Inverse	Parameter $\theta$
	$\varphi_\theta(t)$	$\varphi_\theta^{-1}(t)$	
Clayton	$\frac{1}{\theta}(t^\theta - 1)$	$(1 + \theta)^{-1/\theta}$	$\theta \in [-1, \infty) \setminus \{0\}$
Independence	$-\log(t)$	$\exp(-t)$	

ula constructions as sets of trees. An example of a five-dimensional vine copula is shown in Fig. 1.

A regular vine (R-Vine)  $\mathcal{V}$  is defined as a set of trees  $T_1, \dots, T_{d-1}$ , where  $T_1$  consists of nodes  $N_1 = \{1, \dots, d\}$  and edges  $E_1$ . Every subsequent Tree  $T_j$  uses the edges  $E_{j-1}$  as nodes and connects them with the edges  $E_j$ . The last property needed to define a regular vine is the *proximity property*, stating that if  $a$  and  $b$  are connected by an edge in  $T_j, j \geq 2$ , then  $a$  and  $b$  must share a common node in  $T_{j-1}$ .

A plethora of different structures exist for a  $d$ -dimensional R-Vine based on the definition. Because of this and the fact, that a regular vine possesses  $2^{n-1}$  sampling order we apply D-vines instead, which have a much more restrictive structure. As seen in Fig. 2, a D-Vine is characterized by each node  $n \in N_i$  having a maximum degree of 2. Sampling from D-Vines is a lot simpler and in this work performed by using the MATLAB toolbox VineCopulaMatlab (Kurz 2016).

### 3.4 Common cause of failure

One type of dependent failure tackled in this work is common cause of failure. It is defined as two or more components failing at the same time due to common defects or weaknesses. Causes include but are not limited to: errors in manufacturing, errors during maintenance or operation, and environmental causes such as earthquakes or tsunamis

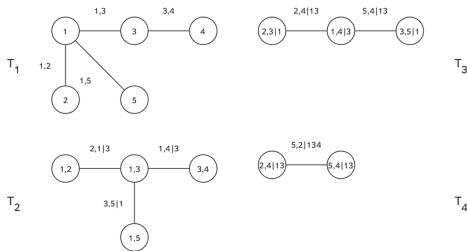


Figure 1. Five-dimensional copula represented as a regular vine.

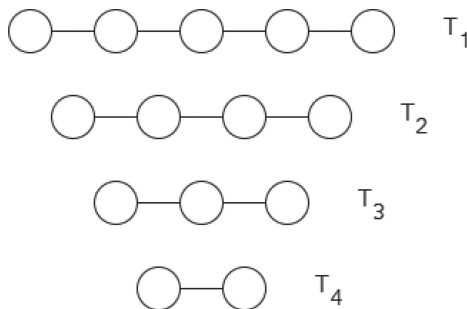


Figure 2. Structure of a five-dimensional D-Vine.

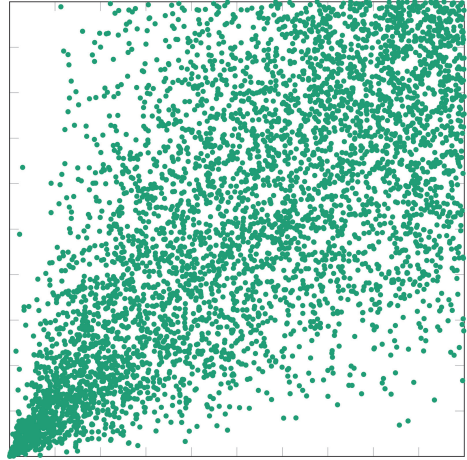


Figure 3. Samples drawn from a Clayton copula with the parameter chosen so  $\tau = 0.5$ .

(Hanks 1998). We model common cause of failure by application of Clayton copulas. This family possesses a property called lower-tail dependence, meaning that dependence is stronger in the lower-left quadrant of  $[0,1]^2$ . By modelling the failures this way, the dependence is much stronger in early component life and as such brings us closer to the traditional bathtub shape or component failure probabilities. Figure 3 shows samples drawn from a Clayton copula. Note, how the stronger dependence in the lower-tail is clearly visible in the scatter plot.

### 3.5 Interdependencies

Interdependencies between nodes and networks are handled by application of Gaussian Copulas. Gaussian copulas possess no tail-dependence and show good results, although other families could be investigated for the same application in the future. In addition to the copula there is one more step required to accurately model the dependencies. We understand interdependencies as the phenomenon of one component failing due to the failure of another. As such, interdependencies imply causality. In order to represent this causality in the model, dependence is introduced in the marginals. During the transformation of the failure times sampled from the copula by the inverse transformation method, the marginal distributions are aggregated from the dependent marginals using Kendall's tau of the random variables  $u_1$  and  $u_2$  as

$$U_1 = (1 - \tau) \cdot F_1^{-1}(u_1) + \tau \cdot F_2^{-1}(u_1), \quad (10)$$

where  $F_1$  and  $F_2$  are the marginals of a copula  $C$ .

#### 4 HANDLING IMPRECISION

Two types of uncertainties must be taken care of during the reliability analysis, namely, aleatory and epistemic uncertainties. Aleatory uncertainty describes the natural randomness inherent in a process, while epistemic uncertainty represents the uncertainty due to vagueness in information or a lack thereof.

Aleatory uncertainty can automatically handled by our reliability analysis technique. Through assuming failure time distributions for the component failures and sampling these during Monte Carlo simulation, the randomness that our model is subject to is fully included. However, the selection appropriate failure time distributions is typically based on either data or expert knowledge, neither of which yield perfect results, in turn introducing epistemic uncertainty into the model. This uncertainty can be reduced by using *probability-boxes* (p-boxes) (Feng et al. 2016).

P-boxes are defined as bounds on the cumulative distribution function of a random variable. The left and right bounds can be found by for example selecting an appropriate distribution and giving the parameters as intervals. As such, a p-box comprises both the aleatory and the epistemic uncertainty. An example of an exponential p-box with parameters  $\lambda \in [1.2, 2.2]$  is shown in Fig. 4.

By feeding the bounds of the p-box into the reliability analysis, the epistemic uncertainty propagates into the result. Thus, instead of one survival function, we obtain an upper and lower bound. Figure 5 shows an example of the upper and lower bounds obtained by performing a reliability analysis of a simple system of two parallel components of the same type, assuming the p-box of Fig. 4 for the failure time distributions.

Similarly to the application of p-boxes to handle epistemic uncertainty in the marginals, we can define the copula parameters as intervals and obtain *imprecise copulas* for the dependencies (Montes et al. 2015). This works especially well since all copula families, including the bivariate Gaussian copula, we apply in the vine copula are defined by a single parameter.

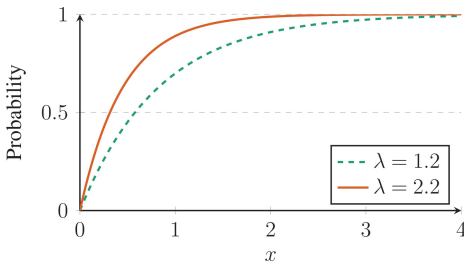


Figure 4. Example of an exponential p-box with  $\lambda \in [1.2, 2.2]$ .

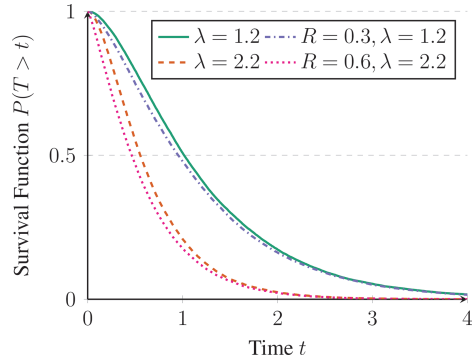


Figure 5. Upper and lower bounds of the reliability resulting from applying the p-box in Fig. 4 to a simple system of two parallel components.

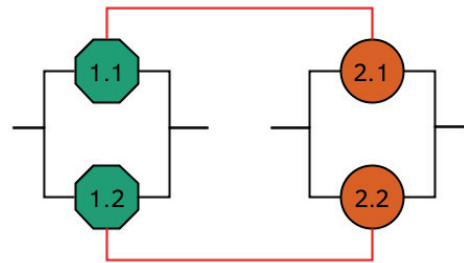


Figure 6. Structure of the example network. The red lines represent interdependencies between the two subsystems.

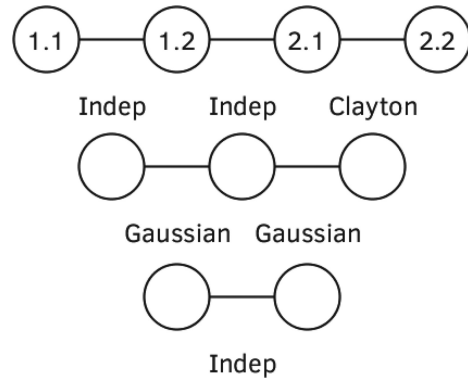


Figure 7. D-Vine used to model the common cause of failure in system 2 and the interdependencies between the two networks.

Finally, the bounds of the p-boxes as well as the bounds of the copula parameters are fed into the reliability analysis. Returning to the simple system of two parallel components and linking the components with an imprecise Gaussian copula  $R \in [0.3, 0.6]$  results in the upper and lower bounds for the reliability as seen in Fig. 5.

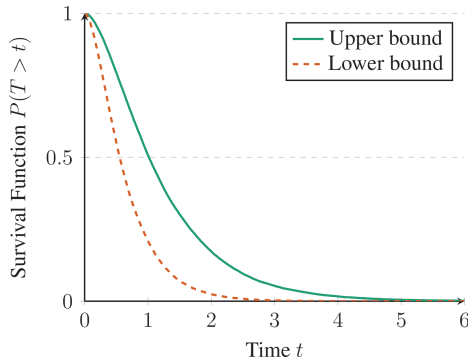


Figure 8. Bounds on the reliability of system 2 based on the assumed imprecisions.

## 5 NUMERICAL EXAMPLE

The methods that were introduced in the previous sections will now be applied to a simple toy example. All marginals and dependencies will be considered as imprecise in order to account for all aleatory and epistemic uncertainties. Figure 6 presents the example system build from two systems of two parallel components where the first and second components respectively are interconnected.

A four-dimension D-Vine copula, including the interdependencies and a common cause of failure shared among the components in system 2, is built in order to sample the component failure times. The structure of the vine is shown in Fig. 7.

The component failure times for systems 1 and 2 are assumed to be exponentially distributed with  $\lambda_1 \in [1.5, 1.7]$  and  $\lambda_2 \in [0.7, 1.1]$ . The copula parameters are chosen such that  $\tau \in [0.2, 0.4]$  for the Clayton copula and  $\tau \in [0.4, 0.6]$  for the Gaussian copulas. The resulting bounds on the reliability of system 2 are plotted in Fig. 8.

## 6 CONCLUSION AND OUTLOOK

In this work we have presented how to perform reliability analyses of complex interdependent networks in a highly imprecise setting. The necessary theory on copulas and vine-copulas and applied to the modelling of dependencies in and between networks. Finally, the modelling of dependencies and a previously introduced method for the reliability analysis of networks were extended to account for both aleatory and epistemic uncertainties. As a result, we obtained bounds on the network reliability. The method was applied to a numerical toy example to prove the functionality.

It is obvious that this paper only serves as a short introduction into future work. The methods must be validated further and applied to complex real world networks in order to ensure usability. Especially the construction of the D-Vine copula for sampling of the component failure times must be further investigated. There exist a plethora of possible vine structures for a given problem and effective automatic construction techniques have to be created.

## REFERENCES

- Aslett, L. (2012). Reliabilitytheory: Tools for structural reliability analysis, r package. <http://www.louisaslett.com>. Accessed: 11.12.2017.
- Beer, M., S. Ferson, & V. Kreinovich (2013). Imprecise probabilities in engineering analyses. *Mechanical systems and signal processing* 37(1), 4–29.
- Behrendorf, J., M. Broggi, S. Brandt, & M. Beer (2017). Numerically efficient reliability analysis of interdependent networks. In *The 27th European Safety and Reliability Conference*.
- Buldyrev, S. V., R. Parshani, G. Paul, H. E. Stanley, & S. Havlin (2010). Catastrophic cascade of failures in interdependent networks. *Nature* 464(7291), 1025–1028.
- Coolen, F. P. & T. Coolen-Maturi (2013). Generalizing the signature to systems with multiple types of components. In *Complex systems and dependability*, pp. 115–130. Springer.
- Feng, G., E. Patelli, M. Beer, & F. P. Coolen (2016). Imprecise system reliability and component importance based on survival signature. *Reliability Engineering & System Safety* 150, 116–125.
- Hanks, B. (1998). An appreciation of common cause failures in reliability. *Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering* 212(1), 31–35.
- Joe, H. (2014). *Dependence modeling with copulas*. CRC Press.
- Kurz, M. (2016). Vinecopulamatlab toolbox. <http://maltekurz.github.io/VineCopulaMatlab/>. Accessed: 11.12.2017.
- Mai, J.F. & M. Scherer (2012). *Simulating copulas: stochastic models, sampling algorithms and applications*. World Scientific.
- Montes, I., E. Miranda, R. Pelessoni, & P. Vicig (2015). Sklar's theorem in an imprecise setting. *Fuzzy Sets and Systems* 278, 48–66.
- Nelsen, R. B. (2006). *An Introduction to Copulas*. Springer Science & Business Media.
- Patelli, E., G. Feng, F. P. Coolen, & T. Coolen-Maturi (2017). Simulation methods for system reliability using the survival signature. *Reliability Engineering & System Safety* 167, 327–337.
- Samaniego, F. J. (2007). System signatures and their applications in engineering reliability. 110.
- Schirmacher, D. & E. Schirmacher (2008). Multivariate dependence modeling using pair-copulas. Technical report, Citeseer.
- Sklar, M. (1959). Fonctions de repartition an dimensions et leurs marges. *Publ. inst. statist. univ. Paris* 8, 229–231.